Structural and mechanistic studies of membrane protein biogenesis and quality control at the endoplasmic reticulum

Thesis by

Vy Ngoc Mai Nguyen

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2025

(Defended May 20, 2025)

Vy Ngoc Mai Nguyen

ORCID: 0000-0002-7563-7386

ACKNOWLEDGEMENT

First and foremost, I want to thank my research advisor, Rebecca Voorhees. Thank you for welcoming me into the lab and giving me a lot of freedom to explore various aspects of my research projects. Your encouragement and support have strengthened me immensely as a scientist. I feel very fortunate to be able to experience and learn so many techniques in this lab and work with many brilliant people. I would also like to thank the members of my committee, Doug Rees, Pamela Bjorkman, and Bil Clemons, for your continued support and advice throughout the years of my graduate career.

I would also like to thank the people who have ignited my passion for research before I arrived at Caltech because I would not be here had it not been for them. During my undergraduate years at UCLA, I took Stan Schein's "Structure: Patterns and Polyhedra" class, where I was first introduced to Todd Yeates' research on protein nanocages and scaffolding. I was fascinated by the structures of living things and obsessed over creating modular origami models of various polyhedra. Stan invited Todd to sit in at my final project presentation for the class, and that eventually gave me the courage to ask Todd to join his lab. I am extremely thankful for my undergraduate research advisor, Todd Yeates. In his lab, I discovered my love for structural biology and was fortunate to work with these wonderful people: Kevin Cannon, Jessica Ochoa, and Mengxiao Nie. I thank Michael Sawaya and Duilio Cascio for teaching me X-ray crystallography.

At Caltech, I was extremely lucky to be working with several talented scientists. I want to thank my mentors when I first joined the lab in the EMC-NOMO team: Giovanni Pinton Tomaleri, Tino Pleiner, Katie Page, Masami Hazu, and Kurt Januszyk. I also want to thank Katie Page, Lena Bögeholz, and Jichen Zhang as I have worked so closely with them for the past 2 years on the TXN project. And I am grateful for the other Voorhees lab members that I had the privilege to share the lab space with for creating such a fun and collaborative working environment (in alphabetical order): Alina Guna, Aline Milach Teixeira, Alison Inglis, Camryn Lee, Charlene DeKalb, Dina Malounda, Esther Leem, Maxine Wang, Melanie Ernst, Natalie Chen, Rob Oania, Sophia Wei, and Taylor Stevens. I want to also

thank the many scientists working at the facilities on Caltech campus for providing technical support and offering valuable scientific insights, especially Songye Chen and Ting-Yu Wang.

A wise person once said, “Maybe the real PhD was the friends we made along the way.” I am incredibly thankful to have met these wonderful people during my time at Caltech (in alphabetical order): Cai Tong Ng, Erini Galatis, Evelyn Li, Maxine Wang, and Victor Garcia. Thank you for spending quality time with me and creating a safe space for me to share my thoughts whenever I’m going through a rough patch during this graduate school career, or to just have fun. I want to give a special shoutout to my Vietnamese friend group (in alphabetical order): Ha Dang, Hai Dong, Khoa Le, Linh Le, Mai Nguyen, and Thu Ton. Thank you for introducing me to all the good food spots around LA, and for giving me a sense of home away from home. Em cảm ơn mọi người nhiều nhé!

I am deeply grateful for BTS. Their meaningful songs and messages have comforted, given me hope, and fueled me to get through this journey. To Namjoon, Seokjin, Yoongi, Hoseok, Jimin, Taehyung, and Jungkook: I cannot wait for your reunion this June. I want to thank Charlene and Mike for being my concert buddies. I also want to give a special shoutout to the ARMY friends I made throughout the years, especially Emi, Krista, Megan, Tiffany, and Vicky. 아포방포!

I am thankful for my family in Vietnam, in Japan, and in America. I would not have been able to get this far without your support.

And last, but not least, I want to thank my husband. Thank you for always being in my corner through ups and downs, and for tolerating my silly quirks. You have given me strength through rough times and remind me that I am enough. I am so lucky to have you!

ABSTRACT

Membrane proteins make up around 30% of the human proteome and carry out essential functions in the cell, including but not limited to signaling, cell adhesion, and metabolic transport. To maintain homeostasis, the eukaryotic cell has evolved complex pathways for membrane protein biogenesis and quality control, as they contain hydrophobic transmembrane domains (TMDs) that can easily aggregate in the aqueous environment of the cytosol causing cellular damage. At the endoplasmic reticulum (ER), the ER membrane protein complex (EMC) co-translationally inserts the first TMD of multipass membrane protein and post-translationally inserts tail-anchored membrane proteins. In this thesis, we are able to show how the EMC can coordinate with the other protein machineries of the multipass translocon (such as the back-of-Sec61/BOS complex) at the ER to accommodate the insertion of diverse multipass membrane protein substrates, depending on the biophysical properties of their N-terminal soluble domain. We also structurally characterize the EMC•BOS holocomplex, highlighting its spatial relation with the other biogenesis factors at the multipass translocon. In addition, this thesis also explores how a novel quality control factor TXNDC15 functions in ER-associated degradation (ERAD) to facilitate degradation of unassembled membrane proteins. We found that TXNDC15 works with the E3 ligase MARCHF6 and recognizes a set of membrane proteins with exposed hydrophobic domain in the ER lumen. Lastly, we were able to also link TXNDC15's quality control function to its implication in ciliopathies, such as Meckel-Gruber syndrome.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Pleiner, T.*, Hazu, M.*, Pinton Tomaleri, G.*, **Nguyen, V. N.**, Januszyk, K., & Voorhees, R. M. (2023). A selectivity filter in the ER membrane protein complex limits protein misinsertion at the ER. *Journal of Cell Biology*, *222*(8), e202212007. https://doi.org/10.1083/jcb.202212007
Contributions: V.N.N. helped analyze data and edit the manuscript.

2. Stevens, T. A., Tomaleri, G. P., Hazu, M., Wei, S., **Nguyen, V. N.**, DeKalb, C., Voorhees, R. M., & Pleiner, T. (2024). A nanobody-based strategy for rapid and scalable purification of human protein complexes. *Nature Protocols*, *19*(1), 127–158. https://doi.org/10.1038/s41596-023-00904-w
Contributions: V.N.N. performed experiments, analyzed data, and helped edit the manuscript.

3. Page, K. R.*, **Nguyen, V. N.***, Pleiner, T.*, Tomaleri, G. P., Wang, M. L., Guna, A., Hazu, M., Wang, T.-Y., Chou, T.-F., & Voorhees, R. M. (2024). Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Molecular Cell*, *84*(17), 3302-3319.e11. https://doi.org/10.1016/j.molcel.2024.08.005
Contribution: V.N.N. participated in designing the project, performed the experiments, analyzed data, helped assemble and write the manuscript, and did the revisions.

4. **Nguyen, V. N.***, Bögeholz, L. A. K.*, Page, K. R.*, Zhang, J., Chen, N., Wang, T.-Y., Mayank, A., Guna, A., Wang, M. L., Wohlschlegel, J., Chou, T.-F., Voorhees, R. M. (2025). A pathway for integral membrane protein quality control within the ER lumen. Manuscript under review.
Contribution: V.N.N. participated in designing the project, performed the experiments, analyzed data, and helped assemble and write the manuscript.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS AND TABLES

Page

CHAPTER 1:

INTRODUCTION

1.1. Overview

Evolved from prokaryotes, eukaryotic cells are distinguished by having membrane-bound organelles. The membrane surrounding these organelles is referred to as the phospholipid bilayer, and the proteins associated with the bilayer are membrane proteins. Membrane proteins make up around 30% of the eukaryotic proteome (Wallin & von Heijne, 1998). They serve various functions in the cell, including but not limited to signaling, transporting metabolites and larger molecules across membranes, and maintaining organelle and cell morphology (von Heijne, 2007). Because of their ubiquitous functions, membrane proteins are major targets for drug development as defects in their biogenesis and quality control have been implicated in numerous diseases (Sanders & Nagy, 2000; Sanders & Myers, 2004; Chiti & Dobson, 2006; Ng et al., 2012).

Membrane proteins are divided into two major categories based on their relative position to the membrane: integral membrane proteins and peripheral membrane proteins. Peripheral membrane proteins are attached to the membrane via interactions with lipids and/or other membrane proteins. On the other hand, integral membrane proteins are permanently anchored or embedded in the bilayer via hydrophobic interactions as these proteins possess one or more transmembrane domains (TMDs). α-helical and β-barrel are the only two types of secondary structures found for these TMDs that are embedded in the membrane (von Heijne, 1996). Because the protein backbone needs to satisfy its hydrogen bonding requirements in the hydrophobic environment of the bilayer, α-helical and β-barrel architectures allow for the protein backbone to hydrogen-bond to itself. The rest of this work will focus on integral membrane proteins with α-helical TMDs.

The hydrophobic nature of these TMDs of integral membrane proteins presents a challenge for the cell as membrane proteins are synthesized by the ribosomes in aqueous cytosolic environment, which can result in aggregation, causing cellular damage. Hence, these membrane proteins with hydrophobic TMDs need to be shielded before they are eventually inserted into the endoplasmic reticulum (ER), where they are folded, assembled into complexes, and trafficked to the correct cellular compartments in the latter steps of the secretion pathway (Shao & Hegde, 2011). In case a membrane protein is not properly assembled or folded, it needs to be rapidly turned over by ER-associated degradation (ERAD) pathways to prevent cytotoxic effects (Hegde & Ploegh, 2010). These checks and balances for the biosynthesis and quality control of membrane proteins at ER help to regulate and maintain cellular homeostasis.

1.2. Membrane protein biogenesis at the ER

As aforementioned, all membrane proteins are translated by the ribosome in the cytosol. When the nascently synthesized membrane protein emerges from the exit tunnel of the ribosome, their hydrophobic patches need to be recognized by targeting machinery and be shielded from the aqueous cytosolic environment (Shao & Hegde, 2011). These membrane proteins are either co-translationally or post-translationally inserted into the ER membrane by insertion machineries, often referred to as ‘insertases’, in the correct orientation in relation to the membrane (i.e., topology). Integral membrane proteins can be classified based on the number of TMDs and their corresponding topologies. Single-pass membrane proteins only contain one single TMD, while multipass membrane proteins contain two or more TMDs. Type I membrane proteins have a signal peptide sequence that get cleaved off upon insertion into the membrane, and their N-termini are translocated into the ER lumen (i.e., N_{exo} topology). Type II membrane proteins adopt the N_{cyt} topology instead, where their C-termini get translocated into the ER lumen and their N-termini remain in the cytosol. Type III membrane proteins also adopt the N_{exo} topology similarly to Type I, but do not contain the cleavable signal peptide (The UniProt Consortium, 2017; Wallin & von Heijne, 1998). Tail-anchored (TA) membrane proteins, which have a single TMD positioned at around 40 amino acids from the C-termini, are in a class of their own as their topology

requires them to be targeted and inserted into the bilayer post-translationally as the newly synthesized TMD only emerges from the ribosomal exit tunnel upon translational termination (Guna, Hazu, et al., 2023).

For many years, the textbook model of how membrane proteins are inserted into the ER membrane is as follows: as the newly synthesized membrane protein emerges from the translating ribosome, the ribosome-interacting signal recognition particle (SRP) recognizes either the cleavable signal peptide or the first TMD of these membrane proteins and further delivers them to the Sec61 translocation complex for insertion at the ER membrane (Grudnik et al., 2009; Keenan et al., 1998; Rapoport et al., 2017; Wild et al., 2004). The Sec61 translocon comprises three subunits: Sec61α, Sec61β, and Sec61γ. The pore formed by Sec61α allows for engagement of the translocating segments, which could trigger the rearrangement of helices within Sec61α to open its 'lateral gate,' allowing for the incorporation of the hydrophobic TMD into the lipid bilayer (Van den Berg et al., 2004; Voorhees & Hegde, 2016).

However, this model is not sufficient to completely explain the insertion requirement of more than 5,000 integral membrane proteins that are encoded in the human genome (The UniProt Consortium, 2017). Additionally, it had been shown that not all TMDs are hydrophobic enough to trigger the opening of the lateral gate of Sec61 for their insertion into the membrane (Enquist et al., 2009; Schorr et al., 2020). Work in recent years have identified additional insertases at the ER, the Oxa1 superfamily. The Oxa1 superfamily — consisting of GET1, EMC3, and TMCO1 — is a group of membrane protein insertases that are evolutionarily conserved and shown to play essential roles in membrane protein biogenesis (Anghel et al., 2017; Hegde & Keenan, 2022). These insertases are subunits of larger membrane complexes that carry out these insertion functions: (1) GET1 is part of the GET1/2 complex in the guided-entry of TA proteins (GET) pathway (Vilardi et al., 2011, 2014); (2) EMC3 is part of the ER membrane protein complex (EMC) (Guna et al., 2018; Jonikas et al., 2009); (3) TMCO1 is part of the GET- and EMC-like (GEL) complex (Lewis & Hegde, 2021; Sundaram et al., 2022).

Previously, our lab determined the molecular structure of the human EMC using cryogenic electron microscopy (cryo-EM) (Pleiner et al., 2020). The EMC is a nine-subunit protein complex that plays a major role in inserting of TA proteins post-translationally and the first TMD of multipass membrane proteins adopting the N_{exo} topology co-translationally (Chitwood et al., 2018; Guna et al., 2018; Jonikas et al., 2009). The localization of TA proteins has been thought to be determined by the biophysical properties of its TMD and C-terminal soluble domain that needs to be translocated to the ER lumen (Costello et al., 2017; Fry et al., 2021). Our lab has also demonstrated that the positively charged residues within the hydrophilic vestibule of the EMC function as a selectivity filter to prevent misinsertion of mitochondrial TA proteins into the ER, as their C-terminal tails are enriched for positive charges (Pleiner et al., 2023). We also found that introducing negatively charged residues into the hydrophilic vestibule of the EMC leads to increased misinsertion of mitochondrial TA proteins into the ER.

More recently, Sec61 has been shown act as a scaffold and coordinate with other protein complexes at the ER to form the multipass translocon (MPT) to facilitate the insertion of multipass membrane proteins (McGilvray et al., 2020; Smalinskaitė et al., 2022; Sundaram et al., 2022). These protein complexes include the GEL complex, the protein associated with the ER translocon (PAT) complex, and the back-of-Sec61 (BOS) complex. The GEL complex is thought to carry out the insertion function at the MPT; it consists of the Oxa1 superfamily insertase TMCO1 and its binding partner OPTI. The PAT complex includes CCDC47 and Asterix, which are thought to help chaperone and guide the TMDs to the MPT. Lastly, the BOS complex is made up of TMEM147, NOMO, and NCLN, and its function at the MPT is not well understood but it was hypothesized that it acts as a scaffold, recruiting the other components to the MPT. However, TMCO1 being the insertase in this model still does not fully explain how multipass membrane proteins of diverse topologies and biophysical properties get inserted into the ER membrane.

1.3. Membrane protein quality control at the ER

As mentioned above, as membrane proteins are synthesized in the cell by the ribosome, they need to be targeted for insertion, folding, and assembly in the ER before being trafficked to their destined location. To maintain proteostasis, misinserted, misfolded, and unassembled membrane proteins need to be rapidly and properly targeted for degradation, as their hydrophobic TMDs can lead to aggregation in the aqueous environment of the cell, resulting in cytotoxic effects. This process where aberrant, misfolded, or unassembled proteins get selected for degradation at the ER is called ER-associated degradation (ERAD) (McCracken & Brodsky, 1996; Neal et al., 2018; Werner et al., 1996).

ERAD has been extensively studied in yeast, where it's divided into three main pathways based on where degradation signal (i.e., degron) recognition occurs: ERAD-C, ERAD-L, and ERAD-M. ERAD-C branch deals aberrant proteins with degron in the cytosol, ERAD-L branch deals with proteins with degron in the ER lumen, and ERAD-M branch deals with proteins with degron within the ER membrane (P. Carvalho et al., 2006; Huyer et al., 2004; Vashist & Ng, 2004). However, the distinction between the branches of ERAD in mammals remains elusive, possibly due to the increased complexity and diversity of the mammalian proteome. Even in yeast, there is overlap between ERAD pathways as certain protein substrates can be targeted by multiple pathways, allowing them to still be degraded even if one of its ERAD pathways becomes defected or overloaded (Vembar & Brodsky, 2008).

The most conserved and well understood ERAD branch is that of the Hrd1 complex in yeast, which can work in both ERAD-L and ERAD-M (Bordallo et al., 1998; Hampton et al., 1996). The Hrd1 complex consists of the E3 ubiquitin ligase Hrd1 (SYVN1 in human), Hrd3, Der1, Usa1, and Yos9. After substrate recognition in the ER lumen or membrane, the substrates are moved across the membrane (i.e., retrotranslocation) by Hrd1, Der1, and Usa1 (P. Carvalho et al., 2010; X. Wu et al., 2020). Once exposed to the cytosol, the protein substrates are ubiquitinated by Hrd1 in a Ubc7-dependent manner, subsequently released from the ER membrane with assistance of the ATPase complex Cdc48 (VCP/p97 in human) and ultimately delivered for degradation at the proteasome (Bodnar & Rapoport, 2017; Cohen et al., 2015; Kikkert et al., 2004). Another key player for ERAD in yeast is the Doa10 complex, which has been shown to function in both ERAD-C and ERAD-M

(Swanson et al., 2001). The Doa10 complex consists of the E3 ubiquitin ligase Doa10 (TEB4/MARCHF6 in human), ubiquitin-conjugating enzymes Ubc6 and Ubc7, and the Ubc7-recruiting protein Cue1 (Hassink et al., 2005). It is thought that Doa10 directly recognizes and facilitates the retrotranslocation of its substrates with the aid of Cdc48/VCP/p97 (Habeck et al., 2015; Schmidt et al., 2020). Ubiquitination of Doa10 substrates requires Ubc6, Ubc7, and the RING domain of Doa10 (Weber et al., 2016).

Previous work in the lab explored how unassembled subunits of membrane protein complexes are differentially targeted for ERAD, specifically the GET1/2 complex in the GET pathway (Inglis et al., 2020). When there is an abundance of each subunit in the cell via overexpression, GET1 and GET2 are independently degraded by separate mechanisms. We found that unassembled GET2 misfolds and exposes its second TMD in the ER lumen. This exposed TMD acts as the degron targeting GET2 for degradation via ERAD-L. When overexpressed, GET1 also gets rapidly degraded; nevertheless, the mechanism by which it gets recognized for degradation is not fully understood.

1.4. Overview of the thesis

The first aim of my thesis is the elucidation of how factors at the ER function to ensure correct biosynthesis of integral membrane proteins. We used a combination of techniques, ranging from genetic, molecular, and cell biology, to structural biology to approach this problem. Chapter 2 describes our effort to understand how the EMC coordinates in these biological hubs at the ER with the components of the multipass translocon to insert various types of multipass membrane proteins. We performed a series of genetic screens on a diverse panel of membrane proteins to identify factors that are important for their biogenesis and found that the EMC has a central role in insertion of these proteins. We then used biochemical and structural techniques to show that another complex at the ‘multipass translocon’, the BOS complex, genetically and physically interacts with the EMC to carry out insertion function. Using cellular assays, we were also able to show that the biophysical properties of the soluble domain of these multipass membrane proteins determine their insertion pathway, requiring different sets of biogenesis factors. We proposed a model

whereby these biogenesis factors coordinate with one another to accommodate the diversity of membrane proteins in the cell that need to be inserted into the membrane.

In contrast, the second aim of my thesis focuses on degradation of membrane proteins, as described in Chapter 3. Using genetic screen, we were able to identify a novel quality control factor responsible for the degradation of orphan membrane protein subunit at the ER. Using biochemical and cell-based assays, we were able to show that this novel ERAD factor can coordinate a set of substrates possessing exposed hydrophobic domain in the ER lumen for degradation. Furthermore, we were able to link this novel factor's quality control function with its implication in human diseases involving defects in the formation of the primary cilia.

Finally, Chapter 4 concludes and summarizes the findings in this thesis, emphasizing how our work contributes to our current understanding in the field of how membrane proteins are made and degraded at the ER.

CHAPTER 2:

ROLE OF A HOLO-INSERTASE COMPLEX IN THE BIOGENESIS OF BIOPHYSICALLY DIVERSE ER MEMBRANE PROTEINS

The following chapter is adapted from Page, Nguyen, Pleiner et al., 2024 and modified according to the Caltech Thesis format.

Page, K. R., Nguyen, V. N., Pleiner, T., Tomaleri, G. P., Wang, M. L., Guna, A., Hazu, M., Wang, T.-Y., Chou, T.-F., & Voorhees, R. M. (2024). Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Molecular Cell*, *84*(17), 3302-3319.e11. https://doi.org/10.1016/j.molcel.2024.08.005

2.1. Summary

Mammalian membrane proteins perform essential physiologic functions that rely on their accurate insertion and folding at the endoplasmic reticulum (ER). Using forward and arrayed genetic screens, we systematically studied the biogenesis of a panel of membrane proteins, including several G-protein coupled receptors (GPCRs). We observed a central role for the insertase, the ER membrane protein complex (EMC), and developed a dual-guide approach to identify genetic modifiers of the EMC. We found that the back of Sec61 (BOS) complex, a component of the multipass translocon, was a physical and genetic interactor of the EMC. Functional and structural analysis of the EMC•BOS holocomplex showed that characteristics of a GPCR's soluble domain determine its biogenesis pathway. In contrast to prevailing models, no single insertase handles all substrates. We instead propose a unifying model for coordination between the EMC, multipass translocon, and Sec61 for biogenesis of diverse membrane proteins in human cells.

2.2. Introduction

Integral membrane proteins are essential in all biological systems, including mammalian cells and their pathogens. Human membrane proteins mediate a range of processes from cell-to-cell signaling to metabolite transport. Similarly, many viruses encode membrane proteins that are critical for fusion with a host cell, organization of their replication machinery, and transport of ions and small molecules (by viroporins) that enhance infectivity and morbidity (Harrison, 2008; Lenard, 2008). In order to carry out these functions, both the soluble and transmembrane domains (TMDs) require distinct charge, hydrophobicity, and length (von Heijne, 2007). The accurate insertion and folding of these topologically and biophysically diverse proteins therefore represents a major challenge in human cells. Despite the importance of this process, how cells regulate biogenesis of the full complexity of the mammalian and viral membrane proteome is not understood.

The majority of membrane proteins, destined for either the plasma membrane or secretory pathway, begin their biogenesis at the endoplasmic reticulum (ER) (Rapoport et al., 1996; Shao & Hegde, 2011). For multipass proteins, the nascent polypeptide is first co-translationally captured in the cytosol by the signal recognition particle (SRP) (Halic & Beckmann, 2005; Shan & Walter, 2005). Once at the ER, substrates must be inserted and folded into the lipid bilayer. Insertion requires two simultaneous processes: (i) transfer of the hydrophobic TMD to a membrane-spanning topology within the lipid bilayer, and (ii) translocation of an associated soluble domain across the membrane into the ER lumen. The latter of which poses an energetic barrier to insertion and is typically catalyzed by a membrane protein insertase (Guna, Hazu, et al., 2023).

The textbook model posits that the Sec61 translocation channel is the major insertase for multipass membrane proteins. It was hypothesized that its unique clam-shell architecture could accommodate all aspects of membrane protein biogenesis: axial opening creates a pore in the membrane for translocation into the ER lumen, while lateral opening would permit partitioning of a TMD into the bilayer (Van den Berg et al., 2004; Voorhees & Hegde, 2016). However, because many multipass proteins contain TMDs that cannot autonomously open the lateral gate (Enquist et al., 2009; Schorr et al., 2020), the simple

model in which each TMD is sequentially inserted by Sec61 alone cannot explain the insertion or folding of most multipass membrane proteins.

Recently, it has instead been proposed that substrates, including TMDs 2-7 of the physiologically essential family of G-protein coupled receptors (GPCRs), are inserted by a 'multipass translocon' that uses Sec61 as a structural scaffold, but does not rely on its lateral gate (McGilvray et al., 2020; Smalinskaitė et al., 2022; Sundaram et al., 2022). The multipass translocon is a dynamic, 8-subunit complex that includes the GET and EMC like (GEL), back of Sec61 (BOS), and PAT complexes (Figure S2.1A). The GEL complex—composed of the Oxa1 superfamily insertase, TMCO1 and its binding partner, OPTI—serves as the dedicated insertase of the multipass translocon. The PAT complex, containing Asterix and CCDC47, has two proposed roles: Asterix chaperones hydrophilic TMDs (Chitwood & Hegde, 2020; Meacock et al., 2002); while CCDC47 is proposed to engage the ribosome and guide TMDs to the multipass translocon (Smalinskaitė et al., 2022; Sundaram et al., 2022). Finally, the function of the BOS complex remains unknown, but is thought to act as scaffold for recruitment of the remaining multipass components. Together, these factors create a protected lipid cavity behind Sec61 to facilitate multipass membrane protein insertion and folding.

Earlier work also established that an additional insertase, the ER membrane protein complex (EMC), was required for biogenesis of many multipass membrane proteins. In mammals, the EMC is a nine-subunit complex that functions as both an insertase and chaperone (Z. Chen et al., 2023; Guna et al., 2018; Shurtleff et al., 2018; Tian et al., 2019). In addition to post-translational insertion of a subset of tail anchored (TA) proteins, the EMC also co-translationally inserts the first TMD of many GPCRs and other multipass membrane proteins that position their N-terminus in the ER lumen or extracellular environment (i.e. N_{exo} topology) (Chitwood et al., 2018). Indeed, expression of rhodopsin, which does not rely on the lateral gate of Sec61 for insertion, is EMC dependent (Satoh et al., 2015). However, the function of all nine of EMC's subunits, particularly those that form its large lumenal domain and are dispensable for insertion, is not known.

Structures of the yeast and human EMC show that substrate TMDs are inserted into the bilayer via a positively charged hydrophilic groove, through which the substrate's soluble N-terminus must also translocate (Bai et al., 2020; Miller-Vedam et al., 2020; O'Donnell et al., 2020; Pleiner et al., 2020). The positioning of positively charged residues within the membrane is a conserved feature of the Oxa1 superfamily of insertases and is required for their activity (Borowska et al., 2015; Kumazaki et al., 2014; McDowell et al., 2020; Pleiner et al., 2020, 2023). It is likely that multipass substrates are therefore directly delivered by SRP to the EMC (H. Wu & Hegde, 2023), leaving the EMC to act upstream of Sec61 and the multipass translocon.

However, this model leaves several central unanswered questions for how human and viral membrane proteins are accommodated by the biogenesis and quality control machinery in the ER. First, whether or how the EMC coordinates with the multipass translocon during multipass biogenesis is not known. Second, if the EMC is responsible for insertion of the first N_{exo} TMD of many membrane proteins (including GPCRs), how substrates are transferred between the EMC, Sec61, and the multipass translocon is not clear. Finally, a systematic analysis of the substrate specificity and cooperation of the suite of biogenesis factors in the ER to ensure insertion and folding of their diverse clients, has not been explored.

2.3. Results

Systematic analysis of membrane protein biogenesis

With the goal of unbiasedly identifying factors required for biogenesis of diverse membrane proteins, we selected a panel of substrates with distinct topologies, biophysical properties, and number of TMDs (Figure 2.1, S2.1B). We included the human GPCR AGTR2, and the viral ORF3a and M from SARS-CoV-2, as well as the post-translationally targeted TA protein, Sec61β. AGTR2 contains TMDs with varying hydrophobicities, as well as small cytosolic and extracellular loops, and a neutrally charged N-terminal soluble domain that must be translocated across the ER membrane.

The viral proteins ORF3a and M from SARS-CoV-2 (Figure 2.1C,D) both adopt an identical three-TMD topology, but their TMDs have distinct biophysical properties and insertion propensities (Dolan et al., 2022; Kern et al., 2021). Further, ORF3a and M have soluble N-termini of different length (42 vs 19 amino acids) and charge (0 vs -2), which we hypothesized could alter the suite of host factors required for their biogenesis. The ability to query two topologically related proteins that are also innocuous upon overexpression was a unique advantage of using viral substrates.

Finally, as a control we included the TA protein Sec61β (Figure 2.1B). TAs contain a single TMD within ~35 amino acids (AA) of their C-terminus, and thus cannot access the co-translational biogenesis pathways typically utilized by multipass proteins (Guna, Hazu, et al., 2023; Kutay et al., 1993). The targeting and insertion of Sec61β has been extensively characterized biochemically, and is one of the few TA proteins known to rely equally on the EMC and GET biogenesis pathways, and therefore serves as a comparison for machinery required for biogenesis of multipass vs singlepass membrane proteins (Guna et al., 2018; Mateja et al., 2009; Stefanovic & Hegde, 2007).

Human K562 cell lines stably expressing these four substrates were generated in which each substrate was expressed along with a translation normalization marker and the CRISPR inhibition (CRISPRi) machinery (Figure 2.1, S2.1C) (Gilbert et al., 2014; Horlbeck et al., 2016). Previous experiments have established that depletion of factors required for targeting, insertion, or folding leads to degradation of the reporters by the ubiquitin-proteasome pathway leading to a decrease in fluorescence (Guna et al., 2018; Pleiner et al., 2020). Conversely, disruption of protein quality control leads to substrate and an increase in fluorescence. Therefore, following transduction with a genome-wide single guide RNA (sgRNA) library, cells that displayed altered substrate levels relative to the normalization control were sorted using fluorescence activated cell sorting (FACS; Figure S2.1C). Deep sequencing of the sgRNAs enriched in both the low and high fluorescing cells was used to identify putative biogenesis and quality control factors, respectively.

In addition to factors related to the physiologic function of each substrate (e.g. lysosomal and vesicular trafficking factors for ORF3a; Figure S2.1D), we identified machinery that differentially affects their biogenesis and degradation (Figure 2.1, Table S2.2, S2.3). For example, amongst the identified quality control factors, loss of the ER-resident E3 ubiquitin ligase HRD1 was found to stabilize only AGTR2, while the more general ERAD component FAF2 was identified in all four screens (Kikkert et al., 2004; Olzmann, Richter, et al., 2013). On the biogenesis side, the GET pathway components had the most pronounced effect on the TA control (Sec61β), consistent with their role in post-translational insertion (Mariappan et al., 2011; Schuldiner et al., 2005, 2008; Vilardi et al., 2011). Conversely, both the translocon associated protein (TRAP) complex and the members of the multipass translocon were only significant hits for biogenesis of AGTR2 and ORF3a, but not Sec61β or M. Interestingly however, the central insertase of the multipass translocon, the GEL complex, was not a significant hit in any screen, despite near-complete depletion under these conditions (Figure S2.1E). While loss of SEC61A affected all four reporters, given the extensive data for Sec61 independent biogenesis of TA proteins, its effect on Sec61β is likely due to an assembly rather than insertion defect. Therefore, one of the only universally identified biogenesis factor was the EMC, which is consistent with its established role in TA and N_{exo} TMD insertion (Chitwood et al., 2018; Guna et al., 2018).

Distinct pathways for biogenesis and quality control of diverse substrates

To delineate how the identified factors affected the biogenesis of a broader range of substrates, we generated a panel of 13 membrane protein reporters (Figure 2.2A). We included substrates with varying numbers of TMDs of distinct lengths and hydrophobicity, as well as those that differ in the structure of the intervening soluble domains. In the panel were multipass proteins in which the N-terminus is translocated across the ER membrane (i.e., N_{exo} topology: GPCRs, ORF3a, and M); multipass proteins with their N-terminus in the cytosol (N_{cyt} topology: TRAM2, EAAT1, GET2, and YIPF1); and single-spanning (Type II: ASGR1; Type I: TRAPα); and TA proteins (SQS, VAMP2, and Sec61β). To allow for direct comparison, all reporters contained a full-length GFP with the exception of

Sec61β, which required use of the split GFP approach to avoid mislocalization (Guna et al., 2022; Guna, Page, et al., 2023; Inglis et al., 2020).

Using an arrayed screen, we tested the depletion of nine factors that represented each of the major biogenesis (EMC, BOS, GEL, PAT, TRAP, GET, UGGT1) and quality control (HRD1, FAF2) complexes identified in the screens (Figure 2.2B). Critically, in most cases knockdown of a single subunit is sufficient to deplete the entire complex (Figure S2.1E) (Colombo et al., 2016; Dettmer et al., 2010; Pleiner et al., 2021; Volkmar et al., 2019). For these experiments, we used the near-diploid human RPE1 cell-line because we postulated that redundancy and compensation would be more pronounced in an aneuploid cell line.

Though some cell-type dependent differences are observed, in general, the arrayed screen both validates the genome-wide screens and suggests substrate features that correlate with biogenesis pathway. As expected, the clearest delineation between TAs and other membrane protein substrates is dependence on the post-translational pathway for targeting and insertion to the ER. Consistent with biochemical data, we find that both our forward and arrayed screens show that TA biogenesis via either the GET or EMC pathway is determined by TMD hydrophobicity (Guna & Hegde, 2018; Rao et al., 2016; Shao et al., 2017; F. Wang et al., 2010). Loss of the GET pathway insertase GET1/2 also appeared to have a small effect on biogenesis of several multipass substrates (e.g., the GPCR AGTR2 and EAAT1) in both the genome-wide and arrayed screens.

In contrast, several factors specific to co-translationally targeted substrates. For example, depletion of the Sec61 associated chaperone, TRAP affects multipass but not TA substrates (Figure 2.2C) (Gemmer et al., 2023; Shao & Hegde, 2011). Though an effect of TRAPβ depletion was observed for the single spanning protein TRAPα, this may be due to an assembly rather than a biogenesis defect. While components of the multipass translocon—including the BOS, PAT, and GEL complexes—are required for substrates with multiple TMDs, they are not required for any of the single spanning membrane proteins. However, our data suggest that the function of the multipass translocon differs across cell-types, because GEL complex dependence for AGTR2 was only observed in RPE1 cells (Figure

2.2B) but not K562s, despite efficient knockdown (Figure S2.1E). It is possible this reflects cell-type specific changes in expression and partial redundancy and/or compensation of biogenesis factors in the ER. For example, we consistently observe that depletion of the EMC leads to a compensatory increase in TMCO1 levels (Figure S2.1E-F).

Critically however, even in this relatively small panel, it is clear that these multipass-specific factors are not required for the biogenesis of all multispanning proteins. For example, we observe variability in dependence on the BOS, GEL, and PAT complex amongst the three related GPCRs tested in the arrayed screen. These data suggest that beyond simply topology, biophysical properties of the TMDs and surrounding regions determine biogenesis pathway. This observation sets the stage for an in-depth study of the relationship between substrate properties and biogenesis requirements.

Identification of genetic modifiers of the EMC genome wide

One commonality across many substrates was a dependence on the EMC for biogenesis. We therefore wondered how the EMC cooperates with other ER factors for insertion and folding of its multipass substrates like GPCRs. Indeed, immunoprecipitation of the human EMC from native membranes suggests it associates with a myriad of ER-resident chaperones (e.g. CNX) (Bergeron et al., 1994), biogenesis machinery (e.g. SRP receptor, glycosylation machinery, and components of the multipass translocon) (Kelleher & Gilmore, 2006; McGilvray et al., 2020; Shan & Walter, 2005), and quality control factors (e.g. p97) (Meyer et al., 2012) (Figure 2.3A, Table S2.4). Recruitment of factors required for the folding and surveillance of nascent proteins to the EMC would ensure that clients are immediately captured for maturation or degradation upon integration into the ER. These results establish the EMC as a central organizing factor for membrane protein biogenesis and quality control within the ER membrane.

To delineate which of these associated factors are physiologically important, we developed a dual-guide CRISPRi approach to systematically identify genetic modifiers of the EMC genome-wide (Figure 2.3B) (Guna, Page, et al., 2023). Briefly, we generated a library that

expresses two sgRNAs on a single plasmid: (1) a genetic anchor guide, targeting the core subunit, EMC2, which when depleted results in loss of the remaining EMC subunits (Pleiner et al., 2021; Volkmar et al., 2019); and (2) a second randomized guide, targeting all open reading frames using the CRISPRi-v2 library (Horlbeck et al., 2016). Transduction of this dual library allows the acute knockdown of both the genetic anchor and a second randomized gene simultaneously in each cell, and is compatible with a standard CRISPRi FACS-based screening and analysis pipeline.

Comparison of the hits identified in the EMC genetic anchor screen with those from a control screen performed with a non-targeting 'anchor' guide library results in three categories of factors (Figure 2.3C, S2.2, Table S2.1). First, are those that have diminished phenotypes when combined with EMC depletion, indicative of an epistatic relationship and potentially a shared pathway with the EMC. Second, are factors that have enhanced phenotypes upon loss of the EMC, likely including factors that represent parallel or partially redundant pathways. Third, are factors that act independently of the EMC, and therefore show no change in phenotype with or without the EMC.

Interrogating both the EMC-dependent TA protein Sec61β and the GPCR AGTR2 allowed us to delineate EMC co-factors that function to support its post vs co-translational biogenesis roles (Figure 2.3C and S2.2). Validating this approach, all EMC subunits have a diminished effect in the EMC2 knockdown background for both TA and GPCR biogenesis. Conversely, the phenotype of known parallel pathways for TA insertion, including the GET components, are enhanced by EMC depletion, particularly in the TA screen (Figure S2.2C-D). Finally, several quality control factors, such as HRD1 and FAF2 exhibit EMC independent effects, suggesting their function may be agnostic to the insertion pathway.

Direct comparison of the biogenesis factors identified in the TA vs the GPCR genetic modifier screens suggest that many more factors are cooperating with the EMC in insertion and folding of multipass membrane proteins than of TAs. We identified many components of the multipass translocon (e.g. BOS and PAT) as epistatic with the EMC for AGTR2

biogenesis (Figure 2.3C-F and S2A-B). To test whether other GPCRs display a similar epistatic dependence, we performed an arrayed screen with dual guides targeting the BOS, PAT, or GEL complexes alone or in combination with a guide targeting the EMC. We included the GPCRs AGTR2, OPRK1 and ADRA1A and the type II membrane protein ASGR1 (Figure 2.3D-F). For the GPCRs but not for ASGR1, EMC displays an epistatic relationship with the BOS, PAT, and GEL complexes, which also co-purify with the EMC under conditions where all components are expressed at endogenous levels (Figure 2.4A-B, S2.3A). These results suggesting both a genetic and physical interaction between these biogenesis complexes.

One potential trivial explanation for genetic epistasis between the EMC and BOS complex is that, following insertion of TMD1 of a GPRC by the EMC, the multipass translocon is responsible for inserting the remaining downstream TMDs. However, we found that addition of a signal sequence or signal anchor to the N-terminus of AGTR2 or the GPCR ADRA1A, which allows them to bypass the EMC and utilize Sec61 for insertion of its first TMD (Chitwood et al., 2018), markedly rescues its dependence on the BOS complex for biogenesis (Figure 2.4C and S3D-E). We postulated that there may be an additional role of the BOS complex as a co-factor of the EMC, beyond its previously reported function as part of the multipass translocon.

The BOS complex is a direct physical interactor of the EMC

Though the BOS and PAT complexes co-immunoprecipitated with the EMC, we first sought to confirm that this reflected a direct physical interaction. To do this we incubated intact cells under conditions where the EMC (Figure S2.3B) and the BOS complex (Figure S2.3C) components are present at endogenous levels with the amine-reactive chemical crosslinker DSP. DSP has a length of ~12Å, such that only factors within close proximity can be covalently crosslinked. The resulting crosslinked species were immunoprecipitated under denaturing conditions in SDS, where we found that subunits of the BOS (NOMO) and PAT (CCDC47) complexes specifically immunoprecipitated with EMC7 under conditions in which other EMC subunits are markedly depleted (Figure 2.4D). NOMO did

not crosslink to EMC5, confirming that the crosslinking and immunoprecipitation conditions were specific.

To confirm that this result did not reflect a long-range interaction between the flexible Ig-like domains of NOMO and the EMC, we performed similar experiments using affinity tags on the other BOS complex subunits that are either more rigid (NCLN) or fully embedded in the bilayer (TMEM147). We found that immunoprecipitation of TMEM147 and NCLN after chemical crosslinking specifically recovered EMC3,7, and 10, but not EMC2 (Figure 2.4E-F, S3F). We therefore concluded that the BOS complex is a direct physical interactor of the EMC in native membranes.

To better understand the interaction between the EMC and the BOS complex, we sought to determine a structure of the 12-subunit holocomplex purified from human cells. Though we have shown that the BOS complex and the EMC interact without exogenous stabilization (Figure 2.4A-B), to increase their local concentration and thereby enable structural analysis, we introduced an ~40 amino acid linker between TMEM147 and EMC2. By using an extremely long and flexible linker, we avoid artificially stabilizing a non-physiologic interaction. Indeed, modeling suggested that this >100 Å linker would not preclude interaction of the EMC and BOS complexes in any orientation or arrangement. Using single particle cryo-EM, we determined the structure of the EMC•BOS holocomplex to an overall resolution of 6.6 Å (Figure 2.5A, S2.4A).

While we could unambiguously fit existing models of the EMC into the holocomplex density, the resolution was insufficient for *de novo* building of the BOS complex. Therefore, using an affinity tag on TMEM147 and more stringent conditions, we purified the BOS complex and determined two structures: (1) BOS(fNOMO), using the full-length NOMO, including its 12 endogenous Ig-like domains (~8 Å resolution) (Figure S2.4B, S2.5A); and (2) BOS(tNOMO) in which we truncated all but the last 3 Ig-like domains of NOMO (3.7 Å resolution) (Figure 2.5B, S2.4C). Notably, the truncated NOMO resulted in a higher purity samples and improved monodispersity upon freezing on EM grids. In both structures, only the 2-3 terminal Ig-like domains of NOMO interact with the lumenal domain of NCLN, suggesting that the remaining Ig-like domains are dynamic. Indeed, structures of

the multipass translocon (McGilvray et al., 2020; Smalinskaitė et al., 2022) could not unambiguously assign density for any of the Ig-like domains. Superposition of the density of the full-length NOMO and truncated NOMO in both complexes suggest that they are qualitatively identical, validating use of the truncated complex for high-resolution model building (Figure S2.5A).

We used this model to unambiguously fit the EM density in the EMC•BOS holocomplex. There are two primary interactions: an intramembrane interface, composed of EMC5 and TMEM147, and a lumenal interface between NCLN and EMC1 totaling a buried surface area of ~855 $Å^2$. Notably, the intermembrane interface is enriched for conserved hydrophobic residues (Figure 2.5C, S2.5B), suggesting that this interface may be the more important than that in the lumen. Conversely, we hypothesize that the absence of crosslinks previously observed between EMC5 and NOMO (Figure 2.4D) can be explained by the lack of primary amines in the membrane-embedded EMC5 subunit. Finally, the interaction surface with the BOS complex is distinct compared to that reported for the chaperone-binding mode of EMC (Figure S2.5C) (Z. Chen et al., 2023).

To validate the holocomplex structure, we site-specifically introduced a photoactivatable amino acid to the lumenal domain of NCLN that directly abuts EMC1. Following UV irradiation of intact cells, such that interactions are captured prior to disruption of the ER membrane, we observed a UV-dependent crosslink between NCLN and EMC1 (Figure S2.5F). These results further support a direct physical interaction between the EMC and BOS complexes, and are consistent with their observed relative orientation in the structure.

Comparison with structures of the multipass translocon bound at Sec61 suggest that BOS binding to the EMC and Sec61 is mutually exclusive (Figure 2.5D) (McGilvray et al., 2020). Conversely, binding of BOS to the GEL and PAT complexes would all be compatible with interaction at the EMC (Figure S2.5D, S2.5G). Indeed, we observed co-immunoprecipitation of both PAT and GEL complex subunits with the EMC by both quantitative proteomics and western blotting, and observed that their interaction with the EMC is independent of EMC-BOS interaction (Figure 2.3A, 2.4A, S2.5E). Further, we

verified using chemical crosslinking that the PAT complex is a physical interactor of the EMC and that its interaction is EMC7-dependent (Figure 2.4D,G). These data therefore suggest that the interaction between the EMC and the PAT complex is highly specific, because loss of a this peripheral subunit (EMC7) abolishes CCDC47 and Asterix binding. Cumulatively, these data suggest that components of the multipass translocon, including the BOS, PAT, and GEL complexes, are also assembled at the EMC in a mutually exclusive manner to their binding to Sec61.

Biophysical properties of substrate soluble domains dictate biogenesis pathway

Having established both a genetic and physical interaction between the EMC and BOS complex, we sought to determine the function of the BOS complex in the biogenesis of EMC-dependent substrates. Analysis of the genome-wide and arrayed screens suggested patterns in the substrate features that confer dependence on the BOS complex. First, while dispensible to singlepass substrates, not all multipass proteins are equally dependent on the BOS complex. We reasoned that if the EMC is required for insertion of the first TMD of these multipass substrates, dependence on the BOS complex may be conferred by properties of this TMD and its surrounding sequence.

Second, M and ORF3a show differential dependence on the BOS complex. The primary difference between these topologically identical proteins, is the length and charge of the N-terminus of ORF3a (Figure 2.1C-D). Indeed, AGTR2 and ADRA1A, which also displayed a clear reliance on the BOS complex when utilizing the EMC (Figure 2.4C, S2.3D-E), contain atypically positive/neutral N-termini, for GPCRs lacking signal sequences(Wallin & von Heijne, 1995). Because the EMC uses a positively charged hydrophilic vestibule for insertion, this limits integration of substrates containing positively charged soluble domains (Pleiner et al., 2023; H. Wu & Hegde, 2023). We therefore hypothesized that when the first TMD of a multipass protein is inserted by the EMC, the charge of the N-terminal soluble domain confers dependence on the BOS complex.

To test this, we determined the dependence of a panel of GPCRs on the EMC, BOS, and PAT complexes for biogenesis (Figure 2.6A, S2.6A-B). We found that those substrates with more positively charged N-termini displayed a stronger dependence on the BOS complex (Figure 2.6A), but observed no connection between charge and EMC dependence (Figure S2.6A). However, this correlation was imperfect, suggesting that features beyond simply charge may play a role.

To interrogate the relationship between N-terminal charge and BOS complex dependence directly, we chose two representative GPCRs, OPRK1 and ADRA1A. Both substrates display strong EMC dependence, but contain distinct N-terminal total charge (-6 vs 0) (Figure 2.6B). ADRA1A, like AGTR2, only requires the BOS complex when utilizing the EMC for insertion of its first TMD (Figure 2.4C, 2.6C, S2.6C-D). In contrast, while addition of a signal sequence to the N-terminus of OPRK1 rescues the effect of EMC depletion, it has minimal effect on its BOS complex dependence. However, an OPRK1 mutant with three additional positive charges within its N-terminus no longer relies on the BOS complex when inserted by Sec61. We therefore concluded that by studying the insertion of these three substrates we could precisely test the effects of the soluble N-terminus of a GPCR on its biogenesis.

Given earlier data suggesting that neutral or positively charged soluble domains are less efficiently translocated by the EMC, we wondered if these substrates might be more likely to rely on alternative insertase pathways. We therefore tested whether GPCRs displayed differential dependence on the GEL complex (i.e. TMCO1) and the insertase activity of Sec61 depending on their N-terminal charge. To specifically query the role of the insertase activity of Sec61, we used the inhibitor Apratoxin A, which prevents opening of the lateral gate (Paatero et al., 2016; Thornburg et al., 2013). We found that substrates containing a neutral N-terminus had increased dependence on the GEL complex and the lateral gate of Sec61. Importantly, while wild type OPRK1 did not depend on GEL or Sec61, introduction of positive charges to its N-terminus increased reliance on both alternative insertases (Figure 2.6D-E, S2.6D-E). Moreover, we observed an even greater dependence of the neutrally charged N-terminal OPRK1 mutant on GEL and Sec61 when both insertases were

impaired simultaneously, both in cells and in vitro (Figure S2.7A,C). We further found that changes to the N-terminal charge of the GPCRs ADRB1 and GPR15 also alters their dependence on the GEL and BOS complexes (Figure 2.6I, S2.7B). We therefore concluded that multipass substrates that cannot be efficiently inserted by the EMC due to charge in their soluble domains instead rely on alternative, partially redundant, pathways for biogenesis.

Additionally, we tested if these substrates might have increased dependence on the methionine-rich loops of the EMC that interact with substrates in the cytosol during their passage into the membrane (Pleiner et al., 2020, 2023). Indeed, GPCRs containing positively charged N-termini have increased dependence on the C-terminus of EMC7 that directly binds substrate TMDs below the hydrophilic vestibule of the EMC (Figure 2.6F-G, S2.6F). An increased reliance on these cytosolic domains of the EMC would be consistent with an increased dwell time at the EMC for neutral or positively charged substrates that are not efficiently translocated through the positively charged vestibule.

Finally, though we observed a correlation between the role of total N-terminal charge and insertase selection, it is clear that charge cannot be the only factor that dictates biogenesis pathway (e.g. Fig. 2.6A). We tested whether other features such as length and secondary structure could alter dependence on each insertase. Indeed, we found that modifying the length of the GPCR CHRM1's N-terminus and altering the secondary structure of CCKAR N-terminus both shift BOS dependence (Fig. S2.6G-H). These data suggest that multiple biophysical properties contribute to the biogenesis pathway required for each membrane protein at the ER.

2.4. Discussion

Topology and biophysical properties dictate biogenesis pathway

By systematically studying the biogenesis of membrane proteins using a series of forward and arrayed genetic screens, we have begun to dissect the substrate specificity of the suite

of biogenesis factors in the ER (Figure 2.1,2.2). Multiple insertases beyond Sec61, including the Oxa1 superfamily members GET1/2, EMC, and the GEL complex, are required for insertion of many single spanning and multipass membrane proteins. As expected, the post-translational insertase GET1/2 had the most pronounced role in TA biogenesis, but its depletion also affected some multipass substrates. This result could be indirect, but it is also possible that under some circumstances, GET1/2 can play a broader role in membrane protein biogenesis at the ER. However, even amongst proteins of identical topology, we observed distinct biogenesis requirements. Therefore, our data suggests that both topology and the biophysical properties of the substrate must determine the biogenesis pathway.

Earlier work has established that the hydrophobicity of a substrate TMD dictates biogenesis requirements, in some cases requiring unique insertion machinery (i.e. for TAs) or chaperones (i.e. Asterix) (Chitwood & Hegde, 2020; Guna, Hazu, et al., 2023; Meacock et al., 2002). Our data suggests that it is not only the TMD that dictates biogenesis pathway, but also features of the associated soluble domains, including charge, length, and secondary structure. As a result, sequences distant from the TMD, in some cases as far as 50 AAs away, have a profound effect on both insertion efficiency and biogenesis pathway into the ER.

The EMC is a central organizing hub for biogenesis and quality control

Having delineated roles for individual factors, we sought to define how these components cooperate to achieve insertion and folding of diverse membrane proteins. The EMC emerged as a central feature of the protein biogenesis and quality control machinery in the ER. We observed coupling of the EMC with chaperones, post-translational modification factors, and quality control machinery, which would ensure that nascent substrates are immediately captured for modification, folding, and triage upon integration into the ER. This type of local recruitment of auxiliary factors to the site of protein insertion is analogous to that observed for Sec61 (Gemmer et al., 2023; Görlich et al., 1992; Hartmann et al., 1993; Jaskolowski et al., 2023; Kalies et al., 1998; Voigt et al., 1996). Further, direct

recruitment of the multipass translocon components to the EMC facilitates organization of ‘biogenesis hubs’ in the ER membrane. The observation that at least some of these factors are stabilized by interaction with EMC7 (Figure 2.4G) provides one potential explanation for the conservation of the large soluble lumenal and cytoplasmic domains of the EMC: recruitment and retention of auxiliary factors to the site of membrane integration.

A working model for multipass membrane protein insertion

The organization of biogenesis machinery into local and dynamic hubs within the ER provides a working model for the insertion and folding of complex multipass membrane proteins. A ribosome nascent chain complex is delivered to the ER by the signal recognition particle (SRP), where substrates first probe the hydrophilic groove of the EMC. Though the molecular details of handover from SRP to the EMC are not yet precisely defined, SRP receptor subunits were recovered in native co-immunoprecipitation of the EMC from ER membranes, suggesting one potential mechanism for recruitment to the EMC (Figure 2.3A). Models consistent with a ‘first refusal’ of substrates by the EMC, best explain data showing the EMC can enforce the correct folding of multipass substrates containing positively charged extracellular domains (Pleiner et al., 2023). Our observation that the PAT complex directly interacts with the EMC, including CCDC47, which binds the ribosome, provides a mechanism for transiently stabilizing ribosome nascent chain complexes at the EMC.

Substrates in which the first TMD can be efficiently inserted by the EMC, such as those with negatively charged and short N-terminal soluble domains, passage through its hydrophilic vestibule into the bilayer. In contrast, substrates that are poorly inserted by the have a longer dwell time at the cytosolic vestibule of the EMC. This is consistent with their increased reliance on the cytosolic C-terminus of EMC7 that contains several conserved hydrophobic residues, previously shown to directly interact with substrates in the cytosol (Pleiner et al., 2023). These data are support the model that the rate limiting step for insertion is translocation of the N-terminal soluble domain through the hydrophilic vestibule of the EMC. Features of the substrate N-terminus combined with the biophysical

properties and architecture of the EMC therefore together dictate the energetic barrier for translocation into the ER lumen.

Those TMDs that are not immediately inserted by the EMC are shuttled to alternative insertases including the GEL complex (i.e., TMCO1) and in some cases Sec61. We hypothesize that handover of those substrates between insertases is facilitated by recruitment of the BOS, GEL, and PAT complexes to the EMC. Indeed, these substrates appear to primarily require the activity of the multipass translocon when using the EMC for insertion of their first TMDs (Figure 2.6C). Charge repulsion between the EMC's positively charged hydrophilic vestibule and clients' N-terminal domains would increase dwell time and facilitate the engagement of multipass translocon factors and transfer to TMCO1 or Sec61.

Structures of the EMC•BOS holocomplex are most consistent with a model in which interactions between biogenesis machineries are dynamic. Rather than a rigid holocomplex that includes both the EMC and Sec61, a dynamic assembly may serve primarily to increase the local concentration of multiple insertases to accommodate diverse substrates as they passage into the bilayer. This would be explain why recruitment of the BOS complex to the EMC is mutually exclusive to its binding to Sec61 (Figure 2.5D), which would potentially facilitate handover between EMC and Sec61. Certainly, recruitment of the multipass translocon components to the EMC provides a putative mechanism for substrates transfer to the back of Sec61 following insertion of their first TMD by the EMC.

A central role for Oxa1 superfamily of insertases throughout all kingdoms of life

Based on differences in their biophysical properties (e.g., charge and size of the hydrophilic vestibule and limitations on translocation of some soluble domains), it is likely that insertases display partial substrate preferences (Figure 2.7A). However, it is clear that Oxa1 superfamily insertases such as the EMC and GEL complex are partially redundant, and loss of any one results in compensation in human cells (Figure S2.1F). Remarkably, recruitment

of the core insertase subunits EMC3/6 to the inner mitochondrial membrane was sufficient to rescue loss of Oxa1 in yeast (Güngör et al., 2022).

This redundancy between Oxa1 superfamily insertases provides a unifying model for membrane protein biogenesis across all kingdoms of life. Superficially, the multipass translocon components appear to be a metazoan specific adaptation, without homologs in fungi and bacteria. But even in humans, while the EMC is strictly required, the GEL insertase complex is not essential, and inactivating mutations in the GEL complex subunits are not under negative selection in healthy adults (S. Chen et al., 2024). In this case, upregulation of the EMC or potentially even GET1/2 may be sufficient to ensure efficient insertion of all required membrane protein substrates. Interestingly, in fungi, which may rely primarily on the EMC for multipass insertion, the EMC is missing one of its soluble subunits (i.e., EMC8/9) (Jonikas et al., 2009). The lack of EMC8/9 would allow the EMC to sit closer to the ribosome, potentially supplanting the GEL complex in the multipass translocon. In bacteria an analogous requirement for Oxa1 superfamily insertases has been reported: YidC is recruited to the ribosome to cooperate with SecYEG in insertion (Sachelaru et al., 2013; Wickles et al., 2014).

These observations however raise questions as to why the biogenesis machinery has expanded from bacteria to mammals. Though the increased size of the membrane proteome may contribute, it is more likely that a decrease in error tolerance in multicellular organisms is the driver for evolution of more complex biogenesis requirements. Unlike a bacterium or yeast cell, mammals rely on many post-mitotic cells that must persist for the lifespan of the organism. The risk of cytotoxicity from misfolding of nascent proteins requires more stringent mechanisms to protect cellular proteostasis. As a result, these additional insertases increase the efficiency of membrane protein insertion and folding such that it can occur even in the face of robust competing degradation pathways.

2.5. Acknowledgements

We thank Songye Chen and Oliver Clarke for technical assistance and T. Stevens and all members of the Voorhees lab for thoughtful discussion. We thank the Caltech Flow Cytometry facility, the Caltech Cryo-EM facility, and the Caltech Millard and Muriel Jacobs Genetics and Genomics Laboratory. We thank Ville Paavilainen and Kerry McPhail for Apratoxin A. Cryo-electron microscopy was performed in the Beckman Institute Center for TEM at Caltech, and data was processed using the Caltech High Performance Cluster, supported by a grant from the Gordon and Betty Moore Foundation.

This work was supported by: the Heritage Medical Research Institute (RMV), the Howard Hughes Medical Institute (RMV), the NIH's National Institute Of General Medical Sciences DP2GM137412 (RMV), the Burrough's Wellcome Fund (RMV), an NSF-CAREER award 2145029 (RMV), a Human Frontier Science Program Fellowship 2019L/LT000858 (AG), the Deutsche Forschungsgemeinschaft (TP), the Tianqiao and Chrissy Chen Institute (TP), a Rosen Family fellowship (KRP), and an Arie Jan Haagen-Smit Fellowship (KRP).

Author contributions: Conceptualization, K.R.P, T.P, and R.M.V.; Investigation, K.R.P., V.N.N., T.P., G.P.T., A.G., T.W.; Resources M.L.W., M.H., T.C.; Writing- Original Draft and Visualization, K.R.P., V.N.N., T.P., R.M.V.; Writing- Review & Editing, all authors; Supervision and Project Administration, R.M.V.

Declaration of interests: RMV is a consultant and equity holder, and GPT is a current employee of Gate Bioscience. The authors have no additional competing financial interests.

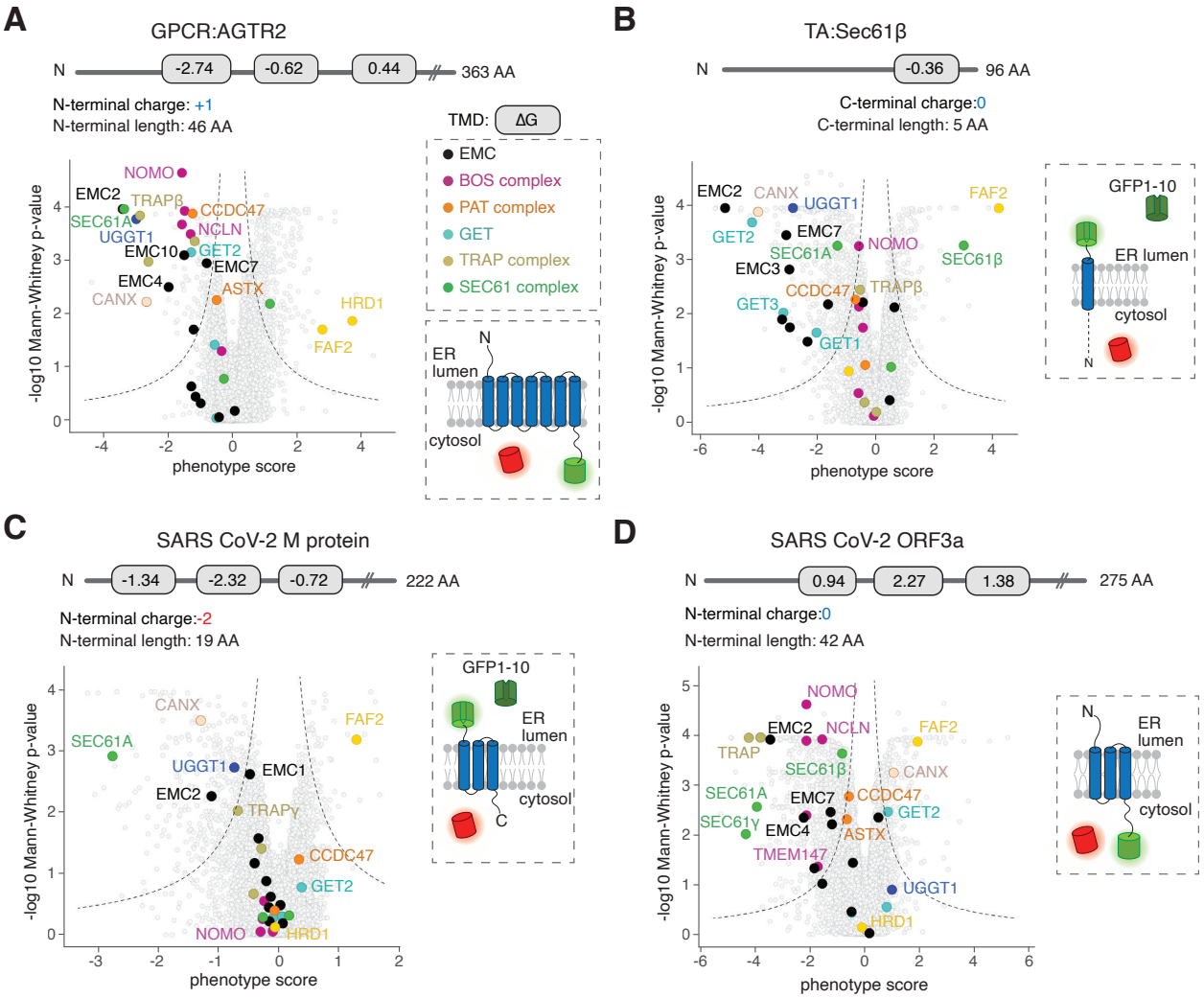

Figure 2.1. Genome-wide CRISPRi screens to systematically query biogenesis factors for diverse ER substrates.

(A) (Top) Schematic of the GPCR AGTR2. The gray rectangles indicates TMDs and their corresponding ΔG value (Hessa et al., 2007). (Right) Topology of the AGTR2 fluorescent reporter used in the genome-wide CRISPRi screen. (Bottom) Volcano plot of GFP:RFP phenotype for the three strongest sgRNAs versus Mann-Whitney p-values from two independent replicates of a genome-wide CRISPRi screen using OMP25-GFP11. Individual genes are displayed in gray, and specific factors that increase or decrease OMP25 mitochondrial integration are highlighted and labeled. Genes that fall outside the indicated dashed lines represent statistically significant hits. **(B)** As in (A), for the tail-anchored protein Sec61β. Here, The GFP11 sequence is appended to the C-terminal of

Sec61β. Upon TA insertion into the ER, GFP11 complements with the GFP1-10 independently localized to the ER lumen, resulting in fluorescence. **(C)** As in (B) for the SARS-CoV-2 M protein. **(D)** As in (A) for the SARS-CoV-2 ORF3a protein.

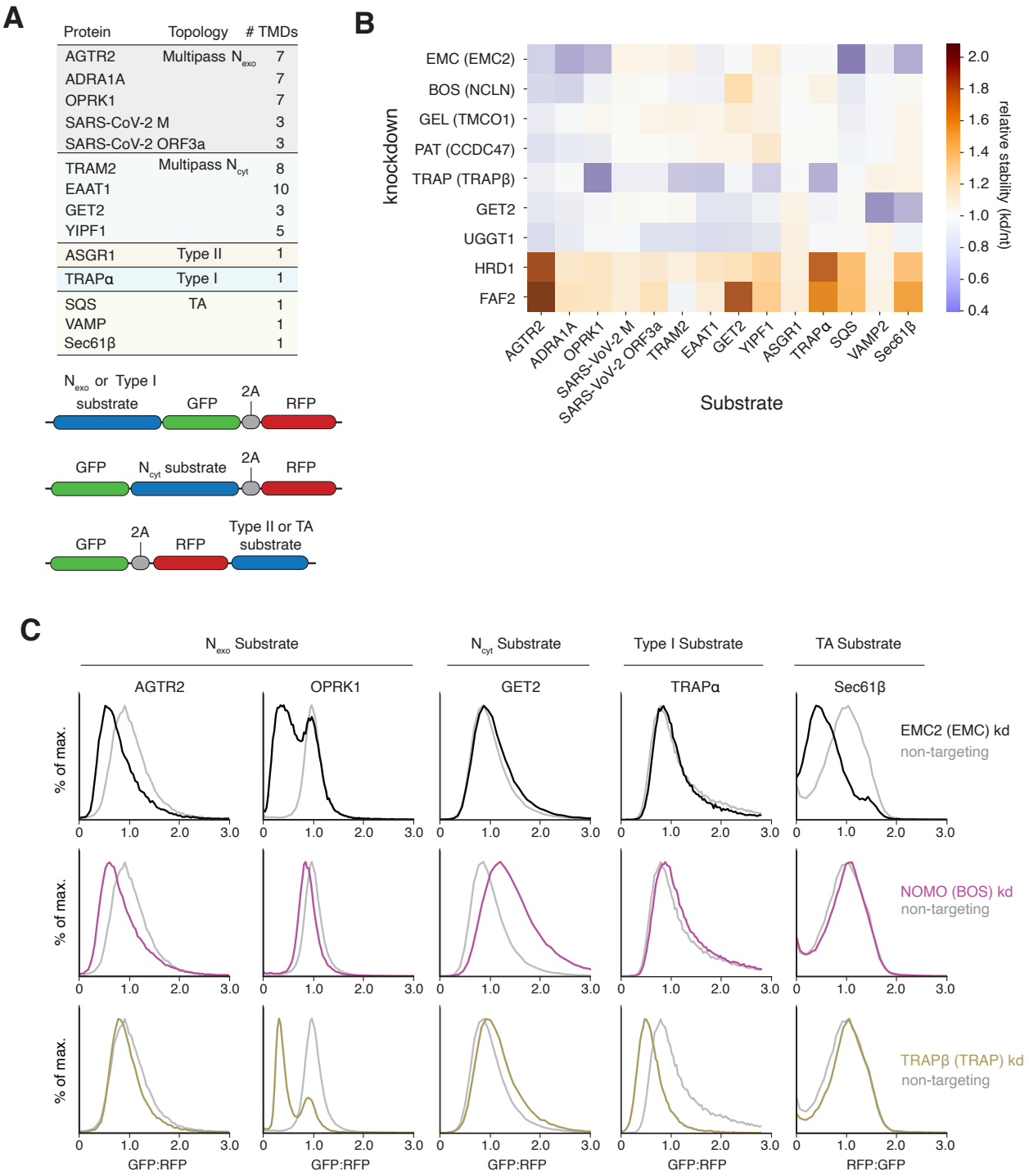

Figure 2.2. Differential effects of ER factors on membrane protein biogenesis.

(A) Properties and schematics of a panel of membrane protein reporters. **(B)** Summary of the arrayed screen performed in RPE1 cells testing the depletion of factors identified in Figure 2.1 on the reporters in (A). Results are displayed as a heatmap indicating relative

stability of each reporter after gene knockdown compared to a non-targeting control. **(C)** Representative analysis of individual data points in (B) displayed as a histogram.

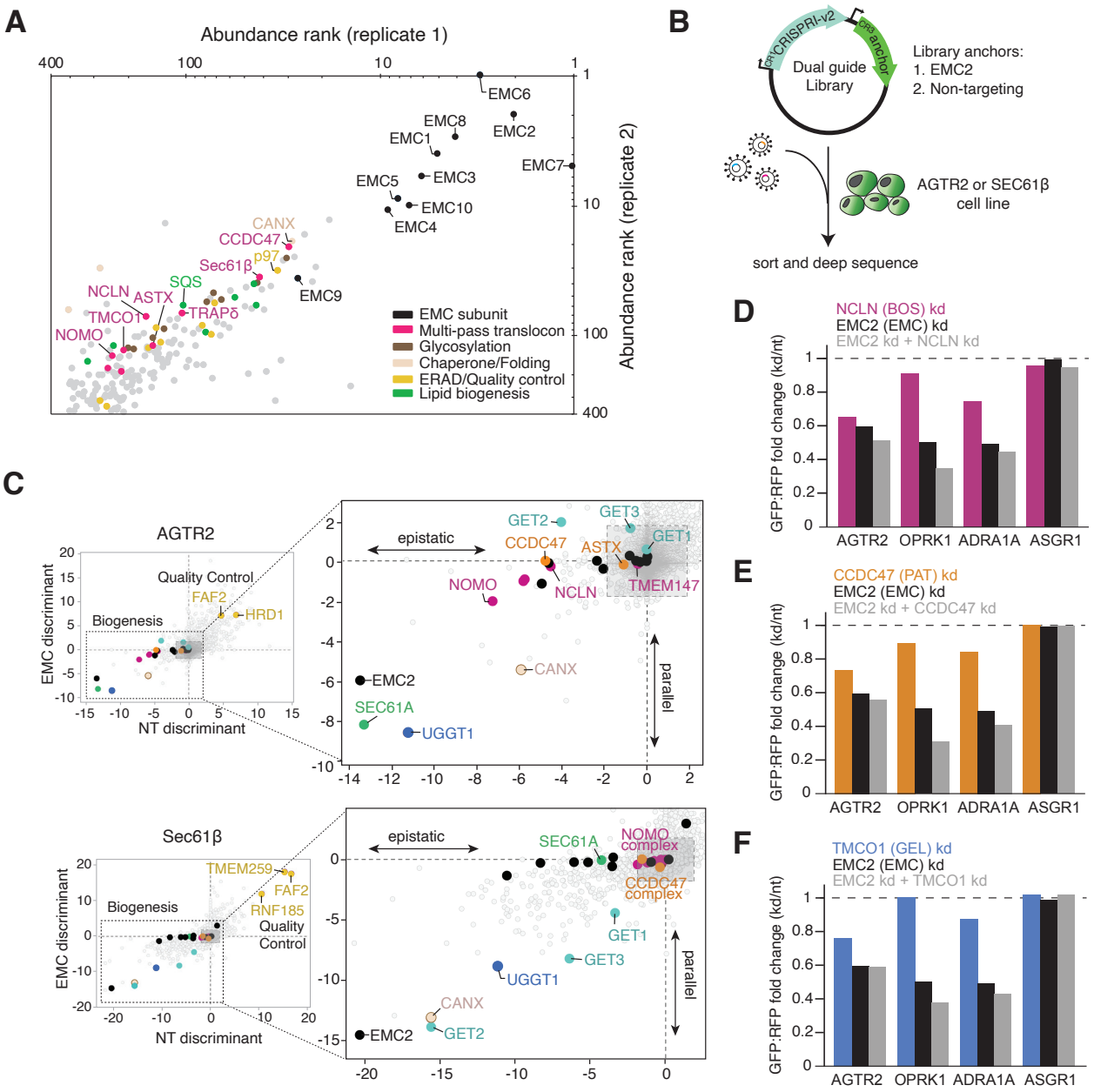

Figure 2.3. Components of the multipass translocon are epistatic with the EMC.

(A) Abundance rank of factors that co-immunopreciptate with the EMC when expressed at endogenous levels, as determined by mass spectrometry from two replicates. **(B)** Schematic of the non-targeting (control) and EMC2 genetic anchor dual guide libraries used for genome-wide EMC genetic modifier screens in AGTR2 and Sec61β reporter cell lines (Guna, Page, et al., 2023). **(C)** Comparison of EMC2- and NT-genetic modifier CRISPRi screens using the discriminant score (calculated using both the phenotype and -$\log_{10}$ Mann-

Whitney p-value for each gene) for AGTR2 (above) and Sec61β (below). Biogenesis factors are boxed and displayed at right in greater detail. Hits that fall off the diagonal are putative epistatic or parallel factors to EMC2, while hits along the diagonal are those independent of the EMC. **(D-F)** Summary of a panel of substrates (GPCRs: AGTR2, OPRK1, and ADRA1A; control: ASGR1) on the indicated factors determined using the fluorescent reporters described in Figure 2.1 (representative histograms Figure S2.2E).

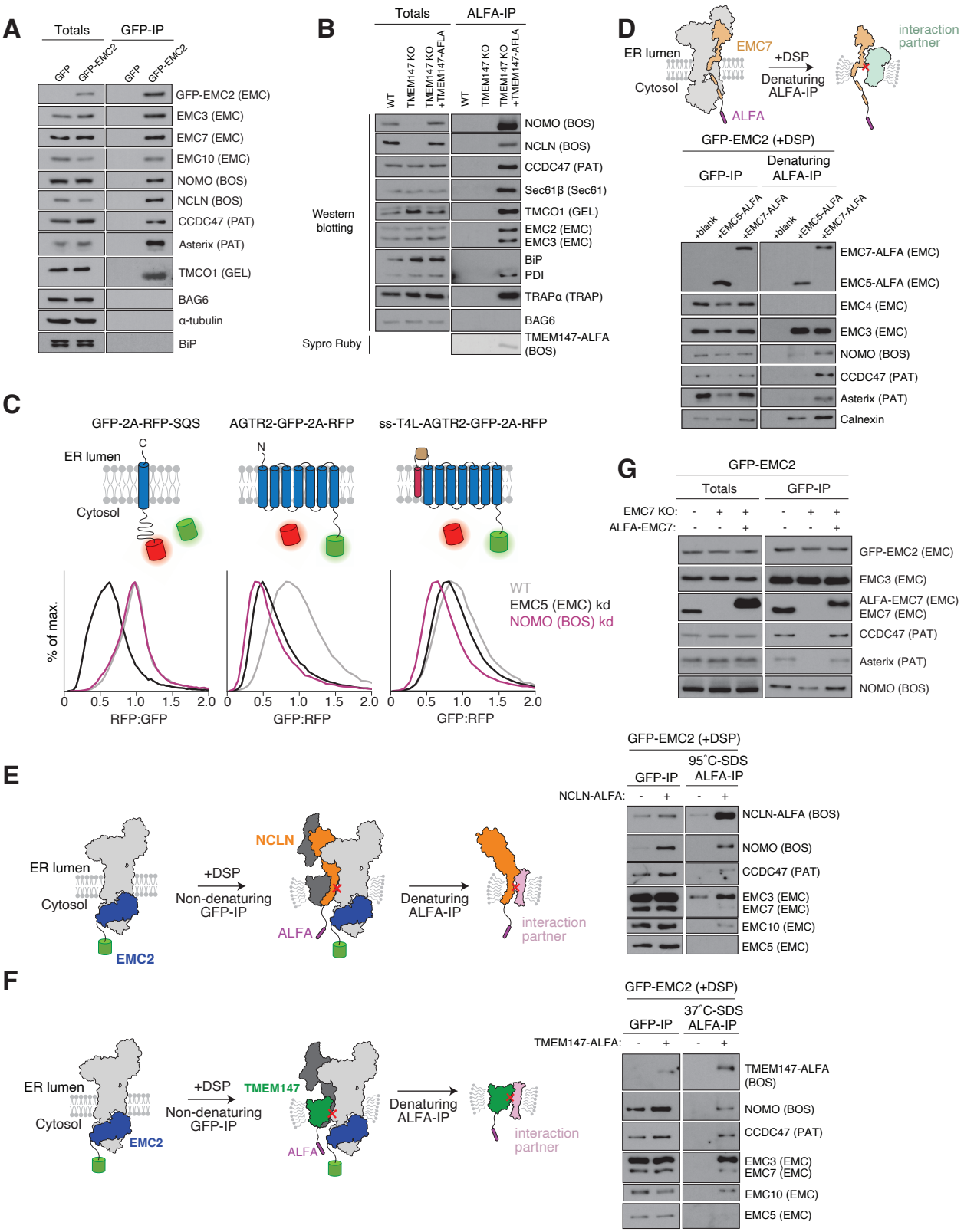

Figure 2.4. The EMC and BOS complex are direct physical interactors.

(A) HEK293T cells stably expressing GFP-EMC2 or GFP (background control) were solubilized and immunoprecipitated under native conditions using the anti-GFP nanobody, and analyzed by SDS-PAGE and western blotting. **(B)** As in (A) except with an ALFA nanobody immunoprecipitation of wild type cells (background control), TMEM147 KO, and TMEM147-ALFA rescue of TMEM147 KO cells. **(C)** Dependence of the TA-protein SQS and the GPCR AGTR2 on the indicated biogenesis factors, as determined using the fluorescent reporter system described in Figures 2.1-2 in RPE1 cells. To target AGTR2 to the Sec61-dependent biogenesis pathway the signal sequence (ss) of preprolactin followed by T4 Lysozyme (T4L) was appended to its N-terminus. Histograms of fluorescence relative to a normalization control was determined by flow cytometry and displayed as a histogram. **(D)** GFP-EMC2 cells were used to generate cell lines expressing either EMC7-ALFA, EMC5-ALFA, or a background control lacking an ALFA-tag. Intact cells were incubated with the chemical crosslinker DSP, and subjected first to a GFP immunoprecipitation under native conditions to isolate all EMC interacting partners (left), followed by a denaturing ALFA immunoprecipitation to isolate factors covalently crosslinked to either EMC5 or EMC7 (right). Prior to analysis by SDS-PAGE and western blotting, the thiol crosslinks were reduced. **(E)** As in (D), but for cell lines stably expressing GFP-EMC2 and NCLN-ALFA. **(F)** As in (D) for cell lines expressing GFP-EMC2 and TMEM147-ALFA. **(G)** EMC7 is required for interaction of the EMC with multipass translocon components. Wild type (WT) or EMC7 KO cells (Δ7) stably expressing GFP-EMC2 were transduced with WT EMC7 or a BFP control and subjected to anti-GFP nanobody purification under native conditions. Samples were analyzed by SDS-PAGE and western blotting.

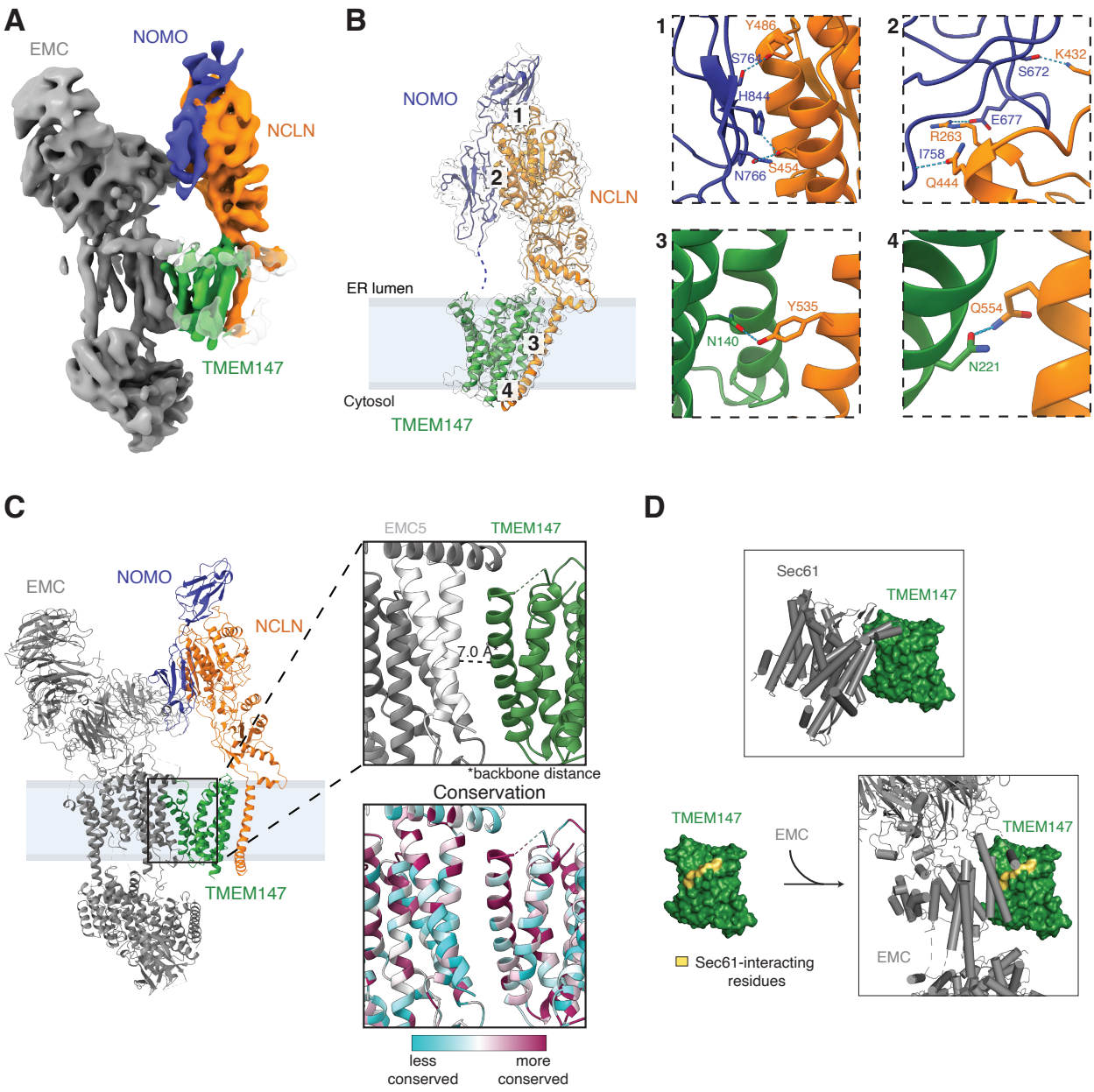

Figure 2.5. Structural analysis of the EMC•BOS holocomplex.

(A) Coulomb potential map of the EMC•BOS complex. **(B)** To enable *de novo* modeling of the BOS complex, structures were determined containing both full-length NOMO with all 12 Ig-like repeats (fNOMO), or a truncated version including only the 3 C-terminal Ig-like domains (tNOMO). The molecular model of the BOS(tNOMO) complex (BOS complex with truncated NOMO) with its corresponding EM density map is shown. Insets 1-2 and 3-4 highlight interactions between subunits in the lumen and within the membrane,

respectively. **(C)** Molecular model of the BOS(fNOMO)•EMC complex. Insets show the interface (top) and conservation (bottom) between EMC5 and TMEM147 within the membrane. **(D)** Comparison of the interaction of TMEM147 (surface filling in green) with Sec61 (McGilvray et al., 2020) versus the EMC (shown as grey cartoon). Residues on TMEM147 interacting with Sec61 are highlighted in yellow.

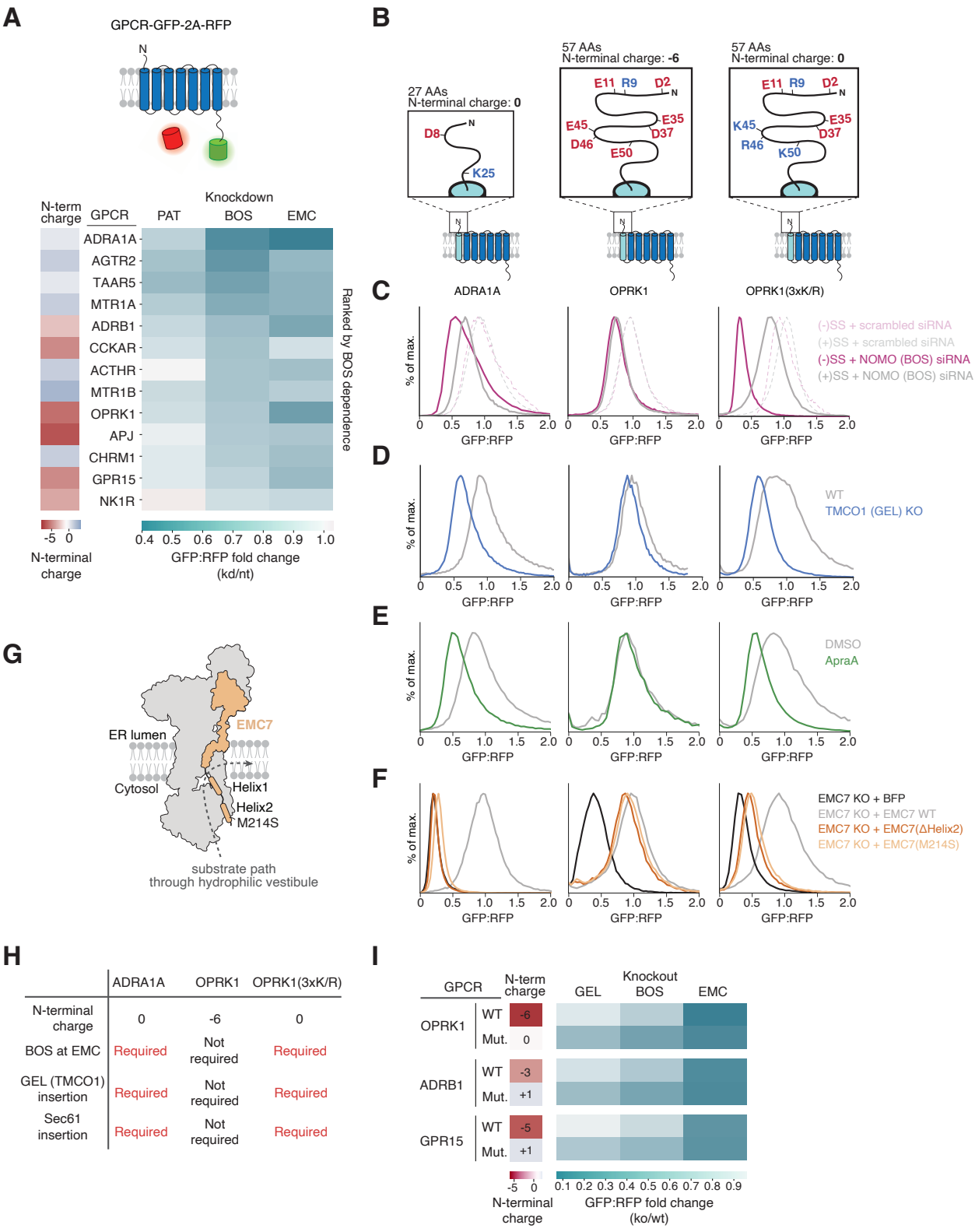

Figure 2.6. Properties of the soluble N-terminus determine biogenesis pathway.

(A) Schematic of the reporter system (top) used in RPE1 cells to test the dependence of the indicated GPCRs on the PAT, BOS, and EMC complexes. (Right) A heatmap summarizing the effect of depletion of each factor ordered by BOS dependence in comparison to the total charge within the N-terminal soluble domain of each GPCR. **(B)** Schematics of GPCRs ADRA1A, OPRK1 and a variant of OPRK1 with a net neutral N-terminal domain (3xKR). **(C)** Stability of the indicated reporters fused to the preprolactin signal sequence (SS) and T4L Lysozyme, as described in Figure 2.4C. Cells were treated with siRNAs targeting NOMO or a scrambled control sequence. **(D)** The reporters in (B) were assayed for dependence on TMCO1 in WT and TMCO1 KO cells. **(E)** The reporters in (B) were assayed for dependence on the lateral gate of Sec61 using the inhibitor Apratoxin A. **(F)** Reporters in (B) were assayed for dependence on features of EMC7. EMC7 KO cells were transduced with either a BFP control, WT EMC7, EMC7 with a deletion of Helix2 (ΔHelix2), or EMC7 containing a single point mutation in Helix2 (M214S) before analysis. **(G)** Cartoon of the EMC with EMC7 highlighted showing a substrate's path of insertion into the ER bilayer (Pleiner et al., 2023). **(H)** Summary of data in (C)-(F). **(I)** The indicated ratiometric wild type and mutant GPCR reporters were expressed and analyzed by flow cytometry in HEK293T WT, TMCO1 (GEL) KO, TMEM147 (BOS) KO, and EMC6 (EMC) KO cells. (Left) A heatmap indicating the charge of each wildtype (WT) and mutant (Mut) GPCR N-terminus. Mutants: OPRK1(E45K,D46R,E50K), ADRB1(E48K,E51K), and GPR15(D11K,D21K,E24K). (Right) Heatmap represents the relative stability of each reporter between the knockout vs wild type control cells.

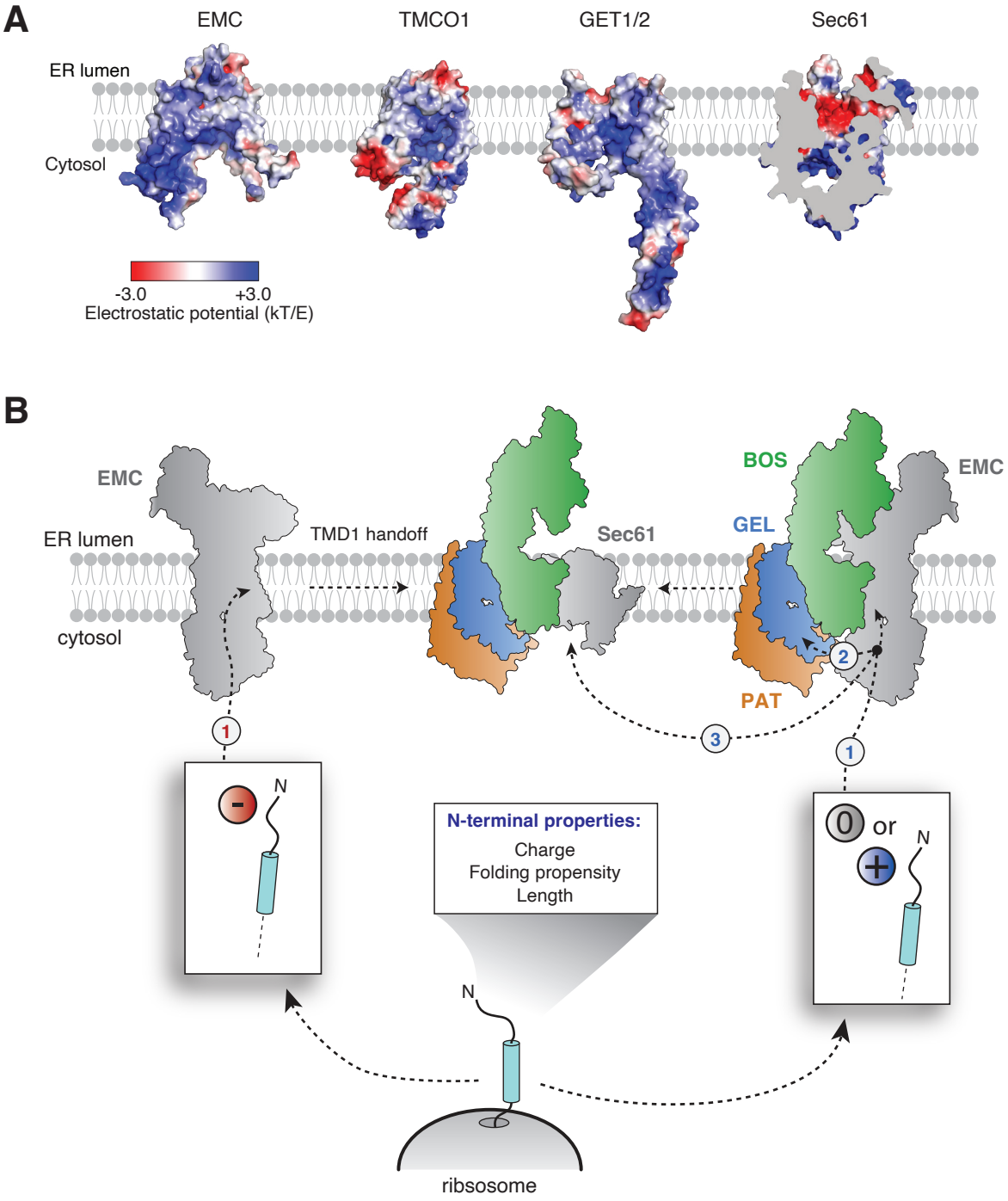

Figure 2.7. Model of multipass membrane protein biogenesis at the ER.

(A) Comparison of the electrostatic surfaces of human ER insertases. Electrostatic potential was generated using APBS and mapped onto surface representation of EMC3/6 (Pleiner et al., 2023), TMCO1 (McGilvray et al., 2020), GET1/2 (McDowell et al., 2020), and Sec61 (Voorhees & Hegde, 2016). **(B)** Model for membrane protein integration. Multipass N_{exo}

proteins are co-translationally targeted to the EMC, where the properties of their N-terminal soluble domain alter the requirements for insertion of the first TMD. (Left) Proteins with a net negative charge N-terminus utilize the EMC for TMD1 insertion without additional co-factors. These proteins are then handed off to the multipass translocon at Sec61 for insertion of the remaining TMDs. (Right) Proteins that have net neutral or positive N-terminal domains require additional factors for insertion of the first TMD. The BOS complex physically associates with the EMC and may facilitate insertion at the EMC or handover to GEL or Sec61. After TMD1 is inserted, the multipass translocon inserts the remaining TMDs.

Table 2.1. CryoEM Data Collection, Refinement, and Validation Statistics

	Full-length BOS (EMD-45293)	Truncated BOS (EMD-45294) (PDB 9C7U)	BOS-EMC (EMD-45295) (PDB 9C7V)
Data Collection and Processing			
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300	300
Camera	Gatan K3	Gatan K3	Gatan K3
Magnification (nominal)	105,000	105,000	105,000
Defocus range (μm)	-1.0 to -3.0	-1.0 to -3.0	-1.0 to -3.0
Calibrated pixel size (Å/pix)	0.416	0.416	0.416
Electron exposure ($e^-/Å^2$/frame)	60	60	60
Number of frames per movie	40	40	40
Automation software	SerialEM	SerialEM	SerialEM
Number of micrographs	11,870	15,929	17,978
Initial particle images (no.)	814,566	1,900,000	3,100,000
Final particle images (no.)	63,018	115,841	45,703
Local resolution range (Å)	7.8-12.6	3.1-8.8	5.5-8.7
Map resolution (Å, FSC=0.143)	8.85	3.65	6.6
Refinement			
Software (*phenix.real_space_refine*)		PHENIX 1.20.1-4487	PHENIX 1.21-5207-000
Initial model used (PDB code)		AlphaFold	PDB 8S9S, 9C7U
Correlation coefficient (CC_{mask})		0.83	0.70
Map sharpening *B* factor ($Å^2$)		-60	-220
Model composition			
Non-hydrogen atoms		6848	24532
Protein residues		896	3144
Ligands		NAG:1	NAG:6
B factors ($Å^2$)		min/max/mean	
Protein		26/170/82	22/221/106
Ligand		83/83/83	30/30/30
R.M.S deviations			
Bond lengths (Å) (# > 4σ)		0.003	0.003
Bond angles (°) (# > 4σ)		0.608	0.659
Validation			
MolProbity score		2.10	2.41
Clashscore		11.43	23.10
Rotamer outliers (%)		0.29	1.45
Cß outliers (%)		0.00	0.00
CaBLAM outliers (%)		5.41	2.63
Ramachandran plot			
Favored (%)		90.9	93.44
Allowed (%)		9.1	6.43
Disallowed (%)		0.00	0.13

2.6. Materials and Methods

Data and Code Availability

- The genome wide FACS screens datasets are provided as Table S2.1, S2.2, S2.3.
- Mass spectrometry dataset of the EMC is provided as Table S2.4 and is deposited to the PRoteomics IDEntifications (PRIDE) Database (PRIDE accession code: PXD045009).
- Original CryoEM maps and structures are deposited to EMDB and PDB databases. EMDB and PDB codes are listed in the Key Resources Table.
- Several figures in this paper make use of existing structures. PDB codes are listed in the Key Resources Table.

Plasmids and antibodies

The sequences used in cell-based assays and structural analysis were derived from UniProtKB/Swiss-Prot. These include: SEC61β (SEC61B; NP_006799.1), squalene synthase isoform 1 (SQS/FDFT1; Q6IAX1), vesicle associated membrane protein 2 (VAMP2; P63027-1), type-2 angiotensin II receptor (AGTR2; P50052), SARS-CoV-2 Membrane protein (VME1_SARS2; P0DTC5), SARS-CoV-2 ORF3a (AP3A_SARS2; P0DTC3), kappa-type opioid receptor (OPRK1; P41145), alpha-1A adrenergic receptor (ADRA1A; P35348),

translocating chain-associated membrane protein 2 (TRAM2; Q15035), excitatory amino acid transporter 1 (SLC1A3/EAAT1; P43003), guided entry of tail-anchored proteins factor CAMLG (GET2/CAMLG; P49069), Yip1 domain family member 1 (YIPF1; Q9Y548), Asialoglycoprotein receptor 1 (ASGR1; P07306), trace amine-associated receptor 5 (TAAR5; O14804); melatonin receptor type 1A (MTR1A/ MTNR1A; P48039), beta-1 adrenergic receptor (ADRB1; P08588), cholecystokinin receptor type A (CCKAR; P32238), adrenocorticotropic hormone receptor (MC2R/ACTHR; Q01718), melatonin receptor type 1B (MTNR1B/MTR1B; P49286), apelin receptor (APLNR/APJ; P35414), muscarinic acetylcholine receptor M1 (CHRM1/ACM1; P11229), G-protein coupled receptor 15 (GPR15; P49685), tachykinin receptor 1 (TACR1/NK1R;), translocon-associated protein subunit alpha (SSR1/SSRA/TRAPA; P43307), ER membrane protein

complex subunit 7 (EMC7; Q9NPA0), ER membrane protein complex subunit 5 (EMC5/MMGT1; Q8N4V1), ER membrane protein complex subunit 2 (EMC2; Q15006), Nicalin (NCLN; Q969V3), transmembrane protein 147 (TMEM147; Q9BVK8), Nodal modulator (NOMO2; Q5JPE7), and mannosyl-oligosaccharide 1,2-alpha-mannosidase IA (MAN1A1; P33908).

The 2nd generation lenti-viral packaging plasmid psPAX2 (Addgene plasmid #12260) was a gift from Didier Trono. The 2nd generation lenti-viral packaging plasmid pCMV-VSV-G was a gift from Bob Weinberg (Addgene plasmid # 8454). The pHAGE2 lenti-viral transfer plasmid was a gift from Magnus A. Hoffmann and Pamela Bjorkman. For inducible expression in K562 cells during CRISPRi screens, the SFFV-tet3G backbone was used (Jost et al., 2017). Though mCherry and EGFP variants were used throughout the study, they are referred to as RFP and GFP, respectively, for clarity. The GFP:RFP reporter system for reporter assays was used as previously described to assess substrate insertion (Guna et al., 2022; Inglis et al., 2020).

For expression in K562 cells during genome-wide CRISPRi screens, AGTR2 and SARS-CoV-2 ORF3a were cloned as N-terminal fusions to GFP, followed by a viral 2A skipping sequence, and RFP. For SARS-CoV-2 M, the reporter was designed using the split GFP system (Cabantous et al., 2005; Kamiyama et al., 2016). Here, the GFP11 tag (RDHMVLHEYVNAAGIT) was inserted at the N-terminal separated by a 3X-GS linker to allow for complementation with GFP1-10. The M protein was designed as a split GFP reporter while AGTR2 and ORF3a were designed to contain full GFP fusions. The latter two substrates are unstable in cells with the additional length of GFP11 fused to the N-termini. Additionally, we note that in the arrayed screen in Figure 2.2, all substrates except for SEC61β contain full GFP or RFP fusions. The SEC61β reporter has been previously described (Guna, Page, et al., 2023). Briefly, the TMD and flanking regions were inserted downstream of the first 70 residues of the flexible cytosolic domain of SEC61β. At the C-terminal of Sec61β, the GFP11 tag (RDHMVLHEYVNAAGIT) was inserted, separated by a 2X-GS linker. To express GFP1-10 in the ER lumen, the human calreticulin signal sequence was appended to the N-terminal of GFP1-10-KDEL as previously described (Guna et al., 2022; Pleiner et al., 2020).

Programmed dual sgRNA guide plasmids were used in assays involving depletion of two genes (Replogle et al., 2020). The sgRNA protospacer sequences used to generate dual guide plasmids are listed in Table S2.5.

To generate knockout cell lines, the following sgRNAs were cloned into pX459 following a standard protocol: TMCO1, CCDC47, TMEM147. The sgRNA sequences are listed in Table S2.5.

The siRNAs used in this study are listed in Table S2.5.

Constructs for expression in rabbit reticulocyte lysate (RRL) were based on the SP64 vector (Promega). For *in vitro* translations, the ALFA epitope (PSRLEEELRRRLTEP) was appended to the C-terminal of SARS-CoV-2 ORF3a and M proteins, separated by a flexible 3X-GS linker (Götzke et al., 2019).

All plasmids are available upon request.

The antibodies used in this study are listed in the Key Resources Table.

Cell culture and cell line construction

All cell lines used in this study are listed in the Key Resources Table.

K562 cells containing KRAB-BFP-dCas9 (Gilbert et al., 2014) were cultured in RPMI-1640 with 25 mM HEPES, 2.0 g/L NaHCO3, and 0.3 g/L L-glutamine supplemented with 10% Tet System Approved FBS, 2 mM glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin. K562 cells were maintained between 0.25×10^6 –1×10^6 cells/mL. HEK 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 100 units/mL penicillin, 100 µg/mL streptomycin, and 10% FBS. RPE1 cells containing the KRAB-BFP-dCas9 machinery (Jost et al., 2017) were cultured in DMEM F12 medium supplemented with 10% FBS and 2 mM glutamine. K562, HEK 293T, and RPE1 cells were grown at 37 °C. Expi293 cells were cultured in Expi293 Expression Medium (Gibco) supplemented with 10% FBS and 2 mM glutamine. Expi293 cells were maintained between 0.5×10^6 – 2×10^6 cells/mL and harvested at 6×10^6 cells/mL.

Clonal knockouts of TMCO1, CCDC47 and TMEM147 were obtained by transfecting HEK 293T cells with pX459 encoding the respective sgRNA using TransIT-293 transfection reagent (Mirus, USA). 72 h post transfection, single cells were sorted into 96-

well plates using a SONY cell sorter (SH800S), and clones were selected following verification of protein depletion by western blotting.

Fluorescent reporter CRISPRi screens

CRISPRi screens were performed as previously described, with minor modifications (Gilbert et al., 2014; Horlbeck et al., 2016). For AGTR2 and Sec61β, screens were performed using either the Non-targeting-dual guide library or the EMC2-dual library. For SARS-CoV-2 M and ORF3a, screens were performed with the single CRISPRi-v2 library. We have previously demonstrated that the additional non-targeting guide in the dual guide cassette does not appreciably alter knockdown efficiency of the second guide in the cassette (Guna, Page, et al., 2023). CRISPRi libraries (single CRISPRi-v2 library, Non-targeting dual library [Addgene #197348], or EMC2 dual library [Addgene #197349]) were transduced at a multiplicity of infection less than one into 300-330 million K562-CRISPRi-Tet-ON cells containing the appropriate reporter. For the duration of the screens, cells were maintained in 1L spinner flasks (Bellco, SKU: 1965-61010) at a volume of 1L. 48 hours post-transfection, BFP positive cells were between 30-40%. After 48h, cells were treated with 1 µg/mL puromycin for three days to select for guide positive cells. Following approximately two days of recovery after puromycin selection, the reporter was induced with doxycycline (100-1000 ng/mL) for 24-48 hours and sorted on a FACSAria Fusion Cell Sorter. Cells were diluted to 0.5×10^6 cells/mL each day to ensure that the culture was maintained at an average coverage of more than 1000 per sgRNA.

For sorting, cells were gated for BFP (selecting guide-positive cells), RFP and GFP (selecting an expressing reporter). Cells were sorted based on the GFP:RFP ratio of the final gated population. Approximately 30 million cells with either the highest or lowest 30% GFP:RFP ratios were collected during sorting, pelleted, and flash-frozen. From cell pellets, genomic DNA was extracted and purified using a Nucleospin Blood XL kit (Takara Bio, #740950.10). The guides were amplified and barcoded by PCR using NEB Next Ultra ii Q5 MM (M0544L). For both single and dual guide CRISPRi screens, a unique forward index primer was used. For single guide CRISPRi screens, a reverse primer that binds downstream of the guide was used (5'-CAAGCAGAAGACGGCATACGAGATCGACTCGGTGCCACTTTTTC). For dual

guide CRISPRi screens, a reverse primer that binds in the hU6 region upstream of the fixed guide was used (5'-CAAGCAGAAGACGGCATACGAGATGGAATCATGGGAAATAGGCCCTC), as previously described (Guna et al., 2023). SPRISelect beads (Beckman Coulter B23317) were used to purify the DNA library (279 or 349 bp), and purified DNA was analyzed on an Agilent 2100 Bioanalyzer prior to sequencing using an Illumina HiSeq2500 using the standard CRISPRi-v2 library sequencing primer (5'-GTGTGTTTTGAGACTATAAGTATCCCTTGGAGAACCACCTTGTTG). Analysis of the sequencing was performed using the pipeline in https://github.com/mhorlbeck/ScreenProcessing (Horlbeck et al., 2016). To ensure coverage, guides with fewer than 50 counts were excluded from analysis. The phenotype score for each gene was calculated from the strongest 3 sgRNA phenotypes. The Mann-Whitney p-value was calculated using the 5 sgRNAs targeting the same gene compared to the negative controls. For screens that were performed in biological duplicate (SARS-CoV-2 M, SARS-CoV-2 ORF3a, and Sec61β), the sgRNA phenotypes were averaged. To calculate the discriminant scores used in Fig. 2.3, each gene's phenotype score was multiplied by its Mann-Whitney p-value.

Lentivirus production

Lentivirus was generated via co-transfection of a transfer plasmid of interest along with packaging plasmid psPAX2 and envelope plasmid VSV.G, using TransIT-293 transfection reagent (Mirus). Lentivirus was harvested 48 hours after transfection, then aliquoted, flash-frozen, and stored for future usage.

Expi293 cell line generation

Cell lines for structural analysis were generated in Expi293 cells. Suspension cells were transduced by mixing 10 million cells with 2.5 ml of harvested lentiviral supernatant in presence of 8 μg/ml polybrene in a final volume of 30 ml in a 125-ml vented flask. For BOS (tNOMO) cell line, 2.5 ml of lentiviral supernatant of each subunit TMEM147-GFP, NCLN-P2A-RFP, NOMO(Δ1-9Ig)-P2A-TagBFP were added during transduction. The cells were grown in a shaking incubator for ~16 hours before being pelleted and resuspended in 50 ml of fresh medium in order to remove lentiviral particles. Then the cells

were continued to be grown for about a week until transduced cells expressing plasmid of interest were sorted with the Sony SH800S cell sorter (Sony Biotechnology).

K562 cell spinfection with programmed guides

K562 dCas9-BFP-KRAB cells were spinfected with lentivirus containing dual sgRNAs targeting two genes of interest or a non-targeting control. Briefly, 250,000 cells were mixed with 200 µl of lentivirus and RPMI medium in the presence of 8 µg/ml polybrene in a total volume of 2 ml in a 12-well plate. Plates containing K562 cells were spun at 1,000 g for 1.5-2 h at 30˚C, resuspended, and cultured in 12-well plates. Approximately 48 h after spinfection, 1 µg/ml puromycin was added for 5 consecutive days to select cells containing the dual guide cassette. To assess the percentage of guide-containing, BFP-positive cells, samples were analyzed using flow cytometry, as described below. After a total of 8 days of knockdown, cells were pelleted, flash frozen, and used in western blot analysis to assess knockdown of individual genes.

Reporter assays

For reporter assays in adherent HEK293 or RPE1 cells, lenti-viral transduction of 50-300 µl lentiviral supernatant and 8 µg/ml polybrene (Millipore-Sigma, USA) were added to ~70% confluent cells in 2.5 ml culture medium in a 6-well. Lenti-viral reporter constructs of all GPCRs, TRAM2, SARS-CoV-2 M, SARS-CoV-2 ORF3a, EAAT1, GET2, YIPF1 for use in HEK 293T and RPE1 cells, contained an upstream CMV promoter followed by the protein fused to GFP, a 2A site, and RFP (Figures 2.2B,C , 2.3D-F, 2.4C, 2.6A, 2.6C-F, S2.6A, S2.6C-F, S2.3D-E, S2.7A). Versions of GPCRs with a signal sequence (Figures 2.4C, 2.6C, and S2.6C) contained N-terminal fusions of the pre-prolactin signal sequence (KGSSQKGSRLLLLLVVSNLLLCQGVV) followed by a T4 Lysozyme soluble domain. In parallel, the first TMD (residues 33-75) of MAN1A1, a membrane protein with N_{cyt} topology, was fused to the N-terminus of GPCRs. Both signal sequence-T4 lysozyme fusions and MAN1A1 fusions behaved similarly (Figure 2.4C and S2.3E). Sec61β, SQS, VAMP2, and ASGR1 lenti-viral reporters for use in HEK 293T and RPE1 cells contained an upstream CMV promoter, followed by GFP, a 2A site and RFP, which was fused to the reporter. The TMD and flanking regions of Sec61β, SQS, and VAMP2 the were fused directly to RFP, as described before (Guna et al., 2018; Guna, Page, et al., 2023; Pleiner et

al., 2020). A charge mutant of OPRK1 (E45K, D46R, E50K) (+0 variant) was used in RPE1 cells in the same GFP-2A-RFP cassette as described above for GPCRs and as previously described (Pleiner et al., 2023) (Figures 2.6C-F, S2.6C, S2.7A).

For CRISPRi knockdown experiments in RPE1 cells, cells were transduced with sgRNA dual guide lenti-viral vectors. After 6 days of knockdown, cells were transduced with fluorescent reporter lenti-viral vectors described above and analyzed ~48h post-transduction (8 days after transduction with guide).

For rescue assay experiments, 300,000 HEK 293T and HEK 293T EMC7 KO cells were seeded into each 6-well plate on Day 1. On Day 2, cells were transduced with 300 µl lentiviral supernatant of rescue construct(s) and 8 µg/ml final concentration of polybrene, marking the start of the 72-hour rescue lentivirus addition. The media was exchanged on Day 3 to remove excess polybrene, and the 48-hour reporter lentivirus addition started by transducing the cells with 150 µl lentiviral supernatant of reporter construct(s) in presence of 8 µg/ml final concentration of polybrene. On Day 4, the cells were split 1:2 into a different set of 6-well plates to be used for western Blot. Lastly, on Day 5, the cells were harvested, washed and resuspended in 500 µl Dulbecco's Phosphate Buffered Saline (Gibco) to be analyzed by flow cytometry or frozen for analysis via western Blot.

Flow cytometry

RPE1 and HEK 293T cells were trypsinized, washed with 1xPBS, and resuspended in 1xPBS for flow cytometry analysis. K562 cells were analyzed directly from 12-well or 6-well cultures. Cells were analyzed using an Attune NxT Flow Cytometer (Thermo Fisher Scientific, USA) or a MACSQuant VYB (Miltenyi Biotec, Germany). Flow cytometry data was analyzed using FlowJo v10.8 Software (BD Life Sciences, USA) or by Python using the FlowCytometryTools package.

The Sec61 inhibitor Apratoxin A was used to analyze the effect of SEC61 inhibition on membrane protein insertion (Paatero et al., 2016; Thornburg et al., 2013). HEK 293T (WT or TMCO1 KO) cells were transduced with reporter lenti-virus, and 48h later, cells were treated with 31.3 nM Apratoxin A in 0.1% DMSO for 12h. Cells were analyzed immediately following treatment with inhibitor. Apratoxin A was a gift from Ville Paavilainen.

Preparation of human ER microsomes

Human derived rough ER microsomes were generated as previously described, with minor modifications (Chitwood et al., 2018). HEK293T cells (WT, NCLN KO, TMCO1 KO, or EMC6 KO) were harvested and washed in 1X PBS. Cells were resuspended in 4 times the pellet volume of sucrose buffer (10 mM HEPES, pH 7.5, 250 mM sucrose, 2 mM magnesium acetate, 1X cOmplete EDTA-free protease inhibitor cocktail [Roche]) and lysed by douncing at 4 °C. Lysed cells were diluted 2X in sucrose buffer and pelleted at 3214 xg for 35 min. at 4 °C. Supernatant was transferred to a new tube and pelleted again at 3214 xg for 35 min. at 4 °C. To isolate the microsomal fraction, samples were pelleted in an ultracentrifuge in an MLA80 rotor (Beckman-Coulter) at 75,000 xg for 1h at 4 °C. Supernatant was removed, and the microsomal pellet was resuspended to an A280 of 75 in microsome buffer (10 mM HEPES, pH 7.5, 250 mM sucrose, 1 mM magnesium acetate, 0.5 mM DTT). To remove contaminating RNAs, microsomes (hRMs) were nucleased. $CaCl_2$ (1 mM) and micrococcal nuclease (0.125 U/μL) were added to hRMs and mixed before incubating for 6 minutes at 25 °C. To quench the reaction, EGTA (2 mM) was added to the sample and the sample was immediately mixed and place on ice. Nucleased hRMs were flash frozen and stored at -80 °C prior to use in *in vitro* translations.

Mammalian *in vitro* translation

Translation extracts were prepared using nucleased rabbit reticulocyte lysate (RRL) supplemented with human derived rough ER microsomes, as previously described (Sharma et al., 2010; Walter & Blobel, 1983). DNA templates for *in vitro* transcription were made by PCR from SP64-based plasmids or directly from double-stranded DNA gene fragments (IDT or Twist Biosciences) using primers within the SP6 promoter (5' end) and following a stop codon and short untranslated region (3' end). Run-off transcription reactions were made by combining 4.8 μL T1 mix (Sharma et al., 2010), 0.1 μL RNasin (Promega), 0.1 μL SP6 polymerase (New England Biolabs) and 50 ng PCR product. Reactions were incubated at 37 °C for 2 hours, and then used directly in translation reactions, which were incubated for 20-45 minutes at 32 °C. To label nascent proteins, radioactive ^{35}S-methionine

(Perkin Elmer) was included in translation reactions, unless otherwise indicated. Samples were then analyzed directly using SDS-PAGE and autoradiography.

For experiments in which the insertion of the first TMD was assessed, substrates were translated in the presence of hRMs derived from HEK 293T cells. The OPRK1 constructs (wildtype or variants with 3xK or 5xK mutations in the N-terminal soluble domain) contain an Asn-Gly-Thr (NGT) glycosylation site at the N-terminus, which allows monitoring of insertion. The VAMP2 control protein contains a C-terminal Opsin tag that gets glycosylated upon insertion of the TA substrate and allows monitoring of insertion. A construct containing the first 85 amino acids of preprolactin was used as a control for signal sequence cleavage. For assays in which Sec61's insertion capacity was assessed, the inhibitor Apratoxin A was used at 1 μM.

Preparation of the ALFA nanobody conjugated to HRP for Western blotting

The ALFA nanobody was coupled to HRP-maleimide through a single engineered C-terminal cysteine residue, as previously described (Pleiner et al., 2017).

DSP crosslinking

Suspension adapted T-REx-293 cells stably expressing either GFP-EMC2 only or GFP-EMC2 plus EMC5-ALFA, EMC7-ALFA, TMEM147-ALFA, or NCLN-ALFA were harvested, washed in PBS, pelleted, and resuspended in PBS containing 1.5 mM final concentration of dithiobis(succinimidyl propionate) (DSP; Thermo Scientific). The cell mixture was incubated at 4°C with head-over-tail rotation for 2 hours. After the incubation, the reaction was quenched by addition of 1M Tris/HCl, pH 7.5 to 20 mM final concentration and incubated for 15 min. Then, the cells were pelleted, weighed, and flash frozen for storage prior to immunoprecipitation, as described below.

Immunoprecipitation

Native immunoprecipitations of the EMC were performed from T-REx-293 cells stably expressing GFP-EMC2 generated as previously described (Pleiner et al., 2021). Cells were isolated and solubilized in solubilization buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 2 mM $MgAc_2$, 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail [Roche], 1% [w/v] glyco-diosgenin [GDN; Anatrace], 1 mM DTT) for 30 minutes on ice. Lysate was cleared

by centrifugation for 10 minutes at 4°C and added to anti-GFP nanobody immobilized on magnetic streptavidin resin. Briefly, to immobilize the nanobody, Pierce™ Streptavidin magnetic resin was incubated with biotinylated His14-Avi-SUMO^{Eu1}-tagged anti-GFP nanobody (Addgene #149336) in wash buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgAc_2$, 0.01% GDN, 1 mM DTT) for 20 min. at 4°C as described previously (Stevens et al., 2023). Following immobilization, the unbound Streptavidin sites were blocked with 50 mM HEPES/KOH pH 7.5 containing 100 μM biotin for 5 min at 4°C. The resin was washed with solubilization buffer and incubated with clarified cell lysate for 1 hour at 4°C with head-over-tail rotation. Resin was washed 3 times in wash buffer and eluted in wash buffer containing 300 nM SUMO^{Eu1} protease (Addgene #149333) for 30 minutes at 4°C with head-over-tail rotation. Input for SDS-PAGE analysis was determined using absorbance at 488nm, in order to normalize the GFP-EMC2 input. Mock control samples were processed using T-REx-293 cells expressing GFP only.

Native immunoprecipitation of TMEM147-ALFA was performed as for GFP-EMC2 described above, with minor modifications. Briefly, WT or TMEM147 KO HEK293T cells were transduced with TMEM147-ALFA-P2A-TagBFP or TagBFP control lentivirus. Cells were solubilized as described above and incubated with magnetic resin immobilized with His14-Avi-SUMOStar-anti-ALFA nanobody. After washing the resin 3 times, elution was performed with 500 nM SUMOStar protease (LifeSensors) (Liu et al., 2008). Samples were analyzed via SDS-PAGE and western blotting.

To show the specificity of the EMC-BOS interaction, chemical crosslinking with DSP was performed followed by sequential immunoprecipitations (Figure 2.4D-F). T-REx-293 cells stably expressing GFP-EMC2 and either EMC5-ALFA, EMC7-ALFA, NCLN-ALFA, or TMEM147-ALFA were harvested and treated with DSP as described above. Cell pellets were solubilized in solubilization buffer without DTT for 30 min. at 4°C with head-over-tail rotation. Immunoprecipitations of GFP-EMC2 were performed under native conditions in GDN as described above and eluted with 300 nM SUMO^{Eu1} protease for 30 min at 4°C. The absorbance at 488 nm was used to normalize the eluate as input for the subsequent immunoprecipitation. For subsequent immunoprecipitations of EMC5-ALFA, EMC7-ALFA, or NCLN-ALFA, elution fractions were brought to 1% SDS and heated at 95°C for 5 min to fully denature the protein samples. To immunoprecipitate TMEM147-ALFA, the

elution fractions were brought to 1% SDS and heated to 37°C for 10 min, as TMEM147 precipitates upon boiling in SDS. After denaturation, samples were diluted in IP buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 0.5% Triton X-100) to 0.05% SDS. Samples were eluted in 500 nM SUMOStar protease in IP buffer for 30 minutes at 4°C. Samples were analyzed using SDS-PAGE and western blotting.

Mass spectrometry analysis of EMC-interacting factors

Native immunoprecipitations of GFP alone (as a negative control) and GFP-EMC2 were performed on T-Rex-293 cells stably either GFP or GFP-EMC2 as described above. Eluted protein samples were precipitated by addition of 1:10 volume of 100% Trichloroacetic acid (TCA) and incubating on ice for 10 minutes before the samples were centrifuged at max speed in a benchtop centrifuge at 4ºC. The pellets were washed in ice cold acetone 2 times before being air dried. TCA-precipitated pellets were resuspended in 50 mM HEPES, pH 8.0 containing 8 M Urea. Samples were reduced by incubation with 4 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Thermo Scientific) for 20 minutes at 37ºC, and alkylated by incubation with 12 mM 2-chloro-acetamide (CAA) (MP Biomedicals) for 15 minutes at 37ºC. Samples were digested with 2ng/µl Lysyl Endoproteinase (Lys-C) (Wako Chemicals) for 4 hours at 37ºC. Samples were diluted 4-fold with 50 mM HEPES, pH 8.0 and $CaCl_2$ was added to 1 mM final concentration. The samples were then digested with 0.6 ng/µl Trypsin (Thermo Scientific) for 18 hours at 37ºC. Samples were desalted using C18 Spin Columns (Thermo Scientific) according to the manufacturer's instruction, and lyophilized prior to mass spectrometry analysis.

LC-MS/MS analysis for the IP experiment (Figure 2.3A) was performed with an EASY-nLC 1200 (ThermoFisher Scientific, San Jose, CA) coupled to a Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific, San Jose, CA). Peptides were separated on an Aurora UHPLC Column (25 cm × 75 µm, 1.6 µm C18, AUR2-25075C18A, Ion Opticks) with a flow rate of 0.35 µL/min for a total duration of 75 min and ionized at 1.6 kV in the positive ion mode. The gradient was composed of 6% solvent B (3.5 min), 6-25% B (42 min), 25-40% B (14.5 min), and 40–98% B (15 min); solvent A: 2% ACN and 0.2% formic acid in water; solvent B: 80% ACN and 0.2% formic acid. MS1 scans were acquired at the resolution of 60,000 from 375 to 1500 m/z, AGC target 3e6, and

maximum injection time 15 ms. The 12 most abundant ions in MS2 scans were acquired at a resolution of 30,000, AGC target 1e5, maximum injection time 60 ms, and normalized collision energy of 28. Dynamic exclusion was set to 30 s and ions with charge +1, +7, +8 and >+8 were excluded. The temperature of ion transfer tube was 275°C and the S-lens RF level was set to 60. MS2 fragmentation spectra were searched with Proteome Discoverer SEQUEST (version 2.5, Thermo Scientific) against *in silico* tryptic digested the UniProt Human proteome Swiss-Prot database (UP000005640). The maximum missed cleavages were set to 2. Dynamic modifications were set to oxidation on methionine (M, +15.995 Da), deamidation on asparagine and glutamine (N and Q, +0.984 Da) and protein N-terminal acetylation (+42.011 Da). Carbamidomethylation on cysteine residues (C, +57.021 Da) was set as a fixed modification. The maximum parental mass error was set to 10 ppm, and the MS2 mass tolerance was set to 0.03 Da. Intensity-based quantification (iBAQ) was performed using the IMP-apQuant PD node (Doblmann et al., 2019; Schwanhäusser et al., 2011). The maximum false peptide discovery rate was specified as 0.01 using the Percolator Node validated by q-value. Data of the abundance rank in each replicate are available in Table S2.4.

Photo-crosslinking and immunoprecipitation

To perform site-specific photo-crosslinking of the BOS complex, (Fig. S2.5G), 3′-azibutyl-N-carbamoyl-lysine (AbK) (Iris Biotech) was incorporated at positions K471 and D472 of NCLN using the amber suppression system as described previously (Ai et al., 2011; Elsässer et al., 2016). In brief, we generated two plasmids: (1) expressing NCLN containing amber mutations at residues K471 and D472 in NCLN along with 4 copies of PyIT(U25C), and the second (2) expressing the WT *Methanosarcina mazei* pyrrolysyl-tRNA synthetase (PyIRS) and an additional 4 copies of the tRNAPylCUA (PyIT(U25C). Expi293 cells (Thermo) were grown in 0.5 mM of AbK, and transiently transfected with the two plasmids at a ratio of 1:4 using PEI “MAX” (Polysciences). Cells were grown for 72 hours before harvesting. Harvested cells were resuspended in PBS and UV-irradiated on ice at 7-10 cm distance from a UVP B-100 series lamp (Analytik Jena) for 15 minutes. The cells were pelleted, weighed, and flash frozen prior to immunoprecipitation.

Immunoprecipitation of NCLN-GFP WT and the AbK-containing mutants (D472amb, K471amb) were performed as described above with some modifications. Cells were solubilized in 1% Triton X-100, sufficient to disrupt the BOS complex, and clarified lysate was incubated with anti-GFP nanobody immobilized on magnetic streptavidin beads. Non-specific binders were removed with wash buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgAc_2$, 0.5% Triton X-100, 1 mM DTT) and the immobilized proteins were eluted using wash buffer supplemented with 300 nM $SUMO^{Eu1}$ protease. The resulting samples were analyzed by western blotting. The totals samples were normalized using the protein absorbance at 280 nm, and the IP samples were normalized using protein absorbance at 488 nm.

Protein purification for structure determination

2 L of Expi293 cells stably expressing the protein(s) of interest by lentiviral transduction were pelleted, washed with PBS and flash-frozen for storage. For BOS (fNOMO), a cell line was generated stably expressing TMEM147-GFP-2A-RFP. For BOS (tNOMO), a cell line stably expressing TMEM147-TEV-GFP, NCLN-RFP, NOMO(Δ1-9Ig)-BFP. For the BOS (fNOMO) • EMC holocomplex, we generated a cell line stably expressing TMEM147-5aa-TEV-GFP-40aa(ALFA)-EMC2. All protein complexes were purified using an anti-GFP nanobody as described previously (Pleiner et al., 2020; Stevens et al., 2023). Briefly, cell pellets were harvested, washed with 1xPBS, and resuspended in solubilization buffer (50 mM HEPES/KOH pH 7.5, 200 mM NaCl, 2 mM $MgAc_2$, 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail [Roche], 1% [w/v] glyco-diosgenin [GDN; Anatrace], 1 mM DTT) at a ratio 6.8 ml solubilization buffer per 1 g cell pellet. Following incubation for 1 hour at 4°C, lysate supernatant was isolated by centrifugation at 18,000 rpm using an SS-34 rotor in a Sorvall RC6+ Superspeed Centrifuge at 4°C for 45 min.

Simultaneously, biotinylated anti-GFP nanobody was immobilized onto streptavidin magnetic beads. Specifically, 80 µl resuspended Pierce™ Streptavidin magnetic beads per 1 g cell pellet were washed and equilibrated in wash buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgAc_2$, 0.0053% GDN, 1 mM DTT). Then His14-Avi-$SUMO^{Eu1}$-tagged anti-GFP nanobody (Addgene #149336) was immobilized onto the washed magnetic beads for 30 min with mixing at 4°C using a ratio of 27 µg for every 80 µl beads. This

immobilization was followed by incubation of beads with 50 mM HEPES/KOH pH 7.5 containing 100 μM biotin for 5 min on ice to block unbound biotin binding sites on the magnetic streptavidin beads. Subsequently, the beads were washed with solubilization buffer and incubated with clarified cell lysate for 1 hour with head-over-tail rotation. After incubation, 4 washes with wash buffer (2 volumes of wash buffer:1 volume of beads) was performed to remove unspecific binding to the beads. To elute the bound proteins, wash buffer containing 500 nM $SUMO^{Eu1}$ protease (Addgene #149333) was added to the beads and left to incubate for 30 minutes with mixing at 4°C. The eluent was further purified using size exclusion chromatography with a 3.5 ml Superose 6 column (GE Life Sciences). For BOS (tNOMO) sample, TEV protease was added (1 mg TEV protease for every 30 mg of BOS (tNOMO) protein) and incubated overnight at 4°C without mixing to remove the GFP-tag before size exclusion chromatography. The fractions corresponding to the protein complexes were concentrated using a 500-μl 30K MWCO concentrator (Millipore-Sigma).

Grid preparation and data collection

For BOS (fNOMO) sample, 3 μL of purified, concentrated protein at 2.48 mg/ml was applied to UltrAuFoil® R 1.2/1.3 holey gold film grid (Ted Pella, Inc.) that had been glow discharged with the PELCO easiGlow™ (Ted Pella, Inc.) at 20 mA for 60 s. The grid was blotted at 6°C, 100% humidity, -4 blot force for 4 seconds and plunged frozen in liquid ethane using the FEI Vitrobot Mark IV (Thermo Fisher Scientific). Data were collected on a Titan Krios operating at 300 keV and equipped with a Gatan K3 direct detector and a 20 eV slit width energy filter. Images were acquired using an automated acquisition pipeline in SerialEM (Mastronarde, 2005) and recorded at 105k magnification with a defocus range of -3.0 to -1.0 μm and total exposure dose of 60 $e^-/Å^2$ in super resolution mode with a pixel size of 0.418 Å/pixel. 11,870 micrographs were collected for this data set.

For BOS (tNOMO), the grid was prepared in a similar manner, except the protein sample was concentrated to 4 mg/ml and mixed with 0.005% 3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO; Sigma Aldrich) immediately before vitrification. 15,929 micrographs were collected for this data set.

For the BOS (fNOMO) • EMC homocomplex sample, the protein concentration was at 2.53 mg/ml and 17,978 micrographs were collected for this data set.

Structure image processing

The workflows for data processing of BOS (fNOMO), BOS (tNOMO), and BOS (fNOMO) • EMC are summarized in Fig. S2.4A, S2.4B, S2.4C, respectively. Data processing was carried out using cryoSPARC v3.2-4.2.1(Punjani et al., 2017). For preprocessing, micrographs were motion-corrected, Fourier-cropped twofold to 0.832 Å/pixel using 'Patch Motion Correction'; then, they were subjected to patch-based contrast transfer function (CTF) estimation with 'Patch CTF Estimation'. Movies were selected based on CTF fit cut-off of 5.0 Å in 'Curate Exposure'. From here on out, details of data processing differ for each structure.

For BOS (fNOMO), 814,566 particles were picked using 'Blob Picker' and extracted with box size = 512 pixels from 7,174 selected movies. Iterative rounds of '2D classifications' performed to remove background and junk particles. 2D classes that resemble BOS (fNOMO) complex were used as template for 'Template Picker' with particle diameter = 190 Å, which resulted in 904,456 picked particles. A round of '2D Classification' was performed to remove background particles, resulting in 246,296 particles, which were subjected to 2 more rounds of 2D classification. Then, the resulting 90,893 particles were used to generate 4 3D classes with 'Ab-Initio Reconstruction'. Using these 4 volumes, we performed 2 rounds of 3D 'Heterogeneous Refinement' on the 246,296 particles from earlier to arrive at the final EM map of BOS (fNOMO), generated from a set of 63,018 particles.

For BOS (tNOMO), 1,900,000 particles were picked using 'Blob Picker' with particle diameter of 120-320 Å and extracted from micrographs using box size = 512 pixels from 13,196 selected movies. After iterative rounds of '2D Classification', 319,567 particles were used to generate a 3D volume using 'Ab-Initio Reconstruction'. This volume was then used for template generation for template picking, resulting in 2,500,000 picked particles. After iterative rounds of 2D classification, 1,265,788 particles were subjected to 'Ab-Initio Reconstruction' into 4 volumes. The particles that correspond to the 2 volumes that best resembled our BOS (tNOMO) complex were put through multiple rounds of 3D

‘Heterogeneous Refinement’ to give us 312,033 particles which were re-extracted using box size of 448 pixels. These particles were used as input in more rounds of 3D ‘Heterogeneous Refinement’ and ‘3D Classification (BETA)’ to arrive at an EM map generated from 115,841 particles. This map was put through 3D ‘Non-uniform Refinement’ and sharpened with a B-factor of -60 $Å^2$ to give us our final map. Additionally, ‘DeepEMhancer’ was also used to aid in model building.

For BOS (fNOMO) • EMC, 3,100,000 particles were picked from 17,586 selected movies using ‘Blob Picker’ with particle diameter of 175-450 Å, which were then extracted from micrographs with box size of 512 pixels, 2x binned, and subjected to 4 rounds of ‘2D Classification’ to remove background particles. 300,442 particles were used to construct 3 ab-initio models with ‘Ab-Initio Reconstruction’ (with default settings except for maximum resolution = 7, initial resolution = 9, and initial minibatch size = 300, final minibatch size = 1000). The map that best resembled EMC with BOS was used to generate 2D templates for template particle picking (‘Template Picker’) with particle diameter of 300 Å, resulting in 1,932,646 picked particles. After 1 round of ‘2D Classification’ and ‘Heterogeneous Refinement’ using 3 junk classes that resemble background and 1 class that resembles EMC with BOS, 595,637 particles were subjected to a first round of ‘Ab-initio Reconstruction’ into 2 classes (default settings except: maximum resolution = 9, initial mini batch size = 400, final mini batch size = 1200). From the 3 replicate runs of the first round of ab-intio reconstruction, the particles associated with the better 3D volume that represented BOS (fNOMO) • EMC were combined and used as input in the second round of ab-initio reconstruction into 2 classes (same settings as previous round except: maximum resolution = 6, initial resolution = 12). This process was repeated for the third time with the second rounds’ particles that correspond to the better 3D volume (similar settings except: maximum resolution = 5, initial resolution = 7). For the subsequent rounds of ab-initio reconstruction, the settings were similar, except for maximum resolution = 4, initial resolution = 6. The particles were unbinned and further classified using 3D heterogeneous refinement twice to achieve a final class of 45,703 particles. To get our final map, local refinement was performed on this map using a mask on the BOS complex generated in Chimera (Pettersen et al., 2004) with soft padding = 18, using pose/shift

gaussian prior during alignment with standard deviation of prior over rotation = 3 degree and standard deviation of prior over shifts = 1 Å.

Model building and refinement

For the BOS (fNOMO) structure, initial models of each subunit (TMEM147, NCLN, NOMO) were generated using AlphaFold2-Multimer ColabFold (AlphaFold2_advanced.ipynb)(Mirdita et al., 2022). Since the map quality was not sufficient for accurate model building and refinement, the initial models of each subunit were only rigid body fitted into the EM density (Fig. 2.5A) and combined in COOT (Casañal et al., 2020; Emsley et al., 2010).

For the BOS (tNOMO) structure, we used the previously generated models from AlphaFold2 to rigid body fit into the density. The models were combined and manually refined in COOT(Casañal et al., 2020; Emsley et al., 2010). The final model was iteratively subjected to *phenix.real_space_refinement* (Afonine et al., 2018; Liebschner et al., 2019) with rigid body and secondary structure restraints.

For BOS (fNOMO) • EMC structure, we used BOS (tNOMO) and EMC structure (PDB: 8S9S) as initial models, which were combined, rigid body fitted and refined manually in COOT (Casañal et al., 2020; Emsley et al., 2010). The model was iteratively subjected to *phenix.real_space_refinement* (Afonine et al., 2018; Liebschner et al., 2019) with rigid body and secondary structure restraints.

CryoEM data collection, refinement, and validation statistics are reported in Table 2.1. Final models were evaluated with MolProbity. All figures in this study were generated with PyMOL (www.pymol.org) and ChimeraX (Goddard et al., 2018; Pettersen et al., 2021).

2.7. Supplementary materials

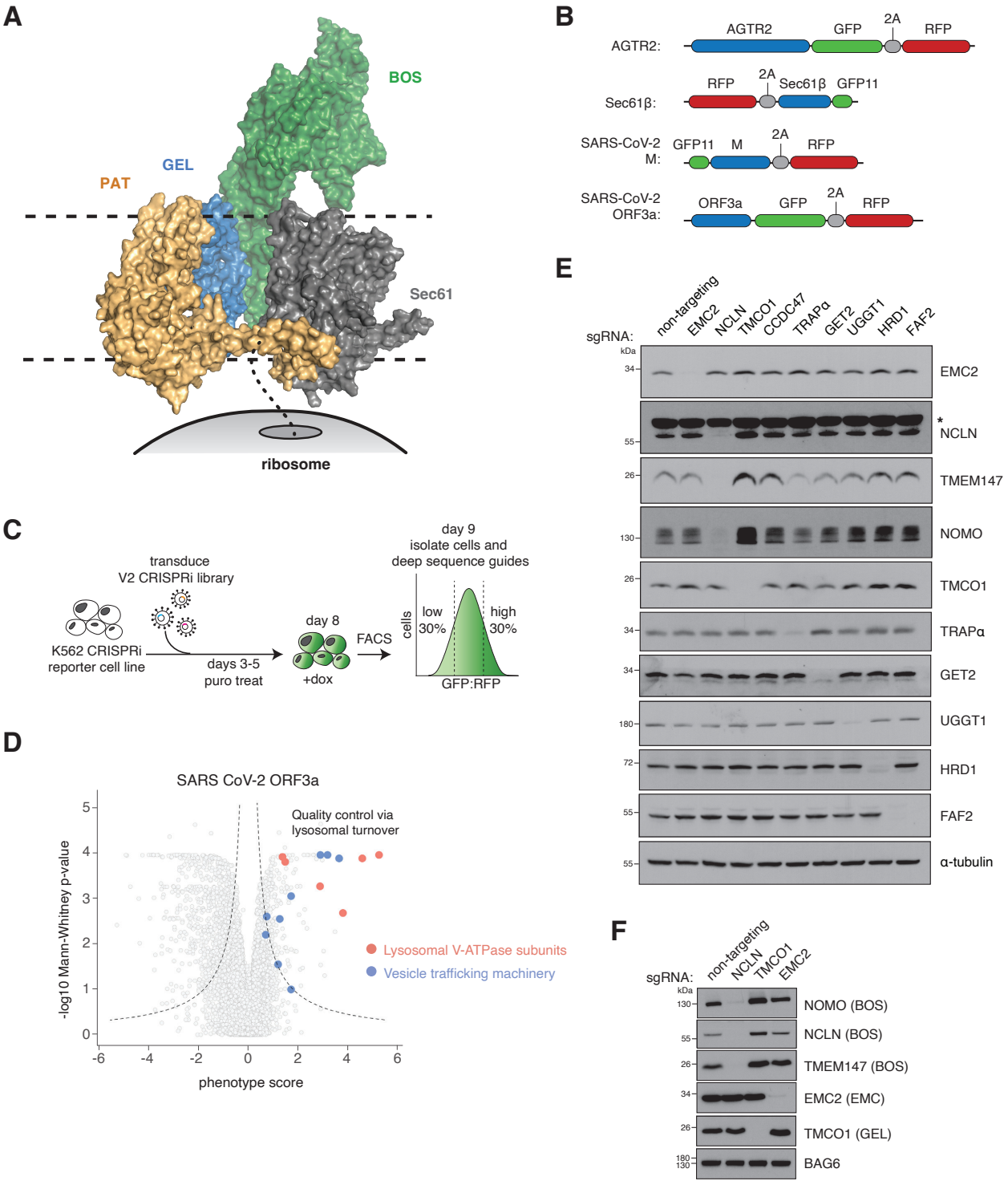

Figure S2.1. Genome-wide analysis of membrane protein biogenesis factors.

(A) Surface representation of the ER MPT (PDB ID: 6W6L) (McGilvray et al., 2020) with respect to the ribosome. The Sec61 translocon is shown in grey, BOS complex in green, GEL complex in blue, and the PAT complex in dark beige. The dotted line emerging from the ribosome represents the nascent chain of protein being inserted by the translocon. (**B**) Schematic of reporters used in CRISPRi screens. **(C)** Schematic of CRISPRi screens performed on the GPCR AGTR2, the TA Sec61β, and the SARS-CoV-2 M, and SARS-CoV-2 ORF3a proteins. **(D)** Volcano plot of factors involved in the turnover of SARS-CoV-2 ORF3a protein. ORF3a is subject to degradation by the lysosome, and therefore knockdown of vesicle trafficking machinery and of the lysosomal V-ATPase subunits stabilizes the ORF3a reporter. **(E)** Analysis of gene knockdown in K562 cells by guides in CRISPRi screens. Samples were subject to SDS-PAGE and western blotting. **(F)** ER insertion machinery shows compensatory effects upon the loss of individual factors. Samples were generated and analyzed as in (E), in K562 cells. Note that loss of the GEL insertase (i.e. TMCO1) results in increased levels of the BOS complex subunits, NOMO, NCLN, and TMEM147. Additionally, loss of EMC2 results in increased levels of the GEL insertase, which includes the subunit TMCO1, an evolutionarily related insertase within the OXA1 superfamily of insertases (Anghel et al., 2017).

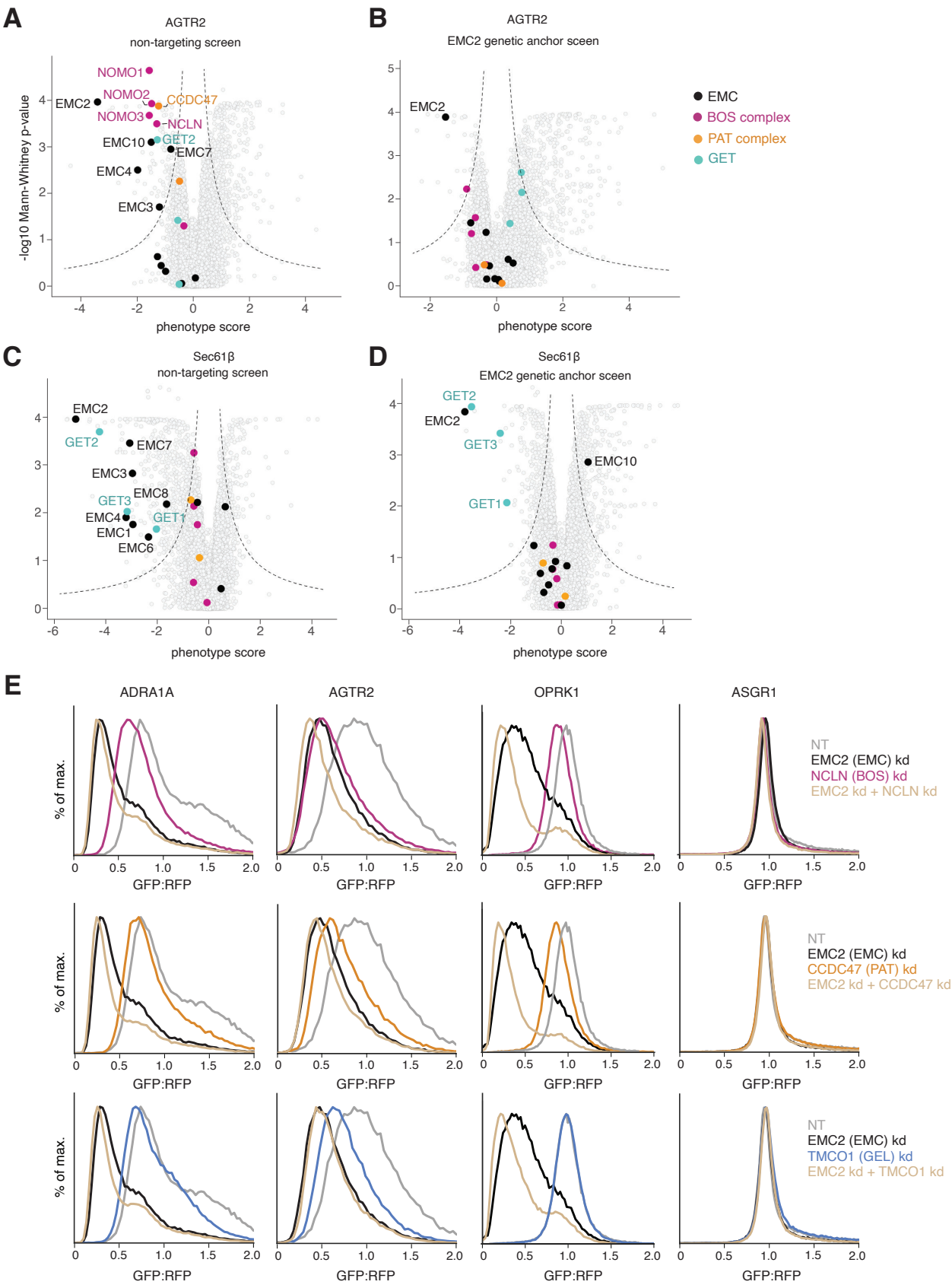
A
AGTR2
non-targeting screen
-log10 Mann-Whitney p-value
phenotype score
NOMO1
NOMO2
CCDC47
EMC2
NOMO3
NCLN
GET2
EMC10
EMC7
EMC4
EMC3
B
AGTR2
EMC2 genetic anchor sceen
phenotype score
EMC2
EMC
BOS complex
PAT complex
GET
C
Sec61β
non-targeting screen
phenotype score
EMC2
GET2
EMC7
EMC3
EMC8
GET3
EMC4
GET1
EMC1
EMC6
D
Sec61β
EMC2 genetic anchor sceen
phenotype score
GET2
EMC2
GET3
EMC10
GET1
E
ADRA1A
AGTR2
OPRK1
ASGR1
% of max.
GFP:RFP
NT
EMC2 (EMC) kd
NCLN (BOS) kd
EMC2 kd + NCLN kd
NT
EMC2 (EMC) kd
CCDC47 (PAT) kd
EMC2 kd + CCDC47 kd
NT
EMC2 (EMC) kd
TMCO1 (GEL) kd
EMC2 kd + TMCO1 kd

Figure S2.2. Genetic modifier screens identify epistatic and parallel pathways with the EMC.

(A) Volcano plot of a dual guide CRISPRi screen for the GPCR AGTR2 with the non-targeting (NT)-dual library. The EMC, MPT components, and the GET pathway components are highlighted. Dotted lines indicate the significance of a gene's effect on the stability of the AGTR2 reporter, and only significant gene hits from the screen are labeled. **(B)** As in (A), but with the EMC2-dual library. Note that components in the same pathway as EMC no longer have a destabilizing effect on the AGTR2 reporter. **(C)** As in (A) for the TA Sec61β with the NT-dual library (Guna, Page, et al., 2023). **(D)** As in (A) for Sec61β with the EMC2-dual library. Note that the EMC components are no longer significant but the GET pathway components, parallel to EMC, become more significant. **(E)** Representative histograms of the reporters assayed in Figure 2.3D-F.

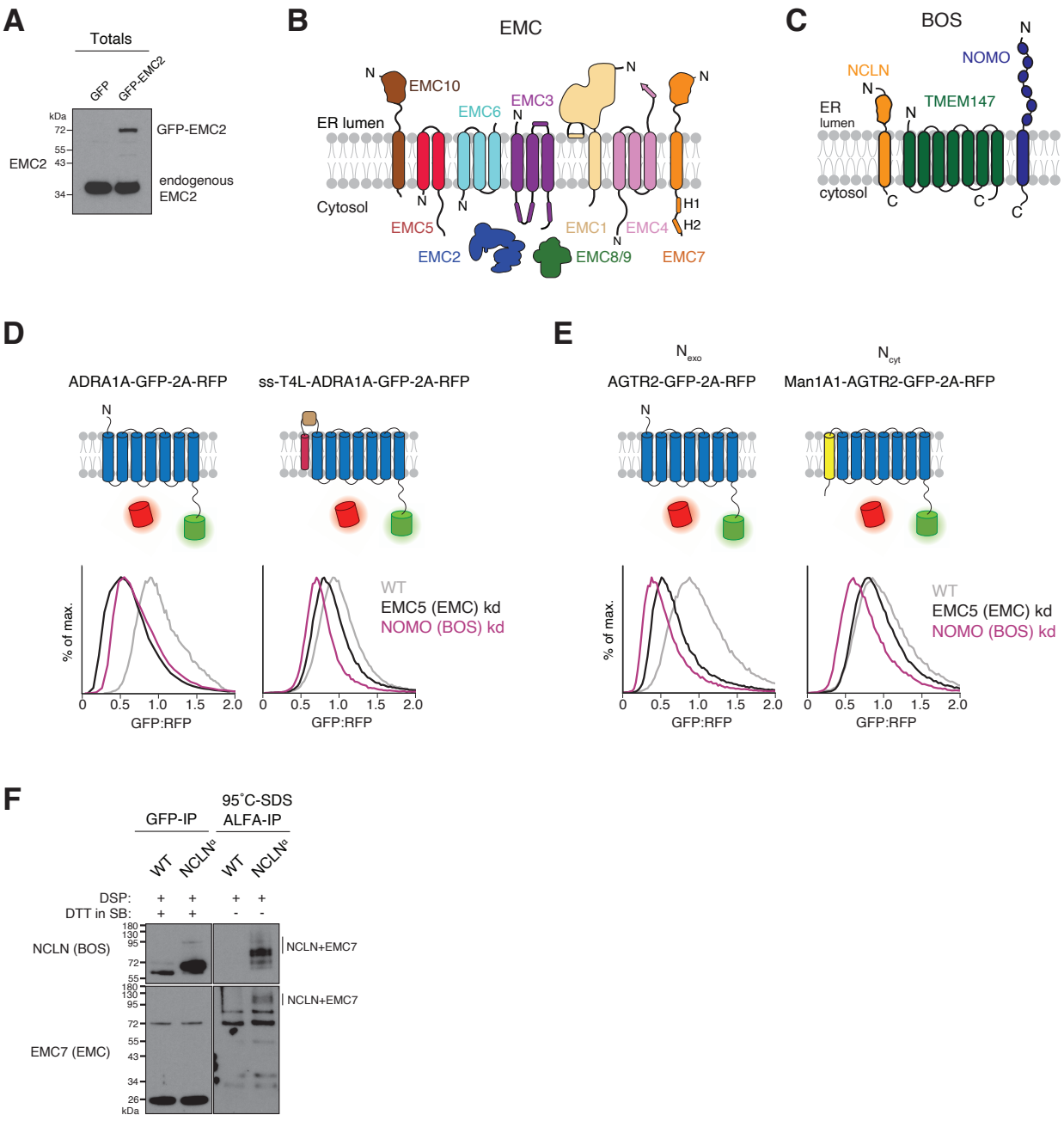

Figure S2.3. The EMC and BOS complex interactions.

(A) Total lysate samples of TREx cells expressing GFP and GFP-EMC2 were subjected to Western blotting with EMC2 antibody, as shown in Fig. 2.4A, with the lower molecular weight band being endogenous EMC2. **(B)** Topology of the nine-subunit human EMC. In mammals, the EMC is composed of seven membrane-spanning and two soluble subunits. The central insertase is composed of the two subunits EMC3 (a homolog of YidC and

member of the Oxa1 superfamily of insertases) and EMC6 (Bai et al., 2020; Miller-Vedam et al., 2020; Pleiner et al., 2020). In both yeast and humans, the EMC also contains a globular cytosolic domain (scaffolded around EMC2 and in mammals containing the redundant paralogs EMC8 or 9) and a large lumenal domain (composed of EMC1, 7, 10, and a single β-strand of EMC4). The function of the lumenal domain of the EMC is not known, but β-propellers, like those conserved in EMC1, are typically considered protein-protein interaction motifs. **(C)** Topology of the heterotrimeric BOS complex. NOMO contains 12 lumenal Ig-like repeats and a single C-terminal TMD. The single-spanning protein NCLN also has a large lumenal domain, and is homologous to nicastrin, a component of the γ-secretase complex (Shah et al., 2005). Finally, TMEM147 contains 7 TMDs and is homologous to APH-1 in the γ-secretase complex (Dettmer et al., 2010). **(D)** Flow cytometry analysis of the ratiometric ADRA1A protein reporter with or without the N-terminal fusion of the signal sequence (ss) of Pre-Prolactin followed by T4 Lysozyme (T4L), as described in Figure 2.4C. **(E)** Flow cytometry analysis of the ratiometric AGTR2 protein reporter with or without an N-terminal fusion of the first TMD of MAN1A1, a membrane protein of N_{cyt} topology. Note that both the N-terminal fusion of MAN1A1 and of (SS)-T4 lysozyme (Figure 2.4C) behave similarly and rescue the NOMO kd and EMC5 kd phenotypes. **(F)** Chemical crosslinking and immunoprecipitations from cell lines stably expressing either GFP-EMC2 alone or GFP-EMC2 and NCLN-ALFA were performed as in Figure 2.4E, except without addition of DTT during analysis of the protein samples by SDS-PAGE and western blot. Note that in both the NCLN and EMC7 western blots, there are multiple bands at ~92 kDa, consistent with a non-reduced band of NCLN+EMC7.

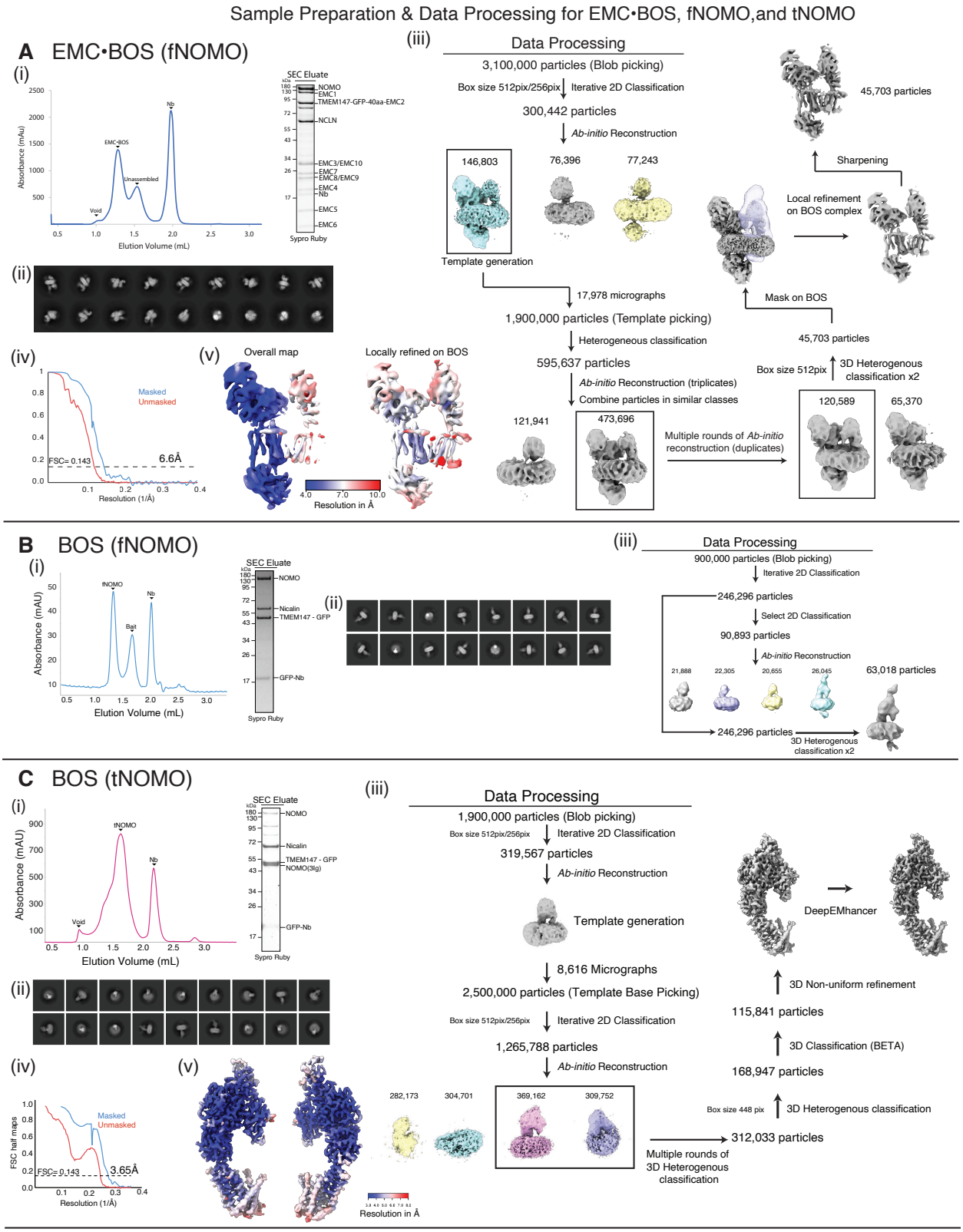

Figure S2.4. Sample preparation and data processing of EMC • BOS (fNOMO), BOS (fNOMO), and BOS (tNOMO).

(A) EMC • BOS (fNOMO)| (**i**) Sample preparation of EMC • fNOMO. Left: Size exclusion chromatogram of EMC • BOS (fNOMO). The main peak corresponds to the EMC • BOS (fNOMO) super-complex, while a smaller side peak likely corresponds to non-interacting EMC and BOS (fNOMO) complexes. Right: SDS-PAGE gel of the size exclusion eluate that contained EMC • BOS (fNOMO), stained by Sypro Ruby. The subunits of EMC and NOMO complexes are labeled. (**ii**) Representative 2D classes of EMC • BOS (fNOMO) complex. (**iii**) CryoSPARC processing scheme used to identify particles corresponding to the EMC • BOS (fNOMO) complex. CryoSPARC processing scheme employed for identifying particles related to the EMC • BOS (fNOMO) complex. All processing was carried out using CryoSPARC; for details, refer to Methods. (**iv**) GFSC curves for the masked and unmasked half maps of the EMC • BOS (fNOMO) complex map that was locally refined on BOS. Nominal resolution of 6.6 Å was determined using gold standard FSC cutoff = 0.143. (**v**) Resolution estimation of the EMC • BOS (fNOMO) complex (Left: overall map; Right: map that was locally refined on BOS complex), generated from cryoSPARC. **(B)** BOS (fNOMO)| (**i**) Sample preparation of BOS (fNOMO). Left: Size exclusion chromatogram of BOS (fNOMO) eluant after GFP-immunoprecipitation. Absorbance at 280 nm of the run was monitored to identify peak fractions that contained BOS (fNOMO). Right: shows the SDS-PAGE gel of the purified BOS (fNOMO) complex, stained by Sypro Ruby. All the components of the BOS (fNOMO) complex were present (NOMO, NCLN, and TMEM147-GFP), along with excess anti-GFP nanobody. (**ii**) Representative 2D classes of the BOS (fNOMO) complex. (**iii**) Data processing scheme in cryoSPARC v3.2-4.2.1 to identify a set of particles that correspond to the BOS (fNOMO) complex; for details, refer to Methods. **(C)** BOS (tNOMO)| (**i**) Sample preparation of BOS (tNOMO). Left: Size exclusion chromatogram of BOS (tNOMO) sample after GFP-immunoprecipitation and TEV cleavage. UV absorbance at 280nm was recorded to track protein fractions that contained BOS (tNOMO) complex. Right: SDS-PAGE gel of the peak fraction containing purified BOS (tNOMO) complex, stained by Sypro Ruby. The components of the BOS (tNOMO) complex were present (TMEM147, NCLN, and tNOMO). Since the bait used for purification was TMEM147-GFP, a portion of endogenous full-length NOMO was also detected. (**ii**) Representative 2D classes of BOS

(tNOMO) complex. (**iii**) Processing scheme used in cryoSPARC to identify particles that represent the BOS (tNOMO) complex; for details, refer to Methods. (iv) Gold Standard Fourier Shell Correlation (GFSC) curves for the masked and unmasked half maps of the overall BOS (tNOMO) complex map. Nominal resolution of 3.6 Å was determined using gold standard FSC cutoff = 0.143. (v) Resolution estimation of the BOS (tNOMO) complex, generated from cryoSPARC.

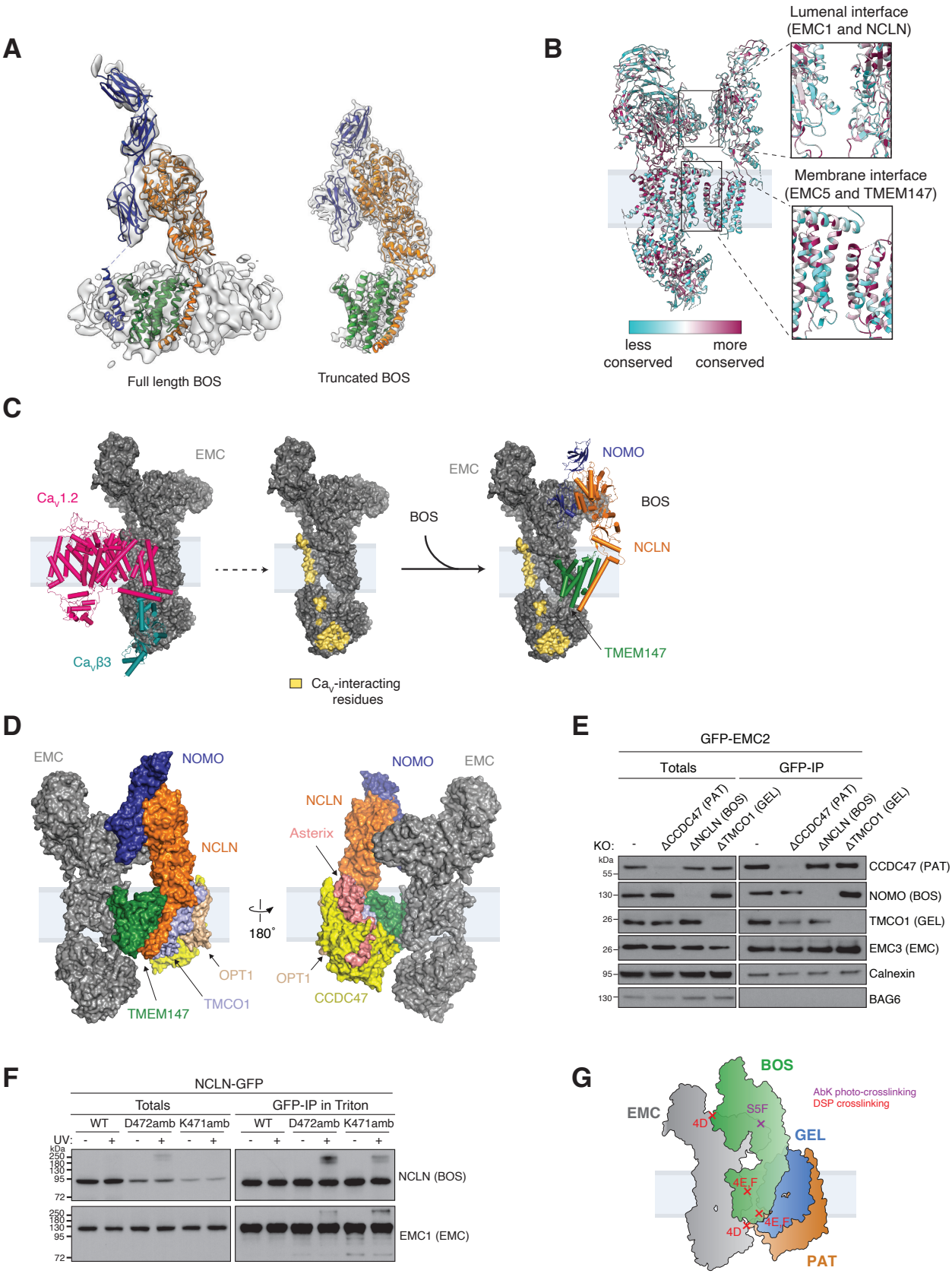

Figure S2.5. Analysis of the BOS complex and EMC • BOS (fNOMO) complex.

(A) Comparison between the full-length and truncated BOS complex models docked in their respective EM density maps. We observed that the truncated BOS complex retained the overall shape of the full-length complex, although only the two most C-terminal Ig-like domains are visible in the truncated BOS complex map. Here, NOMO is colored in blue, TMEM147 in green, and NCLN in orange. **(B)** Sequence conservation analysis of the EMC • BOS (fNOMO) complex. Left: Molecular model of EMC • BOS (fNOMO) with residues being colored by sequence conservation. Less conserved residues are in cyan and more conserved residues are in dark magenta. The sequence alignment contains sequences of EMC • BOS (fNOMO) from the following species: *Homo sapiens*, *Rattus norvegicus*, *Lygus hesperus*, *Callithrix jacchus*, *Chionoecetes opilio*, *Hydra vulgaris*, *Branchiostoma lanceolatum*, *Astyanax mexicanus*, *Castor canadensis*, *Pan troglodytes*, *Neovison vison*, *Macaca mulatta*, *Bactrocera latifrons*, *Bactrocera dorsalis*. Right: Inlets showing conservation of the lumenal interface (between EMC1 and NCLN) and membrane interface (between EMC5 and TMEM147). We observed a conserved patch near the lumenal membrane between EMC5 and TMEM147. **(C)** Comparison of the binding surfaces of EMC with either *Ca*V1.2/*Ca*Vβ3 complex or the BOS complex. The model was generated by aligning the EMC from its structure with *Ca*V1.2/*Ca*Vβ3 complex (PDB ID 8EOI) (Z. Chen et al., 2023) or BOS complex (this study). Left: Cylinder cartoon representation of the *Ca*V1.2/*Ca*Vβ3 complex (pink and dark cyan) (Z. Chen et al., 2023) and surface representation of the EMC (dark grey). Middle, the EMC residues that are interacting with *Ca*V1.2/*Ca*Vβ3 are highlighted in yellow. These residues were determined using the *InterfaceResidues.py* script in PyMOL. Right: Cylinder cartoon representation of the BOS complex (TMEM147 in green, NCLN in orange, and NOMO in blue) and surface representation of the EMC (dark grey) with the *Ca*V-interacting residues being highlighted to show that there is no overlap between the interacting interfaces with EMC by *Ca*V and the BOS complex. **(D)** Superposition of EMC • BOS (fNOMO) complex with the other components of the 'ER translocon' at Sec61. The model was generated by aligning TMEM147 from EMC • fNOMO (this study) and the ER translocon (PDB ID 6W6L) (McGilvray et al., 2020). Sec61 and CCDC47's latch helices were removed from this model for clarity. EMC occupies the same position as Sec61 at the translocon. **(E)** HEK

293T cells (WT, ΔCCDC47, ΔNCLN, or ΔTMCO1) stably expressing GFP-EMC2 were subjected to anti-GFP nanobody purification under native conditions. Samples were analyzed by SDS-PAGE and western blotting. Note that the interaction of the EMC with individual MPT components is independent of the EMC's interactions other MPT components. **(F)** Site-specific photocrosslinking of NCLN with EMC1. 3′-azibutyl-N-carbamoyl-lysine (AbK) was incorporated in a site-specific manner into the lumenal domain of NCLN-GFP at D472 and K471, which are positioned at the interface of the EMC and BOS in the holocomplex structure. Intact cells were exposed to UV light to induce photocrosslinking, and samples were subjected to denaturing immunoprecipitation in the presence of Triton X-100. Note production of a UV-dependent higher molecular weight band consistent in size with an EMC1-NCLN crosslinked species for both amber mutants as identified by Western blot. **(G)** Schematics representing the crosslinks observed in Figures 2.4D-F and S2.5F. Crosslinks in Figures 2.4D-F were achieved through DSP crosslinking, while those in Figure S2.5F were from AbK site-specific photo-crosslinking.

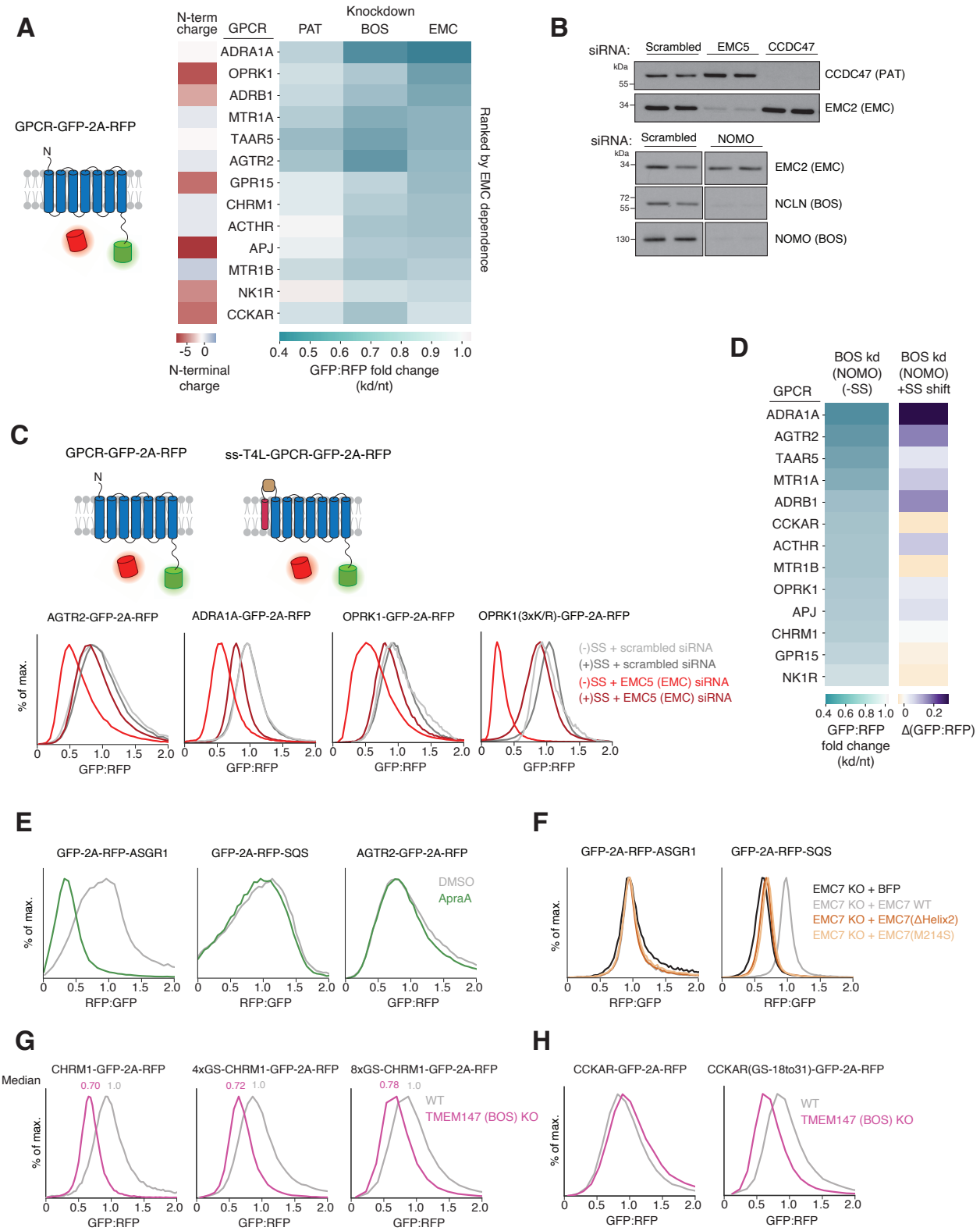

Figure S2.6. Differential effects of biogenesis factors on diverse GPCR substrates.

(A) A panel of GPCRs was assayed for EMC, BOS, or PAT dependence, as described in Figure 2.6. (Left) Schematic of the GPCR reporter constructs in the GFP-2A-RFP cassette. (Right) A heatmap displays dependence of each GPCR on the EMC for biogenesis. The net charge of the N-terminal soluble domain of each GPCR is indicated by the heatmap on the left. Note that there is limited correlation between EMC dependence and N-terminal charge. **(B)** Validation of knockdown experiments in RPE1 cells using siRNAs targeting EMC5, CCDC47 (PAT), NOMO (BOS) used in Figures 2.4 and 2.6. Cells were harvested and analyzed by SDS-PAGE and western blotting. **(C)** (Top) Schematic of the GPCR reporters in the GFP-2A-RFP cassette with or without the N-terminal signal sequence and T4 Lysozyme (ss-T4L). (Below) Flow cytometry analysis as in Figure 2.4C and 2.6C following EMC5 knockdown. **(D)** Bypassing the EMC for insertion of TMD1 of GPCRs reduces the dependence on the BOS complex. GPCR reporter constructs were generated using the GFP-2A-RFP cassette, with and without a signal sequence (SS) and T4 lysozyme fusion to the N-terminus of each GPCR. RPE1 cells were transduced with these reporters after siRNA knockdown of NOMO or a non-targeting control. (Left) A heatmap shows the GFP:RFP fold change of GPCRs without a signal sequence and is ranked by BOS dependence. (Right) A heatmap shows the change in stability between GPCRs containing the appended signal sequence and WT GPCRs. Darker purple indicates a positive shift in stability of the signal sequence-containing GPCR compared to the WT GPCR. **(E)** As in Figure 2.6E, for ASGR1, SQS, and AGTR2. **(F)** As in Figure 2.6F, for ASGR1 and SQS. **(G)** Lengthening the N-terminus of CHRM1 decreases BOS dependence. CHRM1 reporter constructs were generated by introducing a GS-linker of different lengths (4xGS and 8xGS) into its N-terminus. Expression of the resulting ratiometric fluorescent reporters was then tested in either wild type. The median value of each curve is displayed above each histogram. (WT) HEK 293T and TMEM147 (BOS) knockout (KO) cells. **(H)** Altering the secondary structure of the N-terminus of CCKAR increases BOS dependence. The AlphaFold2 (Mirdita et al., 2022) model for CCKAR is predicted to contain a β-hairpin from position 18 to 31 (CCKAR_HUMAN; P32238). We generated a mutant construct in which this hairpin is replaced by a matched glycine serine linker of the same length. The resulting ratiometric fluorescent CCKAR reporter constructs were expressed in either wild

type (WT) HEK 293T and TMEM147 (BOS) knockout (KO) cells. The results are displayed as a histogram.

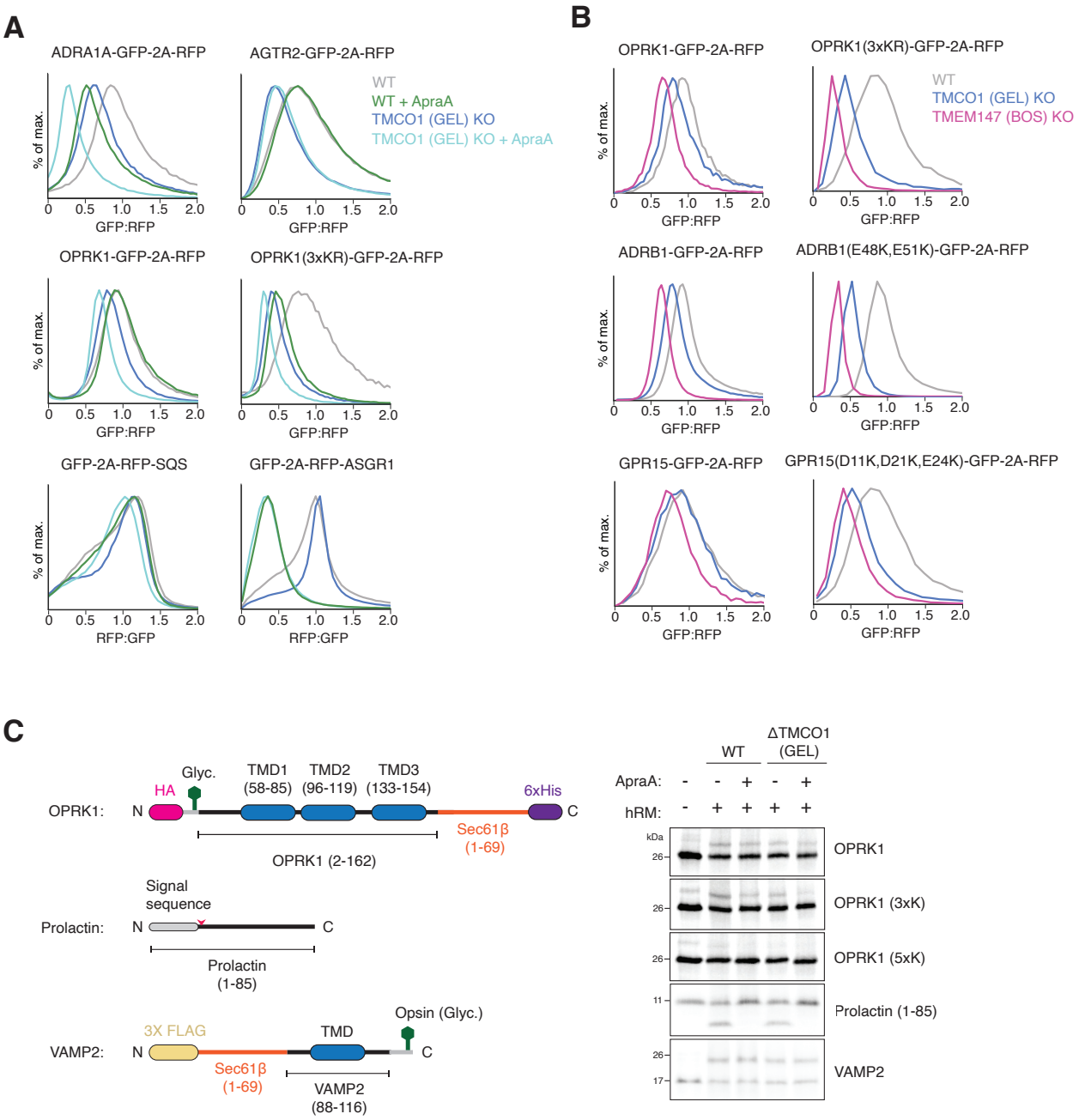

Figure S2.7. Substrates are routed through TMCO1 and Sec61 when unable to access EMC for insertion.

(A) Histograms as in Figure 2.6D-E, for WT and TMCO1 KO HEK 293T cells treated with Apratoxin A or a DMSO control. **(B)** Histograms of the reporters assayed in Figure 2.6I for WT, TMCO1 (GEL) KO, and TMEM147 (BOS) KO HEK 293T cells. **(C)** *In vitro* translation of OPRK1 (wildtype, 3xK, or 5xK), Prolactin, or VAMP2. (Left) Schematic of

constructs used in the translation reactions. The OPRK1 3xK and 5xK variants contain additional positive charges in the N-terminal soluble domain, as described in Figure 2.6B. (Right) Translations were performed in RRL in the presence of either 1 μM Apratoxin A in 0.5% DMSO or a 0.5% DMSO control and in hRMs derived from WT or TMCO1 KO cells. Insertion was determined by the fraction of glycosylated substrate (OPRK1 variants and VAMP2) or by signal sequence cleavage (Prolactin). In vitro insertion reactions tend to result in smaller phenotypes than those observed in cells, in part due to the absence of competing off-pathway reactions, including protein quality control and surveillance. Additionally, the time scales of an in vitro insertion reaction are beyond the physiologic time scale in cells, which gives substrates near-infinite time to be inserted into the ER microsomes.

CHAPTER 3:

A PATHWAY FOR INTEGRAL MEMBRANE PROTEIN QUALITY CONTROL WITHIN THE ER LUMEN

The following chapter is adapted from Nguyen, Bögeholz, Page et al., 2025 and modified according to the Caltech Thesis format.

Nguyen, V. N.*, Bögeholz, L. A. K.*, Page, K. R.*, Zhang, J., Chen, N., Wang, T.-Y., Mayank, A., Guna, A., Wang, M. L., Wohlschlegel, J., Chou, T.-F., Voorhees, R. M. (2025). A pathway for integral membrane protein quality control within the ER lumen. Manuscript in preparation.

3.1. Abstract

At the endoplasmic reticulum (ER), membrane protein biogenesis is tightly monitored to ensure misfolded or unassembled subunits are degraded to protect cellular homeostasis. Using genome-wide CRISPR screens, we identified thioredoxin domain containing protein 15 (TXNDC15), and showed that it is required for the quality control of a class of membrane proteins through binding the E3-ubiquitin ligase MARCHF6. TXNDC15 has evolved to repurpose its thioredoxin reductase domain to promote degradation of substrates with membrane proximal hydrophobic degrons exposed to the ER lumen. Patient mutations to TXNDC15 that cause the ciliopathy disorder Meckel-Gruber syndrome led to defects both in quality control in the ER and primary cilia formation. TXNDC15 therefore defines a distinct fourth branch of ER associated degradation with implications for proteostasis and disease.

3.2. Main text

Nearly seven thousand secreted and integral membrane proteins are encoded by the human genome, the majority of which must be inserted, modified, and assembled at the ER. To prevent cytotoxic effects, ER-associated degradation (ERAD) is responsible for the recognition, extraction and degradation of proteins that fail at any step of this biogenesis process (McCracken & Brodsky, 1996; Werner et al., 1996; Neal et al., 2018). Classically, ERAD has been divided into three branches that specialize in quality control within each of the three ER compartments: the lumen (ERAD-L), membrane (ERAD-M), and cytosol (ERAD-C) (P. Carvalho et al., 2006; Huyer et al., 2004; Vashist & Ng, 2004). However, given the enormous biophysical and topological diversity in the human membrane proteome, the known ERAD factors alone cannot be responsible for surveillance of the thousands of nascent proteins that must transit through the ER.

One class of proteins that pose unique challenges for quality control are multisubunit membrane protein complexes (Huttlin et al., 2017; Marsh & Teichmann, 2015). Complex assembly is highly regulated because unassembled subunits expose thermodynamically unfavorable intersubunit interfaces to the crowded cellular environment, which are prone to aggregation and non-specific interactions (Juszkiewicz & Hegde, 2018; Pleiner et al., 2021). In a healthy cell, a small number of orphan subunits, which are synthesized in excess or remain unassembled, must be identified and degraded (Harper & Bennett, 2016; Sung et al., 2016; Yanagitani et al., 2017). This problem is exacerbated in many diseases, including cancers and trisomies, where abnormal gene copy number results in unusually high levels of excess subunits (Dephoure et al., 2014; Manchado et al., 2012). The cell therefore must have evolved mechanisms to differentiate a nascent subunit contextually: orphan subunits must be efficiently degraded, while the identical subunit in complex with the appropriate binding partners is ignored (Natarajan et al., 2020). In mammals, ~40% of membrane proteins must assemble into higher order oligomeric complexes, which include hundreds of ion channels, receptors, and transporters (Babu et al., 2012). Despite the physiologic importance of the final assembled complexes, and the critical role subunit triage plays in cellular homeostasis, the factors responsible for the recognition and triage of orphan subunits are not known.

Identification of factors required for degradation of orphan GET1 in human cells

To study the quality control of subunit assembly in the ER, we chose a model system that exemplifies many of the common features of membrane protein complexes. GET1/2 is an essential, ubiquitously expressed, ER-resident membrane protein complex that contains only two subunits (Fig. 3.1A) (Vilardi et al., 2011, 2014; Yamamoto & Sakisaka, 2012). Importantly, the stability of both subunits is interdependent: when unassembled, orphan GET1 and GET2 are both efficiently degraded in a proteasome dependent manner (Fig. S3.1A) (Colombo et al., 2016; Rivera-Monroy et al., 2016). Second, the GET1/2 interface relies on both a membrane and lumenal surface, both of which have features common to many multisubunit complexes. For example, its intramembrane interface contains polar and charged residues that must be positioned within the hydrophobic core of the bilayer (McDowell et al., 2020, 2023). These residues are likely required for assembly and function, and are common in many ion channels and transporters (Cymer et al., 2012; F. Wang et al., 2014). As a result, several of GET1 and GET2's transmembrane domains (TMs) are poorly hydrophobic and thereby not predicted to independently integrate into the bilayer (Fig. 3.1B) (Ji et al., 2024). An additional predicted interface between the lumenal domains of GET1/2 is stabilized by hydrophobic interactions, which are common in assembly of many soluble complexes (Fig. 3.1C). Finally, unlike more complex systems, both GET1 and GET2 contain only three TMs, facilitating molecular analysis of their quality control. We chose to initially focus on the quality control of GET1, as earlier work had already outlined the biogenesis of GET2 (Inglis et al., 2020).

To systematically identify factors required for the recognition and degradation of orphan GET1, we generated a human K562 cell line that stably expressed green fluorescent protein (GFP)-tagged GET1 along with a translation normalization marker (red fluorescent protein, RFP) and the CRISPR inhibition (CRISPRi) machinery (Fig. 3.1D, Fig. S3.1B, Table S3.1) (Gilbert et al., 2014; Inglis et al., 2020; Itakura et al., 2016). Because GET1 is overexpressed above its endogenous levels, the majority of the reporter is unassembled, thereby allowing us to specifically study quality control of the orphan subunit (Fig. S3.1C). We reasoned that depletion of factors that affect the biogenesis or quality control of GET1

would result in either a decrease or increase in GET1-GFP:RFP fluorescence, respectively. Following transduction with a genome-wide CRISPRi sgRNA library (Horlbeck et al., 2016), cells displaying altered GFP:RFP fluorescence were isolated by FACS and deep sequenced (Fig. S3.1D).

As expected, depletion of its binding partner, GET2, resulted in decreased levels of our fluorescent GET1 reporter (Fig. 3.1D-E). Conversely, loss of known ERAD factors such as the ubiquitin-activating enzyme 1 UBA1 or the AAA+ ATPase VCP and its cofactor FAF2 stabilized GET1 (Fig. S3.1E-F) (Olzmann, Kopito, et al., 2013; Olzmann, Richter, et al., 2013; Schulman & Wade Harper, 2009). Notably however, the greatest stabilization of GET1 was caused by depletion of a protein of unknown function, thioredoxin domain containing 15 (TXNDC15).

TXNDC15 is a single spanning membrane protein predicted to contain an N-terminal signal sequence followed by a thioredoxin domain and C-terminal TM (Fig. S3.2A). In patients, mutations to TXNDC15 lead to the lethal ciliopathy disorder, Meckel-Gruber Syndrome (Radhakrishnan et al., 2019; Ridnõi et al., 2019; Shaheen et al., 2016). Despite this association with the primary cilia, we confirmed using an arrayed approach that loss of TXNDC15 stabilized our GET1 reporter (Fig. 3.1E). Because this ratiometric reporter acutely expresses both GET1 and a normalization marker from a single open reading frame, we concluded that the effect of TXNDC15 must occur post-translationally on newly synthesized GET1, as would be expected for a quality control factor. We further validated that depletion of TXNDC15 specifically affected the stability of unassembled GET1, but not the assembled GET1/2 complex (Fig. 3.1F, Fig. S3.2B-C). Finally, a panel of other membrane protein substrates with diverse topologies was unaffected by TXNDC15 depletion, excluding generalized dysregulation of the ER (Fig. S3.2D).

To further probe the function of TXNDC15, we next performed a genetic modifier screen (Fig. 3.1G, Fig. S3.3, Table S3.2-3) (Guna, Page, et al., 2023). We observed that depletion of several known ERAD factors — including DERL2 (Derlin-2 in yeast), SYVN1 (Hrd1 in yeast), SEL1L (Hrd3 in yeast) — had an enhanced phenotype in the TXNDC15 knockdown compared to the control screen (Bays et al., 2001; Christianson et al., 2008,

2011; Lilley & Ploegh, 2004; Stein et al., 2014). The increased effects of these canonical ERAD factors upon depletion of TXNDC15 is indicative of function in parallel pathways, consistent with a potential role for TXNDC15 in membrane protein quality control.

TXNDC15 is an ERAD factor required for degradation of orphan GET1

We next reasoned that for TXNDC15 to be a bona fide ERAD factor, it must have the following features. (i) In order to directly function in degradation of an ER-resident substrate like GET1, TXNDC15 must have a population outside of the primary cilia; (ii) because it has no predicted ubiquitin interacting or ligase domains it must work in tandem with other quality control machinery in the ER; and (iii) finally, if it plays a direct role in recognition or degradation, it must physically associate with its substrates.

To test this, we first generated a human cell line in which TXNDC15 was fluorescently tagged at its endogenous locus (Fig. S3.4). We found that the majority of TXNDC15 is localized to the ER (Fig. 3.2A), and we were unable to identify a population in the primary cilia even when overexpressed (Fig. S3.5). This is consistent with an earlier study that reported that overexpressed TXNDC15 was found in the ER (Yamazaki et al., 2024). Second, using this endogenously tagged cell line, we immunoprecipitated TXNDC15 under native conditions to identify co-purifying factors using mass spectrometry (Fig. S3.6A, Fig. 3.2B, Table S3.4). We found that the single most enriched hit is the ER-resident E3-ubiquitin ligase, MARCHF6 (Doa10 in yeast), previously characterized as a component of the ERAD-C pathway (Habeck et al., 2015; Hassink et al., 2005; Kreft et al., 2006; Schmidt et al., 2020).

Immunoprecipitation of exogenously expressed MARCHF6 from whole cells similarly resulted in a pronounced enrichment of TXNDC15 (Fig. 3.2C, Fig. S3.6B, Table S3.5), while recombinant co-expression and purification confirmed that TXNDC15-MARCHF6 form a stable, direct, nearly stoichiometric complex (Fig. 3.2D, Fig. S3.6C). Genetic epistasis experiments suggested that in addition to this physical association, TXNDC15 and MARCHF6 functioned together in degradation of our GET1 reporter (Fig. 3.2E). This

result contrasted with the relationship between TXNDC15 and SYVN1, an ER-resident E3 ubiquitin ligase associated with the ERAD-L pathway (Christianson et al., 2008, 2011; Kikkert et al., 2004). Consistent with our genetic modifier screen (Fig. 3.1G), depletion of TXNDC15 and SYVN1 had an additive effect on GET1 degradation, suggesting they function in parallel or partially redundant pathways.

Finally, to test whether TXNDC15 physically associates with its substrates, we translated orphan GET1 in a cell-free system in the presence of ER-microsomes purified from our endogenously tagged TXNDC15 cell line (Fig. 3.2F, Fig. S3.7A-B). We find that immunoprecipitation of TXNDC15 specifically enriched GET1, but not a membrane protein control that adopts the identical three-TM topology and fold (the SARS-CoV-2 M protein). The population of GET1 interacting with TXNDC15 further appeared to be preferentially ubiquitinated, consistent with a role in membrane protein quality control (Fig. S3.7A). Reporter experiments in cells confirmed that the M protein is not a TXNDC15 substrate (Fig. S3.2D).

Similarly, immunoprecipitation of orphan GET1 in vitro and from cells specifically enriched both TXNDC15 and MARCHF6 (Fig. 3.2G-H, Fig. S3.7C-D, Table S3.6-7). This association is more pronounced for GET1 expressed and purified from ER microsomes in vitro, where membrane extraction and proteasomal degradation do not occur efficiently. This enrichment is consistent with a transient interaction between GET1 and TXNDC15/MARCHF6 occurring in the early stages of substrate recognition and triage in the ER. Based on these results, we therefore concluded that TXNDC15 is an ER-resident quality control factor that functions with MARCHF6 for triage of substrates for degradation by the ubiquitin proteasome pathway.

TXNDC15 does not rely on its thioredoxin reductase activity for function

Having established a role for TXNDC15 in membrane protein quality control, we next sought to define its specific role in substrate recognition and degradation. First, we used a series of mutational experiments to identify which domains of TXNDC15 were required

for its activity (Fig. 3.3A, Fig. S3.8A-B). Leveraging our reporter assay, we established that while anchoring of the TXNDC15 lumenal domain to the ER membrane is required for degradation of orphan GET1, the sequence of the TM itself is dispensable. We found that the positioning of the TXNDC15 lumenal domain relative to the bilayer is also important, because insertion of a flexible linker between the TM and thioredoxin domain was tolerated, while a rigid helical linker was not (Fig. 3.3B, Fig. S3.8C-E). Therefore, the integrity and localization of the TXNDC15 lumenal domain is both necessary and sufficient for its role in membrane protein quality control.

Given the sequence similarity of TXNDC15 to other thioredoxin reductases, we next tested whether enzymatic activity was required for degradation of its substrates. Sequence alignment of TXNDC15 suggested it is missing one of the required catalytic cysteines (Fig. 3.3C), and mutation of its remaining active-site cysteine did not affect its quality control function (Fig. 3.3D, Fig. S3.8F). While we observed a decrease in TXNDC15 activity upon mutation of both this active site cysteine and an additional cysteine at position 210, we cannot differentiate between a potential structural role for these residues and a direct effect on TXNDC15 function (Yamazaki et al., 2024). Therefore, we tested whether TXNDC15 substrates must themselves contain cysteines. Mutation of all three cysteines in GET1 does not decrease its dependence on TXNDC15 for degradation (Fig. 3.3E). Therefore, TXNDC15's role in ERAD cannot rely on its thioreductase activity or require disulfide bond formation with its substrates. Although another thioredoxin domain containing protein, TXNDC11, is an ERAD factor that reduces disulfide bonds to enable degradation of its substrates (Timms et al., 2016), we concluded that TXNDC15 must have evolved a distinct role in recognition and/or degradation of its clients that relies on a non-enzymatic function of its lumenal domain.

TXNDC15 recognizes substrates in the ER lumen

The requirement of the lumenal domain of TXNDC15 was unexpected, because we anticipated that recognition of orphan GET1 would occur via exposed residues within the GET1/2 interface, which is localized primarily within the bilayer. We therefore sought to

define the features of GET1 recognized by TXNDC15. To do this, we leveraged our matched pair of TXNDC15 dependent and independent clients with identical topology: GET1 and M. By generating a series of chimeras, we found that in cells replacing TM3 and the loop between TMs 1 and 2 (loop 2) of GET1 with those from M completely abolished the effect of TXNDC15 on substrate stability (Fig. 3.4A). In vitro, replacing TM3 of M with that from GET1 was sufficient to allow physical association and thereby co-purification with TXNDC15 (Fig. 3.4B, Fig. S3.9A). Together these data defined the GET1 degron as within loop 2 and TM3, which are necessary and sufficient for recognition by TXNDC15.

In the assembled GET1/2 complex, we would expect all three TMs to be inserted into the bilayer, exposing a bipartite subunit interface within the bilayer and ER lumen (Fig. 3.1C). However, we wondered if GET1 may misfold when unassembled, as has been shown for other membrane protein subunits (H. J. F. Carvalho et al., 2019; Inglis et al., 2020; Pleiner et al., 2020). Indeed, consistent with earlier work (Sun & Mariappan, 2020), we found that an engineered glycosylation site on the C-terminus of unassembled GET1 is efficiently modified when expressed in vitro, suggesting that the third, poorly hydrophobic TM becomes exposed to the ER lumen during biogenesis (Fig. 3.4C). Comparison with protease protection experiments suggested that this exposure of TM3 is transient and the majority of GET1 eventually adopts the correct fold (Fig. S3.9B). We found that increasing the hydrophobicity of TM3 in GET1 correlated with both increased dependence on TXNDC15 for degradation in cells (Fig. 3.4D, Table S3.8) and increased association with TXNDC15 in vitro (Fig. 3.4E). Conversely, mutations that disrupted a hydrophobic patch within the GET1 loop 2 decreased TXNDC15 dependence (Fig. 3.4F). These results suggested a model in which TXNDC15 recognizes an exposed, hydrophobic degron in the ER lumen. In the case of GET1, this is composed of loop 2 and TM3, which would be shielded upon assembly with GET2 (Fig. 3.1C), providing a mechanism for differentiating between the assembled and unassembled subunit.

Consistent with this model, while we previously found that increasing the distance of the TXNDC15 lumenal domain from the bilayer disrupts its quality control activity, this can

be partially rescued by a commensurate extension of the length of loop 2 in GET1 (Fig. S3.10). Therefore, the relative positioning between the functional domain of TXNDC15 and the substrate degron is critical for its role in ERAD. Using a series of rescue experiments, we further found that TXNDC15 and MARCHF6 must play unique, non-redundant roles in substrate recognition and degradation (Fig. 3.4G, Fig. S3.11). Overexpression of either TXNDC15 or MARCHF6 cannot replace loss of the other factor. Since TXNDC15 itself has no other predicted enzymatic domains, we propose that TXNDC15 relies on its lumenal domain to recognize substrates, securing them in proximity to MARCHF6 for ubiquitination and subsequent degradation.

To identify additional TXNDC15 substrates beyond GET1, we generated a panel of reporters using our ratiometric fluorescent system (Fig. 3.5A). We found that several other membrane proteins, including other membrane protein subunits, depend on TXNDC15 for their degradation. Notably, degradation of orphan GET2, which had been previously shown to misfold by exposing its TM to the ER lumen, is also partially TXNDC15 dependent (H. J. F. Carvalho et al., 2019; Inglis et al., 2020). Therefore, the recognition of a hydrophobic degron adjacent to the membrane is a unifying feature of TXNDC15-substrates. Further, TXNDC15 appeared to define a distinct and previously undefined branch of ERAD, as the stability of canonical ERAD-C, -M, and -L substrates (Leto et al., 2019) were unaffected by depletion of TXNDC15. We therefore propose that TXNDC15-MARCHF6 are responsible for the surveillance of a unique set of membrane protein clients at the ER.

Together these data suggest a model in which TXNDC15 has repurposed its lumenal thioredoxin domain to identify substrates that expose hydrophobic sequences within a specific distance from the lumenal leaflet of the bilayer (Fig. 3.5B). Some substrates of TXNDC15 include membrane protein subunits that misfold prior to assembly, such as GET1 and GET2, aberrantly presenting hydrophobic interface domains that are shielded upon assembly. The recognition of an exposed subunit interface provides a molecular mechanism for differentiating assembled versus orphan subunits and specifically targeting only the latter for degradation. We hypothesize that these TXNDC15 substrates may not efficiently access the canonical ERAD-L pathway due to their close proximity to the

bilayer, perhaps due to steric or other topological constraints. TXNDC15 therefore functions as an adaptor, bringing clients that do not independently interact with MARCHF6 in proximity to the E3-ligase for ubiquitination, triggering their eventual extraction, and proteasome-dependent degradation. Because MARCHF6 is localized primarily within the membrane, with only short loops in the ER lumen and its functional RING domain in the cytosol, TXNDC15 expands the pool of clients that can rely on this E3-ligase for their surveillance and quality control. Therefore, we propose that TXNDC15 defines a fourth branch of ERAD responsible for surveillance of the membrane proximal region of the ER lumen.

Lastly, considering that imbalances in proteostasis are associated with many diseases, we reasoned that loss of TXNDC15-quality control function could be related to its role in ciliopathies. Earlier work has established that depletion of TXNDC15 in cells leads to defects in cilia function, morphology, and signaling (Breslow et al., 2018). Mutations in TXNDC15 in patients lead to a severe ciliopathy known as Meckel-Gruber syndrome, which is characterized by malformation of the central nervous system, polycystic kidneys, and polydactyly that together lead to death in utero or soon after birth (Radhakrishnan et al., 2019; Ridnõi et al., 2019; Shaheen et al., 2016; Smith et al., 2006). One such patient mutation is characterized by a five amino acid deletion to the TXNDC15 thioredoxin domain (Fig. 3.5C) (Shaheen et al., 2016). We found that expression of wildtype, but not the mutant TXNDC15 was able to rescue defects in cilia formation and morphology in human RPE1 and mouse 3T3 cells, despite accumulation to similar levels (Fig. 3.5D-G, Fig. S3.12A-E). Similarly, using our reporter system we found that this patient mutation was unable to rescue degradation of orphan GET1 (Fig. 3.5H, Fig. S3.11F-G). Therefore, there is a correlation between the function of TXNDC15 in membrane protein quality control at the ER and ciliogenesis at the plasma membrane.

3.3. Conclusions

Because membrane proteins destined for the primary cilia must passage through the ER, we propose that loss of TXNDC15 activity perturbs the ciliary protein landscape, leading

to defects in ciliogenesis and function. Indeed, overexpression of some ciliary proteins have been previously shown to lead to changes in cilia morphology and signaling. A biochemical explanation for TXNDC15 function sets the stage for further studies of ciliogenesis and the associated defects that lead to human disease.

Further, the molecular characterization of this TXNDC15-dependent branch of ERAD exemplifies two important and potentially general features of protein quality control. First, it provides an explanation for how the cell can rely on a small number of E3 ligases to surveille the enormous diversity within the human membrane proteome. The use of adaptors that mediate interactions with specific classes of substrates is an efficient strategy to expand the repertoire of a single E3 ligase. Indeed, this is analogous to the many adaptors that have been characterized in the cytosol that function with HSP70, VCP, and Skp1-Cullin E3 ubiquitin ligases to maintain proteostasis (Madsen et al., 2009; Skowyra et al., 1997; Zuiderweg et al., 2017). Second, this work suggests there is additional spatial specialization in protein quality control beyond that defined by the bilayer alone (Young et al., 2003). An important area for future research is to characterize these membraneless jurisdictions in other organelles and organelle contact sites and delineate how they are defined and integrated. However, it is clear that while a pathway may specialize in a particular region and set of substrates, there is substantial overlap within the protein quality control network. Compensatory and redundant pathways are critical for robustness and may be particularly important in post-mitotic cell types that must survive throughout an eighty- to ninety-year lifespan.

3.4. Acknowledgements

We thank all members of the Voorhees lab, the Caltech Flow Cytometry facility, the Caltech Millard and Muriel Jacobs Genetics and Genomics Laboratory and the Caltech Biological Imaging Facility.

Funding: This work was supported by: the Heritage Medical Research Institute (RMV), an NSF-CAREER award 2145029 (RMV), the Burrough's Wellcome Fund (RMV), the Pew-Stewart Foundation (RMV), the Sontag Foundation (RMV), an NIH K99 Award

(AG), a Rosen Family fellowship (KRP), and an Arie Jan Haagen-Smit Fellowship (KRP). RMV is a Howard Hughes Medical Institute Freeman Hrabowski Scholar.

Competing interests: RMV is a consultant and equity holder of Gate Bioscience. The authors have no additional competing financial interests.

Data and materials availability: The genome wide CRISPR screens datasets are provided as Table S3.1, S3.2, and S3.3. Mass spectrometry datasets are provided as Table S3.4, S3.5, S3.6, S3.7. FACS data plotted in Fig. 3.4D are provided as Table S3.8. All other data are available in the main text or the supplementary materials.

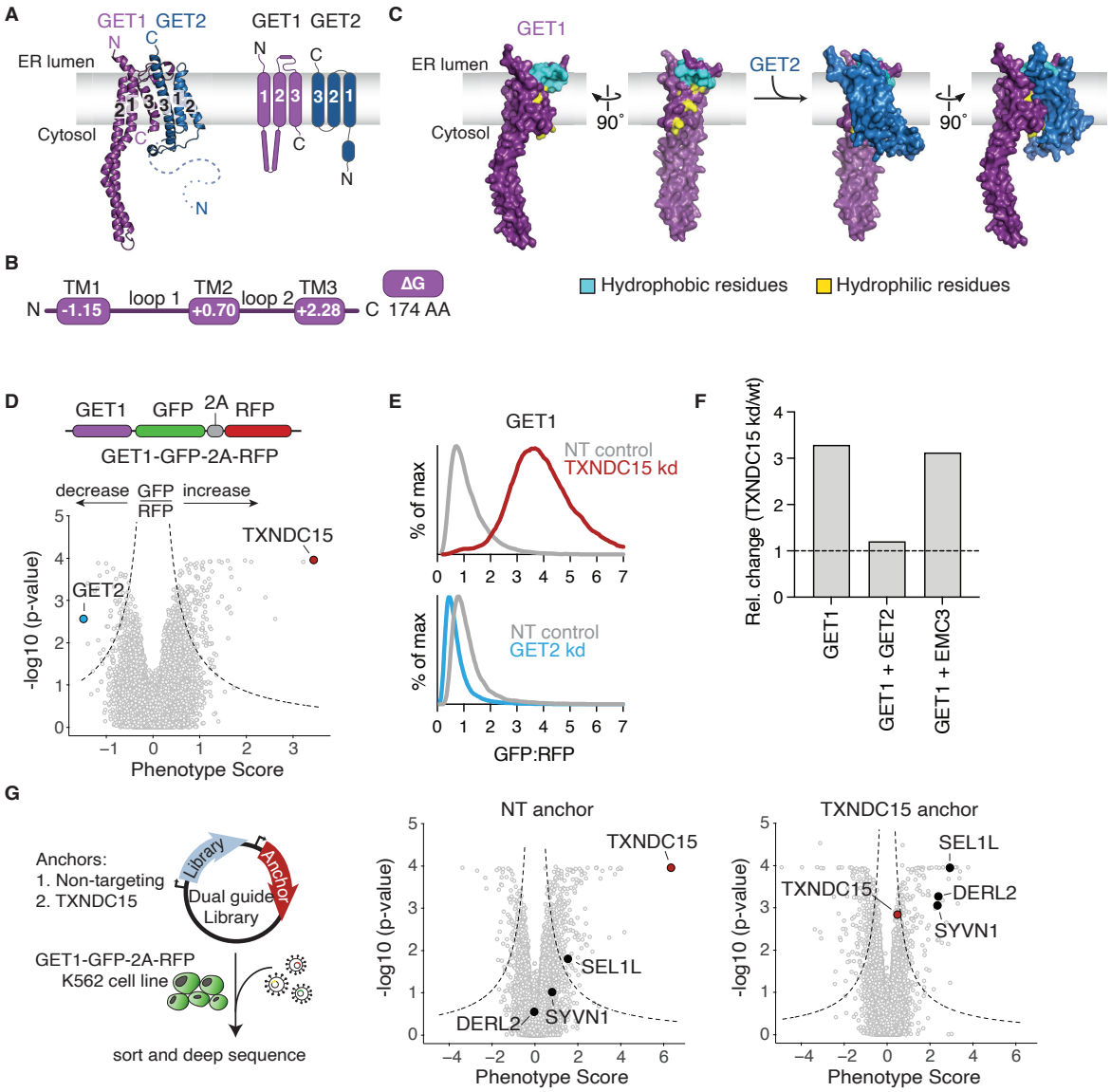

Figure 3.1. Systematic identification of factors required for degradation of orphan GET1.

(A) Schematic of the GET1/2 complex. (Left) AlphaFold 3 predicted model of GET1 (purple) and GET2 (blue) aligned to the GET1/2 structure (PDB: 6SO5) (McDowell et al., 2020). (Right) Topology of the GET1/2 complex. **(B)** Insertion propensity of the TMs of GET1, as calculated using ΔG (Hessa et al., 2007). Note TM3 contains several polar and charged residues while loop2 has a predicted hydrophobic patch positioned at the interface of the GET1/2 complex as shown in (C). **(C)** GET1/2 model as in (A) but represented as a space-filling model. The GET1/2 complex contains a bipartite intersubunit interface

composed of a hydrophilic surface within the membrane (yellow; mediated by GET1 TM3) and a hydrophobic patch in the ER lumen (cyan; mediated by GET1 loop 2). Assembly with GET2 shields these surfaces from exposure to the hydrophobic core of the bilayer and the aqueous lumen. **(D)** Orphan GET1-GFP is expressed along with a normalization marker (RFP) from a single open reading frame to query factors required for its post-translational recognition and degradation (Fig. S3.1B). The volcano plot summarizes the phenotype for the three strongest single guide RNAs (sgRNAs) versus Mann-Whitney p-values calculated from two replicates of the genome wide CRISPRi screen. Individual genes are displayed in grey and specific factors that affect GET1 stability are highlighted and labeled. **(E)** Stability of the GET1 reporter described in (C) was assessed in K562 cells expressing a nontargeting (NT) control or the indicated sgRNA. GET1-GFP expression relative to a normalization control was determined by flow cytometry and displayed as a histogram. **(F)** Using the GET1 reporter shown in (D), the relative change in GET1 stability in TXNDC15 depleted compared to wt HEK293T cells (Fig. S3.2B) was measured in the presence and absence of saturating levels of exogenously expressed GET2. As a specificity control, the effect of EMC3 expression on GET1 stability is also displayed (Fig. S3.2C). The relative change (TXNDC15 kd/wt) was calculated from the medians of the displayed histograms in Fig. S3.2D and is shown here as a bar graph. The dashed line at 1 is the relative change expected for substrates that are not dependent on TXNDC15. **(G)** (Left) Schematic of the dual guide genetic modifier screening approach (Guna, Page, et al., 2023). To identify genetic modifiers of TXNDC15 across the genome we performed two screens using dual libraries expressing either a NT control (middle) or TXNDC15 guide (right) at the anchor position and all guides in the CRISPRi-v2 library at the variable position. Volcano plots summarizing the resulting CRISPRi modifier screens are shown, and statistically significant hits indicated by the dashed lines. TXNDC15 (red) and several known ERAD factors (black) are highlighted.

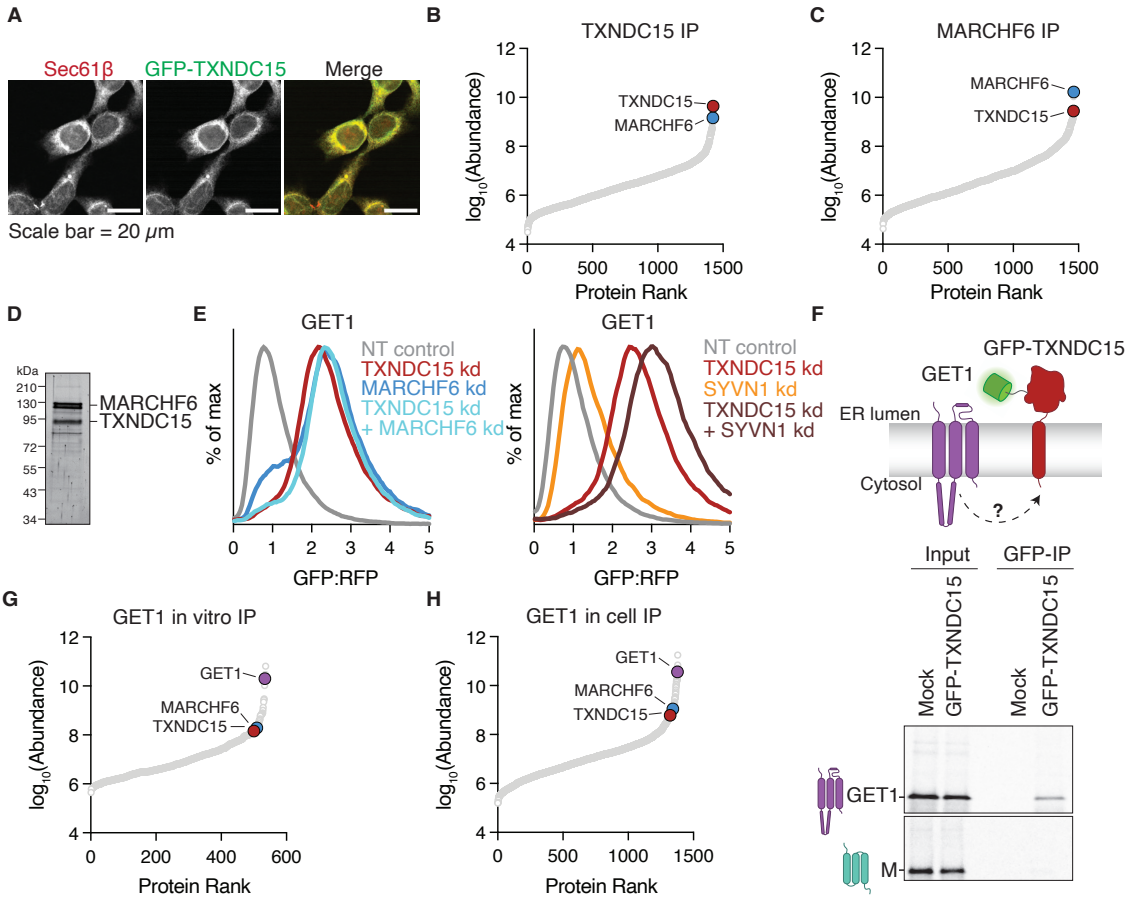

Figure 3.2. TXNDC15 is an ERAD factor that complexes and functions with MARCHF6.

(A) The localization of TXNDC15 was determined by imaging HEK293T cells in which a GFP tag was inserted at the endogenous TXNDC15 locus (Fig. S3.4). Staining for SEC61β was used as an ER marker. Imaging of cilia was subsequently performed using RPE1 cells overexpressing TXNDC15-GFP (Fig. S3.5). **(B)** Analysis of the interactome of endogenously expressed TXNDC15 using mass spectrometry. GFP-TXNDC15 was solubilized and purified from the endogenously tagged HEK293T cell line under native conditions in the detergent GDN. Co-purifying proteins were ranked by average abundance from two replicates, which is plotted against their log10(abundance). Background controls for non-specific interactions are shown in Fig. S3.6A. **(C)** As in (B) for exogenously expressed MARCHF6(C9A)-GFP from Expi293 cells (Fig. S3.6B). MARCHF6(C9A) was used as it is the more stable, catalytically inactive variant of MARCHF6. **(D)** Purification

of a recombinant MARCHF6(C9A)/TXNDC15 complex visualized using Sypro Ruby. Expi293 cells expressing TXNDC15-ALFA and MARCHF6(C9A)-GFP (Fig. S3.6C) were solubilized and immunoprecipitated using first the ALFA-nanobody and then the GFP-nanobody, sequentially. **(E)** Analysis of GET1 stability using the ratiometric fluorescent reporter system described in Fig. 3.1D following depletion of the indicated sgRNAs in K562 cells. Depletion of TXNDC15 and MARCHF6 together do not have an additive effect on GET1 stability in contrast to depletion of TXNDC15 and SYVN1 (Hrd1 in yeast). **(F)** Reconstitution of the TXNDC15 function in vitro. An ^{35}S-methionine labeled GET1 was translated in rabbit reticulocyte lysate (RRL) extract in the presence of human derived ER microsomes (hRMs) prepared from either wt HEK293T cells (mock, which serve as a background control) or HEK293T cells containing endogenously tagged GFP-TXNDC15. In parallel, the M protein from SARS-CoV-2, which adopts the identical 3-TM topology to GET1 but is stable as a monomer, was translated as a control. The membrane fraction was isolated using ultracentrifugation, solubilized in the detergent GDN (input), immunoprecipitated, and eluted under native conditions using the GFP-nanobody (GFP-IP). Samples were analyzed by SDS-PAGE and autoradiography. A longer exposure showing ubiquitination of orphan GET1 is displayed in Fig. S3.7A. Western blotting (WB) of the GFP-TXNDC15 bait in the elution samples is displayed in Fig. S3.7B. **(G)** Analysis of the interaction partners of nascent GET1 as determined by mass spectrometry. Following expression in RRL extract supplemented with hRMs, GET1-ALFA was immunoprecipitated using the ALFA nanobody under native conditions. Background controls are displayed in Fig. S3.7C. **(H)** As in (G), except using GET1-GFP expressed in HEK293T cells (Fig. S3.7D).

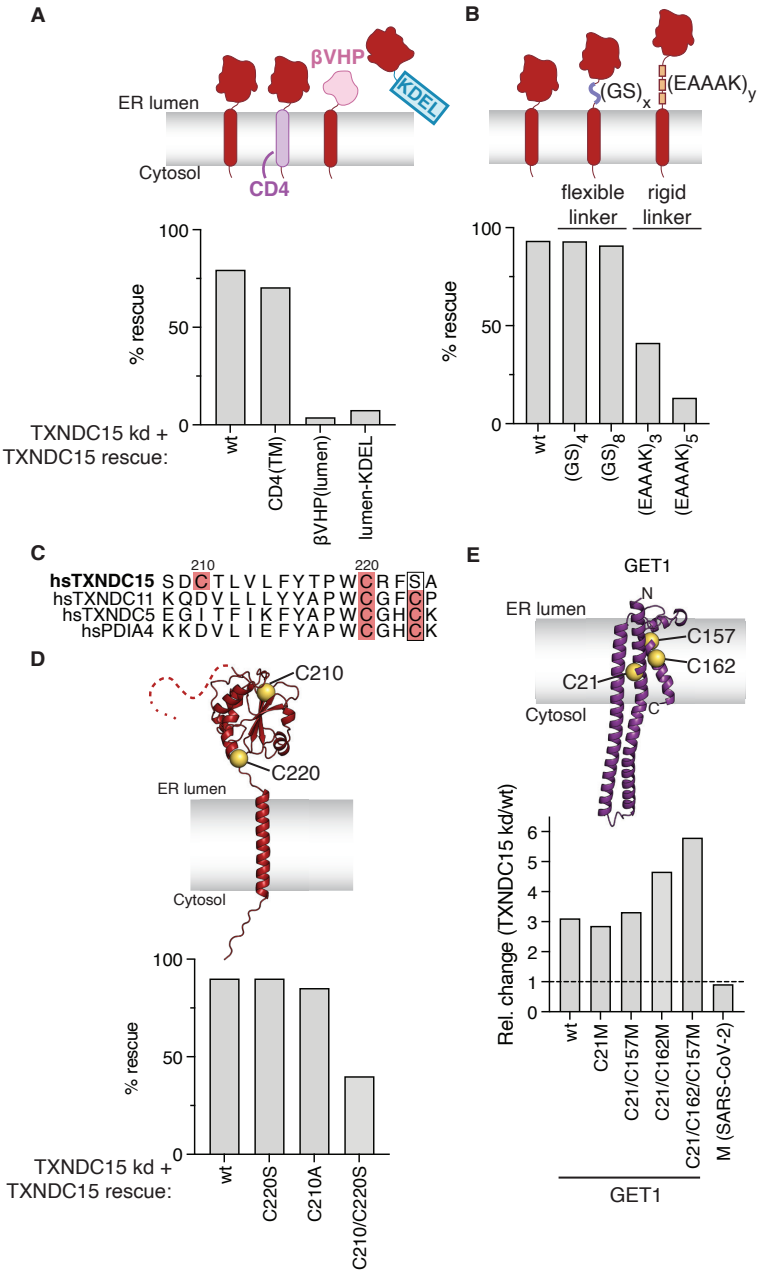

Figure 3.3. TXNDC15 does not require thioredoxin reductase activity for its role in ERAD.

(A) Rescue assay to identify domains essential for TXNDC15's quality control function in HEK 293T cells. The following constructs were transduced in TXNDC15 depleted cells: wt TXNDC15; the TXNDC15 lumenal domain fused to an unrelated TM from CD4 (CD4-TM); TXNDC15 with its lumenal domain replaced with the unrelated globular domain of bVHP (bVHP-lumen); or the TXNDC15 lumenal domain alone fused to a KDEL ER-

retention sequence (lumen-KDEL). Percent rescue (% rescue) was calculated relative to wt TXNDC15, as described in materials and methods. Histograms are shown in Fig. S3.8A. **(B)** As in (A) except for TXNDC15 constructs containing a mutagenized loop between the thioredoxin-like domain and TM. Histograms are shown in Fig. S3.8C-D. $(GS)_x$ denotes flexible linkers of x repeats and $(EAAAK)_y$ denotes rigid linkers of y repeats. **(C)** Multiple sequence alignment of the putative catalytic domain of human TXNDC15 with the indicated proteins in the disulfide isomerase (PDI) family (TXNDC11, TXNDC5, PDIA4). The position of the two catalytic cysteines (residues 220 and 223) are highlighted in red. **(D)** As in (B) for the indicated TXNDC15 cysteine mutants at locations 210 and 220 (shown as yellow spheres). Note that while TXNDC15 contains a cysteine at position 210, this is not within the putative thioredoxin reductase active site. **(E)** To assess whether substrate cysteines are required for TXNDC15's quality control function, we generated the indicated GET1 cysteine mutants (yellow spheres). The relative change in stability of each reporter was calculated as: median GFP:RFP fluorescence in TXNDC15 depleted cells/median GFP:RFP in wt cells. The dashed line at 1 is the relative change we would expect for substrates that are not dependent on TXNDC15.

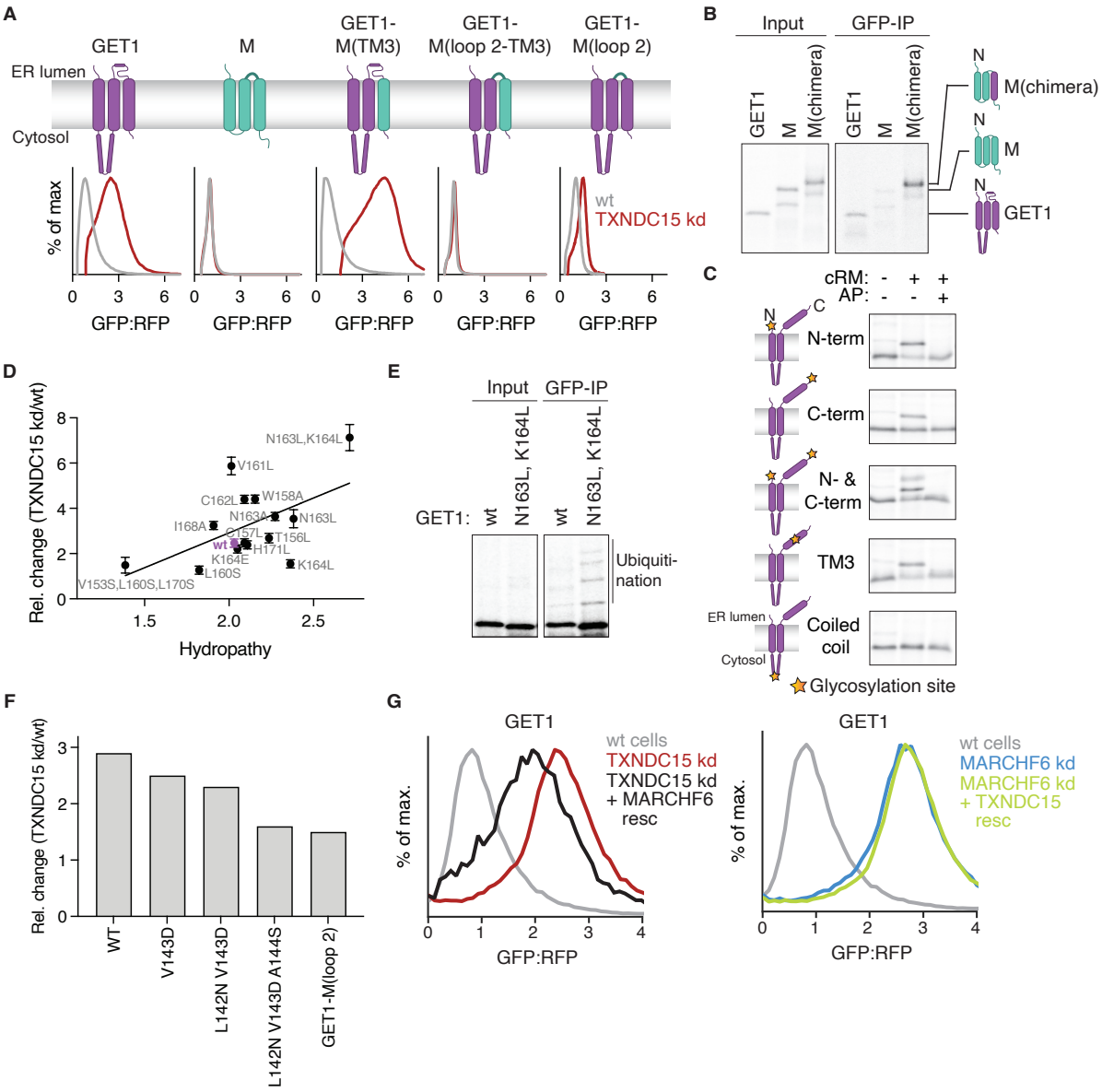

Figure 3.4. TXNDC15 recognizes a hydrophobic degron in the ER lumen.

(A) Analysis of GET1/M chimeras in wt compared to TXNDC15 depleted HEK293T cells by flow cytometry. Replacement of loop 2 and TM3 of GET1 with those from M eliminates the effect of TXNDC15 depletion on GET1 stability. In this experiment all constructs are fused to RFP-VAMP2(TM), to avoid perturbation of the C-terminus while permitting fluorescent detection (as described in the methods section). **(B)** As in Fig. 3.2F, ^{35}S-methionine labeled GET1, M (TXNDC15 independent), and the indicated chimeric construct of M (M(TM1-2) and GET1(TM3)) were translated in RRL in the presence of

hRMs containing endogenously tagged GFP-TXNDC15. The isolated membrane fraction was solubilized (input) and immunoprecipitated under native conditions using GFP-IP. Samples were analyzed by SDS-PAGE and autoradiography. **(C)** ^{35}S-methionine labeled GET1 with the indicated glycosylation tags (N-terminus, C-terminus, both N- and C-termini, TM3, within the cytosolic coiled coil) was translated in RRL in the presence of cRMs. Presence of a glycosylated species indicates exposure of the glycosylation site to the ER lumen. Acceptor peptide (AP) was used to validate that the observed molecular weight shift was specifically due to glycosylation. **(D)** Summary of the effects of TXNDC15 on the stability of a panel of mutations to TM3 of GET1 as determined by flow cytometry in HEK293T cells. Shown is the relative ratio of GET1-GFP:RFP in wt compared to TXNDC15 depleted cells versus the GRAVY hydropathy of TM3 (Kyte & Doolittle, 1982). For comparison, wt GET1 is highlighted in purple. Data were analyzed by linear fitting and error bars represent the standard deviation (n=3). **(E)** As in (B) but comparing wt GET1 with a mutant that both increases TM3 hydrophobicity and TXNDC15 dependence in cells shown in (E) (N163L, K164L). Higher molecular weight bands correspond on ubiquitinated species of GET1, which are enriched in the population of GET1 associated with TXNDC15. **(F)** To assess whether a hydrophobic patch in loop 2 of GET1 is recognized by TXNDC15 (see Fig. 3.1C), hydrophobic amino acids were mutated as follows: L142N, V143D, A144S. For comparison, the chimera where the whole loop is replaced with that from M(loop 2) is shown. Constructs were prepared in the GFP-2A-RFP-VAMP2(TM) cassette as in (A) and reporters were analyzed in HEK293 wt and TXNDC15 kd cell lines by flow cytometry. Disruption of the hydrophobic patch created by these three residues phenocopies replacement of the entire loop 2. **(G)** Analysis of GET1 ratiometric fluorescent reporter stability with MARCHF6 overexpression in TXNDC15 kd cells (left) and TXNDC15 overexpression in MARCHF6 kd cells on the GET1-GFP reporter (right). Expression of MARCHF6 cannot rescue TXNDC15 kd and expression of TXNDC15 cannot rescue MARCHF6 kd in HEK293T cells.

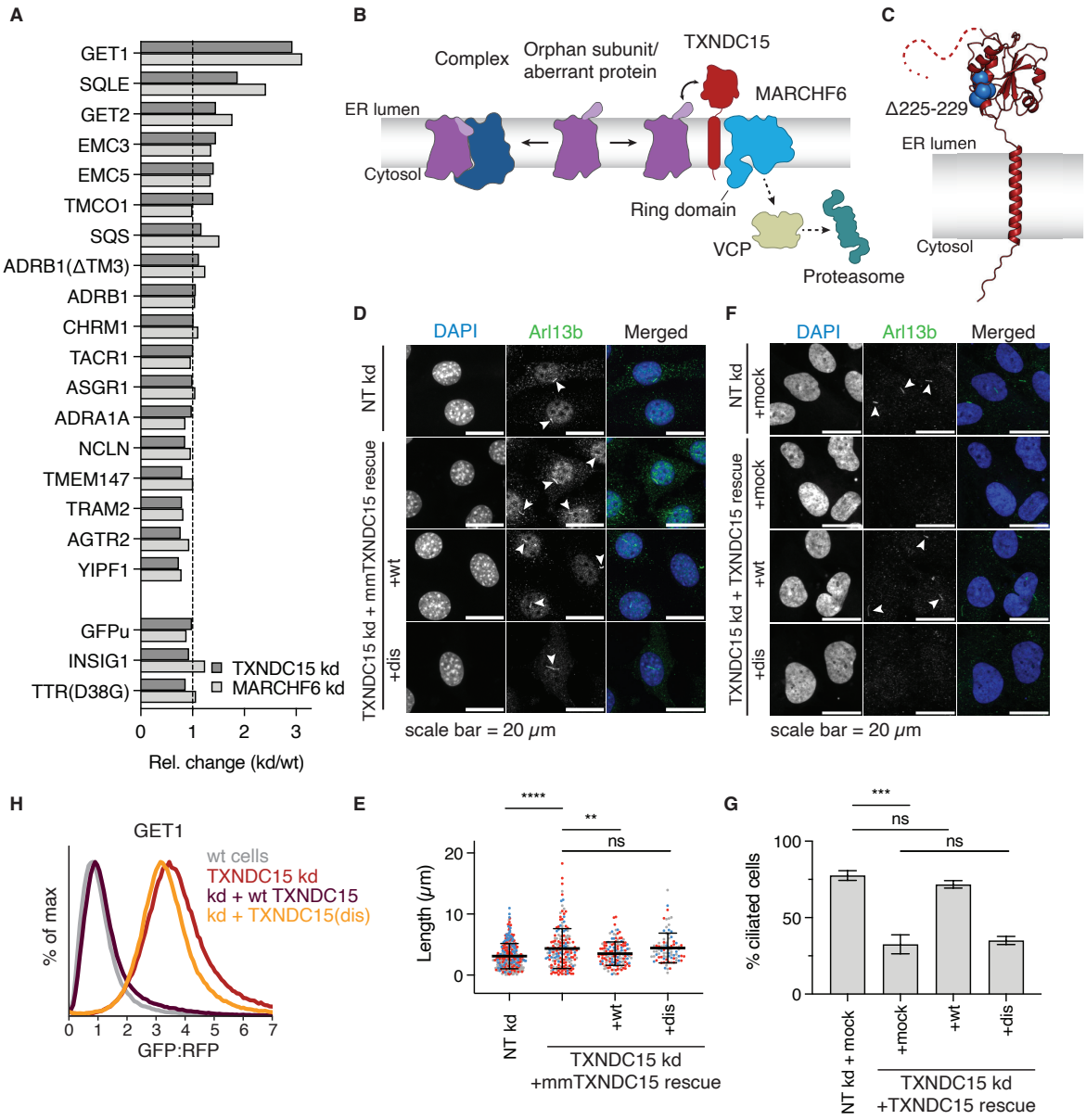

Figure 3.5. Characterization of TXNDC15 clients and its effect on primary cilia.

(A) Summary of the dependence on TXNDC15 and MARCHF6 for the indicated membrane protein substrates determined using the ratiometric fluorescent reporter system described in Fig. 3.1C. The panel includes several representative orphan subunits (GET1, GET2, EMC3, EMC5, TMCO1); tail-anchored proteins (SQS); GPCRs (ADRB1, CHRM1, ADRA1); an unfolded mutant of ADRB1 missing its 3rd TM; and several additional multipass proteins. The effect of loss of TXNDC15 and MARCHF6 on canonical ERAD-L, -C, and -M clients (TTR(D38G), GFPu and INSIG1, respectively) is

depicted below. **(B)** Working model for the function of TXNDC15 as the defining member of a branch of ERAD that surveils the region of the ER lumen adjacent to the bilayer. Substrates are recognized by the lumenal domain of TXNDC15 through exposure of a hydrophobic degron within a particular distance of the lumenal leaflet of the membrane. Recognition of TXNDC15 serves to bring clients in proximity of MARCHF6 for ubiquitination, extraction, and finally proteasome-dependent degradation. **(C)** AlphaFold 3 predicted model of TXNDC15. Residues deleted in the identified patient mutation of TXNDC15 that causes the ciliopathy Meckel-Gruber syndrome (hs: Δ225-229 (Shaheen et al., 2016), mm: Δ209-213) are highlighted as blue spheres. **(D-E)** Confocal microscopy of cilia in mouse NIH 3T3 cells. Cilia length was determined from either wt, TXNDC15 kd, or TXNDC15 kd cells expressing wt murine TXNDC15 or the murine homolog of the human TXNDC15 ciliopathy-mutant (dis, Δ209-213). DAPI (blue) stains nuclear DNA and Arl13b (green) is used as ciliary marker. (E) Cilia lengths of NIH 3T3 cells are reported with the average and standard deviation for wt cells (338 cilia), TXNDC15 kd (182 cilia), murine TXNDC15 rescue (147 cilia) and murine TXNDC15 ciliopathy mutant rescue (85 cilia). Colors represent different replicates (n=3), percentage of ciliated cells is shown in Fig. S3.12C, where we also observed a decrease in cilia in TXNDC15 kd cells. **(F-G)** Confocal microscopy of cilia in RPE1 wt cells and TXNDC15 kd cells with either mock (BFP only), or wt TXNDC15 rescue, or TXNDC15 disease-mutant rescue (dis, Δ225-229). (G) Percent ciliated RPE1 cells, determined by quantifying the number of cilia and number of cells in images collected for each cell sample from confocal microscopy. Mean and standard deviation for 3 replicates are shown (wt: 759 cells; TXNDC15 kd: 742 cells, TXNDC15 wt rescue: 488 cells; TXNDC15(dis) rescue: 520 cells). **(H)** Analysis of the quality control function of the ciliopathy causing mutation of TXNDC15 on GET1 in HEK293T. Rescue experiments in wt and TXNDC15 depleted cells were carried out as described in Fig. 3.3 and analyzed using flow cytometry.

3.5. Materials and Methods

Plasmids

Protein sequences: The sequences used in cell-based and in vitro experiments were derived from UniProtKB/SwissProt. These include: guided entry of tail-anchored proteins factor 1 (GET1; **O00258**), guided entry of tail-anchored proteins factor 2 (GET2; **P49069**), SARS-CoV-2 membrane glycoprotein M (VME1_SARS2; **P0DTC5**), vesicle associated membrane protein 2 (VAMP2; **P63027-1**), ER membrane protein complex subunit 3 (EMC3; **Q9P0I2**), ER membrane protein complex subunit 5 (EMC5/MMGT1; **Q8N4V1**), thioredoxin domain-containing protein 15 (TXNDC15; human: **Q96J42**, mouse: **Q6P6J9**), E3 ubiquitin-protein ligase MARCHF6 (MARCHF6; **O60337**), SARS-CoV-2 ORF3a (AP3A_SARS2; **P0DTC3**), T-cell surface glycoprotein CD4 (CD4; **P01730**), transmembrane and coiled-coil domains 1 (TMCO1; **Q9UM00**), squalene synthase (SQS/FDFT1; **Q6IAX1**), squalene monooxygenase (SQLE; **Q14534**), b-1 adrenergic receptor (ADRB1; **P08588**), muscarinic acetylcholine receptor M1 (CHRM1; **P11229**), tachykinin receptor 1 (TACR1; **P25103**), asialoglycoprotein receptor 1 (ASGR1; **P07306**), alpha-1A adrenergic receptor (ADRA1A; **P35348**), opioid receptor kappa 1 (OPRK1; **P41145**), translocon-associated protein subunit alpha (SSR1/SSRA/TRAPα; **P43307**), nicalin (NCLN; **Q969V3**), transmembrane protein 147 (TMEM147; **Q9BVK8**), translocating chain-associated membrane protein 2 (TRAM2; **Q15035**), type-2 angiotensin II receptor (AGTR2; **P50052**), Yip1 domain family member 1 (YIPF1; **Q9Y548**), insulin-induced gene 1 protein (INSIG1, **O15503**), transthyretin (TTR; **P02766**). The sequences of TTR(D38G, or D18G without signal sequence) and GFPu was designed as previously described (Bence et al., 2001; Christianson et al., 2011).

General plasmid backbones: The 2nd generation lentiviral packaging plasmid pCMV-VSV-G was a gift from Bob Weinberg (Addgene plasmid #8454). The 2nd generation lentiviral packaging plasmid psPAX2 was a gift from Didier Trono (Addgene plasmid #12260). The pHAGE2 lentiviral transfer plasmid was a gift from Magnus A. Hoffmann and Pamela Bjorkman. The dual guide lentiviral vector pJR103 was a gift from Jonathan Weissman (Addgene plasmid #187242). The SFFV-tet3G backbone was used for K562

cell expression during CRISPRi screens (Jost et al., 2017). For transduction of NIH 3T3 cells, we used the plasmid backbone pMCB306 (Addgene plasmid #89360), which was a gift from Michael Bassik. RFP is used in the text and figures and refers to the mCherry variant that was used in this study; GFP refers to the EGFP variant. The ratiometric fluorescence GFP:RFP reporter system was used as previously described (Yanagitani et al., 2017).

GET1 reporter construct in K562 cells in CRISPRi screens: The fluorescent reporter used for genome wide CRISPRi screens contains GET1 cloned downstream of the doxycycline-inducible TET3G promoter with a C-terminal GFP fusion followed by a P2A ribosomal skipping sequence (2A) and RFP normalization marker (Fig. 3.1D, 3.1G, S3.1B, S3.1E, S3.3).

GET1 reporter construct in K562 cells for follow-up experiments: For reporter assay experiments shown in Fig. 3.1E, S3.1A, GET1 reporter construct was the same as GET1 reporter construct used in CRISPRi screens, as described above. For reporter assay experiments shown in Fig. 3.2E, S3.1F, the constitutive reporter construct of GET1 was cloned downstream of an EF1α promoter with a C-terminal GFP fusion followed by a 2A and RFP in the pHAGE2 lentiviral vector.

Reporter constructs in HEK293T cells: For reporter assays in Fig. 3.1F, 3.3A, 3.3B, 3.3D, 3.4G, 3.5A, S3.2C, S3.2D, S3.8, S3.10, S3.11, these following reporter constructs were cloned downstream of a CMV promoter with a C-terminal GFP fusion followed by a 2A and RFP in the pHAGE2 lentiviral vector: GET1, EMC3, EMC5, TMCO1, ADRB1, ADRB1(ΔTM3), ADRA1A, OPRK1, AGTR2, TACR1, CHRM1, INSIG1, TTR(D38G), M, ORF3a, SQLE, TRAM2, TMEM147, NCLN, TRAPα. For experiments in Fig. 3.3E, 3.4D, 3.4F, S3.8F, GET1 point mutants were generated via site-directed mutagenesis. For Fig. S3.10, GET1 mutants with added length to its loop between TM2 and 3 were generated by inserting flexible $(GS)_4$, $(GS)_8$, $(GS)_{12}$ linkers at residue 135 of GET1. For reporter assay experiments shown in Fig. 3.5H, the reporter construct of GET1 was cloned downstream

of an EF1α promoter with a C-terminal GFP fusion followed by a 2A and RFP in the pHAGE2 lentiviral vector.

To retain fluorescent tags in the cytosol, the following reporters were cloned downstream of a CMV promoter with an N-terminal GFP fusion followed by a 2A and RFP in the pHAGE2 lentiviral vector: GET2, YIPF1 (Fig. 3.5A). Tail-anchored membrane protein reporters (VAMP2, SQS, ASGR1) were labelled on the N-terminus using the GFP-2A-RFP cassette in the pHAGE2 lentiviral vector (Fig. 3.5A, S3.4).
For overexpression of GET2 in the presence of GET1-GFP-2A-RFP in Fig. 3.1F, S3.1C, and S3.2C, GET2 was expressed along with a BFP control separated by a 2A site.

To determine the effect of mutations on the C-terminus of GET1 in Fig. 3.4A, chimeric constructs with SARS-CoV-2 M protein (M) were labeled on the N-terminus via fusing an RFP to the TM of VAMP2 and using GFP as translation normalization marker. The full construct used was GFP-2A-RFP-VAMP2-GET1/M/chimera. Additionally, point mutations were introduced to GET1 in the same reporter cassette: V143D, L142N-V143D and L142N-V143D-A144S (Fig. 3.4F).

TXNDC15 rescue constructs in HEK293T cells: The TXNDC15 rescue constructs described in Fig. 3.3A-B, 3.3D, 3.5H, S3.8, S3.10B were cloned downstream of an EF1α promoter with a C-terminal ALFA fusion followed by a 2A and BFP-3xFLAG in the pHAGE2 vector. The ALFA epitope (PSRLEEELRRRLTEP) has been previously described (Götzke et al., 2019). For Fig. 3.3A, S3.8A-B, mutant variants of TXNDC15 were cloned such that the native TXNDC15 TM (residues 322-342) was replaced with the TM of CD4 along with flank residues (GGPGMALIVLGGVAGLLLFIGLGIFFGPGG), the ER lumenal domain of TXNDC15 (residues 33-315) was replaced by the globular, inert, βVHP domain (LSDEDFKAVFGMTRSAFANLPLWKQQNLKKEKGLF), or only the thioredoxin domain residues 1-321 fused to a C-terminal ER-retention sequence KDEL (Munro & Pelham, 1987). For Fig. 3.3B, S3.8C, S3.8E, mutant variants of TXNDC15 with an exogenous flexible linker between the lumenal thioredoxin domain and the TM were cloned such that flexible linkers $(GS)_4$ or $(GS)_8$ were inserted at TXNDC15 residue 322.

Also in Fig. 3.3B, S3.8D-E, mutant variants of TXNDC15 with an exogenous rigid linker between the lumenal thioredoxin domain and the TM were cloned such that residues 322-324 were replaced with rigid linkers $(EAAAK)_3$ or $(EAAAK)_5$. The rigid linker motif $(EAAAK)_n$ (n = 3-5) has been previously characterized to rigidly increase the distance between protein domains (Arai et al., 2001). For Fig. 3.3D, S3.8F, TXNDC15 cysteine mutants (C220S, C210A, C210S C220S) were cloned using site-directed mutagenesis. For Fig. 3.5H, S3.12E, S3.12G, TXNDC15(dis) is a TXNDC15 construct with residues 225-229 deleted using site-directed mutagenesis, as described previously to cause Meckel-Gruber syndrome (Shaheen et al., 2016). In all TXNDC15 rescue variants, its native signal sequence remained intact.

TXNDC15 and MARCHF6 rescue constructs in HEK293T cells (Fig. 3.4G, S3.11B-C): TXNDC15 and MARCHF6 constructs were cloned downstream of a PGK promoter with a C-terminal ALFA fusion followed by a 2A and BFP-3xFLAG in the pHAGE2 vector.

TXNDC15 rescue constructs in murine NIH 3T3 cells (Fig. 3.5D-E, S3.12C, S3.12D, S3.12F): To rescue the TXNDC15 kd phenotype in NIH 3T3 cells using immunofluorescence (Fig. 3.5D-E, S3.12C), the mmTXNDC15 sequence was fused to a C-terminal ALFA-tag followed by a 2A and RFP in the pHAGE2 vector downstream of an EF1α promoter. To rescue the TXNDC15 kd phenotype in NIH 3T3 cells in the rescue reporter assay (fig. S3.12D, S3.12F), the mmTXNDC15 sequence was fused to a C-terminal ALFA-tag followed by a 2A and BFP-3x-FLAG in the pHAGE2 vector downstream of an EF1α promoter. The ciliopathy mutant (mmTXNDC15(dis)) was created by deleting amino acids 209-213 which are equivalent to the human disease mutation as described above.

TXNDC15 rescue constructs in RPE1 cells (Fig. 3.5F-G, S3.12G): TXNDC15 rescue constructs (wt and dis) used for immunofluorescence (Fig. 3.5F-G) were cloned downstream of an EF1α promoter with a C-terminal ALFA fusion followed by a 2A and RFP in the pHAGE2 vector. TXNDC15 rescue constructs used for the TXNDC15 rescue

assay in Fig. S3.12G are the same constructs that were used in HEK293T cells as described above.

Constructs for stable expression in Expi293 cells: For exogenous expression of MARCHF6 to be used for immunoprecipitation-mass spectrometry in Fig. 3.2C, we used the catalytically inactive variant MARCHF6(C9A) – this mutation resides in its RING domain – as MARCHF6 had previously been shown to catalyze its own degradation (Hassink et al., 2005). The sequence of MARCHF6(C9A) was cloned downstream of a CMV promoter with a C-terminal TEV cleavage site (ENLYFQG) followed by a GFP-2A-RFP in the pHAGE2 vector.

For the expression of both TXNDC15 and MARCHF6 to be used for their co-immunoprecipitation shown in Fig. 3.2D, the design of the expression constructs is illustrated in Fig. S3.6C. Construct for "MARCHF6-GFP" is the MARCHF6(C9A)-TEV-GFP-2A-RFP as described above. The TXNDC15 expression construct, "TXNDC15-ALFA", was cloned downstream of a CMV promoter with a C-terminal ALFA followed by a 2A and BFP-3xFLAG in the pHAGE2 vector.

Constructs for in vitro translations: All constructs for expression in rabbit reticulocyte lysate (RRL) were based on the SP64 vector (Promega). For in vitro translations in Fig. 3.2G, 3.2F, 3.4B, the ALFA epitope and a flexible 3xGS linker were appended to the C-terminal of GET1 or the control protein M from SARS-CoV-2. The "M(chimera)" construct was made by replacing residues 78-100 of M with residues 142-174 of GET1 in an M construct with N-terminal HA tag and C-terminal a glycosylation tag NGT and His_6 tag. Alternatively, two point mutations were introduced into GET1 wt to create GET1(N163L, K164L) in Fig. 3.4E.

For in vitro translations in Fig. 3.4C, the canonical NGT acceptor motif for glycosylation was introduced at various locations on GET1 via mutagenesis: "N-term" (HA-NGT-GET1), "C-term" (HA-GET1-NGT-His_6), "N- & C-term" (HA-NGT-GET1-NGT-His_6), "TM3" (HA-GET1(K164G, V165T)-His_6), "Coiled coil" (HA-GET1(GT inserted after N64)-His_6). For in vitro translations to be used for the protease protection assay in Fig. S3.9B, GET1 was cloned with an N-terminal HA tag and a C-terminal His_6 tag.

CRISPRi knockdown guides: Programmed single and dual guides were generated to assay depletion of one or two genes using the CRISPRi system (Replogle et al., 2020). The following sgRNA protospacer sequences were used to clone into pLG1 (single guides) or pJR103 (dual guides):

Gene	**sgRNA protospacer sequence (5'-3')**
Non-targeting control	GACGACTAGTTAGGCGTGTA
TXNDC15	GACGCGGGGCGGTCGTCGAC
mmTXNDC15	GCAGAAAACCGGCACCA
MARCHF6	GCCAGACAAGATGGACACCG
SYVN1/HRD1	GGGAGTGTTGTTAACCGGAG
GET2	GATGTTGGCCGCCGCTGCGA
UBE2G2 (guide 1)	GCGGCTGAGGCGAGGTCGCT
UBE2G2 (guide 2)	GGAAGCAGTCCCCGGTGTCG
UBE2J2 (guide 1)	GGTTCGGTGGGCCCAATCCC
UBE2J2 (guide 2)	GCCGCGAGCGGCCATCTTGG

All plasmids are available upon request.

Antibodies for western blotting and immunofluorescence (catalog number, manufacturer, RRID; dilution)

Primary antibodies: mouse monoclonal anti-α-tubulin (T9026, Sigma-Aldrich, USA, RRID:AB_477593; 1:2,000); rabbit anti-Sec61β (gift from Ramanujan Hegde; 1:50); rabbit polyclonal anti-GFP (A11122, Thermo Scientific, USA, RRID:AB_221569; 1:1,000; rabbit polyclonal anti-Arl13b (17711-1-AP, Proteintech, USA, RRID:AB_2060867; 1:100); mouse monoclonal anti-Arl13b (66739-1-Ig, Proteintech, USA, RRID:AB_2882088; 1:100); mouse monoclonal anti-acetylated-α-tubulin (66200-1-Ig, Proteintech, RRID:AB_2722562; 1:100).

Secondary antibodies used for Western blotting: goat anti-mouse-HRP (172-1011, BioRad, USA, RRID:AB_11125936; 1:2500 or 1:5000); goat anti-rabbit-HRP (170-6515, BioRad, USA, RRID:AB_11125142; 1:2500).

Secondary antibodies used for immunofluorescence: goat anti-mouse IgG Alexa Fluor 488 (115-545-003, Jackson ImmunoResearch Inc, USA, RRID:AB_2338840; 1:500); goat anti-rabbit IgG Alexa Fluor 488 (111-545-003, Jackson ImmunoResearch Inc, USA, RRID:AB_2338046; 1:500); goat anti-mouse IgG Alexa Fluor 647 (115-605-003, Jackson ImmunoResearch Inc, USA, RRID:AB_2338902; 1:500).

Others: anti-FLAG-HRP (A8592, Sigma-Aldrich, USA, RRID:AB_439702; 1:10,000); anti-ALFA nanobody-HRP (this paper; 1:10,000), GFP antibody Alexa Fluor 488 (A-21311, Thermo Scientific, USA, RRID:AB_221477; 1:100).

Conjugation of the ALFA nanobody to HRP for Western blotting

ALFA nanobody was coupled to horseradish peroxidase (HRP)-maleimide via an engineered C-terminal cysteine, as previously described (Pleiner et al., 2017). Briefly, equimolar amounts of the ALFA nanobody and EZ-Link™ maleimide activated HRP (31485, Thermo Scientific, USA) are incubated for 1 hour at room temperature (RT).

Lentiviral production

Lentivirus was produced by co-transfection of packaging plasmid psPAX2 and envelope plasmid VSV.G, along with a transfer plasmid of interest, into HEK293T cells using TransIT-293 transfection reagent (2705, Mirus, USA) as per manufacturer's protocol. Lentivirus was harvested 48 hours after transfection and stored at -80˚C for future use.

Cell culture conditions

Cell Lines	**Source**
K562 KRAB-BFP-dCAS9	(Gilbert et al., 2014)
HEK293T/17	CRL-11268, ATCC, USA
K562-CRISPRi-Tet-ON-(tet-GET1-GFP-2A-RFP)	This study

HEK293T TXNDC15 stable knockdown	This study
HEK293T MARCHF6 stable knockdown	This study
HEK293T SS-GFP-TXNDC15 knockin	This study
HEK293T MARCHF6-TEV-ALFA-2A-BFP knockin	This study
HEK293T SS-GFP-TXNDC15 MARCHF6-TEV-ALFA-2A-BFP knockin	This study
HEK293T MARCHF6-TEV-ALFA-2A-BFP knockin TXNDC15 stable knockdown	This study
Expi293 MARCHF6(C9A)-GFP	This study
Expi293 MARCHF6(C9A)-GFP TXNDC15-ALFA	This study
RPE1 KRAB-BFP-dCAS9	(Jost et al., 2017)
NIH 3T3 Cas9	Gift from Maxence Nachury
NIH 3T3 Cas9 TXNDC15 stable knockdown	This study

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; 11965092, Thermo Scientific, USA) supplemented with 10% fetal bovine serum (FBS; S11150, Bio-Techne, USA), 2 mM glutamine. RPE1 cells were cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12; 11320033, Thermo Scientific, USA) supplemented with 10% FBS, 2 mM glutamine. NIH 3T3 cells were cultured in DMEM supplemented with 10% FBS, 32 mM glutamine, 4.5 g/L D-glucose and 110 mg/L sodium pyruvate (10569010, Thermo Scientific, USA). K562 cells containing KRAB-BFPdCas9 (Gilbert et al., 2014) were cultured in RPMI-1640 with 25 mM HEPES, 2.0 g/L NaHCO3, and 0.3 g/L L-glutamine (SH30255.01, Cytiva, USA) supplemented with 10% Tet System Approved FBS, 2 mM glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin. Expi293 cells were cultured in Expi293 Expression Medium (A1435101, Thermo Scientific, USA). K562 and Expi293 cells were maintained between 0.25×10^6 –1×10^6 cells/mL. Adherent cells were detached from culture dishes (trypsinized) using Trypsin-EDTA (0.25%), phenol red (25200056, Thermo Scientific, USA). HEK293T, RPE1, NIH 3T3, and K562 cells were grown at 37˚C and 5% CO_2. Expi293 cells were grown at 37˚C and 8% CO_2 with shaking at 125 rpm.

For preparations of ER microsomes using HEK293T wt cells and those with endogenously tagged GFP-TXNDC15, these cell lines were adapted for growth in suspension. For adaptation, HEK293T cell lines were trypsinized, pelleted, and resuspended in Expi293 Expression Medium, and maintained between $0.25 \times 10^6 – 1 \times 10^6$ cells/mL.

Endogenously tagged cell line construction and stable knockdown cell lines

Tagging of TXNDC15 at its endogenous locus in HEK293T cells was performed as previously described (Zhang et al., 2017). Briefly, to endogenously tag TXNDC15 with a GFP positioned between its signal sequence and the mature N-terminus, a protospacer within exon two was cloned into pX459 (CCTGGCCCATCGGGTCATG). A vector encoding GFP and 800 bp of homology on either side of the cut site of the TXNDC15 locus was ordered from TWIST Biosciences. The GFP insertion site was designed such that the native signal sequence and signal cleavage site of TXNDC15 would remain intact.

Stable knockdowns of TXNDC15 and MARCHF6 were obtained by transfecting HEK293T cells with pX459 encoding the respective sgRNA using TransIT-293 transfection reagent using the manufacturer's protocol (2705, Mirus, USA). The following sgRNAs were cloned into pX459: TXNDC15 (CGGGCATTTCCAGCTCTTCA), MARCHF6 (CCATGGACTATCTGCTCTCT).

GFP positive single cells of either the stable knockdown or fluorescently tagged knockin were sorted into 96-well plates using a SONY SH800 cell sorter (Sony Biotechnology, USA) 72 hours post transfection. Depletion of TXNDC15 and MARCHF6 were verified with qPCR as described below (Fig. S3.2B, S3.11A). Endogenous tagging of TXNDC15 was verified by Western blotting (Fig. S3.4B).

In NIH 3T3, a stable knockdown cell line was created by transducing NIH 3T3 cells expressing dCas9 with either non-targeting (CGAGGTATTCGGCTCCGCG) or TXNDC15 (GCAGAAAACCGGCACCA) guides in cloned in the pMCB306 backbone. After selection with puromycin (incubation with 3 µg/mL puromycin [A1113803, Thermo Scientific, USA] for three days), knockdown was verified by qPCR (Fig. S3.12A) and polyclonal cell lines were maintained.

Expi293 stable expression cell line generation

Expi293 cells were transduced by adding 2.5 mL of lentivirus of corresponding construct to 10 million cells along with 8 µg/mL final concentration of polybrene in a volume of 30 mL in a 125-mL cell culture flask. After 8-12 hours, the media was exchanged to remove lentiviral particles and prevent cell clumping. Cells were maintained until they were sorted for GFP positive cells using the SONY SH800S cell sorter (Sony Biotechnology, USA).

The MARCHF6(C9A)-GFP cell line was generated by transducing wt Expi293 cells with MARCHF6(C9A)-GFP-2A-RFP lentivirus and sorting for GFP positive cells.
The MARCHF6(C9A)-GFP TXNDC15-ALFA Expi293 cell line was generated by transducing the sorted MARCHF6(C9A)-GFP with TXNDC15-ALFA-2A-BFP lentivirus and sorting for GFP and BFP positive cells.

CRISPRi single sgRNA screen and dual genetic modifier screen

CRISPRi single sgRNA screen: The CRISPRi screen to identify factors for the recognition and degradation of orphan GET1 was performed as previously described (Gilbert et al., 2014; Horlbeck et al., 2016). In short, the TOP5 CRISPRi v2 library was transduced at a multiplicity of infection 0.3 into 330 million K562-CRISPRi-Tet-ON cells stably expressing the GET1-GFP-2A-RFP reporter under an inducible promotor. Transduced cells were enriched using puromycin treatment (addition of 1 µg/mL puromycin for three consecutive days). Cells were maintained at a density of 0.5×10^6 cells/mL throughout the course of the screen. On day 8 the reporter was induced with doxycycline (100-1000 ng/mL) for 36 hours and sorted on a SONY SH800 cell sorter (Sony Biotechnology Inc., USA).
During sorting, cells were first gated based on BFP expression, as a proxy for the expression of an sgRNA, followed by gating for RFP and GFP to select cells expressing the GET1 reporter. From this sgRNA and reporter positive population, cells with perturbed GFP:RFP fluorescence were sorted based on their GFP:RFP ratio. Approximately 25 million cells with GFP:RFP ratios in the highest and lowest 30% were collected, pelleted, and flash-frozen. Genomic DNA was extracted and purified using a Nucleospin Blood XL kit (740950.10, Takara Bio, JPN). Guides were amplified and barcoded by PCR using NEB

Next Ultra ii Q5 MM (M0544L). The DNA library (279 bp for single guide library, 349 bp for dual guide libraries) was purified using SPRISelect beads (B23317, Beckman Coulter, USA), and purified DNA was analyzed on an Agilent 2100 Bioanalyzer prior to sequencing using an Illumina HiSeq2500 using the standard CRISPRi-v2 library sequencing primer (5'-GTGTGTTTTGAGACTATAAGTATCCCTTGGAGAACCACCTTGTTG). The results of the screen were analyzed using published protocols (https://github.com/mhorlbeck/ScreenProcessing) (Horlbeck et al., 2016). To ensure sufficient coverage of the guide libraries, only guides with more than 50 counts were assessed and phenotype scores were computed based on the 3 sgRNAs displaying the strongest phenotypes. The Mann-Whitney p-value was calculated using the 5 sgRNAs targeting the same gene compared to the negative controls.

Cloning the TXNDC15 dual-guide CRISPRi library: Cloning of the genome-wide CRISPRi dual library with a guide targeting TXNDC15 in the anchor position (see Fig 3.1G) was performed as previously described (Guna, Page, et al., 2023). Briefly, the TXNDC15 guide (5'-GACGCGGGGCGGTCGTCGAC-3') was cloned into a hU6-CR3 cassette (pJR152, Addgene #196280). The anchor guide with custom overhangs (see table below) was ligated into the pJR152 vector digested with BstXI/BlpI. This vector containing the hU6-TXNDC15 sgRNA-CR3 element was digested with BamHI and NotI to yield a 400 bp fragment that was gel purified. The CRISPRi-v2 library (Addgene Pooled Libraries #83969) was digested with BamHI and NotI in the presence of shrimp alkaline phosphatase (rSAP) followed by gel purification to yield the 8,800 bp vector. This vector was ligated with the hU6-TXNDC15 sgRNA-CR3 element using T4 (NEB #M0202M) at an insert to vector ratio 1:2 for 16 hours at 16˚C. This construction yielded the mU6-CR1-hU6-CR3 vector, which was purified with SPRISelect beads (Beckman Coulter B23317) and electroporated into MegaX cells at scale (ThermoFisher #C640003). The cells were incubated in 200 mL LB supplemented with 100 µg/mL carbenicillin overnight. The resulting TXNDC15 dual plasmid library was amplified and barcoded for deep sequencing (NGS) by PCR using NEB Next Ultra ii Q5 MM (M0544L) and a reverse primer (CAAGCAGAAGACGGCATACGAGATggaatcatgggaaataggccctc) and index primer. The PCR libraries were purified with SPRISelect beads to yield dual the TXNDC15 dual

DNA library (349 bp), which was sequenced using an Illumina HiSeq2500 with the sequencing primer listed previously. The non-targeting (NT) dual guide genome-wide CRISPRi library was described previously (cite paper) (Addgene Pooled Library #197348). The TXNDC15 dual guide library read counts were compared to the NT dual guide library, which showed consistent guide coverage and distribution between the libraries.

Oligo ID	**Sequence**	**Source**
TXNDC15 guide oligo forward	atgGACGCGGGGCGGTCGTCGACgtttcagagc	This study
TXNDC15 guide oligo reverse	ttagctctgaaacGTCGACGACCGCCCCGCGTCcatgttt	This study

Dual guide genetic modifier screen: The genetic modifier screen was conducted in a similar manner, using a non-targeting dual library or TXNDC15 dual library, as previously described (Guna, Page, et al., 2023). The cells were sorted on a FACSAria Fusion Cell Sorter (BD Biosciences). Following isolation of cells displaying perturbed GET1-GFP:RFP ratios, amplification of the guide region was performed using a unique forward index primer in combination with a reverse primer complementary to the hU6 region upstream of the fixed guide in the dual guide vector (5'-CAAGCAGAAGACGGCATACGAGATGGAATCATGGGAAATAGGCCCTC). Sequencing and analysis were performed using the same strategy as for the single sgRNA screen.

<u>Ratiometric fluorescent reporter assays to test individual depletion of genes and TXNDC15 rescue</u>

Throughout the manuscript we rely on depletion of individual genes to assess the effect of loss of specific factors on a fluorescent reporter. These reporters serve two purposes: First, because the reporters express a substrate (e.g. GET1-GFP) and a normalization marker (e.g.

RFP) from a single open reading frame separated by a viral 2A ribosomal skipping site, changes in GFP:RFP ratio necessarily reflect changes that occur post-translationally. Second, because the reporter is acutely expressed, changes to the GFP:RFP ratio reflect effects to a newly synthesized protein. In both cases, these features are particularly useful for studying the function of quality control factors that must act post-translationally on nascent proteins.

K562 cells: To perform reporter assays in K562 cells, we transduced lentiviral vectors expressing programmed guides into K562 dCas9BFP-KRAB cells. Briefly, 250,000 cells were mixed with 200 µL of lentivirus and RPMI medium in the presence of 8 µg/mL polybrene (TR-1003, Sigma-Aldrich, USA) in a total volume of 1 mL. K562 cells in 12-well plates were spinfected at 1,000 xg for 1.5 hours at 30˚C, resuspended, and cultured in 12-well or 6-well plates. Approximately 48 hours after spinfection, cells containing guides were selected by incubation with 3 µg/mL puromycin for 3 days (treated once). Cells recovered from puromycin treatment for 1.5 days and then were transduced with 50-100 µL of reporter lentivirus. 48 hours after reporter transfection, the cells were analyzed using flow cytometry, as described below.

HEK293T cells: For reporter assays in HEK293T cells, 0.3 x 10^6 of wt or TXNDC15 stable knockdown or MARCHF6 stable knockdown cells were seeded into each 6-well plate. 24 hours later, cells were transduced with 300 µL lentivirus of TXNDC15 rescue construct and 8 µg/mL final concentration of polybrene. 24 hours later, the media was exchanged to remove excess polybrene. 48 hours after reporter transduction, the cells were harvested, washed, pelleted, and resuspended in 500 µL 1x Dulbecco's Phosphate Buffered Saline (1x PBS, Thermo Scientific, USA), and analyzed by flow cytometry.

For rescue assays in HEK293T cells, 0.3 x 10^6 of WT or TXNDC15 depleted or MARCHF6 depleted cells were seeded into each 6-well plate. 24 hours later, cells were transduced with 300 µL lentivirus of each TXNDC15 rescue construct and 8 µg/mL final concentration of polybrene. 24 hours later, the media was exchanged to remove excess polybrene, and cells were transduced with 150 µL reporter lentivirus the in presence of 8

μg/mL final concentration of polybrene. 24 hours later, cells were split 1:2 (half of the sample is used for flow cytometry and the other half is used for Western Blotting). 24 hours later (72 hours after rescue construct transduction and 48 hours after reporter transduction), cells were harvested, pelleted, and resuspended in 500 μL 1x PBS and analyzed by flow cytometry or washed, pelleted, and frozen for analysis via Western blot.

RPE1 cells: For CRISPRi knockdown reporter assay experiment in RPE1, cells were transduced with sgRNA/dual guide lentiviral vectors. 6 days after guide transduction and selection for cells with guides by puromycin treatment (analogous to K562 cell treatment), cells were transduced with reporter lentiviral vectors. After 48 hours, cells were harvested, pelleted, and resuspended in 500 μL 1x PBS and analyzed by flow cytometry. For rescue assays in RPE1, the experiment was performed similarly, except for the addition of rescue lentiviral vectors to the cells 72 hours before analysis as described below. Knockdown of TXNDC15 in RPE1 cells was verified using qPCR (Fig. S3.12B), as described below.

NIH 3T3 cells: Rescue assays in NIH 3T3 cells were performed using polyclonal TXNDC15 stable knockdown cell lines. 48 hours after transduction with lentiviral reporter constructs, cells were harvested and analyzed by flow cytometry as described below.

Flow cytometry to analyze fluorescence reporter assays

HEK293T, RPE1, and NIH 3T3 cells were treated with trypsin, pelleted, and resuspended in 1x PBS for flow cytometry analysis. K562 cells were analyzed directly from 6-well cultures. Cells were analyzed using an Attune NxT Flow Cytometer (Thermo Scientific, USA).

Flow cytometry data was analyzed using FlowJo v10.8 Software (BD Life Sciences, USA) or by Python using the FlowCytometryTools package (Yurtsev & Friedman, 2015). For reporter assays, cells were gated based on RFP expression for cells expressing the reporter constructs. For TXNDC15 rescue assays, cells were first gated based on BFP expression and then RFP expression for cells expressing both the rescue constructs and the reporter. The relative change (kd/wt) of each reporter was calculated by taking the ratio of the

median GFP:RFP intensities between that of the respective knockdown cells and the wildtype/non-targeting control cells. Relative change (kd/wt) is either displayed in bar plots or histograms. Normalized histograms of these assays are generated by setting the GFP:RFP ratio of wildtype or non-targeting control to 1. Histograms, charts, and plots were generated using Python and GraphPad Prism v.10.4.1.

Calculating % rescue of TXNDC15 rescue constructs (Fig. 3.3A-C, S3.10B)

To calculate how well a particular mutant of TXNDC15 can rescue GET1 stability in HEK293T TXNDC15 stable kd compared to wt cells, the following formula (1) was used:

$$\%\, rescue = \frac{Med_{TXNDC15\ kd} - Med_{TXNDC15\ rescue}}{Med_{TXNDC15\ kd} - Med_{wt\ cells}} * 100\%$$

(1)

"Med" represents the median GFP:RFP value of the GET1 reporter, determined from the ratio of GFP to RFP fluorescence measurements obtained from flow cytometry.

K562 treatment with quality control factor inhibitors

To test the effect of inhibition of the VCP AAA ATPase or E1 ubiquitin-activating enzymes on TXNDC15 dependent degradation, K562 wt and TXNDC15 kd cells were prepared as described above. 48 hours after transduction with the GET1-GFP-2A-RFP reporter construct, cells were treated with either DMSO control, or 250 nM E1 inhibitor MLN7243 (S8341, Selleckchem, USA) in DMSO, or 1.25 µM VCP inhibitor CB-5083 (S8101, Selleckchem, USA) in DMSO for 6 hours. Afterwards, cells were analyzed by flow cytometry as described above.

Mass spectrometry sample preparation of TXNDC15 and MARCHF6 immunoprecipitations (IP-MS)

As displayed in Fig. 3.2B-C, native immunoprecipitation of endogenously expressed GFP-TXNDC15 from HEK293T cells or exogenously expressed MARCHF6-GFP from Expi293 cells were performed as described previously (Stevens et al., 2023). The control sample for background binding of TXNDC15 immunoprecipitation in Fig. 3.2B was wt

HEK293T cells that were not endogenously tagged with GFP (Fig. S3.6A). The control sample for background binding of MARCHF6 immunoprecipitation in Fig. 3.2C was the same Expi293 cells overexpressing MARCHF6(C9A)-GFP but the immunoprecipitation was carried out as described below without addition of GFP-nanobody (Fig. S3.6B). Briefly, cells were incubated in Solubilization Buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 2 mM $MgOAc_2$, 1% [w/v] glycol-diosgenin [GDN, GDN101, Anatrace, USA], 1x Roche cOmplete EDTA-free protease inhibitor cocktail [COEDTAF-RO, Sigma-Aldrich, USA], 1 mM DTT) for 1 hour on ice. Solubilized lysate was clarified by centrifugation at 34,540 xg in an SS-34 rotor in a Sorvall RC6+ Superspeed centrifuge for 45 min. at 4˚C. At the same time, biotinylated His14-Avi-$SUMO^{Eu1}$-tagged anti-GFP nanobody was incubated with Pierce™ Streptavidin Magnetic Beads (88817, Thermo Scientific, USA) head-over-tail in Wash Buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 2 mM $MgOAc_2$, 1 mM DTT, 0.0053% GDN) for 30 min. at 4˚C. Beads were washed and incubated with 50 mM HEPES pH 7.5 containing 100 µM biotin on ice for 5 min. to occupy unbound streptavidin sites on the magnetic beads. Then, beads were washed in Solubilization Buffer twice before being incubated with clarified lysate head-over-tail for 1-1.5 hours at 4˚C. After incubation, beads were washed 3 times in Wash Buffer. Proteins were eluted by addition of 0.5 µM $SUMO^{Eu1}$ protease (Addgene #149333) and incubated for 30-45 min. at 4˚C with rotating. Elution protein samples were precipitated by adding 1:10 volume of 100% trichloroacetic acid (TCA), followed by a 10 min. incubation on ice before being pelleted at maximum speed in a benchtop centrifuge at 4˚C. Protein pellets were washed with ice-cold acetone twice before being air dried at RT. TCA-precipitated protein pellets were dissolved in 8 M urea prepared in 50 mM HEPES, pH 8.0. Samples were reduced by incubating with 4 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (20490, Thermo Scientific, USA) for 20 min. at 37˚C, alkylated by incubating with 12 mM 2-chloro-acetamide (CAA) (ICN15495580, MP Biomedicals) for 15 min. at 37˚C. Samples were digested with 2 ng/µL Lysyl Endoproteinase (Lys-C) (125-05061, Wako Chemicals) for 4 hours at 37˚C. Then, samples were diluted 4-fold so that final concentration of urea reached 2 M, and digested overnight with 0.6 ng/µL Trypsin (90057, Thermo Scientific, USA). Samples were desalted using the Pierce™ C18 Spin Columns (89870, Thermo

Scientific, USA) as per manufacturer's instructions. After being eluted off the desalting columns, the samples were lyophilized before mass spectrometry analysis.

Mass spectrometry analysis for TXNDC15 and MARCHF6 interaction partners (Fig. 3.2B-C, S3.6A-B)

LC-MS/MS analysis was performed with an EASY-nLC 1200 (Thermo Scientific, USA) coupled to a Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific, USA). Peptides were separated on an Aurora UHPLC Column (25 cm × 75 µm, 1.7 µm C18, AUR3-25075C18, Ion Opticks, AUS) with a flow rate of 0.35 µL/min. for a total duration of 75 min. and ionized at 1.6 kV in the positive ion mode. The gradient was composed of 6% solvent B (3.5 min), 6-25% B (42 min.), 25-40% B (14.5 min.), and 40–98% B (2 min.) and 98% B (13 min.); solvent A: 2% acetonitrile (ACN; A9554, Fisher Scientific, USA) and 0.2% formic acid (FA, A11750, Fisher Scientific) in LC-MS grade water (W6212, Fisher Scientific, USA); solvent B: 80% ACN and 0.2% FA. MS1 scans were acquired at the resolution of 60,000 from 375 to 1600 m/z, AGC target 3e6, and maximum injection time 15 ms. The 12 most abundant ions in MS1 scans are selected for fragmentation via higher-energy collisional dissociation (HCD) with a normalized collision energy (NCE) of 28. MS2 scans were acquired at a resolution of 30,000, AGC target 1e5 and a maximum injection time 45 ms. Dynamic exclusion was set to 45 s and ions with charge +1, +6, +8 and >+8 were excluded. The temperature of ion transfer tube was 275°C and the S-lens RF level was set to 55. RAW files were searched with Proteome Discoverer SEQUEST (version 2.5, Thermo Scientific, USA) against in silico tryptic digested the UniProt Human proteome Swiss-Prot database (UP000005640). The maximum missed cleavages were set to 2. Dynamic modifications were set to oxidation on methionine (M, +15.995 Da), deamidation on asparagine and glutamine (N and Q, +0.984 Da) and protein N-terminal acetylation (+42.011 Da). Carbamidomethylation on cysteine residues (C, +57.021 Da) was set as a fixed modification. The maximum parental mass error was set to 10 ppm, and the MS2 mass tolerance was set to 0.03 Da. Intensity-based quantification (iBAQ) was performed using the IMP-apQuant PD node (Doblmann et al., 2019; Schwanhäusser et al., 2011). The maximum false peptide discovery rate was specified as 0.01 using the Percolator Node validated by q-value. For each IP-MS sample,

proteins were ranked based on the log10(abundance) and plotted as protein rank versus log10(abundance) accordingly.

Mass spectrometry sample preparation of GET1 interactions in HEK293T cells

As displayed in Fig. 3.2H, interaction partners of GET1 in HEK293T cells were identified using mass spectrometry. The GET1 construct used in this experiment is the same GET1 construct used in the CRISPRi screen, with a C-terminal fusion to GFP-2A-RFP. Control sample was a construct of GFP-2A-RFP (Fig. S3.7D). Inducible K562 cell lines stably expressing protein constructs were harvested and washed in 1x PBS prior to immunoprecipitation. Cell pellets were resuspended in solubilization buffer (50 mM HEPES pH 7.5, 2 mM $MgOAc_2$, 200 mM KOAc, 1% [w/v] digitonin, 1x Roche cOmplete EDTA-free protease inhibitor cocktail [COEDTAF-RO, Sigma-Aldrich, USA], 1 mM DTT) and incubated on ice for 20 min. The lysate was spun at full speed for 20 min. in a table-top centrifuge at 4˚C to remove cell debris. Pierce magnetic Streptavidin beads (88817, Thermo Scientific, USA) were equilibrated in wash buffer (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 100 mM KOAc, 1x Roche cOmplete EDTA-free protease inhibitor cocktail, 0.25% [w/v] digitonin) followed by incubation with biotinylated anti-GFP nanobody in wash buffer for 20 min. at 4˚C. To block free biotin binding sites, magnetic beads were incubated with solubilization buffer containing 50 µM biotin for 5 min. on ice. Beads were washed twice with solubilization buffer before being incubated with clarified cell lysate for 2 hours at 4˚C. Unbound solubilized samples were removed from the beads and then the magnetic beads were washed three times with wash buffer and incubated in wash buffer containing 250 nM $SUMO^{Eu1}$ protease for 30 min. at 4˚C to elute proteins. The elution samples were TCA-precipitated as described above and subjected to mass spectrometry analysis.

Mass spectrometry sample preparation of nascent GET1 interactions in the ER in vitro

As displayed in Fig. 3.2G, putative interaction partners of nascent GET1 in the ER were identified using mass spectrometry. These experiments were performed using GET1 synthesized and purified from cell free RRL in vitro translations in the presence of human ER microsomes (purified from Expi293 cells as described below). To do this, we generated

a construct in which GET1 was fused to a C-terminal ALFA tag. In vitro translations were performed as described below in "Mammalian in vitro translation", except with unlabeled methionine. For control samples, no mRNA was added to the translation reaction (Fig. S3.7C). 750 µL in vitro translation reactions were pelleted through a sucrose cushion (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 100 mM KOAc, 20% sucrose [w/v]) for 20 min. at 55,000 RPM at 4˚C in a TLA-55 rotor (Beckman-Coulter). The microsomal fraction was resuspended in half the translation volume of physiological salt buffer (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 200 mM KOAc) followed by an equal volume of 2x solubilization buffer (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 300 mM KOAc, 2% [w/v] GDN). Samples were solubilized for 20 min. at 4˚C and then pelleted at 14,000 xg for 10 min. at 4°C to remove insoluble debris. Solubilized samples were added to magnetic beads containing anti-ALFA nanobody as described below. Pierce magnetic Streptavidin beads (88817, Thermo Scientific, USA) were equilibrated in wash buffer (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 100 mM KOAc, 0.25% [w/v] GDN) followed by incubation with biotinylated anti-ALFA nanobody in wash buffer for 20 min. at 4˚C. To block free biotin binding sites, magnetic beads were incubated with wash buffer containing 10 µM $dPEG_{24}$-biotin acid (QBD10773, Sigma-Aldrich, USA) for 5-10 min. at 4˚C. Beads were washed 2 times with wash buffer before mixing with solubilized in vitro translation samples. Samples were incubated with head-over-tail mixing at 4˚C for 2 hours. Unbound solubilized samples were removed from the beads and then the magnetic beads were washed three times with wash buffer and resuspended in wash buffer containing 250 nM $SUMO^{Eu1}$ protease in a volume amounting to one tenth or one twentieth of the original in vitro reaction volume. Protease elution proceeded for 30 min. on ice, and the eluate was analyzed using mass spectrometry.

Mass spectrometry analysis for GET1 interaction partners (Fig. 3.2G-H, S3.7C-D)

Acetone precipitates were resuspended in 100 mM Tris-HCl pH 8.5, 8 M urea and reduced and alkylated using 5 mM Tris (2-carboxyethyl) phosphine and 10 mM iodoacetamide, respectively. The samples were proteolytically digested with Lys-C and trypsin overnight at 37˚C. The digestion was quenched by addition of formic acid to a final concentration of 5% and then desalted using C18 tips (87782, Thermo Scientific, USA). Dried peptides were resuspended in 5% formic acid and analyzed by LC-MS/MS. Briefly, peptides were

separated by reversed phase chromatography using 75 µm inner diameter fritted fused silica capillary column packed in-house to a length of 25 cm with Luna C18 3µm reverse phase particles (Jami-Alahmadi et al., 2021). The increasing gradient of acetonitrile was delivered by a Dionex Ultimate 3000 (Thermo Scientific) at a flow rate of 300 nL/min. MS/MS spectra were collected using data dependent acquisition on Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific, USA) with an MS1 resolution (r) of 120,000 followed by sequential MS2 scans at a resolution (r) of 15,000. The data generated by LC-MS/MS were analyzed on MaxQuant bioinformatic pipeline (Cox & Mann, 2008). The Andromeda component of MaxQuant was employed as the peptide search engine and the data were searched against Human reference proteome (Uniprot Reference UP0000056409606). Briefly, a maximum of two missed cleavages was allowed. The maximum false discovery rate for peptide and protein was specified as 0.01. Label-free quantification (LFQ) was enabled with LFQ minimum ratio count of 1. The parent and peptide ion search tolerances were set as 20 and 4.5 ppm respectively. The MaxQuant output files were subsequently processed for statistical analysis of differentially enriched proteins using Provision (Gallant et al., 2020).

Expression and purification of a recombinant TXNDC15-MARCHF6 complex

For Fig. 3.2D, Expi293 cells stably expressing MARCHF6(C9A)-GFP and TXNDC15-ALFA were subjected to first an anti-ALFA immunoprecipitation and then anti-GFP immunoprecipitation. Briefly, 500 mL of cells were harvested at 4500 rpm and washed with ice-cold 1x PBS. Cell pellets were resuspended in solubilization buffer (50 mM HEPES pH 7.5, 2 mM $MgOAc_2$, 200 mM KOAc, 1% [w/v] GDN, 1x Roche cOmplete EDTA-free protease inhibitor cocktail [COEDTAF-RO, Sigma-Aldrich, USA], 1 mM DTT) and incubated at 4˚C for 1 hour. Solubilized lysate was clarified by centrifugation at 34,540 xg in an SS-34 rotor in a Sorvall RC6+ Superspeed centrifuge for 45 min. at 4˚C. Simultaneously, biotinylated His14-Avi-SUMOStar-tagged anti-ALFA nanobody was incubated with Pierce™ Streptavidin Magnetic Beads (88817, Thermo Scientific, USA) head-over-tail in Wash Buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 2 mM $MgOAc_2$, 1 mM DTT, 0.0053% GDN) for 30 min. at 4˚C. After this immobilization, beads were washed and incubated with 50 mM HEPES pH 7.5 containing 100 µM biotin on ice for 5

min. to occupy unbound biotin binding sites on the magnetic beads. Then, beads were washed in Solubilization Buffer twice before being incubated with clarified lysate head-over-tail for 1.5 hours at 4˚C. After incubation, beads were washed 3 times in Wash Buffer. Proteins were eluted by addition of 0.3 µM SUMOStar protease and incubated for 45 min. at 4˚C with rotating. Eluted proteins were then subjected to anti-GFP immunoprecipitation as described above. Finally, eluted proteins were analyzed using SDS-PAGE.

Preparation of human ER microsomes

Expi293 or HEK293T cells (wt or endogenously tagged GFP-TXNDC15) adapted to suspension were harvested and washed in 1x PBS. Cells were resuspended in 4 times the pellet volume of sucrose buffer (10 mM HEPES, pH 7.5, 250 mM sucrose, 2 mM, 1x Roche cOmplete EDTA-free protease inhibitor cocktail [COEDTAF-RO, Sigma-Aldrich, USA]) and lysed by douncing at 4˚C. Cells were diluted to twice the lysis volume in sucrose buffer and pelleted at 3,214 xg for 35 min. at 4˚C. Supernatant from the cell lysate was transferred to a new tube and pelleted again at 3,214 xg for 35 min. at 4˚C. Samples were pelleted in an ultracentrifuge in an MLA80 rotor (Beckman-Coulter, USA) at 75,000 xg for 1 hour at 4˚C to isolate the microsomal fraction. The microsomal pellet was resuspended in microsome buffer (10 mM HEPES, pH 7.5, 250 mM sucrose, 1 mM $MgOAc_2$, 0.5 mM DTT) to an A280 of 75. To remove contaminating RNAs from human rough microsomes (hRMs), hRMs were mixed with $CaCl_2$ (1 mM) and micrococcal nuclease (0.125 U/µL) on ice before incubating for 6 min. at 25˚C. Nucleased hRMs were transferred to ice and immediately quenched with EGTA (2 mM). Nucleased hRMs were flash frozen and stored at -80˚C prior to use in in vitro translations. Only aliquots of nucleased hRMs with fewer than two freeze-thaw cycles were used in experiments.

Mammalian in vitro translations

Translation reactions were prepared using nucleased RRL supplemented with either canine derived pancreatic microsomes (cRMs) or hRMs, as described above and previously (Sharma et al., 2010; Walter & Blobel, 1983). DNA templates for in vitro transcription were made by PCR using primers within the SP6 promoter (5’ end) and after the stop codon (3’ end). Transcription reactions were carried out by combining 4.8 µl T1 mix (Sharma et

al., 2010), 0.1 µl RNAsin (N251, Promega, USA), 0.1 µl SP6 polymerase (M0207L, New England Biolabs, USA) and 50 ng PCR product. Transcription reactions were incubated at 37˚C for 2 hours, and then used directly in a translation reaction, which was incubated for 45 min. at 32˚C. Radioactive ^{35}S-methionine (NEG709A005MC, Perkin Elmer, USA) was used to label nascently translated proteins, unless otherwise indicated. Reactions were analyzed on SDS-PAGE followed by autoradiography. In vitro translation reactions for mass spectrometry were treated with 1 mM puromycin for 10 min at 32 C immediately following translation in order to fully release nascent proteins from the ribosome.

Glycosylation assay: For glycosylation experiment in Fig. 3.4C, translations were performed as described above using cRMs, with addition of a sample containing a glycosylation-inhibiting acceptor peptide (AP) at 0.75 mM final concentration, to confirm the shift in molecular weight observed in the samples containing cRMs was due to insertion into the ER lumen.

Protease protection assay: For the protease protection assay in Fig. S3.9B, translations were performed using cRMs. After incubation for 25 min., 0.5 mg/mL proteinase K (RPROTK-RO, Sigma-Aldrich, USA) was added to the translation reactions and incubated for 1 hour on ice. To quench the digestion, 5 mM PMSF in DMSO (made fresh) was added to each reaction and the samples were immediately transferred to boiling SDS buffer (1% SDS in 0.1 M Tris pH 8.0). Samples were left on the heat block to boil for 3 min. before being removed from heat. The protected fragments were immunoprecipitated in IP buffer (50 mM HEPES 7.5, 1% Triton X-100, with either 250 mM NaCl for the HA IP or 500 mM NaCl and 10 mM imidazole for the His IP) using anti-HA-agarose (A2095, Sigma-Aldrich, USA) or Ni-NTA-agarose resin (30210, Qiagen, DEU). Samples were eluted from resin by boiling in SDS-PAGE sample buffer, containing 50 mM EDTA in His-IPs.

Analysis of TXNDC15 interaction with orphan GET1

For the experiments shown in Fig. 3.2F, S3.7A-B, translation reactions were performed as described above in presence of hRMs prepared from wt or endogenously tagged GFP-TXNDC15 MARCHF6-ALFA HEK293T cells. Translation reactions were pelleted

through a sucrose cushion (50 mM HEPES, pH 7.5, 2 mM $MgOAc_2$, 100mM KOAc, 20% sucrose [w/v]) for 20 min. at 55,000 RPM at 4˚C in a TLA-55 rotor (Beckman-Coulter). Membrane fractions were resuspended in solubilized buffer (50 mM HEPES pH 7.5, 2 mM $MgOAc_2$, 200 mM KOAc, 1% [w/v] GDN) and incubated on ice for 20 min. Solubilized membrane was pelleted at 14,000 xg for 10 min. in a table-top centrifuge for 10 min. at 4˚C and the clarified solubilized membrane was transferred to a fresh tube. Anti-GFP nanobody was immobilized onto magnetic streptavidin beads as described above. Beads were incubated with solubilized membrane for 2 hours at 4˚C with rotating. Unbound fraction was removed by pipetting and beads were washed 3 times before being incubated with wash buffer containing 0.25 µM $SUMO^{Eu1}$ protease for 30 min. at 4˚C. Elution samples were analyzed on SDS-PAGE followed by autoradiography.

Immunofluorescence (IF) confocal microscopy

Preparation and coating of microscopy cover glass: Circular cover glass (72222-01, Electron Microscopy Sciences, USA) was sterilized, placed in the bottom of each 12-well plate, and coated by incubating with 1 mL of 0.01 mg/mL poly-D-lysine hydrobromide (P6407, Sigma-Aldrich, USA) for 1 hour at RT. Coated cover glass was washed 3 times with nanopure water and dried overnight at RT.

Seeding conditions for endogenously tagged GFP-TXNDC15 HEK293T cells: For Fig. 3.2A, 275,000 of endogenously tagged GFP-TXNDC15 HEK293T cells were seeded into each 12-well with coated cover glass. Cells were grown on cover glass for 48 hours before preparation of imaging slides.

Seeding conditions for NIH 3T3 cells: For the imaging rescue assay in Fig. 3.5D-E, S3.12C, 100,000 NIH 3T3 cells (wt or TXNDC15 stable kd) were seeded in 6-well plates and treated with 8 µg/mL final concentration of polybrene or transduced with the respective lentivirus and polybrene. After 48 hours, 65,000 cells were seeded into each 12-well with coated cover glass. To induce cilia formation, NIH 3T3 cells were serum starved by incubation with DMEM supplemented with 0.5% FBS, 32 mM glutamine, 4.5 g/L D-glucose and 110 mg/L sodium pyruvate for 48 hours before preparation of imaging slides.

Seeding conditions for RPE1 cells: For the imaging rescue assay in Fig. 3.5F-G, RPE1 cells were prepared as follows. Cells were transduced with sgRNA guide lentiviral vectors. On Day 2 (24 hours after lentiviral transduction), cells were subjected to puromycin treatment (8 μg/mL final concentration) to select for cells with guides for 48 hours. On Day 7, 100,000 cells were seeded into each 6-well plate. On Day 8, lentivirus of TXNDC15 rescue constructs were added to corresponding 6-well plates. On Day 10, cells were seeded into 12-well with coated cover glass as described below. 90,000 cells (either wildtype or from knockdown-rescue experiment) were seeded into each 12-well with coated cover glass. To induce cilia formation, adherent cells were serum starved by incubation in DMEM/F12 media supplemented with 2 mM glutamine and 0% FBS for 48 hours before preparation of imaging slides.

Imaging slides preparation: Cells were fixed in 4% paraformaldehyde (PFA; 15714, Electron Microscopy Sciences) in 1x PBS for 15 min. and permeabilized cells in 0.1% Triton X-100 in PBS for 5 min. Samples were blocked in IF Blocking Buffer (10% FBS, 5% Normal Donkey Serum in 1x PBS) for 30 min.

For subsequent incubation steps, cover glasses are inverted onto droplets on parafilm in a humidified environment to avoid evaporation. Samples were incubated in 60-μl droplets of primary antibodies diluted in IF Blocking Buffer for 1-2 hours and washed 3 times with 1x PBS. Then, samples were incubated in 60-μl droplets of secondary antibodies diluted in IF Blocking Buffer for 1 hour and washed 3 times with 1x PBS. Specifically, for imaging of the endogenously tagged HEK293T cells in Fig. 3.2A, the samples were first stained with rabbit anti-Sec61β (1:50), washed, stained with goat anti-rabbit IgG Alexa Fluor 488 (1:500), washed, stained with anti-GFP antibody Alexa Fluor 488 (1:100), and washed. For imaging of the RPE1 cells in Fig. S3.5, samples were first stained with mouse anti-Arl13b (1:100), washed, stained with goat anti-mouse IgG Alexa Fluor 647 (1:500), and washed. For imaging of the NIH 3T3 cells in Fig. 3.5D-E, S3.12C, samples were first stained with rabbit anti-Arl13b (1:100) and mouse anti-acetylated tubulin (1:100), washed, stained with goat anti-rabbit IgG Alexa Fluor 488 (1:500) and goat anti-mouse IgG Alexa Fluor 647

(1:500), and washed. For imaging of the RPE1 cells in Fig. 3.5F-G, samples were first stained with rabbit anti-Arl13b (1:100), washed, stained with goat anti-rabbit IgG Alexa Fluor 488 (1:500), and washed.

Samples were stained with 1 µg/mL DAPI (MBD0015, Sigma-Aldrich, USA) in 1x PBS for 5 min. and washed 2 times with 1x PBS, followed by 2 washes with nanopure water. To mount samples onto microscope slides, each stained cover glass was inverted onto 7 µl of SlowFade™ Gold Antifade Mountant (S36936, Invitrogen, USA) was added on a glass microscope slide. Cover glasses were sealed and secured in place on microscope slide with nail polish. Prepared microscope slides were stored at RT overnight to dry and stored at 4˚C until imaging.

Imaging conditions: Samples were imaged using the Leica STELLARIS 8 FALCON confocal fluorescence microscope (Leica Microsystems, USA) using the 63x oil-immersion objective.

Imaging data analysis: All imaging analyses were performed in FIJI (Schindelin et al., 2012). Cilia data analysis was done using CiliaQ package (Hansen et al., 2021).

Quantitative real-time PCR (qPCR) (Fig. S3.2B, S3.11A, S3.12A-B)

qPCR primers

Target gene	**Sequence**	**Source**
hsGAPDH (fwd)	GTCTCCTCTGACTTCAACAGCG	HP205798, Origene, USA
hsGAPDH (rev)	ACCACCCTGTTGCTGTAGCCAA	
hsTXNDC15 (fwd)	GGCATTTCCAGCTCTTCACT	This study
hsTXNDC15 (rev)	GGCCATTGGTTTAGCTCCTT	
hsMARCHF6 (fwd)	TGACGCTGCCATTAGATATGC	This study
hsMARCHF6 (rev)	GGACTATCTGCTCTCTCAACCA	
mmPpia (fwd)	TTCACCTTCCCAAAGACCAC	(Griessl et al., 2017)
mmPpia (rev)	CAAACACAAACGGTTCCCAG	
mmTXNDC15 (fwd)	GTGGCCGAGAGAACCTCTTC	

mmTXNDC15 (rev)	TCCTCACAATTCACCTTGGGG	PrimerBank ID: 52693934c1 (Spandidos et al., 2008, 2010; X. Wang & Seed, 2003)

Primer pairs of hsTXNDC15 and hsMARCHF6 were designed using the NCBI Primer-BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

qPCR experiment: Cells were harvested, washed with 1x PBS, pelleted, and flash frozen. Total RNA from these cells were extracted using the Direct-zol RNA Miniprep Kit (R2050, Zymo Research, USA) and TRIzol Reagent (15596018, Invitrogen, USA). Reverse transcription of the total RNA was performed using the SuperScript™ III First-Strand Synthesis SuperMix (18080400, Invitrogen, USA). qPCR reactions were set up in triplicates using SYBR™ Green Universal Master Mix (4334973, Applied Biosystems, USA) and ran on the StepOnePlus™ Real-Time PCR System (4376600, Applied Biosystems, USA). GAPDH or Ppia were used as the housekeeping gene control and relative expression of each protein was calculated using the Pfaffl method (Pfaffl, 2001). Results are represented as relative expression to control cells (either wt or NT kd) and graphed using GraphPad Prism v.10.4.

Sequence alignment (Fig. 3.3C)

Sequence alignments in Fig. 3.3C were performed using Jalview 2.11.4.1 (Waterhouse et al., 2009). The sequences used for alignment were: hsTXNDC15 (Q96J42), hsTXNDC11 (Q6PKC3), hsTXNDC5 (Q8NBS9), hsPDIA4 (P13667).

Other software

Histograms, charts, and plots were generated using Python and GraphPad Prism v.10.4. Molecular models were generated using the PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

AlphaFold 3 predicted models were generated using the AlphaFold Server (Abramson et al., 2024).

Illustrations were made using Adobe Illustrator by Adobe Inc.

3.6. Supplementary materials

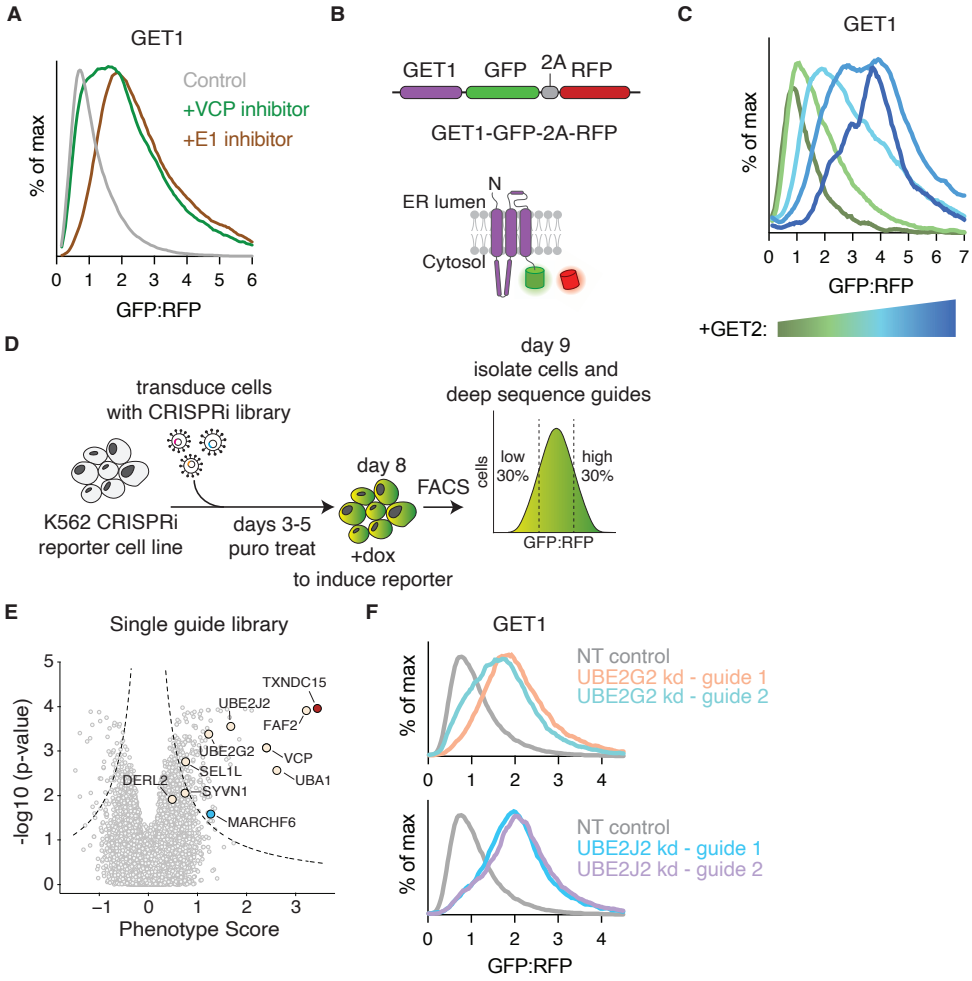

Figure S3.1. A platform for genome wide CRISPRi screens for quality control of orphan GET1.

(A) Unassembled GET1 is stabilized upon inhibition of the ubiquitin proteasome pathway. A GET1-GFP reporter construct was expressed in HEK293T cells along with a normalization control (RFP). Cells were treated with either the VCP inhibitor CB-5083 (250 nM) or the E1 inhibitor MLN7243 (1.25 μM) for 6 hours before analysis by flow cytometry. Stabilization of GET1-GFP relative to a normalization marker (RFP) upon drug treatment is consistent with its degradation by the ubiquitin proteasome system. For all experiments histograms are normalized such that the GFP:RFP ratio in wt cells is 1, as described in the methods. **(B)** Schematic of the GET1 ratiometric fluorescent reporter

construct used for genome wide CRISPR inhibition (CRISPRi) screens. GET1 was expressed as a C-terminal fusion with GFP along with a translation normalization marker, RFP, separated by a viral P2A (2A) peptide skipping sequence. Expression of GET1-GFP and RFP from a single open reading frame ensures that changes in the GFP:RFP ratio must occur post-translationally, facilitating identification of biogenesis and quality control machinery. **(C)** GET2 expression stabilizes the ratiometric GET1 fluorescent reporter. The GET1-GFP-2A-RFP reporter construct used for genome wide screens described in (B) and Fig. 3.1 was transduced in HEK293T cells along with GET2-2A-BFP and analyzed by flow cytometry. All histograms were normalized relative to the median of the GFP:RFP-ratio in wt cells without GET2. Increasing amounts of BFP expression, a proxy for GET2 levels, leads to a commensurate increase in GET1-GFP:RFP ratio. This GET2 dependent stabilization of GET1 suggests that our ratiometric reporter is predominantly unassembled and can therefore be used to study the quality control of orphan GET1. **(D)** Workflow of the FACS-based CRISPRi screen used to identify quality control factors for orphan GET1. A K562 CRISPRi reporter cell line was constructed containing the GET1-GFP reporter under an inducible promoter. Following transduction with a genome wide v2-CRISPRi sgRNA library (Horlbeck et al., 2016), the GET1-GFP reporter was induced with doxycycline for 36 hours prior to cell sorting. Cells were isolated based on ratiometric changes in GFP relative to RFP expression by FACS and the enriched sgRNAs were identified using deep sequencing. **(E)** Volcano plot of the genome wide CRISPRi screen for orphan GET1 highlighting known ERAD factors. Genes falling outside of the indicated dashed lines are considered statistically significant hits, including the indicated ERAD factors. **(F)** Fluorescent reporter assay for GET1 stability upon depletion of the indicated ERAD factors identified in the genome wide screen in (E). sgRNAs targeting the indicated E2 ubiquitin-conjugating enzymes were transduced in K562 cells constitutively expressing the CRISPRi machinery for six days. The GET1 reporter in (B) was transduced and cells were analyzed by flow cytometry after an additional 48 hours. Consistent with the results of the screen, depletion of both ER-resident E2 enzymes leads to stabilization of GET1.

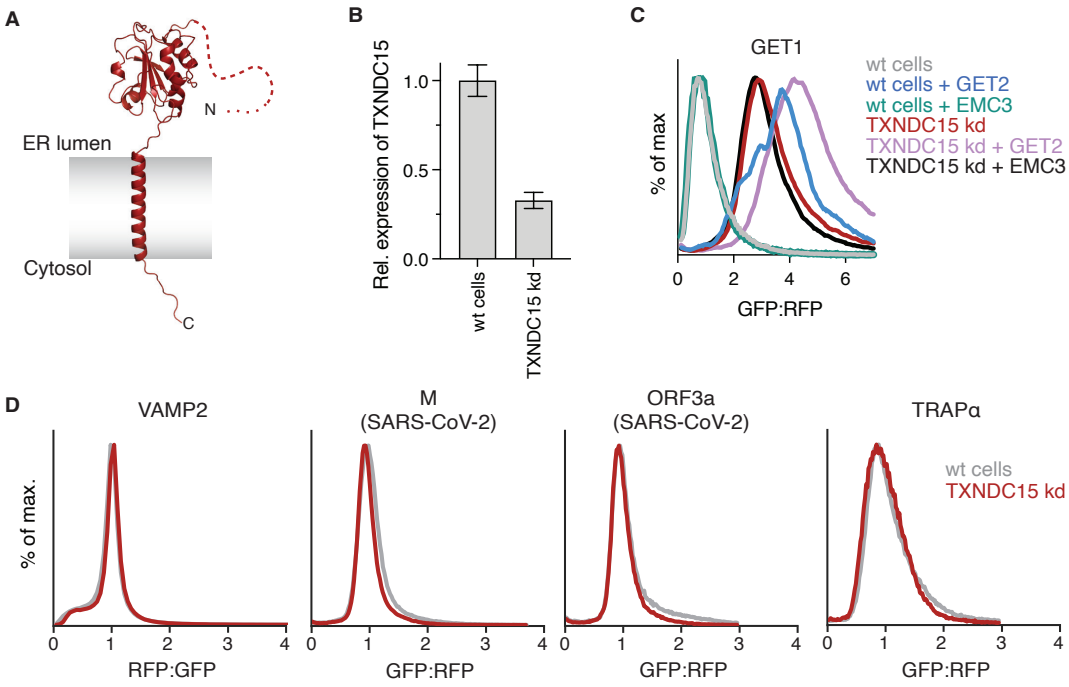

Figure S3.2. TXNDC15 is specifically required for degradation of orphan GET1, but not all ER membrane proteins.

(A) AlphaFold 3 predicted model of TXNDC15. After the signal sequence, the N-terminal 175 residues of the mature protein are predicted to be unstructured followed by a thioredoxin-like domain in the ER lumen and a single C-terminal TM. **(B)** Characterization of a HEK293T TXNDC15 stable knockdown cell line. HEK293T cells were transiently transfected with cas9 and an sgRNA targeting TXNDC15, along with GFP as a transfection marker. Transfected cells were sorted. Relative expression of TXNDC15 in a stable HEK293T TXNDC15 kd cell line compared to a wt control was quantified by qPCR. Errors bars represent errors as determined by error propagation, calculated from the primer efficiency and measurements taken from 3 technical replicates. **(C)** To test whether depletion of TXNDC15 specifically stabilized orphan GET1, but not the assembled GET1/2 complex, we tested whether expression of GET2 would rescue the effect of TXNDC15 depletion on GET1. To do this, we expressed saturating amounts of GET2 in either wt or TXNDC15 kd HEK293T cells along with the ratiometric GET1 fluorescent reporter described in Fig. S3.1B. We have previously shown that GET2 can assemble with GET1-GFP in cells, under similar conditions (Inglis et al., 2020). As a specificity control,

we also expressed the unrelated membrane protein subunit, EMC3. In the presence of GET2, kd of TXNDC15 has a much smaller effect on GET1 stability than when GET1 is expressed on its own and therefore cannot assemble. Overexpression of EMC3 does not affect GET1 stability in wt or TXNDC15 kd cells. Together these data suggested that TXNDC15 specifically acts on orphan GET1 but not the assembled GET1/2 complex. **(D)** To determine if depletion of TXNDC15 caused a general defect on ER proteostasis, we tested the effect of TXNDC15 depletion on a panel of topologically diverse membrane proteins and membrane protein subunits. We generated ratiometric fluorescent reporters for the following substrates: the single spanning tail anchored protein VAMP2; the three-TM containing viral proteins ORF3a and M from SARS-CoV-2, which both have the same topology as GET1; and the two-TM membrane protein subunit of the translocon-associated protein (TRAPa). These reporters were transduced into HEK293T wt and TXNDC15 kd cells, and the GFP:RFP ratio was analyzed by flow cytometry, and the histograms are shown. Given that loss of TXNDC15 did not affect the stability of these other membrane proteins, we concluded that depletion TXNDC15 does not cause non-specific dysregulation of the ER.

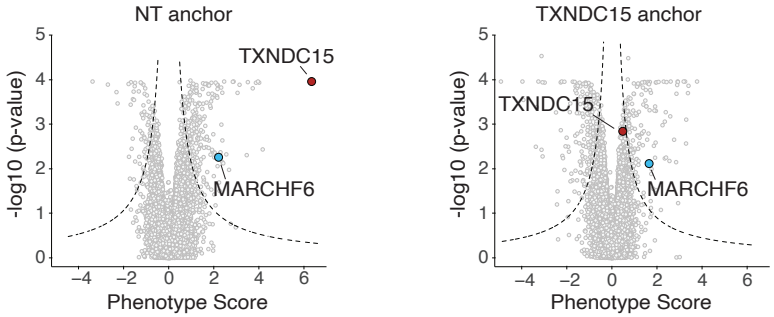

Figure S3.3. Relative phenotype of MARCHF6 and TXNDC15 in the genetic modifier screens described in Fig. 3.1G.

To systematically identify genetic interactors of TXNDC15, and thereby probe its potential function, a genome wide CRISPRi genetic modifier screen was performed. To do this, we prepared two dual guide libraries that expressed either a non-targeting control (left) or TXNDC15 (right) anchor guide in the first position along with a genome wide sgRNA library in the second position (Fig. 3.1G). Volcano plots summarizing the results of these screens, are shown, and MARCHF6 and TXNDC15 are highlighted.

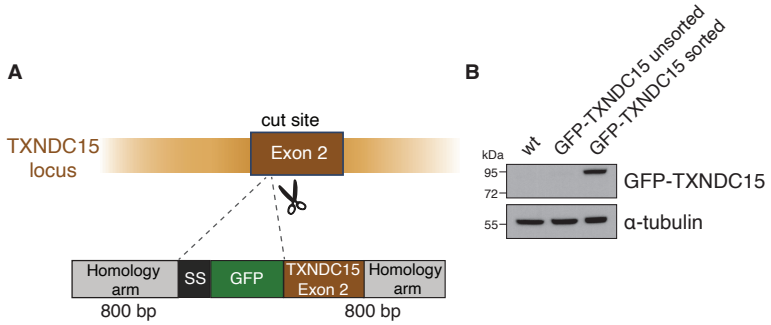

Figure S3.4. Endogenous tagging of TXNDC15 in HEK293T cells.

(A) Schematic of strategy for tagging endogenous TXNDC15 with a GFP at the N-terminus of the mature domain in HEK293T cells. A vector encoding GFP and 800 bp of homology on either side of the PAM cas9 cut site within the TXNDC15 locus was cloned. The GFP insertion site was designed such that the native TXNDC15 signal sequence (SS) and signal cleavage site would remain intact (see materials and methods). **(B)** After homologous recombination, HEK293T cells were sorted for GFP fluorescence, and a monoclonal cell line was established. Western blot of the endogenously tagged GFP-TXNDC15 cell line using the GFP nanobody, with α-tubulin as a loading control.

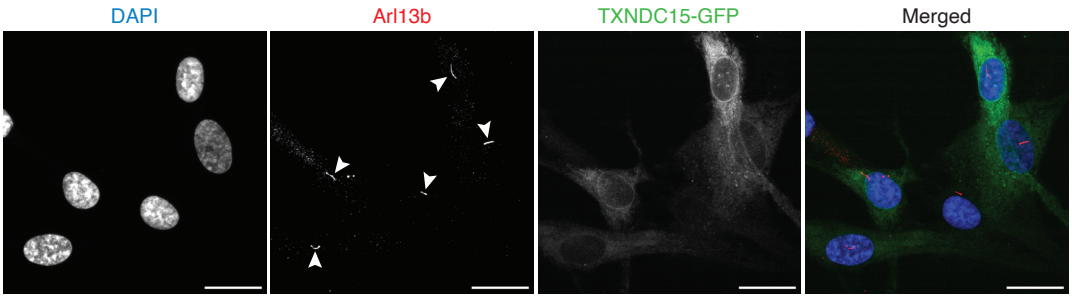

Figure S3.5. Exogenously expressed TXNDC15-GFP does not localize to the primary cilia in human RPE1 cells.

Confocal imaging of RPE1 cells transduced with TXNDC15-GFP (green) subjected to serum starvation to induce cilia formation. Shown is staining with DAPI (blue) for nuclear DNA and Arl13b (red) as a ciliary marker. Cilia are demarcated with arrows in the Arl13b channel. Even upon overexpression of TXNDC15, we did not observe colocalization of TXNDC15 with the Arl13b marker. These data, together with imaging of the endogenously tagged TXNDC15 shown in Fig. 3.2A suggested that TXNDC15 is localized primarily to the ER.

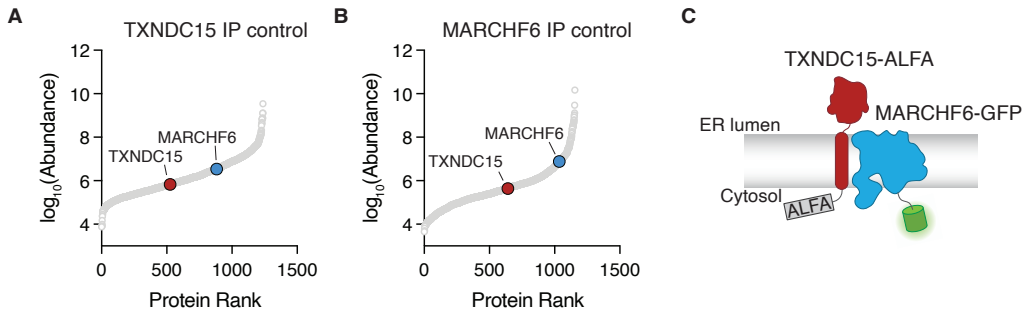

Figure S3.6. The association of TXNDC15 with MARCHF6 is specific.

(A) To ensure that the observed interaction of MARCHF6 with TXNDC15 shown in Fig. 3.2B was specific, we used wt HEK293T cells containing untagged TXNDC15 as the background control for immunoprecipitation of endogenously tagged GFP-TXNDC15. Non-specific interaction partners of the GFP nanobody were analyzed by mass spectrometry and were ranked based on the log10(abundance) and plotted as protein rank versus log10(abundance). No enrichment of TXNDC15 or MARCHF6 was detected. **(B)** Mass spectrometry analysis of the background control for the immunoprecipitation of exogenously expressed MARCHF6 in Fig. 3.2C. Expi293 cells overexpressing MARCHF6(C9A)-GFP was purified in the absence of the GFP nanobody to assess non-specific interactors. Similar to (A), MARCHF6 and TXNDC15 were not enriched under these conditions, suggesting that the interaction between MARCHF6 and TXNDC15 is specific. **(C)** Schematic of TXNDC15-ALFA and MARCHF6-GFP used for the purification of the recombinant TXNDC15/MARCHF6 complex in Fig. 3.2D. A stable Expi293 cell line of the catalytically inactive and thereby more stable MARCHF6(C9A)-GFP and TXNDC15-ALFA was generated. Sequential immunoprecipitation of TXNDC15 followed by MARCHF6 were utilized to purify the complex (see materials and methods).

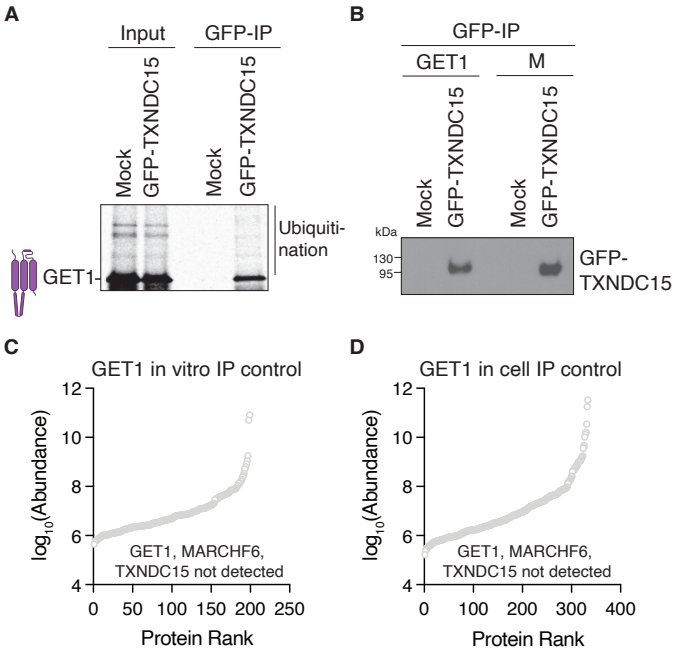

Figure S3.7. Analysis of the physical association between GET1 and TXNDC15.

(A) Depicted is a longer exposure of the experiment described in Fig. 3.2F in order to visualize ubiquitination of orphan GET1. Radioactively labeled GET1 was translated in RRL in the presence of hRMs derived from HEK293T cells that expressed either endogenous wt TXNDC15 (mock) or endogenously tagged GFP-TXNDC15. Membrane fractions were isolated by ultracentrifugation, solubilized in the detergent GDN (input), and immunoprecipitated under native conditions using the GFP-nanobody. Radioactively labeled products were analyzed by SDS-PAGE and autoradiography. The fraction of GET1 that specifically co-purifies with TXNDC15 is enriched for ubiquitinated species. **(B)** Western blot analysis to compare the levels of bait (GFP-TXNDC15) in the elutions of the SARS-CoV-2 M protein and GET1 following the in vitro translation and immunoprecipitations shown in Fig. 3.2F and described in (A). Equal amounts of GFP-TXNDC15 were recovered in both the experimental (GET1) and control (M) ensuring that the decreased signal observed for M was not due to different amounts of GFP-TXNDC15 in the two samples. **(C)** To ensure that the enrichment of MARCHF6 and TXNDC15 with in vitro translated GET1 from hRMs was specific (Fig. 3.2G), we analyzed a mock control. In parallel to translation of GET1 in RRL in the presence of hRMs, we performed an identical reaction without mRNA which was subjected to immunoprecipitation using the

ALFA nanobody (ALFA-IP). Displayed is mass spectrometry analysis of these non-specific interactors in which proteins were ranked based on the log10(abundance) and plotted as protein rank versus log10(abundance). GET1, TXNDC15, MARCHF6 were not detected in this control sample. **(D)** As in (C) for the control sample for immunoprecipitations of GET1-GFP-2A-RFP from HEK293T cells shown in Fig. 3.2H. GFP expressed in HEK293T cells was subjected to immunoprecipitation using the GFP nanobody (GFP-IP) and analyzed by mass spectrometry. Proteins were ranked based on the log10(abundance) and plotted as protein rank versus log10(abundance). GET1, TXNDC15, MARCHF6 were not detected in this control sample.

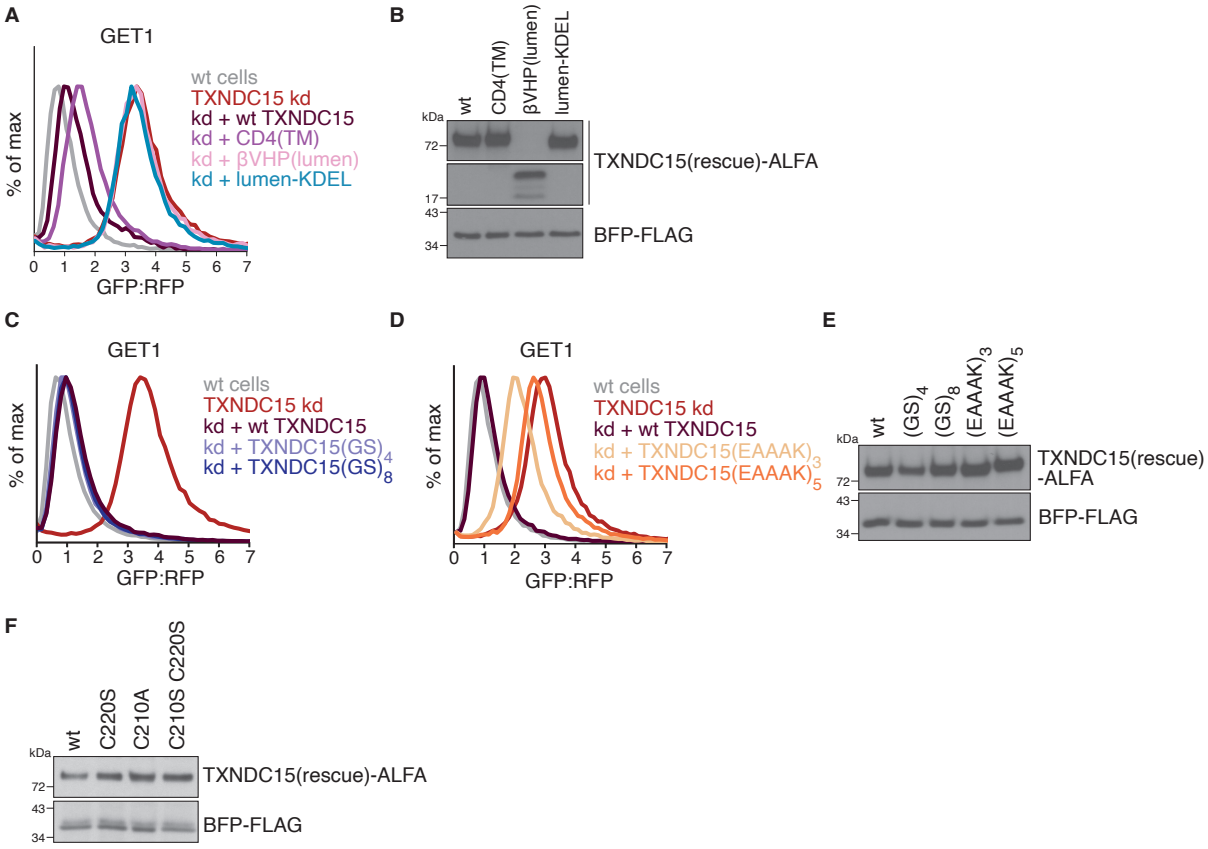

Figure S3.8. Mutational analysis of TXNDC15.

(A) Representative histograms of the TXNDC15 rescue assay shown in Fig. 3.3A used to identify the domains of TXNDC15 required for its quality control function. A series of chimeric TXNDC15 constructs was generated in which the TM or lumenal domain was replaced, along with a BFP expression control separated by a viral 2A sequence. Wt or HEK293T TXNDC15 stable kd cells were transduced with the indicated rescue constructs or a BFP control. 24 hours later, these cells were transduced with the GET1-GFP ratiometric reporter, and analyzed by flow cytometry following 48 hours of reporter expression. While TXNDC15 is still functional when the TM is exchanged with that from an unrelated protein, CD4, exchanging the TXNDC15 lumenal domain with a bVHP globular domain abolished its activity. We also found that the lumenal domain of TXNDC15 must be anchored at the membrane as expression of the TXNDC15 lumenal domain fused to an ER retention sequence (KDEL) in the absence of a TMD could not rescue the TXNDC15 kd phenotype. **(B)** Western blots of cells expressing the TXNDC15

rescue constructs described in (A) to ensure that mutants accumulate to similar levels. Therefore, the observed differences in their quality control activity cannot be explained by differences in expression level or protein stability. TXNDC15 rescue constructs contained an expression marker of 2A-BFP-FLAG (see materials and methods) to ensure similar transduction efficiency. **(C, D)** Representative histograms for the data summarized in Fig. 3.3B in which we test whether the positioning of the TXNDC15 lumenal domain is important for its quality control function. To do this, the loop between the lumenal domain and TM of TXNDC15 was replaced with either a flexible linker of length 4 or 8 [$(GS)_4$, $(GS)_8$] (C) or rigid linkers of length 3 or 5 [$(EAAAK)_3$, $(EAAAK)_5$] (D). Flexible linkers are tolerated while increasing the distance between the TXNDC15 lumenal domain and the membrane using a rigid linker is not. **(E)** Western blot of the TXNDC15 mutant constructs shown in (C) and (D) in HEK293T cells, as described in (B). **(F)** Western blot of the TXNDC15 mutant rescue constructs shown in Fig. 3.3D in HEK293T cells, as described in (B).

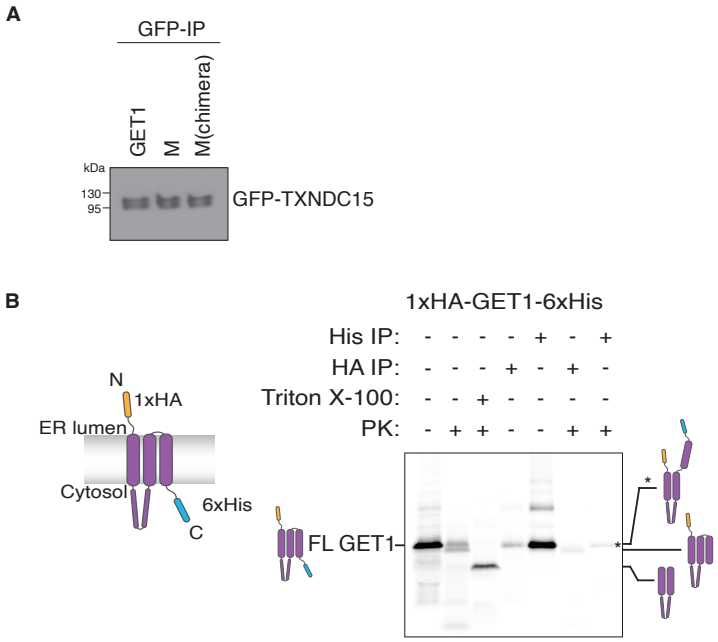

Figure S3.9. In vitro analysis of TXNDC15 substrates binding and GET1 topology.

(A) Western blot analysis of the amount of GFP-TXNDC15 in the GFP-IP elution samples of the experiment in Fig. 3.4B, testing the interaction between GET1, M, and M(chimera) with TXNDC15 in vitro. The efficiency of the GFP-TXNDC15 IP is comparable in all reactions. **(B)** Analysis of GET1 topology via Proteinase K (PK) protection assay. (Left) Schematic of GET1 expression construct with HA- and His-tags on the N- and C-termini respectively. (Right) ^{35}S-methionine labeled HA-GET1-His was translated in RRL in the presence of cRMs. The translation products were then digested with PK in the presence or absence of detergent. Samples were analyzed by SDS-PAGE either directly or following immunoprecipitation using the His or HA tag. Upon treatment with PK, a second species appears along with full length GET1 that lacks the C-terminal His tag, while the coiled coil secondary structure protects the loop between TM1 and TM2 from being cleaved. Only in the presence of detergent (Triton X-100) is the N-terminal HA tag degraded and the loop between TM2 and TM3 cut. After PK digestion, the HA-IP recovers the fragment of GET1 without the C-terminus. These results suggest that the majority of GET1 is properly folded (see schematic on the left), while the His IP after PK treatment displays a small amount of full length GET1 indicating the presence of misfolded GET1 with the C-terminus and TM3 being protected in the ER lumen (*). Therefore, while we observed efficient modification

of a C-terminal glycosylation tag on in vitro translated GET1 due to incorrect folding (Fig. 3.4C), these PK results suggested that in most cases exposure of TM3 to the ER lumen is transient.

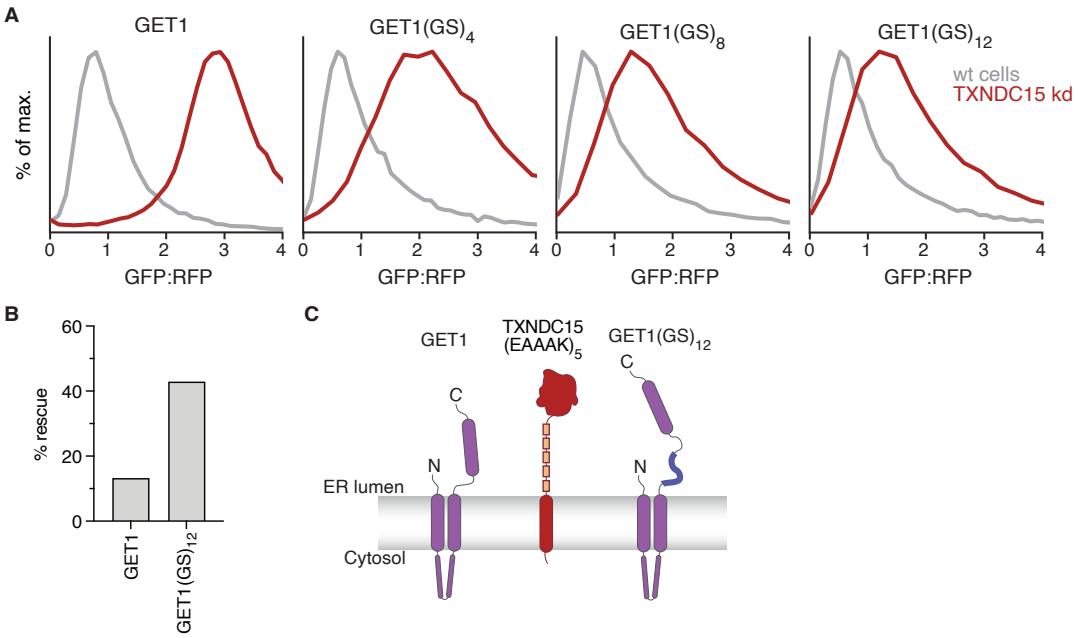

Figure S3.10. The positioning of the GET1 degron and the TXNDC15 lumenal domain from the bilayer is critical for substrate triage by TXNDC15.

(A) To test whether the localization of the substrate degron was important for recognition by TXNDC15, we introduced a linker between TM2 and the GET1 degron identified in Fig. 3.4 (loop2+TM3). We generated a series of GET1 mutants in the GFP-2A-RFP cassette with varying lengths of a flexible (GS) linker of n repeats (n = 0, 4, 8, 12). These linkers would allow the GET1 degron to be access regions of the ER lumen farther from the bilayer. Representative histograms of a reporter assay testing the stability of these GET1 substrates in wt versus TXNDC15 kd HEK293T cells as analyzed by flow cytometry. GET1 constructs with longer linker lengths were less TXNDC15 dependent. **(B)** Given that allowing the GET1 degron to extend further from the bilayer makes it less efficiently recognized by TXNDC15 (A), we tested whether a compensatory increase in the distance of the TXNDC15 lumenal domain might increase TXNDC15 dependent degradation. To do this, we tested whether TXNDC15 containing a rigid $(EAAAK)_5$ linker that increases the distance of its lumenal domain from the bilayer could support degradation of wildtype GET1 versus GET1 with the $(GS)_{12}$ flexible linker. We found that TXNDC15 with a long rigid linker could rescue a larger proportion of substrate degradation when the GET1 degron was localized farther from the membrane. This is consistent with the model that

TXNDC15 uses its lumenal domain to recognize substrates that contain a hydrophobic degron in proximity of the ER membrane. **(C)** Schematic of GET1 and TXNDC15 constructs used in the rescue assay in (B). TXNDC15$(EAAAK)_5$ denotes the construct with a rigid linker that increases the distance between the TXNDC15 lumenal domain and the bilayer. When the lumenal domain of TXNDC15 is further away from the bilayer, it is better able to recognize the GET1 mutant in which its degron can be positioned farther from the membrane.

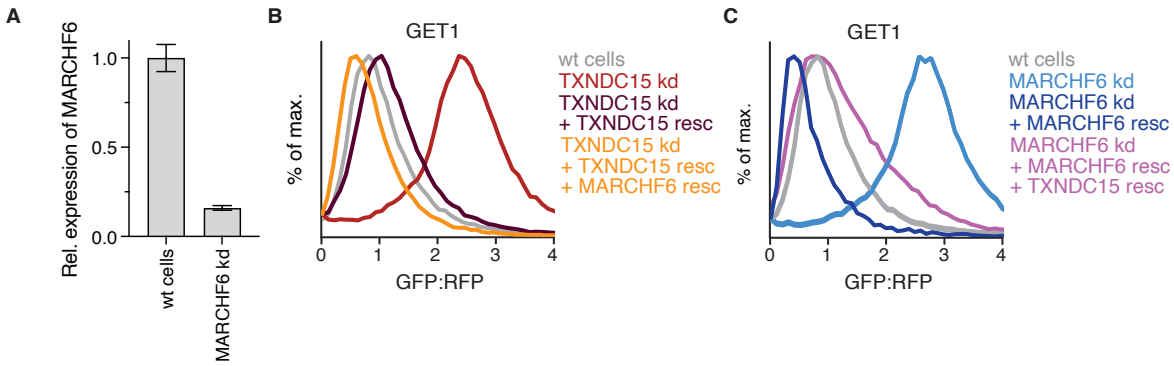

Figure S3.11. TXNDC15 and MARCHF6 play distinct functional roles in membrane protein quality control.

(A) Characterization of a HEK293T MARCHF6 stable knockdown cell line. HEK293T cells were transiently transfected with Cas9 and an sgRNA targeting MARCHF6, along with GFP as a transfection marker. Transfected cells were sorted. Relative expression of MARCHF6 in a HEK293T MARCHF6 stabled kd cell line compared to a wt control was quantified by qPCR. Errors bars represent errors as determined by error propagation, calculated from the primer efficiency and measurements taken from 3 technical replicates. **(B,C)** While data in Fig. 3.4G showed that overexpression of MARCHF6 cannot rescue loss of TXNDC15 in quality control and vice versa, complementary controls show that exogenous expression of TXNDC15 and MARCHF6 can rescue their own depletion. (B) Rescue assays were performed in TXNDC15 kd HEK293T cells using the GET1-GFP-2A-RFP reporter. HEK293T TXNDC15 stable knockdown cells were transduced with either TXNDC15 alone or TXNDC15 and MARCHF6. Expression of TXNDC15 alone or in combination with MARCHF6 can rescue TXNDC15 kd phenotype on GET1 reporter. (C) Rescue assay in MARCHF6 kd HEK293T cells using the GET1-GFP-2A-RFP reporter. HEK293T MARCHF6 stable knockdown cells were transduced with either MARCHF6 alone or MARCHF6 with TXNDC15 which can rescue MARCHF6 kd phenotype on the GET1 reporter. Curves shown for wt cells, TXNDC15 kd and MARCHF6 kd are the same as in Fig 3.4G.

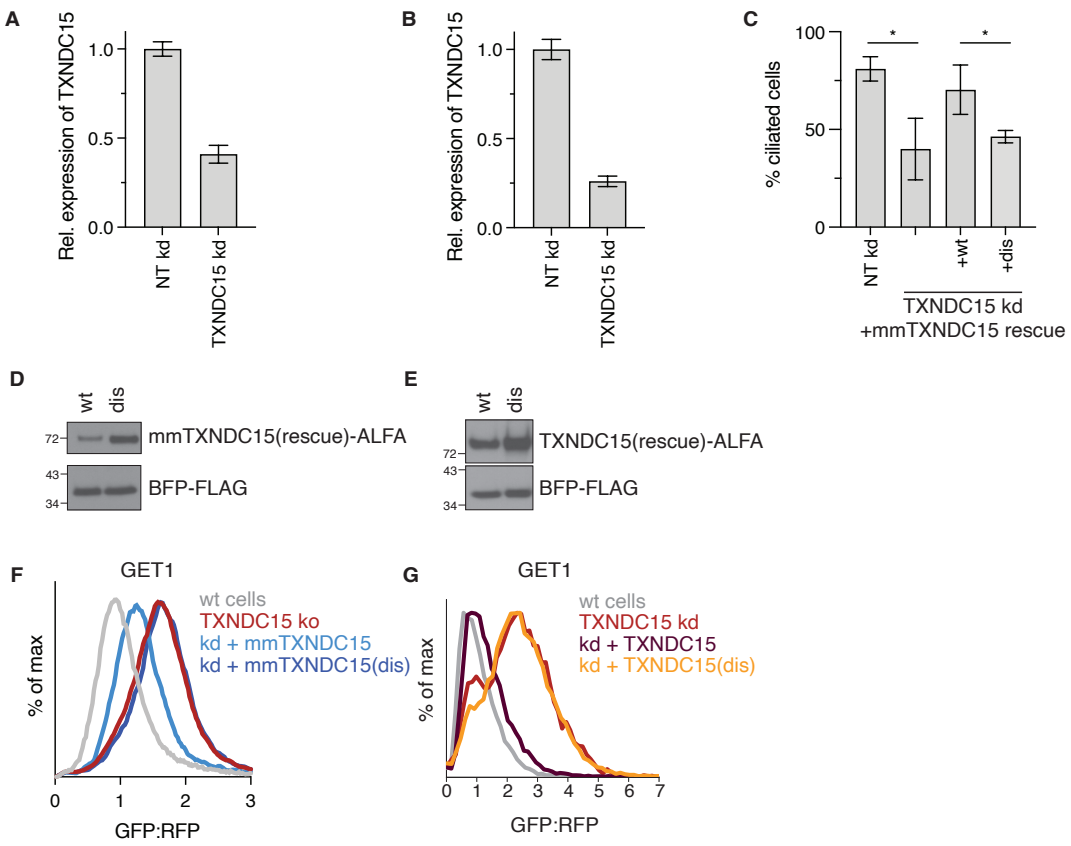

Figure S3.12. Characterization of TXNDC15 disease mutation.

(A) Characterization of a murine NIH 3T3 TXNDC15 stable CRISPRi knockdown cell line used in Fig. 3.5D-E and (D), (F). NIH 3T3 cells expressing dCas9 were transduced with a guide targeting TXNDC15 along with a puromycin selection marker. Transduced cells were subjected to puromycin selection. Relative expression of TXNDC15 in the NIH 3T3 TXNDC15 stable kd cell line compared to a NT control was quantified by qPCR. Errors bars represent errors as determined by error propagation, calculated from the primer efficiency and measurements taken from 3 technical replicates. **(B)** Characterization of RPE1 cells subjected to CRISPRi TXNDC15 knockdown that were used in Fig. 3.5F-G and (H). RPE1 cells expressing dCas9 were transduced with sgRNA targeting TXNDC15 along with a puromycin selection marker. Transduced cells were subjected to puromycin selection. Relative expression of TXNDC15 in the RPE1 transduced with TXNDC15

sgRNA guide compared to a NT control was quantified by qPCR. Errors bars represent errors as determined by error propagation, calculated from the primer efficiency and measurements taken from 3 technical replicates. **(C)** Quantification of ciliated NIH 3T3 cells in wt and TXNDC15 kd cell lines. TXNDC15 kd was rescued by overexpression of either murine mmTXNDC15 wt or the mmTXNDC15 ciliopathy-mutant (dis, Δ209-213). Ciliated cells were imaged by immunofluorescence microscopy (see Fig. 3.5C) and the number of cilia and cells per microscope slide were counted. Mean and standard deviation for 3 replicates are shown (wt: 420 cells; TXNDC15 kd: 445 cells, mmTXNDC15 wt rescue: 207 cells; mmTXNDC15(dis): 184 cells). **(D)** Determination of expression levels by Western blot analysis of overexpressed murine mmTXNDC15 in TXNDC15 kd 3T3 NIH cells as used in (C), (F), and Fig. 3.5D-E. ALFA-tagged mmTXNDC15 was overexpressed from the same mRNA as BFP-FLAG so that the sample was normalized to the number of cells expressing BFP. Western blots were performed using the anti-ALFA-Nb-HRP and anti-FLAG-HRP antibody, respectively. **(E)** Western blot analysis of overexpression levels of TXNDC15 wt and the ciliopathy mutant (dis, Δ225-229) in TXNDC15 kd HEK293T cells as used in Fig. 3.5H. Samples were analyzed as described in (D). **(F)** Analysis of the rescue efficiencies of mmTXNDC15 wt and mmTXNDC15(dis) in NIH 3T3 cells. The stability of the GET1-GFP-2A-RFP reporter was analyzed by flow cytometry and curves were normalized to the median GFP:RFP-ratio of the NT control. mmTXNDC15 wt overexpression leads to a partial rescue of the TXNDC15 kd phenotype while the ciliopathy mutant does not rescue GET1 degradation. These results are consistent with those in Fig. 3.5H, where the same assay was performed in the human HEK293T cells. **(G)** Analysis of rescue efficiencies of the TXNDC15 wt and (dis, Δ225-229) ciliopathy mutant in RPE1 cells as described in (F). TXNDC15 wt can fully rescue the TXNDC15 kd phenotype while the ciliopathy mutant cannot.

CHAPTER 4:

CONCLUSIONS

In recent decades, research in the field of membrane protein biogenesis and quality control has seen significant progress. Developments in techniques in functional genomics, cellular biology, molecular biology, and structural biology have accelerated our understanding of how membrane proteins are made and degraded, as well as how the cell has evolved to have intricate and complex protein machinery to maintain balance in the membrane proteome. Advancements in functional genomics and molecular biological techniques have enabled identification of various factors and mapping of molecular pathways that regulate the insertion and degradation of membrane proteins; and especially in the field of structural biology, the resolution revolution of cryo-EM has allowed for molecular structure determination of various membrane proteins individually as well as in complexes in their native state within these pathways.

The work in this thesis explored how protein machineries coordinate to accommodate the insertion of diverse of membrane protein substrates at the ER. For many years, it was postulated that the Sec61 translocon is responsible for the incorporation of most integral membrane proteins into the membrane. However, the diversity of membrane proteins found in the cell alludes to the hypothesis that Sec61 alone is not sufficient, and many other insertases beyond Sec61 have been discovered (GET1/2, EMC, and the GEL complex). Using functional genomics, immunoprecipitations, and mass spectrometry, we were able to show that the EMC genetically and physically interacts with the BOS complex, which is a component of the multipass translocon at Sec61 that helps to insert multipass membrane proteins, such as GPCRs. Additionally, I was able to structurally characterize the BOS complex at high resolution as well as the EMC • BOS holo-complex at low resolution, highlighting their interaction interface. From the structure, it appeared that EMC occupies the same space as Sec61 in the multipass translocon, suggesting that there is coordination between these biogenesis factors at the ER. Using cell-based assays, we were also able to show that the biophysical properties (e.g., charge, length, secondary

structure) of the soluble domain of an N_{exo} multipass membrane protein dictate the pathway that inserts its first TMD co-translationally, either at the EMC alone or the EMC with additional factors at the multipass translocon, before the protein is handed off to the multipass translocon at Sec61 for the insertion of the remaining TMDs. Our work fills in the knowledge gap of how the EMC coordinates with other biogenesis machineries in the multipass translocon at the ER and how they cooperate to insert diverse protein substrates. Nevertheless, more research needs to be done to define the exact rule set that determines insertase specificity. Moreover, the mechanistic details of how the substrate transfer between these biogenesis factors at the ER membrane remain elusive. Structural determination and biochemical assays looking into the substrate-insertase intermediates will provide valuable insights towards this goal.

In addition, this thesis work also seeks out to investigate a potential mechanism by which orphaned subunits of membrane protein complexes get targeted for ERAD. Using a genome-wide CRISPRi screen, we identified a novel quality control factor of unknown function TXNDC15 that localizes to the ER. We found that TXNDC15 works with the E3 ligase MARCHF6 to degrade a specific set of membrane protein substrates at the ER. I showed that the lumenal domain of TXNDC15 is essential for its quality control function and it recognizes exposed hydrophobic patches in the ER lumen. This is particularly interesting as MARCHF6 had previously only been shown to function on ERAD-M and ERAD-C substrates. As MARCHF6 lacks lumenal soluble domains, we hypothesize that TXNDC15 helps MARCHF6 to recognize a set of membrane proteins with exposed lumenal domains. Although not well-studied, TXNDC15 had been implicated in ciliopathies, such as the lethal Meckel-Gruber syndrome. I demonstrated that the Meckel-Gruber patient mutation in TXNDC15 fails to rescue degradation of our protein substrate and fails to rescue cilia formation in TXNDC15-depleted cells, linking TXNDC15's quality control function to ciliopathies. Nonetheless, more investigation is needed to clarify the recognition mechanism of TXNDC15, and to pinpoint how defect in a quality control factor in the ER leads to defect in cilia formation.

BIBLIOGRAPHY

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A., & Adams, P. D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D: Structural Biology*, *74*(6), Article 6. https://doi.org/10.1107/S2059798318006551

Ai, H., Shen, W., Sagi, A., Chen, P. R., & Schultz, P. G. (2011). Probing Protein–Protein Interactions with a Genetically Encoded Photo-crosslinking Amino Acid. *ChemBioChem*, *12*(12), 1854–1857. https://doi.org/10.1002/cbic.201100194

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), 3708–3716. https://doi.org/10.1016/j.celrep.2017.12.006

Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering, Design and Selection*, *14*(8), 529–532. https://doi.org/10.1093/protein/14.8.529

Babu, M., Vlasblom, J., Pu, S., Guo, X., Graham, C., Bean, B. D. M., Burston, H. E., Vizeacoumar, F. J., Snider, J., Phanse, S., Fong, V., Tam, Y. Y. C., Davey, M., Hnatshak, O., Bajaj, N., Chandran, S., Punna, T., Christopolous, C., Wong, V., … Greenblatt, J. F. (2012). Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. *Nature*, *489*(7417), 585–589. https://doi.org/10.1038/nature11354

Bai, L., You, Q., Feng, X., Kovach, A., & Li, H. (2020). Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature*, *584*(7821), Article 7821. https://doi.org/10.1038/s41586-020-2389-3

Bays, N. W., Gardner, R. G., Seelig, L. P., Joazeiro, C. A., & Hampton, R. Y. (2001). Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nature Cell Biology*, *3*(1), 24–29. https://doi.org/10.1038/35050524

Bence, N. F., Sampat, R. M., & Kopito, R. R. (2001). Impairment of the Ubiquitin-Proteasome System by Protein Aggregation. *Science*, *292*(5521), 1552–1555. https://doi.org/10.1126/science.292.5521.1552

Bergeron, J. J., Brenner, M. B., Thomas, D. Y., & Williams, D. B. (1994). Calnexin: A membrane-bound chaperone of the endoplasmic reticulum. *Trends in Biochemical Sciences*, *19*(3), 124–128. https://doi.org/10.1016/0968-0004(94)90205-4

Bodnar, N. O., & Rapoport, T. A. (2017). Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex. *Cell*, *169*(4), 722-735.e9. https://doi.org/10.1016/j.cell.2017.04.020

Bordallo, J., Plemper, R. K., Finger, A., & Wolf, D. H. (1998). Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. *Molecular Biology of the Cell*, *9*(1), 209–222. https://doi.org/10.1091/mbc.9.1.209

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 (London, England: 1993)*, *23*(9), 1715–1724. https://doi.org/10.1016/j.str.2015.06.025

Breslow, D. K., Hoogendoorn, S., Kopp, A. R., Morgens, D. W., Vu, B. K., Kennedy, M. C., Han, K., Li, A., Hess, G. T., Bassik, M. C., Chen, J. K., & Nachury, M. V. (2018). A CRISPR-based screen for Hedgehog signaling provides insights into ciliary function and ciliopathies. *Nature Genetics*, *50*(3), 460–471. https://doi.org/10.1038/s41588-018-0054-7

Cabantous, S., Terwilliger, T. C., & Waldo, G. S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology*, *23*(1), 102–107. https://doi.org/10.1038/nbt1044

Carvalho, H. J. F., Del Bondio, A., Maltecca, F., Colombo, S. F., & Borgese, N. (2019). The WRB Subunit of the Get3 Receptor is Required for the Correct Integration of

its Partner CAML into the ER. *Scientific Reports*, *9*(1), 11887. https://doi.org/10.1038/s41598-019-48363-2

Carvalho, P., Goder, V., & Rapoport, T. A. (2006). Distinct Ubiquitin-Ligase Complexes Define Convergent Pathways for the Degradation of ER Proteins. *Cell*, *126*(2), 361–373. https://doi.org/10.1016/j.cell.2006.05.043

Carvalho, P., Stanley, A. M., & Rapoport, T. A. (2010). Retro-translocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p. *Cell*, *143*(4), 579–591. https://doi.org/10.1016/j.cell.2010.10.028

Casañal, A., Lohkamp, B., & Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Science*, *29*(4), 1055–1064. https://doi.org/10.1002/pro.3791

Chen, S., Francioli, L. C., Goodrich, J. K., Collins, R. L., Kanai, M., Wang, Q., Alföldi, J., Watts, N. A., Vittal, C., Gauthier, L. D., Poterba, T., Wilson, M. W., Tarasova, Y., Phu, W., Grant, R., Yohannes, M. T., Koenig, Z., Farjoun, Y., Banks, E., … Karczewski, K. J. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. *Nature*, *625*(7993), 92–100. https://doi.org/10.1038/s41586-023-06045-0

Chen, Z., Mondal, A., Abderemane-Ali, F., Jang, S., Niranjan, S., Montaño, J. L., Zaro, B. W., & Minor, D. L. (2023). EMC chaperone–CaV structure reveals an ion channel assembly intermediate. *Nature*, *619*(7969), Article 7969. https://doi.org/10.1038/s41586-023-06175-5

Chiti, F., & Dobson, C. M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry*, *75*(1), 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901

Chitwood, P. J., & Hegde, R. S. (2020). An intramembrane chaperone complex facilitates membrane protein biogenesis. *Nature*, *584*(7822), 630–634. https://doi.org/10.1038/s41586-020-2624-y

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. https://doi.org/10.1016/j.cell.2018.10.009

Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W., & Kopito, R. R. (2011). Defining human ERAD networks through an integrative mapping strategy. *Nature Cell Biology*, *14*(1), 93–105. https://doi.org/10.1038/ncb2383

Christianson, J. C., Shaler, T. A., Tyler, R. E., & Kopito, R. R. (2008). OS-9 and GRP94 deliver mutant α1-antitrypsin to the Hrd1–SEL1L ubiquitin ligase complex for ERAD. *Nature Cell Biology*, *10*(3), 272–282. https://doi.org/10.1038/ncb1689

Cohen, I., Wiener, R., Reiss, Y., & Ravid, T. (2015). Distinct activation of an E2 ubiquitin-conjugating enzyme by its cognate E3 ligases. *Proceedings of the National Academy of Sciences*, *112*(7), E625–E632. https://doi.org/10.1073/pnas.1415621112

Colombo, S. F., Cardani, S., Maroli, A., Vitiello, A., Soffientini, P., Crespi, A., Bram, R. F., Benfante, R., & Borgese, N. (2016). Tail-anchored Protein Insertion in Mammals: FUNCTION AND RECIPROCAL INTERACTIONS OF THE TWO SUBUNITS OF THE TRC40 RECEPTOR *. *Journal of Biological Chemistry*, *291*(29), 15292–15306. https://doi.org/10.1074/jbc.M115.707752

Costello, J. L., Castro, I. G., Camões, F., Schrader, T. A., McNeall, D., Yang, J., Giannopoulou, E.-A., Gomes, S., Pogenberg, V., Bonekamp, N. A., Ribeiro, D., Wilmanns, M., Jedd, G., Islinger, M., & Schrader, M. (2017). Predicting the targeting of tail-anchored proteins to subcellular compartments in mammalian cells. *Journal of Cell Science*, *130*(9), 1675–1687. https://doi.org/10.1242/jcs.200204

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

Cymer, F., Veerappan, A., & Schneider, D. (2012). Transmembrane helix–helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1818*(4), 963–973. https://doi.org/10.1016/j.bbamem.2011.07.035

Dephoure, N., Hwang, S., O'Sullivan, C., Dodgson, S. E., Gygi, S. P., Amon, A., & Torres, E. M. (2014). Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. *eLife*, *3*, e03023. https://doi.org/10.7554/eLife.03023

Dettmer, U., Kuhn, P.-H., 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), 26174–26181. https://doi.org/10.1074/jbc.M110.132548

Doblmann, J., Dusberger, F., Imre, R., Hudecz, O., Stanek, F., Mechtler, K., & Dürnberger, G. (2019). apQuant: Accurate Label-Free Quantification by Quality Filtering. *Journal of Proteome Research*, *18*(1), 535–541. https://doi.org/10.1021/acs.jproteome.8b00113

Dolan, K. A., Dutta, M., Kern, D. M., Kotecha, A., Voth, G. A., & Brohawn, S. G. (2022). Structure of SARS-CoV-2 M protein in lipid nanodiscs. *eLife*, *11*, e81702. https://doi.org/10.7554/eLife.81702

Elsässer, S. J., Ernst, R. J., Walker, O. S., & Chin, J. W. (2016). Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nature Methods*, *13*(2), 158–164. https://doi.org/10.1038/nmeth.3701

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), Article 4. https://doi.org/10.1107/S0907444910007493

Enquist, K., Fransson, M., Boekel, C., Bengtsson, I., Geiger, K., Lang, L., Pettersson, A., Johansson, S., von Heijne, G., & Nilsson, I. (2009). Membrane-integration Characteristics of Two ABC Transporters, CFTR and P-glycoprotein. *Journal of Molecular Biology*, *387*(5), 1153–1164. https://doi.org/10.1016/j.jmb.2009.02.035

Fry, M. Y., Saladi, S. M., Cunha, A., & Clemons, W. M. (2021). Sequence-based features that are determinant for tail-anchored membrane protein sorting in eukaryotes. *Traffic (Copenhagen, Denmark)*, *22*(9), 306–318. https://doi.org/10.1111/tra.12809

Gallant, J. L., Heunis, T., Sampson, S. L., & Bitter, W. (2020). ProVision: A web-based platform for rapid analysis of proteomics data processed by MaxQuant. *Bioinformatics*, *36*(19), 4965–4967. https://doi.org/10.1093/bioinformatics/btaa620

Gemmer, M., Chaillet, M. L., van Loenhout, J., Cuevas Arenas, R., Vismpas, D., Gröllers-Mulderij, M., Koh, F. A., Albanese, P., Scheltema, R. A., Howes, S. C., Kotecha, A., Fedry, J., & Förster, F. (2023). Visualization of translation and protein biogenesis at the ER membrane. *Nature*, *614*(7946), Article 7946. https://doi.org/10.1038/s41586-022-05638-5

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C., Qi, L. S., Kampmann, M., & Weissman, J. S. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, *159*(3), 647–661. https://doi.org/10.1016/j.cell.2014.09.029

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, *27*(1), 14–25. https://doi.org/10.1002/pro.3235

Görlich, D., Hartmann, E., Prehn, S., & Rapoport, T. A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature*, *357*(6373), Article 6373. https://doi.org/10.1038/357047a0

Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., & Frey, S. (2019). The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nature Communications*, *10*(1), 4403. https://doi.org/10.1038/s41467-019-12301-7

Griessl, M., Gutknecht, M., & Cook, C. H. (2017). Determination of suitable reference genes for RT-qPCR analysis of murine Cytomegalovirus in vivo and in vitro. *Journal of Virological Methods*, *248*, 100–106. https://doi.org/10.1016/j.jviromet.2017.06.012

Grudnik, P., Bange, G., & Sinning, I. (2009). *Protein targeting by the signal recognition particle*. *390*(8), 775–782. https://doi.org/10.1515/BC.2009.102

Guna, A., Hazu, M., Pinton Tomaleri, G., & Voorhees, R. M. (2023). A TAle of Two Pathways: Tail-Anchored Protein Insertion at the Endoplasmic Reticulum. *Cold Spring Harbor Perspectives in Biology*, *15*(3), a041252. https://doi.org/10.1101/cshperspect.a041252

Guna, A., & Hegde, R. S. (2018). Transmembrane Domain Recognition during Membrane Protein Biogenesis and Quality Control. *Current Biology: CB*, *28*(8), R498–R511. https://doi.org/10.1016/j.cub.2018.02.004

Guna, A., Page, K. R., Replogle, J. M., Esantsi, T. K., Wang, M. L., Weissman, J. S., & Voorhees, R. M. (2023). A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *BMC Genomics*, *24*(1), 651. https://doi.org/10.1186/s12864-023-09754-y

Guna, A., Stevens, T. A., Inglis, A. J., Replogle, J. M., Esantsi, T. K., Muthukumar, G., Shaffer, K. C. L., Wang, M. L., Pogson, A. N., Jones, J. J., Lomenick, B., Chou, T.-F., Weissman, J. S., & Voorhees, R. M. (2022). MTCH2 is a mitochondrial outer membrane protein insertase. *Science*, *378*(6617), 317–322. https://doi.org/10.1126/science.add1856

Guna, A., Volkmar, N., Christianson, J. C., & Hegde, R. S. (2018). The ER membrane protein complex is a transmembrane domain insertase. *Science (New York, N.Y.)*, *359*(6374), 470–473. https://doi.org/10.1126/science.aao3099

Güngör, B., Flohr, T., Garg, S. G., & Herrmann, J. M. (2022). The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria. *PLoS Biology*, *20*(3), e3001380. https://doi.org/10.1371/journal.pbio.3001380

Habeck, G., Ebner, F. A., Shimada-Kreft, H., & Kreft, S. G. (2015). The yeast ERAD-C ubiquitin ligase Doa10 recognizes an intramembrane degron. *Journal of Cell Biology*, *209*(2), 261–273. https://doi.org/10.1083/jcb.201408088

Halic, M., & Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. *Current Opinion in Structural Biology*, *15*(1), 116–125. https://doi.org/10.1016/j.sbi.2005.01.013

Hampton, R. Y., Gardner, R. G., & Rine, J. (1996). Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. *Molecular Biology of the Cell*, *7*(12), 2029–2044. https://doi.org/10.1091/mbc.7.12.2029

Hansen, J. N., Rassmann, S., Stüven, B., Jurisch-Yaksi, N., & Wachten, D. (2021). CiliaQ: A simple, open-source software for automated quantification of ciliary morphology and fluorescence in 2D, 3D, and 4D images. *The European Physical Journal. E, Soft Matter*, *44*(2), 18. https://doi.org/10.1140/epje/s10189-021-00031-y

Harper, J. W., & Bennett, E. J. (2016). Proteome complexity and the forces that drive proteome imbalance. *Nature*, *537*(7620), 328–338. https://doi.org/10.1038/nature19947

Harrison, S. C. (2008). Viral membrane fusion. *Nature Structural & Molecular Biology*, *15*(7), Article 7. https://doi.org/10.1038/nsmb.1456

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. *European Journal of Biochemistry*, *214*(2), 375–381. https://doi.org/10.1111/j.1432-1033.1993.tb17933.x

Hassink, G., Kikkert, M., Voorden, S. van, Lee, S.-J., Spaapen, R., Laar, T. van, Coleman, C. S., Bartee, E., Früh, K., Chau, V., & Wiertz, E. (2005). TEB4 is a C4HC3 RING finger-containing ubiquitin ligase of the endoplasmic reticulum. *Biochemical Journal*, *388*(2), 647–655. https://doi.org/10.1042/BJ20041241

Hegde, R. S., & Keenan, R. J. (2022). The mechanisms of integral membrane protein biogenesis. *Nature Reviews Molecular Cell Biology*, *23*(2), Article 2. https://doi.org/10.1038/s41580-021-00413-2

Hegde, R. S., & Ploegh, H. L. (2010). Quality and quantity control at the endoplasmic reticulum. *Current Opinion in Cell Biology*, *22*(4), 437–446. https://doi.org/10.1016/j.ceb.2010.05.005

Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S. H., & von Heijne, G. (2007). Molecular code for

transmembrane-helix recognition by the Sec61 translocon. *Nature*, *450*(7172), 1026–1030. https://doi.org/10.1038/nature06387

Horlbeck, M. A., Gilbert, L. A., Villalta, J. E., Adamson, B., Pak, R. A., Chen, Y., Fields, A. P., Park, C. Y., Corn, J. E., Kampmann, M., & Weissman, J. S. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife*, *5*, e19760. https://doi.org/10.7554/eLife.19760

Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M. P., Parzen, H., Szpyt, J., Tam, S., Zarraga, G., Pontano-Vaites, L., Swarup, S., White, A. E., Schweppe, D. K., Rad, R., Erickson, B. K., … Harper, J. W. (2017). Architecture of the human interactome defines protein communities and disease networks. *Nature*, *545*(7655), 505–509. https://doi.org/10.1038/nature22366

Huyer, G., Piluek, W. F., Fansler, Z., Kreft, S. G., Hochstrasser, M., Brodsky, J. L., & Michaelis, S. (2004). Distinct Machinery Is Required in *Saccharomyces cerevisiae* for the Endoplasmic Reticulum-associated Degradation of a Multispanning Membrane Protein and a Soluble Luminal Protein*. *Journal of Biological Chemistry*, *279*(37), 38369–38378. https://doi.org/10.1074/jbc.M402468200

Inglis, A. J., Page, K. R., Guna, A., & Voorhees, R. M. (2020). Differential Modes of Orphan Subunit Recognition for the WRB/CAML Complex. *Cell Reports*, *30*(11), 3691-3698.e5. https://doi.org/10.1016/j.celrep.2020.02.084

Itakura, E., Zavodszky, E., Shao, S., Wohlever, M. L., Keenan, R. J., & Hegde, R. S. (2016). Ubiquilins Chaperone and Triage Mitochondrial Membrane Proteins for Degradation. *Molecular Cell*, *63*(1), 21–33. https://doi.org/10.1016/j.molcel.2016.05.020

Jami-Alahmadi, Y., Pandey, V., Mayank, A. K., & Wohlschlegel, J. A. (2021). A Robust Method for Packing High Resolution C18 RP-nano-HPLC Columns. *Journal of Visualized Experiments: JoVE*, *171*. https://doi.org/10.3791/62380

Jaskolowski, M., Jomaa, A., Gamerdinger, M., Shrestha, S., Leibundgut, M., Deuerling, E., & Ban, N. (2023). Molecular basis of the TRAP complex function in ER

protein biogenesis. *Nature Structural & Molecular Biology*, *30*(6), Article 6. https://doi.org/10.1038/s41594-023-00990-0

Ji, J., Cui, M.-K., Zou, R., Wu, M.-Z., Ge, M.-X., Li, J., & Zhang, Z.-R. (2024). An ATP13A1-assisted topogenesis pathway for folding multi-spanning membrane proteins. *Molecular Cell*, *84*(10), 1917-1931.e15. https://doi.org/10.1016/j.molcel.2024.04.010

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 (New York, N.Y.)*, *323*(5922), 1693–1697. https://doi.org/10.1126/science.1167983

Jost, M., Chen, Y., Gilbert, L. A., Horlbeck, M. A., Krenning, L., Menchon, G., Rai, A., Cho, M. Y., Stern, J. J., Prota, A. E., Kampmann, M., Akhmanova, A., Steinmetz, M. O., Tanenbaum, M. E., & Weissman, J. S. (2017). Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Molecular Cell*, *68*(1), 210-223.e6. https://doi.org/10.1016/j.molcel.2017.09.012

Juszkiewicz, S., & Hegde, R. S. (2018). Quality Control of Orphaned Proteins. *Molecular Cell*, *71*(3), 443–457. https://doi.org/10.1016/j.molcel.2018.07.001

Kalies, K.-U., Rapoport, T. A., & Hartmann, E. (1998). The β Subunit of the Sec61 Complex Facilitates Cotranslational Protein Transport and Interacts with the Signal Peptidase during Translocation. *Journal of Cell Biology*, *141*(4), 887–894. https://doi.org/10.1083/jcb.141.4.887

Kamiyama, D., Sekine, S., Barsi-Rhyne, B., Hu, J., Chen, B., Gilbert, L. A., Ishikawa, H., Leonetti, M. D., Marshall, W. F., Weissman, J. S., & Huang, B. (2016). Versatile protein tagging in cells with split fluorescent protein. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms11046

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*(2), 181–191. https://doi.org/10.1016/S0092-8674(00)81418-X

Kelleher, D. J., & Gilmore, R. (2006). An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*, *16*(4), 47R-62R. https://doi.org/10.1093/glycob/cwj066

Kern, D. M., Sorum, B., Mali, S. S., Hoel, C. M., Sridharan, S., Remis, J. P., Toso, D. B., Kotecha, A., Bautista, D. M., & Brohawn, S. G. (2021). Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. *Nature Structural & Molecular Biology*, *28*(7), Article 7. https://doi.org/10.1038/s41594-021-00619-0

Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V., & Wiertz, E. (2004). Human HRD1 is an E3 ubiquitin ligase involved in degradation of proteins from the endoplasmic reticulum. *The Journal of Biological Chemistry*, *279*(5), 3525–3534. https://doi.org/10.1074/jbc.M307453200

Kreft, S. G., Wang, L., & Hochstrasser, M. (2006). Membrane Topology of the Yeast Endoplasmic Reticulum-localized Ubiquitin Ligase Doa10 and Comparison with Its Human Ortholog TEB4 (MARCH-VI)*. *Journal of Biological Chemistry*, *281*(8), 4646–4653. https://doi.org/10.1074/jbc.M512215200

Kumazaki, K., Chiba, S., Takemoto, M., Furukawa, A., Nishiyama, K., Sugano, Y., Mori, T., Dohmae, N., Hirata, K., Nakada-Nakura, Y., Maturana, A. D., Tanaka, Y., Mori, H., Sugita, Y., Arisaka, F., Ito, K., Ishitani, R., Tsukazaki, T., & Nureki, O. (2014). Structural basis of Sec-independent membrane protein insertion by YidC. *Nature*, *509*(7501), Article 7501. https://doi.org/10.1038/nature13167

Kutay, U., Hartmann, E., & Rapoport, T. A. (1993). A class of membrane proteins with a C-terminal anchor. *Trends in Cell Biology*, *3*(3), 72–75. https://doi.org/10.1016/0962-8924(93)90066-A

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

Lenard, J. (2008). Viral Membranes. *Encyclopedia of Virology*, 308–314. https://doi.org/10.1016/B978-012374410-4.00530-6

Leto, D. E., Morgens, D. W., Zhang, L., Walczak, C. P., Elias, J. E., Bassik, M. C., & Kopito, R. R. (2019). Genome-wide CRISPR Analysis Identifies Substrate-

Specific Conjugation Modules in ER-Associated Degradation. *Molecular Cell*, *73*(2), 377-389.e11. https://doi.org/10.1016/j.molcel.2018.11.015

Lewis, A. J. O., & Hegde, R. S. (2021). A unified evolutionary origin for the ubiquitous protein transporters SecY and YidC. *BMC Biology*, *19*(1), 266. https://doi.org/10.1186/s12915-021-01171-5

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*, *75*(10), Article 10. https://doi.org/10.1107/S2059798319011471

Lilley, B. N., & Ploegh, H. L. (2004). A membrane protein required for dislocation of misfolded proteins from the ER. *Nature*, *429*(6994), 834–840. https://doi.org/10.1038/nature02592

Liu, L., Spurrier, J., Butt, T. R., & Strickler, J. E. (2008). Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expression and Purification*, *62*(1), 21–28. https://doi.org/10.1016/j.pep.2008.07.010

Madsen, L., Seeger, M., Semple, C. A., & Hartmann-Petersen, R. (2009). New ATPase regulators—P97 goes to the PUB. *The International Journal of Biochemistry & Cell Biology*, *41*(12), 2380–2388. https://doi.org/10.1016/j.biocel.2009.05.017

Manchado, E., Guillamot, M., & Malumbres, M. (2012). Killing cells by targeting mitosis. *Cell Death and Differentiation*, *19*(3), 369–377. https://doi.org/10.1038/cdd.2011.197

Mariappan, M., Mateja, A., Dobosz, M., Bove, E., Hegde, R. S., & Keenan, R. J. (2011). The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature*, *477*(7362), 61–66. https://doi.org/10.1038/nature10362

Marsh, J. A., & Teichmann, S. A. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. *Annual Review of Biochemistry*, *84*(Volume 84, 2015), 551–575. https://doi.org/10.1146/annurev-biochem-060614-034142

Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51. https://doi.org/10.1016/j.jsb.2005.07.007

Mateja, A., Szlachcic, A., Downing, M. E., Dobosz, M., Mariappan, M., Hegde, R. S., & Keenan, R. J. (2009). The structural basis of tail-anchored membrane protein recognition by Get3. *Nature*, *461*(7262), Article 7262. https://doi.org/10.1038/nature08319

McCracken, A. A., & Brodsky, J. L. (1996). Assembly of ER-associated protein degradation in vitro: Dependence on cytosol, calnexin, and ATP. *Journal of Cell Biology*, *132*(3), 291–298. https://doi.org/10.1083/jcb.132.3.291

McDowell, M. A., Heimes, M., Enkavi, G., Farkas, Á., Saar, D., Wild, K., Schwappach, B., Vattulainen, I., & Sinning, I. (2023). The GET insertase exhibits conformational plasticity and induces membrane thinning. *Nature Communications*, *14*(1), 7355. https://doi.org/10.1038/s41467-023-42867-2

McDowell, M. A., Heimes, M., Fiorentino, F., Mehmood, S., Farkas, Á., Coy-Vergara, J., Wu, D., Bolla, J. R., Schmid, V., Heinze, R., Wild, K., Flemming, D., Pfeffer, S., Schwappach, B., Robinson, C. V., & Sinning, I. (2020). Structural Basis of Tail-Anchored Membrane Protein Biogenesis by the GET Insertase Complex. *Molecular Cell*, *80*(1), 72-86.e7. https://doi.org/10.1016/j.molcel.2020.08.012

McGilvray, P. T., Anghel, S. A., Sundaram, A., Zhong, F., Trnka, M. J., Fuller, J. R., Hu, H., Burlingame, A. L., & Keenan, R. J. (2020). An ER translocon for multi-pass membrane protein biogenesis. *eLife*, *9*, e56889. https://doi.org/10.7554/eLife.56889

Meacock, S. L., Lecomte, F. J. L., Crawshaw, S. G., & High, S. (2002). Different Transmembrane Domains Associate with Distinct Endoplasmic Reticulum Components during Membrane Integration of a Polytopic Protein. *Molecular Biology of the Cell*, *13*(12), 4114–4129. https://doi.org/10.1091/mbc.e02-04-0198

Meyer, H., Bug, M., & Bremer, S. (2012). Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nature Cell Biology*, *14*(2), Article 2. https://doi.org/10.1038/ncb2407

Miller-Vedam, L. E., Bräuning, B., Popova, K. D., Schirle Oakdale, N. T., Bonnar, J. L., Prabu, J. R., Boydston, E. A., Sevillano, N., Shurtleff, M. J., Stroud, R. M., Craik, C. S., Schulman, B. A., Frost, A., & Weissman, J. S. (2020). Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *eLife*, *9*, e62611. https://doi.org/10.7554/eLife.62611

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, *19*(6), Article 6. https://doi.org/10.1038/s41592-022-01488-1

Munro, S., & Pelham, H. R. B. (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, *48*(5), 899–907. https://doi.org/10.1016/0092-8674(87)90086-9

Natarajan, N., Foresti, O., Wendrich, K., Stein, A., & Carvalho, P. (2020). Quality Control of Protein Complex Assembly by a Transmembrane Recognition Factor. *Molecular Cell*, *77*(1), 108-119.e9. https://doi.org/10.1016/j.molcel.2019.10.003

Neal, S., Jaeger, P. A., Duttke, S. H., Benner, C., K.Glass, C., Ideker, T., & Hampton, R. Y. (2018). The Dfm1 Derlin Is Required for ERAD Retrotranslocation of Integral Membrane Proteins. *Molecular Cell*, *69*(2), 306-320.e4. https://doi.org/10.1016/j.molcel.2017.12.012

Ng, D. P., Poulsen, B. E., & Deber, C. M. (2012). Membrane protein misassembly in disease. *Biochimica Et Biophysica Acta*, *1818*(4), 1115–1122. https://doi.org/10.1016/j.bbamem.2011.07.046

O'Donnell, J. P., Phillips, B. P., Yagita, Y., Juszkiewicz, S., Wagner, A., Malinverni, D., Keenan, R. J., Miller, E. A., & Hegde, R. S. (2020). The architecture of EMC reveals a path for membrane protein insertion. *eLife*, *9*, e57887. https://doi.org/10.7554/eLife.57887

Olzmann, J. A., Kopito, R. R., & Christianson, J. C. (2013). The Mammalian Endoplasmic Reticulum-Associated Degradation System. *Cold Spring Harbor Perspectives in Biology*, *5*(9), a013185. https://doi.org/10.1101/cshperspect.a013185

Olzmann, J. A., Richter, C. M., & Kopito, R. R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proceedings of the*

National Academy of Sciences, *110*(4), 1345–1350. https://doi.org/10.1073/pnas.1213738110

Paatero, A. O., Kellosalo, J., Dunyak, B. M., Almaliti, J., Gestwicki, J. E., Gerwick, W. H., Taunton, J., & Paavilainen, V. O. (2016). Apratoxin Kills Cells by Direct Blockade of the Sec61 Protein Translocation Channel. *Cell Chemical Biology*, *23*(5), 561–566. https://doi.org/10.1016/j.chembiol.2016.04.008

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), e45. https://doi.org/10.1093/nar/29.9.e45

Pleiner, T., Bates, M., & Görlich, D. (2017). A toolbox of anti–mouse and anti–rabbit IgG secondary nanobodies. *Journal of Cell Biology*, *217*(3), 1143–1154. https://doi.org/10.1083/jcb.201709115

Pleiner, T., Hazu, M., Pinton Tomaleri, G., Nguyen, V. N., Januszyk, K., & Voorhees, R. M. (2023). A selectivity filter in the ER membrane protein complex limits protein misinsertion at the ER. *The Journal of Cell Biology*, *222*(8), e202212007. https://doi.org/10.1083/jcb.202212007

Pleiner, T., Hazu, M., Tomaleri, G. P., Januszyk, K., Oania, R. S., Sweredoski, M. J., Moradian, A., Guna, A., & Voorhees, R. M. (2021). WNK1 is an assembly factor for the human ER membrane protein complex. *Molecular Cell*, *81*(13), 2693-2704.e12. https://doi.org/10.1016/j.molcel.2021.04.013

Pleiner, T., Tomaleri, G. P., Januszyk, K., Inglis, A. J., Hazu, M., & Voorhees, R. M. (2020). Structural basis for membrane insertion by the human ER membrane

protein complex. *Science*, *369*(6502), 433–436. https://doi.org/10.1126/science.abb5008

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), Article 3. https://doi.org/10.1038/nmeth.4169

Radhakrishnan, P., Nayak, S. S., Shukla, A., Lindstrand, A., & Girisha, K. M. (2019). Meckel syndrome: Clinical and mutation profile in six fetuses. *Clinical Genetics*, *96*(6), 560–565. https://doi.org/10.1111/cge.13623

Rao, M., Okreglak, V., Chio, U. S., Cho, H., Walter, P., & Shan, S. (2016). Multiple selection filters ensure accurate tail-anchored membrane protein targeting. *eLife*, *5*, e21301. https://doi.org/10.7554/eLife.21301

Rapoport, T. A., Jungnickel, B., & Kutay, U. (1996). Protein Transport Across the Eukaryotic Endoplasmic Reticulum and Bacterial Inner Membranes. *Annual Review of Biochemistry*, *65*(1), 271–303. https://doi.org/10.1146/annurev.bi.65.070196.001415

Rapoport, T. A., Li, L., & Park, E. (2017). Structural and Mechanistic Insights into Protein Translocation. *Annual Review of Cell and Developmental Biology*, *33*, 369–390. https://doi.org/10.1146/annurev-cellbio-100616-060439

Replogle, J. M., Norman, T. M., Xu, A., Hussmann, J. A., Chen, J., Cogan, J. Z., Meer, E. J., Terry, J. M., Riordan, D. P., Srinivas, N., Fiddes, I. T., Arthur, J. G., Alvarado, L. J., Pfeiffer, K. A., Mikkelsen, T. S., Weissman, J. S., & Adamson, B. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nature Biotechnology*, *38*(8), Article 8. https://doi.org/10.1038/s41587-020-0470-y

Ridnõi, K., Šois, M., Vaidla, E., Pajusalu, S., Kelder, L., Reimand, T., & Õunap, K. (2019). A prenatally diagnosed case of Meckel–Gruber syndrome with novel compound heterozygous pathogenic variants in the TXNDC15 gene. *Molecular Genetics & Genomic Medicine*, *7*(5), e614. https://doi.org/10.1002/mgg3.614

Rivera-Monroy, J., Musiol, L., Unthan-Fechner, K., Farkas, Á., Clancy, A., Coy-Vergara, J., Weill, U., Gockel, S., Lin, S.-Y., Corey, D. P., Kohl, T., Ströbel, P., Schuldiner, M., Schwappach, B., & Vilardi, F. (2016). Mice lacking WRB reveal

differential biogenesis requirements of tail-anchored proteins in vivo. *Scientific Reports*, *6*(1), 39464. https://doi.org/10.1038/srep39464

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 *. *Journal of Biological Chemistry*, *288*(23), 16295–16307. https://doi.org/10.1074/jbc.M112.446583

Sanders, C. R., & Myers, J. K. (2004). Disease-related misassembly of membrane proteins. *Annual Review of Biophysics and Biomolecular Structure*, *33*, 25–51. https://doi.org/10.1146/annurev.biophys.33.110502.140348

Sanders, C. R., & Nagy, J. K. (2000). Misfolding of membrane proteins in health and disease: The lady or the tiger? *Current Opinion in Structural Biology*, *10*(4), 438–442. https://doi.org/10.1016/S0959-440X(00)00112-3

Satoh, T., Ohba, A., Liu, Z., Inagaki, T., & Satoh, A. K. (2015). dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. *eLife*, *4*, e06306. https://doi.org/10.7554/eLife.06306

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schmidt, C. C., Vasic, V., & Stein, A. (2020). Doa10 is a membrane protein retrotranslocase in ER-associated protein degradation. *eLife*, *9*, e56945. https://doi.org/10.7554/eLife.56945

Schorr, S., Nguyen, D., Haßdenteufel, S., Nagaraj, N., Cavalié, A., Greiner, M., Weissgerber, P., Loi, M., Paton, A. W., Paton, J. C., Molinari, M., Förster, F., Dudek, J., Lang, S., Helms, V., & Zimmermann, R. (2020). Identification of signal peptide features for substrate specificity in human Sec62/Sec63-dependent ER protein import. *The FEBS Journal*, *287*(21), 4612–4640. https://doi.org/10.1111/febs.15274

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J. F., Weissman, J. S., & Krogan, N. J. (2005). Exploration of the Function and Organization of the Yeast Early Secretory Pathway through an Epistatic Miniarray Profile. *Cell*, *123*(3), 507–519. https://doi.org/10.1016/j.cell.2005.08.031

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H. D., Schwappach, B., & Weissman, J. S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell*, *134*(4), 634–645. https://doi.org/10.1016/j.cell.2008.06.025

Schulman, B. A., & Wade Harper, J. (2009). Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nature Reviews Molecular Cell Biology*, *10*(5), 319–331. https://doi.org/10.1038/nrm2673

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342. https://doi.org/10.1038/nature10098

Shah, S., Lee, S.-F., Tabuchi, K., Hao, Y.-H., Yu, C., LaPlant, Q., Ball, H., Dann, C. E., Südhof, T., & Yu, G. (2005). Nicastrin Functions as a γ-Secretase-Substrate Receptor. *Cell*, *122*(3), 435–447. https://doi.org/10.1016/j.cell.2005.05.022

Shaheen, R., Szymanska, K., Basu, B., Patel, N., Ewida, N., Faqeih, E., Al Hashem, A., Derar, N., Alsharif, H., Aldahmesh, M. A., Alazami, A. M., Hashem, M., Ibrahim, N., Abdulwahab, F. M., Sonbul, R., Alkuraya, H., Alnemer, M., Al Tala, S., Al-Husain, M., … Ciliopathy WorkingGroup. (2016). Characterizing the morbid genome of ciliopathies. *Genome Biology*, *17*(1), 242. https://doi.org/10.1186/s13059-016-1099-5

Shan, S., & Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. *FEBS Letters*, *579*(4), 921–926. https://doi.org/10.1016/j.febslet.2004.11.049

Shao, S., & Hegde, R. S. (2011). Membrane Protein Insertion at the Endoplasmic Reticulum. *Annual Review of Cell and Developmental Biology*, *27*(1), 25–56. https://doi.org/10.1146/annurev-cellbio-092910-154125

Shao, S., Rodrigo-Brenni, M. C., Kivlen, M. H., & Hegde, R. S. (2017). Mechanistic basis for a molecular triage reaction. *Science (New York, N.Y.)*, *355*(6322), 298–302. https://doi.org/10.1126/science.aah6130

Sharma, A., Mariappan, M., Appathurai, S., & Hegde, R. S. (2010). In vitro dissection of protein translocation into the mammalian endoplasmic reticulum. *Methods in Molecular Biology (Clifton, N.J.)*, *619*, 339–363. https://doi.org/10.1007/978-1-60327-412-8_20

Shurtleff, M. J., Itzhak, D. N., Hussmann, J. A., Schirle Oakdale, N. T., Costa, E. A., Jonikas, M., Weibezahn, J., Popova, K. D., Jan, C. H., Sinitcyn, P., Vembar, S. S., Hernandez, H., Cox, J., Burlingame, A. L., Brodsky, J. L., Frost, A., Borner, G. H., & Weissman, J. S. (2018). The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife*, *7*, e37018. https://doi.org/10.7554/eLife.37018

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., & Harper, J. W. (1997). F-Box Proteins Are Receptors that Recruit Phosphorylated Substrates to the SCF Ubiquitin-Ligase Complex. *Cell*, *91*(2), 209–219. https://doi.org/10.1016/S0092-8674(00)80403-1

Smalinskaitė, L., Kim, M. K., Lewis, A. J. O., Keenan, R. J., & Hegde, R. S. (2022). Mechanism of an intramembrane chaperone for multipass membrane proteins. *Nature*, *611*(7934), Article 7934. https://doi.org/10.1038/s41586-022-05336-2

Smith, U. M., Consugar, M., Tee, L. J., McKee, B. M., Maina, E. N., Whelan, S., Morgan, N. V., Goranson, E., Gissen, P., Lilliquist, S., Aligianis, I. A., Ward, C. J., Pasha, S., Punyashthiti, R., Malik Sharif, S., Batman, P. A., Bennett, C. P., Woods, C. G., McKeown, C., … Johnson, C. A. (2006). The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat. *Nature Genetics*, *38*(2), 191–196. https://doi.org/10.1038/ng1713

Spandidos, A., Wang, X., Wang, H., Dragnev, S., Thurber, T., & Seed, B. (2008). A comprehensive collection of experimentally validated primers for Polymerase Chain Reaction quantitation of murine transcript abundance. *BMC Genomics*, *9*(1), 633. https://doi.org/10.1186/1471-2164-9-633

Spandidos, A., Wang, X., Wang, H., & Seed, B. (2010). PrimerBank: A resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Research*, *38*(Database issue), D792-799. https://doi.org/10.1093/nar/gkp1005

Stefanovic, S., & Hegde, R. S. (2007). Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell*, *128*(6), 1147–1159. https://doi.org/10.1016/j.cell.2007.01.036

Stein, A., Ruggiano, A., Carvalho, P., & Rapoport, T. A. (2014). Key Steps in ERAD of Luminal ER Proteins Reconstituted with Purified Components. *Cell*, *158*(6), 1375–1388. https://doi.org/10.1016/j.cell.2014.07.050

Stevens, T. A., Tomaleri, G. P., Hazu, M., Wei, S., Nguyen, V. N., DeKalb, C., Voorhees, R. M., & Pleiner, T. (2023). A nanobody-based strategy for rapid and scalable purification of human protein complexes. *Nature Protocols*, 1–32. https://doi.org/10.1038/s41596-023-00904-w

Sun, S., & Mariappan, M. (2020). C-terminal tail length guides insertion and assembly of membrane proteins. *Journal of Biological Chemistry*, *295*(46), 15498–15510. https://doi.org/10.1074/jbc.RA120.012992

Sundaram, A., Yamsek, M., Zhong, F., Hooda, Y., Hegde, R. S., & Keenan, R. J. (2022). Substrate-driven assembly of a translocon for multipass membrane proteins. *Nature*. https://doi.org/10.1038/s41586-022-05330-8

Sung, M.-K., Porras-Yakushi, T. R., Reitsma, J. M., Huber, F. M., Sweredoski, M. J., Hoelz, A., Hess, S., & Deshaies, R. J. (2016). A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. *eLife*, *5*, e19105. https://doi.org/10.7554/eLife.19105

Swanson, R., Locher, M., & Hochstrasser, M. (2001). A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Matα2 repressor degradation. *Genes & Development*, *15*(20), 2660–2674. https://doi.org/10.1101/gad.933301

Thornburg, C. C., Cowley, E. S., Sikorska, J., Shaala, L. A., Ishmael, J. E., Youssef, D. T. A., & McPhail, K. L. (2013). Apratoxin H and apratoxin A sulfoxide from the

Red Sea cyanobacterium Moorea producens. *Journal of Natural Products*, *76*(9), 1781–1788. https://doi.org/10.1021/np4004992

Tian, S., Wu, Q., Zhou, B., Choi, M. Y., Ding, B., Yang, W., & Dong, M. (2019). Proteomic Analysis Identifies Membrane Proteins Dependent on the ER Membrane Protein Complex. *Cell Reports*, *28*(10), 2517-2526.e5. https://doi.org/10.1016/j.celrep.2019.08.006

Timms, R. T., Menzies, S. A., Tchasovnikarova, I. A., Christensen, L. C., Williamson, J. C., Antrobus, R., Dougan, G., Ellgaard, L., & Lehner, P. J. (2016). Genetic dissection of mammalian ERAD through comparative haploid and CRISPR forward genetic screens. *Nature Communications*, *7*(1), 11786. https://doi.org/10.1038/ncomms11786

The UniProt Consortium. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099

Van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., & Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature*, *427*(6969), 36–44. https://doi.org/10.1038/nature02218

Vashist, S., & Ng, D. T. W. (2004). Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control. *The Journal of Cell Biology*, *165*(1), 41–52. https://doi.org/10.1083/jcb.200309132

Vembar, S. S., & Brodsky, J. L. (2008). One step at a time: Endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology*, *9*(12), Article 12. https://doi.org/10.1038/nrm2546

Vilardi, F., Lorenz, H., & Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. *Journal of Cell Science*, *124*(8), 1301–1307. https://doi.org/10.1242/jcs.084277

Vilardi, F., Stephan, M., Clancy, A., Janshoff, A., & Schwappach, B. (2014). WRB and CAML are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane. *PloS One*, *9*(1), e85033. https://doi.org/10.1371/journal.pone.0085033

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. *Journal of Cell Biology*, *134*(1), 25–35. https://doi.org/10.1083/jcb.134.1.25

Volkmar, N., Thezenas, M.-L., Louie, S. M., Juszkiewicz, S., Nomura, D. K., Hegde, R. S., Kessler, B. M., & Christianson, J. C. (2019). The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis. *Journal of Cell Science*, *132*(2). https://doi.org/10.1242/jcs.223453

von Heijne, G. (1996). Principles of membrane protein assembly and structure. *Progress in Biophysics and Molecular Biology*, *66*(2), 113–139. https://doi.org/10.1016/S0079-6107(97)85627-1

von Heijne, G. (2007). The membrane protein universe: What's out there and why bother? *Journal of Internal Medicine*, *261*(6), 543–557. https://doi.org/10.1111/j.1365-2796.2007.01792.x

Voorhees, R. M., & Hegde, R. S. (2016). Structure of the Sec61 channel opened by a signal sequence. *Science*, *351*(6268), 88–91. https://doi.org/10.1126/science.aad4992

Wallin, E., & von Heijne, G. (1995). Properties of N-terminal tails in G-protein coupled receptors: A statistical study. *Protein Engineering*, *8*(7), 693–698. https://doi.org/10.1093/protein/8.7.693

Wallin, E., & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, *7*(4), 1029–1038. https://doi.org/10.1002/pro.5560070420

Walter, P., & Blobel, G. (1983). Preparation of microsomal membranes for cotranslational protein translocation. *Methods in Enzymology*, *96*, 84–93. https://doi.org/10.1016/s0076-6879(83)96010-x

Wang, F., Brown, E. C., Mak, G., Zhuang, J., & Denic, V. (2010). A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Molecular Cell*, *40*(1), 159–171. https://doi.org/10.1016/j.molcel.2010.08.038

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), 441–444. https://doi.org/10.1038/nature13471

Wang, X., & Seed, B. (2003). A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Research*, *31*(24), e154. https://doi.org/10.1093/nar/gng154

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Weber, A., Cohen, I., Popp, O., Dittmar, G., Reiss, Y., Sommer, T., Ravid, T., & Jarosch, E. (2016). Sequential Poly-ubiquitylation by Specialized Conjugating Enzymes Expands the Versatility of a Quality Control Ubiquitin Ligase. *Molecular Cell*, *63*(5), 827–839. https://doi.org/10.1016/j.molcel.2016.07.020

Werner, E. D., Brodsky, J. L., & McCracken, A. A. (1996). Proteasome-dependent endoplasmic reticulum-associated protein degradation: An unconventional route to a familiar fate. *Proceedings of the National Academy of Sciences*, *93*(24), 13797–13801. https://doi.org/10.1073/pnas.93.24.13797

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. https://doi.org/10.7554/eLife.03035

Wild, K., Halic, M., Sinning, I., & Beckmann, R. (2004). SRP meets the ribosome. *Nature Structural & Molecular Biology*, *11*(11), 1049–1053. https://doi.org/10.1038/nsmb853

Wu, H., & Hegde, R. S. (2023). Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Molecular Cell*, *83*(6), 961-973.e7. https://doi.org/10.1016/j.molcel.2023.01.018

Wu, X., Siggel, M., Ovchinnikov, S., Mi, W., Svetlov, V., Nudler, E., Liao, M., Hummer, G., & Rapoport, T. A. (2020). Structural basis of ER-associated protein

degradation mediated by the Hrd1 ubiquitin ligase complex. *Science*, *368*(6489). https://doi.org/10.1126/science.aaz2449

Yamamoto, Y., & Sakisaka, T. (2012). Molecular Machinery for Insertion of Tail-Anchored Membrane Proteins into the Endoplasmic Reticulum Membrane in Mammalian Cells. *Molecular Cell*, *48*(3), 387–397. https://doi.org/10.1016/j.molcel.2012.08.028

Yamazaki, S., Fujii, T., Chiba, S., Shin, H.-W., Nakayama, K., & Katoh, Y. (2024). TXNDC15, an ER-localized thioredoxin-like transmembrane protein, contributes to ciliary transition zone integrity. *Journal of Cell Science*, *137*(24), jcs262123. https://doi.org/10.1242/jcs.262123

Yanagitani, K., Juszkiewicz, S., & Hegde, R. S. (2017). UBE2O is a quality control factor for orphans of multi-protein complexes. *Science (New York, N.Y.)*, *357*(6350), 472–475. https://doi.org/10.1126/science.aan0178

Young, J. C., Barral, J. M., & Ulrich Hartl, F. (2003). More than folding: Localized functions of cytosolic chaperones. *Trends in Biochemical Sciences*, *28*(10), 541–547. https://doi.org/10.1016/j.tibs.2003.08.009

Yurtsev, E., & Friedman, J. (2015). *FlowCytometryTools* (Version zenodo) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.32992

Zhang, J.-P., Li, X.-L., Li, G.-H., Chen, W., Arakaki, C., Botimer, G. D., Baylink, D., Zhang, L., Wen, W., Fu, Y.-W., Xu, J., Chun, N., Yuan, W., Cheng, T., & Zhang, X.-B. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biology*, *18*(1), 35. https://doi.org/10.1186/s13059-017-1164-8

Zuiderweg, E. R. P., Hightower, L. E., & Gestwicki, J. E. (2017). The remarkable multivalency of the Hsp70 chaperones. *Cell Stress & Chaperones*, *22*(2), 173–189. https://doi.org/10.1007/s12192-017-0776-y