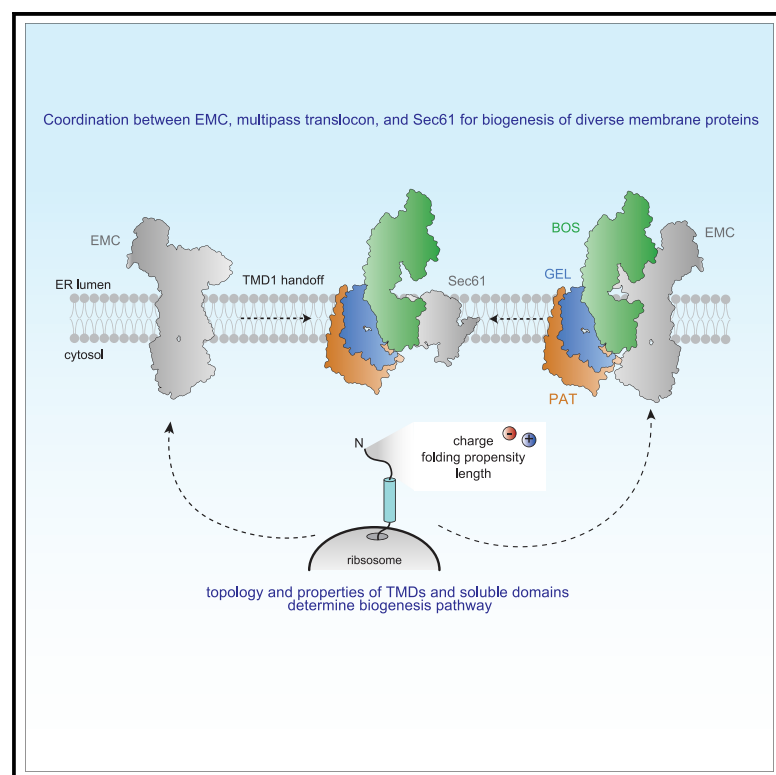


Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins

Graphical abstract



Authors

Katharine R. Page, Vy N. Nguyen, Tino Pleiner, ..., Ting-Yu Wang, Tsui-Fen Chou, Rebecca M. Voorhees

Correspondence

voorhees@caltech.edu

In brief

Page, Nguyen, Pleiner et al. have systematically studied the factors required for the integration and folding of diverse human membrane proteins. The authors report a central role for the EMC in organizing localized hubs that facilitate transfer of substrates between biogenesis pathways as dictated by their topology and biophysical properties.

Highlights

- Genome-wide screens for biogenesis of diverse membrane proteins
- Genetic modifier screen for interactors of the EMC that enable GPCR integration
- Biochemical and structural analysis of biogenesis hubs organized by the EMC
- Analysis of substrate properties that dictate biogenesis pathway



Article

Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins

Katharine R. Page,^{1,3,4} Vy N. Nguyen,^{1,3} Tino Pleiner,^{1,3,5} Giovanni Pinton Tomaleri,^{1,6} Maxine L. Wang,¹ Alina Guna,¹ Masami Hazu,¹ Ting-Yu Wang,¹ Tsui-Fen Chou,¹ and Rebecca M. Voorhees^{1,2,7,*}

¹Division of Biology and Biological Engineering, California Institute of Technology, 1200 E. California Ave., Pasadena, CA 91125, USA

²Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA

³These authors contributed equally

⁴Present address: Altos Labs, Redwood City, CA 94065, USA

⁵Present address: Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA

⁶Present address: Gate Bioscience, Brisbane, CA 94005, USA

⁷Lead contact

*Correspondence: voorhees@caltech.edu

<https://doi.org/10.1016/j.molcel.2024.08.005>

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, the multipass translocon, and Sec61 for the biogenesis of diverse membrane proteins in human cells.

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.^{1,2} In order to carry out these functions, both the soluble and transmembrane domains (TMDs) require distinct charge, hydrophobicity, and length.³ 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).^{4,5} For multipass proteins, the nascent polypeptide is first co-translationally captured in the cytosol by the signal recognition particle (SRP).^{6,7} Once at the ER, substrates must be inserted and folded into the lipid

bilayer. Insertion requires two simultaneous processes: (1) transfer of the hydrophobic TMD to a membrane-spanning topology within the lipid bilayer and (2) translocation of its associated soluble domain across the membrane into the ER lumen. The latter of these processes poses an energetic barrier to insertion and is typically catalyzed by a membrane protein insertase.⁸

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.^{9,10} However, because many multipass proteins contain TMDs that cannot autonomously open the lateral gate,^{11,12} 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.^{13–15} The multipass translocon is a dynamic, 8-subunit complex that includes the GEL (guided

entry of tail-anchored proteins [GET] and ER membrane protein complex [EMC]-like), BOS (back of Sec61), and PAT (protein associated with translocon) complexes (Figure S1A). The GEL complex—composed of the Oxa1 superfamily insertase TMCO1 (transmembrane and coiled-coil domains 1) and its binding partner OPT1 (obligate partner of TMCO1 insertase)—serves as the dedicated insertase of the multipass translocon. The PAT complex, containing Asterix and CCDC47, has two proposed roles: Asterix chaperones hydrophilic TMDs,^{16,17} while CCDC47 is proposed to engage the ribosome and guide TMDs to the multipass translocon.^{14,15} Finally, the function of the BOS complex remains unknown, but is thought to act as a 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 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.^{18–21} 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).²² Indeed, expression of rhodopsin, which does not rely on the lateral gate of Sec61 for insertion, is EMC dependent.²³ However, the function of all nine of EMC's subunits, particularly those that form its large luminal 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.^{24–27} 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.^{27–31} It is likely that multipass substrates are therefore directly delivered by SRP to the EMC,³² 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.

RESULTS

Systematic analysis of membrane protein biogenesis

With the goal of unbiasedly identifying factors required for the biogenesis of diverse membrane proteins, we selected a panel of substrates with distinct topologies, biophysical properties, and number of TMDs (Figures 1 and S1B). We included the hu-

man GPCR AGTR2, and the viral ORF3a and M from severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), as well as the post-translationally targeted TA protein, Sec61 β . 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 (Figures 1C and 1D) both adopt an identical three-TMD topology, but their TMDs have distinct biophysical properties and insertion propensities.^{34,35} Further, ORF3a and M have soluble N termini of different lengths (42 vs. 19 amino acids [aa]) and charges (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, we included the TA protein Sec61 β as a control (Figure 1B). TAs contain a single TMD within ~35 aa of their C terminus and thus cannot access the co-translational biogenesis pathways typically utilized by multipass proteins.^{8,36} The targeting and insertion of Sec61 β has been extensively characterized biochemically and is one of the few TA proteins known to rely equally on the EMC and GET biogenesis pathways. Therefore, the Sec61 β screen serves as a comparison for machinery required for biogenesis of multipass vs. singlepass membrane proteins.^{19,37,38}

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 (Figures 1 and S1C).^{39,40} 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, resulting in a decrease in fluorescence.^{19,27} Conversely, disruption of protein quality control leads to accumulation of 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 S1C). 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 S1D), we identified machinery that differentially affects their biogenesis and degradation (Figure 1; Tables S2 and S3). For example, among the identified quality-control factors, loss of the ER-resident E3 ubiquitin ligase HRD1 was found to stabilize only AGTR2, while the more general ER-associated degradation (ERAD) component FAF2 was identified in all four screens.^{41,42} On the biogenesis side, the GET pathway components had the most pronounced effect on the TA control substrate (Sec61 β), consistent with their role in post-translational insertion.^{43–46} 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 Sec61 β or M. Interestingly,

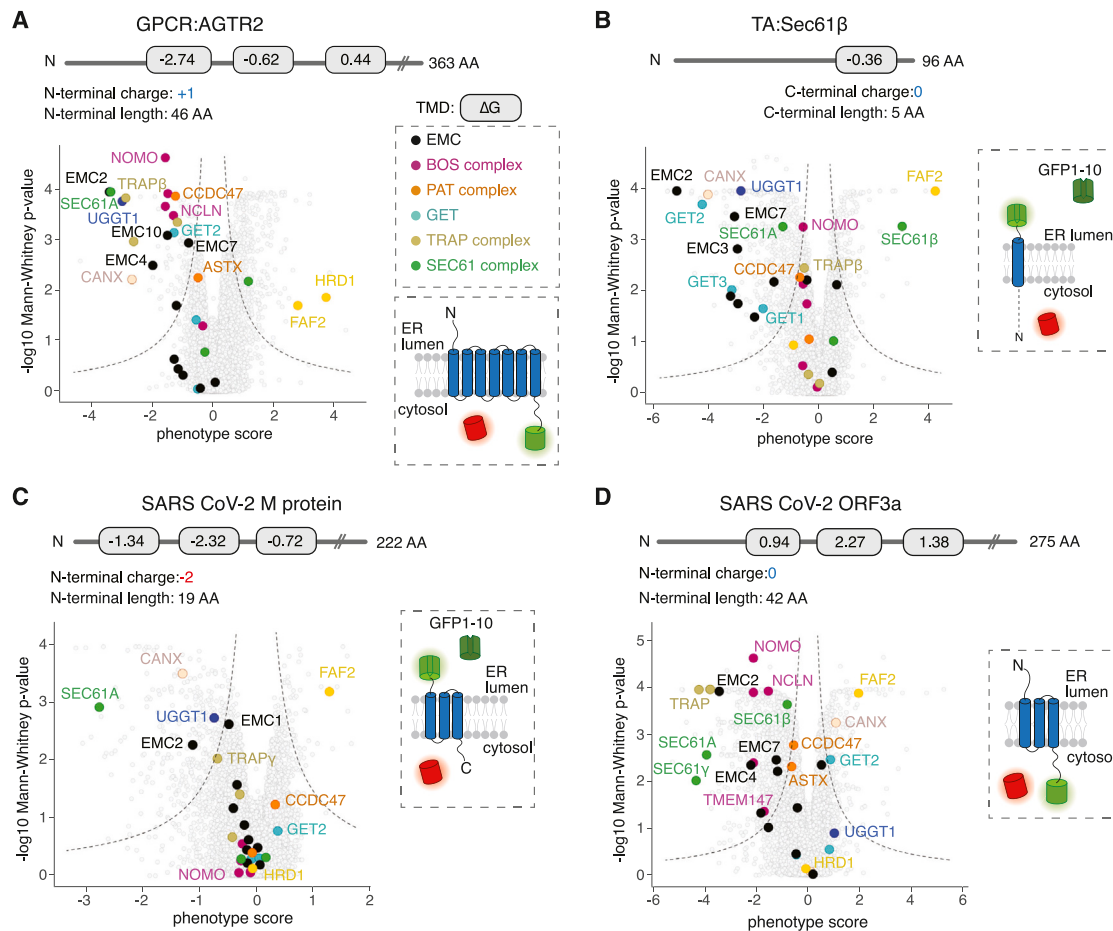


Figure 1. Genome-wide CRISPRi screens to systematically query biogenesis factors for diverse membrane proteins

(A) (Top) Schematic of the GPCR AGTR2. The gray rectangles indicate TMDs and their corresponding ΔG value.³³ (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 vs. Mann-Whitney p values from two independent replicates of a genome-wide CRISPRi screen AGTR2. Individual genes are displayed in gray, and specific factors that increase or decrease AGTR2 stability are highlighted and labeled. Genes that fall outside the dashed lines represent statistically significant hits.

(B) As in (A), for the tail-anchored protein Sec61 β . Here, The GFP11 sequence is appended to the C-terminal of Sec61 β . 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.

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 S1E). Although loss of SEC61A affected all four reporters, given the extensive data for Sec61-independent biogenesis of TA proteins, its effect on Sec61 β is likely due to an assembly rather than insertion defect. Therefore, one of the only universally identified biogenesis factors was the EMC, which is consistent with its established role in TA and N_{exo} TMD insertion.^{19,22}

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 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: TRAP α) and TA proteins (SQS, VAMP2, and Sec61 β). To allow for direct comparison, all reporters contained a full-length GFP, with the exception of Sec61 β , which required use of the split GFP approach to avoid mislocalization.^{47–49}

Using an arrayed screen, we tested the depletion of nine factors that represented each of the major biogenesis (EMC, BOS, GEL, PAT, TRAP, GET, and UGGT1) and quality-control (HRD1 and FAF2) complexes identified in the screens (Figure 2B).

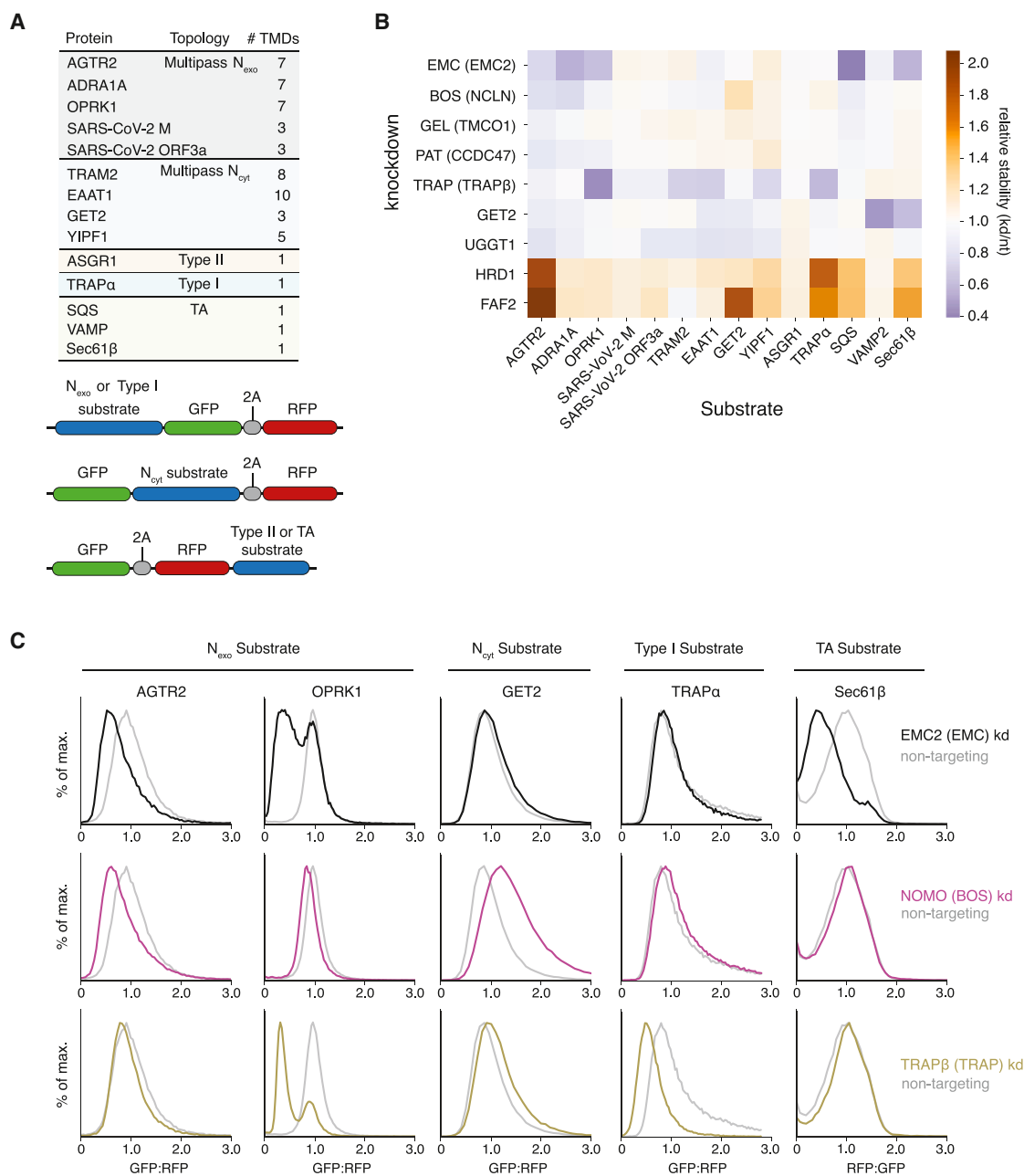


Figure 2. Differential effects of ER factors on membrane protein biogenesis

(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 1 on the reporters in (A). Results are displayed as a heatmap indicating relative stability of each reporter after gene knockdown compared with a non-targeting control.

(C) Representative analysis of individual data points in (B) displayed as a histogram.

Critically, in most cases, knockdown of a single subunit is sufficient to deplete the entire complex (Figure S1E).^{50–53} 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.^{54–57} 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 2C).^{5,58} Though an effect of TRAP β depletion was observed for the single-spanning protein TRAP α , this may be due to an assembly rather than a biogenesis defect. Although 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 2B) but not K562s, despite efficient knockdown (Figure S1E). 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 (Figures S1E and S1F).

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 complexes among 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 the 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),⁵⁹ biogenesis machinery (e.g., SRP receptor, glycosylation machinery, and components of the multipass translocon),^{7,13,60} and quality-control factors (e.g., p97)⁶¹ (Figure 3A; Table S4). 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 3B).⁴⁸ 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^{52,53}; and (2) a second randomized

guide, targeting all open reading frames using the CRISPRi-v2 library.⁴⁰ Transduction of this dual library allows the acute knock-down 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 (Figures 3C and S2; Table S1). 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 Sec61 β and the GPCR AGTR2 allowed us to delineate EMC co-factors that function to support its post- vs. co-translational biogenesis roles (Figures 3C and S2). 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 (Figures S2C and S2D). Finally, several quality-control factors, such as HRD1 and FAF2, exhibit EMC-independent effects, suggesting that 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 (Figures 3C–3F, S2A, and S2B). 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 (Figures 3D–3F). 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 (Figures 4A, 4B, and S3A). 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 GPCR 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,²² markedly rescues its dependence on the BOS complex for biogenesis (Figures 4C, S3D, and S3E). 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.

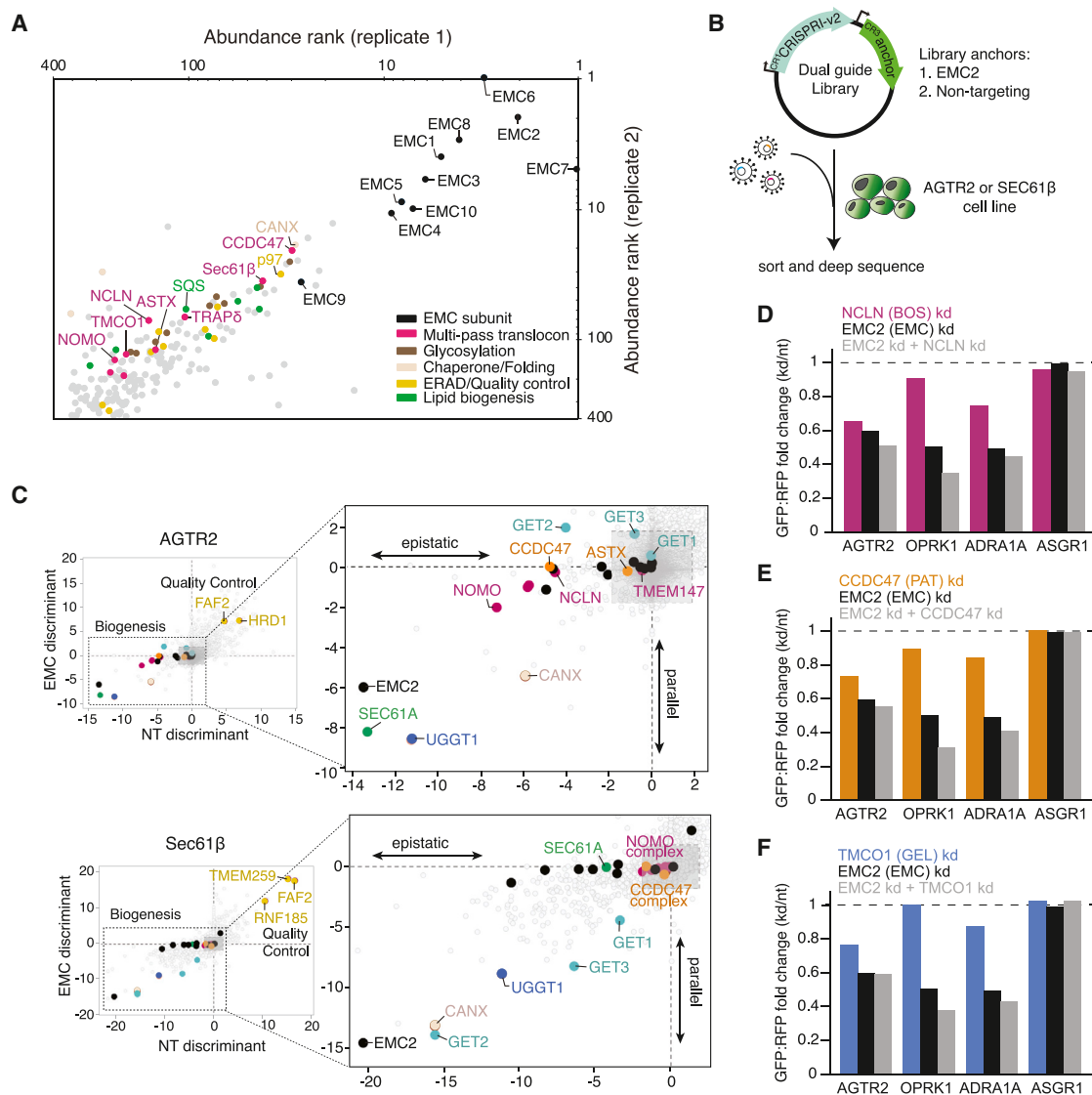


Figure 3. Components of the multipass translocon are epistatic with the EMC

(A) Abundance rank of factors that co-immunoprecipitate 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 Sec61 β reporter cell lines.⁴⁸

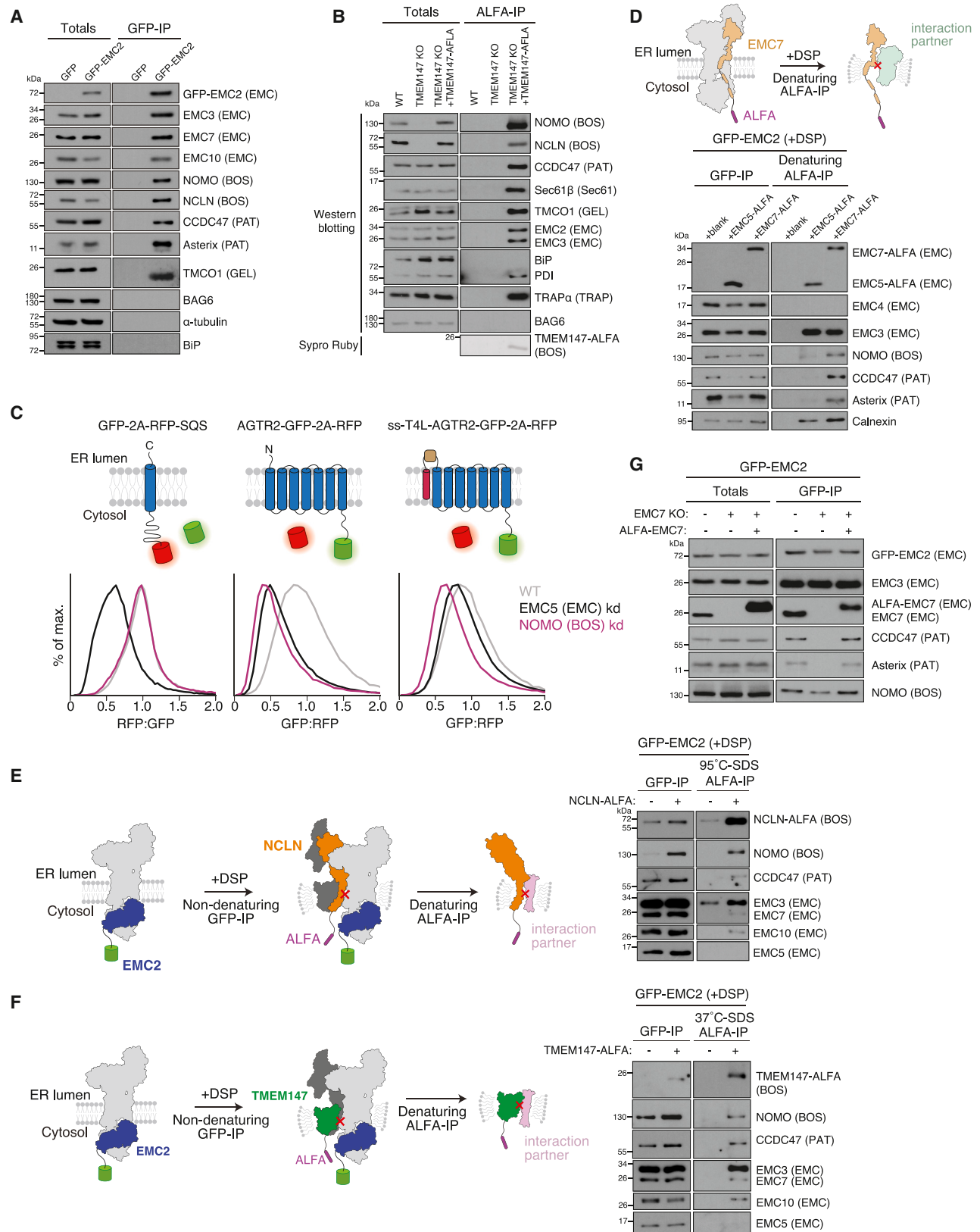
(C) Comparison of EMC2 and non-targeting (NT) genetic modifier CRISPRi screens using the discriminant score (calculated using both the phenotype and $-\log_{10}$ Mann-Whitney p value for each gene) for AGTR2 (above) and Sec61 β (below). Biogenesis factors are boxed and displayed at right in greater detail. 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 1 (representative histograms, Figure S2E).

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 S3B) and the BOS complex (Figure S3C) components are present at endogenous levels with the amine-reactive chemical crosslinker dithiobis-

succinimidyl propionate) (DSP). DSP has a length of $\sim 12\text{\AA}$, 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 4D). NOMO did not crosslink to EMC5, confirming that



(legend on next page)

the crosslinking and immunoprecipitation conditions were specific.

To confirm that this result did not reflect a long-range interaction between the flexible immunoglobulin (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 (nicalin [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 EMC55 (Figures 4E, 4F, and 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 (Figures 4A and 4B), to increase their local concentration and thereby enable structural analysis, we introduced an ~40-aa 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 cryoelectron microscopy (cryo-EM), we determined the structure of the EMC•BOS holocomplex to an overall resolution of 6.6 Å (Figures 5A and S4A and Table 1).

Although 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) (Figures S4B and S5A); and (2) BOS(tNOMO) in which we truncated all but the last 3 Ig-like domains of NOMO (3.7 Å resolution) (Figures 5B and S4C). Notably, the truncated NOMO resulted in 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 luminal domain of NCLN, suggesting that the remaining Ig-like domains are dynamic. Indeed, structures of the multipass translocon^{13,14} 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 S5A).

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 luminal interface between NCLN and EMC1, totaling a buried surface area of ~855 Å². Notably, the intramembrane interface is enriched for conserved hydrophobic residues (Figures 5C and S5B), suggesting that this interface may be more important than that in the lumen. Conversely, we hypothesize that the absence of crosslinks previously observed between EMC5 and NOMO (Figure 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 with that reported for the chaperone-binding mode of EMC (Figure S5C).¹⁸

To validate the holocomplex structure, we site-specifically introduced a photoactivatable amino acid to the luminal 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 S5F). These results further support the case for 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 suggests that BOS binding to the EMC and Sec61 is mutually exclusive (Figure 5D).¹³ Conversely, binding of BOS to the GEL and PAT complexes would all be compatible with interaction at the EMC (Figures S5D and S5G). 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 (Figures 3A, 4A, and S5E). Further, we verified using chemical crosslinking that the PAT complex is a physical interactor of the EMC and that its interaction is EMC7 dependent (Figures 4D and 4G). These data therefore suggest that the interaction between the EMC and the PAT complex is highly specific because loss of this peripheral subunit (EMC7) abolishes

Figure 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 knockout (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 1 and 2 in RPE1 cells. To target AGTR2 to the Sec61-dependent biogenesis pathway, the signal sequence (SS) of prolactin followed by T4 lysozyme (T4L) was appended to its N terminus. Histograms of fluorescence relative to a normalization control were 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 ($\Delta 7$) 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.

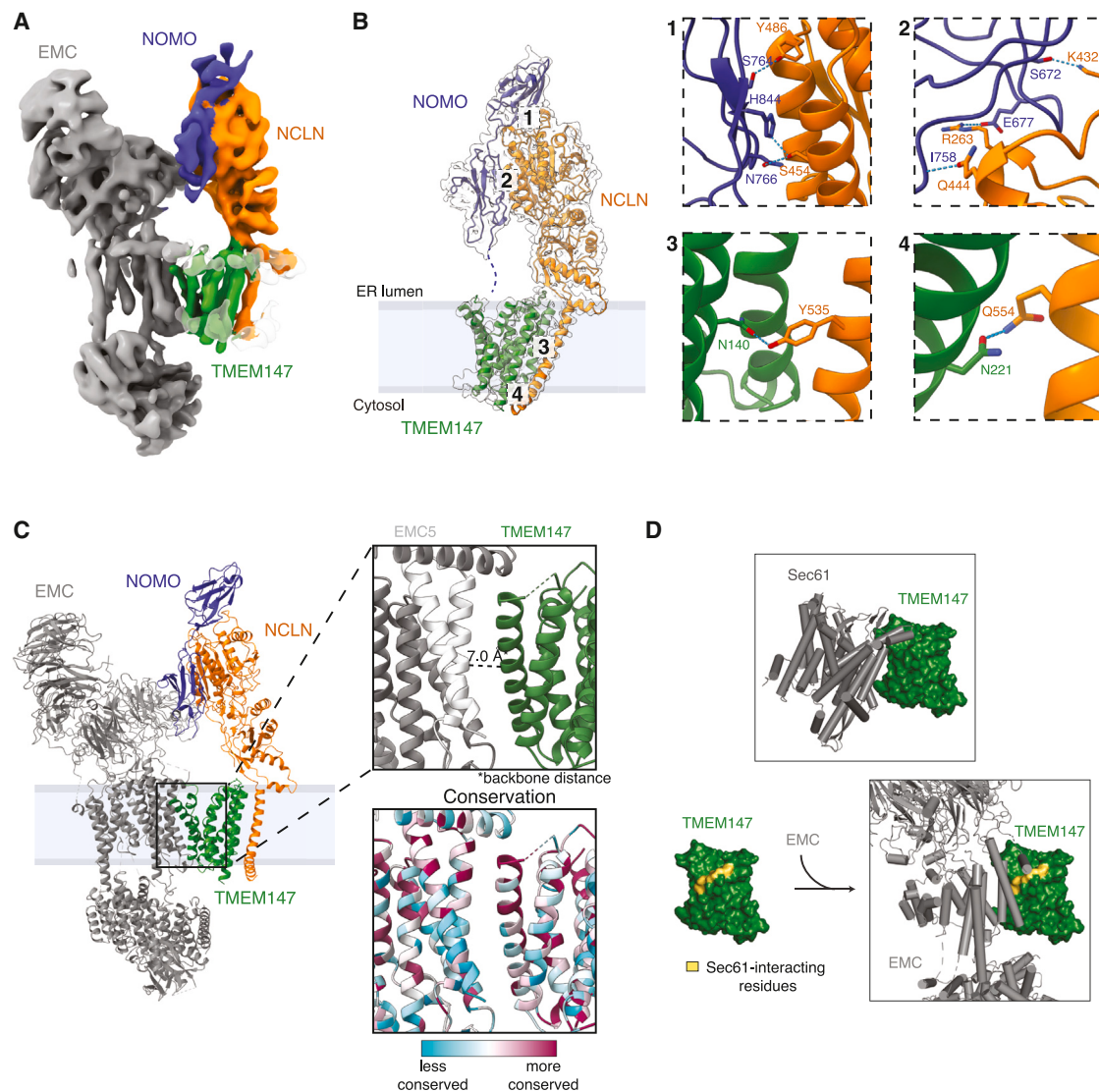


Figure 5. Structural analysis of the EMC-BOS holocomplex

(A) Coulomb potential map of the EMC-BOS complex (EMDB: 45295).

(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 (PDB: 9C7U; EMDB: 45294). 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 (PDB: 9C7V). 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 Sec61¹³ vs. the EMC (shown as gray cartoon). Residues on TMEM147 interacting with Sec61 are highlighted in yellow.

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 dispensable 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.

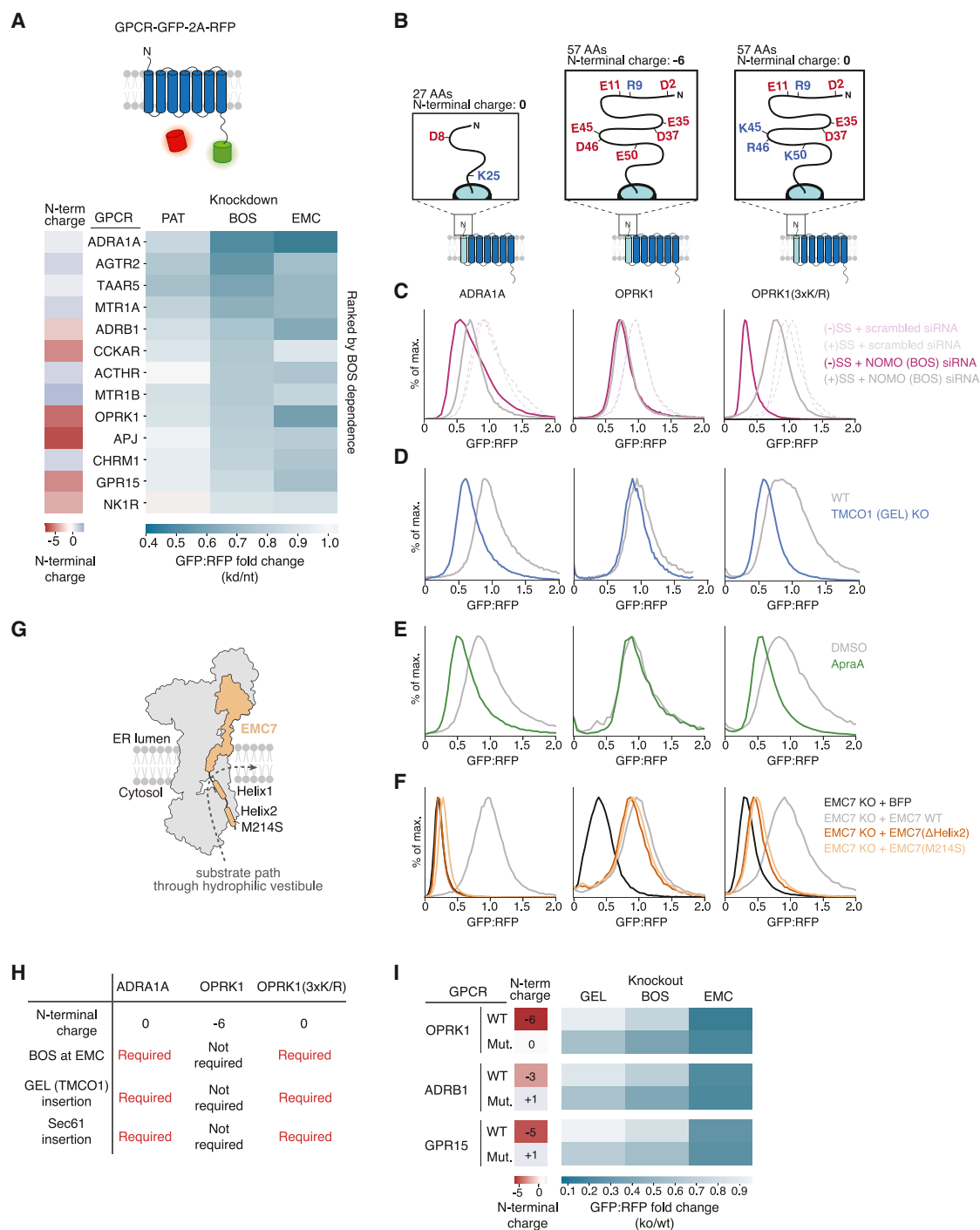


Figure 6. Properties of the soluble N terminus determine biogenesis pathway

(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. (Bottom) A heatmap summarizing the effect of depletion of each factor ordered by BOS dependence in comparison with 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 (3xK/R).

(C) Stability of the indicated reporters fused to the preprolactin signal sequence (SS) and T4L lysozyme, as described in Figure 4C. Cells were treated with small interfering RNAs (siRNAs) targeting NOMO or a scrambled control sequence.

(D) The reporters in (B) were assayed for dependence on TMC01 in WT and TMC01 KO cells.

(E) The reporters in (B) were assayed for dependence on the lateral gate of Sec61 using the inhibitor apratoxin A.

(legend continued on next page)

Second, M and ORF3a show differential dependence on the BOS complex. The primary difference between these topologically identical proteins is the increased length and charge of the N terminus of ORF3a (Figures 1C and 1D). Indeed, AGTR2 and ADRA1A, which also displayed a clear reliance on the BOS complex when utilizing the EMC (Figures 4C, S3D, and S3E), contain atypically positive/neutral N termini for GPCRs lacking signal sequences.⁶² Because the EMC uses a positively charged hydrophilic vestibule for insertion, this limits integration of substrates containing positively charged soluble domains.^{31,32} 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 (Figures 6A, S6A, and S6B). We found that those substrates with more positively charged N termini displayed a stronger dependence on the BOS complex (Figure 6A), but observed no connection between charge and EMC dependence (Figure S6A). 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 6B). ADRA1A, like AGTR2, only requires the BOS complex when utilizing the EMC for insertion of its first TMD (Figures 4C, 6C, S6C, and S6D). 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 whether 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.^{63,64} 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 (Figures 6D, 6E, S6D, and

S6E). 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* (Figures S7A and S7C). We further found that changes to the N-terminal charge of the GPCRs ADRB1 and GPR15 also alter their dependence on the GEL and BOS complexes (Figures 6I, S7B, and S7C). 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 whether 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.^{27,31} 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 (Figures 6F, 6G, and S6F). 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 the biogenesis pathway (e.g., Figure 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 (Figures S6G and S6H). These data suggest that multiple biophysical properties contribute to the biogenesis pathway required for each membrane protein at the ER.

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 (Figures 1 and 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

(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 (Δ Helix2), 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.³¹

(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 wild-type (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.

at the ER. However, even among proteins of identical topology, we observed distinct biogenesis requirements. Therefore, our data suggest 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's TMD dictates biogenesis requirements, in some cases requiring unique insertion machinery (i.e., for TAs) or chaperones (i.e., Asterix).^{8,16,17} Our data suggest that it is not only the TMD that dictates the 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 aa 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.^{58,65–69} 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 4G) provides one potential explanation for the conservation of the large soluble luminal 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 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 3A). Models consistent with a “first refusal” of substrates by the EMC best explain data showing that the EMC can enforce the correct folding of multipass substrates containing positively charged extracellular domains.³¹ Our observation that the PAT complex directly interacts with the EMC, including CDC47, 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 in-

serted by the EMC 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.³¹ These data support the model in which 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 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 explain why recruitment of the BOS complex to the EMC is mutually exclusive to its binding to Sec61 (Figure 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 substrate transfer to the multipass translocon following insertion of their first TMD by the EMC.

A central role for the 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 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 S1F). Remarkably, recruitment of the core insertase subunits EMC3/6 to the inner mitochondrial membrane was sufficient to rescue loss of Oxa1 in yeast.⁷⁰

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, although 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.⁷¹ In this case, upregulation of the EMC or, potentially, even GET1/2 may be sufficient to ensure efficient insertion of all required

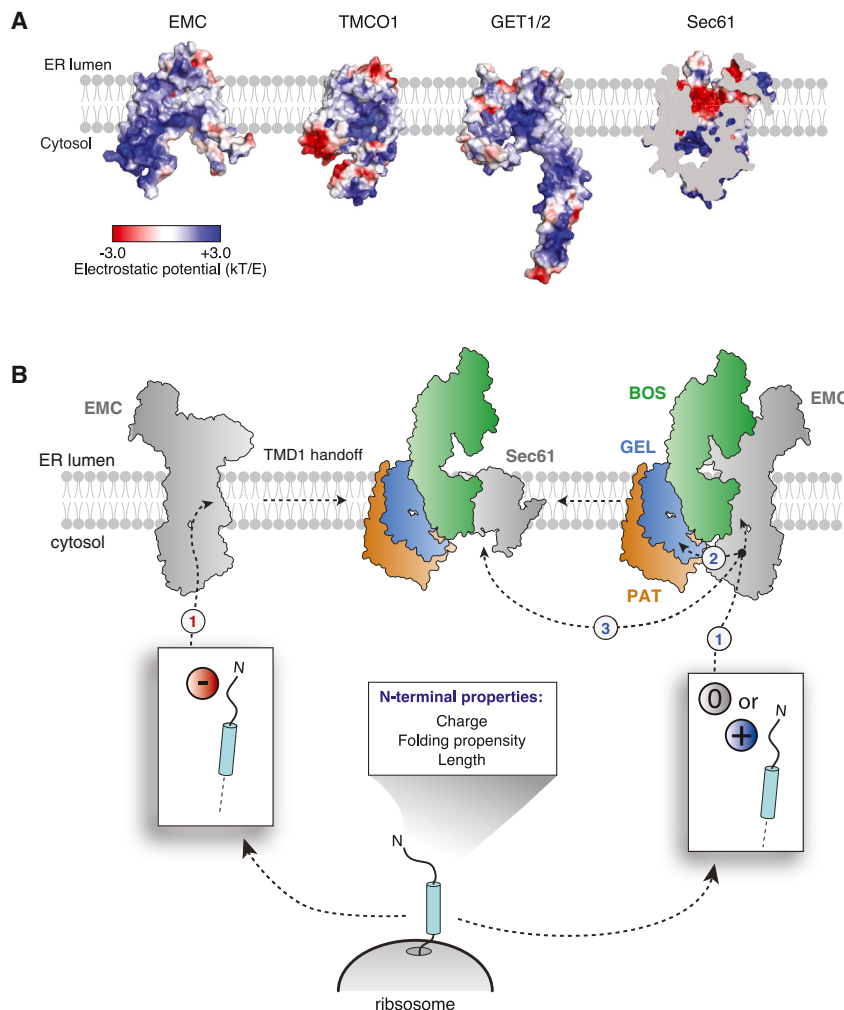


Figure 7. Model of multipass membrane protein biogenesis at the ER

(A) Comparison of the electrostatic surfaces of human ER insertases. Electrostatic potential was generated using Advanced Poisson-Boltzmann Solver (APBS) and mapped onto surface representation of EMC3/6,³¹ TMC01,¹³ GET1/2,³⁰ and Sec61.¹⁰

(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 hand-over to GEL or Sec61. After TMD1 is inserted, the multipass translocon inserts the remaining TMDs.

Limitations of the study

This study represents an important step toward understanding the features that determine the most efficient path for each substrate into the ER membrane. We have found that not only substrate topology but also biophysical features dictate dependence on distinct biogenesis factors (e.g., Figure 2). However, we have only tested a subset of the membrane proteome in this study. Further, although we have identified several biophysical features that dictate the biogenesis pathway,

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).⁷² 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.^{73,74}

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.

delineating a complete set of substrate rules remains an important area for future study. Experiments to test how TMD insertion propensity (e.g., hydrophobicity, helical propensity, length, etc.) and properties of its associated soluble domain (e.g., length, secondary structure, charge, charge density, etc.) dictate insertase specificity would be powerful. Finally, while we found that the EMC and multipass translocon co-localize into biogenesis hubs at the ER, directly mapping the handover of a substrate throughout the complete insertion, folding, and assembly process remains an important next step. Together, these types of mechanistic experiments will reveal how networks of biogenesis and quality-control factors cooperate to ensure synthesis of the full diversity of the membrane proteome.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact

Table 1. Cryo-EM 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 ($\text{\AA}/\text{pix}$)	0.416	0.416	0.416
Electron exposure ($\text{e}^-/\text{\AA}^2/\text{frame}$)	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 ($\text{\AA}$)	7.8–12.6	3.1–8.8	5.5–8.7
Map resolution ($\text{\AA}$, FSC = 0.143)	8.85	3.65	6.6
Refinement			
Software (<i>phenix.real_space_refine</i>)	N/A	PHENIX 1.20.1-4487	PHENIX 1.21-5207-000
Initial model used (PDB code)	N/A	AlphaFold	PDB: 8S9S and 9C7U
Correlation coefficient (CC_{mask})	N/A	0.83	0.70
Map sharpening B factor ($\text{\AA}^2$)	N/A	–60	–220
Model composition			
Non-hydrogen atoms	N/A	6,848	24,532
Protein residues	N/A	896	3,144
Ligands	N/A	NAG:1	NAG:6
B factors ($\text{\AA}^2$)	N/A	min/max/mean	N/A
Protein	N/A	26/170/82	22/221/106
Ligand	N/A	83/83/83	30/30/30
RMSD			
Bond lengths ($\text{\AA}$) (# > 4σ)	N/A	0.003	0.003
Bond angles ($^\circ$) (# > 4σ)	N/A	0.608	0.659
Validation			
MolProbity score	N/A	2.10	2.41
Clashscore	N/A	11.43	23.10
Rotamer outliers (%)	N/A	0.29	1.45
C β outliers (%)	N/A	0.00	0.00
CaBLAM outliers (%)	N/A	5.41	2.63
Ramachandran plot			
Favored (%)	N/A	90.9	93.44
Allowed (%)	N/A	9.1	6.43
Disallowed (%)	N/A	0.00	0.13

- Materials availability
- Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Cell lines
- **METHOD DETAILS**
 - Plasmids and antibodies
 - Cell line construction

- Fluorescent reporter CRISPRi screens
- Lentivirus production
- Expi293 cell line generation
- K562 cell spinfection with programmed guides
- Reporter assays
- Flow cytometry
- Preparation of human ER microsomes

- Mammalian *in vitro* translation
- Preparation of the ALFA nanobody conjugated to HRP for Western blotting
- DSP crosslinking
- Immunoprecipitation
- Mass spectrometry analysis of EMC-interacting factors
- Photo-crosslinking and immunoprecipitation
- Protein purification for structure determination
- Grid preparation and data collection
- Structure image processing
- Model building and refinement
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2024.08.005>.

ACKNOWLEDGMENTS

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. Cryoelectron microscopy was performed in the Beckman Institute Center for TEM at Caltech, and data were 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 (R.M.V.), the Howard Hughes Medical Institute (R.M.V.), the NIH's National Institute of General Medical Sciences DP2GM137412 (R.M.V.) and 1K99GM151478 (A.G.), the Burrough's Wellcome Fund (R.M.V.), an NSF-CAREER award 2145029 (R.M.V.), a Human Frontier Science Program Fellowship 2019/LT000858 (A.G.), the Deutsche Forschungsgemeinschaft (T.P.), the Tianqiao and Chrissy Chen Institute (T.P.), a Rosen Family fellowship (K.R.P.), and an Arie Jan Haagen-Smit Fellowship (K.R.P.).

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., and T.-Y.W.; resources, M.L.W., M.H., and T.-F.C.; writing – original draft and visualization, K.R.P., V.N.N., T.P., and R.M.V.; writing – review & editing, all authors; supervision and project administration, R.M.V.

DECLARATION OF INTERESTS

R.M.V. is a consultant and equity holder, and G.P.T. is a current employee, of Gate Bioscience.

Received: August 30, 2023

Revised: April 19, 2024

Accepted: August 2, 2024

Published: August 21, 2024

REFERENCES

- Harrison, S.C. (2008). Viral membrane fusion. *Nat. Struct. Mol. Biol.* 15, 690–698. <https://doi.org/10.1038/nsmb.1456>.
- Lenard, J. (2008). Viral Membranes. In *Encyclopedia of Virology* (Academic Press), pp. 308–314. <https://doi.org/10.1016/B978-012374410-4.00530-6>.
- von Heijne, G. (2007). The membrane protein universe: what's out there and why bother? *J. Intern. Med.* 261, 543–557. <https://doi.org/10.1111/j.1365-2796.2007.01792.x>.
- Rapoport, T.A., Junnickel, B., and Kutay, U. (1996). Protein Transport Across the Eukaryotic Endoplasmic Reticulum and Bacterial Inner Membranes. *Annu. Rev. Biochem.* 65, 271–303. <https://doi.org/10.1146/annurev.bi.65.070196.001415>.
- Shao, S., and Hegde, R.S. (2011). Membrane Protein Insertion at the Endoplasmic Reticulum. *Annu. Rev. Cell Dev. Biol.* 27, 25–56. <https://doi.org/10.1146/annurev-cellbio-092910-154125>.
- Halic, M., and Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. *Curr. Opin. Struct. Biol.* 15, 116–125. <https://doi.org/10.1016/j.sbi.2005.01.013>.
- Shan, S.O., and Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. *FEBS Lett.* 579, 921–926. <https://doi.org/10.1016/j.febslet.2004.11.049>.
- Guna, A., Hazu, M., Pinton Tomaleri, G., and Voorhees, R.M. (2023). A Tale of Two Pathways: Tail-Anchored Protein Insertion at the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.* 15, a041252. <https://doi.org/10.1101/cshperspect.a041252>.
- Van den Berg, B., Clemons, W.M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427, 36–44. <https://doi.org/10.1038/nature02218>.
- Voorhees, R.M., and Hegde, R.S. (2016). Structure of the Sec61 channel opened by a signal sequence. *Science* 351, 88–91. <https://doi.org/10.1126/science.1244992>.
- Enquist, K., Fransson, M., Boekel, C., Bengtsson, I., Geiger, K., Lang, L., Pettersson, A., Johansson, S., von Heijne, G., and Nilsson, I. (2009). Membrane-integration Characteristics of Two ABC Transporters, CFTR and P-glycoprotein. *J. Mol. Biol.* 387, 1153–1164. <https://doi.org/10.1016/j.jmb.2009.02.035>.
- Schorr, S., Nguyen, D., Haßdenteufel, S., Nagaraj, N., Cavalié, A., Greiner, M., Weissgerber, P., Loi, M., Paton, A.W., Paton, J.C., et al. (2020). Identification of signal peptide features for substrate specificity in human Sec62/Sec63-dependent ER protein import. *FEBS J.* 287, 4612–4640. <https://doi.org/10.1111/febs.15274>.
- McGilvray, P.T., Anghel, S.A., Sundaram, A., Zhong, F., Trnka, M.J., Fuller, J.R., Hu, H., Burlingame, A.L., and Keenan, R.J. (2020). An ER translocon for multi-pass membrane protein biogenesis. *eLife* 9, e56889. <https://doi.org/10.7554/eLife.56889>.
- Smalinskaitė, L., Kim, M.K., Lewis, A.J.O., Keenan, R.J., and Hegde, R.S. (2022). Mechanism of an intramembrane chaperone for multipass membrane proteins. *Nature* 611, 161–166. <https://doi.org/10.1038/s41586-022-05336-2>.
- Sundaram, A., Yamsek, M., Zhong, F., Hooda, Y., Hegde, R.S., and Keenan, R.J. (2022). Substrate-driven assembly of a translocon for multipass membrane proteins. *Nature* 611, 167–172. <https://doi.org/10.1038/s41586-022-05330-8>.
- Chitwood, P.J., and Hegde, R.S. (2020). An intramembrane chaperone complex facilitates membrane protein biogenesis. *Nature* 584, 630–634. <https://doi.org/10.1038/s41586-020-2624-y>.
- Meacock, S.L., Lecomte, F.J.L., Crawshaw, S.G., and High, S. (2002). Different Transmembrane Domains Associate with Distinct Endoplasmic Reticulum Components during Membrane Integration of a Polytopic Protein. *Mol. Biol. Cell* 13, 4114–4129. <https://doi.org/10.1091/mbc.e02-04-0198>.
- Chen, Z., Mondal, A., Abderemane-Ali, F., Jang, S., Niranjana, S., Montañó, J.L., Zaro, B.W., and Minor, D.L. (2023). EMC chaperone-CaV structure reveals an ion channel assembly intermediate. *Nature* 619, 410–419. <https://doi.org/10.1038/s41586-023-06175-5>.
- Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2018). The ER membrane protein complex is a transmembrane domain insertase. *Science* 359, 470–473. <https://doi.org/10.1126/science.aao3099>.
- 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., et al. (2018). The ER membrane protein complex interacts

- p>
cotranslationally to enable biogenesis of multipass membrane proteins.
- eLife*
- 7, e37018.
- <https://doi.org/10.7554/eLife.37018>
- .
21. Tian, S., Wu, Q., Zhou, B., Choi, M.Y., Ding, B., Yang, W., and Dong, M. (2019). Proteomic Analysis Identifies Membrane Proteins Dependent on the ER Membrane Protein Complex. *Cell Rep.* 28, 2517–2526.e5. <https://doi.org/10.1016/j.celrep.2019.08.006>.
 22. Chitwood, P.J., Juszkievicz, S., Guna, A., Shao, S., and Hegde, R.S. (2018). EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* 175, 1507–1519.e16. <https://doi.org/10.1016/j.cell.2018.10.009>.
 23. Satoh, T., Ohba, A., Liu, Z., Inagaki, T., and Satoh, A.K. (2015). dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in *Drosophila* photoreceptors. *eLife* 4, e06306. <https://doi.org/10.7554/eLife.06306>.
 24. Bai, L., You, Q., Feng, X., Kovach, A., and Li, H. (2020). Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature* 584, 475–478. <https://doi.org/10.1038/s41586-020-2389-3>.
 25. Miller-Vedam, L.E., Br  uning, B., Popova, K.D., Schirle Oakdale, N.T., Bonnar, J.L., Prabu, J.R., Boydston, E.A., Sevilano, N., Shurtleff, M.J., Stroud, R.M., et al. (2020). Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *eLife* 9, e62611. <https://doi.org/10.7554/eLife.62611>.
 26. O'Donnell, J.P., Phillips, B.P., Yagita, Y., Juszkievicz, S., Wagner, A., Malinverni, D., Keenan, R.J., Miller, E.A., and 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>.
 27. Pleiner, T., Tomaleri, G.P., Januszyk, K., Inglis, A.J., Hazu, M., and Voorhees, R.M. (2020). Structural basis for membrane insertion by the human ER membrane protein complex. *Science* 369, 433–436. <https://doi.org/10.1126/science.abb5008>.
 28. Borowska, M.T., Dominik, P.K., Anghel, S.A., Kossiakoff, A.A., and Keenan, R.J. (2015). A YidC-like Protein in the Archaeal Plasma Membrane. *Structure* 23, 1715–1724. <https://doi.org/10.1016/j.str.2015.06.025>.
 29. Kumazaki, K., Chiba, S., Takemoto, M., Furukawa, A., Nishiyama, K., Sugano, Y., Mori, T., Dohmae, N., Hirata, K., Nakada-Nakura, Y., et al. (2014). Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* 509, 516–520. <https://doi.org/10.1038/nature13167>.
 30. McDowell, M.A., Heimes, M., Fiorentino, F., Mehmood, S., Farkas,   ., Coy-Vergara, J., Wu, D., Bolla, J.R., Schmid, V., Heinze, R., et al. (2020). Structural Basis of Tail-Anchored Membrane Protein Biogenesis by the GET Insertase Complex. *Mol. Cell* 80, 72–86.e7. <https://doi.org/10.1016/j.molcel.2020.08.012>.
 31. Pleiner, T., Hazu, M., Pinton Tomaleri, G., Nguyen, V.N., Januszyk, K., and Voorhees, R.M. (2023). A selectivity filter in the ER membrane protein complex limits protein misinsertion at the ER. *J. Cell Biol.* 222, e202212007. <https://doi.org/10.1083/jcb.202212007>.
 32. Wu, H., and Hegde, R.S. (2023). Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Mol. Cell* 83, 961–973.e7. <https://doi.org/10.1016/j.molcel.2023.01.018>.
 33. Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H., and von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* 450, 1026–1030. <https://doi.org/10.1038/nature06387>.
 34. Dolan, K.A., Dutta, M., Kern, D.M., Kotecha, A., Voth, G.A., and Brohawn, S.G. (2022). Structure of SARS-CoV-2 M protein in lipid nanodiscs. *eLife* 11, e81702. <https://doi.org/10.7554/eLife.81702>.
 35. Kern, D.M., Sorum, B., Mali, S.S., Hoel, C.M., Sridharan, S., Remis, J.P., Toso, D.B., Kotecha, A., Bautista, D.M., and Brohawn, S.G. (2021). Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. *Nat. Struct. Mol. Biol.* 28, 573–582. <https://doi.org/10.1038/s41594-021-00619-0>.
 36. Kutay, U., Hartmann, E., and Rapoport, T.A. (1993). A class of membrane proteins with a C-terminal anchor. *Trends Cell Biol.* 3, 72–75. [https://doi.org/10.1016/0962-8924\(93\)90066-A](https://doi.org/10.1016/0962-8924(93)90066-A).
 37. Mateja, A., Szlachcic, A., Downing, M.E., Dobosz, M., Mariappan, M., Hegde, R.S., and Keenan, R.J. (2009). The structural basis of tail-anchored membrane protein recognition by Get3. *Nature* 461, 361–366. <https://doi.org/10.1038/nature08319>.
 38. Stefanovic, S., and Hegde, R.S. (2007). Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* 128, 1147–1159. <https://doi.org/10.1016/j.cell.2007.01.036>.
 39. 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., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647–661. <https://doi.org/10.1016/j.cell.2014.09.029>.
 40. 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., et al. (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>.
 41. Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., Voorden, S. van, Thanedar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 Is an E3 Ubiquitin Ligase Involved in Degradation of Proteins from the Endoplasmic Reticulum. *J. Biol. Chem.* 279, 3525–3534. <https://doi.org/10.1074/jbc.M307453200>.
 42. Olzmann, J.A., Richter, C.M., and Kopito, R.R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proc. Natl. Acad. Sci. USA* 110, 1345–1350. <https://doi.org/10.1073/pnas.1213738110>.
 43. Mariappan, M., Mateja, A., Dobosz, M., Bove, E., Hegde, R.S., and Keenan, R.J. (2011). The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature* 477, 61–66. <https://doi.org/10.1038/nature10362>.
 44. Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F., et al. (2005). Exploration of the Function and Organization of the Yeast Early Secretory Pathway through an Epistatic Miniarray Profile. *Cell* 123, 507–519. <https://doi.org/10.1016/j.cell.2005.08.031>.
 45. Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645. <https://doi.org/10.1016/j.cell.2008.06.025>.
 46. Vilardi, F., Lorenz, H., and Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. *J. Cell Sci.* 124, 1301–1307. <https://doi.org/10.1242/jcs.084277>.
 47. 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., et al. (2022). MTCH2 is a mitochondrial outer membrane protein insertase. *Science* 378, 317–322. <https://doi.org/10.1126/science.add1856>.
 48. Guna, A., Page, K.R., Replogle, J.M., Esantsi, T.K., Wang, M.L., Weissman, J.S., and Voorhees, R.M. (2023). A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *BMC Genomics* 24, 651. <https://doi.org/10.1186/s12864-023-09754-y>.
 49. Inglis, A.J., Page, K.R., Guna, A., and Voorhees, R.M. (2020). Differential Modes of Orphan Subunit Recognition for the WRB/CAML Complex. *Cell Rep.* 30, 3691–3698.e5. <https://doi.org/10.1016/j.celrep.2020.02.084>.
 50. Colombo, S.F., Cardani, S., Maroli, A., Vitiello, A., Soffientini, P., Crespi, A., Bram, R.F., Benfante, R., and Borgese, N. (2016). Tail-anchored Protein Insertion in Mammals: FUNCTION AND RECIPROCAL INTERACTIONS OF THE TWO SUBUNITS OF THE TRC40 RECEPTOR. *J. Biol. Chem.* 291, 15292–15306. <https://doi.org/10.1074/jbc.M115.707752>.

51. Dettmer, U., Kuhn, P.-H., Abou-Ajram, C., Lichtenthaler, S.F., Krüger, M., Kremmer, E., Haass, C., and Haffner, C. (2010). Transmembrane Protein 147 (TMEM147) Is a Novel Component of the Nicalin-NOMO Protein Complex. *J. Biol. Chem.* 285, 26174–26181. <https://doi.org/10.1074/jbc.M110.132548>.
52. Pleiner, T., Hazu, M., Tomaleri, G.P., Januszyk, K., Oania, R.S., Sweredoski, M.J., Moradian, A., Guna, A., and Voorhees, R.M. (2021). WNK1 is an assembly factor for the human ER membrane protein complex. *Mol. Cell* 81, 2693–2704.e12. <https://doi.org/10.1016/j.molcel.2021.04.013>.
53. Volkmar, N., Thezenas, M.-L., Louie, S.M., Juskiewicz, S., Nomura, D.K., Hegde, R.S., Kessler, B.M., and Christianson, J.C. (2019). The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis. *J. Cell Sci.* 132, jcs223453. <https://doi.org/10.1242/jcs.223453>.
54. Guna, A., and Hegde, R.S. (2018). Transmembrane Domain Recognition during Membrane Protein Biogenesis and Quality Control. *Curr. Biol.* 28, R498–R511. <https://doi.org/10.1016/j.cub.2018.02.004>.
55. Rao, M., Okreglak, V., Chio, U.S., Cho, H., Walter, P., and Shan, S.O. (2016). Multiple selection filters ensure accurate tail-anchored membrane protein targeting. *eLife* 5, e21301. <https://doi.org/10.7554/eLife.21301>.
56. Shao, S., Rodrigo-Brenni, M.C., Kivlen, M.H., and Hegde, R.S. (2017). Mechanistic basis for a molecular triage reaction. *Science* 355, 298–302. <https://doi.org/10.1126/science.aah6130>.
57. Wang, F., Brown, E.C., Mak, G., Zhuang, J., and Denic, V. (2010). A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Mol. Cell* 40, 159–171. <https://doi.org/10.1016/j.molcel.2010.08.038>.
58. Gemmer, M., Chaillet, M.L., van Loenhout, J., Cuevas Arenas, R., Vimpas, D., Gröblers-Mulderij, M., Koh, F.A., Albanese, P., Scheltema, R.A., Howes, S.C., et al. (2023). Visualization of translation and protein biogenesis at the ER membrane. *Nature* 614, 160–167. <https://doi.org/10.1038/s41586-022-05638-5>.
59. Bergeron, J.J., Brenner, M.B., Thomas, D.Y., and Williams, D.B. (1994). Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem. Sci.* 19, 124–128. [https://doi.org/10.1016/0968-0004\(94\)90205-4](https://doi.org/10.1016/0968-0004(94)90205-4).
60. Kelleher, D.J., and Gilmore, R. (2006). An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47R–62R. <https://doi.org/10.1093/glycob/cwj066>.
61. Meyer, H., Bug, M., and Bremer, S. (2012). Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat. Cell Biol.* 14, 117–123. <https://doi.org/10.1038/ncb2407>.
62. Wallin, E., and von Heijne, G. (1995). Properties of N-terminal tails in G-protein coupled receptors: a statistical study. *Protein Eng.* 8, 693–698. <https://doi.org/10.1093/protein/8.7.693>.
63. Paatero, A.O., Kellosalo, J., Dunyak, B.M., Almaliti, J., Gestwicki, J.E., Gerwick, W.H., Taunton, J., and Paavilainen, V.O. (2016). Apratoxin Kills Cells by Direct Blockade of the Sec61 Protein Translocation Channel. *Cell Chem. Biol.* 23, 561–566. <https://doi.org/10.1016/j.chembiol.2016.04.008>.
64. Thornburg, C.C., Cowley, E.S., Sikorska, J., Shaala, L.A., Ishmael, J.E., Youssef, D.T.A., and McPhail, K.L. (2013). Apratoxin H and apratoxin A sulfoxide from the Red Sea cyanobacterium *Moorea producens*. *J. Nat. Prod.* 76, 1781–1788. <https://doi.org/10.1021/np4004992>.
65. Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T.A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* 357, 47–52. <https://doi.org/10.1038/357047a0>.
66. Hartmann, E., Görlich, D., Kostka, S., Otto, A., Kraft, R., Knespel, S., Bürger, E., Rapoport, T.A., and Prehn, S. (1993). A tetrameric complex of membrane proteins in the endoplasmic reticulum. *Eur. J. Biochem.* 214, 375–381. <https://doi.org/10.1111/j.1432-1033.1993.tb17933.x>.
67. Jaskolowski, M., Jomaa, A., Gamerdinger, M., Shrestha, S., Leibundgut, M., Deuerling, E., and Ban, N. (2023). Molecular basis of the TRAP complex function in ER protein biogenesis. *Nat. Struct. Mol. Biol.* 30, 770–777. <https://doi.org/10.1038/s41594-023-00990-0>.
68. Kalies, K.U., Rapoport, T.A., and Hartmann, E. (1998). The β Subunit of the Sec61 Complex Facilitates Cotranslational Protein Transport and Interacts with the Signal Peptidase during Translocation. *J. Cell Biol.* 141, 887–894. <https://doi.org/10.1083/jcb.141.4.887>.
69. Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T.A. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* 134, 25–35. <https://doi.org/10.1083/jcb.134.1.25>.
70. Güngör, B., Flohr, T., Garg, S.G., and Herrmann, J.M. (2022). The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria. *PLoS Biol.* 20, e3001380. <https://doi.org/10.1371/journal.pbio.3001380>.
71. 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., et al. (2024). A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* 625, 92–100. <https://doi.org/10.1038/s41586-023-06045-0>.
72. Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S., et al. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323, 1693–1697. <https://doi.org/10.1126/science.1167983>.
73. Sachelar, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B., and Koch, H.-G. (2013). YidC Occupies the Lateral Gate of the SecYEG Translocon and Is Sequentially Displaced by a Nascent Membrane Protein. *J. Biol. Chem.* 288, 16295–16307. <https://doi.org/10.1074/jbc.M112.446583>.
74. Wickles, S., Singharoy, A., Andreani, J., Seemayer, S., Bischoff, L., Berninghausen, O., Soeding, J., Schulten, K., van der Sluis, E.O., and Beckmann, R. (2014). A structural model of the active ribosome-bound membrane protein insertase YidC. *eLife* 3, e03035. <https://doi.org/10.7554/eLife.03035>.
75. Chakrabarti, O., and Hegde, R.S. (2009). Functional Depletion of Mahogunin by Cytosolically Exposed Prion Protein Contributes to Neurodegeneration. *Cell* 30, 70–82. <https://doi.org/10.1016/j.cell.2009.03.042>.
76. Mariappan, M., Li, X., Stefanovic, S., Sharma, A., Mateja, A., Keenan, R.J., and Hegde, R.S. (2010). A Ribosome-Associating Factor Chaperones Tail-Anchored Membrane Proteins. *Nature* 466, 1120–1124. <https://doi.org/10.1038/nature09296>.
77. Sharma, A., Mariappan, M., Appathurai, S., and Hegde, R.S. (2010). In vitro dissection of protein translocation into the mammalian endoplasmic reticulum. *Methods Mol. Biol.* 619, 339–363. https://doi.org/10.1007/978-1-60327-412-8_20.
78. Stevens, T.A., Tomaleri, G.P., Hazu, M., Wei, S., Nguyen, V.N., DeKalb, C., Voorhees, R.M., and Pleiner, T. (2023). A nanobody-based strategy for rapid and scalable purification of human protein complexes. *Nat. Protoc.* 19, 127–158. <https://doi.org/10.1038/s41596-023-00904-w>.
79. Liu, L., Spurrier, J., Butt, T.R., and Strickler, J.E. (2008). Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expr. Purif.* 62, 21–28. <https://doi.org/10.1016/j.pep.2008.07.010>.
80. Jost, M., Chen, Y., Gilbert, L.A., Horlbeck, M.A., Krenning, L., Menchon, G., Rai, A., Cho, M.Y., Stern, J.J., Protal, A.E., et al. (2017). Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Mol. Cell* 68, 210–223.e6. <https://doi.org/10.1016/j.molcel.2017.09.012>.
81. Ai, H.W., Shen, W., Sagi, A., Chen, P.R., and Schultz, P.G. (2011). Probing Protein-Protein Interactions with a Genetically Encoded Photo-crosslinking Amino Acid. *ChemBiochem* 12, 1854–1857. <https://doi.org/10.1002/cbic.201100194>.

82. Elsässer, S.J., Ernst, R.J., Walker, O.S., and Chin, J.W. (2016). Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat. Methods* 13, 158–164. <https://doi.org/10.1038/nmeth.3701>.
83. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. <https://doi.org/10.1038/nprot.2013.143>.
84. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* 27, 14–25. <https://doi.org/10.1002/pro.3235>.
85. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
86. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296. <https://doi.org/10.1038/nmeth.4169>.
87. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. <https://doi.org/10.1107/S0907444910007493>.
88. Casañal, A., Lohkamp, B., and Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci.* 29, 1069–1078. <https://doi.org/10.1002/pro.3791>.
89. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
90. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* 74, 531–544. <https://doi.org/10.1107/S2059798318006551>.
91. 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., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877. <https://doi.org/10.1107/S2059798319011471>.
92. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51. <https://doi.org/10.1016/j.jsb.2005.07.007>.
93. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
94. Yurtsev, E., Friedman, J., and Gore, J. (2015). FlowCytometryTools: Version 0.4.5. (v0.4.5). Zenodo. <https://doi.org/10.5281/zenodo.32991>. <https://explore.openaire.eu/search/software?pid=10.5281%2Fzenodo.32991>.
95. Walter, P., and Blobel, G. (1983). Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol.* 96, 84–93. [https://doi.org/10.1016/s0076-6879\(83\)96010-x](https://doi.org/10.1016/s0076-6879(83)96010-x).
96. Cabantous, S., Terwilliger, T.C., and Waldo, G.S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* 23, 102–107. <https://doi.org/10.1038/nbt1044>.
97. 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., et al. (2016). Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.* 7, 11046. <https://doi.org/10.1038/ncomms11046>.
98. 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., et al. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* 38, 954–961. <https://doi.org/10.1038/s41587-020-0470-y>.
99. Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaefer, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., et al. (2019). The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nat. Commun.* 10, 4403. <https://doi.org/10.1038/s41467-019-12301-7>.
100. Pleiner, T., Bates, M., and Görlich, D. (2018). A toolbox of anti-mouse and anti-rabbit IgG secondary nanobodies. *J. Cell Biol.* 217, 1143–1154. <https://doi.org/10.1083/jcb.201709115>.
101. Doblmann, J., Dusberger, F., Imre, R., Hudecz, O., Stanek, F., Mechtler, K., and Dürnberger, G. (2019). apQuant: Accurate Label-Free Quantification by Quality Filtering. *J. Proteome Res.* 18, 535–541. <https://doi.org/10.1021/acs.jproteome.8b00113>.
102. Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2013). Corrigendum: Global quantification of mammalian gene expression control. *Nature* 495, 126–127. <https://doi.org/10.1038/nature11848>.
103. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-EMC2	Proteintech	Cat# 25443-1-AP; RRID:AB_2750836
Mouse monoclonal anti-EMC3	Proteintech	Cat# 67205-1-Ig; RRID:AB_2882498
Rabbit polyclonal anti-EMC4	Proteintech	Cat# 27708-1-AP; RRID:AB_2880950
Rabbit polyclonal anti-EMC5	Bethyl	Cat# A305-833A-M; RRID:AB_2890207
Rabbit polyclonal anti-EMC7/C15orf24	Proteintech	Cat# 27550-1-AP; RRID:AB_2880902
Rabbit polyclonal anti-EMC10/C19orf63	Abcam	Cat# ab180148; RRID:AB_2889936
Rabbit polyclonal anti-TMCO1	Proteintech	Cat# 27757-1-AP; RRID:AB_2880962
Rabbit polyclonal anti-TMEM147	Thermo Fisher Scientific	Cat# PA5-95876; RRID:AB_2807678
Rabbit polyclonal anti-NCLN	Bethyl	Cat# A305-623A-M; RRID:AB_2782781
Goat polyclonal anti-NOMO1	Thermo Fisher Scientific	Cat# PA5-47534; RRID:AB_2607776
Rabbit polyclonal anti-NOMO2	Proteintech	Cat# 14328-1-AP; RRID:AB_2152154
Rabbit polyclonal anti-CCDC47	Bethyl	Cat# A305-100A-M; RRID:AB_2631495
Rabbit polyclonal anti-GET2/CAML	Synaptic Systems	Cat# 359 002; RRID:AB_2620118
Rabbit polyclonal anti-HRD1/SYVN1	Proteintech	Cat# 13473-1-AP; RRID:AB_2287023
Rabbit polyclonal anti-FAF2	Proteintech	Cat# 16251-1-AP; RRID:AB_2262469
Mouse monoclonal anti-UGGT1 (H-9)	Santa Cruz Biotechnology	Cat# sc-374565; RRID:AB_11008073
Rabbit polyclonal anti-Asterix/WDR83OS	Atlas Antibodies	Cat# HPA065685; RRID:AB_2685531
Rabbit polyclonal anti-Calnexin	Proteintech	Cat# 10427-2-AP; RRID:AB_2069033
Mouse monoclonal anti-BiP	BD Biosciences	Cat# 610979; RRID:AB_398292
Mouse monoclonal anti-PDI	Enzo Life Sciences	Cat# ADI-SPA-891-D; RRID:AB_2039449
Mouse monoclonal anti- α -Tubulin	Sigma-Aldrich	Cat# T9026; RRID:AB_477593
Rabbit polyclonal anti-GFP	Gift from Ramanujan Hegde; Chakrabarti and Hegde ⁷⁵	N/A
Rabbit polyclonal anti-TRAP α	Gift from Ramanujan Hegde	N/A
Rabbit polyclonal anti-BAG6	Gift from Ramanujan Hegde; Mariappan et al. ⁷⁶	N/A
Rabbit polyclonal anti-Sec61 β	Gift from Ramanujan Hegde	N/A
HRP-conjugated goat anti-rabbit IgG	BioRad	Cat# 170-6515; RRID:AB_11125142
HRP-conjugated goat anti-mouse IgG	BioRad	Cat# 172-1011; RRID:AB_11125936
HRP-conjugated anti-ALFA Nanobody	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Doxycycline	Sigma-Aldrich	Cat# D9891, CAS: 24390-14-5
GDN	Anatrace	Cat# GDN101, CAS: 1402423-29-3
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Roche	Cat# 11873580001
Pierce Streptavidin Magnetic Beads	Thermo Scientific	Cat# 88817
Biotin	Sigma-Aldrich	Cat# B4501
EasyTag L-[35S]-Methionine	Perkin Elmer	Cat# NEG709A005MC
PEI MAX Mw 40,000	Polysciences	Cat# 24765-1, CAS: 49553-93-7
Puromycin Dihydrochloride	Thermo Scientific	Cat# A1113803
RNasin	Promega	Cat# N251
SP6 Polymerase	New England Biolabs	Cat# M0207L
SPRISelect beads	Beckman Coulter	Cat# B23317

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NEBNext® Ultra™ II Q5® Master Mix	New England Biolabs	Cat# M0544L
TransIT-293 transfection reagent	Mirus	Cat# MIR2705
Lipofectamine 3000	Thermo Scientific	Cat# L3000015
RNAiMAX lipofectamine	Thermo Scientific	Cat# 13778150
Polybrene transfection reagent	Sigma-Aldrich	Cat# TR-1003
Apratoxin A	Gift from Ville Paavilainen and Kerry McPhail	N/A
T1 mix	Sharma et al. ⁷⁷	N/A
DSP (dithiobis(succinimidyl propionate))), Lomant's Reagent	Thermo Scientific	Cat# 22585
His14-Avi-SUMOEl1-tagged anti-GFP nanobody	Stevens et al. ⁷⁸	Addgene ID #149336
His14-Avi-SUMOSTar-anti-ALFA	Pleiner et al. ³¹	N/A
SUMOEu1 Protease	Stevens et al. ⁷⁸	Addgene ID #149333
SUMOSTar Protease	LifeSensors; Liu et al. ⁷⁹	Cat# SP4110
Tris-(2-Carboxyethyl)phosphine, Hydrochloride (TCEP)	Thermo Scientific	Cat# T2556
2-chloro-acetamide (CAA)	MP Biomedicals	SKU: 02154955-CF
Lysyl Endoproteinase (Lys-C)	Wako Chemicals	Cat# 125-05061
Trypsin	Thermo Scientific	Cat# 90057
EZ-Link™ Maleimide Activated Horseradish Peroxidase	Thermo Scientific	Cat# 31485
3'-azibutyl-N-carbamoyl-lysine hydrochloride (AbK)	Iris Biotech	HAA3110

Deposited data

Mass spectrometry data of the EMC	This paper	PRIDE: PXD045009
Structure of the human BOS complex in GDN	This paper	EMDB: 45293
Structure of the human truncated BOS complex in GDN	This paper	PDB: 9C7U, EMDB: 45294
Structure of the human BOS:human EMC complex in GDN	This paper	PDB: 9C7V, EMDB: 45295
Structure of the human ER membrane protein complex (EMC) in GDN	Pleiner et al. ³¹	PDB: 8S9S
Cryo-EM structure of the human ribosome-TMCO1 translocon	McGilvray et al. ¹³	PDB: 6W6L
Homo sapiens WRB/CAML heterotetramer in complex with a TRC40 dimer	McDowell et al. ³⁰	PDB: 6SO5
The structure of the mammalian Sec61 channel opened by a signal sequence	Voorhees and Hegde ¹⁰	PDB: 3JC2
Structure of a human EMC:human Cav1.2 channel complex in GDN detergent	Chen et al. ¹⁸	PDB: 8EOI

Experimental models: cell lines

K562 KRAB-BFP-dCAS9	Gilbert et al. ³⁹	N/A
HEK 293T/17	ATCC	CRL-11268
K562-CRISPRi-Tet-ON-((ER)-GFP1-10)-(tet-RFP-P2A-SEC61β-GFP11)	Guna et al. ⁸	N/A
K562-CRISPRi-Tet-ON-(tet-AGTR2-GFP-P2A-RFP)	This paper	N/A
K562-CRISPRi-Tet-ON-(tet-SARS-CoV-2_ORF3a-GFP-P2A-RFP)	This paper	N/A
K562-CRISPRi-Tet-ON-(tet-SARS-CoV-2_ORF3a-GFP-P2A-RFP)	This paper	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HEK 293T TMCO1 KO	This paper	N/A
HEK 293T TMEM147 KO	This paper	N/A
HEK 293T CCDC47 KO	This paper	N/A
293T TREX Flp-In GFP-EMC2-P2A-RFP	Pleiner et al. ²⁷	N/A
293T TREX Flp-In GFP-EMC2-P2A-RFP + EMC7 KO	Pleiner et al. ³¹	N/A
293T TREX Flp-In GFP-EMC2-P2A-RFP + hsEMC5-ALFA-2A-TagBFP-3xFLAG	This paper	N/A
293T TREX Flp-In GFP-EMC2-P2A-RFP + Pr(ss)-hsEMC7(24-end)-ALFA- P2A-TagBFP-3xFLAG	This paper	N/A
293T TREX Flp-In HEK 293T GFP-EMC2-P2A-RFP + hsNCLN- ALFA-P2A-TagBFP-3xFLAG	This paper	N/A
293T TREX Flp-In GFP-EMC2-P2A-RFP + hsTMEM147-ALFA-P2A-TagBFP-3xFLAG	This paper	N/A
RPE1 KRAB-BFP-dCAS9	Jost et al. ⁸⁰	N/A
Expi293 TMEM147-TEV-GFP-P2A-RFP	This paper	N/A
Expi293 TMEM147-GFP	This paper	N/A
Expi293 NCLN-P2A-RFP	This paper	N/A
Expi293 NOMO2-P2A-TagBFP	This paper	N/A
Expi293 TMEM147-TEV-GFP-40aa-ALFA-EMC2	This paper	N/A
Oligonucleotides		
See Table S5 for oligonucleotides sequences	This paper	Table S5
Recombinant DNA		
CRISPRi-v2-ctrl dual guide library	Guna et al. ⁴⁸	Addgene ID #197348
CRISPRi-v2-EMC2 dual guide library	Guna et al. ⁴⁸	Addgene ID #197349
pHAGE2 lentiviral transfer vector	Gift from Magnus A. Hoffmann and Pamela Bjorkman	N/A
psPAX2 2nd generation lenti-viral packaging plasmid	Gift from Didier Trono	Addgene ID #12260
pCMV-VSV-G 2nd generation lenti-viral packaging plasmid	Gift from Bob Weinberg	Addgene ID #8454
SFFV-tet3G backbone	Jost et al. ⁸⁰	N/A
SP64 vector	Promega	Cat# P1241
GFP-2A-RFP reporter cassette plasmid	Pleiner et al. ³¹	N/A
Amber suppression system plasmids	Ai et al. ⁸¹ ; Elsässer et al. ⁸²	N/A
pSpCas9(BB)-2A-Puro (pX459)	Gift from Feng Zhang; Ran et al. ⁸³	Addgene ID #48139
Software and algorithms		
FlowJo v10.8	BD Biosciences	https://www.bdbiosciences.com/en-us/products/software/flowjo-v10-software
UCSF ChimeraX-1.6.1	Goddard et al. ⁸⁴ ; Pettersen et al. ⁸⁵	https://www.rbvi.ucsf.edu/chimerax/
PyMOL	Schrödinger	https://pymol.org/2/
InterfaceResidues	PyMOL	https://pymolwiki.org/index.php/InterfaceResidues
CryoSPARC v3.2-4.2.1	Punjani et al. ⁸⁶	https://cryosparc.com/
COOT	Emsley et al. ⁸⁷ ; Casañal et al. ⁸⁸	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
AlphaFold2-Multimer ColabFold	Mirdita et al. ⁸⁹	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb
<i>phenix.real_space_refinement</i>	Afonine et al. ⁹⁰ ; Liebschner et al. ⁹¹	https://phenix-online.org/
SerialEM	Mastronarde ⁹²	https://bio3d.colorado.edu/SerialEM/
Adobe Illustrator	Adobe	https://www.adobe.com/uk/creativecloud.html
ImageJ	Schneider et al. ⁹³	https://imagej.nih.gov/ij/
Proteome Discoverer SEQUEST v2.5	Thermo Scientific	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html
Python FlowCytometryTools package	Yurtsev et al. ⁹⁴	https://eyurtsev.github.io/FlowCytometryTools/
Pipeline for analysis of screen sequencing	Horlbeck et al. ⁴⁰	https://github.com/mhorlbeck/ScreenProcessing
Other		
SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific	Cat# 34580
SuperSignal™ West Femto Maximum Sensitivity Substrate	Thermo Fisher Scientific	Cat# 34094
Rabbit Reticulocyte Lysate Mix	Sharma et al. ⁷⁷	N/A
Canine rough microsomes	Walter and Blobel ⁹⁵	N/A
PURExpress ΔRF123 Kit	New England Biolabs	Cat# E6850S
DPBS, no calcium, no magnesium	Thermo Scientific	Cat# 14190144
DMEM, high glucose, GlutaMAX Supplement, pyruvate	Thermo Scientific	Cat# 10569010
DMEM, high glucose, no glutamine, no methionine, no cysteine	Thermo Scientific	Cat# 21013024
DMEM/F-12	Thermo Scientific	Cat# 11320033
Expi293™ Expression Medium	Thermo Scientific	Cat# A1435102
Trypsin-EDTA (0.25%), phenol red	Thermo Scientific	Cat# 25200056
HyClone RPMI-1640 Medium with HEPES, L-glutamine	Cytiva	Cat# SH30255.01
Fetal Bovine Serum	Bio-Techne	Cat# S11150
Fetal Bovine Serum – TET tested	Bio-Techne	Cat# S10350
Penicillin-Streptomycin	Sigma-Aldrich	Cat# P4333
IOB w/ Micro-Carrier Flask 1L	Belco	SKU: 1965-61010
Nucleospin Blood XL kit	Takara Bio	Cat# 740950.10
SH800S Cell Sorter	SONY	N/A
FACSAria™ Fusion Flow Cytometer	BD Biosciences	N/A
HiSeq 2500 System	Illumina	N/A
2100 Bioanalyzer Instrument	Agilent	N/A
MLA-80 Fixed-Angle Rotor	Beckman-Coulter	Cat# 367096
Attune NxT Flow Cytometer	Thermo Scientific	N/A
MACSQuant VYB Flow Cytometer	Miltenyi Biotec	N/A
EASY-nLC 1200	Thermo Scientific	N/A
Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer	Thermo Scientific	N/A
Aurora UHPLC Column	Ion Opticks	Cat# AUR2-25075C18A
UVP B-100 Series UV Lamp	Analytik Jena	N/A
Superose 6 Increase 3.2/300 SEC Column	Cytiva	Cat# 29091598

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Pierce™ Protein Concentrators PES, 30K MWCO	Thermo Scientific	Cat# 88502
UltrAuFoil® R 1.2/1.3 holey gold film grid	Ted Pella, Inc.	Cat# 688-300-AU-50
PELCO easiGlow™ Glow Discharge Cleaning System	Ted Pella, Inc.	Cat# 91000
FEI Vitrobot Mark IV	Thermo Scientific	N/A
FEI Titan Krios	Thermo Scientific	N/A
3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO)	Sigma-Aldrich	Cat# C3649

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rebecca Voorhees (voorhees@caltech.edu).

Materials availability

All unique/stable reagents generated in this study are available from the [lead contact](#) with a completed Material Transfer Agreement.

Data and code availability

- The genome wide FACS screens datasets are provided as [Tables S1](#), [S2](#), and [S3](#). Mass spectrometry dataset of the EMC is provided as [Table S4](#) 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 also listed in the [key resources table](#)). The structure of the human BOS complex in GDN is deposited at EMDB: 45293. The structure of the human truncated BOS complex in GDN is deposited at PDB: 9C7U and EMDB: 45294. The structure of the human BOS:human EMC complex in GDN is deposited at PDB:9C7V and EMDB: 45295. Several figures in this paper make use of existing structures (PDB codes are listed in the [key resources table](#)).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**Cell lines**

All cell lines used in this study are listed in the [key resources table](#).

K562 cells containing KRAB-BFP-dCas9³⁹ 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 machinery⁸⁰ were cultured in DMEM F12 medium supplemented with 10% FBS and 2 mM glutamine. K562, HEK 293T, Expi293, 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.

METHOD DETAILS**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), adrenocorticotrophic 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- α -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.⁸⁰ 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.^{47,49}

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.^{96,97} Here, the GFP11 tag (RDHMLVHEYVNAAGIT) 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, all substrates except for SEC61 β contain full GFP or RFP fusions. The SEC61 β reporter has been previously described.⁴⁸ Briefly, the TMD and flanking regions were inserted downstream of the first 70 residues of the flexible cytosolic domain of SEC61 β . At the C-terminal of Sec61 β , the GFP11 tag (RDHMLVHEYVNAAGIT) was inserted, separated by a 2X-GS linker. To express GFP1-10 in the ER lumen, the human calreticulin signal sequence (MLLSVPLLLGLLGLAVA) was appended to the N-terminal of GFP1-10-KDEL as previously described.^{27,47}

Programmed dual sgRNA guide plasmids were used in assays involving depletion of two genes.⁹⁸ The sgRNA protospacer sequences used to generate dual guide plasmids are listed in Table S5.

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 S5.

The siRNAs used in this study are listed in Table S5.

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.⁹⁹

All plasmids are available upon request.

The antibodies used in this study are listed in the [key resources table](#).

Cell line construction

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.^{39,40} For AGTR2 and Sec61 β , 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.⁴⁸ 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 μ g/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'-CAAGCAGAAGACGGCATACGAGATC

GACTCGGTGCCACTTTTTTC). 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.⁸ 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'-GTGTGTTTTGAGACTATAAGTATCCCTTGGAGAACACCTTGTG). Analysis of the sequencing was performed using the pipeline in <https://github.com/mhorlbeck/ScreenProcessing>.⁴⁰ 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 Sec61 β), the sgRNA phenotypes were averaged. To calculate the discriminant scores used in Figure 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 μ g/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(Δ 1-9lg)-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 spininfected with lentivirus containing dual sgRNAs targeting two genes of interest or a non-targeting control. Briefly, 250,000 cells were mixed with 200 μ l of lentivirus and RPMI medium in the presence of 8 μ g/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 μ g/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 μ l lentiviral supernatant and 8 μ g/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 2B, 2C, 3D–3F, 4C, 6A, 6C–6F, S6A, S6C–S6F, S3D, S3E, and S7A). Versions of GPCRs with a signal sequence (Figures 4C, 6C, and S6C) contained N-terminal fusions of the pre-prolactin signal sequence (KGSSQKGSRLLLLVSNLLLCQGVV) 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 (Figures 4C and S3E). Sec61 β , 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 Sec61 β (SPGLKVGPPVPLVMSLLFIASVFMLHIWGKYT), SQS (SRSHYSPIYLSFVMLLAALSWQYLTLSQVTE), and VAMP2 (KTGKLNKMMIILGVICAILIIIVYFTGSR) were fused directly to RFP, as described before.^{19,27,48} 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 described³¹ (Figures 6C–6F, S6C, and S7A).

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 μ l lentiviral supernatant of rescue construct(s) and 8 μ g/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 μ l lentiviral supernatant of reporter construct(s) in presence of 8 μ g/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 μ l 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.^{63,64} HEK 293T (WT or TMC01 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.²² HEK293T cells (WT, NCLN KO, TMC01 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 cComplete 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 nucleated. CaCl₂ (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 placed on ice. Nucleated 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 nucleated rabbit reticulocyte lysate (RRL) supplemented with human derived rough ER microsomes, as previously described.^{77,95} 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,⁷⁷ 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 ³⁵S-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 by incubating equimolar amounts of ALFA nanobody and maleimide-activated HRP (Thermo Scientific) for 1 hour at room temperature, as previously described.¹⁰⁰

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.⁵² Cells were isolated and solubilized in solubilization buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 2 mM MgAc₂, 1x cCompleteTM EDTA-free Protease Inhibitor Cocktail [Roche], 1% [w/v] glyco-diosgenin [GDN; Anatrache], 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, PierceTM 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

MgAc₂, 0.01% GDN, 1 mM DTT) for 20 min. at 4°C as described previously⁷⁸. 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).⁷⁹ 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 (Figures 4D–4F). 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 4°C. 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 37°C, and alkylated by incubation with 12 mM 2-chloro-acetamide (CAA) (MP Biomedicals) for 15 minutes at 37°C. Samples were digested with 2ng/μl Lysyl Endoproteinase (Lys-C) (Wako Chemicals) for 4 hours at 37°C. Samples were diluted 4-fold with 50 mM HEPES, pH 8.0 and CaCl₂ was added to 1 mM final concentration. The samples were then digested with 0.6 ng/μl Trypsin (Thermo Scientific) for 18 hours at 37°C. 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 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.^{101,102} 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 S4.

Photo-crosslinking and immunoprecipitation

To perform site-specific photo-crosslinking of the BOS complex, (Figure S5G), 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.^{81,82} 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 tRNAPyICUA (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 MgAc₂, 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 total 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(Δ 1–9lg)-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.^{27,78} 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 MgAc₂, 1x cOmpleteTM 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 μ l resuspended PierceTM 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 MgAc₂, 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 μ g for every 80 μ l beads. This immobilization was followed by incubation of beads with 50 mM HEPES/KOH pH 7.5 containing 100 μ M 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- μ l 30K MWCO concentrator (Millipore-Sigma).

Grid preparation and data collection

For BOS (fNOMO) sample, 3 μ L 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 easiGlowTM (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 SerialEM⁹² and recorded at 105k magnification with a defocus range of -3.0 to -1.0 μ m and total exposure dose of 60 e⁻/Å² 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 Figures S4A–S4C, respectively. Data processing was carried out using cryoSPARC v3.2-4.2.1.⁸⁶ 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 $-60 \text{ \AA}^2$ 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-initio 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 Chimera¹⁰³ 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).⁸⁹ 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 (Figure 5A) and combined in COOT.^{88,87}

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.^{88,87} The final model was iteratively subjected to *phenix.real_space_refinement*^{90,91} 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.^{88,87} The model was iteratively subjected to *phenix.real_space_refinement*^{90,91} with rigid body and secondary structure restraints.

CryoEM data collection, refinement, and validation statistics are reported in Table 1. Final models were evaluated with MolProbity. All figures in this study were generated with PyMOL (www.pymol.org) and ChimeraX.^{84,85}

QUANTIFICATION AND STATISTICAL ANALYSIS

For the CRISPRi screens, analysis of the sequencing was performed using the pipeline in <https://github.com/mhorlbeck/ScreenProcessing>.⁴⁰ 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 Figure 3, each gene's phenotype score was multiplied by its Mann-Whitney p-value.