

Supplementary Discussion

Interactions between translocon subcomplexes and the ribosome

Each multipass translocon subcomplex makes contacts with the ribosome. Four contacts between the PAT complex subunit CCDC47 and the ribosome were described earlier¹ and are also seen in our structures: (i) two helices of the α/β -sandwich (residues 311 and 392-396) meet the tip of a 28S rRNA hairpin; (ii) the start of the C-terminal helical extension (residues 400-401) meets the backbone of a 28S rRNA loop; (iii) the site where the C-terminal extension bends (residues 414-433) contacts uL22 and 28S rRNA; (iv) the final C-terminal helix meets eL31 and the 28S rRNA surrounding the exit tunnel. Our structure additionally shows that CCDC47 contacts a flexible extension of ribosomal protein eL6 (residues 234-235) via one helix from CCDC47's α/β -sandwich domain (residues 319 and 323). The second cytoplasmic loop (residues 92-93) of the BOS complex subunit TMEM147 contacts a protuberant hairpin in 5.8S rRNA. GEL complex subunit TMCO1 contacts the ribosome via two sites: one face of the N-terminal tip of its first cytoplasmic helix (residues 38 and 42) contacts uL24, while that of its second cytoplasmic helix (residues 64, 68, and 71) contacts uL24 and 28S rRNA.

Two of the MPT subcomplexes, BOS and PAT, also make contacts with Sec61. PAT subunit CCDC47's latch helices (residues 434, 453, and 457) contact Sec61's N-half at two sites (residues 23, 106-7), while BOS subunit TMEM147's second TMD and adjacent luminal loop (residues 50-59) contact Sec61's hinge loop (residues 220-246), as previously described¹. Finally, the MPT subcomplexes make a few limited contacts with one another. The globular cytoplasmic domain of CCDC47 (residues 284-286) contacts the tip of TMCO1's helical hairpin (residues 58-60), and the globular luminal domain of Nicalin (residues 37, 40) contacts the luminal loops of TMCO1 (residues 110-112, 132, 134). The limited extent of these interactions is consistent with observations that the subcomplexes do not copurify in the absence of ribosomes nor depend on one another for expression², but can partially influence one another's recruitment to the ribosome (Extended Data Fig. 7g,h).

Potential functions of the PAT, GEL, and BOS complexes

Our working model is that the PAT complex is a chaperone (i.e., holdase) for semi-hydrophilic TMDs and the GEL complex is a transporter that facilitates TMD insertion (also often termed an insertase). A chaperone function for the PAT complex is supported by its selective interaction with TMDs of partial hydrophilic character without influencing their insertion³ (Extended Data Fig. 7h), and the interaction of such TMDs with other (presumably irrelevant) ER proteins when the PAT complex is absent (Extended Data Figure 7h). Thus, it seems to hold TMDs and prevent off-pathway interactions, thereby facilitating overall efficiency of substrate maturation³. The GEL complex belongs to the Oxa1 family whose other members are established to be transporters for TMD insertion^{4,5}. Consistent with this role, loss of this complex impairs insertion of TMD2 of Rhodopsin (Extended Data Figure 7f), with which it physically interacts at a point just after insertion (Fig. 4d and Extended Data Figure 7e). A biochemical function for the BOS complex is currently undefined, but it has been noted to be structurally related to parts of the gamma secretase complex¹. The BOS complex may participate in stabilizing the multipass translocon given that the PAT and GEL complexes seem to be recovered slightly less effectively when the BOS complex is deleted (Extended Data Figure 7g). BOS could potentially influence Sec61 function by constraining one of Sec61's luminal loops to a particular configuration. Finally, BOS could contribute to a protected environment for substrate folding. At present, we have not observed direct interactions between substrate and components of the BOS complex using photo-crosslinking probes located in each of the first three membrane-inserted TMDs of Rhodopsin (Fig. 4d and Extended Data Figure 7e).

Location of the substrate in the multipass translocon

Site-specific photo-crosslinking via probes in TMD1, TMD2, and TMD3 of the Rho-4TMD intermediate shows that each of them is in a different location (Fig. 4 and Extended Data

Figure 7e). TMD1 is adjacent to Asterix of the PAT complex, TMD2 is adjacent to TMCO1 of the GEL complex, and TMD3 is on the hinge side of the Sec61 complex just outside the ribosome exit tunnel. Of these, only TMD3 was observed in the cryo-EM reconstruction, possibly because it is partially constrained by tethering by a segment of nascent polypeptide (~40 amino acids) just long enough to reach the peptidyl transferase centre. We speculate the other parts of the substrate are not seen for at least two reasons. First, it seems that the membrane domains of the PAT and GEL complexes, which interact with the non-visualised TMDs, are relatively flexible or conformationally heterogeneous and hence observed only at moderate resolution. Second, each factor probably binds substrate dynamically and in a heterogeneous set of conformations. The combination of flexibility and heterogeneity would preclude visualisation of TMD1 and TMD2, even though their presence and their local environment can be verified biochemically. This flexibility is probably because the substrate is a folding intermediate, which will be unlikely to have a stable defined conformation.

CCDC47 narrows the mouth of the exit tunnel

A striking feature of the C-terminal helix of CCDC47 is its invasion of the ribosome exit tunnel via residues 472-476, effectively narrowing the mouth of the tunnel (Extended Data Fig. 3e,f). An additional seven residues are not visualised in our structure and are either unstructured or heterogeneous. The functional relevance of CCDC47 narrowing the mouth of the exit tunnel and its potential interactions with an emerging nascent chain remain to be experimentally explored. Nonetheless, an interplay between the nascent chain and CCDC47's latch helices could provide an explanation for how substrates can alternate between using the multipass translocon and Sec61 (see model section below).

The TRAP complex in the multipass translocon

The accompanying study² biochemically documents that the TRAP complex is present in the multipass translocon. In our maps, we observe a luminal density similar in shape and size to those previously assigned to the luminal domain of TRAP (Extended Data Fig. 2, 3). However, this putative TRAP density is in a slightly different position relative to the ribosome from where it sits in previously reported TRAP complexes^{6,7}, and the cytoplasmic ribosome-binding helices of TRAP γ are not visible. Our tentative interpretation is, therefore, that TRAP is present in our maps, as one would predict from the biochemical data², but displaced from its canonical site, perhaps by a clash with the luminal domain of the BOS complex, which the TRAP density abuts (Extended Data Fig. 2a).

Model for multipass membrane protein biogenesis

The early steps of multipass membrane protein biogenesis, up to the insertion of a TMD in the N_{exo} topology (that is, with the N-terminal flanking domain facing the exoplasmic side), are thought to be the same as for single-pass membrane proteins. Proteins with a signal sequence will use SRP for targeting⁸, engage the Sec61 lateral gate⁹, and use Sec61's central channel to translocate the downstream segment of polypeptide across the ER¹⁰. The first TMD will therefore enter the Sec61 channel in the N_{exo} topology and pass into the membrane via the lateral gate. Proteins whose first hydrophobic element is an N_{exo} TMD will use this TMD for SRP-mediated targeting and insert in an EMC-dependent and Sec61-independent reaction^{11,12}, with the RNC eventually docking at the Sec61 complex¹³. Proteins whose first hydrophobic element is an N_{cyt} TMD will target via SRP and can engage the Sec61 lateral gate similar to a signal sequence. The polypeptide downstream of the N_{cyt} TMD will translocate through Sec61's channel until the next TMD enters the Sec61 channel in the N_{exo} topology and passes into the membrane via Sec61's lateral gate.

At the end of the early steps, the TMD(s) cannot yet engage the PAT complex because steric constraints imposed by CCDC47's large cytosolic domain would preclude it from approaching the ribosome exit tunnel. Hence, OST can sample its docking site behind Sec61, allowing glycosylation of the nascent chain⁷. Notably, most co-translational glycosylation sites of multipass membrane proteins are located in the translocated domain flanking the first TMD¹⁴. Further elongation allows the early TMD(s) to sample a larger radius

around the exit tunnel, eventually reaching far enough to bind Asterix (Fig. 1b; Extended Data Fig. 1f). This binding is favoured by hydrophilic amino acids in a TMD (Fig. 1c), a feature of nearly all TMDs in multipass membrane proteins.

Substrate binding stabilizes the PAT complex and other multipass translocon components at their respective sites behind the Sec61 complex, precluding OST re-binding to its site². Binding of an early TMD by Asterix redirects the downstream nascent chain in a direction away from Sec61's lateral gate. Furthermore, access to Sec61's cytosolic vestibule is partially blocked by CCDC47's latch helices, which also constrain Sec61 to a closed configuration (Fig. 3c). The next TMD that emerges from the ribosome would therefore be located between the ribosome and the membrane patch partially enclosed by the multipass translocon. The tether downstream of this TMD is initially too short for its insertion but becomes long enough when additional polypeptide emerges from the ribosome.

If this additional polypeptide comprises a short hydrophilic segment and another TMD, the two TMDs could insert as a unit concomitant with translocation of the intervening loop. Insertion seems to occur independently of the Sec61 lateral gate (Fig. 4e,f; Extended Data Fig. 9, 10) into the horseshoe-shaped cavity formed by the multipass translocon components, where such an intermediate is observed by a combination of cryo-EM (Fig. 4c) and crosslinking (Fig. 4d). The interior surface of the multipass translocon is both well-conserved and notably amphiphilic in several areas (Extended Data Fig. 8d). This property, particularly of the GEL complex, may cause membrane thinning in the multipass translocon cavity similar to what happens adjacent to other Oxa1 family members^{15,16}. A thinned membrane and partially hydrophilic surfaces could facilitate translocation of hydrophilic polypeptide, a key barrier to TMD insertion.

In this manner, the multipass translocon, and particularly the GEL complex, would facilitate insertion of internal TMD pairs separated by relatively short translocated loops (typically less than ~50 amino acids). Notably, the vast majority of translocated loops between internal TMD pairs in multipass proteins are short, indicating that many membrane proteins can insert independently of Sec61's lateral gate. Indeed, only a handful of the ~800 GPCRs in the human genome have a translocated internal loop longer than 50 amino acids. Membrane distortion in the area between Asterix and Sec61 might be sufficient to permit loop translocation in some organisms, such as *S. cerevisiae*, that contain the PAT complex but seem to lack the other multipass translocon complexes. Alternatively, these species might use a different Oxa1 family member such as the widely conserved EMC to mediate insertion of internal TMD pairs by a similar mechanism.

During biogenesis, a substrate's TMDs will still contain unsatisfied hydrophilic residues, allowing it to engage the multipass translocon and stabilise its assembled state. Subsequent TMD pairs could therefore use the multipass translocon for insertion similar to the previous pair. Insertion of TMD1 via EMC^{11,12} and all downstream TMDs via the multipass translocon would explain why multipass proteins beginning with an N_{exo} signal anchor seem to be refractory to inhibitors of Sec61's lateral gate (Fig. 4e,f; Extended Data Figs. 9 and 10). Accumulation of multiple TMDs in the protein-lined cavity behind Sec61 also explains why insertion intermediates with up to six TMDs can remain sensitive to urea extraction¹⁷ and how TMDs can shift their positions relative to Sec61 during biogenesis¹⁸.

The multipass translocon cavity seems sufficiently large to accommodate the seven TMDs of a GPCR and may provide a semi-protected environment for membrane protein folding. Successful folding would shield the hydrophilic residues of a substrate's TMDs, thereby diminishing interactions with the amphiphilic multipass translocon components and allowing substrate release into the bulk membrane. Release could occur via either the open side of the multipass translocon between Asterix and Sec61 or after translocon disassembly.

Membrane proteins with more than seven TMDs could be accommodated if early TMDs move out of the cavity to create space for the insertion of downstream TMDs. This seems plausible for most large multipass proteins because they are generally evolved from the fusion of smaller ancestral units^{19,20}. For example, the 12-TMD family of transporters is a

pseudosymmetric assembly of two 6-TMD modules. The first six TMDs might partially assemble in the multipass translocon cavity before egressing into the surrounding membrane. Consistent with this idea, this is the point at which resistance to urea extraction, an indicator of membrane integration, is first observed for a model 12-TMD protein¹⁷.

Multipass proteins with large translocated domains between internal TMDs would need to re-engage the Sec61 lateral gate and use its channel for translocation (Extended Data Fig. 9e). It is currently unknown how Sec61 occlusion by CCDC47 is overcome. Because part of CCDC47 lines the mouth of the exit tunnel (Extended Data Fig. 3f), it can interact with the nascent chain. It is therefore plausible that some nascent chain feature(s) characteristic of long translocated domains displace CCDC47 from the exit tunnel. This could destabilise the final latch helix from its position, thereby removing a major impediment to both Sec61 access by the nascent chain and to Sec61 opening (Fig. 3c). The mechanism of switching between using the multipass translocon for insertion of closely spaced TMDs versus the Sec61 channel for translocation of long loops is an important area for study.

References for Supplementary Discussion

1. McGilvray, P. T. *et al.* An ER translocon for multi-pass membrane protein biogenesis. *Elife* **9**, e56889 (2020).
2. Sundaram, A. *et al.* Substrate-driven assembly of a modular translocon for multipass membrane protein topogenesis. *Submitted* (2022).
3. Chitwood, P. J. & Hegde, R. S. An intramembrane chaperone complex facilitates membrane protein biogenesis. *Nature* **584**, 630–634 (2020).
4. Hennon, S. W., Soman, R., Zhu, L. & Dalbey, R. E. YidC/Alb3/Oxa1 family of insertases. *Journal of Biological Chemistry* **290**, 14866–14874 (2015).
5. McDowell, M. A., Heimes, M. & Sinning, I. Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily. *Nat. Struct. Mol. Biol.* **28**, 234–239 (2021).
6. Pfeffer, S. *et al.* Dissecting the molecular organization of the translocon-associated protein complex. *Nat. Commun.* **8**, (2017).
7. Braunger, K. *et al.* Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* **360**, 215–219 (2018).
8. Akopian, D., Shen, K., Zhang, X. & Shan, S. O. Signal recognition particle: An essential protein-targeting machine. *Annu. Rev. Biochem.* **82**, 693–721 (2013).
9. Voorhees, R. M. & Hegde, R. S. Structure of the Sec61 channel opened by a signal sequence. *Science* **351**, 88–91 (2016).
10. Rapoport, T. A., Li, L. & Park, E. Structural and Mechanistic Insights into Protein Translocation. *Annu. Rev. Cell Dev. Biol.* **33**, 369–390 (2017).
11. Chitwood, P. J., Juszkiwicz, S., Guna, A., Shao, S. & Hegde, R. S. EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* **175**, 1507-1519.e16 (2018).
12. O'Keefe, S. *et al.* An alternative pathway for membrane protein biogenesis at the endoplasmic reticulum. *Commun. Biol.* **4**, 828 (2021).
13. High, S. *et al.* Sec61p is adjacent to nascent type I and type II signal-anchor proteins during their membrane insertion. *J. Cell Biol.* **121**, 743–50 (1993).
14. Cherepanova, N. A., Venev, S. V., Leszyk, J. D., Shaffer, S. A. & Gilmore, R. Quantitative glycoproteomics reveals new classes of STT3A- And STT3B-dependent N-glycosylation sites. *J. Cell Biol.* **218**, 2782–2796 (2019).

15. Wu, X. & Rapoport, T. A. Translocation of Proteins through a Distorted Lipid Bilayer. *Trends Cell Biol.* **31**, 473–484 (2021).
16. Pleiner, T. *et al.* Structural basis for membrane insertion by the human ER membrane protein complex. *Science (80-.).* **369**, 433–436 (2020).
17. Borel, A. C. & Simon, S. M. Biogenesis of polytopic membrane proteins: Membrane segments assemble within translocation channels prior to membrane integration. *Cell* **85**, 379–389 (1996).
18. Sadlish, H., Pitonzo, D., Johnson, A. E. & Skach, W. R. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispinning membrane protein. *Nat. Struct. Mol. Biol.* **12**, 870–8 (2005).
19. Forrest, L. R. Structural Symmetry in Membrane Proteins. *Annu. Rev. Biophys.* **44**, 311–337 (2015).
20. von Heijne, G. Membrane-protein topology. *Nature Reviews Molecular Cell Biology* **7**, 909–918 (2006).