The mechanisms of integral membrane protein biogenesis

Ramanujan S. Hegde[1] ✉ *and Robert J. Keenan*[2] ✉

Abstract | Roughly one quarter of all genes code for integral membrane proteins that are inserted into the plasma membrane of prokaryotes or the endoplasmic reticulum membrane of eukaryotes. Multiple pathways are used for the targeting and insertion of membrane proteins on the basis of their topological and biophysical characteristics. Multipass membrane proteins span the membrane multiple times and face the additional challenges of intramembrane folding. In many cases, integral membrane proteins require assembly with other proteins to form multi-subunit membrane protein complexes. Recent biochemical and structural analyses have provided considerable clarity regarding the molecular basis of membrane protein targeting and insertion, with tantalizing new insights into the poorly understood processes of multipass membrane protein biogenesis and multi-subunit protein complex assembly.

A fundamental feature of all life is a plasma membrane that separates the intracellular and extracellular environments. Transport of metabolites and communication of information across otherwise impermeable membranes is mediated by integral membrane proteins. These membrane-spanning proteins represent ~25% of protein-coding genes in all organisms and they serve numerous crucial functions, including ion and nutrient transport, signalling, pathogenesis, defence and adhesion[1]. Biogenesis of membrane-spanning proteins is therefore one of the most ancient biological processes, and the core machineries that mediate these processes are exceptionally broadly conserved.

In bacteria and archaea, which typically lack intracellular membrane compartments, membrane proteins are inserted directly into the plasma membrane where they function[2]. Eukaryotes insert the vast majority of their membrane proteins into the endoplasmic reticulum (ER) membrane, where these proteins fold and assemble before travelling to their final destination[3]. Because the eukaryotic ER is evolutionarily derived from the prokaryotic plasma membrane[4], the respective insertion machineries are related and share key mechanistic principles. The machineries for insertion into mitochondrial, chloroplast and peroxisomal membranes are reviewed elsewhere[5–9].

The membrane-spanning segments of a protein are typically α-helical, allowing the hydrophilic amide backbone of the protein to be shielded by hydrogen bonding from the hydrophobic membrane interior[10]. A second strategy to accomplish the same goal is by assembly of a β-barrel. This class of membrane proteins, found in the outer membranes of bacteria, mitochondria and plastids, is inserted by a unique machinery that is reviewed elsewhere[11,12]. Here we discuss α-helical membrane protein biogenesis at the eukaryotic ER and evolutionarily related prokaryotic plasma membrane.

Membrane protein biogenesis can be divided into four processes (FIG. 1a). The first is targeting, the process by which a nascent protein is delivered to the membrane where it will be inserted. The second is membrane insertion of transmembrane domains (TMDs) in the appropriate topology, retaining some parts of the protein in the cytosol, while translocating other parts across the membrane. In addition to these universal steps applicable to even the simplest membrane proteins, most polypeptides require additional folding steps in the membrane, and many are assembled with obligate interaction partners. Although these steps of membrane protein biogenesis are typically studied separately, they are intimately linked and often occur simultaneously as the protein is being synthesized.

A full understanding of membrane protein biogenesis requires an appreciation of membrane protein diversity. The human genome contains ~5,000 integral membrane proteins, whose ~20,000 TMDs differ widely in sequence, biophysical properties, location or locations within the protein and topology (FIG. 1b). A major consequence of this diversity is that the machinery tasked with membrane protein biogenesis must be highly accommodating. This is accomplished by different pathways for different classes of substrates (also termed 'clients') based on topology, biophysical property or another distinctive feature. Thus, although the multiple pathways presented in each following section can seem bewildering, their coexistence can be rationalized by the diversity of membrane protein substrates.

[1]*Cell Biology Division, MRC Laboratory of Molecular Biology, Cambridge, UK.*

[2]*Gordon Center for Integrative Science, The University of Chicago, Chicago, IL, USA.*

✉*e-mail: rhegde@mrc-lmb.cam.ac.uk; bkeenan@uchicago.edu*

https://doi.org/10.1038/s41580-021-00413-2

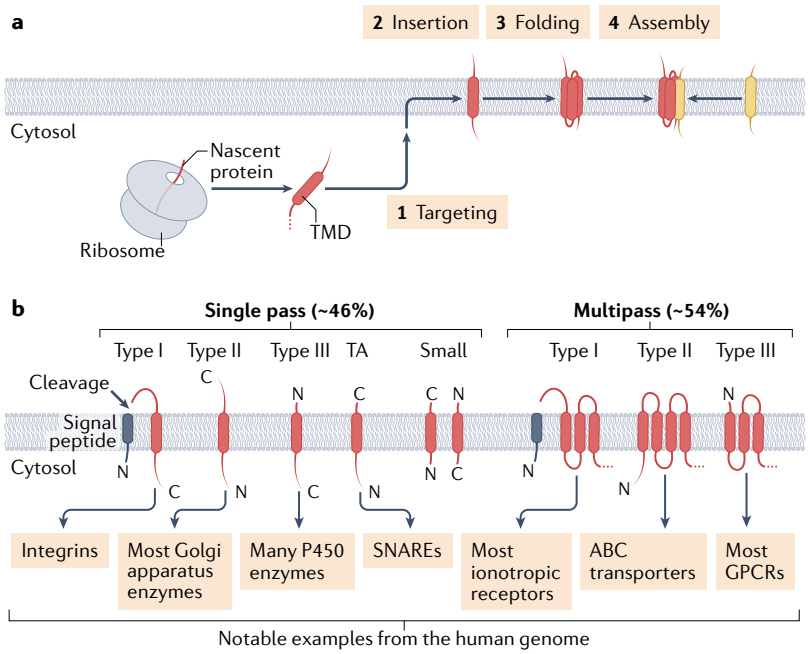


Fig. 1 | **Overview of integral membrane protein biogenesis. a** | The four major steps involved in membrane protein biogenesis. In this figure and subsequent figures, tapered extensions that flank transmembrane domains (TMDs) indicate soluble regions of indeterminate length. When the flanking segment is short, such as the translocated segment of tail-anchored (TA) membrane proteins, a non-tapered line is used. An ellipsis at the end of a polypeptide is used to indicate further polypeptide that contains additional TMDs. **b** | The major classes of integral membrane proteins are indicated, with prominent examples from human cells listed below. ABC, ATP-binding cassette; C, carboxy terminus; GPCR, G-protein-coupled receptor; N, amino terminus.

In this Review, we discuss the four basic steps of membrane protein biogenesis, highlighting the different mechanisms tailored for different types of substrates. Much of our focus is on the mechanistic basis of membrane protein insertion and folding. These areas have recently seen the discovery of new machinery, mechanistic insights from structural and bioinformatics analysis, and the emergence of new concepts. Protein targeting is a comparatively mature area, and thus is summarized in less detail, with references to other, focused reviews. The assembly of membrane protein complexes is a poorly studied problem, so its discussion is necessarily more speculative. In each section, we attempt to synthesize knowledge into a few major concepts and highlight the most pressing directions for future work.

Targeting of membrane proteins

Membrane protein targeting to the eukaryotic ER or bacterial plasma membrane is generally mediated by the most amino-terminal (N-terminal) hydrophobic domain within the substrate. The targeting sequence can either be a cleavable signal peptide, typically at the N terminus, or the first TMD anywhere within the polypeptide (FIG. 1b). The hydrophobic domain of signal peptides is ~7–9 amino acids long[13], whereas TMDs are ~15–25 amino acids long; both are widely variable in sequence and hydrophobicity[14]. This large diversity in targeting sequence location and biophysical properties means that a single targeting pathway or mechanism cannot accommodate all membrane proteins.

Three established targeting pathways are used depending on the position and hydrophobicity of a substrate's targeting sequence (FIG. 2a). A targeting sequence located at least ~65 amino acids from the substrate's carboxy terminus (C terminus) is required for co-translational targeting by the signal recognition particle (SRP)[15–17]. This distance constraint is because the site of targeting sequence recognition by SRP — which occurs at the mouth of the ribosome exit tunnel[18–20] — is positioned ~35 amino acids from the peptidyltransferase centre inside the ribosome (FIG. 2b). After recognition, SRP-mediated targeting to the ER takes ~5–7 seconds[21], during which time another ~30 amino acids can be synthesized. Co-translational targeting occurs only if translation does not terminate during this period, explaining why SRP requires targeting signals to be relatively far from the C terminus.

Membrane proteins whose sole targeting sequence is a TMD closer than ~65 amino acids from the C terminus are termed 'tail-anchored (TA) membrane proteins'[22], and they are targeted post-translationally using either general or specialized cytosolic chaperones. The TMDs of TA proteins differ widely in hydrophobicity (FIG. 2c), a key feature that determines their mechanism of targeting[23]. Those of high hydrophobicity, such as vesicle-associated membrane protein 2 (VAMP2), are targeted by the guided entry of TA protein (GET) pathway. Those of low hydrophobicity, such as the lipid biosynthetic enzyme squalene synthase (SQS), instead use cytosolic chaperones and the ER membrane protein complex (EMC) for their biogenesis. Most TA proteins (one example being SEC61B) can probably use either route for membrane insertion. In the following sections we discuss the targeting pathways that utilize SRP, the GET pathway and EMC in turn. Other poorly studied potential targeting routes are discussed in BOX 1.

Co-translational targeting by SRP. During co-translational targeting, hydrophobic targeting sequences are recognized by SRP at the mouth of the ribosome exit tunnel and delivered to the SRP receptor at the ER membrane (FIG. 2d). The cycle of substrate recognition, targeting to SRP receptor and recycling of SRP back to the cytosol is regulated through the cycle of GTP binding and hydrolysis by two universally conserved GTPases, one in the SRP54 subunit of SRP and the other in the α-subunit of SRP receptor. The structures and molecular details of the SRP system have been extensively reviewed in detail elsewhere[15–17]. Here we consider comparatively recent new insights into the basis of SRP substrate range and specificity.

The minimum requirement for SRP recognition is a predominantly hydrophobic sequence of approximately seven amino acids. Structures of the substrate-binding domain of SRP54 without and with substrate in prokaryotes[24–26] and eukaryotes[20] illustrate a conserved hydrophobic groove that explains the preference of SRP for hydrophobic sequences in its clients. Cryogenic electron microscopy (cryo-EM) analyses of native mammalian SRP just before and immediately after substrate engagement suggest that the substrate-binding groove is normally autoinhibited by a C-terminal amphipathic

Targeting sequence
The sequence element in a protein that directs its delivery to a specific membrane in the cell. For membrane proteins, the targeting sequence is typically a cleavable amino-terminal signal peptide or the first transmembrane domain.

Signal peptide
A targeting sequence that is found at the amino terminus of secretory proteins and some membrane proteins. After they have served their targeting function, signal peptides are cleaved off by an enzyme called 'signal peptidase'.

Tail-anchored (TA) membrane proteins
Membrane proteins whose only transmembrane domain lies within ~65 amino acids of the carboxy terminus and are oriented with the amino terminus facing the cytosol. These are sometimes called 'type IV membrane proteins'.

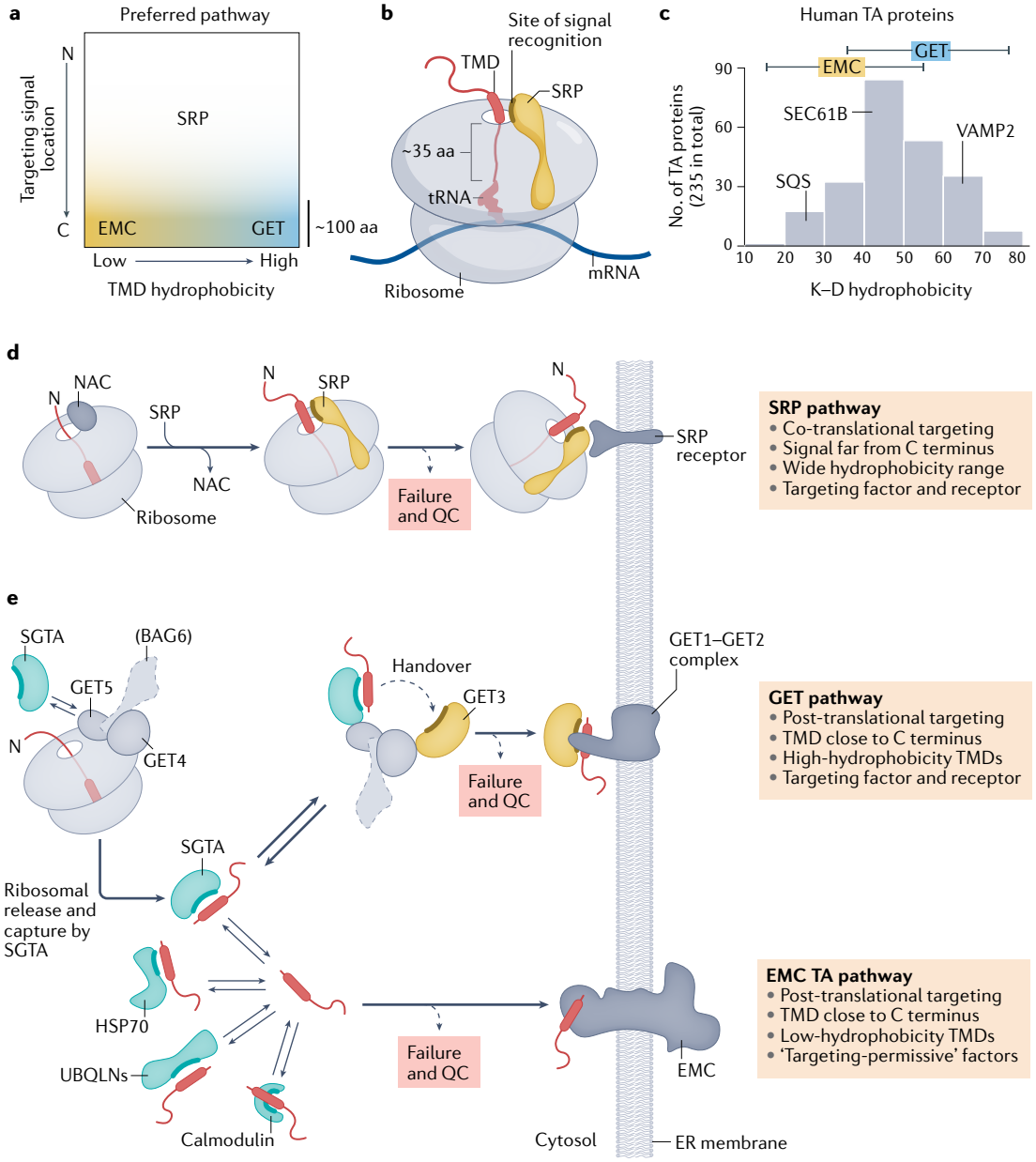


Fig. 2 | **Membrane protein targeting to the endoplasmic reticulum. a** | Both the position and the hydrophobicity of a targeting sequence influence the route that a membrane protein uses for insertion into the endoplasmic reticulum (ER) membrane. The three routes indicated are the signal recognition particle (SRP) targeting pathway, the insertion pathway mediated by the ER membrane protein complex (EMC) and the guided entry of tail-anchored protein (GET) insertion pathway. **b** | The position of SRP on a translating ribosome. SRP recognizes the targeting sequence at the mouth of the ribosome exit tunnel. The site of recognition is ~35 amino acids (aa) from where the nascent polypeptide is attached to the tRNA at the peptidyltransferase centre of the ribosome. **c** | Histogram of all 235 predicted human tail-anchored (TA) proteins plotted by the hydrophobicity of their transmembrane domains (TMDs) calculated using the Kyte–Doolittle (K–D) scale[179]. Squalene synthase (SQS), vesicle-associated membrane protein 2 (VAMP2) and SEC61B are examples of TA proteins — shown in the bins where they are located — that are inserted by the EMC pathway, the GET pathway or either pathway, respectively. **d** | The co-translational SRP pathway for delivering translating ribosomes to the ER membrane. Initially, the nascent polypeptide-associated complex (NAC) prevents SRP from binding to the ribosome. NAC is exchanged for SRP when a hydrophobic targeting sequence, either a signal peptide or a TMD, emerges from the ribosome. SRP then engages the SRP receptor at the ER to mediate targeting. **e** | Post-translational targeting of TA proteins to the ER membrane. TA proteins are initially captured near the ribosome surface by the chaperone SGTA, whose recruitment to this site is facilitated by the GET4–GET5 complex (which in mammals also contains BAG6). If the TMD of the TA protein is of high hydrophobicity, it is transferred to GET3 in a handover reaction mediated by the GET4-GET5 complex. GET3 then targets to a receptor comprising the GET1–GET2 complex. If the TMD is of low hydrophobicity, it is not transferred to GET3. Instead, it is kept soluble in the cytosol by cycles of binding and release from any of several chaperones, including SGTA, HSP70, ubiquilin family proteins (UBQLNs) and calmodulin. The TMD then engages the cytosolic domain of EMC. Quality control (QC) pathways monitor failures during co-translational and post-translational targeting (BOX 2). C, carboxy terminus; N, amino terminus.

Box 1 | **Other potential targeting pathways**

Most membrane proteins in bacteria are targeted co-translationally by the signal recognition particle (SRP) pathway[35] or post-translationally using chaperones[63,64]. However, some membrane proteins seem to be co-translationally targeted directly to SecY by the bacterium-specific ATPase SecA[180,181]. The membrane protein is then presumably inserted via the lateral gate in SecY, with one flanking region being translocated through the SecY channel. The features that determine whether a membrane protein is recognized by SRP or SecA probably involve a combination of the transmembrane domain (TMD) and its flanking regions[182]. Because SecA has a binding site for the signal peptides of secreted proteins[183], it is possible that low-hydrophobicity TMDs that resemble signal peptides are recognized similarly. As with secreted proteins[184], SecA might use repeated cycles of ATP binding and hydrolysis to push the translocated portion of a membrane protein through the central channel of SecY. Precisely how SecA-bound ribosomes are transferred to SecY remains unclear. The considerable overlap in the ribosome-binding site of SecA and SecY suggests that handover may occur via a concerted mechanism similar to that observed with SRP and SecY[182]. As in eukaryotes, it seems that multiple targeting pathways are needed to accommodate the topological and biophysical diversity of the bacterial membrane proteome.

In yeast, some secretory or lipid-anchored proteins that normally engage SRP in wild-type cells can nonetheless be targeted and translocated in SRP-lacking cells[185]. Use of a whole-genome high-content localization screen of one such protein led to the definition of an SRP-independent (SND) targeting pathway composed of three interacting factors: Snd1 in the cytosol and two endoplasmic reticulum membrane proteins Snd2 (TMEM208 in humans) and Snd3 (REF.[186]). Primarily on the basis of synthetic genetic interactions, the SND pathway was proposed to target both SRP-pathway and guided entry of tail-anchored protein (GET)-pathway substrates to the Sec complex (composed in yeast of the Sec61 complex, Sec62, Sec63, Sec71 and Sec72) for translocation and membrane insertion. The membrane protein substrates of the SND pathway were proposed to be those whose first TMD is far from the amino terminus, but not so distal as to be tail-anchored. Because such TMDs would emerge from a translating ribosome, they should be recognized by SRP. It is mechanistically unclear why they seem to need a separate pathway. Additional work is required to assign specific roles to each SND factor, determine the mechanism or mechanisms of substrate engagement by this pathway, determine whether the pathway operates co-translationally or post-translationally, and clarify the role of the human SND2 orthologue[186,187].

The substrate ranges covered by the SRP, GET and ER membrane protein complex (EMC) targeting pathways are challenging to define precisely. Due to substantial overlap in their respective substrate ranges, deletion of any pathway has a strong impact only on the subset of substrates that cannot be accommodated by any other pathway. For this reason, the set of substrates whose biogenesis is impaired in the absence of a targeting factor is often only a small subset of the set of substrates that normally engage that factor. Thus, many substrates that are 'SRP independent' or 'GET independent' may normally engage the SRP pathway or the GET pathway. The substrate range of each pathway is probably best described by direct physical interaction analysis in unperturbed cells when all pathways are available. Although this has been achieved for SRP[35,36], it remains an important goal for the GET, EMC and SND pathways.

helix in SRP54 (REF.[20]). When a signal peptide or TMD emerges from the ribosome exit tunnel, the high local concentration of substrate near SRP54 outcompetes and displaces the autoinhibitory helix, which then serves as a lid to help enclose substrate within the groove. Autoinhibition may minimize promiscuous binding to non-substrates on and off the ribosome, thereby enhancing specificity of SRP towards bona fide signal peptides and TMDs emerging from a translating ribosome.

Promiscuous recognition of less hydrophobic sequences is also antagonized by the nascent polypeptide-associated complex (NAC)[27–29], a highly abundant cytosolic factor that binds near the ribosome exit tunnel of most or all ribosomes[30] (FIG. 2d). In the absence of NAC, SRP can bind to ribosomes translating proteins lacking a signal peptide or TMD, leading to their inappropriate targeting to the ER[27,28,31–34]. Molecular insight into how NAC and SRP coordinate at the ribosome exit tunnel[29] awaits structural analysis. The biological importance of NAC for ensuring targeting specificity is highlighted by extensive mistargeting of numerous proteins to the wrong subcellular location, organelle stress and lethality in *Caenorhabditis elegans* depleted of NAC[28]. Several quality control pathways monitor failures in targeting (BOX 2).

Transcriptome-wide footprinting of SRP-engaged ribosomes has verified conclusions derived from biochemical and structural analyses of SRP's client preference. In bacteria, SRP typically engages the first TMD of membrane proteins[35], whereas in yeast, SRP engages both TMDs and cleavable signal peptides[36]. The analysis of mRNA delocalization from the ER upon rapid SRP depletion in yeast shows that the set of delocalized mRNAs generally matches those engaged by SRP in transcriptome-wide studies[37]. These results all arrive at the consistent conclusion that in eukaryotes SRP engages and mediates the co-translational targeting of ribosomes that display a signal peptide or TMD at the exit tunnel. In prokaryotes, SRP's substrate preference seems to be shifted towards higher hydrophobicity, making it more specific for TMDs than signal peptides[35,38].

Post-translational targeting by the GET pathway. TA proteins of moderate to high hydrophobicity are recognized and targeted by the GET pathway (reviewed in detail elsewhere[39–41]) (FIG. 2e). Here we use the nomenclature originally described for components of the yeast pathway, which has recently been applied to other organisms. Specificity of the GET pathway for hydrophobic TMDs is imparted by the homodimeric, ATP-dependent chaperone GET3 (REFS[42–44]). The TMD-binding site spans the GET3 dimer and is a deep, ~35-Å-long hydrophobic groove that can accommodate an α-helix of ~20 amino acids[45–47]. GET3 is restricted to C-terminal TMDs because the association of SRP with the ribosome ensures that SRP has priority for non-C-terminal TMDs that emerge from the ribosome during translation. Thus, GET pathway substrates are defined by a combination of negative selection of non-TA proteins by SRP and positive selection of high-hydrophobicity TMDs by GET3. Proteins that engage GET3 are targeted to the ER via a receptor composed of GET1 and GET2 (REFS[48,49]).

Loading of substrates onto GET3 occurs within a pretargeting complex that additionally contains the TMD-binding chaperone SGTA (Sgt2 in yeast) and the GET4–GET5 complex (which in metazoans also contains the quality control factor BAG6)[43,44,50,51]. SGTA engages a TMD shortly after nascent protein release from the ribosome[43–45,51,52]. TMD capture might occur at the ribosome surface near the exit tunnel, where the yeast Get4–Get5 complex was recently found to bind[52], consistent with the ribosome-binding capacity of the mammalian GET4–GET5–BAG6 complex observed earlier[44]. Rapid and reversible binding of SGTA to GET5 (REF.[53]) would allow it to sample the ribosome exit tunnel region to capture TMDs there. SGTA can also bind TMD substrates released into the cytosol[45,51] or acquire

Sec61 complex
A heterotrimeric protein complex that translocates hydrophilic polypeptide segments across the membrane through an aqueous channel and inserts hydrophobic domains into the membrane through a lateral gate. It is called the 'SecY complex' in prokaryotes.

Insertases
Transmembrane proteins containing a hydrophilic vestibule that facilitates translocation of short polypeptide segments across the membrane concomitant with transmembrane domain insertion.

them from HSP70 family members via an HSP70–SGTA interaction[54]. Although most HSP70 clients are soluble proteins engaged in folding, these are not transferred to SGTA because the latter has preference for longer hydrophobic segments that typify TMDs[43,55]. Structural modelling of the SGTA family suggests its substrate-binding site is more hydrophobic and larger than in HSP70 but smaller than in GET3 (REF.[56]).

SGTA-associated TMDs are then transferred to GET3 in a handover reaction coordinated by the pretargeting complex[43,50,51,57]. TMDs that do not match the preference of GET3 for high hydrophobicity are not transferred[43,55], providing a potential mechanism to prevent loading of mitochondrial membrane proteins whose TMDs are typically less hydrophobic. Thus, a series of chaperones with preference for increasingly longer and more hydrophobic sequences (GET3 > SGTA > HSP70) effectively 'filters' potential substrates, loading onto GET3 only those TMDs of high hydrophobicity intended for the ER. Lower-hydrophobicity TMDs remain on SGTA or engage some other cytosolic chaperone, as discussed next.

Chaperone-facilitated targeting. TA proteins with TMDs of low hydrophobicity probably get captured by SGTA similarly to TA proteins of the GET pathway[23,43]. This is because SGTA has a broader substrate range and can bind TMDs that are shorter and less hydrophobic than GET3 (REFS[51,55,56]). However, as discussed already, these TMDs are not loaded effectively into the hydrophobic groove of GET3 (REF.[23]). Substrate binding by SGTA is highly dynamic[51], and upon release, the TMD of such TA proteins seems to be recognized directly by EMC at the ER membrane[23,58,59]. Until their recognition by EMC, these proteins are prevented from aggregation by dynamic binding and release from SGTA or other TMD-binding chaperones (FIG. 2e). Such chaperones are considered 'targeting-permissive' factors and include calmodulin[3,23], members of the ubiquilin family[60] and heat shock proteins such as HSP70 (REFS[61,62]), the last of which may serve a similar role in bacteria[63,64].

The lack of a requirement for a dedicated targeting pathway may reflect the lower hydrophobicity of these proteins. This means they are less prone to aggregation and can be kept soluble by abundant chaperones that dynamically bind and release them in the cytosol. Chaperone binding prevents inappropriate interactions, while chaperone release allows opportunities for targeting via direct ER engagement. Dynamic substrate release from the chaperone is crucial for recognition by EMC because impeding release impairs targeting and membrane insertion[23]. A conceptually similar mechanism is currently thought to facilitate mitochondrial and chloroplast targeting of nuclear-encoded proteins[65–67].

TMD insertion

Once a polypeptide arrives at the membrane, its TMD or TMDs need to be inserted into the lipid bilayer. The partitioning of a hydrophobic segment of a polypeptide into the hydrophobic membrane is an energetically favourable reaction[68–70]. The two critical impediments to this process are off-pathway interactions before insertion and translocation of TMD-flanking hydrophilic polypeptide across the bilayer (FIG. 3a, left). Off-pathway interactions such as aggregation are minimized during substrate delivery to the membrane by targeting factors and chaperones as discussed earlier. The second barrier, translocation of TMD-flanking hydrophilic domains concomitant with TMD insertion, is the key role of TMD insertion factors.

The ER contains at least four insertion factors: the Sec61 complex, the GET1–GET2 complex, EMC and the TMCO1 complex. The last three factors contain subunits (GET1, EMC3 and TMCO1) that were recently appreciated to be evolutionarily and structurally related to each other and to members of the Oxa1 family of protein insertases[71,72]. These collectively form the Oxa1 superfamily, an ancient class of proteins having arisen before the divergence of archaea and bacteria, and

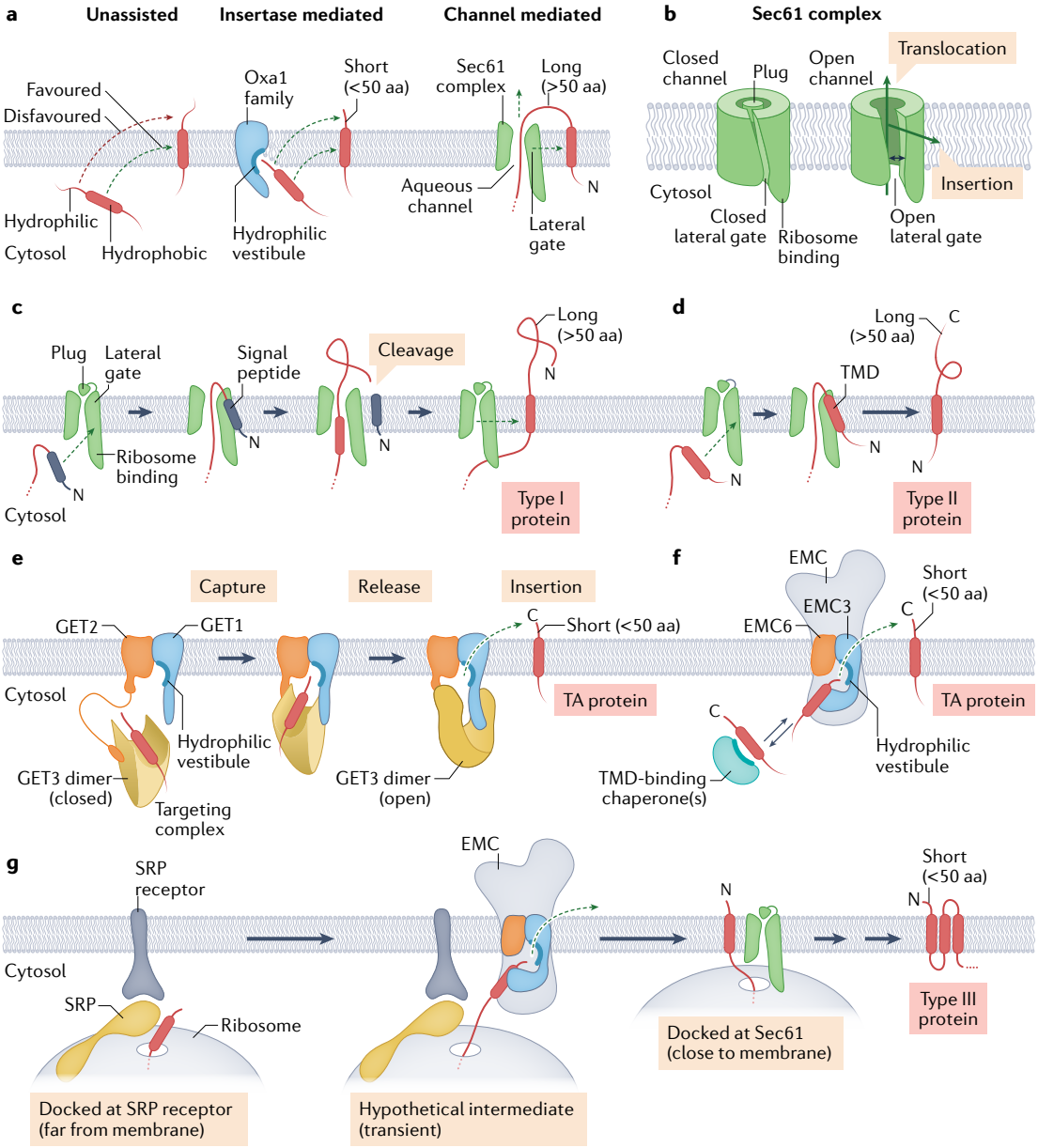
Box 2 | Quality control of membrane proteins

The biogenesis of membrane proteins can fail at any of various steps: translation, targeting, insertion, folding and assembly. Each step is monitored by a combination of both general and transmembrane domain (TMD)-specific quality control pathways. Incomplete translation due to a stalled ribosome is recognized by the ribosome-associated quality control pathway (reviewed elsewhere[188–191]). Ribosome-associated quality control serves to target the partially synthesized protein for ubiquitin-mediated degradation, recycle or degrade the ribosome and, in some cases, degrade the mRNA. This pathway operates in the cytosol and at organelle membranes, including the endoplasmic reticulum (ER).

Failure at the targeting or insertion step can lead to a membrane protein that is mislocalized to the cytosol. Several quality control factors deal with mislocalized membrane proteins depending on features of their TMDs. The cytosol contains several TMD-binding factors that also interact with E3 ubiquitin ligases. These include BAG6, which associates with the ubiquitin ligase RNF126 (REFS[192,193]), and the ubiquilin family of proteins, which interact with as yet unidentified E3 ligases[60]. Mammals have four ubiquilin family members, perhaps with different substrate preferences that would also differ from the substrate preference of BAG6, which seems to prefer particularly hydrophobic TMDs that normally use the guided entry of tail-anchored protein (GET) pathway. Yeast does not have an obvious BAG6 homologue, but contains the ubiquilin family member Dsk2, which might have a similar role. In addition, mislocalized membrane proteins can be directly recognized by the ER-resident E3 ligases Doa10 in yeast[194,195] and perhaps MARCH6 and TRC8 in mammals[196,197].

In addition to cytosolic mislocalization, membrane proteins can be misinserted into the wrong organelle or in the wrong orientation. For example, the TMDs of nuclear-encoded mitochondrial membrane proteins can be similar to ER-destined membrane proteins. Similarly, many N_{cyt} signal anchors (for type II transmembrane proteins) and signal peptides are similar to ER membrane protein complex (EMC)-dependent N_{exo} signal anchors (for type III transmembrane proteins), differing primarily in their TMD-flanking charged residues. Thus, it seems likely that EMC sometimes inadvertently inserts terminal TMDs intended for mitochondria or N_{cyt} orientation. Recently, the P5A-ATPase family member ATP13A1 (Spf1 in yeast) was found to dislocate moderately hydrophobic TMDs containing flanking positive charges facing the ER lumen[198]. The mitochondrial outer membrane contains an analogous (although mechanistically distinct) ATP-dependent TMD dislocase termed 'ATAD1' (Msp1 in yeast)[199–201]. Thus, TMD mistargeting and misorientation appear to be sufficiently frequent occurrences to warrant robust and highly conserved quality control pathways for their mitigation.

Finally, membrane proteins that are inserted but fail to fold or assemble properly are recognized by organelle quality control pathways. At the ER, these pathways are collectively termed 'ER-associated degradation' (reviewed elsewhere[202,203]). One mechanism of membrane protein recognition in ER-associated degradation seems to involve TMDs that expose hydrophilic side chains to the membrane. When correctly folded or assembled, such hydrophilic residues are shielded from the membrane, perhaps explaining why their persistent exposure is a reliable indicator of failed membrane protein biogenesis. This mechanism of recognition is conceptually similar to how exposure of hydrophobic patches in the aqueous environments of the cytosol or ER lumen is used as a cue for soluble protein misfolding[204].



hence long before emergence of eukaryotes. Thus, its members are found in the bacterial and archaeal plasma membranes, the inner membranes of mitochondria and plastids (which are evolutionarily related to the inner membrane of endosymbiont bacteria, from which they evolved) and the eukaryotic ER.

As argued in detail later, all Oxa1 superfamily members may share a core mechanism of TMD insertion involving a hydrophilic vestibule to aid flanking domain translocation (FIG. 3a, middle). This insertion-coupled translocation reaction is typically restricted to fewer than ~50 amino acids because Oxa1 superfamily proteins lack a membrane-spanning channel[73]. By contrast, an aqueous channel housed in the Sec61 complex allows it to translocate hydrophilic polypeptide segments of unlimited length[74] (FIG. 3a, right). This is why TMDs flanked by long hydrophilic regions use the Sec61 complex for their insertion. This basic concept — that short TMD-flanking domains are translocated via a hydrophilic vestibule within an Oxa1 superfamily member and long TMD-flanking domains are translocated through the Sec61 channel — is a recurrent theme in the discussion of each insertion pathway and their respective substrates below.

TMD insertion by the Sec61 complex. The Sec61 complex (termed the 'SecY complex' in bacteria and archaea) is a universally conserved protein conducting channel used for both secretion and membrane insertion of proteins (reviewed extensively elsewhere[74]). The channel is housed in the largest subunit, SecY or Sec61α, with two small subunits located peripherally. The Sec61 complex is able to open axially across the membrane for polypeptide translocation and laterally into the membrane for TMD insertion (FIG. 3b). Structures of SecY and Sec61 show them to be a pseudosymmetric membrane protein

Oxa1 superfamily
An evolutionarily related group of membrane protein insertases that includes Oxa1 in the inner mitochondrial membrane, YidC in the bacterial inner membrane, Ylp1 in the archaeal plasma membrane, Alb3 in the chloroplast inner membrane, and GET1, EMC3 and TMCO1 in the eukaryotic endoplasmic reticulum.

◀ Fig. 3 | **Membrane protein insertion at the endoplasmic reticulum. a** | Comparison of unassisted, insertase-mediated and channel-mediated insertion of transmembrane domains (TMDs). Insertion of a hydrophobic TMD into the hydrophobic membrane is energetically favoured, whereas translocation of a hydrophilic flanking domain across the membrane is disfavoured. Insertases lower the energy barrier for flanking domain translocation to facilitate TMD insertion. Channels allow long flanking domains to be translocated by providing a continuous aqueous conduit across the membrane. **b** | The open Sec61 complex containing an aqueous channel for protein translocation and a lateral gate for membrane insertion of hydrophobic segments. In the closed state, a plug domain occludes the channel. **c** | A type I membrane protein contains a signal peptide that engages the lateral gate of Sec61. The following segment of polypeptide is threaded through the Sec61 channel, after which the signal peptide is cleaved. A downstream TMD enters the Sec61 channel, then moves into the membrane through the lateral gate of Sec61. **d** | A type II membrane protein uses its first TMD to engage the lateral gate of Sec61 similarly to a signal peptide. The TMD moves into the membrane through the lateral gate, and the downstream polypeptide is translocated through the Sec61 channel. **e** | Tail-anchored (TA) protein insertion by the guided entry of TA protein (GET) pathway. A targeting complex comprising GET3 and a TA protein is captured by the cytosolic domain of GET2. Next, the cytosolic domain of GET1 releases the TA protein from GET3. The TMD of the TA protein then inserts into the membrane, using the hydrophilic vestibule in GET1 to facilitate translocation of the short carboxy-terminal tail. **f** | TA protein insertion by the ER membrane protein complex (EMC). The TMD of the TA protein, kept soluble in the cytosol by a chaperone, engages the cytosolic domain of EMC. From here, the TMD is inserted into the membrane, using the hydrophilic vestibule in EMC3 to facilitate translocation of the short carboxy-terminal tail. **g** | Type III membrane protein insertion by EMC. A ribosome translating a TMD-containing protein is targeted to the ER by the signal recognition particle (SRP) and its receptor. At this point, the ribosome is far enough from the membrane to transiently allow EMC to sample the region near the nascent protein. The TMD engages the cytosolic domain of EMC and is inserted into the membrane, using the hydrophilic vestibule in EMC3 to facilitate translocation of the short amino-terminal tail. The ribosome is then brought close to the membrane by docking onto Sec61, where the remainder of the protein is synthesized and inserted. aa, amino acids; C, carboxy terminus; N, amino terminus.

with its N-terminal and C-terminal halves surrounding a central hourglass-shaped pore[75,76]. The two halves come together like a clamshell, with a hinge at the back and a frontside lateral gate that can open towards the membrane. In the inactive state, the central pore is occluded by a short helix known as the plug.

Inactive and closed Sec61 can be opened for translocation when a signal peptide (or TMD) binds to and parts the lateral gate[77,78]. Lateral gate opening leads to plug displacement, creating an open conduit across the membrane. The signal peptide binds to Sec61 with its N terminus facing the cytosol and eventual cleavage site facing the lumen (FIG. 3c). This position and orientation causes the polypeptide downstream of the signal peptide to be pulled into the pore of Sec61. From this point, further translational elongation results in translocation, and the hydrophobic signal peptide diffuses into the membrane, where it is cleaved by the lumen-oriented active site of signal peptidase[79].

TMDs can be inserted into the membrane by Sec61 in two modes. In the first mode, a cleavable signal peptide has already threaded the polypeptide through the central pore within Sec61 when a TMD emerges from the ribosome (FIG. 3c). The TMD therefore moves into an open Sec61 channel. It is thought that a dynamic lateral gate that is constantly sampling the open conformation allows the nascent TMD inside the Sec61 channel to access the surrounding membrane. The hydrophobicity of the TMD would favour the membrane environment, causing TMD insertion by a simple partitioning mechanism[10,80]. Proteins that initiate insertion in this way are called 'type I membrane proteins' (see FIG. 1b) and represent roughly one third of all membrane proteins made at the ER.

In the second mode, a TMD engages, opens and then passes through the Sec61 lateral gate using the mechanism described above for a signal peptide[81] (FIG. 3d, compare with FIG. 3c). Hence, the N-terminal flanking domain is retained in the cytosol, and the C-terminal flanking domain is threaded through the central channel. The TMD is initially positioned at a parted lateral gate before passing through the gate into the membrane. Proteins that initiate their insertion by this mechanism are termed 'type II membrane proteins' (see FIG. 1b), with their first TMD termed a 'type II' or 'N_{cyt}' signal anchor (meaning the N terminus faces the cytosol).

Proteins that initiate their membrane insertion using a TMD preceded by fewer than ~50 amino acids are called 'type III membrane proteins' (see FIG. 1b), with their first TMD termed a 'type III' or 'N_{exo}' signal anchor. Almost two-thirds of all membrane proteins begin with either an N_{exo} signal anchor or an N_{cyt} signal anchor. N_{exo} signal anchors have long been thought to be inserted via the lateral gate in Sec61, similarly to N_{cyt} signal anchors[82]. However, inhibitors of the lateral gate in Sec61 that block the insertion of signal peptides and N_{cyt} signal anchors do not inhibit the insertion of N_{exo} signal anchors[83–85]. More strikingly, immunodepletion of Sec61 had little effect on the insertion of all tested N_{exo} signal anchors, yet completely precluded signal peptide and N_{cyt} signal anchor function[86]. These observations suggest that N_{exo} signal anchors are inserted by a qualitatively different mechanism than either signal peptides or N_{cyt} signal anchors. As discussed later, recent findings indicate that N_{exo} signal anchors can be inserted by EMC, providing one explanation for these otherwise unexpected findings[86].

Finally, it should be noted that the eukaryotic Sec61 complex associates with several proteins, but their role in TMD insertion remains unclear. The translocon-associated protein (TRAP) complex and the translocating chain-associated membrane protein (TRAM) are found in many eukaryotes and facilitate the ability of weakly hydrophobic signal peptides to initiate co-translational translocation through Sec61 in a mammalian cell-free translation system[87,88]. A similar function is ascribed to the Sec62–Sec63 complex during post-translational translocation, which also relies on weakly hydrophobic signal peptides[89–91]. Of these factors, TRAM has also been observed to interact with low-hydrophobicity TMDs during or shortly after insertion[92,93], but a functional role for this interaction has not been demonstrated. Some or all of these factors might perhaps assist insertion of certain TMDs similarly to how they facilitate signal peptide function, but this has not been studied.

TA protein insertion by the GET complex. Membrane proteins engaged by GET3 are delivered to an ER-localized receptor composed of GET1 and GET2, both of which are three-TMD proteins containing cytosolic GET3-binding domains[48]. Reconstitution experiments

Type I membrane proteins
Signal peptide-containing membrane proteins oriented with their mature amino terminus facing the lumen (a topology that is also termed 'N_{exo}') following signal peptide cleavage.

Type II membrane proteins
Membrane proteins oriented with their amino terminus facing the cytosol (a topology that is also termed 'N_{cyt}').

Type III membrane proteins
Membrane proteins oriented with their amino terminus facing the lumen (a topology that is also termed 'N_{exo}'); these proteins typically possess a short (fewer than 50 amino acids) amino-terminal flanking region.

using yeast components have rigorously established that the Get1–Get2 complex is both necessary and sufficient for insertion of Get3-targeted TA proteins[94–96]. The insertion reaction involves three main steps (FIG. 3e): initial engagement of the TA–GET3 targeting complex by the GET1–GET2 receptor; receptor-mediated release of TA protein from GET3; TMD insertion into the lipid bilayer.

Engagement of the TA–Get3 targeting complex and TA protein displacement from Get3 are mediated by the cytosolic domains of Get1 and Get2, both of which have partially overlapping binding sites on Get3 (REFS[94,97]). Because a single TA protein binds to a Get3 homodimer[45], there are two potential binding sites on the TA–Get3 targeting complex. The binding of the cytosolic domain of Get2 to Get3 does not disrupt the TA–Get3 interaction, whereas the Get1 cytosolic domain can dislodge TA proteins from Get3 (REFS[94,96–98]). Although these individual activities of GET1 and GET2 are well established, their precise order of interactions with GET3 is not resolved. In one model (shown in FIG. 3e), GET2, whose GET3-binding domain resides on a long flexible tether, engages GET3 initially. This interaction brings the targeting complex close to GET1, which then engages the other binding site on GET3 to release TA protein.

After TA protein release, the TMD of the client protein is inserted into the lipid bilayer in a reaction that requires the TMD regions of the GET1–GET2 complex[95]. A recent cryo-EM structure of the human GET1–GET2–GET3 complex observed two copies of the GET1–GET2 heterodimer bound to a single GET3 homodimer[99]. Notably, GET1 in this structure showed an overall architecture, including a cytosol-facing hydrophilic vestibule, consistent with its earlier assignment as an Oxa1 superfamily member[71]. Although this structure contains two GET1–GET2 heterodimers, the stoichiometry of the GET1–GET2 complex during insertion is currently unclear. Biochemical experiments show that single Get1–Get2 heterodimers from yeast reconstituted into liposomes seem to be sufficient for TA protein insertion[100]. This result suggests that the simplest model is one where the GET3–TA targeting complex engages one GET1–GET2 heterodimer such that the hydrophilic tail of the TA protein can access the hydrophilic vestibule in GET1 (FIG. 3e). From this intermediate, GET1-mediated dislodging of substrate from GET3 allows the substrate's TMD to enter the membrane, with the barrier to hydrophilic tail translocation eased by the hydrophilic vestibule in GET1. A single GET1–GET2 complex mediating insertion would be consistent with the likely stoichiometry of other Oxa1 superfamily members during TMD insertion[58,59,101–106].

Terminal TMD insertion by EMC. Several observations had long indicated that the GET pathway is not the only route for TA protein insertion at the ER. In both yeast and mammalian systems, the extent of GET pathway dependence for insertion differs widely among TA proteins[107]. Mechanistically, TMD hydrophobicity of many natural TA proteins can be lower than what is efficiently accommodated by GET3 (REFS[45,55]). At least one explanation for these discrepancies came with the discovery that a widely conserved and large protein complex termed the EMC[108] can insert low-hydrophobicity and moderate-hydrophobicity TA proteins[23]. The observation that many TA proteins have a TMD of intermediate hydrophobicity compatible with either pathway (FIG. 2c) may explain why neither EMC nor GET subunits are essential in yeast but their deletion is lethal in combination[108].

EMC is an ER-resident eight- or nine-subunit complex (depending on the species) with large cytosolic and lumenal domains connected by a membrane domain containing 14 TMDs (reviewed elsewhere[109–111]). Purified EMC reconstituted into synthetic liposomes was sufficient to mediate insertion of a model TA protein upon release from a chaperone[23] (FIG. 3f). The current model is that the cytosolic domain of EMC has one or more transient binding sites for moderately hydrophobic TMDs, from which the membrane can be accessed for insertion concomitant with translocation of the hydrophilic tail.

Recent structures of yeast and mammalian EMC reveal that, as predicted[71], EMC3 is an Oxa1 superfamily member with a cytosol-facing hydrophilic vestibule. On the basis of this similarity, one model for TA protein insertion is that this vestibule facilitates translocation of the C-terminal hydrophilic tail (FIG. 3f). Consistent with this idea, mutations in or near this vestibule impair insertion[59,105,106].

EMC was also noted to contain a large membrane-embedded hydrophobic groove on the side opposite the hydrophilic vestibule[58,59,105,106]. This groove is exposed to the surrounding membrane and is sufficiently large to potentially accommodate a substrate TMD. In mammalian EMC structures, this groove is continuous with a shallow cytosolic cradle formed primarily by the EMC2–EMC8 subcomplex. The observation that this isolated subcomplex can bind the TMD of a model TA protein led to the proposal that TMD insertion might occur from here into the membrane-exposed hydrophobic groove[58]. Having two different routes into the membrane might broaden the substrate range of EMC by accommodating TMDs with different biophysical features, an idea worth exploring in future studies.

In addition to TA proteins, EMC can also mediate co-translational insertion of N_{exo} signal anchors displayed on translating ribosomes[86] (FIG. 3g). Although N_{exo} signal anchors are of the opposite topology to TA proteins, they are nonetheless similar in containing a short translocated domain of fewer than ~50 amino acids. In addition, most EMC-dependent N_{exo} signal anchors also have partial hydrophilic character, similarly to EMC-dependent TA proteins. These observations suggest that one general class of substrates for EMC-mediated insertion are TMDs close to either terminus of a protein.

A major unresolved issue with N_{exo} signal anchor insertion is precisely when EMC acts in the process[110]. The most likely possibility is after SRP-mediated targeting of the polypeptide to SRP receptor but before ribosome docking on Sec61 (FIG. 3g). Because SRP and Sec61 occupy overlapping sites on the ribosome, the handover reaction of the polypeptide between SRP and

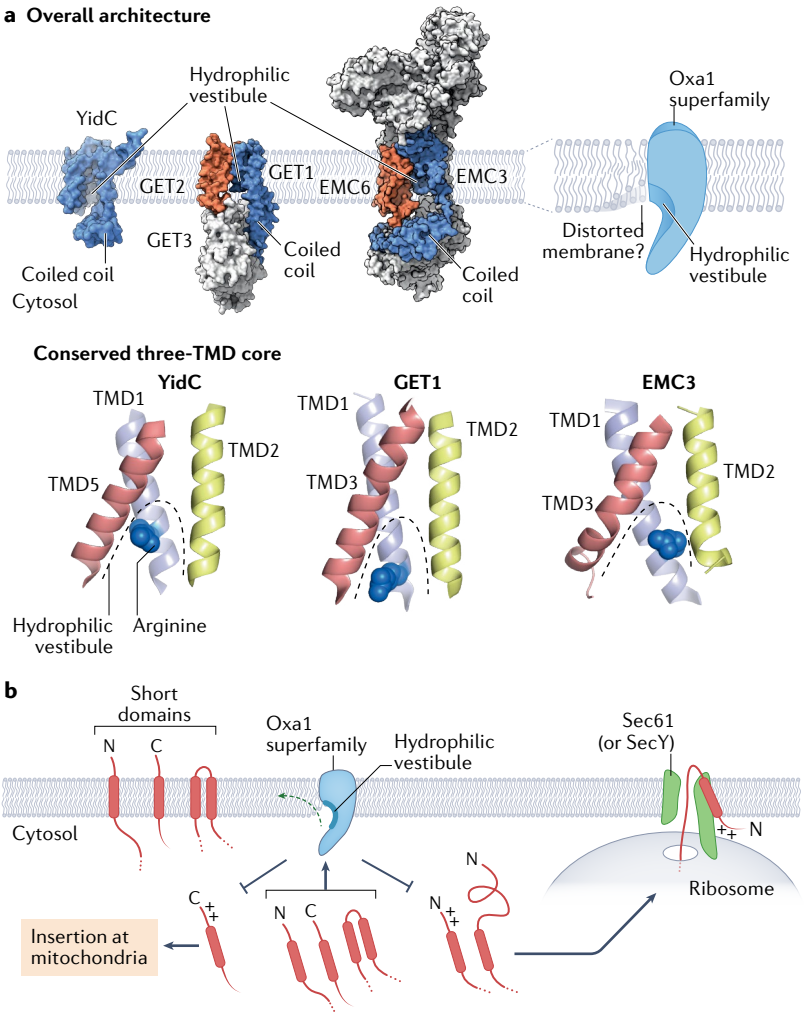


Fig. 4 | **Structure and function of the Oxa1 superfamily insertases. a** | Some examples of Oxa1 superfamily members (blue) found in the prokaryotic plasma membrane (YidC) and eukaryotic endoplasmic reticulum (GET1 and EMC3). Each of these proteins contains a conserved membrane-embedded hydrophilic vestibule open towards the cytosol. The vestibule facilitates translocation of the hydrophilic segment of polypeptide that flanks a transmembrane domain (TMD) during TMD insertion. Close-up views of the conserved core of each Oxa1 superfamily member comprising three TMDs that define the hydrophilic vestibule (dotted line). A conserved arginine is found in the vestibule of most, but not all, Oxa1 superfamily members. The cartoon at the top right illustrates a cross section at the plane of the hydrophilic vestibule where the membrane might be distorted or thinned. A thin membrane might lower the energy barrier for polypeptide translocation. YidC is a single polypeptide, whereas GET1 forms an obligate complex with GET2 (orange), both of which participate in recruiting the targeting factor GET3 (grey). Only one of two GET1–GET2 complexes observed in the structure is shown; the other is behind the displayed complex facing the opposite direction. It is currently unclear whether the functional unit during TMD insertion consists of one or two copies of the GET1–GET2 complex. EMC3 is part of a nine-protein complex that includes EMC6 (orange), which is structurally and evolutionarily related to GET2. Other EMC subunits are shown in grey. The membrane domains of EMC4, EMC7 and EMC10 probably reside in front of EMC3 in this view but were not resolved in this structure. Protein Data Bank IDs 3WO6 for YidC, 6SO5 for human GET complex and 6WW7 for human EMC. **b** | Substrates favoured and disfavoured for insertion by Oxa1 superfamily insertases such as YidC, the GET1–GET2 complex and EMC. Tail-anchored proteins with positive charges (shown with '+') in the translocated tail are inserted into mitochondria. Signal anchors with a positively charged or lengthy amino-terminal flanking domain are inserted by the Sec61 complex. C, carboxy terminus; N, amino terminus.

Sec61 must necessarily involve an intermediate when the signal anchor is released from its binding site on SRP, but cannot engage Sec61 until SRP dissociates from the ribosome. Importantly, the large size of SRP and its receptor[112] necessarily means the ribosome is ~100 Å from the membrane at this step. This space is compatible with the cytosolic domain of EMC approaching the ribosome exit tunnel to access the substrate.

This putative intermediate would necessarily be short-lived because of the high affinity of Sec61 for its binding site at the ribosome exit tunnel[113]. After ribosome binding to Sec61, the ribosome exit tunnel is positioned very close to the membrane, sterically preventing EMC from accessing the emerging nascent polypeptide. It is attractive to postulate that the limited time available for EMC to act precludes suboptimal hydrophobic sequences (for example, signal peptides and TMDs intended for N_{cyt} insertion by Sec61) from promiscuous insertion in the incorrect N_{exo} orientation. The mechanism of co-translational triage of TMDs for insertion by EMC versus Sec61 remains to be resolved.

General principles of TMD insertion by the Oxa1 superfamily. Oxa1 resides in the mitochondrial inner membrane and mediates insertion of many inner membrane proteins, including components of the respiratory chain complexes. The earliest studies on Oxa1 recognized its similarity to Alb3 in the inner chloroplast membrane and YidC in the bacterial plasma membrane, both of which were later shown to function in membrane protein insertion[114–118]. More recently, GET1, EMC3, TMCO1 and archaeal proteins of unknown function were found to be evolutionarily related to the Oxa1 family[71]. Recent structures demonstrate that the conserved core of YidC[119,120] is found in an archaeal protein, Ylp1 (REF.[121]), yeast and human EMC3 (REFS[58,59,105,106]), TMCO1 (REF.[104]) and GET1 (REF.[99]). These observations, beyond their relevance to protein evolution, allow experimental results from the study of one family member to be interpreted in light of mechanistic insights from other members. In what follows, we attempt to synthesize the biochemical and structural analysis of various insertase family members into a set of shared core principles.

The conserved element of all Oxa1 superfamily members is a three-TMD core arranged to form a cytosol-facing hydrophilic vestibule embedded in the membrane (shown in FIG. 4a for YidC, GET1 and EMC3) (recently reviewed[122]). Most but not all family members contain one or more basic amino acids in the vestibule. The first and second TMDs of the core are linked by a cytosolic coiled coil that is highly variable in length and conformation among family members. In most cases, Oxa1 superfamily members form a complex with evolutionarily conserved partners whose functions are unclear. Although TMCO1 and Ylp1 initially seemed to contain only the core subunit, co-evolutionary analyses suggest they may interact with C20orf24 and Mj0606, respectively, to form complexes structurally and evolutionarily related to the GET1–GET2 and EMC3–EMC6 complexes[72]. Beyond these shared features, different Oxa1 superfamily members harbour different additional domains or associated subunits. These additional

elements can include extra TMDs (for example, YidC and Oxa1), large domains in the lumen or cytosol, or associated subunits that contribute to larger protein assemblies (for example, EMC and TMCO1). Elements common to all (or nearly all) Oxa1 superfamily members are likely to participate in the conserved function of TMD insertion, whereas the divergent elements may provide regulatory, member-specific or organism-specific functions.

The hydrophilic vestibule can be considered a partial channel across the membrane where a segment of the substrate's flanking hydrophilic domain can move part of the way towards the lumen. Partial translocation of the polypeptide within the vestibule could reduce the energetic barrier to its complete translocation because the lipid bilayer might be locally distorted to shorten the distance between the head groups of each leaflet (reviewed elsewhere[73]). Such a mechanism would allow short polypeptide segments (typically shorter than ~50 amino acids), but not long or folded domains, to cross the membrane concomitant with TMD insertion. The one or more basic amino acids inside the vestibule would impose the additional constraint of disfavouring vestibule entry of substrates with multiple basic amino acids. Thus, substrates for Oxa1 superfamily members are TMDs flanked by relatively short unstructured segments of polypeptide with relatively few basic amino acids; TMDs whose context does not meet these criteria would be rejected and instead inserted by alternative machinery such as the Sec61 complex (FIG. 4b).

Consistent with this model, terminal TMDs that are known to be directly inserted by YidC, EMC and the GET complex have short unfolded translocated regions, and increasing the length of this tail impedes insertion in the GET pathway[42] and probably the other two pathways as well[123]. The need for a hydrophilic vestibule with an accompanying distorted membrane would explain why the cytosolic domains of the Get1–Get2 complex at the membrane, while sufficient for Get3 targeting and substrate release, do not mediate effective TMD insertion. Oxa1 superfamily members also have the capacity to translocate hydrophilic loops between sequential TMDs[117,124]. Such loops are presumably translocated through the hydrophilic vestibule concomitant with insertion of the two adjacent TMDs, either in rapid succession or together as a hairpin. Reconstitution studies of YidC-mediated insertion of several different multipass membrane proteins containing translocated loops shorter than ~35 amino acids support this mechanism of hairpin insertion[124–126]. A multipass protein containing an ~300 amino acid loop could not be translocated by YidC unless this loop was shortened to ~15 amino acids[124].

The discrimination against flanking basic residues might explain why N_{cyt} signal anchors close to the N terminus are not inappropriately inserted in the N_{exo} orientation by EMC (or YidC in bacteria) even if such clients are initially targeted to EMC or YidC. The same mechanism could be used by EMC and the GET1–GET2 complex to reject TA proteins intended for mitochondria, which typically have positive charges in their translocated tail. Thus, Oxa1 superfamily members may contribute to the long-observed 'positive-inside' rule at the bacterial plasma membrane and eukaryotic ER, in which the cytosolic flanking domains of TMDs are enriched in positively charged amino acids[127]. Such rejection is unlikely to be perfect, but any inappropriately inserted TMDs might be rectified by quality control mechanisms (see BOX 2).

Substrate access to the insertases may be regulated. Given the simple paradigm of TMD insertion by a membrane-thinning hydrophilic vestibule, why are Oxa1 superfamily members typically part of larger complexes or embellished with additional modules (FIG. 4a)? One possibility might be to regulate substrate access to the hydrophilic vestibule. This seems to be the case in the GET pathway, where the cytosolic domain of Get2 is important for delivering substrates to Get1 (REFS[94,96]). In addition, it is attractive to posit that the hydrophilic vestibule in GET1 is normally occluded, either by GET2 or by a yet unseen 'closed' conformation of GET1, until engaged by substrate-bound GET3. Thus, the vestibule would be contextually gated rather than constitutively residing in the ER in an energetically unfavourable 'open' conformation.

In the case of EMC, the hydrophilic vestibule of EMC3 seems to be partially occluded by EMC4, EMC6, EMC7 and possibly EMC10 (REFS[58,59,105,106,111]). An attractive model is one where one or more of these subunits regulate access to the insertase module of EMC3. How this regulation could be achieved is unclear, but one possibility is that substrate binding to the cytosolic domains of EMC subunits induces their conformational changes that better expose the hydrophilic vestibule of EMC3. This is consistent with the observation that mutations in the cytosolic domains of EMC that in the structure are relatively far from the hydrophilic vestibule of EMC3 impede insertase function of the complex[105,106].

Whether YidC or Oxa1 undergoes any conformational changes that regulate their hydrophilic cavities is unclear. Crosslinking studies with YidC suggest that the cytosolic coiled coil connected to two TMDs of the three-TMD core can potentially interact with SRP[128] and the ribosome[102,129]. Perhaps SRP or ribosome binding to the coiled coil is coupled to an opening of a normally closed hydrophilic vestibule. TMCO1, about which very little is known, also has a cytosolic coiled coil that interacts with the ribosome[104] and seems likely to have an interaction partner (C20orf24) related to GET2 and EMC6. Whether these features regulate its putative insertase activity is unknown. Conformational changes in the Oxa1 superfamily insertases during their functional cycle are poorly understood and warrant biophysical and structural analysis.

Membrane protein folding

TMDs of multipass membrane proteins are often packed together using non-hydrophobic amino acids (FIG. 5a). Furthermore, TMDs that contribute to channels, ligand-binding sites, or catalytic sites typically contain polar or charged amino acids. Calculations of insertion propensity[130] suggest that the TMDs of multipass membrane proteins are typically less hydrophobic than the TMDs of single-pass membrane proteins (FIG. 5b). Furthermore, exposed hydrophilic side chains within

Multipass membrane proteins
Proteins spanning the membrane more than once.

Single-pass membrane proteins
Proteins spanning the membrane once.

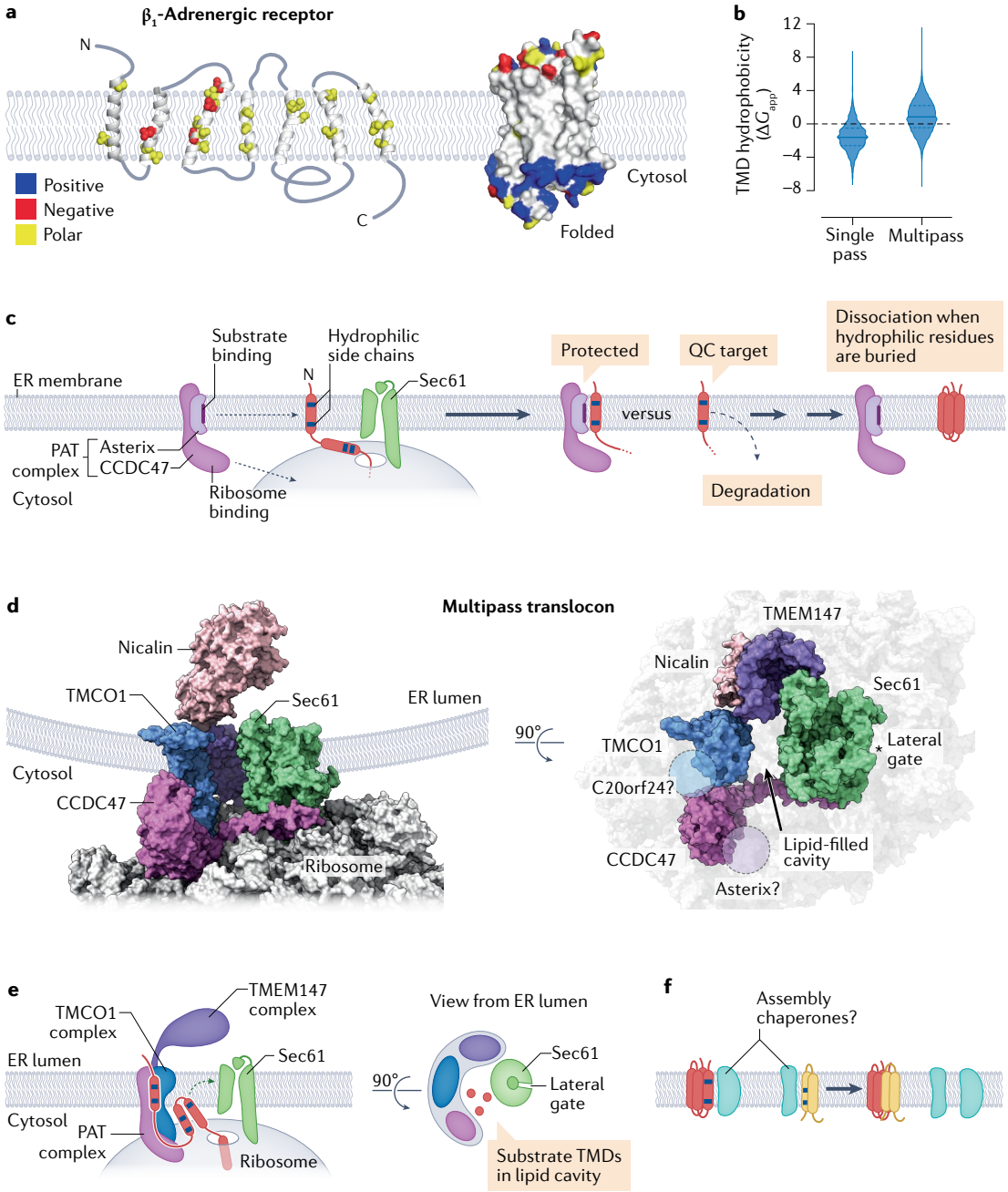


Fig. 5 | **Biogenesis of multipass membrane proteins. a** | The linear and folded states of the β_1-adrenergic receptor (ADRB1; Protein Data Bank ID 2VT4) illustrate how most of the hydrophilic side chains (blue, red and yellow) in the transmembrane domain (TMD) become buried upon folding. **b** | Violin plots showing the hydrophobicity of TMDs in single-pass and multipass proteins. Hydrophobicity was calculated as the predicted energy of membrane insertion, where a negative value indicates higher hydrophobicity and favours insertion and a positive value indicates lower hydrophobicity and disfavours insertion[130]. **c** | The PAT complex engages and protects nascent TMDs with hydrophilic residues until they are buried in the protein interior upon folding. Substrates that are not shielded are potential targets for quality control (QC). PAT complex shielding could occur regardless of the route of TMD insertion. **d** | Cryogenic electron microscopy structure of the ribosome-bound multipass translocon (Protein Data Dank ID 6W6L). The view from the endoplasmic reticulum (ER) lumen also indicates the potential positions of C20orf24 (adjacent to its likely interaction partner TMCO1) and Asterix (adjacent to its interaction partner CCDC47). The lateral gate is indicated by an asterisk. **e** | Protein biogenesis by the multipass translocon. In the intermediate depicted, the first TMD is being held by the PAT complex until TMD2 and TMD3 are inserted as a unit using the hydrophilic vestibule of the TMCO1 complex. The lipid cavity may be the site where multiple substrate TMDs can be accommodated and undergo folding while being protected from aggregation and inappropriate interactions. **f** | For the assembly of multiprotein complexes, hypothetical assembly factors may act as membrane chaperones that temporarily shield individual subunits of the complex in their unassembled state. This may serve to stabilize these intermediates in the membrane until their assembly with interaction partners. C, carboxy terminus; N, amino terminus.

the membrane would be recognized for quality control by ER-associated degradation pathways (BOX 2). Thus, successful folding of multipass membrane proteins requires the biogenesis machinery to not only insert partially hydrophilic TMDs but also temporarily stabilize and shield them in the membrane until their successful assembly with other TMDs. The mechanisms of multipass membrane protein biogenesis are just beginning to be defined.

Partially hydrophilic TMDs engage the PAT complex. Pioneering experiments investigating multi-pass membrane protein biogenesis used chemical crosslinking to find proteins adjacent to various insertion intermediates of the seven-TMD protein rhodopsin[131,132]. In addition to the ribosome-associated Sec61 complex, prominent crosslinks were seen between TMD1 of rhodopsin and an unidentified ~10-kDa protein provisionally termed 'PAT10' (for protein associated with the ER translocon of 10 kDa). This crosslinking partner was recently identified to be Asterix and was shown to tightly interact with another protein, CCDC47, to form the PAT complex[133] (FIG. 5c). Site-specific photocrosslinking experiments indicate that Asterix directly engages substrate TMDs inside the membrane[133], whereas parallel structural studies described later show that CCDC47 binds to Sec61-bound ribosomes[104].

Mutagenesis experiments suggest that hydrophilic amino acids within TMDs of the substrate are essential for PAT complex engagement[133]. This is noteworthy because essentially all multipass membrane proteins contain TMDs that have partial hydrophilic character that must be packed against other TMDs in the final structure (FIG. 5a). The PAT complex seems to selectively engage and presumably protect semihydrophilic TMDs until their intramolecular interaction partners are synthesized. Consistent with this function, depletion of either PAT complex subunit impairs stable expression of several multipass membrane proteins without any obvious effect on single-pass membrane proteins[133].

How exactly the absence of the PAT complex impacts membrane protein biogenesis is not known. The simplest explanation is that in the absence of a chaperoning activity of the PAT complex, a partially hydrophilic TMD might engage quality controls factors[134] (BOX 2) or even slip out of the membrane[135–137], leading to promiscuous degradation. Such off-pathway fates would compete with on-pathway folding, explaining why the phenotypic effects of PAT complex subunit loss on membrane proteins are partial. This is analogous to how many cytosolic proteins can fold to at least some degree without chaperones, whose primary role is to minimize off-pathway outcomes of secondary and tertiary structure acquisition.

The PAT complex seems to be essential at the organismal level[138,139], causes ER stress (presumably due to excess membrane protein misfolding) when deleted in cells[138] and has been conserved broadly across eukaryotes. The remote yeast homologues of Asterix (YPR063C) and CCDC47 (YNR021W) have not been studied yet but could provide a useful system for large-scale genetic and phenotypic analyses. Future studies should investigate the mechanistic and structural basis of substrate recognition by Asterix and the precise consequences of its absence for insertion and folding. As discussed next, the PAT complex probably cooperates with other putative chaperones, insertases and Sec61 to ensure membrane protein folding. Among these additional factors, members of the Oxa1 superfamily may have chaperone functions in addition to their insertase roles (see BOX 3).

A specialized translocon for multipass membrane proteins. The initial study of the PAT complex did not investigate its interactions with or position relative to the Sec61 complex. Instead, unexpected insight into this issue has

Box 3 | **Oxa1 superfamily members may also operate as chaperones**

The observation that newly inserted transmembrane domains (TMDs) near the prokaryotic SecY complex can crosslink to the Oxa1 superfamily member YidC has long suggested that YidC might have a potential chaperone function[205]. This idea is attractive because the hydrophilic vestibule in YidC used for polypeptide translocation serves as an ideal binding site for a partially hydrophilic TMD. In vivo experiments showing an effect of YidC depletion on insertion of some substrates[63,116,125,206–209] and folding of others[210–212] support this dual-function model for YidC. Whereas the insertion function of YidC is strongly supported by reconstitution studies in vitro[213–219], its chaperone function is less well defined and difficult to disentangle from the insertase role.

The best studied chaperone substrate for YidC is the sugar transporter LacY. Folding of LacY, as monitored by conformation-specific antibodies and intramolecular disulfide crosslinking, was observed to be impaired in *Escherichia coli* acutely depleted of YidC[211,212]. LacY was found to have been inserted in its normal topology on the basis of resistance to alkaline extraction and cysteine accessibility of inter-TMD loops. Because LacY can physically interact with YidC, the effect on folding seems to be direct. However, it can be challenging to exclude indirect effects of YidC depletion — which impairs production of various membrane proteins, causes a loss of membrane potential, induces a stress response and eventually leads to cell death — or an impairment in some aspect of YidC-mediated LacY insertion[220]. Despite these caveats, the concept of YidC acting as a chaperone is attractive and might also apply to other Oxa1 superfamily members.

In this regard, eukaryotic endoplasmic reticulum membrane protein complex (EMC) is noteworthy. Proteomic analyses of EMC-depleted cells show changes in the levels of various types of membrane proteins[221,222]. Because many of these are neither tail-anchored nor begin with an N_{exo} signal anchor (type III membrane protein orientation), it has been speculated that EMC performs functions beyond insertion of terminal TMDs[105,222]. Although it is difficult to exclude indirect effects of this depletion, one explanation could be that EMC also mediates insertion of TMD pairs linked by a short lumenal loop. A second non-mutually exclusive possibility is that EMC, either by itself or via interaction partners, acts as a chaperone during membrane protein biogenesis.

Co-immunoprecipitation of EMC with various multipass membrane proteins, and in some cases their substrate-specific maturation factors, supports a post-translational role in early biogenesis[222]. Genetic and proximity-biotinylation experiments have also suggested a co-translational function for EMC during biogenesis of certain multipass membrane proteins[222]. Although this idea is attractive and analogous to co-translational function of YidC near SecY[128,205,223,224], the large cytosolic domain of EMC limits its access to ribosome-associated Sec61. Hence, EMC is able to engage a nascent membrane protein only after it has diffused ~100 Å from the ribosome exit tunnel or after the ribosome has been detached from Sec61. This steric constraint might help explain why the best candidate co-translational substrates of EMC (for example, Yor1, Fks1 and viral polyproteins) are large multipass proteins with one or more long inter-TMD loops.

A second line of evidence for an insertase-independent function for EMC comes from mutagenesis studies[105]. EMC mutations in regions far from the EMC3–EMC6 insertase module can impair maturation of a membrane protein (TMEM97) without affecting known insertase substrates. Seeking a mechanistic explanation for such substrate-specific mutants may reveal currently unexplored aspects of EMC function. Towards this end, it will be important to analyse different co-translational and post-translational biogenesis intermediates for their direct interaction with EMC using approaches such as site-specific crosslinking and, eventually, structural methods.

Intramembrane chaperone
A factor that promotes folding in the membrane by temporarily shielding partially hydrophilic transmembrane domains of nascent polypeptides until their successful assembly with other transmembrane domains.

Chaperonin
A family of ATP-driven multimeric chaperone complexes characterized by a cylindrical structure with an internal chamber. The interior of the chaperonin cylinder provides a protected environment within which nascent proteins can fold.

come from a parallel seemingly unrelated study investigating the interaction partners and structure of the Oxa1 superfamily member TMCO1 (REF.[104]). Purification and proteomic analysis of TMCO1-containing ribosomes revealed the presence of the Sec61 complex, CCDC47 and a three-protein complex composed of TMEM147, nicalin and NOMO1. The absence of Asterix, the obligate partner of CCDC47, in this proteomic experiment is probably explained by technical limitations associated with the paucity of Asterix-derived peptides generated by trypsin digestion[133]. Thus, one can reasonably postulate that this TMCO1-containing translocon contains several key elements for membrane protein biogenesis: a protein conducting channel (Sec61 complex), an Oxa1 superfamily insertase complex (TMCO1 and C20orf24) and an intramembrane chaperone (the PAT complex).

Consistent with this function, sequencing of the mRNAs recovered with TMCO1-containing ribosomes revealed an exceptional enrichment of sequences coding for multipass membrane proteins[104]. Analysis of EAAT1, a trimeric multipass membrane protein whose mRNA was enriched in TMCO1-containing ribosomes, showed markedly reduced levels in cells depleted of CCDC47, TMCO1 or components of the TMEM147 complex. A moderate-resolution cryo-EM reconstruction of TMCO1-containing ribosomes revealed the positions of CCDC47, TMCO1 and the TMEM147 complex relative to the Sec61 complex[104] (FIG. 5d). These components (and potentially yet unidentified factors) constitute a translocon that appears to be specialized for multipass membrane protein biogenesis.

The three factors (TMCO1 complex, PAT complex and TMEM147 complex) define a membrane-exposed cavity, presumably containing lipids, on the hinge side of Sec61 opposite its lateral gate (FIG. 5d). The exit tunnel of the ribosome is positioned where this cavity abuts Sec61, which likely allows the nascent polypeptide to access the lateral gate of Sec61 as well as the multipass translocon components. Thus, TMDs emerging from a ribosome bound to this multipass translocon would have (at least) two insertion routes, Sec61 or TMCO1, depending on the substrate.

For example, TMDs followed by a lengthy lumenal domain might engage Sec61 so its channel can be utilized for translocation of this large soluble domain (FIG. 3d). By contrast, TMDs flanked by only short translocated segments could use TMCO1 by a mechanism typical for Oxa1 superfamily members. Notably, there seems to be sufficient space between the ribosome and the membrane to accumulate two TMDs and a short loop, thereby allowing their concerted insertion via TMCO1 (FIG. 5e). In this way, poorly hydrophobic internal TMDs of multipass membrane proteins, many of which cannot engage Sec61 effectively, could nonetheless be inserted by TMCO1.

The mammalian multipass translocon may be analogous to the prokaryotic 'holotranslocon' containing the SecY complex (which also contains SecE and SecG), YidC, the SecD–SecF complex and YajC[140]. Both translocons contain a SecY family channel, an Oxa1 superfamily insertase (for example, YidC) and a lipid-filled cavity. Although many TMDs might be able to use either the channel or the insertase for insertion, the unique functionality of the SecY family is its capacity to translocate lengthy soluble domains across the membrane. Conversely, Oxa1 superfamily members might be better at other reactions, such as insertion of two-TMD pairs or low-hydrophobicity terminal TMDs, that are less suited for the SecY family. A translocon with both a protein translocation channel and an insertase may therefore provide the requisite flexibility to accommodate the diverse TMDs and loops of multipass membrane proteins for their efficient biogenesis.

A potential site for membrane protein folding. The lipid-filled cavity of the mammalian multipass translocon[104] is large enough to accommodate multiple substrate TMDs with their exposed hydrophilic parts being temporarily chaperoned by the PAT complex or other translocon components (FIG. 5e). A protected cavity would facilitate intramembrane folding of nascent multipass proteins without off-pathway interactions with bulk ER proteins or quality control factors. This mechanism is analogous to cytosolic protein folding inside a chaperonin chamber[141]. In the chaperonin example, the surface properties of the chamber's interior promote folding of substrates. It will be interesting to explore whether the interior features of the multipass translocon cavity similarly facilitate productive TMD–TMD interaction to catalyse intramembrane folding. At present, the bacterial holotranslocon has been visualized only at low resolution, so the placement of its constituents must be considered provisional[140]. Nevertheless, it seems to also contain a lipid-filled cavity, potentially providing a protected site for membrane protein folding[142].

Crosslinking analyses of multipass membrane proteins suggest that different TMDs bind to and are released from multiple translocon proteins at different stages of synthesis[131,143]. Furthermore, it has been observed that up to six TMDs of a multipass membrane protein can remain in an easily extractable state at the translocon before their eventual membrane integration[144]. These findings can potentially be rationalized by a model where multiple substrate TMDs are inserted into and fold within the multipass translocon cavity. Directly testing this idea will require increasingly precise assays for different steps in multipass membrane protein insertion and folding, combined with the capacity to generate, analyse and structurally characterize key intermediates in the process.

Membrane protein assembly

Roughly half of membrane proteins are part of multiprotein complexes, many with other membrane proteins[145–147]. The mechanism of assembly of two or more membrane proteins within the lipid bilayer is poorly understood. The simplest model is that the individual subunits diffuse until they encounter their partner or partners. This mechanism, although plausible for certain simple complexes or in reconstituted systems, is unlikely to be the primary strategy used in a crowded cellular environment. The main reason is that unassembled orphan subunits are prone to aggregation and are recognized by quality control systems, as evident from

Assembly factor
A factor that promotes the assembly of two or more proteins, possibly by temporarily shielding their inter-subunit interfaces.

the well-known phenomenon of subunit degradation when its interaction partner is eliminated[148–150]. Avoiding these off-pathway fates between completion of synthesis of all subunits of a multiprotein complex and their assembly probably requires chaperones to temporarily shield assembly intermediates from recognition by quality control pathways.

A 'placeholder' mechanism for subunit stabilization. Although assembly chaperones for membrane proteins are not well defined, hypotheses about their plausible characteristics and mechanism can be gleaned from known factors for soluble protein complex assembly. For example, haemoglobin assembly is facilitated by an assembly factor that specifically binds to and temporarily shields the surface of the α-subunit intended for interaction with the β-subunit[151,152]. During EMC assembly, cytosolic EMC2 temporarily interacts with a different assembly factor until the factor is displaced by cytosolic EMC8 binding to the same site[153]. Similarly, subunits of the 19S proteasome base interact with factors that shield regions which eventually dock onto the 20S proteasome core[154]. In each of these cases, the assembly factor acts as a temporary 'placeholder' for the subunit or subunits that will eventually bind to the surfaces covered by the assembly factor. Thus, membrane protein assembly factors may similarly bind temporarily to inter-subunit interfaces (FIG. 5f). A worthwhile endeavour is to search for such factors and analyse their role in assembly of membrane protein complexes.

Such placeholder chaperones might be generic in many cases given that intramembrane interfaces between subunits often share similar features such as polar side chains. Thus, a factor capable of dynamically covering such surfaces may be sufficient to temporarily shield a newly synthesized subunit for long enough to find its partner. Acquisition of such a chaperone at the site of synthesis (for example, as part of the multipass translocon) would provide an initial opportunity for complex assembly instead of recognition by quality control factors that should only be engaged in the case of biosynthesis failure. Thus, factors implicated in membrane protein folding are reasonable candidates for also aiding assembly of protein complexes by remaining associated with their substrates. Indeed, both EMC and YidC have been implicated in biogenesis of multiprotein complexes[155–157], although how they mediate these functions remains speculative.

In addition to shielding proteins from premature degradation, assembly chaperones may also prevent poorly hydrophobic TMDs from membrane dislocation. The single TMDs of some T cell receptor subunits[137] and certain TMDs of multipass protein subunits have been observed in the cytosol or lumen in their unassembled state[135,136,158]. Because topologically incorrect subunits would not be able to assemble, it is attractive to posit the existence of factors that stabilize metastable topologies that are conducive to associating with their stabilizing interaction partners. This too could be performed potentially by multipass translocon subunits that remain associated after translocon disassembly at the end of substrate synthesis. As with cytosolic protein complexes, some of these stabilizing chaperones might be substrate specific. Indeed, T cell receptor subunits engage a yet unidentified T cell-specific ER factor during, but not after, assembly[159].

Assembly factor discovery from neurobiology and virology. In addition to insertases and chaperones potentially doubling as assembly factors, other candidates for assembly factors can be mined from orthogonal fields. For example, ion channels, neurotransmitter receptors and sensory receptors have been extensively analysed for factors involved in their productive expression. Proteomic, loss-of-function and gain-of-function screens have found a variety of genes, many of which operate in the secretory and endocytic pathways[155,160–167]. Conserved ER-resident protein hits are excellent candidates for factors involved in the early biogenesis of receptors or channels. In support of this idea, EMC subunits, PAT complex subunits and other components of the multipass translocon were hits in such screens[155,165–168]. Dissection of other hits from these and similar studies may yield insights that have general applicability.

For different reasons, another field from which biogenesis factors consistently emerge is virology. Numerous genetic screens for host factors involved in the life cycles of enveloped viruses also identified ER biogenesis factors such as the signal peptidase complex, oligosaccharyltransferase complex and EMC[169–172]. Viral membrane proteins are sometimes complicated, produced from polyproteins and therefore likely to require host chaperones and assembly factors for their correct biogenesis. Indeed, there is a long history of investigating protein biogenesis in the ER by the study of virus glycoproteins[173] such as influenza haemagglutinin[174,175]. Mining the many functional screens for the involvement of ER-resident host factors in propagating viral infection, then analysing them in focused biochemical assays, may yield one route to understanding assembly of multiprotein complexes.

Conclusions and outlook

The steps of membrane protein biogenesis from targeting to assembly are understood in decreasing mechanistic depth. The initial concept of membrane protein biogenesis[176], where a single linear pathway and uniform machinery sequentially interprets hydrophobic elements as they emerge from the ribosome, has been diversified into multiple pathways at each step. Revealing how the nascent polypeptide is routed towards one machinery versus another on the basis of its sequence features, has emerged as a key goal. The eventual aim is to explain in molecular terms the specific steps taken by each major class of membrane protein to achieve its final assembled state.

This aim seems to be nearing completion for most types of single-pass membrane proteins. Biochemical reconstitution, structures of key factors, in vivo analysis and evolutionary considerations now lead to a mostly unified and consistent view. Nevertheless, small membrane proteins, many of which were overlooked in early genome annotations[177,178], are poorly studied, and the basis of their biogenesis in the correct topology is

not clear. Even for seemingly settled substrates such as N_{exo} signal anchors of type III proteins, a role for EMC (and the uncertain relationship to Sec61) has only recently emerged. Furthermore, some alternative pathways for membrane protein biogenesis have been reported but remain poorly understood (BOX 1). For example, we have essentially no mechanistic information on components of the yeast SRP-independent (SND) targeting pathway, whose deletion impacts the localization of some proteins for unclear reasons.

The richest avenues for future work lie in the steps of multipass membrane protein biogenesis, including the assembly of multi-subunit complexes. Here, our understanding is still at the stage of compiling a reasonably complete list of contributing factors. Given the highly pleiotropic effects of perturbing core protein biogenesis pathways in cells, caution is warranted in assigning molecular functions to factors primarily on the basis of end point phenotypes upon their deletion in vivo. Complementary biochemical studies will be needed to demonstrate a direct effect at a specific step of substrate biogenesis. This proven two-pronged strategy that was so successful in dissecting simpler biogenesis pathways will undoubtedly be valuable for multipass membrane proteins. Thus, devising sensitive quantitative cellular reporters and setting up biochemical reconstitution systems where protein topology and folding can be precisely evaluated will be important.

Published online: 23 September 2021

1. von Heijne, G. The membrane protein universe: what's out there and why bother? *J. Intern. Med.* **261**, 543–557 (2007).
2. Xie, K. & Dalbey, R. E. Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. *Nat. Rev. Microbiol.* **6**, 234–244 (2008).
3. Shao, S. & Hegde, R. S. Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* **27**, 25–56 (2011).
4. Baum, D. A. & Baum, B. An inside-out origin for the eukaryotic cell. *BMC Biol.* **12**, 76 (2014).
5. Wiedemann, N. & Pfanner, N. Mitochondrial machineries for protein import and assembly. *Annu. Rev. Biochem.* **86**, 685–714 (2017).
6. Gupta, A. & Becker, T. Mechanisms and pathways of mitochondrial outer membrane protein biogenesis. *Biochim. Biophys. Acta Bioenerg.* **1862**, 148323 (2021).
7. Kim, J., Na, Y. J., Park, S. J., Baek, S. H. & Kim, D. H. Biogenesis of chloroplast outer envelope membrane proteins. *Plant. Cell Rep.* **38**, 783–792 (2019).
8. Xu, X., Ouyang, M., Lu, D., Zheng, C. & Zhang, L. Protein sorting within chloroplasts. *Trends Cell Biol.* **31**, 9–16 (2021).
9. Mayerhofer, P. U. Targeting and insertion of peroxisomal membrane proteins: ER trafficking versus direct delivery to peroxisomes. *Biochim. Biophys. Acta* **1863**, 870–880 (2016).
10. White, S. H. & von Heijne, G. Transmembrane helices before, during, and after insertion. *Curr. Opin. Struct. Biol.* **15**, 378–386 (2005).
11. Walther, D. M., Rapaport, D. & Tommassen, J. Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell Mol. Life Sci.* **66**, 2789–2804 (2009).
12. Noinaj, N., Gumbart, J. C. & Buchanan, S. K. The beta-barrel assembly machinery in motion. *Nat. Rev. Microbiol.* **15**, 197–204 (2017).
13. Almagro Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
14. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
15. Nyathi, Y., Wilkinson, B. M. & Pool, M. R. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim. Biophys. Acta* **1833**, 2392–2402 (2013).
16. Keenan, R. J., Freymann, D. M., Stroud, R. M. & Walter, P. The signal recognition particle. *Annu. Rev. Biochem.* **70**, 755–775 (2001).
17. Zhang, X. & Shan, S. O. Fidelity of cotranslational protein targeting by the signal recognition particle. *Annu. Rev. Biophys.* **43**, 381–408 (2014).
18. Halic, M. et al. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* **427**, 808–814 (2004).
19. Schaffitzel, C. et al. Structure of the E. coli signal recognition particle bound to a translating ribosome. *Nature* **444**, 503–506 (2006).
20. Voorhees, R. M. & Hegde, R. S. Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. *eLife* https://doi.org/10.7554/eLife.07975 (2015).
Cryo-EM structures of biochemically defined states of the mammalian SRP–ribosome complex are presented.
21. Goder, V., Crottet, P. & Spiess, M. In vivo kinetics of protein targeting to the endoplasmic reticulum determined by site-specific phosphorylation. *EMBO J.* **19**, 6704–6712 (2000).
22. Guna, A. & Hegde, R. S. Transmembrane domain recognition during membrane protein biogenesis and quality control. *Curr. Biol.* **28**, R498–R511 (2018).
23. Guna, A., Volkmar, N., Christianson, J. C. & Hegde, R. S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **359**, 470–473 (2018).
Cell-based assays and biochemical reconstitution reveal EMC as a bona fide insertase for TA membrane proteins with moderately hydrophobic TMDs.
24. Keenan, R. J., Freymann, D. M., Walter, P. & Stroud, R. M. Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell* **94**, 181–191 (1998).
25. Hainzl, T., Huang, S., Merilainen, G., Brannstrom, K. & Sauer-Eriksson, A. E. Structural basis of signal-sequence recognition by the signal recognition particle. *Nat. Struct. Mol. Biol.* **18**, 389–391 (2011).
26. Janda, C. Y. et al. Recognition of a signal peptide by the signal recognition particle. *Nature* **465**, 507–510 (2010).
27. Wiedmann, B., Sakai, H., Davis, T. A. & Wiedmann, M. A protein complex required for signal-sequence-specific sorting and translocation. *Nature* **370**, 434–440 (1994).
28. Gamerdinger, M., Hanebuth, M. A., Frickey, T. & Deuerling, E. The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. *Science* **348**, 201–207 (2015).
29. Hsieh, H. H., Lee, J. H., Chandrasekar, S. & Shan, S. O. A ribosome-associated chaperone enables substrate triage in a cotranslational protein targeting complex. *Nat. Commun.* **11**, 5840 (2020).
30. Gamerdinger, M. et al. Early scanning of nascent polypeptides inside the ribosomal tunnel by NAC. *Mol. Cell* **75**, 996–1006 e1008 (2019).
31. Moller, I. et al. Unregulated exposure of the ribosomal M-site caused by NAC depletion results in delivery of non-secretory polypeptides to the Sec61 complex. *FEBS Lett.* **441**, 1–5 (1998).
32. Beatrix, B., Sakai, H. & Wiedmann, M. The alpha and beta subunit of the nascent polypeptide-associated complex have distinct functions. *J. Biol. Chem.* **275**, 37838–37845 (2000).
33. Powers, T. & Walter, P. The nascent polypeptide-associated complex modulates interactions between the signal recognition particle and the ribosome. *Curr. Biol.* **6**, 331–338 (1996).
34. Lauring, B., Sakai, H., Kreibich, G. & Wiedmann, M. Nascent polypeptide-associated complex protein prevents mistargeting of nascent chains to the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA* **92**, 5411–5415 (1995).
35. Schibich, D. et al. Global profiling of SRP interaction with nascent polypeptides. *Nature* **536**, 219–223 (2016).
Selective ribosome profiling in *Escherichia coli* shows that SRP targets inner membrane proteins but not periplasmic or outer membrane proteins.
36. Chartron, J. W., Hunt, K. C. & Frydman, J. Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* **536**, 224–228 (2016).
Selective ribosome profiling in yeast cells demonstrates SRP binding to nascent secretory proteins containing signal peptides or TMDs.
37. Costa, E. A., Subramanian, K., Nunnari, J. & Weissman, J. S. Defining the physiological role of SRP in protein-targeting efficiency and specificity. *Science* **359**, 689–692 (2018).
Acute SRP depletion and proximity-specific ribosome profiling in yeast reveal SRP-dependent targeting of nascent secretory proteins with signal peptides or TMDs.
38. Lee, H. C. & Bernstein, H. D. The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc. Natl Acad. Sci. USA* **98**, 3471–3476 (2001).
39. Hegde, R. S. & Keenan, R. J. Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* **12**, 787–798 (2011).
40. Shan, S. O. Guiding tail-anchored membrane proteins to the endoplasmic reticulum in a chaperone cascade. *J. Biol. Chem.* **294**, 16577–16586 (2019).
41. Borgese, N., Coy-Vergara, J., Colombo, S. F. & Schwappach, B. The ways of tails: the GET pathway and more. *Protein J.* **38**, 289–305 (2019).
42. Stefanovic, S. & Hegde, R. S. Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* **128**, 1147–1159 (2007).
43. Wang, F., Brown, E. C., Mak, G., Zhuang, J. & Denic, V. A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Mol. Cell* **40**, 159–171 (2010).
44. Mariappan, M. et al. A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466**, 1120–1124 (2010).
45. Mateja, A. et al. Protein targeting. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science* **347**, 1152–1155 (2015).
Biochemical reconstitution and crystallographic analysis reveal how the TMD of a TA membrane protein binds to a large and dynamic hydrophobic groove in Get3.
46. Mateja, A. et al. The structural basis of tail-anchored membrane protein recognition by Get3. *Nature* **461**, 361–366 (2009).
47. Gristick, H. B. et al. Crystal structure of ATP-bound Get3-Get4-Get5 complex reveals regulation of Get3 by Get4. *Nat. Struct. Mol. Biol.* **21**, 437–442 (2014).
48. Schuldiner, M. et al. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* **134**, 634–645 (2008).
49. Yamamoto, Y. & Sakisaka, T. Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol. Cell* **48**, 387–397 (2012).
50. Mock, J. Y. et al. Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. *Proc. Natl Acad. Sci. USA* **112**, 106–111 (2015).
51. Shao, S., Rodrigo-Brenni, M. C., Kivlen, M. H. & Hegde, R. S. Mechanistic basis for a molecular triage reaction. *Science* **355**, 298–302 (2017).

52. Zhang, Y. et al. Ribosome-bound Get4/5 facilitates the capture of tail-anchored proteins by Sgt2 in yeast. *Nat. Commun.* **12**, 782 (2021).
53. Chartron, J. W., VanderVelde, D. G. & Clemons, W. M. Jr Structures of the Sgt2/SGTA dimerization domain with the Get5/UBL4A UBL domain reveal an interaction that forms a conserved dynamic interface. *Cell Rep.* **2**, 1620–1632 (2012).
54. Cho, H. & Shan, S. O. Substrate relay in an Hsp70-cochaperone cascade safeguards tail-anchored membrane protein targeting. *EMBO J.* https://doi.org/10.15252/embj.201899264 (2018).
55. Rao, M. et al. Multiple selection filters ensure accurate tail-anchored membrane protein targeting. *eLife* https://doi.org/10.7554/eLife.21301 (2016).
56. Lin, K. F., Fry, M. Y., Saladi, S. M. & Clemons, W. M. Jr Molecular basis of tail-anchored integral membrane protein recognition by the cochaperone Sgt2. *J. Biol. Chem.* https://doi.org/10.1016/j.jbc.2021.100441 (2021).
57. Chio, U. S., Chung, S., Weiss, S. & Shan, S. O. A chaperone lid ensures efficient and privileged client transfer during tail-anchored protein targeting. *Cell Rep.* **26**, 37–44 e37 (2019).
58. O'Donnell, J. P. et al. The architecture of EMC reveals a path for membrane protein insertion. *eLife* https://doi.org/10.7554/eLife.57887 (2020).
Structural and biochemical analysis of human EMC identifies a TMD-binding site in the cytosolic domain of EMC, and potential routes into the bilayer.
59. Pleiner, T. et al. Structural basis for membrane insertion by the human ER membrane protein complex. *Science* **369**, 433–436 (2020).
Cryo-EM structure and functional analysis of human EMC demonstrate the importance of the membrane-embedded hydrophilic vestibule of EMC3 for TA protein insertion.
60. Itakura, E. et al. Ubiquilins chaperone and triage mitochondrial membrane proteins for degradation. *Mol. Cell* **63**, 21–33 (2016).
61. Abell, B. M., Rabu, C., Leznicki, P., Young, J. C. & High, S. Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. *J. Cell Sci.* **120**, 1743–1751 (2007).
62. Rabu, C., Wipf, P., Brodsky, J. L. & High, S. A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. *J. Biol. Chem.* **283**, 27504–27513 (2008).
63. Aschtgen, M. S., Zoued, A., Lloubes, R., Journet, L. & Cascales, E. The C-tail anchored TssL subunit, an essential protein of the enteroaggregative Escherichia coli Sci-1 type VI secretion system, is inserted by YidC. *Microbiologyopen* **1**, 71–82 (2012).
64. Peschke, M. et al. SRP, FtsY, DnaK and YidC are required for the biogenesis of the E. coli tail-anchored membrane proteins DjlC and Flk. *J. Mol. Biol.* **430**, 389–403 (2018).
65. Lee, J., Kim, D. H. & Hwang, I. Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. *Front. Plant. Sci.* **5**, 173 (2014).
66. Becker, T., Song, J. & Pfanner, N. Versatility of preprotein transfer from the cytosol to mitochondria. *Trends Cell Biol.* **29**, 534–548 (2019).
67. Bykov, Y. S., Rapaport, D., Herrmann, J. M. & Schuldiner, M. Cytosolic events in the biogenesis of mitochondrial proteins. *Trends Biochem. Sci.* **45**, 650–667 (2020).
68. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).
69. Wimley, W. C. & White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **3**, 842–848 (1996).
70. Engelman, D. M., Steitz, T. A. & Goldman, A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys Chem.* **15**, 321–353 (1986).
71. Anghel, S. A., McGilvray, P. T., Hegde, R. S. & Keenan, R. J. Identification of Oxa1 homologs operating in the eukaryotic endoplasmic reticulum. *Cell Rep.* **21**, 3708–3716 (2017).
Reports the discovery of Get1, EMC3 and TMCO1 as remote Oxa1 homologues in the ER, defining the evolutionarily ancient Oxa1 superfamily.
72. Lewis, A. J. O. & Hegde, R. S. A unified evolutionary origin for SecY and YidC. Preprint at *bioRxiv* https://doi.org/10.1101/2020.12.20.422553 (2020).
73. Wu, X. & Rapoport, T. A. Translocation of proteins through a distorted lipid bilayer. *Trends Cell Biol.* https://doi.org/10.1016/j.tcb.2021.01.002 (2021).
74. Rapoport, T. A., Li, L. & Park, E. Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* **33**, 369–390 (2017).
75. Van den Berg, B. et al. X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).
76. Voorhees, R. M., Fernandez, I. S., Scheres, S. H. & Hegde, R. S. Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* **157**, 1632–1643 (2014).
77. Li, L. et al. Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* **531**, 395–399 (2016).
Crystal structure of SecY engaged with SecA and an engineered secretory protein segment suggests how hydrophobic signals exit the lateral gate.
78. Voorhees, R. M. & Hegde, R. S. Structure of the Sec61 channel opened by a signal sequence. *Science* **351**, 88–91 (2016).
Cryo-EM structure of ribosome-bound mammalian Sec61 engaged with a signal peptide suggests how hydrophobic signals access the membrane.
79. Liaci, A. M. et al. Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. *bioRxiv* https://doi.org/10.1101/2020.11.11.378711 (2020).
80. Hessa, T., White, S. H. & von Heijne, G. Membrane insertion of a potassium-channel voltage sensor. *Science* **307**, 1427 (2005).
81. Spiess, M. & Lodish, H. F. An internal signal sequence: the asialoglycoprotein receptor membrane anchor. *Cell* **44**, 177–185 (1986).
82. 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**, 743–750 (1993).
83. McKenna, M., Simmonds, R. E. & High, S. Mycolactone reveals the substrate-driven complexity of Sec61-dependent transmembrane protein biogenesis. *J. Cell Sci.* **130**, 1307–1320 (2017).
Biochemical analysis shows that the effect on membrane protein biogenesis of mycolactone, a small-molecule Sec61 inhibitor, depends on how the nascent chain initially engages with Sec61.
84. Zong, G. et al. Ipomoeassin F binds sec61alpha to inhibit protein translocation. *J. Am. Chem. Soc.* **141**, 8450–8461 (2019).
85. Tranter, D. et al. Coibamide A targets Sec61 to prevent biogenesis of secretory and membrane proteins. *ACS Chem. Biol.* **15**, 2125–2136 (2020).
86. Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S. & Hegde, R. S. EMC is required to initiate accurate membrane protein topogenesis. *Cell* **175**, 1507–1519 e1516 (2018).
Cell-based assays and biochemical reconstitution define the role of EMC as a co-translational insertase for the first TMD of many multipass membrane proteins, including G-protein-coupled receptorss.
87. Fons, R. D., Bogert, B. A. & Hegde, R. S. Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J. Cell Biol.* **160**, 529–539 (2003).
88. Voigt, S., Jungnickel, B., Hartmann, E. & Rapoport, T. A. 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 (1996).
89. Itskanov, S., Kuo, K. M., Gumbart, J. C. & Park, E. Stepwise gating of the Sec61 protein-conducting channel by Sec63 and Sec62. *Nat. Struct. Mol. Biol.* **28**, 162–172 (2021).
90. Weng, T. H. et al. Architecture of the active post-translational Sec translocon. *EMBO J.* **40**, e105643 (2021).
91. Schorr, S. et al. Identification of signal peptide features for substrate specificity in human Sec62/Sec63-dependent ER protein import. *FEBS J.* **287**, 4612–4640 (2020).
92. Do, H., Falcone, D., Lin, J., Andrews, D. W. & Johnson, A. E. The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**, 369–378 (1996).
93. Heinrich, S. U., Mothes, W., Brunner, J. & Rapoport, T. A. The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* **102**, 233–244 (2000).
94. Mariappan, M. et al. The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature* **477**, 61–66 (2011).
95. Wang, F., Chan, C., Weir, N. R. & Denic, V. The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* **512**, 441–444 (2014).
Cell-based assays and biochemical reconstitution establish the Get1–Get2 complex as a bona fide insertase that guides the TMDs of TA proteins into the lipid bilayer.
96. Wang, F., Whynot, A., Tung, M. & Denic, V. The mechanism of tail-anchored protein insertion into the ER membrane. *Mol. Cell* **43**, 738–750 (2011).
97. Stefer, S. et al. Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. *Science* **333**, 758–762 (2011).
98. Rome, M. E., Chio, U. S., Rao, M., Gristick, H. & Shan, S. O. Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. *Proc. Natl Acad. Sci. USA* **111**, E4929–4935 (2014).
99. McDowell, M. A. et al. Structural basis of tail-anchored membrane protein biogenesis by the GET insertase complex. *Mol. Cell* **80**, 72–86 e77 (2020).
Cryo-EM structures of the yeast and human GET1–GET2–GET3 complexes reveal a heterotetrameric membrane assembly with hydrophilic vestibules into the bilayer.
100. Zalisko, B. E., Chan, C., Denic, V., Rock, R. S. & Keenan, R. J. Tail-anchored protein insertion by a single Get1/2 heterodimer. *Cell Rep.* **20**, 2287–2293 (2017).
101. Kedrov, A. et al. Elucidating the native architecture of the YidC: ribosome complex. *J. Mol. Biol.* **425**, 4112–4124 (2013).
102. Wickles, S. et al. A structural model of the active ribosome-bound membrane protein insertase YidC. *eLife* **3**, e03035 (2014).
103. Itoh, Y. et al. Mechanism of membrane-tethered mitochondrial protein synthesis. *Science* **371**, 846–849 (2021).
104. McGilvray, P. T. et al. An ER translocon for multi-pass membrane protein biogenesis. *eLife* https://doi.org/10.7554/eLife.56889 (2020).
Discovery that TMCO1 operates as part of a large, multi-subunit translocon involved in the biogenesis of most mammalian multipass membrane proteins.
105. Miller-Vedam, L. E. et al. Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *eLife* https://doi.org/10.7554/eLife.62611 (2020).
Cryo-EM structures of yeast and human EMC coupled with mutational analysis suggest how different regions of the multifuctional complex function during biogenesis of its membrane protein clients.
106. Bai, L., You, Q., Feng, X., Kovach, A. & Li, H. Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature* **584**, 475–478 (2020).
Cryo-EM structure and functional analysis of yeast EMC implicate the hydrophilic vestibule of EMC3 in TMD insertion and highlight a potential gating role for the EMC4 subunit.
107. Rivera-Monroy, J. et al. Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins in vivo. *Sci. Rep.* **6**, 39464 (2016).
108. Jonikas, M. C. et al. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* **323**, 1693–1697 (2009).
109. Volkmar, N. & Christianson, J. C. Squaring the EMC - how promoting membrane protein biogenesis impacts cellular functions and organismal homeostasis. *J. Cell Sci.* https://doi.org/10.1242/jcs.243519 (2020).
110. Chitwood, P. J. & Hegde, R. S. The role of EMC during membrane protein biogenesis. *Trends Cell Biol.* **29**, 371–384 (2019).
111. Bai, L. & Li, H. Cryo-EM structures of the endoplasmic reticulum membrane complex. *FEBS J.* https://doi.org/10.1111/febs.15786 (2021).
112. Kobayashi, K. et al. Structure of a prehandover mammalian ribosomal SRP.SRP receptor targeting complex. *Science* **360**, 323–327 (2018).
113. Prinz, A., Behrens, C., Rapoport, T. A., Hartmann, E. & Kalies, K. U. Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. *EMBO J.* **19**, 1900–1906 (2000).
114. Bonnefoy, N., Chalvet, F., Hamel, P., Slonimski, P. P. & Dujardin, G. OXA1, a Saccharomyces cerevisiae nuclear gene whose sequence is conserved from prokaryotes to eukaryotes controls cytochrome oxidase biogenesis. *J. Mol. Biol.* **239**, 201–212 (1994).

115. Sundberg, E. et al. ALBINO3, an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. *Plant. Cell* **9**, 717–730 (1997).
116. Samuelson, J. C. et al. YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637–641 (2000).
117. Hell, K., Neupert, W. & Stuart, R. A. Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *EMBO J.* **20**, 1281–1288 (2001).
118. Moore, M., Harrison, M. S., Peterson, E. C. & Henry, R. Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. *J. Biol. Chem.* **275**, 1529–1532 (2000).
119. Kumazaki, K. et al. Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Sci. Rep.* **4**, 7299 (2014).
120. Kumazaki, K. et al. Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509**, 516–520 (2014).
The high-resolution crystal structure of bacterial YidC reveals a membrane-embedded hydrophilic vestibule proposed to function during TMD insertion.
121. Borowska, M. T., Dominik, P. K., Anghel, S. A., Kossiakoff, A. A. & Keenan, R. J. A YidC-like protein in the archaeal plasma membrane. *Structure* **23**, 1715–1724 (2015).
122. 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).
123. Shanmugam, S. K. et al. New insights into amino-terminal translocation as revealed by the use of YidC and Sec depletion strains. *J. Mol. Biol.* **431**, 1025–1037 (2019).
124. Welte, T. et al. Promiscuous targeting of polytopic membrane proteins to SecYEG or YidC by the Escherichia coli signal recognition particle. *Mol. Biol. Cell* **23**, 464–479 (2012).
125. Facey, S. J., Neugebauer, S. A., Krauss, S. & Kuhn, A. The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of Escherichia coli. *J. Mol. Biol.* **365**, 995–1004 (2007).
126. Samuelson, J. C. et al. Function of YidC for the insertion of M13 procoat protein in Escherichia coli: translocation of mutants that show differences in their membrane potential dependence and Sec requirement. *J. Biol. Chem.* **276**, 34847–34852 (2001).
127. von Heijne, G. & Gavel, Y. Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* **174**, 671–678 (1988).
128. Petriman, N. A. et al. The interaction network of the YidC insertase with the SecYEG translocon, SRP and the SRP receptor FtsY. *Sci. Rep.* **8**, 578 (2018).
129. Kedrov, A. et al. Structural dynamics of the YidC:ribosome complex during membrane protein biogenesis. *Cell Rep.* **17**, 2943–2954 (2016).
130. Hessa, T. et al. Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030 (2007).
131. Meacock, S. L., Lecomte, F. J., Crawshaw, S. G. & High, S. Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Mol. Biol. Cell* **13**, 4114–4129 (2002).
132. Ismail, N., Crawshaw, S. G., Cross, B. C., Haagsma, A. C. & High, S. Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. *Biochem. J.* **411**, 495–506 (2008).
133. Chitwood, P. J. & Hegde, R. S. An intramembrane chaperone complex facilitates membrane protein biogenesis. *Nature* **584**, 630–634 (2020).
Discovery of the PAT complex as an intramembrane chaperone that shields hydrophilic TMDs during multipass membrane protein biogenesis.
134. Sato, B. K., Schulz, D., Do, P. H. & Hampton, R. Y. Misfolded membrane proteins are specifically recognized by the transmembrane domain of the Hrd1p ubiquitin ligase. *Mol. Cell* **34**, 212–222 (2009).
135. Inglis, A. J., Page, K. R., Guna, A. & Voorhees, R. M. Differential modes of orphan subunit recognition for the WRB/CAML complex. *Cell Rep.* **30**, 3691–3698 e3695 (2020).
136. Sun, S. & Mariappan, M. C-terminal tail length guides insertion and assembly of membrane proteins. *J. Biol. Chem.* **295**, 15498–15510 (2020).
137. Feige, M. J. & Hendershot, L. M. Quality control of integral membrane proteins by assembly-dependent membrane integration. *Mol. Cell* **51**, 297–309 (2013).
138. Yamamoto, S. et al. Contribution of calumin to embryogenesis through participation in the endoplasmic reticulum-associated degradation activity. *Dev. Biol.* **393**, 33–43 (2014).
139. Morimoto, M. et al. Bi-allelic CCDC47 variants cause a disorder characterized by woolly hair, liver dysfunction, dysmorphic features, and global developmental delay. *Am. J. Hum. Genet.* **103**, 794–807 (2018).
140. Botte, M. et al. A central cavity within the holo-translocon suggests a mechanism for membrane protein insertion. *Sci. Rep.* **6**, 38399 (2016).
141. Hayer-Hartl, M., Bracher, A. & Hartl, F. U. The GroEL-GroES chaperonin machine: a nano-cage for protein folding. *Trends Biochem. Sci.* **41**, 62–76 (2016).
142. Martin, R. et al. Structure and dynamics of the central lipid pool and proteins of the bacterial holo-translocon. *Biophys. J.* **116**, 1931–1940 (2019).
143. Sadlish, H., Pitonzo, D., Johnson, A. E. & Skach, W. R. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat. Struct. Mol. Biol.* **12**, 870–878 (2005).
144. 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).
145. Krogan, N. J. et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**, 637–643 (2006).
146. Gavin, A. C. et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006).
147. Babu, M. et al. Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. *Nature* **489**, 585–589 (2012).
148. Lippincott-Schwartz, J., Bonifacino, J. S., Yuan, L. C. & Klausner, R. D. Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins. *Cell* **54**, 209–220 (1988).
149. Juszkiewicz, S. & Hegde, R. S. Quality control of orphaned proteins. *Mol. Cell* **71**, 443–457 (2018).
150. McShane, E. et al. Kinetic analysis of protein stability reveals age-dependent degradation. *Cell* **167**, 803–815 e821 (2016).
151. Feng, L. et al. Molecular mechanism of AHSP-mediated stabilization of alpha-hemoglobin. *Cell* **119**, 629–640 (2004).
152. Kihm, A. J. et al. An abundant erythroid protein that stabilizes free alpha-haemoglobin. *Nature* **417**, 758–763 (2002).
153. Pleiner, T. et al. WNK1 is an assembly factor for the human ER membrane protein complex. *Mol. Cell* https://doi.org/10.1016/j.molcel.2021.04.013 (2021).
154. Rousseau, A. & Bertolotti, A. Regulation of proteasome assembly and activity in health and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 697–712 (2018).
155. Richard, M., Boulin, T., Robert, V. J., Richmond, J. E. & Bessereau, J. L. Biosynthesis of ionotropic acetylcholine receptors requires the evolutionarily conserved ER membrane complex. *Proc. Natl Acad. Sci. USA* **110**, E1055–1063 (2013).
156. Pop, O. I. et al. YidC is required for the assembly of the MscL homopentameric pore. *FEBS J.* **276**, 4891–4899 (2009).
157. Nishikawa, H., Kanno, K., Endo, Y. & Nishiyama, K. I. Ring assembly of c subunits of F0 F1 -ATP synthase in Propionigenium modestum requires YidC and UncI following MPIase-dependent membrane insertion. *FEBS Lett.* **595**, 647–654 (2021).
158. Coelho, J. P. L. et al. A network of chaperones prevents and detects failures in membrane protein lipid bilayer integration. *Nat. Commun.* **10**, 672 (2019).
159. Bonifacino, J. S. et al. Association and dissociation of the murine T cell receptor associated protein (TRAP). Early events in the biosynthesis of a multisubunit receptor. *J. Biol. Chem.* **263**, 8965–8971 (1988).
160. Schwenk, J. et al. An ER assembly line of AMPA-receptors controls excitatory neurotransmission and its plasticity. *Neuron* **104**, 680–692 e689 (2019).
161. Schwenk, J. et al. High-resolution proteomics unravel architecture and molecular diversity of native AMPA receptor complexes. *Neuron* **74**, 621–633 (2012).
162. Gu, S. et al. Brain alpha7 nicotinic acetylcholine receptor assembly requires NACHO. *Neuron* **89**, 948–955 (2016).
163. Halevi, S. et al. The C. elegans ric-3 gene is required for maturation of nicotinic acetylcholine receptors. *EMBO J.* **21**, 1012–1020 (2002).
164. Eimer, S. et al. Regulation of nicotinic receptor trafficking by the transmembrane Golgi protein UNC-50. *EMBO J.* **26**, 4313–4323 (2007).
165. Satoh, T., Ohba, A., Liu, Z., Inagaki, T. & Satoh, A. K. dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. *eLife* https://doi.org/10.7554/eLife.06306 (2015).
166. Gottschalk, A. et al. Identification and characterization of novel nicotinic receptor-associated proteins in Caenorhabditis elegans. *EMBO J.* **24**, 2566–2578 (2005).
167. Talbot, B. E., Vandorpe, D. H., Stotter, B. R., Alper, S. L. & Schlondorff, J. S. Transmembrane insertases and N-glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. *J. Biol. Chem.* **294**, 12655–12669 (2019).
Genome-wide screen for factors that impair surface expression of a multipass transient receptor potential channel mutant yields components of the EMC, the SND pathway, the PAT complex and the multipass translocon.
168. Kamat, S., Yeola, S., Zhang, W., Bianchi, L. & Driscoll, M. NRA-2, a nicalin homolog, regulates neuronal death by controlling surface localization of toxic Caenorhabditis elegans DEG/ENaC channels. *J. Biol. Chem.* **289**, 11916–11926 (2014).
169. Savidis, G. et al. Identification of zika virus and dengue virus dependency factors using functional genomics. *Cell Rep.* **16**, 232–246 (2016).
170. Marceau, C. D. et al. Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. *Nature* **535**, 159–163 (2016).
171. Labeau, A. et al. A genome-wide CRISPR-Cas9 screen identifies the dolichol-phosphate mannose synthase complex as a host dependency factor for dengue virus infection. *J. Virol.* https://doi.org/10.1128/JVI.01751-19 (2020).
172. Schneider, W. M. et al. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **184**, 120–132 e114 (2021).
173. Doms, R. W., Lamb, R. A., Rose, J. K. & Helenius, A. Folding and assembly of viral membrane proteins. *Virology* **193**, 545–562 (1993).
174. Daniels, R., Kurowski, B., Johnson, A. E. & Hebert, D. N. N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. *Mol. Cell* **11**, 79–90 (2003).
175. Maggioni, M. C., Liscaljet, I. M. & Braakman, I. A critical step in the folding of influenza virus HA determined with a novel folding assay. *Nat. Struct. Mol. Biol.* **12**, 258–263 (2005).
176. Blobel, G. Intracellular protein topogenesis. *Proc. Natl Acad. Sci. USA* **77**, 1496–1500 (1980).
177. Kemp, G. & Cymer, F. Small membrane proteins - elucidating the function of the needle in the haystack. *Biol. Chem.* **395**, 1365–1377 (2014).
178. Chen, J. et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140–1146 (2020).
179. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
180. Huber, D. et al. SecA cotranslationally interacts with nascent substrate proteins in vivo. *J. Bacteriol.* https://doi.org/10.1128/JB.00622-16 (2017).
181. Wang, S., Yang, C. I. & Shan, S. O. SecA mediates cotranslational targeting and translocation of an inner membrane protein. *J. Cell Biol.* **216**, 3639–3653 (2017).
182. Wang, S. et al. The molecular mechanism of cotranslational membrane protein recognition and targeting by SecA. *Nat. Struct. Mol. Biol.* **26**, 919–929 (2019).
Cryo-EM structure and biochemical analysis provide mechanistic insight into how SecA mediates co-translational targeting of certain membrane proteins in *E. coli*.
183. Gelis, I. et al. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**, 756–769 (2007).
184. Chatzi, K. E., Sardis, M. F., Economou, A. & Karamanou, S. SecA-mediated targeting and translocation of secretory proteins. *Biochim. Biophys. Acta* **1843**, 1466–1474 (2014).
185. Ast, T., Cohen, G. & Schuldiner, M. A network of cytosolic factors targets SRP-independent proteins

to the endoplasmic reticulum. *Cell* **152**, 1134–1145 (2013).
186. Aviram, N. et al. The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* **540**, 134–138 (2016).
187. Hassdenteufel, S. et al. hSnd2 protein represents an alternative targeting factor to the endoplasmic reticulum in human cells. *FEBS Lett.* **591**, 3211–3224 (2017).
188. Phillips, B. P. & Miller, E. A. Ribosome-associated quality control of membrane proteins at the endoplasmic reticulum. *J. Cell Sci.* https://doi.org/10.1242/jcs.251983 (2020).
189. Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nat. Struct. Mol. Biol.* **23**, 7–15 (2016).
190. Yip, M. C. J. & Shao, S. Detecting and rescuing stalled ribosomes. *Trends Biochem. Sci.* https://doi.org/10.1016/j.tibs.2021.03.008 (2021).
191. Joazeiro, C. A. P. Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.* **20**, 368–383 (2019).
192. Hessa, T. et al. Protein targeting and degradation are coupled for elimination of mislocalized proteins. *Nature* **475**, 394–397 (2011).
193. Rodrigo-Brenni, M. C., Gutierrez, E. & Hegde, R. S. Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to Bag6. *Mol. Cell* **55**, 227–237 (2014).
194. Dederer, V. et al. Cooperation of mitochondrial and ER factors in quality control of tail-anchored proteins. *Elife* **8**, https://doi.org/10.7554/eLife.45506 (2019).
195. Matsumoto, S. et al. Msp1 clears mistargeted proteins by facilitating their transfer from mitochondria to the ER. *Mol. Cell* **76**, 191–205 e110 (2019).
196. Stefanovic-Barrett, S. et al. MARCH6 and TRC8 facilitate the quality control of cytosolic and tail-anchored proteins. *EMBO Rep.* https://doi.org/10.15252/embr.201745603 (2018).
197. Leto, D. E. et al. Genome-wide CRISPR analysis identifies substrate-specific conjugation modules in ER-associated degradation. *Mol. Cell* **73**, 377–389 e311 (2019).
198. McKenna, M. J. et al. The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science* https://doi.org/10.1126/science.abc5809 (2020).
199. Chen, Y. C. et al. Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. *EMBO J.* **33**, 1548–1564 (2014).
200. Okreglak, V. & Walter, P. The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proc. Natl Acad. Sci. USA* **111**, 8019–8024 (2014).
201. Wohlever, M. L., Mateja, A., McGilvray, P. T., Day, K. J. & Keenan, R. J. Msp1 is a membrane protein dislocase for tail-anchored proteins. *Mol. Cell* **67**, 194–202 e196 (2017).
202. Wu, X. & Rapoport, T. A. Mechanistic insights into ER-associated protein degradation. *Curr. Opin. Cell Biol.* **53**, 22–28 (2018).
203. Christianson, J. C. & Ye, Y. Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nat. Struct. Mol. Biol.* **21**, 325–335 (2014).
204. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
205. Beck, K. et al. YidC, an assembly site for polytopic Escherichia coli membrane proteins located in immediate proximity to the SecYE translocon and lipids. *EMBO Rep.* **2**, 709–714 (2001).
206. Yi, L. et al. YidC is strictly required for membrane insertion of subunits a and c of the F_1F_0 ATP synthase and SecE of the SecYEG translocase. *Biochemistry* **42**, 10537–10544 (2003).
207. van Bloois, E., Jan Haan, G., de Gier, J. W., Oudega, B. & Luirink, J. F_1F_0 ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane. *FEBS Lett.* **576**, 97–100 (2004).
208. Celebi, N., Yi, L., Facey, S. J., Kuhn, A. & Dalbey, R. E. Membrane biogenesis of subunit II of cytochrome bo oxidase: contrasting requirements for insertion of N-terminal and C-terminal domains. *J. Mol. Biol.* **357**, 1428–1436 (2006).
209. van Bloois, E., Haan, G. J., de Gier, J. W., Oudega, B. & Luirink, J. Distinct requirements for translocation of the N-tail and C-tail of the Escherichia coli inner membrane protein CyoA. *J. Biol. Chem.* **281**, 10002–10009 (2006).
210. Wagner, S. et al. Biogenesis of MalF and the $MalFGK_2$ maltose transport complex in Escherichia coli requires YidC. *J. Biol. Chem.* **283**, 17881–17890 (2008).
211. Zhu, L., Kaback, H. R. & Dalbey, R. E. YidC protein, a molecular chaperone for LacY protein folding via the SecYEG protein machinery. *J. Biol. Chem.* **288**, 28180–28194 (2013).
212. Nagamori, S., Smirnova, I. N. & Kaback, H. R. Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.* **165**, 53–62 (2004).
213. du Plessis, D. J., Nouwen, N. & Driessen, A. J. Subunit a of cytochrome o oxidase requires both YidC and SecYEG for membrane insertion. *J. Biol. Chem.* **281**, 12248–12252 (2006).
214. van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N. & Driessen, A. J. F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J. Cell Biol.* **165**, 213–222 (2004).
215. Serek, J. et al. Escherichia coli YidC is a membrane insertase for Sec-independent proteins. *EMBO J.* **23**, 294–301 (2004).
216. Stiegler, N., Dalbey, R. E. & Kuhn, A. M13 procoat protein insertion into YidC and SecYEG proteoliposomes and liposomes. *J. Mol. Biol.* **406**, 362–370 (2011).
217. Robinson, P. J. & Woolhead, C. A. Post-translational membrane insertion of an endogenous YidC substrate. *Biochim. Biophys. Acta* **1833**, 2781–2788 (2013).
218. Price, C. E., Kocer, A., Kol, S., van der Berg, J. P. & Driessen, A. J. In vitro synthesis and oligomerization of the mechanosensitive channel of large conductance, MscL, into a functional ion channel. *FEBS Lett.* **585**, 249–254 (2011).
219. Ernst, S., Schonbauer, A. K., Bar, G., Borsch, M. & Kuhn, A. YidC-driven membrane insertion of single fluorescent Pf3 coat proteins. *J. Mol. Biol.* **412**, 165–175 (2011).
220. Serdiuk, T. et al. Insertion and folding pathways of single membrane proteins guided by translocases and insertases. *Sci. Adv.* **5**, eaau6824 (2019).
221. Tian, S. et al. Proteomic analysis identifies membrane proteins dependent on the ER membrane protein complex. *Cell Rep.* **28**, 2517–2526 e2515 (2019).
222. Shurtleff, M. J. et al. The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife* https://doi.org/10.7554/eLife.37018 (2018).
223. Houben, E. N., ten Hagen-Jongman, C. M., Brunner, J., Oudega, B. & Luirink, J. The two membrane segments of leader peptidase partition one by one into the lipid bilayer via a Sec/YidC interface. *EMBO Rep.* **5**, 970–975 (2004).
224. Sachelaru, I. et al. YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. *J. Biol. Chem.* **288**, 16295–16307 (2013).

Acknowledgements
The authors thank members of the Hegde and Keenan laboratories for productive discussions that influenced this Review. R.S.H. is funded by the UK Medical Research Council (MC_UP_A022_1007). R.J.K. is funded by grants from the US National Institutes of Health (R01 GM086487 and R01 GM130051).

Author contributions
Both authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Molecular Cell Biology thanks the anonymous reviewers for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.