Spf1 and Ste24: quality controllers of transmembrane protein topology in the eukaryotic cell

Donald J. Tipper[1]* and Carol A. Harley[2,3]

[1]University of Massachusetts Medical School, Worcester, MA, United States, [2]i3S-Instituto de Investigação e Inovação em Saude, Universidade do Porto, Porto, Portugal, [3]IBMC-Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal

DNA replication, transcription, and translation in eukaryotic cells occur with decreasing but still high fidelity. In contrast, for the estimated 33% of the human proteome that is inserted as transmembrane (TM) proteins, insertion with a non-functional inverted topology is frequent. Correct topology is essential for function and trafficking to appropriate cellular compartments and is controlled principally by responses to charged residues within 15 residues of the inserted TM domain (TMD); the flank with the higher positive charge remains in the cytosol (inside), following the positive inside rule (PIR). Yeast (*Saccharomyces cerevisiae*) mutants that increase insertion contrary to the PIR were selected. Mutants with strong phenotypes were found only in *SPF1* and *STE24* (human cell orthologs are *ATP13A1* and *ZMPSte24*) with, at the time, no known relevant functions. Spf1/Atp13A1 is now known to dislocate to the cytosol TM proteins inserted contrary to the PIR, allowing energy-conserving reinsertion. We hypothesize that Spf1 and Ste24 both recognize the short, positively charged ER luminal peptides of TM proteins inserted contrary to the PIR, accepting these peptides into their large membrane-spanning, water-filled cavities through interaction with their many interior surface negative charges. While entry was demonstrated for Spf1, no published evidence directly demonstrates substrate entry to the Ste24 cavity, internal access to its zinc metalloprotease (ZMP) site, or active withdrawal of fragments, which may be essential for function. Spf1 and Ste24 comprise a PIR quality control system that is conserved in all eukaryotes and presumably evolved in prokaryotic progenitors as they gained differentiated membrane functions. About 75% of the PIR is imposed by this quality control system, which joins the UPR, ERAD, and autophagy (ER-phagy) in coordinated, overlapping quality control of ER protein function.

KEYWORDS

transmembrane proteins, topology, positive inside rule, quality control, topology error recognition

Introduction

Essential background to transmembrane (TM) protein insertion mechanisms can be found in reviews (von Heijne, 2006; Shao and Hegde, 2011; Aviram et al., 2017; Lewis and Hegde, 2021; Dederer and Lemberg, 2021; Hegde and Keenan, 2022) and in recent publications (McKenna et al., 2020; McKenna et al., 2022; Smalinskaite et al., 2022). A review of algorithms for predicting TM protein topology from sequence profiles was published by Jones (2007). Alphafold (Heaven, 2020) predicts TM protein structures well (Hegedus et al., 2022) but may not yet consider the effects of membrane lipid composition. A recent review (Chen et al., 2023) describes the multiple quality control

systems that recognize misfolded and aggregated endoplasmic reticulum (ER) luminal proteins: the unfolded protein response (UPR) comprises ER lumen–nucleus signaling mechanisms and re-export mechanisms; the ER-associated protein degradation (ERAD) system uses multiple ER membrane ubiquitin ligases to remove proteins that escape the UPR; and the ER-associated autophagy (ER-phagy) system delivers ERAD-resistant and aggregated ER proteins to the lysosome for destruction. The proposed PIR quality control system is potentially an important additional component.

TM protein insertion, topology, functionality, and trafficking

The overall turnover of nascent proteins in mammalian cells is estimated at 30% (Schubert et al., 2000) and, given the error-prone nature of TM protein insertion, the turnover of this 33% of the proteome may exceed this average, so one-third or more of discarded nascent proteins, about 10% of the proteome, is fodder for error correction by ERAD and, principally, by Spf1/Ste24-mediated PIR quality control. The first stage of nuclear-encoded TM protein insertion is co-translational or post-translational targeting to the ER, the outside mitochondrial membrane (OMM), chloroplasts and other plastids, or peroxisomes (Aviram et al., 2017; McKenna et al., 2020; McKenna et al., 2022; Dahan et al., 2022; Hegde and Keenan, 2022). TM protein functions are dependent on the insertion of their hydrophobic TMDs in the ER with the correct topology, which is necessary for trafficking to their cellular sites of function, where they become markers for organelle recognition. ER insertion requires cooperation between cytoplasmic chaperones preventing TMD aggregation, such as the signal recognition particle, SRP (Walter and Blobel, 1981a; Walter and Blobel, 1981b; Akopian et al., 2013), ER receptors for these chaperones, and the TM protein complexes that mediate protein insertion and secretion. The evolution of these sorting components commenced in prokaryotes before the separation of bacteria and

Archaea (Lewis et al., 2021; Hegde and Keenan, 2022). Intramembrane folding and assembly of TMDs of multiple TM proteins, or within proteins with multiple TMDs (multi-pass), is required to produce functional TM complexes (Hegde and Keenan, 2022; Smalinskaite et al., 2022). Schlebach and Sanders (2014) surveyed 470 known pathogenic mutations in five misfolding-prone TM proteins; about 10% had predicted and likely pathogenic effects on topogenesis.

Topology and folding of TM proteins

The most N-terminal TMD of a TM protein (TMD1) can insert either Ncyt with the N-terminus remaining in the cytoplasm (Rapoport et al., 2017), or Nexo, with the N-terminus inserted into the lumen of the ER. These terms are pictorially defined in Figure 1. Orientation depends on responses to topogenic signals, which are principally provided by charged residues within 15 residues of the TMD (von Heijne, 1986). The charge difference spanning the TMD (C-terminal charges minus N-terminal charges) is the determinant of the positive inside rule (PIR) (von Heijne, 1986; von Heijne, 2006; Hartmann et al., 1989; Krogh et al., 2001). *In vivo* and *in vitro* studies in bacteria with genetically modified phospholipid composition demonstrated the role of local membrane composition in topogenesis (Vitrac et al., 2013). Local membrane surface charge, principally the ratio of phosphatidylserine (negative charge) to phosphatidylethanolamine (uncharged), can affect the influence of negatively peptide charges on the PIR (White and von Heijne, 2005; Vitrac et al., 2013; Baker et al., 2017), potentially allowing topological inversion on transit to a different membrane environment. Variation can lead to dual topologies, as for diacylglycerol transferase in HepG2 cells (Wurie et al., 2011). Other TM proteins are known to have dual topology (Rapp et al., 2006). N-glycosylation can also affect topology after insertion (Goder et al., 1999). These influences
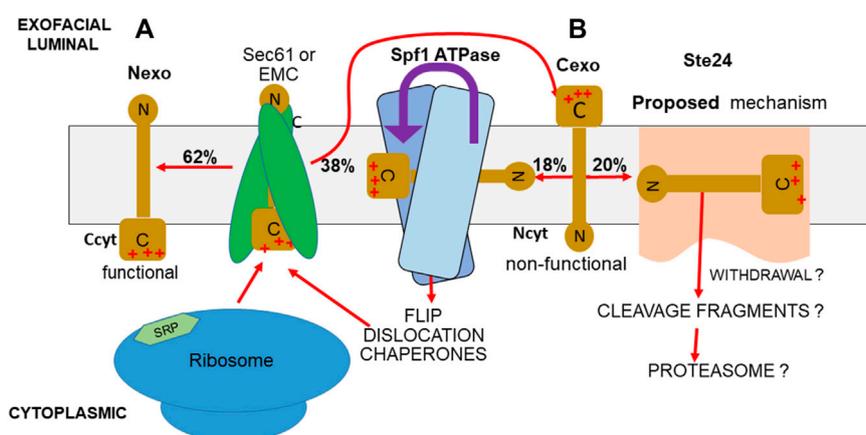
FIGURE 1
PIR quality control system. **(A)** If the more positively charged TMD-flanking peptide of a TM protein is the C-terminus, as in SPßla and SPInv fusions, the PIR requires that this flank remains in the cytoplasm (Ccyt/Nexo), in the functional orientation during insertion. **(B)** Any of this TM protein inserted with inverted topology (Ncyt/Cexo) is non-functional. Approximately half of this inverted protein is accepted by the Spf1 cavity for ATPase-driven dislocation, allowing reinsertion (McKenna et al., 2020; McKenna et al., 2022; center). Most of the rest is eliminated by Ste24. A hypothetical mechanism is the charge-directed entry into the Ste24 cavity and ZMP cleavage, possibly aided by removal by the Dfm1-recruited cdc48 AAA ATPase, followed by proteasomal disassembly (right).

on topology are incorporated into The Charge Balance Rule (Dowhan et al., 2019), a more general version of the PIR that gives equal weight to positive and negative charges while allowing for the effects of the lipid environment.

The PIR/Charge Balance rule states that insertion should leave the more positively charged TMD-flanking peptide in the cytosol. The charge difference determines the strength of the bias (Harley and Tipper, 1996; Harley et al., 1998). The PIR was deduced from empirical observations in bacteria (von Heijne, 1986), where, as in Archaea, the membrane's positive outside potential gradient provides a clear rationale. The complexity of TM protein insertion, topology determination, and effects on the topology of intramembrane complex assembly were reviewed by Shao and Hegde (2011). The hydrophobicity of the TMD modulates the response to the PIR flanking charges, and this effect is predicted by an algorithm based on amino acid insertion-energy profiles (Elazar et al., 2016). A recent review, "Folding and Misfolding of Human Membrane Proteins in Health and Disease" discussed the effects of the thermodynamics and kinetics of TM protein folding on TM protein structure (Marinko et al., 2019). While eukaryotes have multiple membranous organelles, they all lack measurable potential gradients, with the exception of the vacuole, whose proton pump maintains a positive inside-to-outside potential, rendering it topologically outside by prokaryotic criteria. Yet, the insertion of TM proteins into eukaryotic cell membranes still follows the PIR (Harley et al., 1998; Baker et al., 2017). This is an enigma.

ER translocon: Sec61

The Sec61 translocon provides the sole TM channel for protein secretion and the major channel for TM protein insertion (Hegde and Keenan, 2022; Smalinskaite et al., 2022). Sec61 evolved from the SecY complex in bacteria and Archaea (Lewis and Hegde, 2021) and exists in all cells in several functionally differing complexes (Aviram and Schuldiner, 2017). The SRP recognizes the short N-terminal, seven-to-nine residue signal peptides in secreted proteins, if sufficiently hydrophobic, and the much longer first TMD of TM proteins as they exit the ribosome, if they are not too far from the N-terminus (Halic et al., 2004). The binding of the SRP to its receptor displaces the SRP and allows the signal peptide, or TMD, to enter Sec61, initiating ribosome-driven translocation. Signal peptide recognition and characteristics were recently reviewed (Nielsen et al., 2019). Disease-associated mutations in signal peptides cause failure in SRP interaction, leading to translation failure and mRNA decay (Tikhonova et al., 2019). Charges on extra-membranous loops of Sec61 confer a modest PIR response on entry (Goder et al., 2004), presumably responsible for the PIR response at Sec61 not subsequently provided by Ste24 and Spf1. If the TMD is too far from the N-terminus or if a signal peptide's hydrophobicity is weak, post-translational protein insertion occurs using alternate chaperone complexes (Osborne et al., 2005; Aviram and Schuldiner, 2017).

Single-pass and multi-pass TM proteins

Approximately 45% of all TM proteins have a single TMD (single-pass), and these fall into several classes (Hegde and Keenan, 2022): Type 1 proteins have a signal peptide, and their TMDs exit Sec61 via the lateral gate; after signal peptide removal, they are inserted Nexo as for secreted proteins (Hegde and Kenan, 2022). In Type 2, the TMD is inserted Ncyt. The TMD of Type 3, with N-termini no longer than 50 residues, is inserted Nexo (Aviram et al., 2017), usually by the EMC complex (Chitwood and Hegde, 2019). All types are potential substrates for the PIR quality control system. SRP binding has a broad affinity range that the ER insertases must accommodate, requiring multiple, often functionally overlapping insertion mechanisms (Aviram and Schuldiner, 2017; Hegde and Keenan, 2022).

All TM classes can be multi-pass; the topology of their TMDs alternates and is determined by that of the first TMD, TMD1. After insertion, the TMD1 of a multi-pass TM protein is translocated to an insertase complex that is associated with, but separate from, Sec61 for pairwise insertion of subsequent TMDs by the GEL complex (Smalinskaite et al., 2022). Presumably, these multi-pass proteins can only be monitored by the PIR quality control system before additional TMDs are added, as the Spf1 and Ste24 cavities cannot internalize or accommodate multi-pass TM proteins. Monitoring of TMD1 topology during translocation may be particularly important as inverted multi-pass TMD1 insertion would impair subsequent TMD insertion, requiring correction, probably by ERAD.

The EMC, GET, and GEL insertases: Oxa1 family members

The ER membrane protein complex (EMC) was discovered in a screen of a yeast deletion library (Giaver et al., 2002) for genes whose deletion resulted in the accumulation of misfolded proteins in the ER lumen, leading to activation of the Hac1p transcription factor and expression of UPR genes (Jonikas et al., 2009). Components of the GET pathway (guided entry of tail-anchored proteins) were first described by Schuldiner et al. (2005). The Sec61, EMC, and GET insertase complexes, together with variants of the Sec61 complex and the GEL complex for polytopic TM protein insertion, comprise the ER TM protein insertases. The EMC and GET insertases are more efficient than Sec61 for the insertion of specific substrate classes, but the loss of neither pathway is lethal, so insertases are selective but not exclusive. The EMC, GET, and GEL insertases are members of the Oxa1 family. Oxa1 insertase (Lurink et al., 2001). Translocation often results in Nexo insertion, though Cexo insertion can also occur (Chitwood et al., 2018; Wu and Rapoport, 2018; Wu et al., 2020).

The TM components of the EMC are encoded by a set of six–seven genes, depending on the species, whose products can be immunoprecipitated as a stoichiometric TM protein complex (Jonikas et al., 2009; Chitwood and Hegde, 2019). The structure of this complex was described by Pleiner et al. (2020). Spf1 is commonly associated with the EMC (Ast et al., 2016). TM domains exiting the ribosome at the ER have a transient opportunity to dock at the EMC before docking at Sec61 (Chitwood et al., 2019; Hegde and Kenan, 2022). The EMC Oxa1 insertase (EMC3) prefers TM domains of lower hydrophobicity than does Sec61, but the functions of the EMC and Sec61 overlap. In G protein-coupled receptors (GPCRs),

negatively charged N-terminal peptides preceding TMD1 will insert in the EMC vestibule and be translocated to Nexo (Chitwood and Hegde, 2019), obeying the PIR. TMD1 is then translocated to the PAT insertase complex associated with Sec61; the subsequent six TMDs are inserted pairwise after transfer to the PAT-associated GEL complex (Smalinskaite et al., 2022). The role of the EMC in the insertion of the β1 adrenergic receptor (β1AR, a GPCR) was studied in human rough microsomes (Chitwood et al., 2018). Nexo insertion with correct topology and folding was efficient, as shown by the binding of alprenolol, a β1AR inhibitor. In *EMC5Δ* cells, in which the EMC fails to assemble, default insertion at Sec61 resulted in about 50% β1AR destruction, presumably a result of high rates of topologically erroneous Cexo insertion. Correspondingly, alprenolol binding in EMC5Δ rough microsomes also decreased by 50%, presumably due to non-functional Cexo insertion at Sec61. In a construct that contained only β1AR TMD1 and its following short cytosolic loop, *in vitro* insertion was Nexo. In the absence of the EMC, insertion was ~50% Cexo, again illustrating the high error rate of default insertion at Sec61.
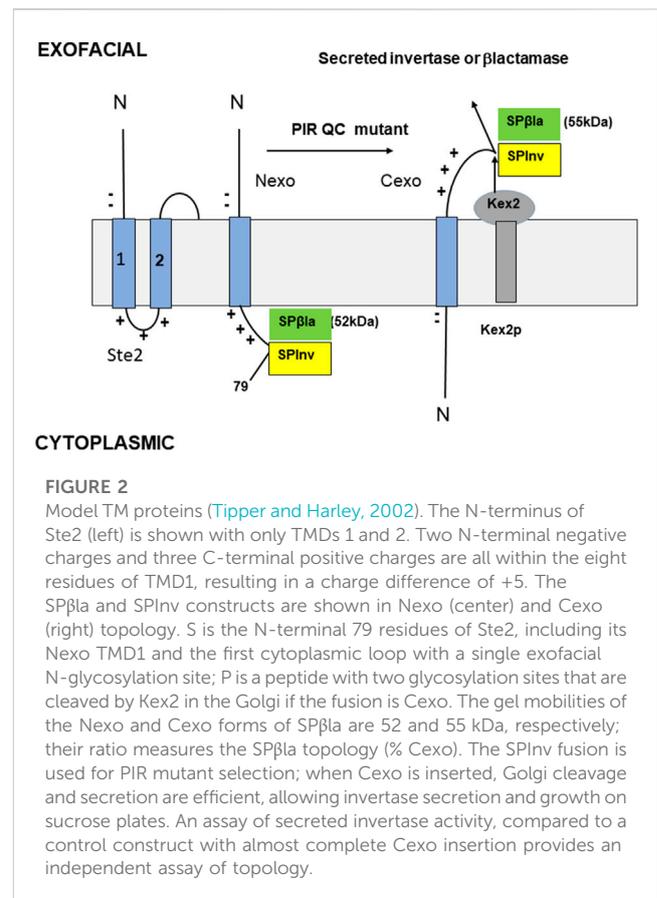
GET insertion of tail-anchored proteins

The GET pathway inserts tail-anchored proteins with strongly hydrophobic TMDs and proteins with similarly hydrophobic GPI C-terminal anchors, presenting them post-translationally for import at the GET1–GET3 complex (Schuldiner et al., 2008; Ast et al., 2016), another Oxa1 family insertase. The EMC plays a major role as an alternate to the GET pathway for tail-anchored TM protein insertion and post-translational secretion of proteins whose signals do not efficiently associate with the SRP (Baum and Baum, 2014; Wang et al., 2014; Aviram et al., 2017). Cytosolic aggregation is prevented by the binding of abundant chaperones such as HSP70s, eventually guiding export at Sec61 using ER luminal Kar2/Bip for ATP-driven ratcheted entry (Jung et al., 2013; Dederer et al., 2021).

Selection for yeast mutants unresponsive to PIR signals; the PIR-quality control system

Yeast mutants unresponsive to PIR signals were selected to discover the mechanisms for response to TMD-adjacent peptide charge signals (Tipper and Harley, 2002). At that time, Sec61 was the only established translocon; the fascinating complexity of TM protein insertion is now apparent (Smalinskaite et al., 2022). Insertion of the model TM proteins used in mutant selection involves the EMC in addition to Sec61; our data, however, demonstrate that this is irrelevant since the major determinant of the PIR is a post-insertion quality control system independent of the insertion mechanism.

The Sec61-secreted *SUC2* gene product, invertase (Inv), cleaves sucrose into glucose and fructose. Yeast lacks a sucrose transporter, so invertase is essential for growth in sucrose media. *SUC2* was expressed via a Golgi-cleavable fusion to the C-terminus of a TM protein with a strong PIR signal directing Nexo insertion. In a *SUC2Δ* strain, Nexo insertion of the fusion would prevent invertase

FIGURE 2
Model TM proteins (Tipper and Harley, 2002). The N-terminus of Ste2 (left) is shown with only TMDs 1 and 2. Two N-terminal negative charges and three C-terminal positive charges are all within the eight residues of TMD1, resulting in a charge difference of +5. The SPβla and SPInv constructs are shown in Nexo (center) and Cexo (right) topology. S is the N-terminal 79 residues of Ste2, including its Nexo TMD1 and the first cytoplasmic loop with a single exofacial N-glycosylation site; P is a peptide with two glycosylation sites that are cleaved by Kex2 in the Golgi if the fusion is Cexo. The gel mobilities of the Nexo and Cexo forms of SPβla are 52 and 55 kDa, respectively; their ratio measures the SPβla topology (% Cexo). The SPInv fusion is used for PIR mutant selection; when Cexo is inserted, Golgi cleavage and secretion are efficient, allowing invertase secretion and growth on sucrose plates. An assay of secreted invertase activity, compared to a control construct with almost complete Cexo insertion provides an independent assay of topology.

secretion (Figure 2). Growth on sucrose plates should then be selected for the desired mutations. The TM protein chosen was S (Figure 2), a 79-residue N-terminal fragment of Ste2; the GPCR receptor for the yeast α-mating pheromone encompassing its 50-residue N-terminus, TMD1, and the 8-residue following cytoplasmic loop. S is very similar to the β1AR fragment described by Chitwood et al. (2018). The charge difference across TMD1 is strong (+5, Figure 2). A ß-lactamase (ßla) fusion to TMD1 is inserted Nexo (Harley and Tipper, 1996). P is a doubly N-glycosylated fragment of the K1 killer preprotoxin, cleaved in the Golgi by the Kex2 protease (Bostian et al., 1980). The fusion product is SPInv (Tipper and Harley, 2002; Figure 2). Cexo insertion should result in invertase secretion. Insertion topology was determined by the secreted invertase activity. SPInv expression resulted in just 3% of the secreted invertase activity of the *SUC2* parent strain (Tipper and Harley, 2002), which is insufficient for growth on sucrose as this requires sufficient invertase production to form cell wall-bound aggregates. In contrast, a SPInv fusion with a −4 charge difference, predicted to insert Cexo, expressed 95% of the parental strain invertase activity, so it was inserted 95% Cexo and is efficiently cleaved in the Golgi (Tipper and Harley, 2002).

Topology was independently determined by the expression of C-terminal ßla fusion to SP (SPßla, Figure 2). SPßla topology is measured by the ratio of 52 kDa (Nexo) and 55 kDa (Cexo) forms (Figure 2); after pulse-labelling and immunoprecipitation. In normal cells, SPßla was inserted with 3% Cexo, as for SPInv (Tipper and Harley, 2002). *SUC2Δ* cells expressing SPInv were selected for growth on

sucrose after severe ethyl methanesulfonate-induced or transposon insertion mutagenesis. Mutants allowing strong growth were detected at very low frequencies in only two genes, *SPF1* and *STE24*. Deletion mutants had the same phenotype as the strongest selected mutants; *SPF1Δ* and *STE24Δ* strains produced 22% and 24%, respectively, of the *SUC2* parent strain invertase activity. These data were corroborated by pulse-labelled analysis of SPßla insertion (Figure 2); in the *SPF1Δ* and *STE24Δ* strains, insertion was 22% and 25% Cexo, respectively (Tipper and Harley, 2002). Since complete loss of PIR topology control should result in a 50/50 Cexo/Nexo ratio, about half of this control is lost in either deletion mutant. A double mutant strain was highly stress-sensitive, was both cold- and heat-sensitive, and had 38% of parent strain invertase activity, indicating a 38/50 or about 76% loss of topology control; substrate recognition by Spf1 and Ste24 appears to overlap about 50%, but their other substrates are not shared, so effects of deletion are partially additive (Tipper and Harley, 2002).

Nexo SPInv insertion occurs at the EMC, and Cexo insertion occurs at Sec61

Nexo forms of SPInv and SPßla fusions have 50-residue N-terminal peptides flanking the TMD (Figure 2) and, thus, can translocate at the EMC, as described for the TMD1 of other GPCRs (Chitwood et al., 2018). If insertion were Cexo, however, the bulky C-terminus (Figure 2) would preclude insertion into the EMC3 Oxa1 vestibule, so insertion would necessarily occur at Sec61. The residual bias of approximately 24% toward correct topology in the double mutant would then reflect weak selection at the Sec61 translocon (Goder et al., 2004). In our hypothetical PIR quality control model, Spf1 and Ste24, by eliminating most of the positive external Cexo inserts, are the major effectors of the PIR for single-pass TM proteins, and also for multipass TM proteins if topology of TMD1 is effectively monitored. Spf1 and Ste24 are conserved in all eukaryotes and so must have evolved as this cell type evolved to minimize OMM protein insertion at the ER (an exclusive function of Spf1) and dysfunction caused by highly error-prone TM protein insertion at the ER (a function of both Spf1 and Ste24). Spf1, by dislocating Cexo inserts (McKenna et al., 2020; McKenna et al., 2022), allows reinsertion, a frugal, energy-conserving mechanism; Ste24, by contrast, presumably imposes a major strain on cell resources by fragmenting incorrectly inserted TM proteins for proteasomal disassembly (Figure 1).

Role of Spf1 in eliminating unfolded proteins from the ER

Spf1 deletion in yeast causes synthetic lethality with the deletion of Hac1 (Cronin et al., 2002), the transcription factor that induces the UPR (Ron and Walter, 2007). The comprehensive yeast deletion library (Giaver et al., 2002) was screened for deletions causing induction of the UPR (Jonikas et al., 2009); the range was 2–18 fold. The strongest were *SCJ1*, a DnaJ chaperone homolog that interacts with Kar2/Bip (Jung et al., 2013) for post-translational import to the ER lumen, and *ARV1Δ*, essential for glycophosphatidylinositol (GPI) anchor synthesis. Strong

induction of the UPR is predictable for both deletions. The next strongest inducers were *SPF1Δ* and *STE24Δ*, both of which caused 12–13 fold induction, higher than BST1 (adenyl cyclase 2) orLAS21 (also equired for GPI anchor synthesis), and a full 50% higher than any of the remaining deletions, such as ERAD, EMC and GET component deletions, none of which caused more than 8-fold induction (Jonikas et al., 2009). The similarly massive accumulation of unfolded proteins in the ER lumen caused by either *SPF1Δ* or *STE24Δ* is consistent with the existence of many substrates for their quality control activities, estimated to be one third of the 33% average proteome turnover (Schubert et al., 2000), and with the massive induction of Hac1 expression caused by their deletion (Jonikas et al., 2009).

Spf1 and OMM protein insertion

Spf1 is a P5A-ATPase, a transmembrane helix dislocase that is particularly important for expelling nuclear-encoded tail-anchored OMM proteins erroneously inserted at the ER (McKenna et al., 2020) for translocation to the OMM, which is their proper location. These OMM proteins are inserted post-translationally as they have a single, relatively short, C-terminus-proximal TMD of relatively low hydrophobicity; the following short C-termini contain several positive charges (Chio et al., 2017). These proteins bind abundant chaperones such as HSP70 family members and calmodulin after exiting the ribosome. There are no known chaperones for specific targeting to the OMM, the default location for insertion (Hegde and Kenan, 2022). Their C-terminus inserts Cexo into the mitochondrial inter-membrane space (Krumpe et al., 2012). However, nearly half insert Cexo at the ER (McKenna et al., 2020); this brings their C-terminal positive charges into the ER lumen, contrary to the PIR. Import of OMP25, a yeast OMM protein containing a UV-activated cross-linker tag in its TMD, into yeast membrane vesicles identified several interacting OMM proteins but only a single and prominent ER interactor, Spf1 (McKenna et al., 2020). Spf1 is the only P5A-ATPase in yeast, and its human homolog, Atp13A1, is the sole P5A-ATPase in human cells. P-type ATPase families were reviewed by Andersen et al. (2016). Closely related homologs of Spf1 exist in all eukaryotes. In human rough-ER microsomes, Atp13A1 dislocated OMP25, but only when Atp13A1 retained its ATPase activity and an ATP source was present, demonstrating that the ATPase-dependent conformational shift of Atp13A1 is essential for the dislocation of the mislocalized OMM protein to the cytosol (McKenna et al., 2020). The mitochondrial OMM protein Msp1 performs an analogous function by dislocating mislocalized ER proteins from the OMM for ER insertion (Matsumoto et al., 2019).

As shown in Figure 3, seven of Spf1's 10 TM helices enclose an unusually large, membrane-spanning, water-filled cavity. Conserved kinks in TM helices keep fenestrations open to the lipid bilayer in either conformation, potentially allowing TM protein substrate entry. The conformation with a binding pocket open to the lumen (Figure 3) allows positively charged luminal Nexo peptides to lead TM protein substrates into Spf1's cavity, presumably guided by interaction with negative charges on the cavity's inner surface (McKenna et al., 2020; McKenna et al., 2022). Internalized substrates presumably rearrange rapidly to an energetically minimal state with a hydrophobic core exposing hydrophilic residues for interaction with the
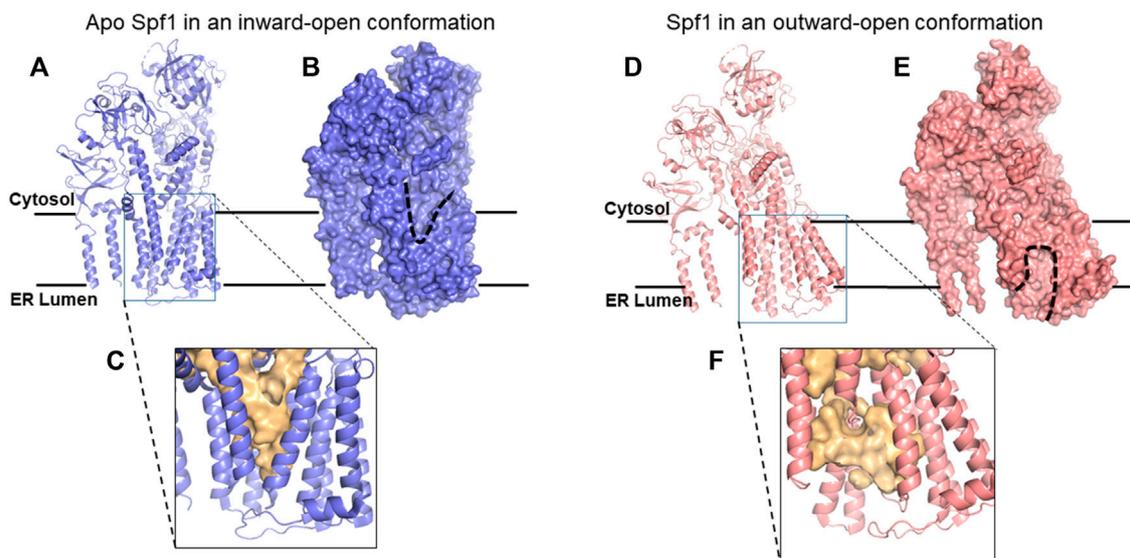
FIGURE 3
Cryo-EM structures of Spf1. **(A)** Ribbon representation of apo Spf1 in a conformation open to the cytosol (inward). **(B)** Surface representation with the "V"-shaped substrate-binding pocket outlined (dashed line); bound substrates can exit the cytosol. **(C)** Enlarged view of the substrate-binding pocket (light gold surface) with a closed luminal gate. **(D)** Ribbon representation of BeF3-bound Spf1 in a conformation open to the lumen (outward). **(E)** Surface representation with the substrate-binding pocket outlined (dashed line). **(F)** Enlarged view of the substrate-binding pocket (light gold surface) with an open luminal gate. This cavity should preferentially bind ER-TM proteins with positively charged luminal peptides. Reproduced with permission from McKenna et al. (2020).
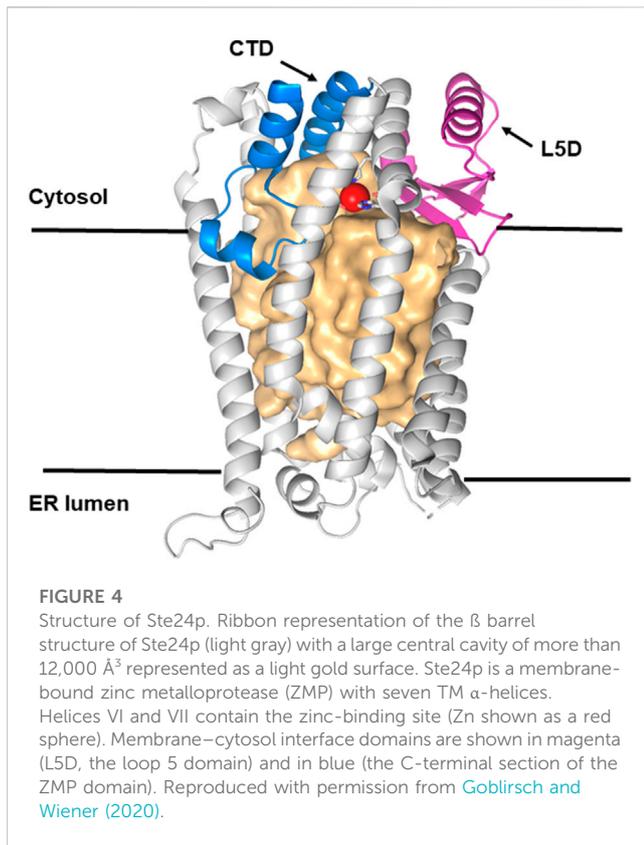
cavity's inner surface. ATP hydrolysis results in substrate exit to the cytosol (Figure 3), causing substrate dislocation. In the conformation that is open to the lumen (Figure 3), the cavity contains an additional helical component (McKenna et al., 2020), which apparently represents an averaged image of an internalized single-pass TM protein substrate; this is not visible in Figure 3, as the relevant data were described but not published (McKenna et al., 2020; McKenna et al., 2022). Type P4 ATPases, the closest relatives of type 5, transport lipids (Palmgren and Nissen, 2011), the first report of P-type ATPases transporting substrates other than ions. P4 ATPases were later called lipid flippases (Andersen et al., 2016). The human P5B-ATPase, Atp13A2, is a lysosomal polyamine transporter (van Veen et al., 2020), a secondary macromolecule transporter. Spf1/ATP13A1 is a third example, demonstrated to dislocate single-pass TM proteins, a novel, and important observation (McKenna et al., 2020; McKenna et al., 2022). An Spf1/ATP13A1-dislocated OMM protein presumably has repeated chances for insertion in its appropriate OMM location or to re-enter the ER. In the absence of Spf1/Atp13A1, OMM proteins inserted at the ER are hydrolyzed by signal peptide protease (SPP), a member of the ancient family of aspartyl proteases with active sites buried in the membrane; these include the presenilins, the active components of γ−secretase (Weihofen et al., 2002; Yucel and Lemberg, 2020). The fragments they produce are delivered to the proteasome by ERAD after ubiquitination (Dederer et al., 2020; McKenna et al., 2022). Ste24 does not recognize these misplaced OMM proteins, allowing Spf1 to maximize its energy-conserving relocation of OMM proteins to the OMM. SPP also cleaves other tail-anchored TM proteins in the ER, such as HO-1 (heme oxygenase), a component of squalene synthase, and a few other tail-anchored TM protein substrates (Boname et al., 2014; Yucel and Lemberg, 2020; McKenna et al., 2022). We propose that Ste24 recognizes and

fragments many of the single-pass TM proteins erroneously inserted at the ER in positive outside topology, including about half of those recognized by Spf1 and about half of those not recognized by Spf1 (Tipper and Harley, 2002). We propose that the PIR quality control system complements the multiple mechanisms for eliminating misfolded ER TM proteins (Chen et al., 2023). The major roles played by Spf1 and Ste24 are illustrated by their dominant roles in eliminating misfolded ER proteins (Jonikas et al., 2009).

In *SPF1Δ* cells, mitochondrial OMM proteins were not only massively mislocalized to the ER, but they were also markedly depleted from the mitochondria, illustrating the major role of Spf1 in maintaining normal ER-OMM OMM protein distribution (Krumpe et al., 2012). Sterol content is much lower in the OMM than in the ER, and OMM TA proteins have a strong preference for *in vitro* insertion into membranes of low sterol content (Krumpe et al., 2012). In *SPF1Δ* cells, the ergosterol content of the OMM and the ER is similar (Krumpe et al., 2012), suggesting that OMM protein content controls ergosterol content, which is more probable than control of protein distribution by ergosterol content ((Krumpe et al., 2012). Ste24 does not contribute to the control of OMM protein localization, which is entirely dependent on Spf1 or SPP.

Ste24 and ZMPSte24: transmembrane zinc metalloproteases

Yeast *STE24* encodes a TM zinc metalloprotease (ZMP), required for processing the precursor of the yeast's mating pheromone (Tam et al., 1998), and since this precursor contains a C-terminal isoprenylated CAAX box, Ste24 was presumed to

FIGURE 4
Structure of Ste24p. Ribbon representation of the ß barrel structure of Ste24p (light gray) with a large central cavity of more than 12,000 Å³ represented as a light gold surface. Ste24p is a membrane-bound zinc metalloprotease (ZMP) with seven TM α-helices. Helices VI and VII contain the zinc-binding site (Zn shown as a red sphere). Membrane−cytosol interface domains are shown in magenta (L5D, the loop 5 domain) and in blue (the C-terminal section of the ZMP domain). Reproduced with permission from Goblirsch and Wiener (2020).

cleave these C-termini. However, this was incorrect; Ste24 cuts the a-factor precursor at other sites (Goblirsch and Wiener, 2020). The yeast CAAX box peptidases have no direct role in PIR responses (Tipper and Harley, 2002). The human homolog, ZMPSte24, can restore Ste24 function in *STE24Δ* yeast cells, demonstrating functional conservation (Ast et al., 2016). Ste24 structure was determined by Pryor et al. (2013). Ste24 evolved from prokaryotic TM ZMP ancestors such as HtpX of *E. coli* (Arolas et al., 2014). Comparison of the structures of the yeast *Saccharomyces mikatae* Ste24 and ZMPSte24 showed strict conservation over this huge evolutionary gap, with an RMSD (relative mean squared deviation) for amino acid α carbon positions of only 1.7 Å (Goblirsch et al., 2019). Sequences among a group of 58 Ste24 orthologs covering widely divergent species contain 38 invariant residues and 30% similarity, so not only are they highly conserved, but they all share the same structure (Goblirsch et al., 2019). Ste24 has multiple specific substrates, such as prelamin A in animals, which is processed into lamin A at the inner nuclear membrane. Lamin A then assembles into the nuclear lamina, which provides mechanical stability to the nuclear envelope, maintaining nuclear architecture and controlling gene expression dependent on that architecture (Smith et al., 2021).

Ste24 structure

The structure of ZMPSte24 (Pryor et al., 2013) can be divided into three domains (Goblirsch and Wiener, 2020; Goblirsch et al., 2019; Figure 4): its ZMP protease structure lies at the cytosolic

interface and is similar to that of soluble ZMP proteases such as thermolysins; it comprises 79 residues, with 18 (23%) completely conserved among the 58 Ste24 orthologs; the adjacent L5D domain also lies at the cytosolic interface and contributes a β-strand to the ZMP active site. It consists of 135 residues, with 13 (10%) completely conserved (Goblirsch et al., 2020) (Figure 4). The unique α-barrel TM domain contains 247 residues; although only seven of these are completely conserved, these include the aromatic residues that mark the termini of specific TM helices and three proline residues that introduce kinks in these helices (Goblirsch et al., 2020). This implies that conservation of the precise α-barrel shape, with its several fenestrations, is essential for function, potentially allowing access to TM-protein substrates from the surrounding membrane, as for Spf1. The internal, roughly cylindrical cavity of Ste24 is very large (12,000 (23^3) Å³), water-filled, and lined with relatively hydrophilic residues, many of which are negatively charged. Some of these are located close to the L5D domain (Figure 4) (Goblirsch and Wiener, 2020) and they include the fully conserved aspartate 164. Determining the effect of a D164T or a potentially more disruptive D164K mutation on Ste24 function could prove informative; substrate selection might be significantly altered.

Partially reduced function in ZMPSte24 causes lipodystrophies (maldistribution of lipid tissue; Guénantin et al., 2014). Lipodystrophy is potentially a consequence of the treatment of HIV patients with inhibitors of the HIV aspartyl protease since ZMPSte24 may be sensitive to these inhibitors (Goblirsch and Wiener, 2020), but this is not established. Mutations that cause severe reductions in ZMPSte24 activity cause laminopathies, such as progerias, as do lamin A mutations (Ahn et al., 2022). Progeria severity in ZMPSTE24 mutants correlates inversely with the residual Ste24 proteolytic activity (Goblirsch and Wiener, 2020).

Roles of Ste24 in relieving clogged insertases and in type 2 diabetes

Proteins that fail to bind the SRP because they are too short, their TMD is too far beyond the N-terminus, or their signal peptides have relatively low hydrophobicity and are secreted post-translationally. Secretion requires the Kar2 (yeast) or Bip (metazoan) ER luminal chaperone to bind these substrates and, by its DNAj-stimulated ATPase action, ratchet them into the ER lumen using a distinct Sec62 translocon complex (Johnson et al., 2013; Pobre and Hendershot, 2019; Runnebohm et al., 2020). Substrate contact with Kar2 causes attached chaperones to dissociate from C-terminal domains; these are then prone to fold before insertion, clogging the insertase. Ste24 clears these clogs (Ast et al., 2016), functionally overlapping with Hrd1 ubiquitin ligase and ERAD for clearance (Runnebohm et al., 2020). Clog cleavage by Ste24 requires substrate withdrawal, apparently using the Dfm1-recruited AAA ATPase Cdc48 (Dederer et al., 2020), since a Dfm1 deletion prevents unclogging (Ast et al., 2016). Fragments of clogged substrates or other Ste24 substrates would be degraded by the proteasome. Dfm1 functions were studied by Sato and Hampton (2006). Effects of a *DFM1Δ* mutation or expression in a CDC48[ts] strain (Ast et al., 2016) would test the potential role of this withdrawal mechanism in Ste24 function. The hexameric Cdc48 structure was described by Bodnar et al. (2018).

Type 2 diabetes causes pancreatic β-cell failure and is a burgeoning worldwide scourge. In patients with the disease, islet amyloid polypeptide (IAPP), a 37-amino acid hormone that is co-secreted with insulin, misfolds and clogs the translocon secretion channel, contributing to β-cell failure and causing the formation of amyloid deposits in pancreatic islets (Clark et al., 1988). Mouse IAPP does not form amyloid, but mice overexpressing human IAPP develop spontaneous type 2 diabetes (Kayatekin et al., 2018). Expression of a human IAPP hexamer (6xIAPP), designed to accelerate amyloid formation, accelerated type 2 diabetes development in mice. In yeast cells, inducible expression of 6xIAPP caused a strong growth defect, and the 6xIAPP oligomer accumulated in the secretory pathway and the cytoplasm, consistent with clog formation, providing a model for IAPP toxicity (Kayatekin et al., 2018). Overexpression of Ste24 suppressed 6xIAPP toxicity, whereas deletion of Ste24 exacerbated it; expression of ZMPSte24 suppressed 6xIAPP toxicity in a *STE24Δ* mutant, illustrating a conserved protease function in clog destruction; expression of Ste24 protected rat insulinoma cells against toxicity mediated by 6xIAPP expression (Ast et al., 2016; Kayatekin et al., 2018). Deletion of Ste24 or Dfm1 enhanced 6xIAPP toxicity in yeast (Kayatekin et al., 2018), consistent with essential roles for Dfm1 and Cdc48 in IAPP clog extraction by Ste24, although direct interaction was not demonstrated (Kayatekin et al., 2018; Dederer et al., 2021). The mechanism of access for the withdrawn clog to the Ste24 ZMP active site is unknown.

Ste24 and control of secretion from the cytosol

Cpy, a yeast glycopeptide protease, is secreted to the vacuole post-translationally. After insertion at Sec61 and glycosylation in the Golgi, pro-Cpy normally transits to the vacuole, where cleavage by Pep4 completes maturation. If the *Cpy* signal peptide is completely deleted, about 65% accumulates in un-glycosylated form in the cytoplasm, although 35% still reaches the vacuole in fully mature form (Blachly-Dyson and Stevens, 1987; Hosomi et al., 2020). This increases to 95% in *STE24Δ* cells. Pep4ΔN23, also lacking its signal peptide, behaves similarly; 23% reaches the vacuole in WT cells and >90% in *STE24Δ* cells. Ste24, thus, recognizes and eliminates the aberrant secretion of Cpy and Pep4 from the cytosol. Ste24 recruitment would presumably require recognition of such aberrant secretion substrates and their withdrawal from Sec61, potentially involving Dfm1 and Cdc48, as for insertase clog clearance. Preventing these non-functional secretion events may be a significant part of Ste24's quality control functions.

Roles of Spf1 and Ste24 in eliminating TM proteins inserted in the ER contrary to the PIR

While Spf1 and Ste24 have very different effector mechanisms, both eliminate TM proteins inserted in an inverted topology contrary to the PIR with similar efficiency (Tipper and Harley, 2002). Both have large, membrane-spanning, water-filled cavities and while Ste24's cavity is considerably larger

than that of Spf1, they have basic structural similarities specifically related to the demonstrated substrate recognition and entry functions of Spf1 (McKenna et al., 2020). Both cavities are enclosed by seven TM helices with conserved kinks that provide stable fenestrations, potentially allowing TM protein entry and both internal surfaces have distinctly negative electrostatic potentials (Figures 3, 4), consistent with similar mechanisms for substrate recognition. OMM proteins are tail-anchored TM proteins recognized in the ER only by Spf1, while about half of other Spf1 and Ste24 substrates overlap (Tipper and Harley, 2002), so substrate recognition mechanisms must also partially overlap. Mechanisms following recognition are unrelated; for Spf1, ATP hydrolysis results in a series of conformational shifts between the E2 state, open to the lumen, and the E1 state, open to the cytosol (Figure 3) (McKenna et al., 2020). In the E2 conformation, electronegative potential in the cavity presumably promotes entry of the short, positively charged luminal substrate peptides, guiding lateral entry of the entire TM proteins via fenestrations. Ste24 ZMP activity is required for the PIR quality control function (Tipper and Harley, 2002); internalized substrates presumably rearrange to an energetically minimal state, as for Spf1. This might allow diffusion to Ste24's cytosolic face and interaction with the L5D and ZMP modules; however, efficient hydrolysis may require an external motive force. We hypothesize that this may be provided by Dfm1-recruited Cdc48 AAA ATPase (Dederer et al., 2021) and insertase clog removal. Substrate fragments would be pulled into the cytosol for delivery to the proteasome (Figure 1). However, no direct evidence demonstrates substrate entry into the Ste24 cavity, subsequent internal access to the ZMP site, cdc48 withdrawal, or proteasomal destruction.

The PIR quality control system

The proposed PIR quality control system, consisting of the effector pairs Spf1/Ste24 (yeast) and *ATP13A1/ZMPSte24* (human), presumably evolved along with the eukaryotic cell to address the cellular stress caused by mitochondrial OMM proteins inserted into the ER (Spf1-specific substrates) and the massive accumulation of inverted TM proteins in the ER resulting from high rates of aberrant ER insertion. The PIR quality control system supplements ERAD, but with specific recognition and substrate criteria; substrates are recognized by their defining positive outside ER luminal peptide inserts; the resulting correction imposes the PIR. The positive outside topology of PIR quality control substrates is required for entry into the Spf1's cavity (McKenna et al., 2020, McKenna et al., 2022) and presumably for the proposed entry into the larger Ste24's cavity. The sharing of some substrates with ERAD (Runnebohm et al., 2020) is an example of normal redundancy in cellular functions. This PIR quality control system is highly effective but never 100% effective, as illustrated by the 97% Nexo insertion of the model SPInv and SPβla TM proteins with −4 PIR signals (Tipper and Harley, 2002). A weaker PIR response is apparent in a SPInv fusion with a +2 PIR signal, entirely due to the two TMD-adjacent N-terminal negative charges in S (Figure 2); insertion was 8% Nexo in WT cells, increasing to about 30% in *SPF1Δ* and *STE24Δ* cells and 43% in

the double-mutant strain, so this higher rate of erroneous insertion is still effectively corrected by the PIR quality control system.

ZMPSte24 in antiviral responses

ZMPSte24 plays an important role, apparently as a major effector in a broad-spectrum constitutive defense against fusion between the membranes of host cells and the membranes of viral pathogens such as influenza, coronaviruses, HIV, Ebola, and Zika (Brass et al., 2009; Narayana et al., 2015; Fu et al., 2017; Majdoul et al., 2022). Ste24 is also effective against SARS-Cov-2 infection (Shilagardi et al., 2022). Screening for genes affecting γ-interferon-induced responses to influenza revealed a role for the inducible transmembrane protein (IFITM) family (Brass et al., 2009). ZMPSte24 interacts with IFITM3 (Fu et al., 2017), the best-studied and most active family member, which may chaperone ZMPSte24 to its viral fusion targets. Antiviral activity did not require ZMPSte24 protease activity but may result from the recognition of virus–host cell membrane hemifusions (Hosomi et al., 2020). Disruption of trafficking is suggested to cause the transit of virus-infected endosomes to the vacuole for destruction, but the mechanisms remain conjectural (Majdoul et al., 2022).

Identification of "constitutive" PIR quality control substrates by affinity labeling

Recent progress toward identifying Atp13A1 and ZMPSte24 substrates comes from the use of proximity labeling techniques in which a protein bait, coupled to a promiscuous form of biotin ligase, is used to tag transient protein interactants within a range of ~10 nm (average protein size) for their subsequent identification (Go et al., 2021). A total of 192 proteins or peptides, representing all internal cellular compartments of human HEK293 cells, were converted into bait. These were then expressed in HEK293 cells. The proteins identified as interacting with Atp13A1 are presumably frequently contacted and are essentially constitutive cellular targets for dislocation; these included a particular signal peptide fragment, several ER TM proteins, TM proteins in an ER/Golgi intermediate compartment, subunits of the ER Ca^{2+} channel, components of a lipid droplet export compartment, and Sec62 and Sec61B (Go et al., 2021). All, except the signal peptide, are potential orphan (excess) components of major ER TM protein complexes, which are plausible Atp13A1 substrates. Other interactants were components of endosomes, the plasma membrane, and nuclear pores, perhaps also orphans or products of trafficking errors (Go et al., 2021). None of the identified interactants had a plausible role in substrate recognition (Go et al., 2021).

For ZMPSte24, the conclusions were similar (Go et al., 2021). Approximately 40% of the interactants were shared with Atp13A1, consistent with the quantitative results of Tipper and Harley (2002), showing substantial but incomplete substrate overlap in the PIR quality control. The remaining 60% are likely to represent orphan components

of TM protein complexes, or proteins that failed to transit to their proper destinations; these included ER luminal proteins, mitochondrial matrix proteins, and an OMM protein that is not an identified Spf1 substrate (Go et al., 2021). As is probably true for dislocation by Atp13A1, ZMPSte24 hydrolysis of random products of aberrant TM protein insertion or post-translational secretion probably far exceeds hydrolysis of the constitutive substrates identified by proximity labeling (Go et al., 2021). None of the identified Ste24 interactors had a plausible role in substrate recognition.

The eukaryotic PIR quality control system: a hangover from prokaryotic ancestors?

Basic similarities in the structures of Spf1 and Ste24 are the only clue to a common mode of substrate selection, independent of the obvious differences in substrate removal mechanisms. The proposed shared mechanisms for substrate selection by recognition of TMD-adjacent positively charged luminal peptides may explain how this quality control system imposes approximately 75% of the PIR after error-prone insertion. The role of membrane lipid charge on PIR responses needs to be taken into account (Dowhan et al., 2019). The role of Ste24 in the PIR quality control mechanism requires considerable further investigation.

The question remains as to why the PIR is still relevant in a eukaryotic cell. Plausible prokaryotic progenitors of eukaryotic cells, such as the archaeal Asgard family (Gabaldon, 2021; Liu et al., 2021), still had a positive outside prokaryotic plasma membrane potential as the internal membrane complexity and trafficking machinery evolved. This increasing complexity probably resulted in an increasing frequency of TM protein insertion errors, driving the selection of quality control effectors capable of recognizing and eliminating nonfunctional TM proteins inserted "positive outside." Such erroneous insertions were defined over many millennia in their prokaryotic ancestors by the PIR, the determinant of topological "right and wrong." The acquisition of mitochondrial precursors, presumably from other cells, would have provided the energy supply necessary for this increasingly expensive cellular lifestyle, but after the OMM protein gene migrated to the nucleus, OMM proteins would have been included as a major source of additional PIR quality control substrates. It is possible that Spf1 evolved to address this specific issue, while Ste24 evolved much earlier to meet more general housekeeping problems. Thus, the persistence of dependence on the PIR for quality control of TM protein topology in eukaryotic descendants, as ER and cytoplasmic membrane functions diverged, could therefore be considered a hangover or relict from their ancient prokaryotic ancestors. A search we made of the Asgard family sequences for critical sequences of Spf1 (ATPase domain) or Ste24 (ZMP domain) did not prove informative.

ER-phagy is a system that removes ERAD-resistant misfolded proteins and aggregates from the ER for lysosomal destruction, apparently working in concert with ERAD and the UPR to control TM protein quality in the ER (Chen et al., 2023). The PIR quality control system is another important component essential for maintaining ER TM protein quality. Defects in the UPR, ERAD, and ER-phagy have been shown to play major roles in cancer and

neurodegenerative diseases (Chen et al., 2023). Defects in PIR quality control may cause similar pathology.

Conclusion

We propose that Spf1/Atp13A1 and Ste24/ZMPSte24 share a common mechanism for PIR quality control substrate recognition, although this has only been demonstrated for Spf1/Atp13A1 (McKenna et al., 2020) and many aspects of Ste24 function remain unresolved. This hypothesis is based on basic similarities in structures that are essential at least for Spf1/Atp13A1 function (McKenna et al., 2020, McKenna et al., 2022). Direct analysis of the mechanism of Ste24 is a top priority to test the PIR quality control hypothesis and evaluate its importance. Testing the potential role of the conserved Ste24 aspartate 164 in substrate cleavage should be straightforward. Analysis of the potential roles of Dfm1 and cdc48 in substrate removal from Ste24 may substantially clarify Ste24's mechanism. Dfm1 is a candidate for the unidentified mutant with about half of Spf1's effect isolated by Tipper and Harley (2002). To test whether Ste24 substrates do pass through the Ste24 cavity to reach its ZMP domain, a strategically located biotin ligase fusion (Fenech et al., 2023), perhaps linked to a cytosol-accessible L5D residue (Figure 4), may identify transiently adjacent effectors or substrate fragments. A Ste24 substrate with a suitably positioned UV-activated cross-linker (McKenna et al., 2020) may potentially identify both STE24-interacting effectors and substrates, perhaps identifying features that distinguish them from Spf1 substrates.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahn, J., Jeong, S., Kang, S. M., Jo, I., Park, B. J., and Ha, N. C. (2022). Crystal structure of progeria mutant S143F lamin A/C reveals increased hydrophobicity driving nuclear deformation. *Commun. Biol.* 5, 267. doi:10.1038/s42003-022-03212-3

Akopian, D., Shen, K., Zhang, X., and Shan, S. o. (2013). Signal recognition particle: an essential protein-targeting machine. *Annu. Rev. Biochem.* 82, 693–721. doi:10.1146/annurev-biochem-072711-164732

Andersen, J. P., Vestergaard, A. L., Mikkelsen, S. A., Mogensen, L. S., Chalat, M., and Molday, R. S. (2016). P4-ATPases as phospholipid flippases-structure, function, and enigmas. *Front. Physiol.* 7, 275. doi:10.3389/fphys.2016.00275

Arolas, J., García-Castellanos, R., Goulas, T., Akiyama, Y., and Gomis-Rüth, F. X. (2014). Expression and purification of integral membrane metallopeptidase HtpX. *Protein Expr. Purif.* 99, 113–118. doi:10.1016/j.pep.2014.04.008

Ast, T., Michaelis, S., and Schuldiner, M. (2016). The protease Ste24 clears clogged translocons. *Cell* 164, 103–114. doi:10.1016/j.cell.2015.11.053

Aviram, N., and Schuldiner, M. (2017). Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J. Cell Sci.* 130, 4079–4085. doi:10.1242/jcs.204396

Baker, J. A., Wong, W. C., Eisenhaber, B., Warwicker, J., and Eisenhaber, F. (2017). Charged residues next to transmembrane regions revisited: "Positive-inside rule" is complemented by the "negative inside depletion/outside enrichment rule. *BMC Biol.* 15, 66. doi:10.1186/s12915-017-0404-4

Baum, D. A., and Baum, B. (2014). An inside-out origin for the eukaryotic cell. *BMC Biol.* 12, 76. doi:10.1186/s12915-014-0076-2

Blachly-Dyson, E., and Stevens, T. H. (1987). Yeast carboxypeptidase Y can be translocated and glycosylated without its amino-terminal signal sequence. *J. Cell Biol.* 104, 1183–1191. doi:10.1083/jcb.104.5.1183

Bodnar, N. O., Kim, K. H., Ji, Z., Wales, T. E., Svetlov, V., Nudler, E., et al. (2018). Structure of the Cdc48 ATPase with its ubiquitin-binding cofactor ufd1-npl4. *Nat. Struct. Mol. Biol.* 25, 616–622. doi:10.1038/s41594-018-0085-x

Boname, J. M., Bloor, S., Wandel, M. P., Nathan, J. A., Antrobus, R., Dingwell, K. S., et al. (2014). Cleavage by signal peptide peptidase is required for the degradation of selected tail-anchored proteins. *J. Cell Biol.* 205, 847–862. doi:10.1083/jcb.201312009

Bostian, K., (1980). Translation analysis of the killer-associated virus-like particle dsRNA genome of *Saccharomyces cerevisiae*: M-dsRNA encodes toxin. *Cell* 19, 404–414.

Brass, A. L., Huang, I. C., Benita, Y., John, S. P., Krishnan, M. N., Feeley, E. M., et al. (2009). The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell* 139, 1243–1254. doi:10.1016/j.cell.2009.12.017

Chen, G., Wei, T., Ju, F., and Li, H. (2023). Protein quality control and aggregation in the endoplasmic reticulum: from basic to bedside. *Front. Cell Dev. Biol.* 11, 1156152. doi:10.3389/fcell.2023.1156152

Chio, U. S., Cho, H., and Shan, S. O. (2017). Mechanisms of tail-anchored membrane protein targeting and insertion. *Annu. Rev. Cell Dev. Biol.* 33, 417–438. doi:10.1146/annurev-cellbio-100616-060839

Chitwood, P. J., and Hegde, R. S. (2019). The role of EMC during membrane protein biogenesis. *Trends Cell Biol.* 29, 371–384. doi:10.1016/j.tcb.2019.01.007

Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S., and Hegde, R. S. (2018). EMC is required to initiate accurate membrane protein topogenesis. *Cell* 175, 1507–1519. doi:10.1016/j.cell.2018.10.009

Clark, A., Wells, C. A., Buley, I. D., Cruickshank, J. K., Vanhegan, R. I., Matthews, D. R., et al. (1988). Islet amyloid, increased A-cells, reduced B-cells and exocrine fibrosis: quantitative changes in the pancreas in type 2 diabetes. *Diabetes Res.* 9, 151–159.

Cronin, S. R., Rao, R., and Hampton, R. Y. (2002). Cod1p/Spf1p is a P-type ATPase involved in ER function and Ca2+ homeostasis. *J. Cell Biol.* 157, 1017–1028. doi:10.1083/jcb.200203052

Dahan, N., Bykov, Y. S., Boydston, E. A., Fadel, A., Gazi, Z., Hochberg-Laufer, H., et al. (2022). Peroxisome function relies on organelle-associated mRNA translation. *Sci. Adv.* 8, eabk2141. doi:10.1126/sciadv.abk2141

Dederer, V., Khmelinskii, A., Huhn, A. G., Okreglak, V., Knop, M., and Lemberg, M. K. (2019). Cooperation of mitochondrial and ER factors in quality control of tail-anchored proteins. *Elife* 8, e45506. doi:10.7554/eLife.45506

Dederer, V., and Lemberg, M. K. (2021). Transmembrane dislocases: a second chance for protein targeting. *Trends Cell Biol.* 31, 898–911. doi:10.1016/j.tcb.2021.05.007

Dowhan, W., Vitrac, H., and Bogdanov, M. (2019). Lipid-assisted membrane protein folding and topogenesis. *Protein J.* 38 (3), 274–288. doi:10.1007/s10930-019-09826-7

Elazar, A., Weinstein, J. J., Prilusky, J., and Fleishman, S. J. (2016). Interplay between hydrophobicity and the positive-inside rule in determining membrane-protein topology. *Proc. Natl. Acad. Sci. U S A.* 113, 10340–10345. doi:10.1073/pnas.160588811

Fenech, E. J., Cohen, N., Kupervaser, M., Gazi, Z., and Schuldiner, M. (2023). A toolbox for systematic discovery of stable and transient protein interactors in baker's yeast. *Mol. Syst. Biol.* 19, e11084. doi:10.15252/msb.202211084

Fu, B., Wang, L., Li, S., and Dorf, M. E. (2017). ZMPSTE24 defends against influenza and other pathogenic viruses. *J. Exp. Med.* 214, 919–929. doi:10.1084/jem.20161270

Gabaldon, T. (2021). Origin and early evolution of the eukaryotic cell. *Annu. Rev. Microbiol.* 75, 631–647. doi:10.1146/annurev-micro-090817-062213

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391. doi:10.1038/nature00935

Go, C. D., Knight, J. D. R., Rajasekharan, A., Rathod, B., Hesketh, G. G., Abe, K. T., et al. (2021). A proximity-dependent biotinylation map of a human cell. *Nature* 595, 120–124. doi:10.1038/s41586-021-03592-2

Goblirsch, B. R., Pryor, E. E., and Wiener, M. C. (2019). The tripartite architecture of the eukaryotic integral membrane protein zinc metalloprotease Ste24. *Proteins* 88, 604–615. doi:10.1002/prot.25841

Goblirsch, B. R., and Wiener, M. C. (2020). Ste24: an integral membrane protein zinc metalloprotease with provocative structure and emergent Biology. *J. Mol. Biol.* 432, 5079–5090. doi:10.1016/j.jmb.2020.03.016

Goder, V., Bieri, C., and Spiess, M. (1999). Glycosylation can influence topogenesis of mem-brane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *J. Cell Biol.* 147 (2), 257–266. doi:10.1083/jcb.147.2.257

Goder, V., Junne, T., and Spiess, M. (2004). Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Mol. Biol. Cell* 15, 1470–1478. doi:10.1091/mbc.e03-08-0599

Guenantin, A. C., Briand, N., Bidault, G., Afonso, P., Béréziat, V., Vatier, C., et al. (2014). Nuclear envelope-related lipodystrophies. *Semin. Cell Dev. Biol.* 29, 148–157. doi:10.1016/j.semcdb.2013.12.015

Halic, M., Becker, T., Pool, M. R., Spahn, C. M. T., Grassucci, R. A., Frank, J., et al. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427, 808–814. doi:10.1038/nature02342

Harley, C. A., Holt, J. A., Turner, R., and Tipper, D. J. (1998). Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution. *J. Biol. Chem.* 273, 24963–24971. doi:10.1074/jbc.273.38.24963

Harley, C. A., and Tipper, D. J. (1996). The role of charged residues in determining transmembrane protein insertion orientation in yeast. *J. Biol. Chem.* 271, 24625–24633. doi:10.1074/jbc.271.40.24625

Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5786–5790. doi:10.1073/pnas.86.15.5786

Heaven, W. D. (2020). *DeepMind's protein-folding AI has solved a 50-year-old grand challenge of biology.* Cambridge, MA, USA: MIT Technology Review.

Hegde, R. S., and Keenan, R. J. (2022). The mechanisms of integral membrane protein biogenesis. *Nat. Rev. Mol. Cell Biol.* 23, 107–124. doi:10.1038/s41580-021-00413-2

Hegedűs, T., Geisler, M., Lukács, G. L., and Farkas, B. (2022). Ins and outs of AlphaFold2 transmembrane protein structure predictions. *Cell Mol. Life Sci.* 15, 73. doi:10.1007/s00018-021-04112-1

Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J.* 5, 3021–3027. doi:10.1002/j.1460-2075.1986.tb04601.x

Hosomi, A., Iida, K., Cho, T., Iida, H., Kaneko, M., and Suzuki, T. (2020). The ER-associated protease Ste24 prevents N-terminal signal peptide-independent translocation into the endoplasmic reticulum in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 295, 10406–10419. doi:10.1074/jbc.RA120.012575

Johnson, N., Powis, K., and High, S. (2013). Post-translational translocation into the endoplasmic reticulum. *Biochim. Biophys. Acta* 1833, 2403–2409. doi:10.1016/j.bbamcr.2012.12.008

Jones, D. T. (2007). Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics* 23, 538–544. doi:10.1093/bioinformatics/btl677

Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., et al. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323, 1693–1697. doi:10.1126/science.1167983

Jung, K. W., Kang, H. A., and Bahn, Y. S. (2013). Essential roles of the Kar2/BiP molecular chaperone downstream of the UPR pathway in Cryptococcus neoformans. *PLoS One* 8, e58956. doi:10.1371/journal.pone.0058956

Kayatekin, C., Amasino, A., Gaglia, G., Flannick, J., Bonner, J. M., Fanning, S., et al. (2018). Translocon declogger Ste24 protects against IAPP oligomer-induced proteotoxicity. *Cell* 173, 62–73 e9. doi:10.1016/j.cell.2018.02.026

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi:10.1006/jmbi.2000.4315

Krumpe, K., Frumkin, I., Herzig, Y., Rimon, N., Özbalci, C., Brügger, B., et al. (2012). Ergosterol content specifies targeting of tail-anchored proteins to mitochondrial outer membranes. *Mol. Biol. Cell* 23, 3927–3935. doi:10.1091/mbc.E11-12-0994

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

Liu, Y., Makarova, K. S., Huang, W. C., Wolf, Y. I., Nikolskaya, A. N., Zhang, X., et al. (2021). Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature* 593, 553–557. doi:10.1038/s41586-021-03494-3

Luirink, J., Samuelsson, T., and De Gier, J. W. (2001). YidC/Oxa1p/Alb3: evolutionarily conserved mediators of membrane protein assembly. *FEBS Lett.* 501, 1–5. doi:10.1016/s0014-5793(01)02616-3

Majdoul, S., and Compton, A. A. (2022). Lessons in self-defence: inhibition of virus entry by intrinsic immunity. *Nat. Rev. Immunol.* 22, 339–352. doi:10.1038/s41577-021-00626-8

Marinko, J., Huang, H., Penn, W. D., Capra, J. A., Schlebach, J. P., and Sanders, C. R. (2019). Folding and misfolding of human membrane proteins in Health and disease: from single molecules to cellular proteostasis. *Chem. Rev.* 119 (9), 5537–5606. doi:10.1021/acs.chemrev.8b00532

Matsumoto, S., Nakatsukasa, K., Kakuta, C., Tamura, Y., Esaki, M., and Endo, T. (2019). Msp1 clears mistargeted proteins by facilitating their transfer from mitochondria to the ER. *Mol. Cell* 76, 191–205. doi:10.1016/j.molcel.2019.07.006

McKENNA, M. J., Adams, B. M., Chu, V., Paulo, J. A., and Shao, S. (2022). ATP13A1 prevents ERAD of folding-competent mislocalized and misoriented proteins. *Mol. Cell* 82, 4277–4289 e10. doi:10.1016/j.molcel.2022.09.035

McKENNA, M. J., Sim, S. I., Ordureau, A., Wei, L., Harper, J. W., Shao, S., et al. (2020). The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science* 369, eabc5809. doi:10.1126/science.abc5809

Narayana, S. K., Helbig, K. J., McCartney, E. M., Eyre, N. S., Bull, R. A., Eltahla, A., et al. (2015). The interferon-induced transmembrane proteins, IFITM1, IFITM2, and IFITM3 inhibit hepatitis C virus entry. *J. Biol. Chem.* 290, 25946–25959. doi:10.1074/jbc.M115.657346

Nielsen, H., Tsirigos, K. D., Brunak, S., and von Heijne, G. (2019). A brief history of protein sorting prediction. *Protein J.* 38 (3), 200–216. doi:10.1007/s10930-019-09838-3

Osborne, A. R., Rapoport, T. A., and van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* 21, 529–550. doi:10.1146/annurev.cellbio.21.012704.133214

Palmgren, M. G., and Nissen, P. (2011). P-type ATPases. *Annu. Rev. Biophys.* 40, 243–266. doi:10.1146/annurev.biophys.093008.131331

Pleiner, T., Tomaleri, G. P., Januszyk, K., Inglis, A. J., Hazu, M., and Voorhees, R. M. (2020). Structural basis for membrane insertion by the human ER membrane protein complex. *Science* 369, 433–436. doi:10.1126/science.abb5008

Pobre, K. F. R., Poet, G. J., and Hendershot, L. M. (2019). The endoplasmic reticulum (ER) chaperone BiP is a master regulator of ER functions: getting by with a little help from ERdj friends. *J. Biol. Chem.* 294, 2098–2108. doi:10.1074/jbc.REV118.002804

PryorHoranyi, P. S., Clark, K. M., Fedoriw, N., Connelly, S. M., Koszelak-Rosenblum, M., et al. (2013). Structure of the integral membrane protein CAAX protease Ste24p. *Science* 339, 1600–1604. doi:10.1126/science.1232048

Rapoport, T. A., Li, L., and Park, E. (2017). Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* 33, 369–390. doi:10.1146/annurev-cellbio-100616-060439

Rapp, M., Granseth, E., Seppälä, S., and von Heijne, G. (2006). Identification and evolution of dual-topology membrane proteins. *Nat. Struct. Mol. Biol.* 13 (2), 112–116. doi:10.1038/nsmb1057

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. *Nat. Rev. Mol. Cell Biol.* 8, 519–529. doi:10.1038/nrm2199

Runnebohm, A. M., Richards, K. A., Irelan, C. B., Turk, S. M., Vitali, H. E., Indovina, C. J., et al. (2020). Overlapping function of Hrd1 and Ste24 in translocon quality control provides robust channel surveillance. *J. Biol. Chem.* 295 (47), 16113–16120. doi:10.1074/jbc.AC120.016191

Sato, B. K., and Hampton, R. Y. (2006). Yeast derlin Dfm1 interacts with Cdc48 and functions in ER homeostasis. *Yeast* 23, 1053–1064. doi:10.1002/yea.1407

Schlebach, J. P., and Sanders, C. R. (2015). Influence of pathogenic mutations on the energetics of translocon-mediated bilayer integration of transmembrane helices. *J. Membr. Biol.* 248 (3), 371–381. doi:10.1007/s00232-014-9726-0

Schubert, U., Antón, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404, 770–774. doi:10.1038/35008096

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T., et al. (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* 123, 507–519. doi:10.1016/j.cell.2005.08.031

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

Shao, S., and Hegde, R. S. (2011). Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* 27, 25–56. doi:10.1146/annurev-cellbio-092910-154125

Shilagardi, K., Spear, E. D., Abraham, R., Griffin, D. E., and Michaelis, S. (2022). The integral membrane protein ZMPSTE24 protects cells from SARS-CoV-2 spike-mediated pseudovirus infection and syncytia formation. *mBio* 13 (5), e0254322. doi:10.1128/mbio.02543-22

Smalinskaite, L., Kim, M. K., Lewis, A. J. O., Keenan, R. J., and Hegde, R. S. (2022). Mechanism of an intramembrane chaperone for multipass membrane proteins. *Nature* 611, 161–166. doi:10.1038/s41586-022-05336-2

Smith, C. L., Poleshko, A., and Epstein, J. A. (2021). The nuclear periphery is a scaffold for tissue-specific enhancers. *Nucleic Acids Res.* 49, 6181–6195. doi:10.1093/nar/gkab392

Tam, A., Nouvet, F. J., Fujimura-Kamada, K., Slunt, H., Sisodia, S. S., and Michaelis, S. (1998). Dual roles for Ste24p in yeast a-factor maturation: NH2-terminal proteolysis and COOH-terminal CAAX processing. *J. Cell Biol.* 142 (3), 1998 635–649. doi:10.1083/jcb.142.3.635

Tikhonova, E., Karamysheva, Z. N., von Heijne, G., and Karamyshev, A. L. (2019). Silencing of aberrant secretory protein expression by disease-associated mutations. *J. Mol. Biol.* 431 (14), 2567–2580. doi:10.1016/j.jmb.2019.05.011

Tipper, D. J., and Harley, C. A. (2002). Yeast genes controlling responses to topogenic signals in a model transmembrane protein. *Mol. Biol. Cell* 13, 1158–1174. doi:10.1091/mbc.01-10-0488

Van VeenMartin, S., Van den Haute, C., Benoy, V., Lyons, J., Vanhoutte, R., et al. (2020). ATP13A2 deficiency disrupts lysosomal polyamine export. *Nature* 578, 419–424. doi:10.1038/s41586-020-1968-7

Vitrac, H., Bogdanov, M., and Dowhan, W. (2013). *In vitro* reconstitution of lipid-dependent dual topology and postassembly topological switching of a membrane protein. *Proc. Natl. Acad. Sci. U S A.* 110, 9338–9343. doi:10.1073/pnas.1304375110

Von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14, 4683–4690. doi:10.1093/nar/14.11.4683

Von Heijne, G. (2006). Membrane-protein topology. *Nat. Rev. Mol. Cell Biol.* 7 (12), 909–918. doi:10.1038/nrm2063

Walter, P., and Blobel, G. (1981a). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* 91, 557–561. doi:10.1083/jcb.91.2.557

Walter, P., and Blobel, G. (1981b). Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of *in-vitro*-assembled polysomes synthesizing secretory protein. *J. Cell Biol.* 91, 551–556. doi:10.1083/jcb.91.2.551

Wang, F., Chan, C., Weir, N. R., and Denic, V. (2014). The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* 512, 441–444. doi:10.1038/nature13471

Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K., and Martoglio, B. (2002). Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* 296, 2215–2218. doi:10.1126/science.1070925

White, S. H., and Von Heijne, G. (2005). Do protein-lipid interactions determine the recognition of transmembrane helices at the ER translocon? *Biochem. Soc. Trans.* 33, 1012–1015. doi:10.1042/BST20051012

Wu, X., and Rapoport, T. A. (2018). Mechanistic insights into ER-associated protein degradation. *Curr. Opin. Cell Biol.* 53, 22–28. doi:10.1016/j.ceb.2018.04.004

Wu, X., Siggel, M., Ovchinnikov, S., Mi, W., Svetlov, V., Nudler, E., et al. (2020). Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. *Science* 368, eaaz2449. doi:10.1126/science.aaz2449

Wurie, H. R., Buckett, L., and Zammit, V. A. (2011). Evidence that diacylglycerol acyltransferase 1 (DGAT1) has dual membrane topology in the endoplasmic reticulum of HepG2 cells. *J. Biol. Chem.* 286 (42), 36238–36247. doi:10.1074/jbc.M111.251900

Yucel, S. S., and Lemberg, M. K. (2020). Signal peptide peptidase-type proteases: versatile regulators with functions ranging from limited proteolysis to protein degradation. *J. Mol. Biol.* 432, 5063–5078. doi:10.1016/j.jmb.2020.05.014