Signal Sequences Encode Information for Protein Folding in the Endoplasmic Reticulum

Sha Sun[1], Xia Li[1], and Malaiyalam Mariappan[1]
Department of Cell Biology, Nanobiology Institute, Yale University West Campus.

Department of Cell Biology
Nanobiology Institute
Yale School of Medicine
Yale West Campus
West Haven, CT 06516, USA

[*] Correspondence: malaiyalam.mariappan@yale.edu

Summary: Roughly one-third of newly synthesized proteins enter into the endoplasmic reticulum (ER) via the Sec61 translocon. It is unclear how nascent chains bind correct chaperones and properly fold upon entering the ER lumen. We find that signal sequences harbor information to recruit specific chaperones for protein folding in the ER. Using a substrate-trapping proteomic approach, we identify that marginally hydrophobic signal sequences are transiently clogged at the Sec61 translocon, which recruits BiP chaperone through Sec63 to bind onto nascent chains. Surprisingly, this privileged BiP binding not only releases clogged nascent chains into the ER lumen but also prevent inappropriate interactions and promotes folding and maturation. Signal sequence swapping bypasses BiP-dependent unclogging and translocation, but the translocated nascent chain is terminally misfolded after binding the wrong chaperone in the ER lumen. Thus, signal sequence-dependent chaperone recruitment explains why signal sequences are paradoxically diverse and use multiple protein translocation pathways in cells.

KEYWORDS
Endoplasmic Reticulum, Signal Sequence, Chaperone, Protein folding and misfolding

INTRODUCTION

More than 30% of newly synthesized proteins from the cytosolic ribosomes are delivered to the endoplasmic reticulum (ER) by the co-translational protein targeting pathway (Juszkiewicz and Hegde, 2018). These nascent polypeptides typically contain a hydrophobic targeting sequence that is either cleavable N-terminal signal sequence or non-cleavable first transmembrane domain (TMD). As the targeting sequences emerge from cytosolic ribosomes, they are co-translationally recognized by the signal recognition particle (SRP) (Shan and Walter, 2005). The SRP-bound ribosome-associated nascent chains are then delivered to the ER membrane via an interaction with the SRP receptor. The targeting sequence is then transferred from the SRP to the heterotrimeric Sec61 translocon complex (Rapoport, 2007). The targeting sequences are again recognized, this time, by the Sec61 translocon, allowing the translocation of the following nascent polypeptide across the ER membrane with concomitant cleavage of the signal peptide or integration of the TMD into the lipid bilayer.

The hydrophobicity of targeting sequences plays a pivotal role in recognition by both the SRP and the Sec61 translocon (Keenan et al., 1998; Li et al., 2016; Voorhees and Hegde, 2016). While SRP can efficiently recognize a broad range of hydrophobicity in targeting sequences in mammals (Hegde and Kang, 2008; Kim et al., 2002; Voorhees and Hegde, 2015), the Sec61 translocon appears to be more stringent in that it only prefers targeting sequences with sufficient hydrophobicity (Jungnickel and Rapoport, 1995; Plumb et al., 2015). Nevertheless, hydrophobicity of targeting sequences is remarkably variable, but they can be broadly classified into two categories: marginally hydrophobic (“weak”) and strong hydrophobic (“strong”) sequences (Hegde and Bernstein, 2006; Hegde and Kang, 2008; Voorhees and

Hegde, 2015). A strong targeting sequence tightly engages with the Sec61 translocon as a hairpin loop in which the N-terminus is localized in the cytosol and the following nascent chain is inserted into the aqueous pore of the Sec61 translocon (Jungnickel and Rapoport, 1995). By contrast, a weak targeting sequence loosely engages the Sec61 translocon due to its reduced hydrophobicity. Many translocon associated factors have been implicated in assisting the translocation of proteins with weak targeting sequences (Hegde and Kang, 2008). Despite their inefficiency in mediating protein translocation into the ER, weak targeting signal sequences are paradoxically conserved in mammals. One of the functions of weak targeting sequences is that protein translocation can be selectively attenuated during ER stress conditions to reduce the ER folding burden (Kang et al., 2006). In some cases, the signal sequence regulates the timing of signal peptide cleavage and thereby enhancing the folding and maturation of proteins (Snapp et al., 2017).

Similar to targeting sequences, TMDs of polytopic membrane proteins are recognized by the Sec61 translocon and laterally partition into the lipid bilayer via the later gate of the translocon. Again, TMD hydrophobicity plays a crucial role in partitioning from the lateral gate of the translocon to the hydrophobic core lipid bilayer. Yet, TMDs of membrane proteins greatly vary in their hydrophobicity. It is estimated that ~25% of TMDs of polytopic membrane proteins are marginally hydrophobic, and they are not suitable for insertion into the membrane by themselves in isolation (Elofsson and von Heijne, 2007; White and von Heijne, 2008). Recent studies have shown that weak TMDs can stall in the interior pore of the Sec61 translocon (Kida et al., 2016; Kida and Sakaguchi, 2018). In some cases, weak TMDs can be completely translocated into the ER lumen. From there, they can be either retrieved by neighboring TMDs or degraded by ER-associated quality control pathways (Feige and Hendershot, 2013; Lu et al., 2000; Onishi et al., 2013; Skach et al., 1994). It is poorly understood about the factors and mechanisms involved in clearing the translocon clogged with weak TMDs during the co-translational membrane protein insertion. This problem is likely distinct from the translocon unclogging by Ste24, which is shown to act on post-translationally clogged translocons in yeast (Ast et al., 2016). Conversely, many soluble secretory proteins contain long hydrophobic segments but translocate through the translocon (Hessa et al., 2007). Why do they not get 'stuck' as they are being translocated? One possibility is that there are proteins that constantly monitor protein translocation through the Sec61 channel and are clearing the clogged nascent chains, but the identities of these proteins are unknown.

During their translocation into the ER lumen, nascent chains have a propensity for binding with many proteins non-specifically in the highly crowded lumen, thus leading to misfolding and aggregation. Earlier studies have shown that a variety of molecular chaperones and folding enzymes are involved in the folding of newly synthesized proteins in the ER (Braakman and Hebert, 2013; Fewell et al., 2001; Helenius et al., 1992; Hendershot, 2004). Given the broad spectrum of chaperone system in the ER, the recruitment of specific chaperones to translocating nascent chains must be carefully orchestrated to preclude non-specific associations and off-pathway aggregates. This is particularly important for two highly abundant luminal chaperones, the ER-resident Hsp70 chaperone BiP and protein disulfide isomerase (PDI) (Gidalevitz et al., 2013). BiP prefers to bind short hydrophobic sequences in

nascent chains, and its binding is tightly regulated by both its ATPase activity and co-chaperones (Flynn et al., 1991; Pobre et al., 2019). By contrast, PDI rapidly binds and releases from substrates with no amino acid specificity (Irvine et al., 2014; Noiva et al., 1993). It is unknown how the recruitment of these and other chaperones to translocating nascent chains are orchestrated in the ER lumen. Studies of cytosolic proteins have shown that the ribosome-associated factors recruit specific chaperones for the folding of newly synthesized proteins in the cytosol (Doring et al., 2017b; Willmund et al., 2013). However, such a mechanism is not feasible for the ER luminal proteins, as the ribosome is separated from the translocating nascent chain by the ER membrane. This physical barrier suggests that the Sec61 translocon and its associated factors (Gemmer and Forster, 2020) might function as a hub to recruit selective chaperones to the translocating nascent chains for their proper folding and maturation.

In this study, our substrate-trapping proteomic approach revealed that nascent proteins with weak signal sequences can transiently clog the Sec61 translocon. We identified that the translocon associated membrane protein Sec63 is responsible for clearing the clogged translocon by recruiting BiP ATPase through its luminal J-domain to bind onto the clogged nascent chain. Surprisingly, we found that the Sec63/BiP-mediated clearance of the clogged translocon is directly coupled with protein folding and maturation in the ER. Replacing a weak signal sequence with a strong signal sequence can bypass the Sec63/BiP dependence, but the translocated nascent chain becomes terminally misfolded and aggregated by binding with the incorrect chaperone PDI. Thus, we propose that signal sequences evolved to recruit appropriate chaperones to meet the unique challenges of folding the secretome in the ER.

Results

The Sec63/BiP complex clears translocons clogged with marginally hydrophobic sequences.

We began our studies by investigating how the Sec61 translocon deals with marginally hydrophobic sequences in proteins during their translocation into the ER. Earlier studies have shown that marginally hydrophobic sequences are neither hydrophobic enough to be inserted into the lipid bilayer nor hydrophilic enough to be translocated into the ER lumen, thus transiently stalling within the Sec61 translocon channel (Kida et al., 2016; Kida and Sakaguchi, 2018). To investigate the mechanism by which the translocon stalled nascent chain is cleared, we used our model C-terminally Venus-tagged substrates derived from WRB, which is a subunit of the tail-anchored membrane protein insertase (Figure S1A) (Yamamoto and Sakisaka, 2012). We have recently shown that the C-terminal TMD of WRB-Venus is translocated into the ER lumen due to the presence of the positively charged lysine residue (Figure S1A) (Sun and Mariappan, 2019). The translocation of the C-TMD of WRB into the ER lumen can be either further increased by incorporating more charged residues or reduced by adding more hydrophobic residues (Figure S1B) (Sun and Mariappan, 2019). Consistent with our recent findings, substrates with marginally hydrophobic C-terminal TMDs (K, D, RRK, and DDRRK), provisionally termed weak TMDs, were poorly inserted into the ER membrane and were translocated into the ER lumen as evidenced by the glycosylation (Figure 1A and 1B and Figure S1B). By contrast, the substrate with sufficient hydrophobicity (L) was efficiently inserted into the

ER membrane as shown by the mostly non-glycosylated form (Figure 1A and 1B). We further confirmed the translocation of weak TMDs into the ER lumen by the AviTag-based biotinylation assay. Consistent with the glycosylation assay, the weak TMD substrate (D) evaded biotinylation by cytosolic-localized biotin ligase since it was translocated into the ER lumen. By contrast, the strong TMD substrate (4L) was efficiently biotinylated, as it was properly inserted and thereby exposing the C-terminal AviTag to the cytosolic biotin ligase (Figure S1C). Furthermore, the pulse-chase experiment revealed that translocated weak TMD substrates were quickly degraded by the ER-associated quality control pathway (Figure S1D).

We hypothesized that weak TMD substrates may transiently clog the Sec61 translocon before their release into the ER lumen. To test this hypothesis, we pulled down the chromosomally tagged 2XStrep-tagged Sec61α from cells expressing weak TMD substrates. We found that the non-glycosylated form, but not the translocated and glycosylated form, selectively associated with the Sec61 translocon (Figure 1C). Of note, the translocon enrichment with the TMD substrate (D) was significantly more pronounced than the TMD substrate (K), suggesting that the former can efficiently clog the translocon and translocate into the ER lumen (Figure 1B and Figure 1C). Next, we asked how clogged substrates are cleared from the Sec61 translocon. We surmised that force might be required to release the clogged nascent chain from the Sec61 translocon into the ER lumen. We hypothesized that the abundant luminal BiP ATPase chaperone might exert a force to release the clogged Sec61 translocon. The transient expression of BiP ATPase mutants (G227D and T37G) (Awad et al., 2008) in cells inhibited the translocation of the clogged TMD into the ER lumen (Figure 1D). Co-immunoprecipitation studies further revealed a selective interaction between BiP and substrates with weak TMDs (Figure 1E).

The BiP requirement for the clearance of translocon clogged TMDs raises the question of how BiP is locally recruited to the clogged translocon. The translocon associated membrane protein Sec63 is known to recruit and activate BiP ATPase in order to facilitate post-translational protein translocation into the ER lumen (Deshaies et al., 1991; Matlack et al., 1999; Meyer et al., 2000). To test if Sec63 is required for clearing clogged TMDs from translocon into the ER lumen, we used knockout cells of either Sec63 or its interacting protein Sec62 generated by CRISPR/cas9. Although both Sec62 and Sec63 are required for BiP-mediated post-translational translocation into the ER, Sec63 was sufficient to clear the translocon clogged with a weak TMD in cells (Figure 1F). To determine biochemical features of TMDs that depend on the Sec63 activity, we tested the translocation of TMDs with varying hydrophobicity in Sec63-/- cells (Figure 1G and Figure S1B). The translocation of substrates with weak TMDs (K, D, RRK, and DDRRK) was significantly impaired as demonstrated by reduced glycosylated forms in Sec63-/- cells relative to control cells. By contrast, strong TMDs (3L and 4L) were mostly inserted into the membrane as evidenced by the appearance of mostly non-glycosylated forms in both control and Sec63-/- cells (Figure 1G). Of note, the extremely charged TMD (DDRRK) was only partially dependent on Sec63 (Figure 1G). These results suggest that a weak TMD transiently clogs at the Sec61 translocon and thereby relies on the Sec63 activity for its release into the lumen, whereas Sec63 activity is dispensable if the TMD is sufficiently hydrophobic or it is changed to a nearly hydrophilic sequence. Consistent with our data from cells, the translocation of a charged

TMD into the ER lumen was dependent on Sec63 activity in an in vitro assay performed using microsomes derived from wild type or Sec63-/- cells (Figure S1E). Surprisingly, Sec63 was dispensable for translocation of a weak TMD with a small C-terminal tail containing 26 amino acids (Figure S1F). This result indicated that the C-terminal tail length after the TMD is crucial for Sec63-dependent translocation into the ER lumen. Truncation studies further revealed that the translocation of the C-TMD with a tail longer than 100 amino acids was dependent on Sec63 activity in cells (Figure S1F). Taken together, these data suggest that Sec63 is required for clearing the Sec61 translocon clogged with weak TMDs with C-terminal tails longer than 100 amino acids.

A substrate-trapping approach identifies endogenous translocon-clogged substrates
Next, we asked the role of Sec63 in clearing the translocon clogged with endogenous substrates. To identify translocon clogging endogenous substrates, we developed a substrate trapping strategy. We hypothesized that a Sec63 J-domain mutant might be recruited to clogged translocons but may not be able to clear translocons due to its inability to activate BiP ATPase activity. This may result in trapping translocons clogged with nascent chains. To this hypothesis, we stably complemented either WT Sec63, Sec63 J-mutant (HPD/AAA), or the translocon interaction defective Sec63 mutant (Δ230-300) into Sec63-/- cells (Figure S2A). Using our model substrates, we first tested whether Sec63 J-domain or its interaction with the Sec61 translocon is essential for unclogging the Sec61 translocon. WT Sec63 complemented cells cleared translocons clogged with weak TMDs by releasing them into the ER lumen as demonstrated by the glycosylated form of substrates (Figure 2A and Figure S2B). By contrast, Sec63 J-mutant or the Sec61 interaction defective mutant (Δ230-300) was not able to clear the translocon clogging as evidenced by the mostly non-glycosylated form of substrates. Using these stable cell lines, we first performed metabolic labeling and immunoprecipitation to determine if Sec63 J-mutant could trap translocons clogged with endogenous substrates and if the trapping is dependent on its interaction with the translocon. We found that all three subunits of the newly synthesized Sec61 translocon, α, β, and γ were associated with both WT Sec63 and J-mutant but were not detected from either Sec63 knockout cells or the Sec61 interaction defective mutant (Δ230-300) expressing cells (Figure 2B). Quantification of bands further revealed that newly synthesized radiolabeled proteins were significantly enriched with Sec63 J-mutant relative to WT Sec63 or Sec63 (Δ230-300) (Figure 2B), supporting our conclusion that both Sec63 J-domain and its interaction with the translocon are important for clearing clogged translocons in cells.

We next performed a large-scale affinity purification of Sec63 as depicted in Figure 2C and subjected immunoprecipitants to mass spectrometry to identify translocon clogged endogenous clients. Our analysis of the mass spectrometry data revealed that a significant number of proteins were found in both WT Sec63 and Sec63 J-mutant relative to control Sec63-/- cells (Table S1). Consistent with our prediction, 165 secretory and membrane proteins were at least two-fold enriched in Sec63 J-mutant, and of those 101 clients were only found in Sec63 J-mutant (Figure 2D, Table S2). Also, there were a few proteins that were highly enriched in either WT or J-mutant. Notably, ACLY (ATP-citrate synthase) was enriched in the WT Sec63 sample, while PDI chaperones (PDIA3 and PDIA4) were exclusively detected in the Sec63 J-mutant

sample (Figure 2E, Table S2). Our bioinformatics analysis of secretory and membrane proteins further revealed that many Sec63's clients contained marginally hydrophobic signal sequences compared to prolactin hormone and BiP that are known to have a strong signal sequence (Figure S2C) (Kang et al., 2006). Importantly, around one-third of substrates has been previously implicated in human genetic disorders caused by mutations in corresponding genes (Table S2).

Proteins with weak signal sequences depend on Sec63 for their translocation into the ER lumen.

To determine why some substrates are prone to clogging within the translocon channel during their translocation into the ER, we constructed C-terminally FLAG-tagged expression plasmids for two substrates, AGAL and CD44 (Figure S3A and S3B). The *AGAL* gene encodes for a lysosomal enzyme called α-galactosidase A, the activity of which is altered in Fabry disease due to genetic mutations (Germain, 2010). The *CD44* gene encodes the plasma membrane-localized single spanning membrane protein with an N-terminal signal sequence. We first examined if these substrates depended on Sec63 for their translocation into the ER. Immunoblotting of transiently expressed AGAL exhibited mostly glycosylated form, with a small fraction of non-glycosylated AGAL (Figure 3A). The glycosylation of AGAL was verified by digestion with Endo H, which removes high mannose N-glycans from glycoproteins. In sharp contrast, the translocation of AGAL as mirrored by the glycosylated form was significantly impaired in Sec63-/- cells relative to wild type cells (Figure 3A). The expression of CD44 in wild type cells showed two glycosylated bands. While the faster migrating glycosylated band was sensitive to Endo H, but the slower migrating complex type of N-glycan was less sensitive to Endo H, indicating that it is localized to the post-ER compartment (Figure 3A). Similar to AGAL, the translocation of CD44 was significantly reduced as it displayed mostly the non-glycosylated form in Sec63-/- cells. We then asked if the translocation defects associated with AGAL and CD44 could be rescued by WT Sec63. Indeed, both AGAL and CD44 were translocated at similar levels to that of wild type cells when Sec63-/- cells were complemented with WT Sec63 (Figure 3B). By contrast, Sec63 J-mutant complemented cells showed a severe defect in translocation of both AGAL and CD44 relative to WT Sec63 complemented cells. (Figure 3B). This result suggests that the activation of BiP ATPase by the J-domain of Sec63 is critical for the translocation of these proteins. Since Sec63 functions along with Sec62 in post-translational protein translocation, we wondered whether Sec62 is also required for translocation of AGAL and CD44. While the knockout of Sec62 had a significant effect on the translocation of CD44 as mirrored by less glycosylated forms, it had a subtle effect on the translocation of AGAL, suggesting that Sec62 plays a substrate-dependent role in promoting translocation of secretory proteins (Figure 3C).

To understand the biochemical features that define Sec63 substrates, we specifically focused on the hydrophobicity of signal sequences because many of them contain less hydrophobic helical (h) region (Figure S2C). Since we found that weak TMDs can transiently clog the Sec61 translocon, we hypothesized that weak signal sequences may also clog the Sec61 translocon, and thereby recruiting the Sec63/BiP complex for the translocation of their following nascent chains. To test this, we swapped the signal sequence of AGAL or CD44 with

the strong signal sequence from prolactin and tested its Sec63 dependency by expressing it in Sec63-/- cells. In support of our hypothesis, Prl-ss-AGAL bypassed the Sec63 requirement as reflected by a fully glycosylated form of AGAL in both WT and Sec63-/- cells, whereas AGAL with its endogenous signal sequence poorly glycosylated in Sec63-/ cells (Figure 3D). We also obtained a similar result for CD44 where Prl-ss-CD44 was mostly matured independent of Sec63 activity (Figure 3E). Furthermore, we independently tested Sec63 dependency on the ENaC complex, which is a Na+ re-absorption channel composed of three membrane subunits, α, β, and γ. (Figure S3C). The translocation of ENaC α as indicated by glycosylation was dependent on Sec63, but not Sec62, whereas ENaC β and γ subunits were translocated independent of Sec63 (Figure 3F and Figure S3D). This result is consistent with our prediction that the targeting sequence of ENaC α (1st TMD) is less hydrophobic than targeting sequences of the other two subunits β and γ (Figure S3C). Indeed, replacing the targeting TMD of ENaC α with the targeting TMD of ENaC β resulted in partial rescue of translocation defects observed in Sec63-/- cells (Figure S3E). Collectively, our results suggest that a weak targeting sequence in the form of either a signal sequence or TMD requires Sec63/BiP for efficient translocation of its mature domain into the ER lumen.

Sec63 co-translationally assists the translocation of secretory proteins into the ER

Previous studies have established that Sec63 is predominantly involved in post-translational translocation of substrates into the ER lumen by recruiting and activating BiP ATPase through its J-domain (Matlack et al., 1999; Misselwitz et al., 1998). Sec63 also appears to be required for the co-translational protein translocation of select substrates (Brodsky et al., 1995; Conti et al., 2015; Jung et al., 2014; Lang et al., 2012; Young et al., 2001), but its precise role in this pathway is ill-defined. The recent structural studies of the Sec61/Sec63/Sec63 complex from yeast suggest that both Sec63 and ribosome share similar binding sites in the Sec61 translocon (Itskanov and Park, 2019; Wu et al., 2019). This means both Sec63 and the ribosome cannot bind to the Sec61 translocon at the same time, questioning the role of Sec63 in the co-translational translocation process. We therefore wanted to determine whether our novel Sec63' clients are translocated via co- or post-translational translocation pathway. To test this, we performed in vitro translocation and protease protection assays. AGAL was sensitive to proteinase K when it was translated in the rabbit reticulocyte lysate (RRL) based translation system in the absence of the ER-derived rough microsomes. AGAL was translocated and glycosylated in the presence of microsomes and thereby insensitive to the protease digestion (Figure 4A, and 4B). In sharp contrast, AGAL was completely digested by proteinase K when microsomes were post-translationally incubated with the supernatant of the translation reaction in which ribosomes were removed by centrifugation (Figure 4A and 4B). To directly detect the co-translational translocation of AGAL, we in vitro translated transcripts encoding AGAL lacking the stop codon either in the absence or presence of microsomes. Due to the lack of stop codon, the nascent polypeptide was still attached to the ribosome via tRNA peptidyl linkage (Figure 4C). AGAL translated in the absence of microsomes was completely digested by proteinase K. By contrast, AGAL or Prl-ss-AGAL translated in the presence microsomes was translocated into the ER lumen as demonstrated by glycosylation, which was resistant to the protease digestion. The fact that the ribosome-associated nascent chain could be translocated and glycosylated

strongly suggests that AGAL is a substrate of the co-translational translocation pathway where Sec63 acts to help with the translocation of AGAL.

The aforementioned data show that Sec63 dependent substrate (AGAL) and Sec63 independent substrate (Prl-ss-AGAL) are co-translationally translocated into the ER lumen. We therefore wanted to determine at which step during the biosynthesis of AGAL, Sec63 would assist the translocation of AGAL. To address this, we in vitro translated transcripts lacking stop codon encoding defined length of AGAL either in the presence or absence of microsomes (Figure 4D). We observed minimal protease protected fragments of nascent chains of AGAL-80, -100, and -130, but nascent chains of 220 and 429 (full-length) were protease protected (Figure 4D). This result suggests that the AGAL nascent chain up to the length of 130 amino acids is looped into the cytosol through the gap between the ribosome and the translocon, thus becoming accessible to proteinase K. In sharp contrast, Prl-ss-AGAL, which is not dependent on Sec63, showed protease protected nascent chains with all different lengths of nascent chains, indicating that the successful translocation is initiated right after the signal sequence engages the Sec61 translocon (Figure 4E). Taken together, our data suggest that Sec63 co-translationally acts on translocons clogged with substrates containing weak signal sequences and assists with the translocation of their following polypeptides into the ER lumen.

Signal sequences dictate the chaperone recruitment to translocating nascent chains in the ER

Next, we asked why weak signal sequences are frequently found in secretory and membrane proteins despite their inefficiency in mediating translocation (Rane et al., 2004), and why they were not evolved into strong signal sequences. For instance, marginal hydrophobicity of the AGAL signal sequence is conserved in mammals, particularly the charged arginine residue within the hydrophobic core region (Figure 5A). We hypothesized that signal sequence-dependent modulation of Sec61 translocon components may be responsible for recruiting specific downstream luminal chaperones to bind onto nascent chains. Accordingly, we predicted that AGAL nascent chain would preferentially bind to BiP chaperone since its weak signal sequence modulates the Sec61 components such that Sec63 is recruited to the translocon. Consistent with our prediction, co-immunoprecipitation of AGAL revealed a strong binding to BiP relative to Sec63-independent substrate prolactin, which showed less binding to BiP (Figure 5B). This result also suggested that Sec63 independent substrates may bind to other luminal chaperones.

We reasoned that PDI chaperones may bind to Sec63-independent substrates because PDIs are a family of highly abundant chaperones in the ER and are known to bind substrates with less specificity (Noiva et al., 1993). This is supported by our observation that PDIs were highly enriched in the affinity purification of Sec63 J-mutant, which cannot activate BiP to bind onto translocating nascent chains (Figure 2E). Besides, chemical cross-linking data revealed that PDI chaperone selectively interacted with prolactin, which has a strong signal sequence (Figure 5C). By contrast, almost no cross-linking was observed between PDI and AGAL, which has a weak signal sequence (Figure 5C). We next asked whether signal sequences are solely responsible for recruiting correct chaperones to nascent chains. To address this, we

metabolically labeled cells expressing AGAL carrying either its weak endogenous signal sequence or a strong signal sequence from prolactin and analyzed by autoradiography after chemical cross-linking and immunoprecipitation. The result showed that the cross-linked pattern of AGAL was dependent on whether it carried a weak or strong signal sequence (Figure 5D). Our conclusion of signal sequence-dependent chaperone selection was further supported by the observation that PDI chaperone preferentially cross-linked with Prl-ss-AGAL compared to AGAL (Figure 5E). This result also implies that the weak signal sequence of AGAL is crucial for recruiting BiP chaperone via Sec63 to shield its nascent chains from binding to PDI. Collectively, our results suggest that proteins containing weak signal sequences preferentially bind to BiP that is recruited via Sec63, while proteins containing strong signal sequences bind to PDI chaperones.

Signal sequence-dependent chaperone recruitment is crucial for protein folding and maturation in the ER.

We hypothesized that the privileged BiP binding to nascent chains via Sec63 may not only facilitate their translocation into the ER but may also promote their folding and maturation by shielding them from inappropriate interactions. To test this, we monitored the localization of AGAL in cells by confocal microscopy. Properly folded AGAL should be trafficked to lysosomes, where it is responsible for breaking down of a fatty substance, globotriaosylceramide. We found that AGAL was mostly localized to lysosomes as demonstrated by its co-localization with the lysosomal marker protein LAMP1 (Figure 6A and 6B). By contrast, Prl-ss-AGAL was sparsely localized to lysosomes, suggesting that its trafficking was impaired in cells. Since AGAL is also known to be secreted into the medium, when properly folded (Ioannou et al., 1992), we performed a pulse-chase analysis to determine if AGAL or Prl-ss-AGAL is secreted (Figure 6C). The intracellular level of AGAL was reduced during the chase period with a concomitant appearance in the extracellular medium. Treating cells with the proteasome inhibitor MG132 had a little effect on AGAL secretion, suggesting that most AGAL is properly folded and secreted without being routed for proteasomal degradation. Strikingly, nearly no secretion was detected for Prl-ss-AGAL during the chase period, which suggests that Prl-ss-AGAL is terminally misfolded and accumulated in the ER lumen. Surprisingly, the intracellular misfolded Prl-ss-AGAL was not even recognized for ER-associated degradation as there was a little further accumulation of Prl-ss-AGAL upon inhibiting the proteasome. Furthermore, the terminally misfolded Prl-ss-AGAL was aggregated in cells since a significant proportion of Prl-ss-AGAL was localized in detergent-insoluble pellet compared to AGAL (Figure 6D). These results emphasize the importance of Sec63-mediated privileged BiP binding to AGAL nascent chains precluding inappropriate interactions and off-pathway aggregates in the ER.

To directly determine the role of signal sequence in protein folding and maturation, we monitored the α-galactosidase activity of both AGAL and Prl-ss-AGAL. Remarkably, Prl-ss-AGAL showed a little enzymatic activity compared to wild type AGAL, which exhibited high enzymatic activity (Figure 6E). This was not due to transient overexpression of AGAL as even cells stably expressing Prl-ss-AGAL also displayed a very little α-galactosidase activity despite Prl-ss-AGAL being highly accumulated compared to wild type AGAL (Figure S4A). This result suggests that the loss of signal sequence-dependent recruitment of BiP chaperone to nascent

chains of Prl-ss-AGAL caused their terminal misfolding in the ER lumen. Another explanation for this result is that the endogenous signal sequence of AGAL had a specific feature that is important for protein folding and maturation of α-galactosidase. To exclude this possibility, we replaced the endogenous weak signal sequence of AGAL with another weak signal sequence from CD44 (Figure 3E) and tested its folding status by monitoring α-galactosidase activity. AGAL containing the weak signal sequence from CD44 was active at the same level as wild type AGAL (Figure 6E), arguing that any weak signal sequence that recruits BiP chaperone via Sec63 can restore the folding and maturation of AGAL. Furthermore, we asked whether protein folding is disrupted when a protein with a strong signal sequence is swapped with a weak signal sequence. To address this, we swapped the strong signal sequence of prolactin with the weak signal sequence of AGAL and tested its folding status by monitoring the secretion of prolactin. As expected, the translocation of AGAL-ss-prolactin now became dependent on the Sec63 activity, whereas wild type prolactin was less dependent on Sec63 (Figure S5B). Of note, the expression of AGAL-ss-prolactin was significantly reduced in control cells compared to prolactin. The pulse-chase experiment further revealed that the secretion of prolactin was noticeably reduced for AGAL-ss-prolactin compared to prolactin, suggesting that strong signal sequences are important for faster maturation and secretion of proteins (Figure S5C). Thus, our data suggest that signal sequence-dependent chaperone selection is crucial for protein folding and maturation in the ER.

Discussion

In this study, we uncover a new function for signal sequences. In addition to targeting roles (Blobel and Dobberstein, 1975), they can recruit correct chaperones for folding and maturation of nascent chains in the ER. We found that proteins with a weak signal sequence or TMD can clog the Sec61 translocon (Figure 7). This translocon clogging is sensed by the translocon associated protein Sec63, which recruits BiP chaperone via its luminal J-domain to bind onto the nascent chain. This privileged BiP binding not only facilitates the release of the clogged nascent chain from the translocon but also prevents non-productive interactions and promotes protein folding. Conversely, proteins with strong signal sequences bypass the Sec63/BiP-mediated folding pathway and preferentially bind to PDI chaperones for their folding and maturation. Thus, our studies suggest that evolution has exploited diversity in signal sequences to recruit selective luminal chaperones through modulating translocon-associated factors for proper folding and maturation of secretory proteins.

The decisive step in the biogenesis of secretory and membrane proteins is the interaction of hydrophobic targeting sequence with the Sec61 translocon in a hairpin looped orientation. This allows the lateral exit of the hydrophobic targeting sequence from the translocon to the lipid phase with the concomitant translocation of the succeeding nascent chain into the ER lumen. Despite the pivotal role of hydrophobicity in this process, weak targeting sequences are frequently found in secretory and polytopic membrane proteins (Hegde and Kang, 2008). To address how the Sec61 translocon deals with these suboptimal hydrophobic sequences, we devised a series of model substrates whose hydrophobicity of C-terminal TMDs was varied by introducing charged or hydrophobic amino acids. Consistent with the free energy prediction (Hessa et al., 2007), TMDs with strong hydrophobicity are efficiently inserted into the

ER membrane. By contrast, TMDs with insufficient hydrophobicity are transiently clogged at the Sec61 translocon and are eventually released into the ER lumen. This is consistent with the previous observation that insufficient hydrophobic sequences are stalled in the interior pore of the Sec61 translocon channel (Kida and Sakaguchi, 2018).

Our data suggest that Sec63 monitors the Sec61 translocon status and clears the translocon clogged with nascent chains. The key function of Sec63 is to recruit the luminal chaperone BiP ATPase through its J-domain to bind onto the clogged nascent chain. This results in the release of the clogged nascent chain into the ER lumen. Our findings could explain why the translocation of the C-terminus of membrane proteins is impaired upon disrupting the Sec62/63 complex in yeast (Jung et al., 2014). Future studies are required to understand how Sec63-recruited BiP pulls the translocon clogged nascent that is extruded into the cytosol. One possible mechanism is that BiP binding to the nascent chain near the translocon pore may release the nascent chain by the entropic pulling force (De Los Rios and Goloubinoff, 2016). To identify the endogenous translocon clogging substrates, we developed a “substrate-trapping” approach in which the Sec63 J-domain mutant fails to unclog the clogged translocon. This unbiased proteomic approach identified numerous secretory and membrane proteins that are clogged at the Sec61 translocon complex when Sec63 is not able to activate BiP ATPase via its J-domain. Our analysis of mass spectrometry data suggest that many of the substrates harbor weak signal sequences or signal anchor sequences (1st TMD). Our results are consistent with a recent independent proteomic study performed using Sec63 depleted cells (Schorr et al., 2020). Although many substrates are found in both approaches, a number of them are only found in one of the approaches. One explanation for this discrepancy is that our approach was devised based on the function of the Sec63 J-domain, which is responsible for clearing the clogged translocon. In addition to marginally hydrophobic sequences, other features in signal sequences such as length and flanking amino acids might also influence clogging at the Sec61 translocon (Ziska et al., 2019).

Earlier studies have established the role of Sec63 in the post-translational translocation pathway (Matlack et al., 1999; Misselwitz et al., 1998). Growing evidence suggests that Sec63 is also co-translationally involved in the translocation of select substrates into the ER lumen (Brodsky et al., 1995; Conti et al., 2015; Jung et al., 2014; Lang et al., 2012; Young et al., 2001). However, recent structural studies of the yeast post-translational translocation complex suggest that Sec63 binding sites in the Sec61 translocon highly overlaps with the ribosome binding, which occurs during the co-translational protein translocation (Itskanov and Park, 2019; Wu et al., 2019). These studies raise the question of whether Sec63’ clients follow the co- or post-translational translocation pathway. Three of our observations suggest that Sec63’ substrates use the co-translational protein translocation pathway. First, Sec63 clients identified in this study are larger than 160 amino acids in length, which appear to be a limitation for post-translational protein translocation in mammalian cells (Lakkaraju et al., 2012). Second, our in vitro experiment shows that the Sec63 substrate AGAL could not be post-translationally translocated into the ER. Third, AGAL was efficiently translocated and glycosylated even when it is still bound to the ribosome. These observations also raise the intriguing question of how Sec63 can access the clogged translocon that is bound to the ribosome. Our protease protection assays

indicate that during the initial phase of translocation, the nascent chain is extruded into the cytosol through the gap of ribosome and translocon, likely due to a non-hairpin engagement of weak signal sequence with the Sec61 translocon (Figure 7). We propose that Sec63 may sense the clogged Sec61 translocon by directly binding to the extruded cytosolic nascent chain from the clogged translocon. Once recruited to the clogged channel, Sec63 may take advantage of the weakened interaction between Sec61 and the ribosome and displace the ribosome to engage with the translocon. Our findings of Sec63 recruitment to the clogged translocon is consistent with the recent observations that Sec62/Sec63 is selectively enriched with the Sec61 translocon when the translocation is blocked by a passenger domain folding (Conti et al., 2015).

The weak signal sequences are not as efficient as strong signal sequences in mediating protein translocation into the ER. It is estimated that more than 5% of newly synthesized proteins with weak signal sequences are not translocated and are routed for proteasomal degradation (Hegde and Zavodszky, 2019; Rane et al., 2004). Surprisingly, despite this inefficiency, weak signal sequences are commonly found and conserved across all mammals. One function for weak signal sequences is that they can selectively impair the translocation of their mature domains into the ER during ER stress conditions, thus preventing cell death caused by protein misfolding and aggregation (Kang et al., 2006). Our study suggests that substrates that depend on BiP for protein folding have likely evolved with weak signal sequences, thus ensuring the co-translational recruitment of BiP via the Sec63 J-domain. The privileged BiP binding to nascent chains protects them from inappropriate interactions and promotes folding in the ER. This idea is consistent with our observation that skipping Sec63-mediated recruitment of BiP by replacing the weak signal sequence of AGAL with a strong signal sequence leads to terminal misfolding and aggregation. We predict that the privileged recruitment of BiP to a nascent chain may not only occur during the initial phase of protein translocation but may also occur at any time during translocation as long as nascent chains contain a translocon clogging element such as a pseudo transmembrane segment (Hessa et al., 2007).

Our results suggest that PDI can opportunistically bind to translocating nascent chains in the absence of BiP recruitment to nascent chains. Moreover, proteins with a strong signal sequence predominantly interact with PDI as their nascent chains are not coated by BiP. Given that PDI is a highly abundant chaperone family and can dynamically bind and release substrates on a sub-millisecond time scale (Gidalevitz et al., 2013; Irvine et al., 2014), it could be a default chaperone for nascent chains that are entering into the ER lumen. This promiscuous binding nature of PDI poses a challenge for secretory proteins that strictly depend on BiP to shield their linear hydrophobic stretches until their translation is terminated. We speculate that evolution solved this problem by tagging these proteins with weak sequences, which ensure the privileged recruitment of BiP via Sec63 to nascent chains. Our data showing BiP independent translocation of prolactin explains the earlier observation that why prolactin nascent chain fails to interact with BiP (Kassenbrock et al., 1988). Our studies find that a strong signal sequence containing prolactin preferentially interacts with PDI. In contrast to AGAL, PDI binding to prolactin does not impede the folding and secretion of prolactin. This suggests that PDI binding to some nascent chains can allow co-translational protein folding since its ability to bind and release substrates dynamically (Irvine et al., 2014). This idea is in line with the

observation that PDI binds to a substrate that co-translationally folds in the ER lumen (Robinson et al., 2017).

In the cytosol, the ribosome-associated factors can recruit specific chaperones to bind onto nascent chains (Doring et al., 2017a; Willmund et al., 2013). Since ribosomes-bound to translocons cannot directly access the ER lumen, we speculate that the translocon-associated accessory proteins function as a hub to recruit specific luminal chaperones to bind onto translocating nascent chains. Our findings suggest that signal sequences contain information to modulate translocon-associated factors, which in turn would recruit specific luminal chaperones for protein folding. In addition to the weak signal sequence, we propose other features in signal sequences may also be responsible for modulating translocon accessory proteins and their downstream chaperones. This interpretation is supported by the recent findings that glycine and proline-rich signal sequences depend on the translocon-associated protein (TRAP) complex (Nguyen et al., 2018). Curiously, the endogenous TRAP complex is recently shown to form a complex with Calnexin chaperone (Lee et al., 2020), suggesting a potential role in the folding of glycoproteins (Molinari and Helenius, 2000). We propose that multiple translocation pathways (Aviram and Schuldiner, 2017) in cells may be evolved to meet the challenges of folding the complex secretome by recruiting protein-specific chaperones. This may explain why most secretory proteins in yeast use the Sec63/BiP-mediated post-translational translocation pathway (Ast et al., 2013), likely their dependency on BiP for folding and maturation.

Earlier studies have shown that mutations in the Sec63 gene can cause autosomal dominant polycystic kidney diseases (PKD) due to defects in protein maturation and cell surface localization of polycystin 1 and presumably other proteins (Besse et al., 2017; Davila et al., 2004). Our studies suggest that around one-third of our Sec63's clients have been previously implicated in human genetic disorders, including lysosomal storage disorders, and PKD, which suggest that these proteins are particularly vulnerable to misfolding and aggregation due to genetic mutations. In light of our findings, selective induction of BiP by small molecules, such as 147 (Plate et al., 2016), may be a potential therapeutic approach to treat these diseases. Furthermore, our studies may guide the design of customized signal sequences for the production of a growing number of therapeutic recombinant secretory proteins, including enzymes, growth factors, and antibodies.

ACKNOWLEDGMENTS

We are grateful to Jean Kanyo and TuKiet Lam from the Yale Keck Protromic facility for assisting with analyzing mass spectrometry data. We thank Suhila Appathurai for generating HEK293 Sec61α-2x Strep-tag cells. We thank Jacob Culver for the discussion and comments. We thank Zai-Rong Zhang for comments on the manuscript. We want to thank Matthew Jordan for suggesting the GRAVY tool. This work is funded by NIH grants R01GM117386 (M.M), R21AG056800 (M.M), and Yale Start-up fund.

CONTRIBUTIONS

SS designed and performed most of the experiments. XL performed BiP interaction studies. MM supervised the project. SS and MM wrote the manuscript with inputs from XL.

DECLARATION of INTERESTS

The authors declare no competing interests.

Figure Legends

Figure 1. Sec63/BiP is required for clearing translocons clogged with marginally hydrophobic TMDs

(A) A diagram illustrating the translocon clogged with the marginally hydrophobic C-terminal TMD of WRB-Venus. While the cleared substrate is translocated into the ER lumen and glycosylated, the properly inserted substrate is shown as a non-glycosylated form. The red circle in the C-TMD indicates a charged residue.

(B) Cells expressing C-TMDs of substrates bearing the indicated charged residues (shown in red color) or hydrophobic residue (indicated in blue color) were radiolabeled and analyzed by SDS-PAGE and autoradiography after immunoprecipitation of substrates with anti-GFP antibodies.

(C) The indicated substrates were expressed in chromosomally 2X Strep-tagged Sec61α cells, lysed, and pulled down with Strep-Tactin beads and analyzed by immunoblotting for the indicated antigens. The substrates were immunoblotted with anti-GFP antibodies.

(D) Cells co-expressing the indicated versions of BiP and the substrate were analyzed by Immunoblotting with anti-BiP antibody or anti-GFP antibodies for the substrate.

(E) Cells expressing the indicated FLAG-tagged substrates were immunoprecipitated with anti-FLAG antibody and analyzed by immunoblotting with an anti-BiP antibody or anti-GFP antibodies for substrates.

(F) Control HEK293, Sec62-/-, or Sec63-/- cells expressing the indicated substrates were radiolabeled and analyzed as in (B). Immunoblots show the knockout of either Sec62 or Sec63.

(G) Control HEK293 or Sec63-/- cells expressing the indicated substrates were immunoprecipitated with anti-GFP antibodies and analyzed by autoradiography.

See also Figure S1

Figure 2. Identification of translocon clogged endogenous substrates of Sec63

(A) Control HEK293 or Sec63-/- cells stably expressing the indicated Sec63 constructs were transfected with the C-TMD of WRB substrate carrying either lysine (K) or aspartic acid (D). The transfected cells were radiolabeled and analyzed by autoradiography after immunoprecipitation with anti-GFP antibodies for substrates. The immunoblot shows the expression of Sec63 from the indicated cell lines.

(B) The indicated Sec63-/- complemented cell lines were radiolabeled and analyzed by autoradiography after immunoprecipitation with anti-FLAG beads. The traces are densitometry profiles of lanes and shown at the right side of the autoradiograph, illustrating a selective enrichment of signals with Sec63 J-mutant.

(C) A scheme describing affinity purification and identification of translocon clogged substrates by mass spectrometry.
(D) Venn diagram of protein distribution among the three groups. The diagram was built from mass spectrometry data from Table S1.
(E) The scatter plot shows the normalized spectral counts comparing Sec63 and Sec63 J-mutant samples. The plot was created from mass spectrometry data from Table S1 using the Scaffold software.
See also Figure S2

Figure 3. The translocation of proteins with weak targeting sequences depend on Sec63
(A) Control HEK293 or Sec63-/- cells were transfected with AGAL-FLAG or CD44-FLAG. The cell lysates were either untreated or treated with Endo H and analyzed by immunoblotting with an anti-FLAG antibody for substrates and anti-Sec63 antibodies. The open circle indicates the signal sequence cleaved and non-glycosylated form of AGAL. The closed circle indicates the non-glycosylated form of AGAL with the signal sequence.
(B) Control HEK293 or Sec63-/- cells stably expressing the indicated constructs were transfected with AGAL-FLAG or CD44-FLAG. The cell lysates were directly analyzed by immunoblotting as in (A). Star indicates FLAG-tagged Sec63 proteins.
(C) The indicated cell lines were transfected with plasmids expressing AGAL-FLAG or CD44-FLAG, and the cell lysates were directly analyzed as in (A).
(D) Top: A cartoon depicting the signal sequence of AGAL is replaced with the signal sequence of prolactin, termed Prl-ss-AGAL. Bottom: The indicated cell lines were transfected with AGAL-FLAG or Prl-ss-AGAL-FLAG and analyzed by immunoblotting with an anti-FLAG antibody.
(E) Top: A cartoon depicting the signal sequence of CD44 is swapped with the signal sequence of prolactin, termed Prl-ss-CD44. The indicated cells were transfected and analyzed as in (D).
(F) The plasmids expressing ENaC α-HA, ENaC β-HA, or ENaC γ-HA were transfected into control HEK293 or Sec63-/- cells. The cells were radiolabeled and immunoprecipitated with anti-HA antibodies for ENaCs and analyzed by autoradiography.
See also Figure S3

Figure 4. Sec63 substrates are co-translationally translocated into the ER lumen
(A) Scheme depicting the protocol for proteinase K (PK) treatment for the co- or post-translational protein translocation assay.
(B) For the co-translational protein translocation, transcripts encoding AGAL-FLAG were in vitro translated in the rabbit reticulocyte lysate (RRL) based translation system including either buffer or canine rough microsomes (CRM). After translation, samples were digested with PK and analyzed by autoradiography after immunoprecipitation with anti-FLAG antibody beads. For the post-translational translocation, transcripts encoding AGAL-FLAG were translated and centrifuged to remove the ribosomes. The supernatant was incubated with RM and digested with PK and analyzed after immunoprecipitation with anti-FLAG antibody beads.
(C) AGAL or Prl-ss-AGAL transcripts lacking stop codon were in vitro translated in RRL including either buffer or CRM. The reactions were treated with PK and analyzed as in (B). Note that transcripts lacking stop codon produce nascent chains that remain attached to ribosomes via a covalent peptidyl-tRNA bond.

(D). The ribosome-associated nascent chains (RNCs) of the indicated lengths of AGAL were produced in RRL including either buffer or CRM and digested with PK before analyzing by autoradiography. AGAL was analyzed after immunoprecipitation with anti-FLAG antibody beads. The star symbol indicates a distortion of the AGAL-130 nascent chain caused by the co-migration with abundant hemoglobin from RRL.
(E) RNCs of the indicated lengths of Prl-ss-AGAL were produced and analyzed as in (D).

Figure 5. Signal sequence-dependent interaction between translocated proteins and luminal chaperones.
(A) Conservation of AGAL signal sequence in different mammalian species. Blue shade indicates the hydrophobic region (h-region) of the signal sequence. Charged and hydrophilic amino acids in the h-region are indicated in red color. Note that the positively charged arginine residue is conserved among different mammalian species.
(B) Cells were transfected with AGAL-FLAG or prolactin-FLAG and immunoprecipitated with anti-FLAG antibody. The immunoprecipitants were analyzed by immunoblotting (IB) with an anti-FLAG antibody for substrates and an anti-BiP antibody.
(C) Cells expressing AGAL-FLAG or prolactin-FLAG were subjected to chemical cross-linking with the cysteine reactive bismaleimidohexane (BMH), immunoprecipitated with anti-FLAG beads, and analyzed by immunoblotting with the indicated antibodies.
(D) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were radiolabeled and subjected to increasing concentration of a lysine reactive crosslinker, DSS (disuccinimidyl suberate). The cross-linked samples were immunoprecipitated with anti-FLAG beads and analyzed by autoradiography. Star denotes a non-specific band.
(E) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were subjected to increasing concentration of BMH crosslinker and analyzed as in (C).

Figure 6. Signal sequence-dependent chaperone recruitment promotes protein folding and maturation
(A) AGAL-FLAG or Prl-ss-AGAL-FLAG was co-transfected with LAMP1-CFP in U2OS cells and processed for immunostaining with anti-FLAG antibody for AGAL and anti-GFP antibodies for LAMP1. Scale bars are 10 μm.
(B) The percentage of cells from (A) containing AGAL puncta that colocalize with the lysosomal marker LAMP1-GFP. Error bar represents the standard deviation (n=3).
(C) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were radiolabeled and chased for the indicated time points. Both cells and medium were collected at each time point and analyzed by autoradiography after immunoprecipitation with an anti-FLAG antibody for AGAL.
(D) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were solubilized with the Triton X-100 buffer. Soluble and insoluble fractions were separated by centrifugation and analyzed by immunoblotting with an anti-FLAG antibody for AGAL.
(E) The upper panel diagram shows the constructs where the AGAL signal sequence was swapped with either the signal sequence of prolactin or CD44. Bottom panel: Cells expressing AGAL-FLAG, Prl-ss-AGAL-FLAG, or CD44-ss-AGAL-FLAG were lysed and assayed for α-galactosidase activity. AGAL immunoblot was probed with an anti-FLAG antibody. The activity

was presented as arbitrary units (a.u) after subtracting the signal from cells transfected with an empty vector. Error bar represents the standard deviation (n=3).
See also Figure S4

Figure 7. A model for signal sequence-dependent chaperone recruitment and protein folding in the ER.
(A) Nascent chains that contain marginally hydrophobic or weak signal sequences such as AGAL transiently clog the Sec61 translocon. This leads to the privileged recruitment of BiP to the clogged translocon via Sec63 for releasing as well as shielding nascent chains from non-productive interactions, thus promoting folding and maturation.
(B) Swapping a weak signal sequence with a strong signal sequence bypasses the privileged BiP recruitment to the translocating nascent chain, but the translocated nascent chain becomes terminally misfolded and aggregated after binding the wrong chaperone PDI.

STAR★Methods

KEY RESOURCES TABLE

REAGEMT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat monoclonal anti-FLAG	Biolegend	637303
Rabbit polyclonal anti-HA	From the Hegde Lab	
Mouse monoclonal anti-HA	Biolegend	161312
Mouse monoclonal anti-Tubulin	Abcam	ab11312
Rabbit polyclonal anti-GFP	From the Hegde Lab	
Mouse monoclonal anti-BiP	BD Bioscience Pharmingen	610979
Rabbit polyclonal anti-Sec62	From the Hegde Lab	
Rabbit polyclonal anti-Sec63	From the Hegde Lab	
Rabbit polyclonal anti-Sec61α	From the Hegde Lab	
Rabbit polyclonal anti-Sec61β	From the Hegde Lab	
Rabbit polyclonal anti-PDIA1	Enzo Life Sciences	ADI-SPA-890-D
Rabbit polyclonal anti-PDIA3	Cell signaling	2881
Avidin-HRP	Jackson Immuno reserach	016-030-084
Goat anti-mouse HRP	Jackson Immuno reserach	11-035-003
Goat anti-rabbit HRP	Jackson Immuno reserach	111-035-144

Goat anti-rat HRP	Cell signaling	7077S
Chemicals, peptides, and recombinant proteins		
Strep-Tactin XT beads	IBA, Germany	2-4010-010
Rat anti-FLAG L5 affinity gel	BioLegend	651503
Mouse anti-HA magnetic beads	Pierce	88837
Goat anti-Rabbit Cy3	Jackson Immuno reserach	711-165-152
Goat anti-Rat Cy2	Jackson Immuno reserach	112-225-167
Protein A agarose	Repligen	CA-PRI-0100
DMEM	Corning	10–013-CV
FBS	Gibco	16000044
Penicillin	Gibco	15140122
Lipofectamine 2000	Invitrogen	11668-019
Doxycycline	Clontech	631311
Hygromycin	Invitrogen	10687010
Blasticidin	RPI Research Products	B12150-0.01
EDTA-free Protease inhibitor cocktail	Roche	11873580001
Poly L-lysine	Peptides International	OKK-3056
Express ^{35}S protein labeling Mix	Perkin Elmer	NEG072007MC
^{35}S Methionine (for in vitro translation)	Perkin Elmer	NEG009T005MC
DMEM (-Glut/-Cys/-Met)	Gibco	21013024
Dialysed FBS	Thermo Fisher Scientific	26400044
Digitonin	Millipore	300410
EndoH	NEB	P0702
Apyrase	NEB	M0398
Trichloroacetic acid	Sigma	T6399
Cycloheximide	Sigma	7077S
Rabbit Reticulocyte Lysate	Green Hectares	Ph:1-800-GHLYSAT
RNAsin	Promega	N251B

SP6 polymerase	NEB	#M0207S
Proteinase K	Roche	3115879001
BMH	Thermo Fisher Scientific	22330
DSS	Thermo Fisher Scientific	21555
Fluoromount G	SouthernBiotech	0100–01
MG132	Millipore	133407-82-6
4-MU-a-galactosidase substrate	Sigma	M7633
Luciferase Reporter Gene Assay Kit	Roche	11669893001
SuperSignal West Pico Chemiluminescent substrate	Thermo Fisher Scientific	34580
SuperSignal West Femto Chemiluminescent substrate	Thermo Fisher Scientific	34096
SuperSignal West Chemiluminescent substrate	Thermo Fisher Scientific	32106
Experimental Models: Cell lines		
HEK293 Flip-In T-Rex cells	Invitrogen	R78007
HEK293 Sec61α-2x Strep-tag cells	This study	
HEK293 Sec63-/- cells	(Sun and Mariappan, 2019)	
HEK293 Sec62-/- cells	(Li et al., 2020)	
HEK293 Sec63-/- cells plus Sec63-FLAG	(Li et al., 2020)	
HEK293 Sec63-/- plus Sec63 J-mutant-FLAG (H132A/P133A/D134A)	(Li et al., 2020)	
HEK293 Sec63-/- plus Sec63 J-mutant-FLAG (Δ230-300)	(Li et al., 2020)	
U2OS cells	The Schlieker Lab	
HeLa cells stably expressing AGAL-FLAG	This study	
HeLa cells stably expressing Prl-AGAL-FLAG	This study	
Recombinant DNA		
pCDNA 5/FRT/TO WRB-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM551
pCDNA 5/FRT/TO WRB (K164L)-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM553
pCDNA 5/FRT/TO WRB (3L)-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM642

pCDNA 5/FRT/TO WRB (4L)-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM640
pCDNA 5/FRT/TO WRB (K164D)-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM556
pCDNA 5/FRT/TO WRB-RRK-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM560
pCDNA 5/FRT/TO WRB-DDRRK-HA-NFT-Venus	(Sun and Mariappan, 2019)	MM641
pCDNA/Zeo WRB (K164D)-HA-NFT-Venus-AviTag	This study	MM816
pCDNA/Zeo WRB (4L)-HA-NFT-Venus-AviTag	This study	MM817
pCDNA 5/FRT/TO WRB (K164D)-HA-NFT- 26aa	This study	MM596
pCDNA 5/FRT/TO WRB (K164D)-HA-NFT-Venus-truncation-63aa	This study	MM783
pCDNA 5/FRT/TO WRB (K164D)-HA-NFT-Venus-truncation-103aa	This study	MM785
pCDNA/rat-BiP	The Hegde Lab	MM810
pCDNA/rat-BiP-T37G	This study	MM811
pCDNA/rat-BiP-G227D	This study	MM812
pCDNA 5/FRT/TO Sec63-FLAG	This study	MM793
pCDNA 5/FRT/TO Sec63-FLAG-J-mutant (H132A/P133A/D134A)	This study	MM794
pCDNA 5/FRT/TO Sec63-FLAG-Δ230-300	This study	MM805
pCDNA 5/FRT/TO AGAL-FLAG	This study	MM1017
pCDNA 5/FRT/TO Prl-ss-AGAL-FLAG	This study	MM1037
pCDNA 5/FRT/TO CD44-ss-AGAL-FLAG-swap	This study	MM1107
pCDNA 5/FRT/TO CD44-FLAG	This study	MM1015
pCDNA 5/FRT/TO Prl-ss-CD44-FLAG	This study	MM1036
pCDNA 5/FRT/TO prolactin-FLAG	This study	MM1111
pCDNA 5/FRT/TO AGAL-ss-prolactin-FLAG	This study	MM1112
pCDNA 5/FRT/TO ENaC α-HA	This study	MM800
pCDNA 5/FRT/TO β ENaC α-HA	This study	MM1032
pCDNA 5/FRT/TO ENaC β-HA	This study	MM801
pCDNA 5/FRT/TO ENaC γ-HA	This study	MM802

pSpCas9(BB)-2A-Puro	The Zhang Lab	Addgene # 62988 MM286
pOG44	Invitrogen	MM449
pCDNA3.1 BirA-FLAG	The Hegde Lab	MM376
LAMP1-CFP expressing plasmid	The De Camilli Lab	
Software and algorithms		
ΔG prediction server v1.0	http://dgpred.cbr.su.se/index.php?p=Tmpred	ΔG prediction server v1.0
Uniprot	https://www.uniprot.org/	
Prism 7		
GRAVY	http://www.gravy-calculator.de/index.php	Hydrophobicity score
Image J		
Photoshop and Illustrator		
MacVector		
Scaffold		

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Malaiyalam Mariappan (malaiyalam.mariappan@yale.edu).

METHOD DETAILS

DNA constructs

Most of the cDNAs used in this study were cloned into pcDNA/FRT/TO vector (Invitrogen). Mouse Sec63-FLAG (template mouse Sec63 plasmid was a kind gift from Dr. Stefan Somlo, Yale School of Medicine), Human AGAL-FLAG, Human CD44-FLAG, Rat ENaC α-HA, Rat ENaC β-HA, Rat ENaC γ-HA (template ENaCs were a gift from Dr. Cecilia Canessa, Yale School of Medicine), Bovine prolactin-FLAG, and BirA-FLAG were cloned into pcDNA/FRT/TO vector using standard molecular biology methods. pcDNA/FRT/TO containing WRB-HA-Venus substrates were previously described (Sun and Mariappan, 2019). WRB-HA-Venus containing C-terminus AviTag (AGLNDIFEAQKIEWHA) was constructed using an overlap PCR method. pcDNA/FRT/TO containing Prl-ss-AGAL-FLAG (ss: signal sequence), CD44-ss-AGAL-FLAG were constructed by replacing the AGAL signal sequence (1-31) with Bovine-prolactin signal sequence (1-30) or CD44 signal sequence (1-20) using phosphorylated primers and Phusion Site-Directed Mutagenesis procedure. pcDNA/FRT/TO containing AGAL-ss-prolactin-TEV-FLAG was cloned by replacing Rat-prolactin signal sequence (1-30) with AGAL signal sequence

(1-31) using the above procedure. Sec63-FLAG (Δ230-300) was also created using phosphorylated primers. Sec63-FLAG with J domain mutation (H132A/P133A/D134A) and Rat-BiP mutant plasmids, G227D and T37G, were created using Pfu polymerase (Agilent Technologies) based site directed mutagenesis method. The coding sequences of all constructs were verified by sequencing (Yale Keck DNA Sequencing).

Cell lines and CRISPR/Cas9 genome editing

HEK293-Flp-In T-Rex cells (Invitrogen, HeLa-Flp-In T-Rex cells and U2OS cells (a kind gift from Dr. Christian Schlieker, Yale University) were cultured in high glucose DMEM containing 10% FBS at 5% CO_2. The knockout cell lines of Sec63 (Sun and Mariappan, 2019) or Sec62 were previously described (Li et al., 2020). The following Sec61α gRNA sequence: 5' AAAGCGAGGTTGGCAGCATGGGG 3' was cloned into the gRNA expression vector(Mali et al., 2013). The single-strand DNA oligonucleotide homology-directed repair (HDR) sequence (CATCTACCAGTACTTTGAGATCTTCGTTAAGGAGCAAAGCGAGGTTGGCAGCTCTGCCTG GAGCCACCCCCAGTTCGAGAAGGGCGGCGGCAGCGGCGGCGGCAGCGGCGGCGGCAG CTGGAGCCACCCCCAGTTCGAGAAGGCCTCTATGGGGGCCCTGCTCTTCTGAGCCCGTCT CCCGGACAGGTTGAGGAAGCTGC) was synthesized (IDT) for inserting 2X Strep-tag into the C-terminus of Sec61α encoding gene in HEK293 cells. 300 pmol of HDR oligonucleotide was electroporated into one million HEK293 cells along with 2.5 µg each of pSpCas9(BB)-2A-Puro (Ran et al., 2013) and Sec61α gRNA plasmid (Amaxa kit R, program A-24; Lonza). Immediately after electroporation, cells were plated in a 6 well plate. After 24 h of electroporation, the expression of Cas9 was selected by puromycin treatment (2.5µg/ml) for 48 h. Cells were then returned to non-selecting media and grown for 2 days. Cells were replated at 0.5 cell/well in 96 well plates and expanded for 3 weeks. Individual clones were examined for the presence of 2X Strep-tagged Sec61α by immunoblotting with anti-Sec61α antibodies. Positive clones containing 2X Strep-tagged Sec61α exhibited a slower migrating band compared to Sec61α present in control HEK293 cells. To establish HeLa cells stably expressing AGAL-FLAG or Prl-ss-AGAL-FLAG, HeLa-Flp-In T-Rex cells (0.6 million/well) were plated in a 6 well plate and transfected with 1.6µg of AGAL plasmid and 0.4µg of pOG44. The transfected cells were selected with hygromycin (150µg /ml) and blasticidin (10µg /ml) for about three weeks. The individual colonies were picked, and the expression of recombinant AGAL variants was analyzed by immunoblotting with an anti-FLAG antibody.

Immunoprecipitations

To test the interaction between WRB-Venus substrates and the Sec61 translocon complex, the chromosomally Sec61α-2x Strep-tagged HEK293 cells were plated on a poly-lysine (0.15mg/ml) coated 6 well plates at 0.6 million/well. The cells were transiently transfected with 2µg WRB-Venus substrates using 5µL of Lipofectamine 2000 and treated with 200 ng/ml doxycycline to induce protein expression. After 24 hours of transfection, cells were harvested in 1XPBS and centrifuged for 2 min at 10,000g. The cell pellet was lysed in 200µl of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc, 2% digitonin, and 1X Roche EDTA-free protease inhibitor cocktail) by incubating on ice for 30 min. The 5% digitonin (Millipore) stock was boiled for 5 min just before adding into the lysis buffer to avoid digitonin precipitation during IPs. The supernatants were collected by centrifugation at 15,000g for 15 min at 4°C. For co-

immunoprecipitation, the supernatant was rotated with StrepTactin XT beads (IBA, Germany) for 2 h at 4°C. The beads were washed 3 times with 1 ml of wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc, 0.1% digitonin). The bound material was eluted from the beads by directly boiling in 50 µl of 2x SDS sample buffer for 5 min and analyzed by immunoblotting.

To test the interaction between Sec63-FLAG and Sec61 translocon, similar digitonin IP was followed but using HEK293 cells stably expressing variants of Sec63-FLAG and anti-FLAG beads. In this experiment, 20ng/mL doxycycline was used to induce the protein expression for 24 h before harvesting for IPs. To capture BiP interaction with WRB-Venus substrates, AGAL-FLAG, or prolactin-FLAG, HEK293 cells were plated on a 6 well plate at 0.6 million/well and transiently transfected with 2 µg of indicated plasmids. Cells were then washed and harvested in 1XPBS and centrifuged for 2 min at 10,000g. The cell pellets were lysed in NP40 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% sodium deoxycholate, and 0.5% NP40) supplemented with apyrase (10 U/ml) and 10 mM $CaCl_2$ for 30 min on ice. The cell lysates were centrifuged at 15,000g for 15 min. The supernatant was incubated with anti-GFP antibodies for 1.5 h and subsequently incubated with Protein A beads for 1 hour at 4°C. The beads were washed three times with 1 ml of NP40 buffer and eluted by directly boiling in 50 µl of 2x SDS sample buffer for 5 min and analyzed by immunoblotting.

Endo H treatment

Indicated cell lines (0.15 million cells/well) were plated on poly-lysine (0.15mg/mL) coated 24-well plates and transiently transfected with 400ng indicated plasmids. Protein expression was induced with 200ng/mL doxycycline. After 24 h transfection and induction, cells were collected with 150µL 2X SDS sample buffer for immunoblotting. For Endo H treatment, the cells were collected in 150 µL SDS buffer (0.1M Tris pH8.0, 1% SDS buffer) and diluted with 60µL of 2% BME, 0.1M Tris (pH 6.8) and boiled at 100°C for 5 min. The SDS buffer dissolved samples were then divided into two aliquots (60µL for each), and were treated with or without 2µL (1000 units) of Endo H (NEB) in a 100µL reaction volume including 1x G5 buffer (NEB) at 37°C for 4 h. The reaction was terminated by adding 50µL 5X SDS sample buffer and boiled for 5 min before analyzing by immunoblotting with indicated antibodies.

Metabolic labeling

Indicated cells (0.6 million cells/well) were plated on poly-lysine coated (0.15mg/mL) 6-well plates and transiently transfected with 2µg indicated plasmids. The expression of indicated proteins was induced with doxycycline (200ng/mL) for 24 hours. The cells were starved with Met/Cys-free medium including 10% dialyzed FBS and doxycycline for 30 min. Cells were then labeled with 80µCi/mL Express-^{35}S protein labeling Mix (NEG772007MC) for 30 min. The labeled cells were directly harvested in 1ml of RIPA buffer including 1X EDTA-free protease inhibitor cocktail (Roche). The lysate was cleared by centrifugation at 20,000g for 15 min. The supernatant was mixed with anti-FLAG beads or anti-GFP antibodies for 1.5 h with rotation in the cold room. After incubation, protein-A agarose was added to the samples containing anti-GFP antibodies and incubated for another 1 hour. The beads were washed three times with 1 ml of RIPA buffer, and proteins were eluted with 2X SDS sample buffer and run on 7.5% or 10% Tris-Tricine SDS-PAGE gels. The gels were dried and analyzed by autoradiography. The above Endo H treatment protocol was used for metabolically labeled samples.

For isolating translocon-clogged endogenous substrates, Sec63-/- and Sec63-/- cells complemented with WT Sec63, Sec63 J-mutant (HPD/AAA) or the translocon interaction defective Sec63 mutant (Δ230-300) were labeled as described above. The labeled cells were washed and harvested 1XPBS and centrifuged at 10,000g for 2 min. The cell pellets were lysed in 200μL of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc, 2% digitonin, and 1X Roche protease inhibitor cocktail) by incubating on ice for 30 min. The supernatant was collected by centrifugation at 15,000g for 15 min and incubated with anti-FLAG beads (BioLegend) for 2 h at 4°C. The beads were washed 3 times with 1 ml of wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc, 0.1% digitonin). The bound material was eluted from beads by directly boiling in 50 μl of 2x SDS sample buffer for 5 min and analyzed by SDS-PAGE and autoradiography.

For pulse-chase analysis, transfected cells were labeled as above for 30 min and chased in complete DMEM medium supplemented with label-free 2mM Methionine and 2mM Cysteine for indicated time points. The cells were directly harvested in a RIPA buffer with protease inhibitor cocktail and cleared by centrifugation at 15,000g for 15 min. The supernatant was incubated with indicated antibody beads for 2 h. The immunoprecipitated proteins were then eluted with the 2X SDS sample buffer and analyzed by SDS-PAGE/autoradiography. For checking the biotinylation status of WRB-Venus-Avi-Tag variants, cells were co-transfected with 1 μg WRB-Venus-Avi-Tag plasmid and 1 μg BirA-FLAG plasmid. After 24 h of transfection, cells were labeled and chased for 1 h. Note that the residual biotin present in the cell culture medium is sufficient for biotinylation in cells. The cells were harvested in a RIPA buffer and centrifuged at 15,000g for 15 min. The supernatant was either incubated with Avidin-agarose (Thermo Scientific) or anti-GFP antibodies for 1.5 h. Anti-GFP antibodies were recovered by incubating with Protein A beads for 1 h. The beads were washed, eluted, and analyzed as described above. For checking the secretion of AGAL or prolactin variants, the cells were transfected, labeled, and chased as above, but the medium was also collected at each time point and supplemented with salt and detergent to adjust the composition to match the RIPA buffer. The samples were immunoprecipitated with anti-FLAG beads and analyzed as above.

Affinity purification of translocon clogged substrates and mass spectrometry

Three different HEK293 cell lines (Sec63-/-, Sec63-/- complemented with WT Sec63-FLAG, Sec63-/- complemented with Sec63-FLAG J-mutant (HPD/AAA) were plated in 10cm plates at 10 million cells/plate. Protein expression was induced by 20ng/mL doxycycline. After 24 h of induction, the cells were transferred from one 10cm plate to two 15cm plates for each cell line including 20ng/mL doxycycline. After culturing two days, cells at ~100% confluence were treated with 100ug/mL cycloheximide for 5 min. The cells were then collected using ice-cold 1XPBS and centrifuged at 1500 g for 10 min to obtain the cell pellets. The cell pellets were resuspended with lysis buffer (50mM HEPES pH 7.4, 100mM KAC, 50mM NaCl, 10mM MgAC, 1% digitonin, 1X Roche protease inhibitor cocktail, 1mM DTT, 100ug/mL cycloheximide). The lysates were further homogenized using a manual dounce by passing through 10 times. The lysates were then centrifuged at 35,000g for 30 min at 4°C in an MLA80 rotor. The supernatant was incubated with 300μL of pre-washed mouse anti-FLAG beads (Sigma) slurry for 2 hours at 4°C. The beads were washed for 6 times with 5 mL of wash buffer (50mM HEPES pH 7.4, 100mM KAC, 50mM NaCl, 10mM MgAC, 0.1% digitonin, 1mM DTT). The samples were eluted

with the elution buffer (0.1M Glycine pH2.3 and 0.1% Triton X-100). The eluted materials were TCA precipitated and subjected to mass spectrometry analysis.

The immunoprecipitants were digested with trypsin and analyzed by LC-MS/MS using an LTQ Orbitrap Elite equipped with a Waters nanoACQUITY ultra-performance liquid chromatography (UPLC) system at the Mass Spectrometry & Proteomics Resource of the W.M. Keck Foundation Biotechnology Resource Laboratory, Yale University. Proteins were searched against the homo sapiens SwissProt protein database using the Mascot Search Engine (Matrix Science, LLC, Boston, MA; version. 2.6.0). Scaffold program (version Scaffold_4.8.7; Proteome Software, Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Scaffold program was also used to generate normalized spectral counts for each sample. In the supplemental Table 1, we listed protein hits from all three samples that have protein threshold of 5.0% FDR, at least 2 peptides, and peptide threshold of 95%. In the supplemental Table 2, we listed protein hits from Sec63 J-mutant that are at least 2-fold enriched relative to WT Sec63 and 5-fold enriched relative to negative control Sec63-/- cells. The information of signal sequences or the 1st TMD sequences, subcellular localization, and association with human diseases were extracted from UniProt.Org. In Figure S2C, we listed the top 23 clients of Sec63 from the supplemental Table 2 that localize other than the ER in the secretory pathway. The hydrophobic region was defined using the MacVector program by selecting the Kyte-Doolittle hydrophilicity option in windows of 3-19 amino acids. The defined hydrophobic region was analyzed by grand average hydropathy (GRAVY) to obtain a hydrophobicity score.

In vitro transcription and translation

The open reading frames were PCR amplified from pcDNA/FRT/TO based constructs using a forward primer bearing Sp6 promoter sequence and annealing sequence to the CMV promoter and a reverse primer that anneals to the poly-A region of the vector. For generating ribosome-associated nascent chains (RNCs), PCR products were amplified using the above forward primer, but the reverse primer anneals to 3' of the template cDNA without the stop codon. In vitro transcription reactions including PCR product, SP6 polymerase (NEB), and RNasin (Promega) were performed at 40°C for 1.5h as described previously (Sharma et al., 2010). The transcripts were translated in the rabbit reticulocyte lysate (RRL) translation system including ^{35}S-methionine and with or without rough microsomes (RM). Rough microsomes were either prepared from the canine pancreas as described previously (Walter and Blobel, 1983) or cultured HEK293 cells as described previously (Zhang et al., 2013). Transcripts encoding WRB-Venus variants were translated in the presence or absence of HEK or Sec63-/- microsomes for 30 min at 32°C. The translated products (10μL) were denatured with 4X volume of 2XSDS sample buffer and analyzed by SDS-PAGE and autoradiography.

For the co-translational translocation assay, transcripts encoding AGAL-FLAG or Prl-ss-AGAL were translated in RRL in the presence or absence of canine pancreas microsomes for 30 min at 32°C. The digestion was terminated by adding 5 mM PMSF and incubating on ice for 15 min. The translation products were treated or untreated with 0.5mg/mL Proteinase K for 45 min on ice. The resulting samples were transferred into the boiling SDS buffer (100 mM Tris pH 8.0 and 1% SDS) and heated for an additional 5 min. The samples were diluted 10 times with the Triton buffer (50mM Hepes pH7.4, 150mM NaCl, 5mM MgAC, 1% Triton X-100). The diluted samples were incubated with anti-FLAG beads (Bio-legend) for 2 hours at 4°C. The

beads were washed three times with 1ml of Triton buffer and eluted by directly boiling in 50 µl of 2XSDS sample buffer for 5 min and analyzed by SDS-PAGE and autoradiography. For the post-translational translocation assay, transcripts encoding AGAL-FLAG were translated in RRL for 30 min at 32°C. The translation reactions were centrifuged at 70,000 rpm in a TLA120.1 rotor for 30 min at 4°C. The post-ribosomal supernatant was incubated with or without microsomes for 30 min at 32°C. The resulting samples treated with proteinase K and analyzed as above.

RNCs of indicated lengths were produced from transcripts lacking stop codon in the presence or absence of microsomes. The resulting translated products were split into two and incubated with or without 0.5mg/mL Proteinase K for 45 min on ice. The digestion was terminated by adding 5 mM PMSF and incubating on ice for 15 min. The samples were transferred into boiling the 2x SDS sample buffer and incubated for an additional 5 min. For RNCs of AGAL or Prl-ss-AGAL, the translated products were analyzed by autoradiography after denaturation and immunoprecipitation with anti-FLAG beads as described above. In Figure 4D and E, the denatured samples were treated with RNase A for 15 min at 37°C to remove tRNAs from nascent chains and analyzed by SDS-PAGE and autoradiography.

Subcellular fractionation

HEK 293 cells (0.3 x 10^6/well) were plated on a poly-lysine coated 12-well plate and transiently transfected with 800g AGAL-FLAG variant plasmid. Protein expression was induced with 200ng/mL doxycycline. After 24 hours of induction, the cells were harvested with KHM and centrifuged to collect the cell pellet. The cell pellet was resuspended with 300µL of the Triton buffer (50mM Hepes pH7.4, 150mM NaCl, 5mM MgAC, 0.5% Triton X-100) and centrifuged for 15 min at 14,000 rpm at 4°C. The supernatant was taken and 150µL of 5XSDS buffer was added, while the pellet was resuspended well with 300µL of SDS buffer (1% SDS, 0.1M Tris pH 8.0) and then mixed with 150 µL 5XSDS buffer. The resulting samples were analyzed by immunoblotting with Rat anti-FLAG antibody.

α-galactosidase-A activity assay

Cells (0.15 x 10^6/well) were plated on poly-lysine (0.15mg/mL) coated 24-well plates and transiently transfected with 400ng AGAL-FLAG versions. After 48 hours of transfection, the cells were washed once with 1XPBS and harvested in 1XPBS. After centrifugation at 10000 rpm for 2 min, the cell pellet was resuspended in the 100µL of NP40 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% NP40) and incubated on ice for 30min. The lysate was centrifuged for 5 min at 10,000g at 4°C. The assay was performed as described previously with the following modifications (Lukas et al., 2016). 50µL of 2mM artificial substrate 4-MU-α-D-galactopyranoside (Sigma) dissolved in 0.06 M phosphate citrate buffer (pH 4.7) was mixed well with 1µL of the supernatant in a black 96 well plate (Greiner, #655076) for fluorescence reading, and incubated for 5min at 37°C. The reaction was terminated by the addition of 200µL 1 M glycine buffer (pH 10.5). The released 4-MU was measured by a fluorescence reader (TECAN) using an excitation filter, 360nm, and emission filter, 465nm. The readings were subtracted from empty vector-transfected cells, and AGAL variant activities were normalized to wild type AGAL activity.

Immunofluorescence

U2OS cells (0.1 X 10^6) were plated on 12 mm round glass coverslips (Fisher Scientific) coated with 0.15 mg/mL poly-lysine for 1 hour in 24-well plates. The cells were transfected with 400ng AGAL-FLAG variants and 400 ng LAMP-CFP plasmid (LAMP1-CFP plasmid was a generous gift from Dr. Pietro De Camilli, Yale School of Medicine), and the protein expression was induced by adding 200ng/mL doxycycline. After 24 h of transfection, the cells were washed once with 1XPBS and fixed with 3.7% formaldehyde for 10 min and permeabilized with 0.1% Triton X-100 for 5 min. The non-specific binding sites were blocked with Buffer A (1XPBS containing 10% Horse Serum and 0.1% Saponin) for 1 hour. 150µL of rat anti-FLAG (Bio-legend) and rabbit anti-GFP antibodies were added at 1:100 dilutions in Buffer A and incubated for 1 hour. The wells were then washed with Buffer A for 5 times with 5min incubation for each time. 150µL of the secondary antibodies Goat anti-Rat Cy2 and Goat anti-Rabbit Cy3 were added at 1:100 dilutions in Buffer A and incubated for 1 hour before washing five times with Buffer A. Then the coverslips were mounted using Fluoromount G (SouthernBiotech). Cells were imaged on Leica scanning confocal microscopy consisting of an inverted microscope (Leica SP6/SP8), and an HC PL APO 63X (CS2 No: 11506350) oil objective lens (Leica, Wetzlar, Germany). Sequential image scanning at 2x zoom, 100 Hz, and 1024 pixels X 1024 pixels and line averaging of 3 was used for displayed images. To quantify the number of cells with co-localization of AGAL and LAMP1, we first identified the total number of cells of around 50, and then we counted cells with the two-channel signal co-localization. The percentage of cells with co-localization was calculated. Data were graphed using GraphPad Prism and represented with the standard error of the mean from three independent experiments.

Chemical crosslinking

Cells were plated on a poly-lysine coated 6 well plate at 0.6 million cells/well. The cells were transiently transfected with 2µg AGAL-FLAG variants or prolactin-FLAG using 5µL of Lipofectamine 2000 and treated with 200 ng/ml doxycycline to induce the protein expression. 24 hours after transfection, cells were washed once with KHM buffer (20mM Hepes pH 7.4, 110mM KAC, 2mM MgAC), then the cells were incubated with KHM buffer containing the indicated concentration of BMH (Thermo Scientific) for 30 min on ice. The crosslinking reaction was terminated by incubation with 10mM DTT for 15 min on ice. The crosslinked cell samples were then collected in RIPA buffer and the lysate was cleared by centrifugation at 14,000 rpm for 15 min. The supernatant was incubated with anti-FLAG beads for 2 hours at 4°C. The beads were washed 3 times with 1mL of RIPA buffer. The bound material was eluted from the beads by boiling in 50µL of 2XSDS sample buffer for 5 min and analyzed by SDS-PAGE and immunoblotting. In Figure 5D, metabolically labeled cells expressing AGAL or Prl-ss-AGAL were subjected to DSS crosslinking for 30min at room temperature. The reaction was quenched by incubating with 100mM Tris-HCl pH 8.0 for 15 min. The samples were lysed and immunoprecipitated as above and analyzed by SDS-PAGE and autoradiography.

Quantification and statistical analysis

Quantification of bands in immunoblots and autoradiographs were performed using the ImageJ software (NIH). The graphs were created by Prism7, and the error bars reflect the standard error of the mean from three independent experiments.

Supplemental Information

Figure S1 (related Figure 1) Characterization of Sec63-mediated clearance of clogged translocons

(A) A cartoon depicting the predicted topology of WRB fused with the C-terminus Venus tag. The red circle in the C-TMD (3rd TMD) indicates the charged lysine amino acid. The amino acid sequence and free energy prediction for all three TMDs of WRB are indicated. Note that the C-TMD has a high free energy value because it contains the positive charge lysine residue.
(B) The C-TMD of WRB-Venus is mutated to either increase the free energy by introducing charged residues as specified in red color or reduce the free energy by introducing the hydrophobic leucine residue as indicated in blue color. The free energy of all C-TMD variants was predicted as previously described (Hessa et al., 2007).
(C) The indicated substrates were co-transfected with BirA-FLAG, radiolabeled, and chased for the indicated time point. The cell lysates were either immunoprecipitated with anti-GFP antibodies for substrates or used for the Avidin pull down to selectively isolate the biotinylated substrates. The samples were eluted from beads with SDS-sample buffer and analyzed by autoradiography. The diagram on the right illustrates that the substrate with marginal hydrophobicity (Subs. D) was poorly biotinylated by the cytosolic BirA since its C-terminal Avi-TAG is translocated into the ER lumen. By contrast, the C-terminal Avi-TAG of the substrate with strong hydrophobicity (Subs. 4L) was efficiently biotinylated by the cytosolic BirA as it was properly inserted into the membrane and localized in the cytosol.
(D) Cells expressing the indicated substrates were radiolabeled, chased for the indicated time points, and analyzed by SDS/autoradiography after immunoprecipitation with anti-GFP antibodies for substrates.
(E) Transcripts encoding the indicated substrates were translated in RRL in the presence of rough microsomes prepared from either Control HEK293 or Sec63-/- cells. The reactions were directly analyzed by SDS-PAGE and autoradiography.
(F) Control HEK293 or Sec63-/- cells were transfected with the marginal hydrophobicity substrate (Subs. D) that contains the indicated length of amino acids after their C-TMDs. The cell lysates were immunoprecipitated and analyzed by immunoblotting with anti-HA antibodies.

Figure S2 (related Figure 2) Sec63 substrates contain marginally hydrophobic targeting sequences.

(A) Control HEK293 or Sec63-/- cells stably expressing the indicated constructs were lysed, immunoprecipitated with anti-FLAG beads and analyzed by immunoblotting for indicated antigens.
(B) The indicated cell lines were transfected with C-TMDs of substrates carrying charged amino acid(s) of either D, or RRK. The transfected cells were directly analyzed by immunoblotting with anti-GFP antibodies for substrates.
(C) A partial list of Sec63 clients extracted from Table S2 that are localized in the post-ER

compartments. The signal sequence of bovine prolactin or human BiP serves as a strong signal sequence (Kang et al., 2006). Blue color shade denotes the hydrophobic region (h-region) of the signal sequence. Hydrophilic amino acids and glycine are indicated in red color. Hydrophobicity of h-regions was analyzed by grand average hydropathy (GRAVY). Protein localization, molecular weight, and associated human diseases were extracted from Uniprot.org.

Figure S3 (related to Figure 3). Sec63 assists the translocation of proteins with weak targeting sequences.
(A, B) Amino acid sequences of AGAL and CD44 are annotated to indicate the peptides (yellow) identified by mass spectrometry. The signal sequences were underlined.
(C) A cartoon depicting the topology of ENaC α, ENaC β, or ENaC γ. The table contains the amino acid sequence of the first TMD of ENaC α, ENaC β, or ENaC γ with its predicted free energy value. Note that ENaC α has a high free energy value relative to β and γ subunits.
(D) ENaC α-HA was transfected into either control HEK293, Sec63-/-, or Sec62-/- cells. The transfected cells were directly analyzed by immunoblotting with anti-HA antibodies.
(E) ENaC α-HA or β-ENaC α-HA (1st TMD of α subunit was swapped with the 1st TMD of β subunit) transfected into the indicated cell lines and analyzed by immunoblotting with anti-HA antibodies for substrates.

Figure S4 (related to Figure 6) Efficient secretion of prolactin depends on its signal sequence.
(A) HeLa cells stably expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were assayed for α-galactosidase activity and plotted as arbitrary units (a.u) after subtracting the signal from control HeLa cells. Error bar represents the standard deviation (n=2). Note that Prl-AGAL was significantly more accumulated in cells relative to AGAL due to its misfolding and aggregation.
(B) Control HEK293 cells, Sec63-/- cells, or Sec63-/- cells stably expressing wild type Sec63 and or J-domain mutant were transfected with plasmids expressing prolactin-FLAG or AGAL-ss-prolactin-FLAG. The cell lysates were analyzed by immunoblotting with an anti-FLAG antibody for substrates or Sec63 antibodies.
(C) Cells expressing either prolactin-FLAG or AGAL-ss-prolactin-FLAG were radiolabeled and cells and medium were collected at indicated time point. The samples were then immunoprecipitated with an anti-FLAG antibody and analyzed by SDS-PAGE autoradiography. Star symbol indicates the uncleaved signal sequence of prolactin.

Table S1 (related to Figure 2). Proteins identified from Sec63-/-, WT Sec63, or Sec63 J-domain mutant sample.
The scaffold program was used to generate normalized spectral counts for each sample. The table shows protein hits from all three samples that have protein threshold of 5.0% FDR, at least 2 peptides, and peptide threshold of 95%.

Table S2 (Related to Figure 2) Characterization of signal sequences of Sec63's clients.
The table shows protein hits from Sec63 J-mutant that is at least 2-fold enriched relative to WT Sec63 and 5-fold enriched relative to negative control Sec63-/- cells. The information of signal sequences or the 1st TMD sequences, subcellular localization, and association with human

diseases were extracted from UniProt.Org. The red color labeled proteins were also found in Sec63 substrates identified from Schorr et al, the FEBS Journal, 2020. Note that the known Sec63 interaction proteins, Sec62 and Sec61β, were manually removed from the list.

REFERENCES

Ast, T., Cohen, G., and Schuldiner, M. (2013). A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. Cell *152*, 1134-1145.

Ast, T., Michaelis, S., and Schuldiner, M. (2016). The Protease Ste24 Clears Clogged Translocons. Cell *164*, 103-114.

Aviram, N., and Schuldiner, M. (2017). Targeting and translocation of proteins to the endoplasmic reticulum at a glance. Journal of Cell Science *130*, 4079-4085.

Awad, W., Estrada, I., Shen, Y., and Hendershot, L.M. (2008). BiP mutants that are unable to interact with endoplasmic reticulum DnaJ proteins provide insights into interdomain interactions in BiP. Proc Natl Acad Sci U S A *105*, 1164-1169.

Besse, W., Dong, K., Choi, J., Punia, S., Fedeles, S.V., Choi, M., Gallagher, A.R., Huang, E.B., Gulati, A., Knight, J.*, et al.* (2017). Isolated polycystic liver disease genes define effectors of polycystin-1 function. J Clin Invest *127*, 1772-1785.

Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol *67*, 835-851.

Braakman, I., and Hebert, D.N. (2013). Protein Folding in the Endoplasmic Reticulum. Cold Spring Harbor Perspectives in Biology *5*.

Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. Proc Natl Acad Sci U S A *92*, 9643-9646.

Conti, B.J., Devaraneni, P.K., Yang, Z., David, L.L., and Skach, W.R. (2015). Cotranslational stabilization of Sec62/63 within the ER Sec61 translocon is controlled by distinct substrate-driven translocation events. Mol Cell *58*, 269-283.

Davila, S., Furu, L., Gharavi, A.G., Tian, X., Onoe, T., Qian, Q., Li, A., Cai, Y., Kamath, P.S., King, B.F.*, et al.* (2004). Mutations in SEC63 cause autosomal dominant polycystic liver disease. Nat Genet *36*, 575-577.
De Los Rios, P., and Goloubinoff, P. (2016). Hsp70 chaperones use ATP to remodel native protein oligomers and stable aggregates by entropic pulling. Nat Struct Mol Biol *23*, 766-769.

Deshaies, R.J., Sanders, S.L., Feldheim, D.A., and Schekman, R. (1991). Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. Nature *349*, 806-808.

Doring, K., Ahmed, N., Riemer, T., Suresh, H.G., Vainshtein, Y., Habich, M., Riemer, J., Mayer, M.P., O'Brien, E.P., Kramer, G.*, et al.* (2017a). Profiling Ssb-Nascent Chain Interactions Reveals Principles of Hsp70-Assisted Folding. Cell *170*, 298-311 e220.

Doring, K., Ahmed, N., Riemer, T., Suresh, H.G., Vainshtein, Y., Habich, M., Riemer, J., Mayer, M.P., O'Brien, E.P., Kramer, G.*, et al.* (2017b). Profiling Ssb-Nascent Chain Interactions Reveals Principles of Hsp70-Assisted Folding. Cell *170*, 298-+.

Elofsson, A., and von Heijne, G. (2007). Membrane protein structure: prediction versus reality. Annu Rev Biochem *76*, 125-140.

Feige, M.J., and Hendershot, L.M. (2013). Quality control of integral membrane proteins by assembly-dependent membrane integration. Mol Cell *51*, 297-309.

Fewell, S.W., Travers, K.J., Weissman, J.S., and Brodsky, J.L. (2001). The action of molecular chaperones in the early secretory pathway. Annual Review of Genetics *35*, 149-191.

Flynn, G.C., Pohl, J., Flocco, M.T., and Rothman, J.E. (1991). Peptide-Binding Specificity of the Molecular Chaperone Bip. Nature *353*, 726-730.

Gemmer, M., and Forster, F. (2020). A clearer picture of the ER translocon complex. Journal of Cell Science *133*.

Germain, D.P. (2010). Fabry disease. Orphanet J Rare Dis *5*, 30.

Gidalevitz, T., Stevens, F., and Argon, Y. (2013). Orchestration of secretory protein folding by ER chaperones. Biochimica Et Biophysica Acta-Molecular Cell Research *1833*, 2410-2424.
Hegde, R.S., and Bernstein, H.D. (2006). The surprising complexity of signal sequences. Trends Biochem Sci *31*, 563-571.

Hegde, R.S., and Kang, S.W. (2008). The concept of translocational regulation. J Cell Biol *182*, 225-232.

Hegde, R.S., and Zavodszky, E. (2019). Recognition and Degradation of Mislocalized Proteins in Health and Disease. Cold Spring Harb Perspect Biol *11*.

Helenius, A., Marquardt, T., and Braakman, I. (1992). The endoplasmic reticulum as a protein-folding compartment. Trends Cell Biol *2*, 227-231.

Hendershot, L.M. (2004). The ER function BiP is a master regulator of ER function. Mt Sinai J Med *71*, 289-297.

Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H., and von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature *450*, 1026-1030.

Ioannou, Y.A., Bishop, D.F., and Desnick, R.J. (1992). Overexpression of human alpha-galactosidase A results in its intracellular aggregation, crystallization in lysosomes, and selective secretion. J Cell Biol *119*, 1137-1150.

Irvine, A.G., Wallis, A.K., Sanghera, N., Rowe, M.L., Ruddock, L.W., Howard, M.J., Williamson, R.A., Blindauer, C.A., and Freedman, R.B. (2014). Protein Disulfide-Isomerase Interacts with a Substrate Protein at All Stages along Its Folding Pathway. Plos One *9*.

Itskanov, S., and Park, E. (2019). Structure of the posttranslational Sec protein-translocation channel complex from yeast. Science *363*, 84-87.

Jung, S.J., Kim, J.E., Reithinger, J.H., and Kim, H. (2014). The Sec62-Sec63 translocon facilitates translocation of the C-terminus of membrane proteins. J Cell Sci *127*, 4270-4278.

Jungnickel, B., and Rapoport, T.A. (1995). A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. Cell *82*, 261-270.

Juszkiewicz, S., and Hegde, R.S. (2018). Quality Control of Orphaned Proteins. Mol Cell *71*, 443-457.
Kang, S.W., Rane, N.S., Kim, S.J., Garrison, J.L., Taunton, J., and Hegde, R.S. (2006). Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. Cell *127*, 999-1013.

Kassenbrock, C.K., Garcia, P.D., Walter, P., and Kelly, R.B. (1988). Heavy-chain binding protein recognizes aberrant polypeptides translocated in vitro. Nature *333*, 90-93.

Keenan, R.J., Freymann, D.M., Walter, P., and Stroud, R.M. (1998). Crystal structure of the signal sequence binding subunit of the signal recognition particle. Cell *94*, 181-191.

Kida, Y., Ishihara, Y., Fujita, H., Onishi, Y., and Sakaguchi, M. (2016). Stability and flexibility of marginally hydrophobic-segment stalling at the endoplasmic reticulum translocon. Molecular Biology of the Cell *27*, 930-940.

Kida, Y., and Sakaguchi, M. (2018). Interaction mapping of the Sec61 translocon identifies two Sec61 regions interacting with hydrophobic segments in translocating chains. Journal of Biological Chemistry *293*, 17050-17060.

Kim, S.J., Mitra, D., Salerno, J.R., and Hegde, R.S. (2002). Signal sequences control gating of the protein translocation channel in a substrate-specific manner. Dev Cell *2*, 207-217.

Lakkaraju, A.K., Thankappan, R., Mary, C., Garrison, J.L., Taunton, J., and Strub, K. (2012). Efficient secretion of small proteins in mammalian cells relies on Sec62-dependent posttranslational translocation. Mol Biol Cell *23*, 2712-2722.

Lang, S., Benedix, J., Fedeles, S.V., Schorr, S., Schirra, C., Schauble, N., Jalal, C., Greiner, M., Hassdenteufel, S., Tatzelt, J.*, et al.* (2012). Different effects of Sec61alpha, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. J Cell Sci *125*, 1958-1969.

Lee, S., Shin, Y., Kim, K., Song, Y., Kim, Y., and Kang, S.W. (2020). Protein Translocation Acquires Substrate Selectivity Through ER Stress-Induced Reassembly of Translocon Auxiliary Components. Cells *9*.

Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. Nature *531*, 395-399.

Li, X., Sun, S., Appathurai, S., Sundaram, A., Plumb, R., and Mariappan, M. (2020). The Sec63/BiP complex suppresses higher-order oligomerization and RNase activity of IRE1α during ER stress. bioRxiv, 2020.2004.2003.024356.

Lu, Y., Turnbull, I.R., Bragin, A., Carveth, K., Verkman, A.S., and Skach, W.R. (2000). Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. Molecular Biology of the Cell *11*, 2973-2985.

Lukas, J., Scalia, S., Eichler, S., Pockrandt, A.M., Dehn, N., Cozma, C., Giese, A.K., and Rolfs, A. (2016). Functional and Clinical Consequences of Novel alpha-Galactosidase A Mutations in Fabry Disease. Hum Mutat *37*, 43-51.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. Cell *97*, 553-564.

Meyer, H.A., Grau, H., Kraft, R., Kostka, S., Prehn, S., Kalies, K.U., and Hartmann, E. (2000). Mammalian Sec61 is associated with Sec62 and Sec63. J Biol Chem *275*, 14550-14557.

Misselwitz, B., Staeck, O., and Rapoport, T.A. (1998). J proteins catalytically activate Hsp70 molecules to trap a wide range of peptide sequences. Mol Cell *2*, 593-603.

Molinari, M., and Helenius, A. (2000). Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. Science *288*, 331-333.

Nguyen, D., Stutz, R., Schorr, S., Lang, S., Pfeffer, S., Freeze, H.H., Forster, F., Helms, V., Dudek, J., and Zimmermann, R. (2018). Proteomics reveals signal peptide features determining the client specificity in human TRAP-dependent ER protein import. Nat Commun *9*, 3765.

Noiva, R., Freedman, R.B., and Lennarz, W.J. (1993). Peptide Binding to Protein Disulfide-Isomerase Occurs at a Site Distinct from the Active-Sites. Journal of Biological Chemistry *268*, 19210-19217.
Onishi, Y., Yamagishi, M., Imai, K., Fujita, H., Kida, Y., and Sakaguchi, M. (2013). Stop-and-Move of a Marginally Hydrophobic Segment Translocating across the Endoplasmic Reticulum Membrane. Journal of Molecular Biology *425*, 3205-3216.

Plate, L., Cooley, C.B., Chen, J.J., Paxman, R.J., Gallagher, C.M., Madoux, F., Genereux, J.C., Dobbs, W., Garza, D., Spicer, T.P.*, et al.* (2016). Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation. Elife *5*.

Plumb, R., Zhang, Z.R., Appathurai, S., and Mariappan, M. (2015). A functional link between the co-translational protein translocation pathway and the UPR. Elife *4*.

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. Journal of Biological Chemistry *294*, 2098-2108.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.

Rane, N.S., Yonkovich, J.L., and Hegde, R.S. (2004). Protection from cytosolic prion protein toxicity by modulation of protein translocation. EMBO J *23*, 4550-4559.

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663-669.

Robinson, P.J., Pringle, M.A., Woolhead, C.A., and Bulleid, N.J. (2017). Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation. J Biol Chem *292*, 6978-6986.
Schorr, S., Nguyen, D., Hassdenteufel, S., Nagaraj, N., Cavalie, A., Greiner, M., Weissgerber, P., Loi, M., Paton, A.W., Paton, J.C.*, et al.* (2020).

Identification of signal peptide features for substrate specificity in human Sec62/Sec63-dependent ER protein import. FEBS J.

Shan, S.O., and Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. FEBS Lett *579*, 921-926.

Sharma, A., Mariappan, M., Appathurai, S., and Hegde, R.S. (2010). In vitro dissection of protein translocation into the mammalian endoplasmic reticulum. Methods Mol Biol *619*, 339-363.

Skach, W.R., Shi, L.B., Calayag, M.C., Frigeri, A., Lingappa, V.R., and Verkman, A.S. (1994). Biogenesis and Transmembrane Topology of the Chip28 Water Channel at the Endoplasmic-Reticulum. Journal of Cell Biology *125*, 803-815.

Snapp, E.L., McCaul, N., Quandte, M., Cabartova, Z., Bontjer, I., Kallgren, C., Nilsson, I., Land, A., von Heijne, G., Sanders, R.W.*, et al.* (2017). Structure and topology around the cleavage site regulate post-translational cleavage of the HIV-1 gp160 signal peptide. Elife *6*.

Sun, S., and Mariappan, M. (2019). The C-terminal tail guides assembly and degradation of membrane proteins. bioRxiv, 843326.

Voorhees, R.M., and Hegde, R.S. (2015). Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. Elife *4*.

Voorhees, R.M., and Hegde, R.S. (2016). Structure of the Sec61 channel opened by a signal sequence. Science *351*, 88-91.

Walter, P., and Blobel, G. (1983). Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol *96*, 84-93.

White, S.H., and von Heijne, G. (2008). How translocons select transmembrane helices. Annual Review of Biophysics *37*, 23-42.

Willmund, F., del Alamo, M., Pechmann, S., Chen, T.T., Albanese, V., Dammer, E.B., Peng, J.M., and Frydman, J. (2013). The Cotranslational Function of Ribosome-Associated Hsp70 in Eukaryotic Protein Homeostasis. Cell *152*, 196-209.

Wu, X., Cabanos, C., and Rapoport, T.A. (2019). Structure of the post-translational protein translocation machinery of the ER membrane. Nature *566*, 136-139.

Yamamoto, Y., and Sakisaka, T. (2012). Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. Mol Cell *48*, 387-397.

Young, B.P., Craven, R.A., Reid, P.J., Willer, M., and Stirling, C.J. (2001). Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. EMBO J *20*, 262-271.

Zhang, Z.R., Bonifacino, J.S., and Hegde, R.S. (2013). Deubiquitinases sharpen substrate discrimination during membrane protein degradation from the ER. Cell *154*, 609-622.

Ziska, A., Tatzelt, J., Dudek, J., Paton, A.W., Paton, J.C., Zimmermann, R., and Hassdenteufel, S. (2019). The signal peptide plus a cluster of positive charges in prion protein dictate chaperone-mediated Sec61 channel gating. Biol Open *8*.

Sun et al., Figure 1

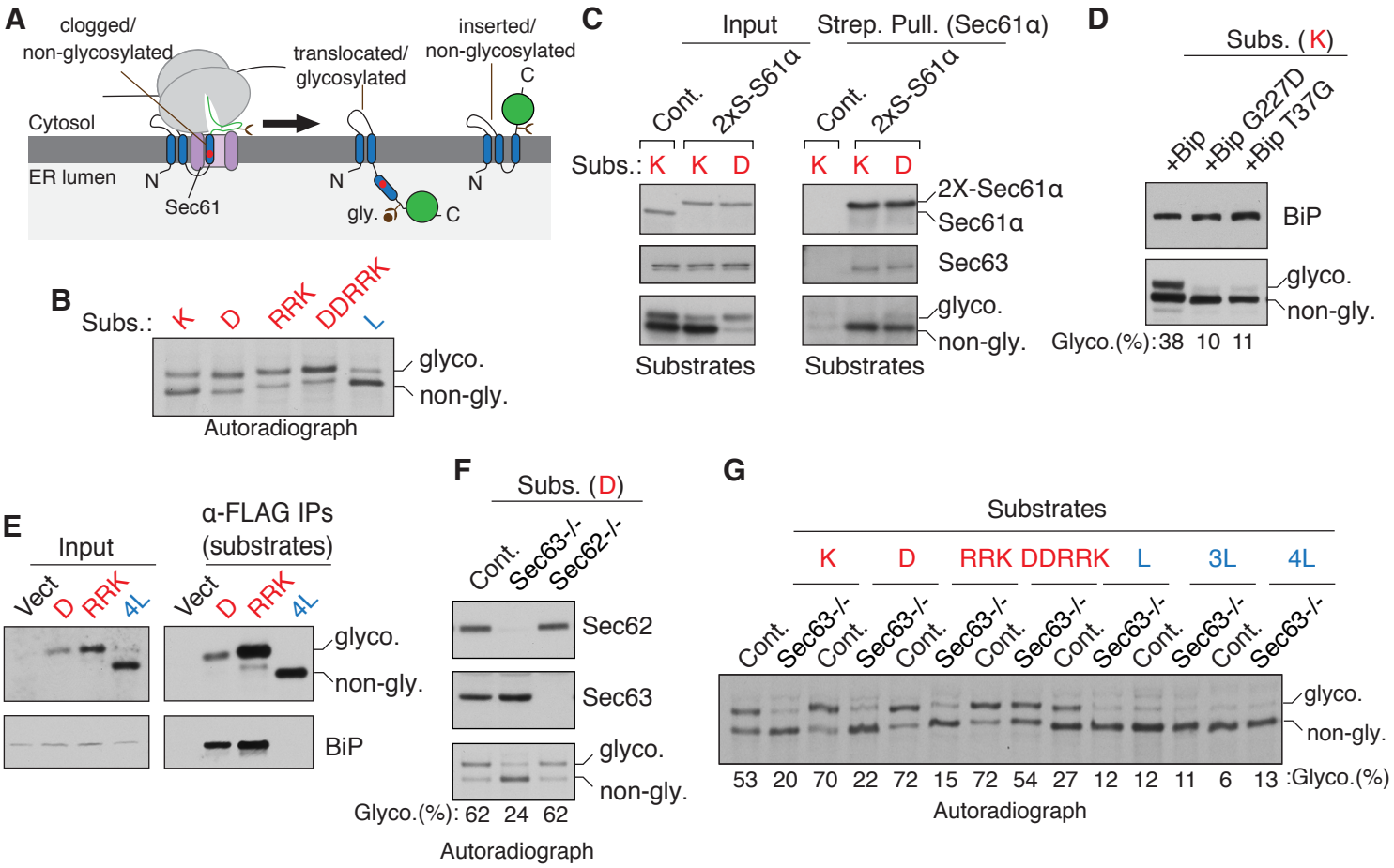


Figure 1. Sec63/BiP is required for clearing translocons clogged with marginally hydrophobic TMDs

(A) A diagram illustrating the translocon clogged with the marginally hydrophobic C-terminal TMD of WRB-Venus. While the cleared substrate is translocated into the ER lumen and glycosylated, the properly inserted substrate is shown as a non-glycosylated form. The red circle in the C-TMD indicates a charged residue.

(B) Cells expressing C-TMDs of substrates bearing the indicated charged residues (shown in red color) or hydrophobic residue (indicated in blue color) were radiolabeled and analyzed by SDS-PAGE and autoradiography after immunoprecipitation of substrates with anti-GFP antibodies.

(C) The indicated substrates were expressed in chromosomally 2X Strep-tagged Sec61α cells, lysed, and pulled down with Strep-Tactin beads and analyzed by immunoblotting for the indicated antigens. The substrates were immunoblotted with anti-GFP antibodies.

(D) Cells co-expressing the indicated versions of BiP and the substrate were analyzed by Immunoblotting with anti-BiP antibody or anti-GFP antibodies for the substrate.

(E) Cells expressing the indicated FLAG-tagged substrates were immunoprecipitated with anti-FLAG antibody and analyzed by immunoblotting with an anti-BiP antibody or anti-GFP antibodies for substrates.

(F) Control HEK293, Sec62-/-, or Sec63-/- cells expressing the indicated substrates were radiolabeled and analyzed as in (B). Immunoblots show the knockout of either Sec62 or Sec63.

(G) Control HEK293 or Sec63-/- cells expressing the indicated substrates were immunoprecipitated with anti-GFP antibodies and analyzed by autoradiography.

Sun et al., Figure 2

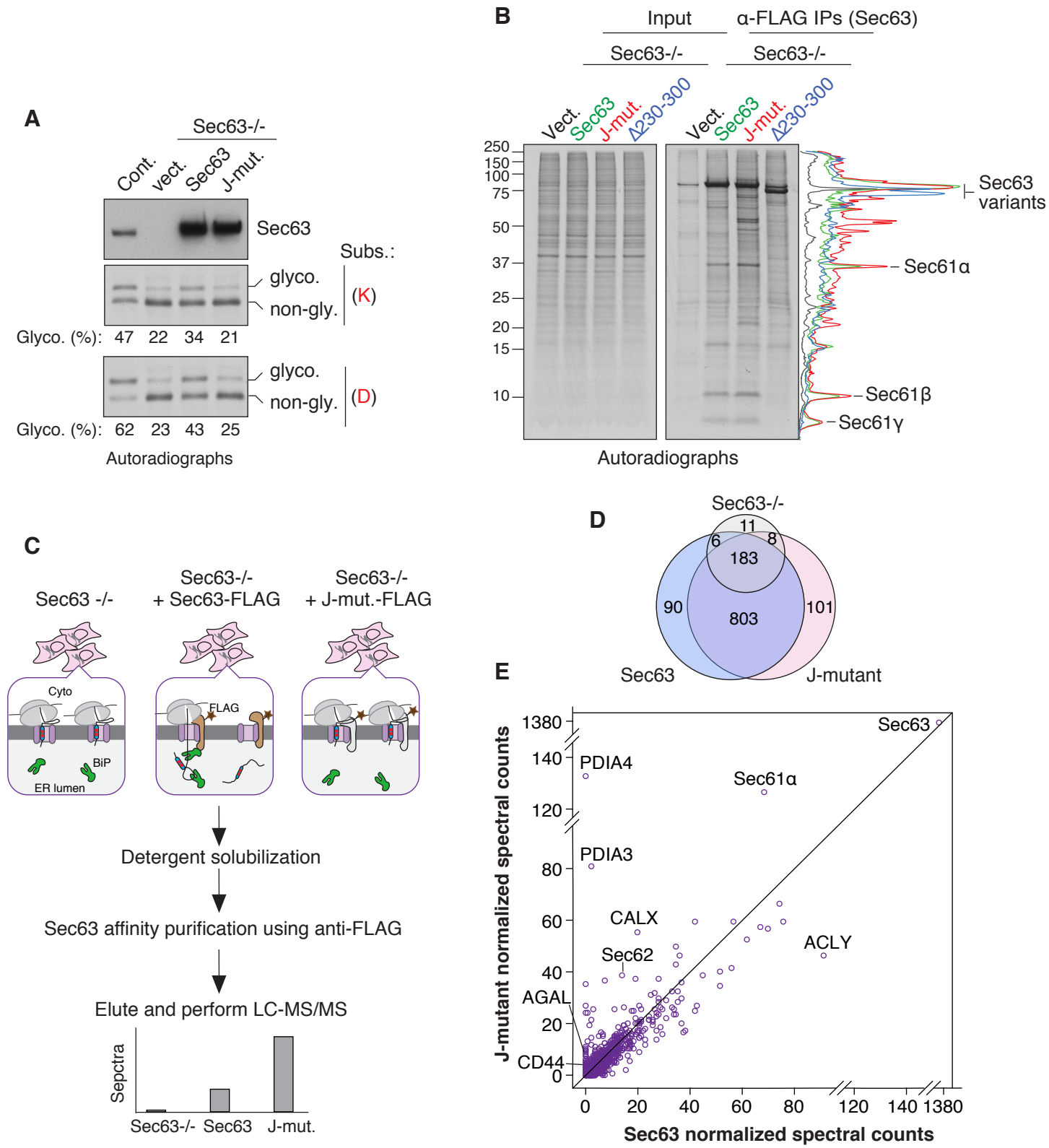


Figure 2. Identification of translocon clogged endogenous substrates of Sec63

(A) Control HEK293 or Sec63-/- cells stably expressing the indicated Sec63 constructs were transfected with the C-TMD of WRB substrate carrying either lysine (K) or aspartic acid (D). The transfected cells were radiolabeled and analyzed by autoradiography after immunoprecipitation with anti-GFP antibodies for substrates. The immunoblot shows the expression of Sec63 from the indicated cell lines.

(B) The indicated Sec63-/- complemented cell lines were radiolabeled and analyzed by autoradiography after immunoprecipitation with anti-FLAG beads. The traces are densitometry profiles of lanes and shown at the right side of the autoradiograph, illustrating a selective enrichment of signals with Sec63 J-mutant.

(C) A scheme describing affinity purification and identification of translocon clogged substrates by mass spectrometry.

(D) Venn diagram of protein distribution among the three groups. The diagram was built from mass spectrometry data from Table S1.

(E) The scatter plot shows the normalized spectral counts comparing Sec63 and Sec63 J-mutant samples. The plot was created from mass spectrometry data from Table S1 using the Scaffold software.

Sun et al., Figure 3

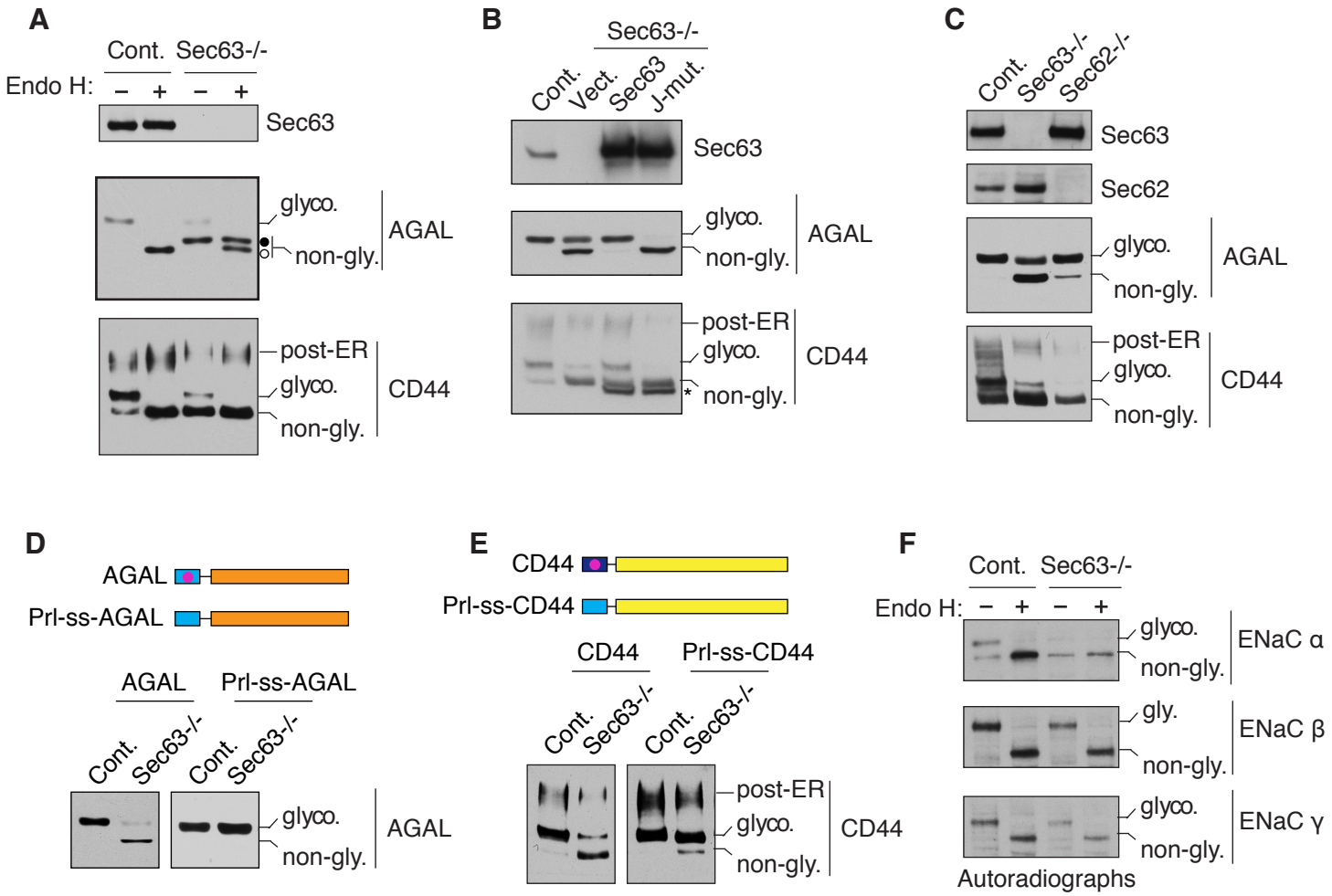


Figure 3. The translocation of proteins with weak targeting sequences depend on Sec63
(A) Control HEK293 or Sec63-/- cells were transfected with AGAL-FLAG or CD44-FLAG. The cell lysates were either untreated or treated with Endo H and analyzed by immunoblotting with an anti-FLAG antibody for substrates and anti-Sec63 antibodies. The open circle indicates the signal sequence cleaved and non-glycosylated form of AGAL. The closed circle indicates the non-glycosylated form of AGAL with the signal sequence.
(B) Control HEK293 or Sec63-/- cells stably expressing the indicated constructs were transfected with AGAL-FLAG or CD44-FLAG. The cell lysates were directly analyzed by immunoblotting as in (A). Star indicates FLAG-tagged Sec63 proteins.
(C) The indicated cell lines were transfected with plasmids expressing AGAL-FLAG or CD44-FLAG, and the cell lysates were directly analyzed as in (A).
(D) Top: A cartoon depicting the signal sequence of AGAL is replaced with the signal sequence of prolactin, termed Prl-ss-AGAL. Bottom: The indicated cell lines were transfected with AGAL-FLAG or Prl-ss-AGAL-FLAG and analyzed by immunoblotting with an anti-FLAG antibody.
(E) Top: A cartoon depicting the signal sequence of CD44 is swapped with the signal sequence of prolactin, termed Prl-ss-CD44. The indicated cells were transfected and analyzed as in (D).
(F) The plasmids expressing ENaC α-HA, ENaC β-HA, or ENaC γ-HA were transfected into control HEK293 or Sec63-/- cells. The cells were radiolabeled and immunoprecipitated with anti-HA antibodies for ENaCs and analyzed by autoradiography.

Sun et al., Figure 4

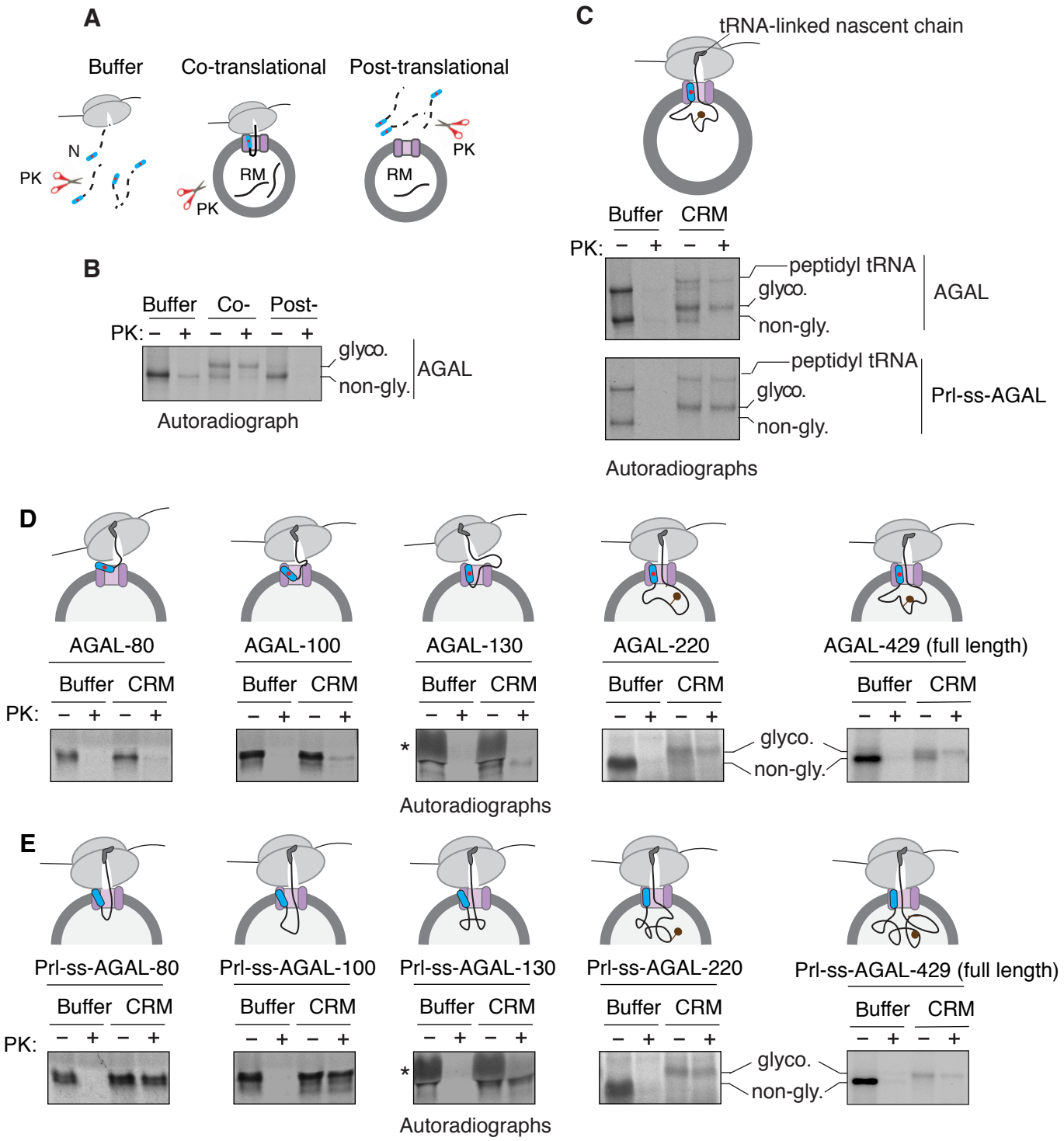


Figure 4. Sec63 substrates are co-translationally translocated into the ER lumen

(A) Scheme depicting the protocol for proteinase K (PK) treatment for the co- or post-translational protein translocation assay.

(B) For the co-translational protein translocation, transcripts encoding AGAL-FLAG were in vitro translated in the rabbit reticulocyte lysate (RRL) based translation system including either buffer or canine rough microsomes (CRM). After translation, samples were digested with PK and analyzed by autoradiography after immunoprecipitation with anti-FLAG antibody beads. For the post-translational translocation, transcripts encoding AGAL-FLAG were translated and centrifuged to remove the ribosomes. The supernatant was incubated with RM and digested with PK and analyzed after immunoprecipitation with anti-FLAG beads.

(C) AGAL or Prl-ss-AGAL transcripts lacking stop codon were in vitro translated in RRL including either buffer or CRM. The reactions were treated with PK and analyzed as in (B). Note that transcripts lacking stop codon produce nascent chains that remain attached to ribosomes via a covalent peptidyl-tRNA bond.

(D). The ribosome-associated nascent chains (RNCs) of the indicated lengths of AGAL were produced in RRL including either buffer or CRM and digested with PK before analyzing by autoradiography. AGAL was analyzed after immunoprecipitation with anti-FLAG antibody beads. The star symbol indicates a distortion of the AGAL-130 nascent chain caused by the co-migration with abundant hemoglobin from RRL.

(E) RNCs of the indicated lengths of Prl-ss-AGAL were produced and analyzed as in (D).

Sun et al., Figure 5

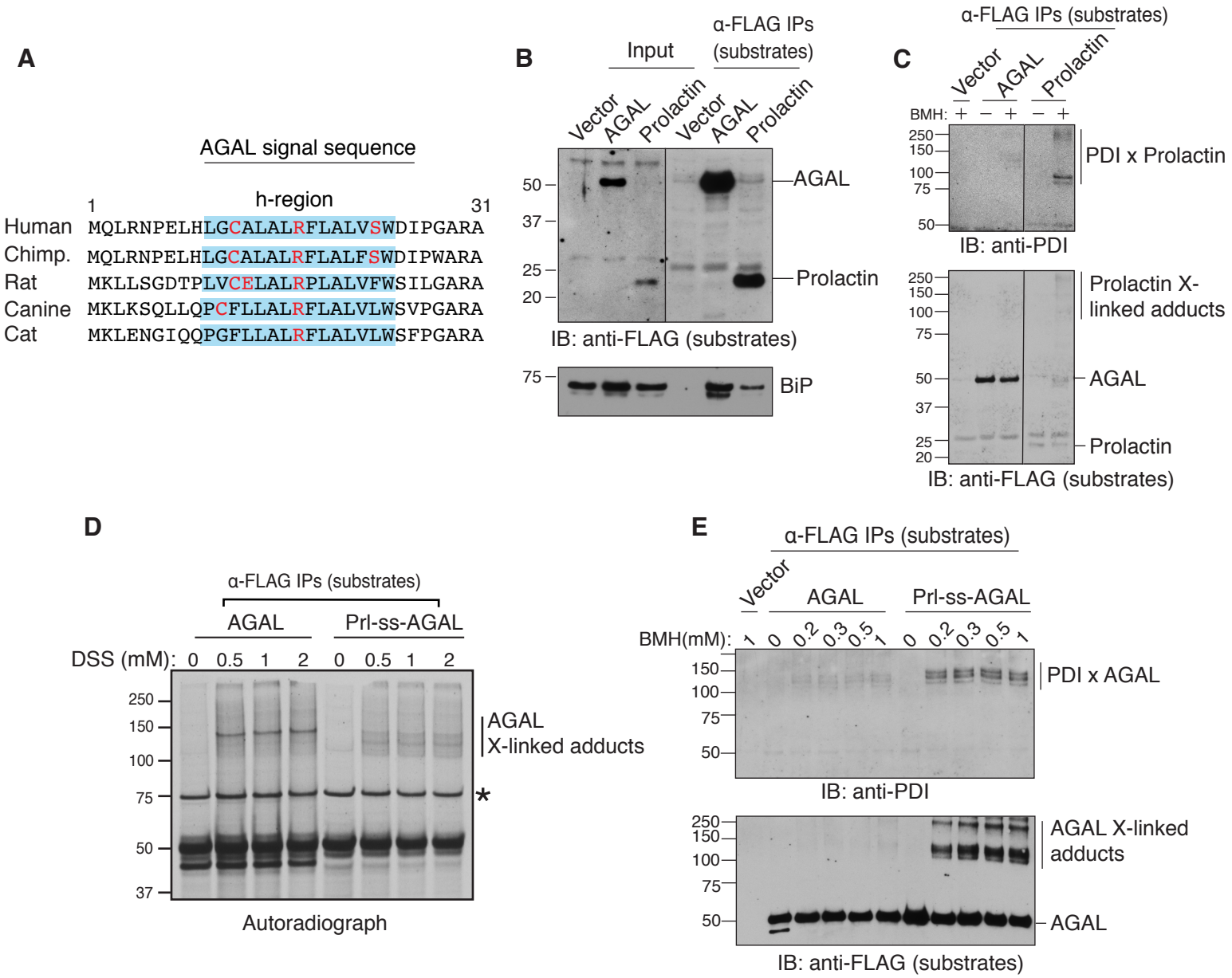


Figure 5. Signal sequence-dependent interaction between translocated proteins and luminal chaperones
(A) Conservation of AGAL signal sequence in different mammalian species. Blue shade indicates the hydrophobic region (h-region) of the signal sequence. Charged and hydrophilic amino acids in the h-region are indicated in red color. Note that the positively charged arginine residue is conserved among different mammalian species.
(B) Cells were transfected with AGAL-FLAG or prolactin-FLAG and immunoprecipitated with anti-FLAG antibody. The immunoprecipitants were analyzed by immunoblotting (IB) with an anti-FLAG antibody for substrates and an anti-BiP antibody.
(C) Cells expressing AGAL-FLAG or prolactin-FLAG were subjected to chemical cross-linking with the cysteine reactive bismaleimidohexane (BMH), immunoprecipitated with anti-FLAG beads, and analyzed by immunoblotting with the indicated antibodies.
(D) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were radiolabeled and subjected to increasing concentration of a lysine reactive crosslinker, DSS (disuccinimidyl suberate). The cross-linked samples were immunoprecipitated with anti-FLAG beads and analyzed by autoradiography. Star denotes a non-specific band.
(E) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were subjected to increasing concentration of BMH crosslinker and analyzed as in (C).

Sun et al., Figure 6

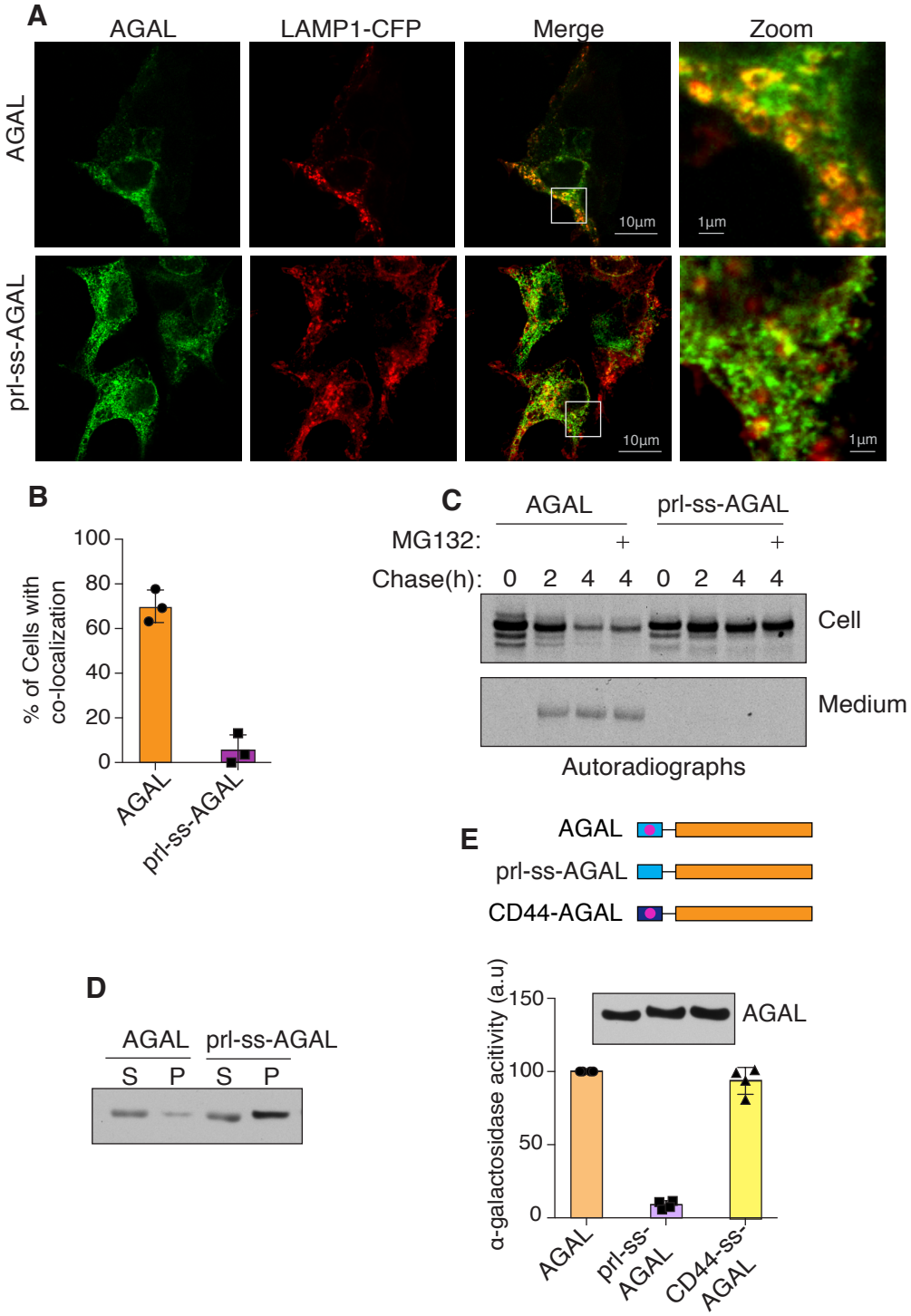


Figure 6. Signal sequence-dependent chaperone recruitment promotes protein folding and maturation
(A) AGAL-FLAG or Prl-ss-AGAL-FLAG was co-transfected with LAMP1-CFP in U2OS cells and processed for immunostaining with anti-FLAG antibody for AGAL and anti-GFP antibodies for LAMP1. Scale bars are 10 m.
(B) The percentage of cells from (A) containing AGAL puncta that colocalize with the lysosomal marker LAMP1-GFP. Error bar represents the standard deviation (n=3).
(C) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were radiolabeled and chased for the indicated time points. Both cells and medium were collected at each time point and analyzed by autoradiography after immunoprecipitation with an anti-FLAG antibody for AGAL.
(D) Cells expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were solubilized with the Triton X-100 buffer. Soluble and insoluble fractions were separated by centrifugation and analyzed by immunoblotting with an anti-FLAG antibody for AGAL.
(E) The upper panel diagram shows the constructs where the AGAL signal sequence was swapped with either the signal sequence of prolactin or CD44. Bottom panel: Cells expressing AGAL-FLAG, Prl-ss-AGAL-FLAG, or CD44-ss-AGAL-FLAG were lysed and assayed for -galactosidase activity. AGAL immunoblot was probed with an anti-FLAG antibody. The activity was presented as arbitrary units (a.u) after subtracting the signal from cells transfected with an empty vector. Error bar represents the standard deviation (n=3).

Sun et al., Figure 7

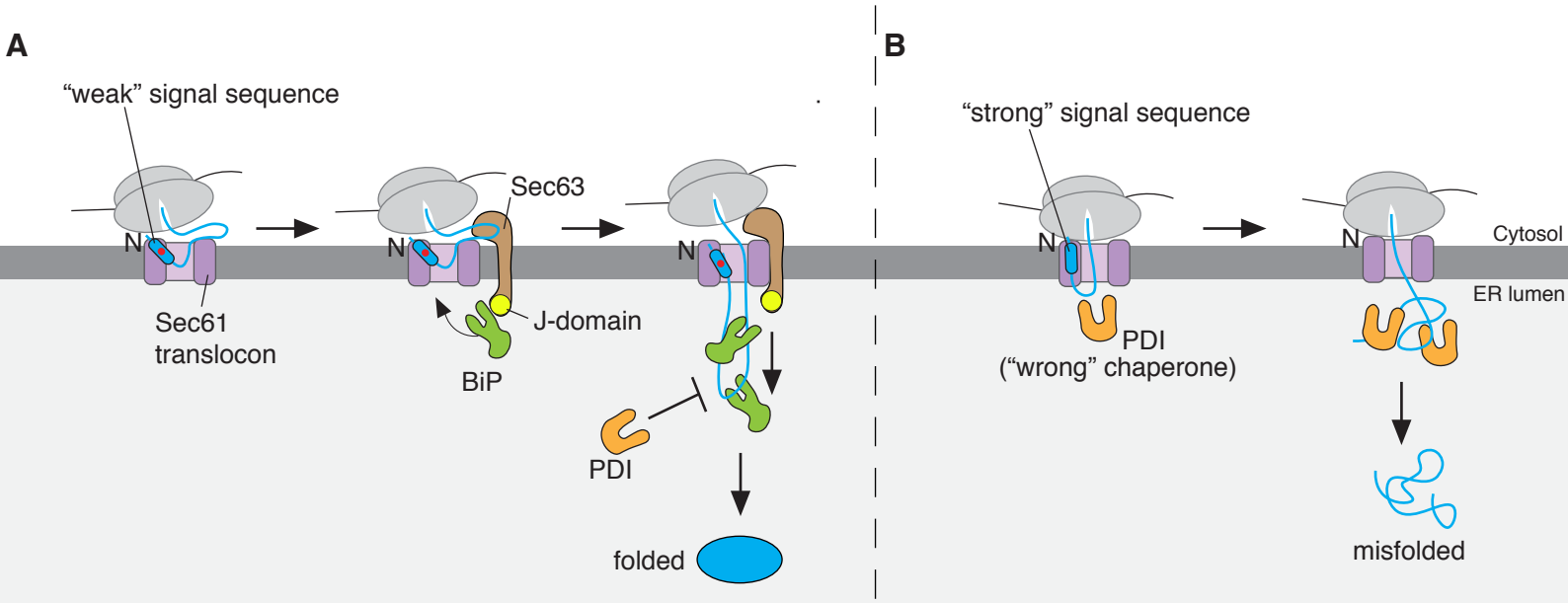


Figure 7. A model for signal sequence-dependent chaperone recruitment and protein folding in the ER.
(A) Nascent chains that contain marginally hydrophobic or weak signal sequences such as AGAL transiently clog the Sec61 translocon. This leads to the privileged recruitment of BiP to the clogged translocon via Sec63 for releasing as well as shielding nascent chains from non-productive interactions, thus promoting folding and maturation.
(B) Swapping a weak signal sequence with a strong signal sequence bypasses the privileged BiP recruitment to the translocating nascent chain, but the translocated nascent chain becomes terminally misfolded and aggregated after binding the wrong chaperone PDI.

Sun et al., Supplemental Figure S1

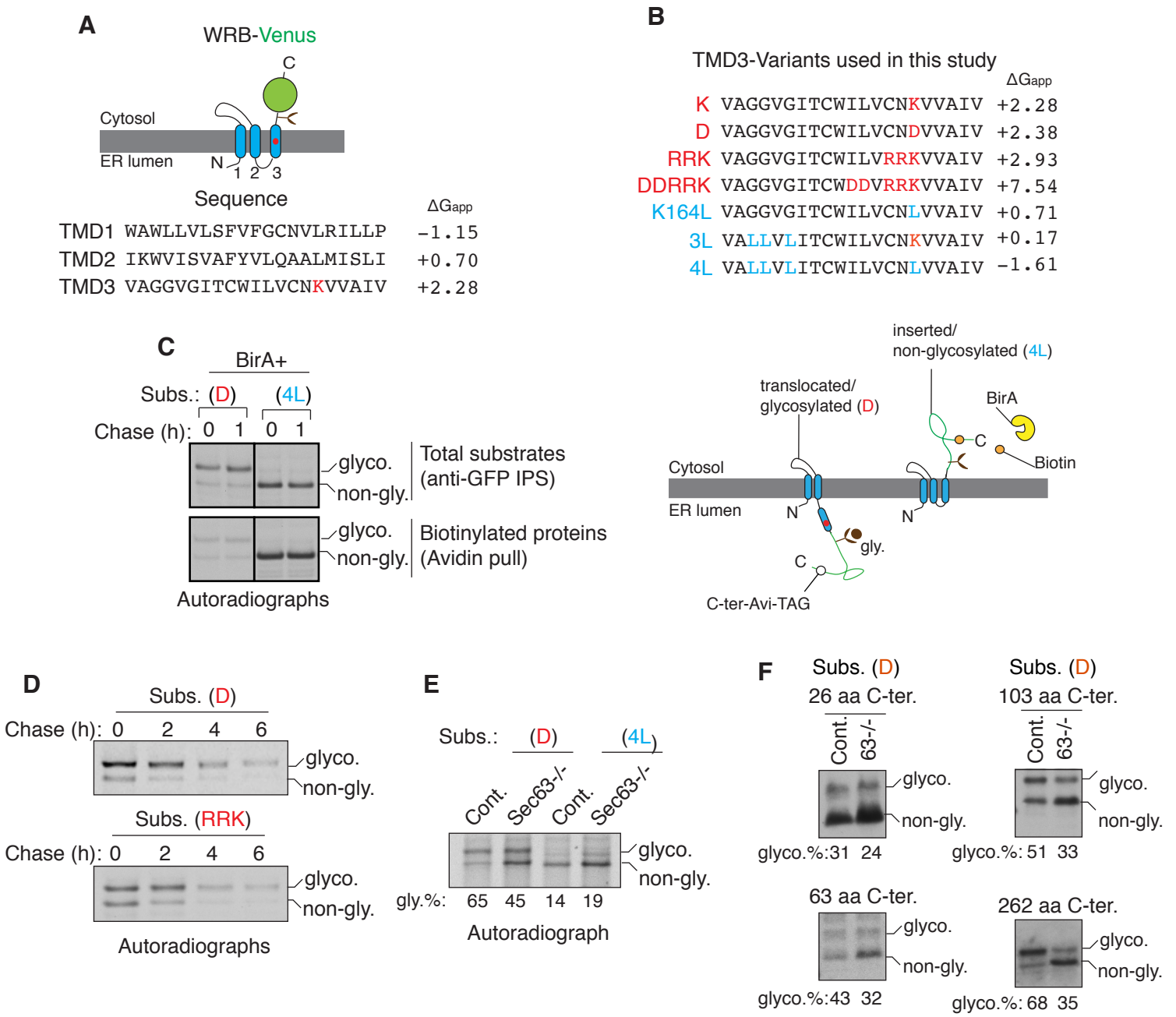


Figure S1 (related Figure 1) Characterization of Sec63-mediated clearance of clogged translocons

(A) A cartoon depicting the predicted topology of WRB fused with the C-terminus Venus tag. The red circle in the C-TMD (3rd TMD) indicates the charged lysine amino acid. The amino acid sequence and free energy prediction for all three TMDs of WRB are indicated. Note that the C-TMD has a high free energy value because it contains the positive charge lysine residue.

(B) The C-TMD of WRB-Venus is mutated to either increase the free energy by introducing charged residues as specified in red color or reduce the free energy by introducing the hydrophobic leucine residue as indicated in blue color. The free energy of all C-TMD variants was predicted as previously described (Hessa et al., 2007).

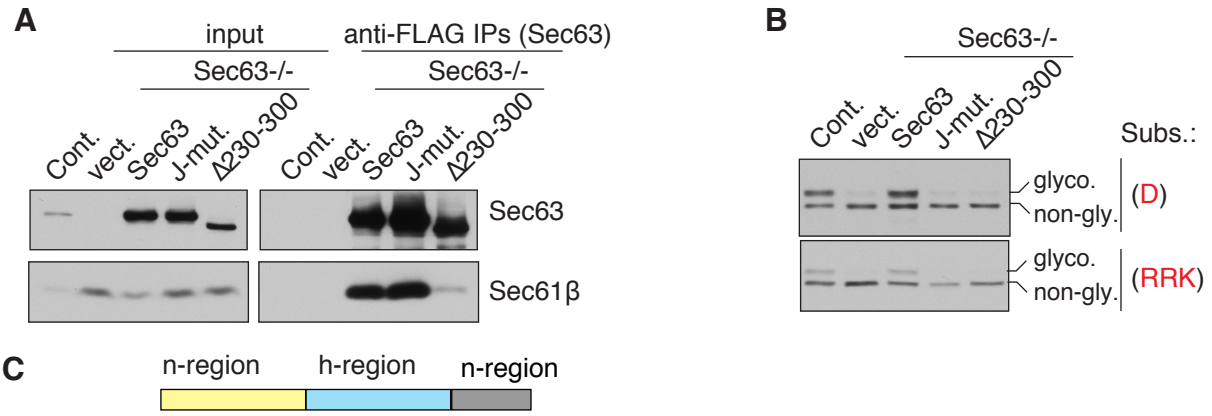
(C) The indicated substrates were co-transfected with BirA-FLAG, radiolabeled, and chased for the indicated time point. The cell lysates were either immunoprecipitated with anti-GFP antibodies for substrates or used for the Avidin pull down to selectively isolate the biotinylated substrates. The samples were eluted from beads with SDS-sample buffer and analyzed by autoradiography. The diagram on the right illustrates that the substrate with marginal hydrophobicity (Subs. D) was poorly biotinylated by the cytosolic BirA since its C-terminal Avi-TAG is translocated into the ER lumen. By contrast, the C-terminal Avi-TAG of the substrate with strong hydrophobicity (Subs. 4L) was efficiently biotinylated by the cytosolic BirA as it was properly inserted into the membrane and localized in the cytosol.

(D) Cells expressing the indicated substrates were radiolabeled, chased for the indicated time points, and analyzed by SDS/autoradiography after immunoprecipitation with anti-GFP antibodies for substrates.

(E) Transcripts encoding the indicated substrates were translated in RRL in the presence of rough microsomes prepared from either Control HEK293 or Sec63-/- cells. The reactions were directly analyzed by SDS-PAGE and autoradiography.

(F) Control HEK293 or Sec63-/- cells were transfected with the marginal hydrophobicity substrate (Subs. D) that contains the indicated length of amino acids after their C-TMDs. The cell lysates were immunoprecipitated and analyzed by immunoblotting with anti-HA antibodies.

Sun et al., Supplemental Figure S2



C

Protein	Signal sequence	Signal sequence length	H-region hydrophobicity	Cellular localization	Molecular weight (KD)
APLP2	MAATGTAAAAATGRLLLLLLVGLTAPALALA	31	1.85	Cell membrane/Nucleus	87
LG3BP	MTPPRLFWVWLLVAGTQG	18	2.22	Secreted	65
GANAB	MAAVAAVAARRRRSWASLVLAFLGVCLG	28	2.61	ER/Golgi	107
CO6A1	MRAARALLPLLLQACWTAA	19	1.74	Secreted	109
KIRR1	MLSLLVWILTLSDTFS	16	2.47	Cell membrane	84
AGAL	MQLRNPELHLGCALALRFLALVSWDIPGARA	31	1.82	Lysosome	49
CD166	MESKGASSCRLLFCLLISATVFRPGLG	27	2.89	Cell membrane	65
PTPRF	MAPEPAPGRTMVPLVPALVMLGLVAGAHG	29	2.19	Membrane	213
LDLR	MGPWGWKLRWTVALLLAAAGT	21	1.89	Cell membrane	95
CD44	MDKFWWHAAWGLCLVPLSLA	20	1.82	Cell membrane	82
CPVL	MVGAMWKVIVSLVLLMPGPCDG	22	2.17	Secreted	54
HLAA	MAVMAPRTLLLLLSGALALTQTWA	24	2.64	Cell membrane	22
EPCAM	MAPPQVLAFGLLLAAATATFAAA	23	2.15	Cell membrane	35
FUCO2	MRPQELPRLAFPLLLLLLLLLPPPPCPA	28	2.52	Secreted	54
CRIM1	MYLVAGDRGLAGCGHLLVSLLGLLLLLARSGTRA	34	2.18	Secreted/Cell membrane	114
IPRI	MAMGLFRVCLVVVTA	15	2.39	Cell membrane	62
SPIT2	MAQLCGLRRSRAFLALLGSLLLSGVLA	27	2.32	Membrane	28
CNTN1	MKMWLLVSHLVIISITTCLA	20	2.01	Cell membrane	113
LEPR	MICQKFCVVLLHWEFIYVITA	21	1.84	Cell membrane	132
SULF2	MGPPSLVLCLLSATVFSLLGGSSA	24	2.20	ER/Golgi/Cell membrane	100
KALM	MVPGVPGAVLTLCLWLAASSGCLA	24	1.70	Cell membrane/Secreted	76
CRTAP	MEPGRRGAAALLALLCVACALRAGRA	26	2.86	Secreted	47
TXD15	MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRG	32	1.68	Cell Membrane	40
Strong signal sequence controls					
PRL	MDSKGSSQKGSRLLLLLVVSNLLLCQGVVS	30	2.44	ER	26
BIP	MKLSLVAAMLLLLSAARA	18	3.19	ER	72

Figure S2 (related Figure 2) Sec63 substrates contain marginally hydrophobic targeting sequences.
(A) Control HEK293 or Sec63-/- cells stably expressing the indicated constructs were lysed, immunoprecipitated with anti-FLAG beads and analyzed by immunoblotting for indicated antigens.
(B) The indicated cell lines were transfected with C-TMDs of substrates carrying charged amino acid(s) of either D, or RRK. The transfected cells were directly analyzed by immunoblotting with anti-GFP antibodies for substrates.
(C) A partial list of Sec63 clients extracted from Table S2 that are localized in the post-ER compartments. The signal sequence of bovine prolactin or human BiP serves as a strong signal sequence (Kang et al., 2006). Blue color shade denotes the hydrophobic region (h-region) of the signal sequence. Hydrophilic amino acids and glycine are indicated in red color. Hydrophobicity of h-regions were analyzed by grand average hydropathy (GRAVY). Protein localization, molecular weight, and associated human diseases were extracted from Uniprot.org.

A

AGAL_HUMAN (100%), 48,768.3 Da
Alpha-galactosidase A OS=Homo sapiens OX=9606 GN=GLA PE=1 SV=1
11 exclusive unique peptides, 14 exclusive unique spectra, 14 total spectra, 152/429 amino acids (35% coverage)

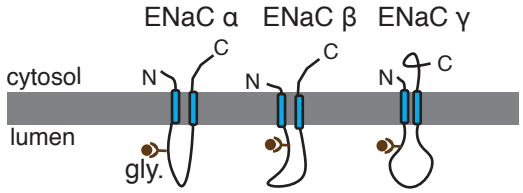
```
MQLRNPELHL GCALALRFLA LVSWDIPGAR ALDNGLARTP TMGWLHWERF
MCNLDCQEEP DSCISEKLFM EMAELMVSEG WKDAGYEYLC IDDCWMAPQR
DSEGRLQADP QRFPHGIRQL ANYVHSKGLK LGIYADVGNK TCAGFPGSFG
YYDIDAQTFA DWGVDLLKFD GCYCDSLENL ADGYKHMSLA LNRTGRSIVY
SCEWPLYMWP FQKPNYTEIR QYCNHWRNFA DIDDSWKSIK SILDWTSFNQ
ERIVDVAGPG GWNDPDMLVI GNFGLSWNQQ VTQMALWAIM AAPLFMSNDL
RHISPQAKAL LQDKDVIAIN QDPLGKQGYQ LRQGDNFEVW ERPLSGLAWA
VAMINRQEIG GPRSYTIAVA SLGKGVACNP ACFITQLLPV KRKLGFYEWT
SRLRSHINPT GTVLLQLENT MQMSLKDLL
```

B

CD44_HUMAN (100%), 81,537.5 Da
CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3
6 exclusive unique peptides, 6 exclusive unique spectra, 6 total spectra, 63/742 amino acids (8% coverage)

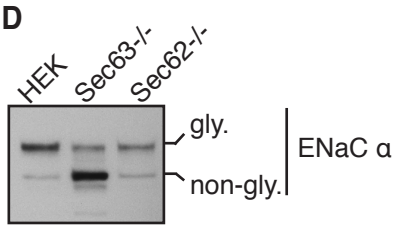
```
MDKFWWHAAW GLCLVPLSLA QIDLNITCRF AGVFHVEKNG RYSISRTEAA
DLCKAFNSTL PTMAQMEKAL SIGFETCRYG FIEGHVVIPR IHPNSICAAN
NTGVYILTSN TSQYDTYCFN ASAPPEEDCT SVTDLPNAFD GPITITIVNR
DGTRYVQKGE YRTNPEDIYP SNPTDDDVSS GSSSERSSTS GGYIFYTFST
VHPIPDEDSP WITDSTDRIP ATTLMSTSAT ATETATKRQE TWDWFSWLFL
PSESKNHLHT TTQMAGTSSN TISAGWEPNE ENEDERDRHL SFSGSGIDDD
EDFISSTIST TPRAFDHTKQ NQDWTQWNPS HSNPEVLLQT TTRMTDVDRN
GTTAYEGNWN PEAHPPLIHH EHHEEEETPH STSTIQATPS STTEETATQK
EQWFGNRWHE GYRQTPKEDS HSTTGTAAAS AHTSHPMQGR TTPSPEDSSW
TDFFNPISHP MGRGHQAGRR MDMDSSHSIT LQPTANPNTG LVEDLDRTGP
LSMTTQQSNS QSFSTSHEGL EEDKDHPTTS TLTSSNRNDV TGGRRDPNHS
EGSTTLLEGY TSHYPHTKES RTFIPVTSAK TGSFGVTAVT VGDSNSNVNR
SLSGDQDTFH PSGGSHTTHG SESDGHSHGS QEGGANTTSG PIRTPQIPEW
LIILASLLAL ALILAVCIAV NSRRRCGQKK KLVINSGNGA VEDRKPSGLN
GEASKSQEMV HLVNKESSET PDQFMTADET RNLQNVDMKI GV
```

C



	TMD1 sequence	ΔG_{app}^{pred}(Kcal·mol)
ENaC α	86 FWAVLWLCTFGMMYWQFGLLF 106	-0.841
ENaC β	51 MWFLLTLLFAALVCWQWGIFI 71	-2.594
ENaC γ	56 WIGFTLTAVALILWQCALLVF 76	-1.398

D



E

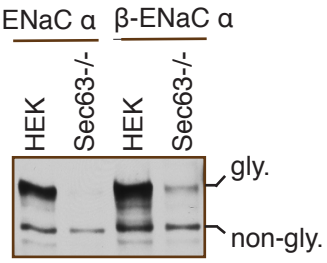


Figure S3 (related to Figure 3). Sec63 assists the translocation of proteins with weak targeting sequences.
(A, B) Amino acid sequences of AGAL and CD44 are annotated to indicate the peptides (yellow) identified by mass spectrometry. The signal sequences were underlined.
(C) A cartoon depicting the topology of ENaC α, ENaC β, or ENaC γ. The table contains the amino acid sequence of the first TMD of ENaC α, ENaC β, or ENaC γ with its predicted free energy value. Note that ENaC α has a high free energy value relative to β and γ subunits.
(D) ENaC α-HA was transfected into either control HEK293, Sec63-/-, or Sec62-/- cells. The transfected cells were directly analyzed by immunoblotting with anti-HA antibodies.
(E) ENaC α-HA or β-ENaC α-HA (1st TMD of α subunit was swapped with the 1st TMD of β subunit) transfected into the indicated cell lines and analyzed by immunoblotting with anti-HA antibodies for substrates.

Sun et al., Supplemental Figure 4

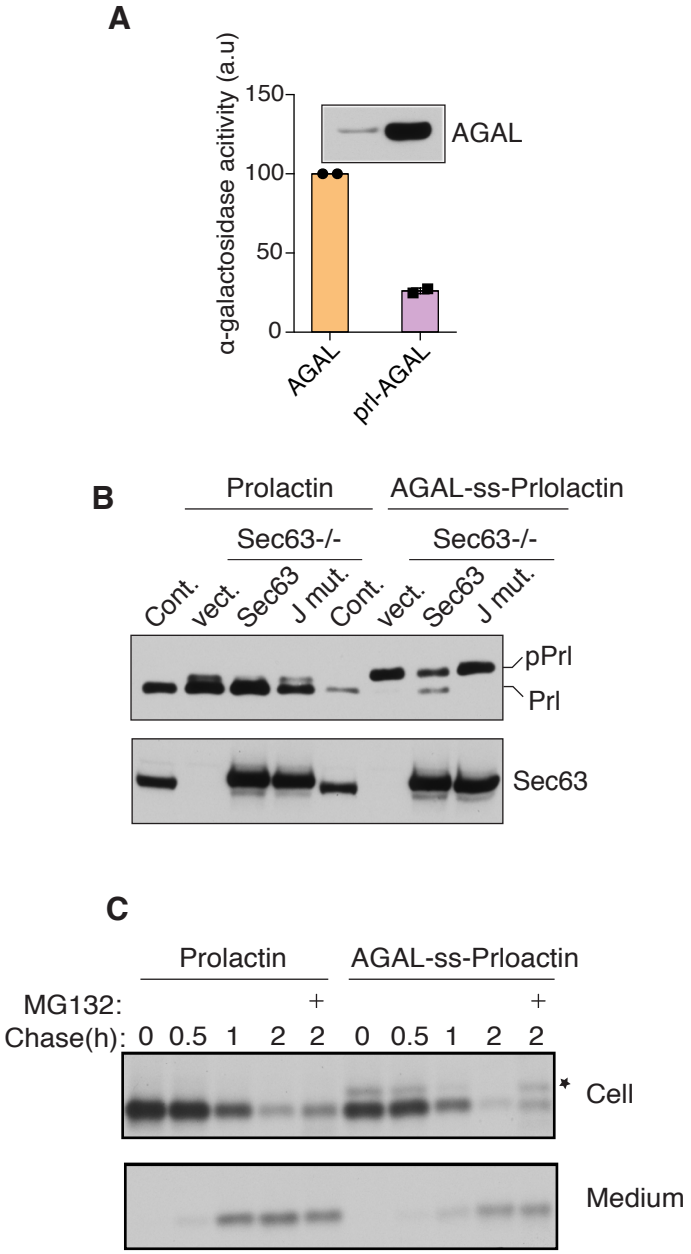


Figure S4 (related to Figure 6) Efficient secretion of prolactin depends on its signal sequence.
(A) HeLa cells stably expressing AGAL-FLAG or Prl-ss-AGAL-FLAG were assayed for α-galactosidase activity and plotted as arbitrary units (a.u) after subtracting the signal from control HeLa cells. Error bar represents the standard deviation (n=2). Note that Prl-AGAL was significantly more accumulated in cells relative to AGAL due to its misfolding and aggregation.
(B) Control HEK293 cells, Sec63-/- cells, or Sec63-/- cells stably expressing wild type Sec63 and or J-domain mutant were transfected with plasmids expressing prolactin-FLAG or AGAL-ss-prolactin-FLAG. The cell lysates were analyzed by immunoblotting with an anti-FLAG antibody for substrates or Sec63 antibodies.
(C) Cells expressing either prolactin-FLAG or AGAL-ss-prolactin-FLAG were radiolabeled and cells and medium were collected at indicated time point. The samples were then immunoprecipitated with an anti-FLAG antibody and analyzed by SDS-PAGE autoradiography. Star symbol indicates the uncleaved signal sequence of prolactin.