Signal sequences encode information for protein folding in the endoplasmic reticulum

Sha Sun[1, 2], Xia Li[1], and Malaiyalam Mariappan[1]
Department of Cell Biology, Yale School of Medicine, Nanobiology Institute, Yale University West Campus.

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

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

[2]Present address: National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing, China.

Abstract: One-third of newly synthesized proteins in mammals are translocated into the endoplasmic reticulum (ER) through the Sec61 translocon. It remains unclear how protein translocation coordinates with the chaperone availability to promote protein folding in the ER. We find that signal sequences cause a translocation pause at the Sec61 translocon until nascent chains engage with luminal chaperones for efficient translocation and folding in the ER. Using a substrate-trapping proteomic approach, we identify that nascent proteins with marginally hydrophobic signal sequences accumulate on the cytosolic side of the Sec61 translocon, which recruits BiP chaperone through Sec63 to bind onto nascent chains. Surprisingly, BiP binding not only releases translocationally paused nascent chains into the ER lumen but also ensures protein folding. Increasing hydrophobicity of signal sequences bypasses Sec63/BiP-dependent protein translocation but translocated nascent chains misfold and aggregate under conditions of limited BiP availability in the ER. Thus, signal sequence-dependent protein folding explains why signal sequences are diverse and use multiple protein translocation pathways.

Keyword

Endoplasmic reticulum; Signal sequence; Protein translocation; Chaperone; Protein folding; Protein misfolding

Introduction

About 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 either cleavable N-terminal signal sequences (SSs) or non-cleavable first transmembrane domains (TMDs). As SSs emerge from cytosolic ribosomes, they are co-translationally recognized and captured by the signal recognition particle (SRP) (Shan and Walter, 2005). The SRP-bound ribosome-associated nascent chains (RNCs) are then delivered to the ER membrane by an interaction with the SRP receptor (Jiang et al., 2008). The SS is then transferred from the SRP to the heterotrimeric Sec61 translocon complex (Rapoport, 2007). The SS is again recognized, and bound, this time, by the Sec61 translocon, allowing the translocation of the following nascent polypeptide across

68 the ER membrane with concomitant cleavage of a signal peptide or integration of a TMD into
69 the lipid bilayer.
70
71 SSs are diverse and significantly differ in length, hydrophobicity, charge, and flanking
72 mature domain, but all of them contain a core hydrophobic region of at least six non-hydrophilic
73 residues. Earlier studies have shown that SRP can efficiently recognize a broad range of
74 hydrophobic SSs of RNCs in the cytosol and target them to the ER of mammalian cells (Hegde
75 and Kang, 2008; Kim et al., 2002; Voorhees and Hegde, 2015). However, the interaction
76 between the Sec61 translocon and SSs greatly varies. Surprisingly, only a minority of signal
77 sequences, such as the signal sequence from prolactin, tightly engages with the Sec61
78 translocon. This results in a looped orientation in which the N-terminus of the SS is positioned at
79 the cytosolic face of the translocon, and the following the hydrophobic portion of the SS
80 becomes intercalated into the lateral gate of the Sec61 translocon (Jungnickel and Rapoport,
81 1995; Kim et al., 2002; Voorhees and Hegde, 2016a). Hence, these SSs mediate efficient
82 protein translocation through the Sec61 translocon. Surprisingly, most SSs inefficiently interact
83 with the Sec61 translon, leading to a nonlooped orientation that forces the mature domain of a
84 nascent chain extruded into the cytosol (Hegde and Kang, 2008). These signal sequences
85 require accessory factors including TRAM, TRAP, Sec63/Sec63, and luminal chaperones to
86 improve their interaction with the Sec61 translocon and thereby promote the translocation of
87 their mature domains into the ER (Brodsky et al., 1995; Fons et al., 2003; Nguyen et al., 2018;
88 Schorr et al., 2020b; Voigt et al., 1996). If the role of signal sequences is merely to mediate
89 targeting and translocation into the ER, it is difficult to rationalize why signal sequences are
90 evolved to use different accessory factors for their efficient translocation into the ER. One
91 important function of SSs is that protein translocation into the ER can be selectively attenuated
92 in a signal sequence-dependent manner during ER stress to reduce the ER folding burden
93 (Kang et al., 2006). In some cases, the SS can regulate the timing of signal peptide cleavage,
94 thereby enhancing protein maturation (Rehm et al., 2001; Rutkowski et al., 2003; Snapp et al.,
95 2017). However, the factors and mechanisms involved in the SS-dependent protein maturation
96 are less understood.
97
98 Similar to SSs, TMDs of polytopic membrane proteins are recognized by the Sec61
99 translocon and laterally partition into the lipid bilayer using the lateral gate of the translocon.
100 Again, TMD hydrophobicity plays a crucial role in partitioning from the lateral gate of the
101 translocon to the hydrophobic 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, meaning 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 marginally hydrophobic can transiently stall in the interior pore of the Sec61 translocon (Kida et al., 2016; Kida and Sakaguchi, 2018). It is unclear how and which factors assist in clearing the marginally hydrophobic TMD that is transiently retained at the Sec61 transocon during the co-translational membrane protein insertion. In some cases, less hydrophobic TMDs are completely translocated into the ER lumen (Feige and Hendershot, 2013; Ojemalm et al., 2012; Skach et al., 1994; Sun and Mariappan, 2020). They can then either be retrieved by interactions with the membrane embedded TMDs or degraded by ER-associated quality control pathways (Buck et al., 2017).

Upon translocation into the ER lumen, nascent chains are folded with the assistance of molecular chaperones and folding enzymes (Braakman and Hebert, 2013; Fewell et al., 2001; Helenius et al., 1992; Ma and Hendershot, 2004). Among the broad spectrum of chaperones in the ER, the member of the Hsp70 family chaperone BiP ATPase plays a central role in shielding exposed hydrophobic residues in nascent proteins and thereby preventing protein misfolding and aggregation. BiP also assists in protein folding by cycling between high-affinity and low-affinity substrate binding through ATP-dependent conformational changes that are regulated by co-chaperones and nucleotide exchange factors in the ER lumen (Pobre et al., 2019). Considering the essential roles of BiP in preventing protein misfolding and promoting folding, it is less understood how the translocating nascent chains are accurately matched to the availability of active BiP molecules in the ER lumen. Studies of cytosolic proteins have shown that the ribosome-associated complex (RAC) recruits Hsp70 to assist the folding of aggregation-prone nascent polypeptides (Doring et al., 2017; Willmund et al., 2013). It is unclear whether a similar mechanism exists in the ER since the ribosome is separated from the translocating nascent chain by the ER membrane bilayer. This physical barrier suggests that the Sec61 translocon and its associated factors (Gemmer and Forster, 2020) might recruit the luminal the Hsp70 chaperone BiP to the aggregation-prone nascent polypeptides for their proper folding and maturation in the ER.

In this study, we used a series of substrates that vary in TMD hydrophobicity to demonstrate that the Sec61 translocon-associated membrane protein Sec63 is responsible for releasing marginally hydrophobic (or weak) TMDs that are transiently retained at the Sec61

136 translocon by recruiting and activating BiP ATPase. We developed a substrate-trapping
137 proteomic approach and identified endogenous secretory and membrane proteins that are
138 transiently retained at the Sec61 translocon in the absence of Sec63 activity. We demonstrate
139 that the nascent chain with a marginally hydrophobic (or weak) SS is co-translationally recruited
140 to the ER membrane, but the initial translocation of ~160 amino acids is paused and is
141 accumulated on the cytosolic side of the Sec61 translocon. Sec63 releases the translocation
142 pause through recruiting and activating BiP ATPase to bind and facilitate the nascent chain
143 translocation into the ER. Surprisingly, the BiP-mediated co-translational translocation also
144 guarantees subsequent protein folding in the ER. Increasing the hydrophobicity of a weak SS
145 can bypass the Sec63/BiP-dependent translocation, but the translocated nascent chain is at a
146 high risk of misfolding and aggregating under conditions of limited BiP availability in the ER.
147 Thus, our studies suggest that SSs have evolved to recruit appropriate chaperones through
148 modulating translocon-associated factors to meet the unique challenges of folding secretory and
149 membrane proteins in the ER.
150
151 **Results**
152
153 **The Sec63/BiP complex clears translocons transiently retained with marginally**
154 **hydrophobic sequences.**
155 We first investigated how the Sec61 translocon responds to marginally hydrophobic sequences
156 in proteins during their translocation into the ER. Earlier studies have shown such sequences
157 are neither hydrophobic enough to be inserted into the lipid bilayer nor hydrophilic enough to be
158 translocated into the ER lumen, leading to their transient retention at the Sec61 translocon (Kida
159 et al., 2016; Kida and Sakaguchi, 2018). To investigate the mechanism by which the transiently
160 retained nascent chain at the Sec61 translocon is cleared, we used our model C-terminally
161 Venus-tagged substrates derived from WRB, a subunit of the tail-anchored membrane protein
162 insertase (Fig. S1A) (Yamamoto and Sakisaka, 2012). We have recently shown that the C-
163 terminal TMD of WRB-Venus is translocated into the ER lumen because of a positively charged
164 lysine residue (Fig. S1A) (Sun and Mariappan, 2020). The translocation of the C-TMD of WRB
165 into the ER lumen can be further increased by either incorporating more charged residues or
166 reduced by adding more hydrophobic residues (Sun and Mariappan, 2020) (Fig. S1B).
167 Consistent with our recent findings, substrates with marginally hydrophobic C-terminal TMDs (K,
168 D, RRK, and DDRRK), provisionally termed weak TMDs, were poorly inserted into the ER
169 membrane and were translocated into the ER lumen as evidenced by the glycosylation (Fig. 1A

170 and 1B and Fig. S1B). By contrast, the substrate with sufficient hydrophobicity (L) was efficiently
171 inserted into the ER membrane as shown by the mostly non-glycosylated form (Fig. 1A and 1B).
172
173 We hypothesized that weak TMD substrates may be transiently retained at the Sec61
174 translocon before their translocation into the ER lumen. To test this, we precipitated the
175 chromosomally 2xStrep-tagged Sec61α from cells expressing weak TMD substrates and found
176 that the unglycosylated form was highly enriched with the Sec61 translocon (Fig. 1C). However,
177 the glycosylated form was not significantly associated with the Sec61 translocon since it was
178 already translocated into the ER lumen and likely moved away from the translocon. Next, we
179 asked how the transiently retained substrates are cleared from Sec61 translocons. We
180 hypothesized that the abundant luminal chaperone BiP ATPase might promote the translocation
181 of transiently retained substrates from Sec61 translocons. Indeed, the transient expression of
182 BiP mutants (G227D and T37G), which are defective in ATP-induced release of substrates (Wei
183 et al., 1995), in cells inhibited the translocation of weak TMDs into the ER lumen (Fig. 1D).
184
185 The BiP involvement in this process raises the question of how BiP is locally recruited to
186 the Sec61 translocon to release weak TMDs into the ER lumen. The translocon-associated
187 membrane protein Sec63 is known to recruit and activate BiP ATPase in order to facilitate post-
188 translational protein translocation into the ER lumen in yeast (Deshaies et al., 1991; Matlack et
189 al., 1999; Meyer et al., 2000). To test if Sec63 is required in this process, we used knockout
190 cells of either Sec63 or its interacting protein Sec62 generated by CRISPR/Cas9. Although both
191 Sec62 and Sec63 are required for BiP-mediated post-translational translocation into the ER,
192 Sec63 was sufficient to translocate the translocon-retained weak TMD into the ER lumen (Fig.
193 1E). To determine biochemical features of TMDs that depend on Sec63 activity, we tested the
194 translocation of TMDs with varying hydrophobicity in Sec63-/- cells (Fig. 1F and S1B). The
195 translocation of substrates with weak TMDs (K, D, and RRK) was significantly impaired as
196 demonstrated by reduced glycosylated forms in Sec63-/- cells relative to that of wild-type (WT)
197 cells. By contrast, strong TMDs (L, 3L and 4L) were mostly inserted into the membrane as
198 evidenced by the appearance of predominantly non-glycosylated forms in both WT and Sec63-/-
199 cells (Fig. 1F). Of note, the extremely charged TMD (DDRRK) was only partially dependent on
200 Sec63 (Fig. 1F). These results suggest that a weak TMD transiently retains at the Sec61
201 translocon and thereby relies on Sec63 activity for its translocation into the lumen, whereas
202 Sec63 activity is dispensable if the TMD is either sufficiently hydrophobic or it is changed to a
203 nearly hydrophilic sequence. Consistent with our data from cells, the translocation of a charged

204 TMD into the ER lumen depended on Sec63 activity in an in vitro assay performed using
205 microsomes derived from WT or Sec63-/- cells (Fig. S1C). Surprisingly, Sec63 was dispensable
206 for translocation of a weak TMD with a small C-terminal tail containing 26 amino acids (Fig.
207 S1D). This result indicated that the C-terminal tail length after the TMD is crucial for its Sec63-
208 dependent translocation into the ER lumen. Truncation studies further revealed that the
209 translocation of the C-TMD with a tail longer than 100 amino acids was dependent on Sec63
210 activity in cells (Fig. S1D). Taken together, these data suggest that Sec63 is required for
211 clearing the Sec61 translocon transiently retained with weak TMDs containing the C-terminal
212 tails longer than 100 amino acids.
213
214 **A substrate-trapping approach identifies the transiently retained endogenous substrates**
215 **at the Sec61 translocon**
216 To determine the role of Sec63 in clearing the translocon of retained endogenous substrates,
217 we developed a substrate trapping strategy. We predicted that a Sec63 J-domain mutant could
218 trap the Sec61 translocon-retained with nascent chains because the mutant J-domain cannot
219 recruit and activate BiP ATPase. To examine this, we stably complemented either WT Sec63,
220 Sec63 J-mutant (HPD/AAA), or the translocon interaction defective Sec63 mutant (Δ230-300)
221 into Sec63-/- cells (Fig. S2A). Using our model substrates, we first tested whether Sec63 J-
222 domain or its interaction with the translocon is essential for clearing the transiently retained
223 substrates at the Sec61 translocon. WT Sec63 complemented cells cleared translocons
224 retained with weak TMDs by releasing them into the ER lumen as illustrated by the glycosylated
225 form of substrates (Fig. 2A and S2B). By contrast, Sec63 J-mutant and the Sec61 interaction
226 defective mutant (Δ230-300) were not able to clear the translocon-retained substrates as
227 evidenced by the primarily unglycosylated form of substrates (Fig. 2A and S2B). Using these
228 stable cell lines, we next performed metabolic labeling and immunoprecipitation to determine if
229 Sec63 J-mutant could trap translocons-retained with endogenous substrates, and if the trapping
230 depends on its interaction with the translocon. We found that all three subunits of the newly
231 synthesized Sec61 translocon, α, β, and γ were associated with both WT Sec63 and J-mutant
232 but were not detected from either Sec63-/- cells or the Sec61 interaction defective mutant
233 (Δ230-300) expressing cells (Fig. 2B). Quantification of bands further revealed that newly
234 synthesized radiolabeled proteins were significantly enriched with Sec63 J-mutant relative to
235 WT Sec63 or Sec63 (Δ230-300) (Fig. 2B). These findings support our conclusion that both
236 Sec63 J-domain and its interaction with the translocon are important for clearing translocons
237 transiently retained with nascent polypeptides in cells.

238

239 We next performed a large-scale affinity purification of Sec63 as depicted in Fig. 2C and
240 subjected immunoprecipitants to mass spectrometry to identify translocon-associated
241 endogenous clients. Our analysis showed many proteins associated with both WT Sec63 and
242 Sec63 J-mutant relative to control Sec63-/- cells (Fig. 2D and Table S1). Consistent with our
243 prediction, about 117 secretory and membrane proteins were at least two-fold enriched in
244 Sec63 J-mutant relative to WT Sec63 and five-fold enriched than Sec63-/- cells (Fig. 2E and
245 Table S2). Of the 117 proteins, about 65% of clients were non-resident ER proteins, indicating
246 that these clients interacted with Sec63 J-mutant during their retention at the Sec61 translocon
247 in the ER. Interestingly, about one-third of these candidate substrates have been previously
248 implicated in human genetic disorders caused by mutations in corresponding genes (Table S2).
249 Many of our Sec63 clients were also found in the independent proteomic study performed using
250 Sec63 depleted cells (Fig. 2E and Table S2) (Schorr et al., 2020a). We found about 86
251 secretory and membrane proteins were enriched with WT Sec63, but about 60% of them are
252 ER-resident proteins, suggesting that many of these proteins are likely interacting proteins of
253 Sec63 in the ER (Table S3). Thus, our substrate-trapping proteomic approach identified
254 potential endogenous secretory and membrane proteins that are transiently retained at the
255 Sec61 translocon in the absence of Sec63 activity.
256

257 **Proteins containing weak signal sequences depend on Sec63 for their translocation into**
258 **the ER lumen.**

259 To validate our Sec63's substrates, we constructed C-terminally FLAG-tagged expression
260 plasmids for two substrates, AGAL and CD44 (Fig. S3A and S3B). The *AGAL* gene encodes for
261 a lysosomal enzyme called α-galactosidase A, the activity of which is altered by genetic
262 mutations in Fabry disease (Germain, 2010). The *CD44* gene encodes a plasma membrane-
263 localized single spanning membrane protein with an N-terminal SS. We first examined if these
264 substrates depended on Sec63 to translocate into the ER. Immunoblotting of AGAL from
265 transiently expressing cells revealed that AGAL was fully glycosylated (Fig. 3A). AGAL
266 glycosylation was verified by digestion with Endo H, which removes high mannose N-linked
267 glycans from glycoproteins. AGAL glycosylation was markedly reduced in Sec63-/- cells relative
268 to WT cells (Fig. 3A). Surprisingly, Endo H digestion revealed that unglycosylated AGAL in
269 Sec63-/- cells migrated slower than Endo H-deglycosylated AGAL (Fig. 3A and Fig. S3C), which
270 suggests that the unglycosylated AGAL in Sec63-/- cells still contains the uncleaved SS. The
271 unglycosylated AGAL carrying the SS was mislocalized to the cytosol of Sec63-/- cells, as it was

272 accessible to proteinase K (PK) added to the semipermeabilized cells, but the
273 translocated/glycosylated form was mostly protected from PK digestion (Fig. S3D). Expression
274 of CD44 in WT cells showed two glycosylated bands. The faster migrating core glycosylated
275 band was sensitive to Endo H but the slower migrating complex type of N-glycan was less
276 sensitive to Endo H, indicating that it is localized to post-ER compartments (Fig. 3A). Similar to
277 AGAL, the translocation of CD44 was significantly reduced as it displayed mostly the
278 unglycosylated form in Sec63-/- cells. We then asked if the translocation defects associated with
279 AGAL and CD44 could be rescued by WT Sec63. Indeed, both AGAL and CD44 were
280 translocated at similar levels to that of WT cells when Sec63-/- cells were complemented with
281 WT Sec63 (Fig. 3B). By contrast, Sec63 J-mutant (HPD/AAA) complemented cells showed
282 almost no translocation of AGAL and CD44 relative to WT Sec63 complemented cells and
283 Sec63-/- cells. (Fig. 3B). This result implies that the activation of BiP ATPase by the J-domain of
284 Sec63 is critical for the translocation of these proteins. Since there is no glycosylated form of
285 AGAL in the J-mutant lane, this result also suggests that Sec63 J-mutant completely inhibits
286 Sec63 independent translocation in Sec63-/- cells (Fig. 3B), presumably by blocking
287 spontaneous BiP binding to the translocating nascent chains. Since Sec63 functions along with
288 Sec62 in post-translational protein translocation, we wondered whether Sec62 is also required
289 for translocation of AGAL and CD44. While the knockout of Sec62 had a significant effect on the
290 translocation of CD44 as mirrored by less glycosylated forms, it had a subtle effect on the
291 translocation of AGAL, suggesting that Sec62 plays a substrate-dependent role in promoting
292 translocation of secretory proteins (Fig. 3C).
293
294 Since our data above showed that weak TMDs relied on Sec63 activity (Fig. 1), we
295 wanted to determine if our new Sec63 substrates AGAL and CD44 also contain weak SSs. We
296 therefore swapped the SS of AGAL or CD44 with the strong SS from secretory hormone
297 prolactin (prl) and tested its Sec63 dependency by expressing it in Sec63-/- cells (Fig. 3D). Prl-
298 AGAL bypassed the Sec63 requirement as demonstrated by a fully glycosylated form of AGAL
299 in both WT and Sec63-/- cells, whereas AGAL with its endogenous SS poorly glycosylated in
300 Sec63-/- cells (Fig. 3D). Interestingly, some of the hydrophilic residues within the core
301 hydrophobic region of AGAL SS are conserved in mammals, suggesting an evolutionary
302 pressure to maintain Sec63 dependency (Fig. S4A). Indeed, the hydrophobicity is the key
303 determinant for Sec63 dependency since the translocation could occur independent of Sec63
304 activity for 3L-AGAL in which less hydrophobic amino acids of AGAL SS were replaced with
305 hydrophobic leucine amino acids (Fig. 3D). We also obtained a similar result for CD44 where

306 Prl-CD44 and 3L-CD44 mostly translocated independent of Sec63 activity compared to WT
307 CD44 (Fig. 3E and Fig. S4B). Lastly, replacing prolactin SS with less hydrophobic AGAL SS led
308 to Sec63 dependent translocation of prolactin into the ER, suggesting that the weak SS of
309 AGAL is sufficient to render Sec63 dependency for a protein (Fig. S4C). Collectively, our results
310 show that weak SSs rely on Sec63/BiP for efficient translocation of their nascent chains into the
311 ER lumen.
312
313 **Sec63 co-translationally releases the translocation pause caused by a marginally**
314 **hydrophobic signal sequence**
315 Previous studies have established that Sec63 is predominantly involved in post-translational
316 translocation of substrates into the ER lumen by recruiting and activating BiP ATPase through
317 its J-domain (Corsi and Schekman, 1997; Matlack et al., 1999). Studies have shown that Sec63
318 is also involved for the co-translational protein translocation of select substrates (Brodsky et al.,
319 1995; Conti et al., 2015; Jung et al., 2014; Lang et al., 2012; Young et al., 2001), but its precise
320 role in this pathway is ill-defined. The recent structural studies of the Sec61/Sec62/Sec63
321 complex from yeast suggest that both Sec63 and ribosome share the same binding site in the
322 Sec61 translocon (Itskanov and Park, 2019; Wu et al., 2019). This means both the ribosome
323 and Sec63 cannot bind to the Sec61 translocon at the same time, questioning the role of Sec63
324 in the co-translational translocation process. To determine whether our novel Sec63's
325 substrates are co- or post-translationally translocated into the ER, we performed in vitro
326 translation of AGAL transcripts using the rabbit reticulocyte (RRL) translation system including
327 [^{35}S] methionine in either the presence or absence of ER-derived rough microsomes. AGAL was
328 digested by proteinase K (PK) without microsomes, but it was translocated and glycosylated in
329 the presence of microsomes and thereby insensitive to protease digestion (Fig. 4A, and 4B). In
330 sharp contrast, AGAL was completely digested by PK when microsomes were post-
331 translationally incubated with the supernatant of the translation reaction in which ribosomes
332 were removed by centrifugation (Fig. 4A and 4B). To directly detect the co-translational
333 translocation of AGAL, we translated AGAL transcripts lacking the stop codon in vitro in either
334 the absence or presence of microsomes. The absence of stop codon left the nascent
335 polypeptide attached to the ribosome by a tRNA peptidyl linkage (Fig. 4C). AGAL translated in
336 the absence of microsomes was completely digested by PK. By contrast, AGAL or Prl-AGAL
337 translated in the presence microsomes was translocated into the ER lumen as demonstrated by
338 protease resistant glycosylation (Fig. 4C). We noticed that glycosylation of AGAL into
339 microsomes was slightly less efficient than that of Prl-AGAL. The fact that the nascent chain of

340 AGAL while covalently attached to the ribosome could be translocated and glycosylated
341 supports the conclusion that AGAL is co-translationally translocated into the ER.
342
343 The data above show that Sec63-dependent substrate (AGAL) and Sec63-independent
344 substrate (Prl-AGAL) are co-translationally translocated into the ER lumen. We therefore
345 wanted to determine at which step during the biosynthesis of AGAL, Sec63 would assist the
346 translocation of AGAL. To address this, we in vitro translated transcripts lacking the stop codon
347 encoding defined chain lengths of AGAL in either the presence or absence of microsomes (Fig.
348 4D). Translation of transcripts without stop codon yielded translationally arrested radiolabeled
349 ribosome-associated nascent chains (RNCs), representing physiological translocation
350 intermediates. RNCs of AGAL-80, -100, -130, and 160 amino acids in length were digested by
351 PK, as minimal protease protected fragments were observed (Fig. 4D). However, the RNC of
352 AGAL-220 was protected from PK digestion. These results suggest that the AGAL nascent
353 chain of ~160 amino acids is translocationally paused and accumulated to the cytosolic side of
354 the Sec61 translocon and thus accessible to PK digestion. The translocation pause was
355 released when the nascent chain reaches the length of 220 amino acids, as shown by the
356 appearance of PK protected nascent chains. Conversely, Prl-AGAL, which is not dependent on
357 Sec63, showed protease protected fragments with all different lengths of nascent chains,
358 illustrating that the successful translocation is initiated right after the SS engages the Sec61
359 translocon (Fig. 4E).
360
361 The protease protection experiment with 3L-AGAL nascent chains showed a similar
362 result to that of Prl-AGAL (Fig. 4F), which demonstrates that hydrophobicity of the SS
363 determines the efficient translocation of nascent chains into the ER. The translocation defects
364 associated with AGAL nascent chains below 220 residues were not caused by their inability to
365 target to the ER membrane since both AGAL-100 and AGAL-220 nascent chains were recruited
366 to microsomes with a comparable efficiency to nascent chains of 3L-AGAL (Fig. S5). We next
367 asked whether Sec63 is required for assisting co-translational translocation of AGAL nascent
368 chain into the ER in vitro. Indeed, nascent chains of AGAL-220 were protease sensitive when
369 incubated with microsomes derived from Sec63-/- cells compared to microsomes derived from
370 WT cells (Fig. 4G). In contrast, 3L-AGAL showed protease protected fragments for both WT and
371 Sec63-/- microsomes since its translocation into the ER was independent of Sec63. Collectively,
372 these results suggest that Sec63 substrates are co-translationally translocated into the ER
373 lumen with an initial translocation pause due to their weak SSs. Our data also suggest that

374 Sec63 can release the translocationally paused nascent chain into the ER even in the presence
375 of translating ribosome.
376
377 **Sec63 mediates BiP binding to translocating nascent chains in the ER**
378 Next, we asked why weak SSs are frequently found in secretory proteins despite their
379 inefficiency in mediating translocation. We hypothesized that SSs may direct the recruitment of
380 luminal chaperones to bind onto nascent chains through modulating the Sec61 translocon
381 associated factors. Accordingly, we predicted that the weak SS of AGAL nascent chain would
382 preferentially bind to BiP chaperone since its translocation depends on the J-domain of Sec63.
383 To test this, we introduced a C-terminal 3xFLAG-tag on endogenous genes encoding BiP using
384 CRISPR/Cas9 because commercial antibodies against BiP could not efficiently pulldown the
385 endogenous BiP. The C-terminal 3xFLAG introduced downstream of the KDEL retention signal
386 of BiP did neither significantly change BiP expression levels nor its localization to the ER (Fig.
387 5A, B). To determine if the newly synthesized AGAL nascent chain is a substrate of BiP, BiP-
388 3xFLAG HEK293 cells expressing AGAL were radiolabeled with ^{35}S methionine/cysteine for 2
389 min, chased, and immunoprecipitated with an anti-FLAG antibody. Consistent with our
390 prediction, AGAL nascent chains were co-immunoprecipitated with BiP. The interaction reduced
391 with the chase time because matured AGAL presumably left the ER (Fig. 5C). BiP binding to
392 AGAL and other endogenous substrates was sensitive to ATP since the incubation of cell
393 lysates with ATP prior to immunoprecipitation completely abolished BiP binding to all nascent
394 chains including AGAL (Fig. 5D).
395
396 To determine if BiP binding to AGAL relies on Sec63, we generated Sec63 knockout in
397 BiP-3xFLAG HEK293 cells using CRISPR/Cas9 (Fig. S6A). Consistent with our model, BiP
398 binding to AGAL was markedly reduced in Sec63-/- cells, although about two-fold more BiP was
399 immunoprecipitated in Sec63-/- cells than WT cells due to BiP upregulation in Sec63-/- cells
400 (Fig. 5E and 5F). Lastly, we tested whether BiP binding to nascent chains of AGAL containing
401 strong SSs also depends on Sec63. To our surprise, similar to AGAL, we found nascent chains
402 of Prl-AGAL and 3L-AGAL also relied on Sec63 activity for binding to BiP, although they were
403 translocated independent of Sec63 (Fig. S6B). This result implies that Sec63 is also required for
404 mediating BiP binding to these nascent chains after their translocation into the ER. Considered
405 together, these results suggest that AGAL nascent chains depend on Sec63 activity for their
406 translocation into the ER as well as their binding to BiP. However, AGAL nascent chains

407 containing strong SSs translocate independent of Sec63 activity but still requires Sec63 for BiP
408 binding after their translocation into the ER lumen.
409
410 **BiP-dependent protein translocation is coupled with protein folding in the ER**
411 We hypothesized that BiP binding to AGAL nascent chains via Sec63 may not only facilitate
412 their translocation into the ER but also promote their folding and maturation in the ER. To test
413 this, we monitored the secretion of AGAL since properly folded AGAL is secreted into the
414 extracellular medium upon overexpression (Ioannou et al., 1992). Pulse-chase analysis
415 revealed that intracellular levels of both AGAL and Prl-AGAL in WT cells were reduced during
416 the chase period with a concomitant increase in their extracellular levels (Fig. 6A). By contrast,
417 the secretion of both AGAL and Prl-AGAL was about two-fold reduced in Sec63-/- cells (Fig.
418 6A). The non-translocated/unglycosylated population of AGAL and Prl-AGAL were degraded by
419 proteasomes in Sec63-/- cells since they were accumulated upon treatment with the
420 proteasomal inhibitor MG132 (Fig. 6A). Moreover, AGAL carrying either a weak or strong SS
421 showed a similar α-galactosidase activity in WT cells, but the activity was significantly reduced
422 for all AGAL variants in Sec63-/- cells (Fig. S7A). Reduced activity and secretion in Sec63-/-
423 cells could be attributed to the weak SS of AGAL that relies on Sec63 for efficient translocation
424 into the ER. However, these effects were also observed for even AGAL carrying a strong SS
425 that translocated into the ER mostly independent of Sec63. This indicates that AGAL with a
426 strong SS is likely misfolded after translocation into the ER due to its inefficient binding to BiP in
427 Sec63-/- cells (Fig. S6B).
428
429 To determine if AGAL carrying a strong SS is misfolded after translocation into the ER of
430 Sec63-/- cells, we separated the detergent soluble fraction from the insoluble fraction, an
431 indicator of misfolded/aggregated proteins. We found a subtle difference between AGAL
432 variants in WT cells. By contrast, a significant fraction of translocated/glycosylated form of Prl-
433 AGAL and 3L-AGAL was sedimented as detergent-insoluble fractions in Sec63-/- cells,
434 supporting our notion that AGAL carrying a strong SS are misfolded and aggregated after
435 translocation into the ER lumen (Fig. 6B). To our surprise, most of the translocated/glycosylated
436 form of AGAL moved to the soluble fraction, but non-translocated/unglycosylated AGAL with the
437 uncleaved SS sedimented as detergent-insoluble aggregates (Fig. 6B and S7B). This result
438 implies that the AGAL population that was translocated even in the absence of Sec63 could be
439 properly folded and matured in Sec63-/- cells. We reasoned that this is due to the upregulation
440 of BiP in Sec63-/- cells, which could partly promote translocation as well as folding of AGAL

441 nascent chains. Sec63-independent but BiP dependent translocation of AGAL in Sec63-/- was
442 supported by the observation that overexpression of BiP ATPase mutants significantly impaired
443 translocation of AGAL in Sec63-/- cells with little effects on the translocation of AGAL with a
444 strong SS (Fig. S7C). Thus, our results suggest that strictly BiP-dependent translocation of
445 AGAL in Sec63-/- ensures subsequent folding and maturation of AGAL in the ER, whereas BiP-
446 independent translocation of AGAL carrying a strong SS is at high risk for misfolding after
447 translocation into the ER of Sec63-/- cells.
448
449 Our data thus far suggest a model in which proteins with weak SSs can sense BiP
450 availability in the ER lumen and translocate only upon binding to BiP molecules, which also
451 subsequently assist with protein folding and maturation. In the absence of BiP binding, however,
452 these proteins are synthesized in the cytosol and degraded by proteasomes. To test this model,
453 we co-expressed AGAL variants with BiP mutants (G227D and T37G) that can bind substrates
454 but fail to release them upon ATP binding (Wei et al., 1995). AGAL translocation was markedly
455 reduced in cells expressing BiP mutants and that primarily the non-translocated/unglycosylated
456 AGAL was found in the insoluble fraction (Fig. 6C). However, the population of AGAL that
457 translocated into the ER was completely soluble as shown by the presence of glycosylated form
458 only in the soluble fraction. This result suggests that AGAL either translocates and folds well by
459 binding to active BiP molecules or it does not translocate into the ER when it failed to bind
460 active BiP molecules. Conversely, Prl-AGAL was efficiently translocated/glycosylated
461 irrespective of the BiP functional status in the ER, and thus a significant fraction of Prl-AGAL is
462 misfolded after translocation into the ER and moved to the insoluble fraction (Fig. 6C).
463 Furthermore, our in vitro experiment demonstrated that the translocation of AGAL nascent chain
464 was significantly inhibited in microsomes depleted of luminal proteins compared to that of AGAL
465 containing a strong SS (Fig. S7D).
466
467 We further examined whether the weak SS of AGAL can sense BiP availability in the ER
468 by depleting BiP in cells using siRNA (Fig. 6D). The translocation of AGAL but not Prl-AGAL
469 and 3L-AGAL was impaired in BiP depleted cells, which resembles the result seen in Sec63-/-
470 cells (Fig. S7B and S7E). In BiP depleted cells, primarily the non-translocated/unglycosylated
471 form of AGAL moved to the insoluble fraction, whereas a significant population of
472 translocated/glycosylated form of Prl-AGAL and 3L-AGAL went to insoluble fractions (Fig. 6E).
473 These results suggest that the weak SS of AGAL can respond to BiP deficiency by not entering
474 the ER where they would misfold. We then investigated whether the weak SS plays a role in

475 preventing misfolding of AGAL under ER stress conditions, which is known to reduce BiP
476 availability due to the increased load of misfolded proteins in the ER (Kang et al., 2006; Vitale et
477 al., 2019). Indeed, a fraction of Prl-AGAL and 3L-AGAL but not AGAL were readily misfolded
478 and aggregated during ER stress-induced with DTT (Fig. S7F). Collectively, our data suggest
479 that proteins with weak SSs can sense BiP chaperone deficiency in the ER and respond by
480 aborting translocation into the ER to prevent misfolding and aggregation inside the ER.
481
482 **Discussion**
483 In addition to targeting roles (Blobel and Dobberstein, 1975), our study presented here suggest
484 that SSs also contain information for protein folding in the ER by coupling protein translocation
485 with the availability of chaperone binding. Our data show that marginally hydrophobic (or weak)
486 SSs containing proteins cause a translocation pause at the Sec61 translocon to recruit luminal
487 BiP chaperone via Sec63 for efficient folding in the ER (Figure 7). This translocation pausing
488 mechanism also allows aborting translocation under conditions of BiP deficiency to prevent
489 protein misfolding and aggregation in the ER. Thus, our studies suggest that evolution has
490 exploited diversity in SSs to recruit selective luminal chaperones for protein folding and
491 maturation of secretory and membrane proteins.
492
493 Our experiments with model WRB substrates demonstrate that marginally hydrophobic
494 C-terminal TMDs carrying longer C-terminal tails are transiently retained at the Sec61
495 translocon and are eventually released into the ER lumen. Our data indicate that Sec63
496 monitors the Sec61 translocon status and directs the clearance of nascent chains transiently
497 retained at the translocon. The key function of Sec63 is to activate the luminal chaperone BiP
498 ATPase through its J-domain so that ADP-bound BiP can bind and release the translocon-
499 retained nascent chain into the ER lumen (Fig. 7). Our findings could explain why the
500 translocation of the C-terminus of membrane proteins is impaired upon disrupting the Sec62/63
501 complex in yeast (Jung et al., 2014). Our “substrate-trapping” proteomic approach identified
502 numerous secretory and membrane proteins that are transiently retained at the Sec61
503 translocon complex. Although our data analyzing AGAL and CD44 reveal that marginally
504 hydrophobic SSs plays an important role in determining Sec63 dependency, other SS features
505 such as length and flanking amino acids might also contribute to transient retention of nascent
506 chains at the Sec61 translocon and thus recruiting Sec63/BiP (Schorr et al., 2020b; Ziska et al.,
507 2019).
508

509 The role of Sec63 in the post-translational translocation pathway is well established,
510 however, its function in the co-translational protein translocation remains unclear due to the fact
511 that Sec63 and ribosome share the same binding site in the Sec61 translocon (Brodsky et al.,
512 1995; Conti et al., 2015; Itskanov and Park, 2019; Jung et al., 2014; Lang et al., 2012; Matlack
513 et al., 1999; Misselwitz et al., 1998; Wu et al., 2019; Young et al., 2001). Three of our
514 observations suggest that Sec63's substrates, which we identified here, use the co-translational
515 protein translocation pathway. First, our Sec63 clients are larger than 160 amino acids in length,
516 which appear to be a limitation for post-translational protein translocation in mammalian cells
517 (Lakkaraju et al., 2012). Second, our *in vitro* experiment shows that the Sec63 substrate AGAL
518 could not be post-translationally translocated into the ER. Third, AGAL was efficiently
519 translocated and glycosylated even when it is still attached to the ribosome. Our data also
520 raises the intriguing question of how Sec63 and ribosome can bind to the Sec61 translocon at
521 the same time as they share the same binding site. Our protease protection data suggest that
522 the translocation of nascent chain containing a weak SS is initially paused ~160 amino acids in
523 length and is accumulated on the cytosolic side of the membrane (Fig. 7). This is presumably by
524 caused by a loose interaction between the ribosome nascent chain carrying a weak SS and the
525 Sec61 translocon. We speculate that Sec63 is recruited to the Sec61 translocon by sensing and
526 binding to the cytosolically accumulated nascent chain (Fig. 7). Once recruited, Sec63 J-domain
527 triggers ATP hydrolysis of BiP ATPase to bind recalcitrant nascent chains entering through the
528 Sec61 translocon, thus facilitating their translocation into the ER. Our findings suggest that cells
529 can compensate the loss of Sec63-mediated recruitment of BiP to translocating nascent chains
530 by upregulating BiP levels in Sec63-/- cells. We predict that once the cytosolically accumulated
531 nascent chain at the Sec61 translocon is pulled into the ER lumen, the translating ribosome may
532 establish a tight complex with the Sec61 translocon by displacing Sec63. Our findings may
533 explain why Sec62/Sec63 is selectively enriched with the Sec61 translocon hosting nascent
534 chains of prion protein, which also has a weak signal sequence (Conti et al., 2015; Kang et al.,
535 2006). Our data suggest that Sec62 has a less significant role in the translocation of AGAL,
536 however, additional substrates need to be analyzed to determine its role in translocation and
537 folding of secretory and membrane proteins.
538
539 Despite inefficiency in mediating protein translocation, weak SSs are commonly found
540 and conserved across all mammals. One important function for weak SSs is that they can
541 selectively impair the translocation of their mature domains into the ER during ER stress
542 conditions (Kang et al., 2006). Earlier studies have also shown that the nature of the interaction

543 between a SS and the Sec61 translocon can influence downstream events such as the timing of
544 SS cleavage and N-linked glycosylation (Rutkowski et al., 2001; Rutkowski et al., 2003), which
545 in turn may direct protein maturation (Lingappa et al., 2002). Our study now suggests that
546 Sec63/BiP-mediated release of initial translocation pausing of nascent chains carrying weak
547 SSs is coupled with the BiP-assisted protein folding and maturation in the ER. This mechanism
548 may be particularly important for the folding of AGAL since it contains the hydrophobic N-
549 terminal region right after the signal sequence that is buried into the structure, which is also the
550 hotspot for genetic mutations causing Fabry disease (Garman, 2007). When functional BiP is
551 scarce in the ER under conditions such as stress and starvation, the translocation of substrates
552 like AGAL is prevented and thus misfolding and aggregation in the ER lumen is avoided.
553 Although the detergent soluble assay used here provides evidence of protein folding and
554 misfolding, it does not provide quantitative information. New experimental approaches such as a
555 folding sensitive fluorescence reporter are required to quantitatively determine signal sequence-
556 dependent folding in live cells. The non-translocated substrates are likely eliminated by the
557 Bag6-mediated cytosolic quality control pathway (Hessa et al., 2011).
558
559 Our data show that increasing the hydrophobicity of weak SSs is sufficient to bypass the
560 BiP-dependent translocation, but translocated nascent chains are at risk of misfolding and
561 aggregation under conditions of BiP deficiency. Our findings may explain why about 40% of the
562 analyzed yeast secretome use the Sec63/Sec62-mediated post-translational translocation
563 pathway (Ast et al., 2013), which likely reflects their dependency on BiP for folding and
564 maturation. We predict that the recruitment of BiP may occur at any time during translocation as
565 long as nascent chains contain a translocon retention element such as marginally hydrophobic
566 sequences. Consistent with this notion, earlier studies have observed translocation pausing
567 events in which the nascent chain transiently exposes to the cytosol during the translocation of
568 apolipoprotein B and immunoglobulin G (Hegde and Lingappa, 1996; Rutkowski et al., 2001).
569 We suggest that these translocation pausing events may promote maturation of proteins by
570 recruiting specific luminal chaperones to nascent chains.
571
572 We propose that the translocon-associated accessory proteins may function as a hub to
573 recruit specific luminal chaperones and enzymes to bind onto translocating nascent chains. In
574 addition to marginally hydrophobic SSs that recruit BiP through Sec63, we propose other
575 features in SSs may also modulate translocon accessory proteins and their downstream
576 chaperones. This notion is supported by the recent findings that glycine and proline-rich SSs

577 containing proteins depend on the translocon-associated protein (TRAP) complex for their
578 efficient translocation into the ER (Nguyen et al., 2018). Curiously, the endogenous TRAP
579 complex was recently shown to form a complex with Calnexin chaperone (Lee et al., 2020),
580 suggesting a potential role in chaperone selection and glycoprotein folding. It is tempting
581 suggest that multiple cellular translocation pathways (Aviram and Schuldiner, 2017) in cells may
582 have evolved to meet the challenges of folding the complex secretome by recruiting protein-
583 specific chaperones.
584
585 Earlier studies have shown that mutations in the Sec63 gene can cause autosomal
586 dominant polycystic kidney diseases (PKD) due to defects in protein maturation and cell surface
587 localization of polycystin 1 and presumably other proteins (Besse et al., 2017; Davila et al.,
588 2004). Around one-third of our Sec63 clients have been previously implicated in human genetic
589 disorders, including many lysosomal storage disorders and PKD. This suggests that these
590 proteins are particularly vulnerable to misfolding and aggregation due to genetic mutations. In
591 light of our findings, selective induction of BiP by small molecules, such as AA147 (Plate et al.,
592 2016), may be a potential therapeutic approach to treat these diseases. Furthermore, our
593 studies may guide the design of customized SSs to produce a growing number of therapeutic
594 recombinant secretory proteins, including enzymes, growth factors, and antibodies.
595
596 **Materials and Methods**
597
598 **DNA constructs**
599
600 Most of the cDNAs used in this study were cloned into pcDNA/FRT/TO vector (Invitrogen).
601 pcDNA/FRT/TO containing WRB-HA-Venus substrates were previously described (Sun and
602 Mariappan, 2020). Mouse Sec63-FLAG (template mouse Sec63 plasmid was a kind gift from Dr.
603 Stefan Somlo, Yale School of Medicine). Sec63 HPD/AAA (H132A/P1133A/D134A) and Sec63
604 Δ230-330 were previously described (Li et al., 2020). Prl-AGAL-FLAG and CD44-FLAG
605 constructs were generated by replacing AGAL SS (1-31aa) and CD44 SS (1-20aa) with bovine
606 prolactin SS (1-31aa) using phosphorylated primers and Phusion Site-Directed Mutagenesis
607 procedure. Since the flanking threonine (T) of prolactin SS that remained after the SS cleavage
608 inhibited AGAL activity, Prl-AGAL-HA used in Fig. 5 and Fig. 6 was created by replacing AGAL
609 SS (1-31aa) with bovine prolactin SS (1-30 aa) using the above procedure. 3L-AGAL and 3L-
610 CD44 constructs were created using Pfu polymerase (Agilent Technologies) based site-directed

611 mutagenesis method. AGAL-prolactin-FLAG was cloned by replacing bovine prolactin SS (1-30)
612 with AGAL SS (1-31) using the above procedure. BiP mutants (G227D and T37G) were
613 introduced into pcDNA Rat BiP using the above site-directed mutagenesis method. The coding
614 sequences of all constructs were verified by sequencing (Yale Keck DNA Sequencing).
615

616 **Antibodies and reagents**

617 Antibodies used for immunoblotting are as follows: Rat α-FLAG L5 (BioLegend #12775), Rabbit
618 α-HA, Rabbit α-GFP, Rabbit α-Sec63, Rabbit α-Sec63, Rabbit α-Sec61β, Rabbit α-BAG6, and
619 Rabbit α-Sec61α are a gift from Dr. Ramanujan Hegde. Rabbit α-BiP (Proteintech #11587-1-
620 AP), Mouse α-HA (Biolegend #901513), Mouse α-PDI (Affinity Bioreagents #MA3-018), Goat α-
621 Rat-HRP (Cell Signaling, #7077), Goat α-Mouse-HRP (Jackson ImmunoResearch, #115-035-
622 003), Goat α-Rabbit-HRP (Jackson ImmunoResearch, #111-035-003), Goat anti-rat HRP (Cell
623 signaling, #7077S) Goat α-RAT IgG-Cy2 (Jackson ImmunoResearch # 112-225-167), Goat α-
624 Mouse IgG-Alexa657 (Invitrogen # A-21235). Beads were purchased as follows: Strep-Tactin
625 XT beads (IBA, Germany #2-4010-010), Rat anti-FLAG L5 affinity gel (Biolegend, #651503),
626 Protein A agarose (Repligen, #CA-PRI-0100), Mouse anti-HA magnetic beads (Pierce, #88837),
627 Poly L-lysine (Peptides International, # OKK-3056). Rabbit Reticulocyte Lysate was purchased
628 from Green Hectares (Ph:1-800-GHLYSAT). Detergents were purchased as follows: Digitonin
629 (EMD Millipore, Billerica, MA), Triton X-100 (Thermo Fisher), Sodium Deoxycholate (Sigma),
630 and Sodium Dodecyl Sulfate (Sigma), Saponin (Sigma), 37% Formaldehyde (Avantor), and
631 Hoechst 33342 stain (Cell Signaling # 4082S), Tween 20 (American Bioanalytical).

632 **Cell culture and CRISPR/Cas9 genome editing**
633 HEK293-Flp-In T-Rex cells (Invitrogen) were cultured in high glucose DMEM containing 10%
634 FBS at 5% CO_2. The knockout cell lines of Sec63 or Sec62 as well as 2xStrep-Sec61α,cells
635 were previously described (Li et al., 2020). To engineer chromosomal BiP 3xFLAG HEK293
636 cells, the BiP gRNA sequence (5′ CCTCTTCACCAGTTGGGGG 3′) was cloned into the gRNA
637 expression vector (Mali et al., 2013). The single-strand DNA oligonucleotide homology-directed
638 repair (HDR) sequence
639 (CTGGAAGAAATTGTTCAACCAATTATCAGCAAACTCTATGGAAGTGCAGGCCCTCCCCCAg
640 ccGACTACAAGGACCACGACGGCGATTATAAGGATCACGACATCGACTACAAAGACGACGA
641 TGACAAGgcaACTGGTGAAGAGGATACAGCAGAAAAAGATGAGTTGTAGACACTGATCTGC
642 TAGTGCTGTAA) was synthesized (IDT) for inserting 3xFLAG-tag into the C terminus of BiP

643 encoding gene in HEK293 cells. While the sequence indicated in blue encodes for 3xFLAG,
644 green denotes the C-terminus KDEL peptide encoding sequence. Pink indicates the stop codon.
645 300 pmol of HDR oligonucleotide was electroporated into one million HEK293 Flp-In T-Rex cells
646 along with 2.5 µg each of pSpCas9(BB)-2A-Puro and BiP gRNA plasmid (Amaxa kit R, program
647 A-24; Lonza). Immediately after electroporation, cells were plated in a 6 well plate. After 24 h of
648 electroporation, the expression of Cas9 was selected by puromycin treatment (2.5 µg/ml) for 48
649 h. Most cells were died of electroporation and puromycin treatment, but about 5% survived cells
650 were grown for a week without puromycin. Cells were replated at 0.5 cell/well in 96 well plates
651 and expanded for 3 weeks. Individual clones were examined for the presence of 3xFLAG BiP by
652 immunoblotting with anti-BiP antibodies. Positive clones containing 3xFLAG BiP exhibited a
653 slower migrating band compared to BiP present in control HEK293 cells.
654
655 **Immunoprecipitations**
656 To test the interaction between WRB-Venus substrates and the Sec61 translocon complex, the
657 chromosomally Sec61α-2x Strep-tagged HEK293 cells were plated on a poly-lysine (0.15mg/ml)
658 coated 6 well plates at 0.6 million/well. The cells were transiently transfected with 2µg WRB-
659 Venus substrates using 5µL of Lipofectamine 2000 and treated with 200 ng/ml doxycycline to
660 induce protein expression. After 24 hours of transfection, cells were harvested in 1XPBS and
661 centrifuged for 2 min at 10,000g. The cell pellet was lysed in 200µl of lysis buffer (50 mM Tris
662 pH 8.0, 150 mM NaCl, 5 mM MgAc, 2% digitonin, and 1X Roche EDTA-free protease inhibitor
663 cocktail) by incubating on ice for 30 min and then diluted to 1ml with the lysis buffer containing
664 0.1% digitonin. The 5% digitonin (Millipore) stock was boiled for 5 min just before adding into the
665 lysis buffer to avoid digitonin precipitation during IPs. The supernatants were collected by
666 centrifugation at 15,000g for 15 min at 4°C. For co-immunoprecipitation, the supernatant was
667 rotated with StrepTactin XT beads (IBA, Germany) for 2h at 4°C. The beads were washed 3
668 times with 1 ml of wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc, 0.1% digitonin).
669 The bound material was eluted from the beads by directly boiling in 50 µl of 2x SDS sample
670 buffer for 5 min and analyzed by immunoblotting.
671 To test the interaction between Sec63-FLAG and Sec61 translocon shown in Fig. S2A,
672 the above digitonin IP procedure was followed but using HEK293 Sec63-/- cells stably
673 expressing variants of Sec63-FLAG. 20ng/mL doxycycline was used to induce the protein
674 expression for 24 h before harvesting for IPs.
675
676 **Metabolic labeling and immunoprecipitation**

677 Indicated cells (0.6 million cells/well) were plated on poly-lysine coated (0.15mg/mL) 6-well
678 plates and transiently transfected with 2μg indicated plasmids. The expression of indicated
679 proteins was induced with doxycycline (200ng/mL) for 24 hours. The cells were starved with
680 Met/Cys-free medium including 10% dialyzed FBS and doxycycline for 30 min. Cells were then
681 labeled with 80μCi/mL Express-^{35}S protein labeling Mix (NEG772007MC) for 30 min. The
682 labeled cells were directly harvested in 1ml of RIPA buffer including 1X EDTA-free protease
683 inhibitor cocktail (Roche). The lysate was cleared by centrifugation at 20,000g for 15 min. The
684 supernatant was mixed with the unconjugated or antibodies conjugated to beads indicated in
685 the figure legends for 1.5 h in the cold room. In the case of unconjugated anti-rabbit antibodies,
686 20ul slurry of protein-A agarose was added to samples and incubated for another 1 hour. The
687 beads were washed three times with 1 ml of RIPA buffer, and proteins were eluted with 50ul of
688 2X SDS sample buffer and run on 7.5% or 10% Tris-Tricine SDS-PAGE gels. The gels were
689 dried and analyzed by autoradiography.
690 For isolating endogenous substrates transiently retained at the Sec61 tranlocon, Sec63-
691 /- and Sec63-/- cells complemented with WT Sec63, Sec63 J-mutant (HPD/AAA) or the
692 translocon interaction defective Sec63 mutant (Δ230-300) were labeled as described above.
693 The labeled cells were washed and harvested 1XPBS and centrifuged at 10,000g for 2 min. The
694 cell pellets were lysed in 200μL of lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgAc,
695 2% digitonin, and 1X Roche protease inhibitor cocktail) by incubating on ice for 30 min and then
696 diluted to 1ml with the lysis buffer containing 0.1% digitonin. The supernatant was collected by
697 centrifugation at 15,000g for 15 min and incubated with Rat anti-FLAG beads (BioLegend) for
698 2h at 4°C. The beads were washed 3 times with 1 ml of wash buffer (50 mM Tris pH 8.0, 150
699 mM NaCl, 5 mM MgAc, 0.1% digitonin). The bound material was eluted from beads by directly
700 boiling in 50 μl of 2x SDS sample buffer for 5 min and analyzed by SDS-PAGE and
701 autoradiography.
702 For pulse-chase analysis, transfected cells were labeled as above for 30 min and
703 chased in complete DMEM medium supplemented with label-free 2mM Methionine and 2mM
704 Cysteine for indicated time points. The cells were directly harvested in RIPA buffer including 1X
705 EDTA free protease inhibitor cocktail (Roche) and cleared by centrifugation at 15,000g for 15
706 min. The supernatant was incubated with the indicated antibody conjugated to beads for 2 h.
707 For checking the secretion of AGAL or Prl-AGAL, the medium was collected at each time point,
708 adjusted to 50 mM Tris pH 8.0, 150 mM NaCl, 1% Triton X-100, and immunoprecipitated and
709 analyzed as above.
710

711 **Endogenous BiP-3xFLAG immunoprecipitation**

712 BiP-3xFLAG HEK293 cells (0.9 million cells/well) were plated on poly-lysine coated (6-well
713 plates and transiently transfected with 2μg of indicated plasmids and induced with doxycycline
714 (200ng/mL) for 24 hours. The cells were starved with Met/Cys-free medium including 2.5%
715 dialyzed FBS and doxycycline for 45 min. Cells were then labeled with 90 μCi/mL Express-^{35}S
716 protein labeling Mix (NEG772007MC) for 2 min at 37°C. The labeled cells were immediately
717 placed on ice and harvested in 1ml of ice-cold 1X PBS and centrifuged at 10,000g for 2 min.
718 Cell pellets were solubilized in 200 ul of lysis buffer (50 mM Tris, pH 7.5, 160 mM NaCl, 1%
719 Digitonin, 1mM CaCl2, Apyrase (10 U/mL)) for 30 min on ice. The lysate was diluted to 1ml with
720 lysis buffer containing 0.1% Digitonin before centrifugation at 20,000g for 15 min. While 800ul of
721 supernatant was incubated with 12 ul of mouse anti-FLAG beads (Sigma), the remaining 200ul
722 was diluted to 800ul with NP40 buffer and incubated 15 ul mouse anti-HA magnetic beads
723 (Thermo Scientific). After 90 min of incubation, anti-FLAG beads were washed 3x with 1ml of
724 lysis buffer containing 0.1% Digitonin but anti-HA magnetic beads were washed with 3x with 1ml
725 of NP40 buffer. The washed beads were boiled with 50ul of 2x SDS samples buffer and
726 analyzed by 7.5% Tris-Tricine SDS PAGE and autoradiography. For the 2nd IP, anti-FLAG
727 immunoprecipitants (~40 ul in 2x SDS sample buffer) were diluted to 1ml with Triton buffer (50
728 mM Tris, pH 8, 150 mM NaCl, 1% Triton X-100) and incubated with 15ul of anti-HA magnetic
729 beads for 90 min. The beads were washed 3x with 1ml of Triton buffer and analyzed as above.
730

731 **Affinity purification of translocon-retained endogenous substrates and mass**
732 **spectrometry**

733 Three different HEK293 cell lines (Sec63-/-, Sec63-/- stably complemented with WT Sec63-
734 FLAG, Sec63-/- complemented with Sec63-FLAG J-mutant (HPD/AAA) were plated in 10cm
735 plates at 10 million cells/plate. Protein expression was induced by 20ng/mL doxycycline. After
736 24 h of induction, the cells were transferred from one 10cm plate to two 15cm plates for each
737 cell line including 20ng/mL doxycycline. After culturing two days, cells at ~100% confluence
738 were treated with 100ug/mL cycloheximide for 5 min. The cells were then collected using ice-
739 cold 1XPBS and centrifuged at 1500 g for 10 min to obtain the cell pellets. The cell pellets were
740 resuspended in lysis buffer (50mM HEPES pH 7.4, 100mM KAc, 50mM NaCl, 10mM MgAc, 1%
741 digitonin, 1X Roche protease inhibitor cocktail, 1mM DTT, 100ug/mL cycloheximide). The
742 lysates were further homogenized using a dounce by passing through manually for 10 times.
743 The lysates were then centrifuged at 35,000g for 30 min at 4°C in an MLA80 rotor. The
744 supernatant was incubated with 300μL of pre-washed mouse anti-FLAG beads (Sigma) slurry

745 for 2 hours at 4°C. The beads were washed for 6 times with 5 mL of wash buffer (50mM HEPES
746 pH 7.4, 100mM KAC, 50mM NaCl, 10mM MgAC, 0.1% digitonin, 1mM DTT). The samples were
747 eluted with the elution buffer (0.1M Glycine pH2.3 and 0.1% Triton X-100). The eluted materials
748 were TCA precipitated and subjected to mass spectrometry analysis.
749 The immunoprecipitants were digested with trypsin and analyzed by LC-MS/MS using an
750 LTQ Orbitrap Elite equipped with a Waters nanoACQUITY ultra-performance liquid
751 chromatography (UPLC) system at the Mass Spectrometry & Proteomics Resource of the W.M.
752 Keck Foundation Biotechnology Resource Laboratory, Yale University. Proteins were searched
753 against the homo sapiens SwissProt protein database using the Mascot Search Engine (Matrix
754 Science, LLC, Boston, MA; version. 2.6.0). Scaffold program (version Scaffold_4.8.7; Proteome
755 Software, Portland, OR) was used to validate MS/MS-based peptide and protein identifications.
756 Scaffold program was also used to generate normalized spectral counts for each sample. In the
757 supplemental Table 1, we listed protein hits from all three samples that have a protein threshold
758 of 5.0% FDR, at least 2 peptides, and a peptide threshold of 95%. In the supplemental Table 2,
759 we listed protein hits of secretory/membrane proteins from Sec63 J-mutant that are at least 2-
760 fold enriched relative to WT Sec63 and 5-fold enriched relative to negative control Sec63-/-
761 cells. The information of SSs or the 1st TMD sequences, subcellular localization, and association
762 with human diseases were extracted from UniProt.Org. In the supplemental Table 3, we listed
763 secretory/membrane protein hits from WT Sec63 that are at least 2-fold enriched relative to
764 Sec63 J-mutant and 5-fold enriched relative to negative control Sec63-/- cells. The hydrophobic
765 region of a signal sequence shown in Fig. 3D, E was defined using the MacVector program by
766 selecting the Kyte-Doolittle hydrophilicity option in windows of 3-19 amino acids. The defined
767 hydrophobic region was analyzed by grand average hydropathy (GRAVY) to obtain a
768 hydrophobicity score.
769
770 **Endo H treatment**
771 Indicated cell lines (0.15 million cells/well) were plated on poly-lysine (0.15mg/mL) coated 24-
772 well plates and transiently transfected with 400ng indicated plasmids. Protein expression was
773 induced with 200ng/mL doxycycline. After 24 h of transfection, the cells were collected in 150μL
774 of 150ul 2x SDS sampled buffer and boiled for 10min. 20ul sample was mixed with 65ul of
775 water, 5ul of 20% Triton X-100, 10 ul of 10X G5 buffer (NEB), and 1ul of Endo H HF (NEB). The
776 reactions were incubated for 4h at and were terminated by adding 50μL 5x SDS sample buffer
777 and boiled for 5 min before analyzing by immunoblotting. The above procedure was used for
778 Endo H digestion of immunoprecipitated samples that were already boiled in a 2x SDS sample.

779

780 ***In vitro* transcription and translation**

781 The open reading frames were PCR amplified from pcDNA/FRT/TO based constructs using a
782 forward primer bearing Sp6 promoter sequence and annealing sequence to the CMV promoter
783 and a reverse primer that anneals to the poly-A region of the vector. For generating ribosome-
784 associated nascent chains (RNCs), PCR products were amplified using the above forward
785 primer, but the reverse primer anneals to 3' of the template cDNA without the stop codon. In
786 vitro transcription reactions including PCR product, SP6 polymerase (NEB), and RNasin
787 (Promega) were performed at 37°C for 1.5h as described previously (Sharma et al., 2010). The
788 transcripts were translated in the rabbit reticulocyte lysate (RRL) translation system including
789 ^{35}S-methionine and with or without rough microsomes. Rough microsomes were either prepared
790 from the canine pancreas as described previously (Walter and Blobel, 1983) or cultured
791 HEK293 or Sec63-/- cells as described previously (Zhang et al., 2013). Transcripts encoding
792 WRB-Venus variants were translated in the presence or absence of HEK or Sec63-/-
793 microsomes for 30 min at 32°C. The translated products (10μL) were denatured with 5x volume
794 of 2x SDS sample buffer and analyzed by SDS-PAGE and autoradiography.

795 For the co-translational translocation assay, transcripts encoding AGAL-FLAG, Prl-
796 AGAL-FLAG were translated in RRL in the presence or absence of canine pancreas
797 microsomes (CRM) for 30 min at 32°C. The translation products were treated or untreated with
798 0.5mg/mL Proteinase K (PK) for 60 min on ice. PK digestion was terminated by adding 5 mM
799 PMSF and incubating on ice for 15 min. The resulting samples were transferred into the boiling
800 SDS buffer (100 mM Tris pH 8.0 and 1% SDS) and heated for an additional 5 min. The samples
801 were diluted 10 times with the Triton buffer (50 mM Tris pH 8.0, 150mM NaCl, 1% Triton X-100).
802 The diluted samples were incubated with anti-FLAG beads (Bio-legend) for 2 hours at 4°C. The
803 beads were washed three times with 1ml of Triton buffer and eluted by directly boiling in 50 μl of
804 2XSDS sample buffer for 5 min and analyzed by SDS-PAGE and autoradiography. For the post-
805 translational translocation assay, transcripts encoding AGAL-FLAG were translated in RRL for
806 30 min at 32°C. The translation reactions were centrifuged at 70,000 rpm in a TLA120.1 rotor for
807 30 min at 4°C. The post-ribosomal supernatant was incubated with or without microsomes for
808 30 min at 32°C. The resulting samples were treated with proteinase K and analyzed as above.

809 RNCs of indicated lengths were produced from transcripts lacking stop codon in the
810 presence or absence of the indicated microsomes. The resulting translated products were split
811 into two and incubated with or without 0.5mg/mL Proteinase K (PK) for 60 min on ice. The PK
812 digestion was terminated by adding 5 mM PMSF and incubating on ice for 15 min. The samples

813 were transferred into boiling the 2x SDS sample buffer and incubated for an additional 5 min.
814 For RNCs of AGAL or Prl-ss-AGAL, the translated products were analyzed by autoradiography
815 after denaturation and immunoprecipitation with anti-FLAG beads as described above. In some
816 cases, the denatured samples were treated with RNase A for 15 min at 37°C to remove tRNAs
817 from nascent chains before analyzing by SDS PAGE and autoradiography.
818 Luminal proteins depleted CRM shown in Fig. 6C was prepared as described previously
819 (Kang et al., 2006), 100ul CRM at a concentration of 1eq/ul was diluted with 400ul with
820 membrane buffer (MB: 50 HEPES 7.4, 250mM Sucrose, 2mM MgCl2, 1mM DTT, and 0.075%
821 deoxy Big CHAP (DBC)) for 10min on ice. The detergent permeabilized microsomes were
822 centrifuged for 70,000 rpm for 30min in a TLA100.3 rotor. The resulting pellet was resuspended
823 in 500ul of MB without DBC and centrifuged as above. The final pellet was resuspended in 50ul
824 of MB without DBC and used for in vitro translation reactions as indicated above.
825
826 **Membrane recruitment assay**
827 We followed the previously published procedure (Yanagitani et al., 2011) with the following
828 modifications. Transcripts encoding versions of AGAL lacking a stop codon were translated in
829 RRL including ^{35}S-methionine in the presence of either trypsin (20ug/ml) digested canine rough
830 microsomes (CRM) (Plumb et al., 2015) or CRM for 20 min at 32°C. 25ul translation reactions
831 were layered on 100ul of 1 M sucrose prepared in 50mM HEPES pH 7.5, 100 mM NaCl, 5mM
832 MgCl2. After sedimentation for 15 min at 20,000×g, 100 ul supernatants were carefully collected
833 and pellets were resuspended in 75ul of 1M sucrose. Aliquots of both supernatants and pellets
834 were analyzed by SDS-PAGE and autoradiography.
835
836 **Isolation of detergent-soluble and -insoluble fractions**
837 The indicated cells (0.35 x 10^6/well) were plated on a poly-lysine coated 12-well plate and
838 transiently transfected with 0.8ug of AGAL-HA or its variants. Protein expression was induced
839 with 200ng/mL doxycycline. After 24 hours of induction, the cells were harvested with 1X PBS
840 and centrifuged to collect the cell pellet. The cell pellet was resuspended with 400μL of NP40
841 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% sodium deoxycholate, and 0.5% NP40) and
842 centrifuged for 10 min at 14,000 rpm at 4°C. The detergent soluble supernatant was collected,
843 and the detergent-insoluble pellet was directly boiled in 200 ul of 2XSDS buffer for 10min. Both
844 soluble and insoluble fractions were analyzed by immunoblotting.
845
846 **Immunofluorescence**

847 BiP-3xFLAG HEK293 cells (0.12 X 10^6) were plated on 12 mm round glass coverslips (Fisher
848 Scientific) coated with 0.15 mg/mL poly-lysine in 24-well plates. The following day, cells were
849 fixed and immunostained as previously described (Sundaram et al., 2017). Rat anti-FLAG
850 antibody (BioLegend) and mouse anti-PDI antibody (Affinity Bioreagents) were used as primary
851 antibodies. Goat anti-mouse Alexa 647 and Goat anti-rat Cy2 (Jackson Immuno Research)
852 were used as secondary antibodies. The cells were imaged on a Leica scanning confocal
853 microscope as previously described (Sundaram et al., 2017).
854
855 **α-galactosidase-A activity assay**
856 Cells (0.35 x 10^6/well) were plated on poly-lysine (0.15mg/mL) coated 12-well plates and
857 transiently transfected with 800ng AGAL-FLAG versions. After 24 hours of transfection, the cells
858 were washed once with 1XPBS and harvested in 1XPBS. After centrifugation at 10000 g for 2
859 min, the cell pellet was resuspended in the 100μL of NP40 buffer (50 mM Tris, pH 7.5, 150 mM
860 NaCl, 0.5% NP40, and 0.5% deoxycholate) and incubated on ice for 20min. The lysate was
861 centrifuged for 10 min at 14000 g at 4°C. The assay was performed as described previously
862 with the following modifications (Lukas et al., 2016). 50μL of 2mM artificial substrate 4-MU-α-D-
863 galactopyranoside (Sigma) dissolved in 0.06 M phosphate citrate buffer (pH 4.7) was mixed well
864 with 1μL of the supernatant in a black 96 well plate (Greiner, #655076) for fluorescence reading,
865 and incubated for 5min at 37°C. The reaction was terminated by the addition of 200μL 1 M
866 glycine buffer (pH 10.5). The released 4-MU was measured by a fluorescence reader (TECAN)
867 using 360nm excitation filter and 465nm emission filter. The readings were subtracted from
868 empty vector-transfected cells, and AGAL variant activities were normalized to wild type AGAL
869 activity.
870
871 **Quantification and statistical analysis**
872 Quantification of bands in immunoblots and autoradiographs were performed using the ImageJ
873 software (NIH). The graphs were created by Prism7, and the error bars reflect the standard
874 error of the mean from three independent experiments.

875

876 **Online supplemental material**
877 **Fig. S1** shows the characterization of Sec63-mediated clearance of translocon-retained
878 marginally hydrophobic TMDs. **Fig. S2** demonstrates both J-domain and translocon interacting
879 regions are required for releasing translocon-retained substrates. **Fig. S3** supports the

880 mislocalization of AGAL in the cytosol of Sec63-/- cells. **Fig. S4** shows conserved weak signal
881 sequences determine Sec63 dependency. **Fig. S5** shows the weak signal sequence of AGAL
882 can efficiently target its ribosome-nascent chain to the ER membrane. **Fig. S6** shows Sec63
883 mediates BiP binding to translocating nascent chains containing both weak and strong signal
884 sequences. **Fig. 7** demonstrates the effect of signal sequences on protein translocation and
885 protein folding. **Table S1** shows proteins identified from Sec63-/-, WT Sec63, or Sec63 J-
886 domain mutant sample. **Table S2** shows secretory and membrane proteins that are enriched in
887 Sec63 J-mutant**. Table S3** shows secretory and transmembrane proteins that are enriched in
888 WT Sec63.
889
890
891 **Acknowledgements**
892 We are grateful to Jean Kanyo and TuKiet Lam from the Yale Keck Protromic facility for
893 assisting with analyzing mass spectrometry data. We thank Jacob Culver and Matthew Jordan
894 for the discussion. We thank Jacob Culver, Sang-Wook Kang, and Zai-Rong Zhang for
895 comments on the manuscript. We are thankful to Fred Gorelick for the comments and edits on
896 the manuscript. This work is funded by NIH grant R01GM117386 (M.M).
897
898 The authors declare no competing financial interests.
899
900 **Author contributions**
901 SS designed and performed most of the experiments. XL performed imaging and BiP binding
902 and dissociation experiments. MM generated BiP-3xFLAG HEK293 cells, performed BiP
903 interaction studies, and supervised the project. SS and MM wrote the manuscript with inputs
904 from XL.
905
906 **<u>Figure Legends</u>**
907
908 **Figure 1. Sec63 is required for releasing marginally hydrophobic TMDs retained at the**
909 **Sec61 translocon**
910 (A) A diagram illustrating the translocon retained with the marginally hydrophobic C-terminal
911 TMD of WRB-Venus, which is then translocated into the ER lumen and is glycosylated. The C-
912 TMD with sufficient hydrophobicity is properly inserted into the ER membrane and thereby is not
913 glycosylated. The red circle in the C-TMD indicates a charged residue.

914 (B) Cells expressing C-TMDs of substrates bearing indicated charged residues (shown in red
915 color) or hydrophobic residue (indicated in blue color) were radiolabeled and analyzed by SDS-
916 PAGE and autoradiography after immunoprecipitation of substrates with anti-GFP antibodies.
917 (C) The indicated substrates were expressed in chromosomally 2x Strep-tagged Sec61α cells,
918 lysed, and pulled down with Strep-Tactin beads and analyzed by immunoblotting for the
919 indicated antigens. The substrates were immunoblotted with anti-GFP antibodies.
920 (D) Cells co-expressing the indicated versions of BiP and the substrate were analyzed by
921 immunoblotting with anti-BiP antibodies or anti-GFP antibodies for substrates. The percent of
922 glycosylation is indicated below each panel.
923 (E) WT HEK293, Sec62-/-, or Sec63-/- cells expressing the indicated substrates were
924 radiolabeled and analyzed as in (B). Immunoblots show the knockout of either Sec62 or Sec63.
925 (F) WT HEK293 or Sec63-/- cells expressing the indicated substrates were immunoprecipitated
926 with anti-GFP antibodies and analyzed by autoradiography.
927
928 **Figure 2. Identification of translocon-retained endogenous substrates of Sec63**
929 (A) WT HEK293 or Sec63-/- cells stably expressing the indicated Sec63 constructs were
930 transfected with the C-TMD of WRB substrate carrying either lysine (K) or aspartic acid (D). The
931 transfected cells were radiolabeled and analyzed by autoradiography after immunoprecipitation
932 with anti-GFP antibodies for substrates. The immunoblot shows the expression of Sec63 from
933 the indicated cell lines.
934 (B) The indicated Sec63-/- complemented cell lines were radiolabeled and analyzed by
935 autoradiography after immunoprecipitation with anti-FLAG beads. The traces are densitometry
936 profiles of lanes and shown at the right side of the autoradiograph, illustrating a selective
937 enrichment of signals with the Sec63 J-mutant (HPD/AAA).
938 (C) A scheme describing affinity purification and identification of translocon-retained substrates
939 by mass spectrometry.
940 (D) Venn diagram of protein distribution among the three groups. The diagram was built from
941 mass spectrometry data from Table S1.
942 (E) The top 50 secretory and membrane proteins from Table S2 that were enriched in Sec63 J-
943 mutant. The red color labeled proteins are found also in the study by Schorr et al., 2020.
944
945 **Figure 3. Translocation of proteins with less hydrophobic signal sequences depend on**
946 **Sec63**

947 (A) WT HEK293 or Sec63-/- cells were transfected with AGAL-FLAG or CD44-FLAG. Cell
948 lysates were either untreated or treated with Endo H and analyzed by immunoblotting with an
949 anti-FLAG antibody for substrates and anti-Sec63 antibodies. “gly” indicates glycosylated forms.
950 “ungly” denotes non-translocated/unglycosylated forms with uncleaved SSs. “degly” indicates
951 Endo H digested de-glycosylated forms. Post-ER form of CD44 represents the Endo H resistant
952 glycosylated form.
953 (B) WT HEK293 or Sec63-/- cells stably expressing the indicated constructs were transfected
954 with AGAL-FLAG or CD44-FLAG. Cell lysates were directly analyzed by immunoblotting as in
955 (A). # denotes FLAG-tagged Sec63 proteins.
956 (C) WT HEK293, Sec63-/-, or Sec62-/- cells were transfected with plasmids expressing AGAL-
957 FLAG or CD44-FLAG, and the cell lysates were directly analyzed as in A.
958 (D) Top: Cartoons depicting the signal sequence (SS) of AGAL is either replaced with the SS of
959 prolactin (Prl-AGAL) or mutated to generate 3L-AGAL. Blue color shade denotes the
960 hydrophobic region (h-region) of the signal sequence. Hydrophobicity scores of h-regions were
961 analyzed by grand average hydropathy (GRAVY). Bottom: The indicated cell lines were
962 transfected with the mentioned constructs and analyzed by immunoblotting with an anti-FLAG
963 antibody.
964 (E) Top: A cartoon depicting the SS of CD44 that is either replaced with the SS of prolactin (Prl-
965 CD44) or mutated to generate 3L-CD44. Hydrophobicity scores were analyzed as in (D). The
966 indicated cell lines were transfected and analyzed as in D.
967
968 **Figure 4. Sec63 substrates are co-translationally translocated with an initial pause**
969 (A) Scheme depicting the protocol for proteinase K (PK) treatment for the co- or post-
970 translational protein translocation assay.
971 (B) For the co-translational protein translocation, transcripts encoding AGAL-FLAG were *in vitro*
972 translated in the rabbit reticulocyte lysate (RRL) based translation system including either buffer
973 or canine rough microsomes (CRM). After translation, samples were digested with PK and
974 analyzed by autoradiography after immunoprecipitation with anti-FLAG antibody. For the post-
975 translational translocation, transcripts encoding AGAL-FLAG were translated and centrifuged to
976 remove ribosomes. The supernatant was incubated with CRM and digested with PK and
977 analyzed after immunoprecipitation with anti-FLAG antibody.
978 (C) AGAL or Prl-AGAL transcripts lacking stop codon were *in vitro* translated in RRL including
979 either buffer or CRM. The reactions were treated with PK and analyzed as in (B). Note that

980 transcripts lacking stop codon produce nascent chains that remain attached to ribosomes
981 through a covalent peptidyl-tRNA bond.
982 (D). The ribosome-associated nascent chains (RNCs) of the indicated lengths of AGAL were
983 produced in RRL including either buffer or CRM, digested with PK, treated with RNase A to
984 remove tRNA and analyzed by autoradiography. The percentage of protease protected
985 fragments (PF) is shown below each panel. The star symbol indicates a distortion of the AGAL-
986 130 nascent chain caused by the co-migration with abundant hemoglobin from RRL.
987 (E and F) RNCs of the indicated lengths of Prl-AGAL or 3L-AGAL were produced and analyzed
988 as in D.
989 (G) Transcripts of AGAL-220 or 3L-AGAL-220 were translated in the presence of either rough
990 microsomes (RM) derived from WT HEK293 cells or Sec63-/- cells and digested with PK before
991 analyzing by autoradiography. Note that protease protected fragments were mostly disappeared
992 in samples including 1% Triton X-100.
993
994 **Figure 5. Sec63 mediates BiP binding to translocating nascent chains in the ER**
995 (A) WT HEK293 cells or HEK293 cells that contain a 3xFLAG tag at the C-terminus of
996 endogenous BiP were analyzed by immunoblotting with an anti-BiP antibody. Bag6 serves as a
997 loading control.
998 (B) WT HEK293 and BiP-3xFLAG cell lines were immunostained with anti-FLAG for BiP and
999 anti-PDI antibodies and imaged using a confocal microscope. Scale bars are 10 μm.
1000 (C) Cells expressing AGAL-HA were radiolabeled for 2 minutes and chased for the indicated
1001 time points. The cell lysates were immunoprecipitated with anti-FLAG antibodies for BiP and
1002 analyzed by autoradiography.
1003 (D) BiP-3xFLAG HEK293 or BiP-3xFLAG HEK293 Sec63-/- cells were transfected with either
1004 empty vector or AGAL-HA. Cells were radiolabeled for 2 minutes, lysed with or without ATP,
1005 and immunoprecipitated with anti-FLAG beads before analyzing by autoradiography.
1006 (E) BiP-3xFLAG HEK293 or BiP-3xFLAG HEK293 Sec63-/- cells were transfected and
1007 radiolabeled as in D. Radiolabeled cells were lysed and immunoprecipitated with either anti-
1008 FLAG antibodies for BiP or anti-HA antibodies for AGAL and analyzed by autoradiography.
1009 (F) A coomassie stained gel showing immunoprecipitated 3xFLAG tagged endogenous BiP from
1010 either BiP-3xFLAG HEK293 cells or BiP-3xFLAG HEK293 Sec63-/- cells. Ig represents
1011 immunoglobulins.
1012

1013 **Figure 6. Weak signal sequences sense BiP deficiency and prevents protein misfolding**
1014 **in the ER**
1015 (A) WT or Sec63-/- cells expressing AGAL or Prl-AGAL were radiolabeled for 30min and chased
1016 in the presence or absence of 20μM MG132 for the indicated time points. Both cells and
1017 medium were collected at each time point and analyzed by autoradiography after
1018 immunoprecipitation with an anti-HA antibody for AGAL. Secreted AGAL is given as percentage
1019 of total signals from both secreted and intracellular bands. "gly" indicates glycosylated forms.
1020 "ungly" denotes unglycosylated forms with uncleaved SSs.
1021 (B) Cells expressing the indicated constructs were lysed in NP40 buffer, and the aliquots of total
1022 lysates and detergent soluble fractions were analyzed by immunoblotting. The detergent-
1023 insoluble fractions were lysed in SDS sample buffer and analyzed by immunoblotting. BiP,
1024 Sec63, and Bag6 were blotted from soluble fractions. The percent translocated/glycosylated
1025 protein values were calculated by insoluble glycosylated signal divided by the sum of
1026 glycosylated and unglycosylated signals from both soluble and insoluble fractions. The percent
1027 of non-translocated/unglycosylated form was calculated by insoluble unglycosylated signal
1028 divided by the sum of glycosylated and unglycosylated signals from both soluble and insoluble
1029 fractions.
1030 (C) HEK293 cells were co-transfected with the indicated AGAL constructs and vector, BiP
1031 (G227D), or BiP (T37G). Cells were treated with 10μM MG132 for 6h before harvesting and
1032 analyzing as in B.
1033 (D) HEK293 cells were treated with either control siRNA or BiP siRNA and analyzed by
1034 immunoblotting for the indicated antigens.
1035 (E) BiP siRNA treated HEK293 cells were transfected with the AGAL constructs and harvested
1036 after treating the cells with 10μM MG132 for 6h. Cells were lysed and analyzed as in B.
1037
1038 **Figure 7. Model for signal sequence-dependent protein folding in the ER**
1039 The ribosome-associated nascent chain containing a weak signal sequence (SS) is efficiently
1040 recruited to the ER membrane but loosely engages the Sec61 translocon. This causes a pause
1041 in the initial translocation of ~160 amino acids in length that accumulates on the cytosolic side of
1042 the membrane. Sec63 is likely recruited to the site by sensing the accumulated nascent chains
1043 on the cytosolic side of the Sec61 translocon and activates ATPase activity of BiP through its
1044 luminal J-domain. Now, ADP-bound BiP can bind onto nascent chains and facilitate their
1045 translocation into the ER lumen, thus ensuring subsequent BiP-assisted folding of nascent
1046 chains in the ER. Under conditions of BiP deficiency, nascent chains with weak SSs still depend

1047 on BiP for translocation, and thus their BiP-assisted folding is guaranteed. However, if nascent
1048 chains fail to engage luminal BiP, their translocation is aborted, and they are mislocalized in the
1049 cytosol and eliminated by proteasomal degradation. Conversely, even under conditions of BiP
1050 deficiency, nascent chains with strong SSs are efficiently translocated in a BiP-independent
1051 manner and thus are at a high risk of misfolding and aggregating after translocation into the ER.
1052
1053 **Table S1. Proteins identified from Sec63-/-, WT Sec63, or Sec63 J-domain mutant sample.**
1054 The scaffold program was used to generate normalized spectral counts for each sample. The
1055 table shows protein hits from all three samples that have protein threshold of 5.0% FDR, at least
1056 2 peptides, and peptide threshold of 95%.
1057
1058 **Table S2. Secretory and membrane proteins that are enriched in Sec63 J-mutant.**
1059 The table shows protein hits from Sec63 J-mutant that is at least 2-fold enriched relative to WT
1060 Sec63 and 5-fold enriched relative to negative control Sec63-/- cells. The information of SSs or
1061 the 1st TMD sequences, subcellular localization, and association with human diseases were
1062 extracted from UniProt.Org. The red color labeled proteins were also found in the list of Sec63
1063 substrate candidates identified by Schorr et al, the FEBS Journal, 2020. Note that the known
1064 Sec63 interaction proteins, Sec62 and Sec61β, were manually removed from the list.
1065
1066 **Table S3. Secretory and transmembrane proteins that are enriched in WT Sec63.**
1067 The table shows protein hits from Sec63 WT that is at least 2-fold enriched relative to Sec63 J-
1068 mutant and 5-fold enriched relative to negative control Sec63-/- cells. The information of
1069 subcellular localizations is extracted from UniProt.Org. The red color labeled proteins were also
1070 found in the list of Sec63 substrate candidates identified by Schorr et al, the FEBS Journal,
1071 2020. Note that the mitochondrial proteins were manually removed from the list.
1072
1073
1074 **REFERENCES**
1075
1076 Ast, T., G. Cohen, and M. Schuldiner. 2013. A network of cytosolic factors targets SRP-
1077 independent proteins to the endoplasmic reticulum. *Cell*. 152:1134-1145.
1078 Aviram, N., and M. Schuldiner. 2017. Targeting and translocation of proteins to the endoplasmic
1079 reticulum at a glance. *Journal of Cell Science*. 130:4079-4085.
1080 Besse, W., K. Dong, J. Choi, S. Punia, S.V. Fedeles, M. Choi, A.R. Gallagher, E.B. Huang, A. Gulati,
1081 J. Knight, S. Mane, E. Tahvanainen, P. Tahvanainen, S. Sanna-Cherchi, R.P. Lifton, T.

1082 Watnick, Y.P. Pei, V.E. Torres, and S. Somlo. 2017. Isolated polycystic liver disease genes
1083 define effectors of polycystin-1 function. *J Clin Invest*. 127:1772-1785.
1084 Blobel, G., and B. Dobberstein. 1975. Transfer of proteins across membranes. I. Presence of
1085 proteolytically processed and unprocessed nascent immunoglobulin light chains on
1086 membrane-bound ribosomes of murine myeloma. *J Cell Biol*. 67:835-851.
1087 Braakman, I., and D.N. Hebert. 2013. Protein Folding in the Endoplasmic Reticulum. *Cold Spring*
1088 *Harbor Perspectives in Biology*. 5.
1089 Brodsky, J.L., J. Goeckeler, and R. Schekman. 1995. BiP and Sec63p are required for both co- and
1090 posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc Natl*
1091 *Acad Sci U S A*. 92:9643-9646.
1092 Buck, T.M., A.S. Jordahl, M.E. Yates, G.M. Preston, E. Cook, T.R. Kleyman, and J.L. Brodsky. 2017.
1093 Interactions between intersubunit transmembrane domains regulate the chaperone-
1094 dependent degradation of an oligomeric membrane protein. *Biochem J*. 474:357-376.
1095 Conti, B.J., P.K. Devaraneni, Z. Yang, L.L. David, and W.R. Skach. 2015. Cotranslational
1096 stabilization of Sec62/63 within the ER Sec61 translocon is controlled by distinct
1097 substrate-driven translocation events. *Mol Cell*. 58:269-283.
1098 Corsi, A.K., and R. Schekman. 1997. The lumenal domain of Sec63p stimulates the ATPase
1099 activity of BiP and mediates BiP recruitment to the translocon in Saccharomyces
1100 cerevisiae. *J Cell Biol*. 137:1483-1493.
1101 Davila, S., L. Furu, A.G. Gharavi, X. Tian, T. Onoe, Q. Qian, A. Li, Y. Cai, P.S. Kamath, B.F. King, P.J.
1102 Azurmendi, P. Tahvanainen, H. Kaariainen, K. Hockerstedt, O. Devuyst, Y. Pirson, R.S.
1103 Martin, R.P. Lifton, E. Tahvanainen, V.E. Torres, and S. Somlo. 2004. Mutations in SEC63
1104 cause autosomal dominant polycystic liver disease. *Nat Genet*. 36:575-577.
1105 Deshaies, R.J., S.L. Sanders, D.A. Feldheim, and R. Schekman. 1991. Assembly of yeast Sec
1106 proteins involved in translocation into the endoplasmic reticulum into a membrane-
1107 bound multisubunit complex. *Nature*. 349:806-808.
1108 Doring, K., N. Ahmed, T. Riemer, H.G. Suresh, Y. Vainshtein, M. Habich, J. Riemer, M.P. Mayer,
1109 E.P. O'Brien, G. Kramer, and B. Bukau. 2017. Profiling Ssb-Nascent Chain Interactions
1110 Reveals Principles of Hsp70-Assisted Folding. *Cell*. 170:298-+.
1111 Elofsson, A., and G. von Heijne. 2007. Membrane protein structure: prediction versus reality.
1112 *Annu Rev Biochem*. 76:125-140.
1113 Feige, M.J., and L.M. Hendershot. 2013. Quality control of integral membrane proteins by
1114 assembly-dependent membrane integration. *Mol Cell*. 51:297-309.
1115 Fewell, S.W., K.J. Travers, J.S. Weissman, and J.L. Brodsky. 2001. The action of molecular
1116 chaperones in the early secretory pathway. *Annual Review of Genetics*. 35:149-191.
1117 Fons, R.D., B.A. Bogert, and R.S. Hegde. 2003. Substrate-specific function of the translocon-
1118 associated protein complex during translocation across the ER membrane. *J Cell Biol*.
1119 160:529-539.
1120 Garman, S.C. 2007. Structure-function relationships in alpha-galactosidase A. *Acta Paediatr*.
1121 96:6-16.
1122 Gemmer, M., and F. Forster. 2020. A clearer picture of the ER translocon complex. *Journal of*
1123 *Cell Science*. 133.
1124 Germain, D.P. 2010. Fabry disease. *Orphanet J Rare Dis*. 5:30.

1125 Hegde, R.S., and S.W. Kang. 2008. The concept of translocational regulation. *J Cell Biol*. 182:225-
1126 232.
1127 Hegde, R.S., and V.R. Lingappa. 1996. Sequence-specific alteration of the ribosome-membrane
1128 junction exposes nascent secretory proteins to the cytosol. *Cell*. 85:217-228.
1129 Helenius, A., T. Marquardt, and I. Braakman. 1992. The endoplasmic reticulum as a protein-
1130 folding compartment. *Trends Cell Biol*. 2:227-231.
1131 Hessa, T., A. Sharma, M. Mariappan, H.D. Eshleman, E. Gutierrez, and R.S. Hegde. 2011. Protein
1132 targeting and degradation are coupled for elimination of mislocalized proteins. *Nature*.
1133 475:394-397.
1134 Ioannou, Y.A., D.F. Bishop, and R.J. Desnick. 1992. Overexpression of human alpha-
1135 galactosidase A results in its intracellular aggregation, crystallization in lysosomes, and
1136 selective secretion. *J Cell Biol*. 119:1137-1150.
1137 Itskanov, S., and E. Park. 2019. Structure of the posttranslational Sec protein-translocation
1138 channel complex from yeast. *Science*. 363:84-87.
1139 Jiang, Y., Z. Cheng, E.C. Mandon, and R. Gilmore. 2008. An interaction between the SRP
1140 receptor and the translocon is critical during cotranslational protein translocation. *J Cell*
1141 *Biol*. 180:1149-1161.
1142 Jung, S.J., J.E. Kim, J.H. Reithinger, and H. Kim. 2014. The Sec62-Sec63 translocon facilitates
1143 translocation of the C-terminus of membrane proteins. *J Cell Sci*. 127:4270-4278.
1144 Jungnickel, B., and T.A. Rapoport. 1995. A posttargeting signal sequence recognition event in
1145 the endoplasmic reticulum membrane. *Cell*. 82:261-270.
1146 Juszkiewicz, S., and R.S. Hegde. 2018. Quality Control of Orphaned Proteins. *Mol Cell*. 71:443-
1147 457.
1148 Kang, S.W., N.S. Rane, S.J. Kim, J.L. Garrison, J. Taunton, and R.S. Hegde. 2006. Substrate-
1149 specific translocational attenuation during ER stress defines a pre-emptive quality
1150 control pathway. *Cell*. 127:999-1013.
1151 Kida, Y., Y. Ishihara, H. Fujita, Y. Onishi, and M. Sakaguchi. 2016. Stability and flexibility of
1152 marginally hydrophobic-segment stalling at the endoplasmic reticulum translocon.
1153 *Molecular Biology of the Cell*. 27:930-940.
1154 Kida, Y., and M. Sakaguchi. 2018. Interaction mapping of the Sec61 translocon identifies two
1155 Sec61 regions interacting with hydrophobic segments in translocating chains. *Journal of*
1156 *Biological Chemistry*. 293:17050-17060.
1157 Kim, S.J., D. Mitra, J.R. Salerno, and R.S. Hegde. 2002. Signal sequences control gating of the
1158 protein translocation channel in a substrate-specific manner. *Dev Cell*. 2:207-217.
1159 Lakkaraju, A.K., R. Thankappan, C. Mary, J.L. Garrison, J. Taunton, and K. Strub. 2012. Efficient
1160 secretion of small proteins in mammalian cells relies on Sec62-dependent
1161 posttranslational translocation. *Mol Biol Cell*. 23:2712-2722.
1162 Lang, S., J. Benedix, S.V. Fedeles, S. Schorr, C. Schirra, N. Schauble, C. Jalal, M. Greiner, S.
1163 Hassdenteufel, J. Tatzelt, B. Kreutzer, L. Edelmann, E. Krause, J. Rettig, S. Somlo, R.
1164 Zimmermann, and J. Dudek. 2012. Different effects of Sec61alpha, Sec62 and Sec63
1165 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian
1166 cells. *J Cell Sci*. 125:1958-1969.

1167 Lee, S., Y. Shin, K. Kim, Y. Song, Y. Kim, and S.W. Kang. 2020. Protein Translocation Acquires
1168 Substrate Selectivity Through ER Stress-Induced Reassembly of Translocon Auxiliary
1169 Components. *Cells*. 9.
1170 Li, X., S. Sun, S. Appathurai, A. Sundaram, R. Plumb, and M. Mariappan. 2020. A Molecular
1171 Mechanism for Turning Off IRE1α Signaling during Endoplasmic Reticulum Stress. *Cell*
1172 *Rep*. 33:108563.
1173 Lingappa, V.R., D.T. Rutkowski, R.S. Hegde, and O.S. Andersen. 2002. Conformational control
1174 through translocational regulation: a new view of secretory and membrane protein
1175 folding. *Bioessays*. 24:741-748.
1176 Lukas, J., S. Scalia, S. Eichler, A.M. Pockrandt, N. Dehn, C. Cozma, A.K. Giese, and A. Rolfs. 2016.
1177 Functional and Clinical Consequences of Novel alpha-Galactosidase A Mutations in Fabry
1178 Disease. *Hum Mutat*. 37:43-51.
1179 Ma, Y., and L.M. Hendershot. 2004. ER chaperone functions during normal and stress
1180 conditions. *J Chem Neuroanat*. 28:51-65.
1181 Mali, P., L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. DiCarlo, J.E. Norville, and G.M. Church. 2013.
1182 RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*. 339:823-
1183 826.
1184 Matlack, K.E., B. Misselwitz, K. Plath, and T.A. Rapoport. 1999. BiP acts as a molecular ratchet
1185 during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell*.
1186 97:553-564.
1187 Meyer, H.A., H. Grau, R. Kraft, S. Kostka, S. Prehn, K.U. Kalies, and E. Hartmann. 2000.
1188 Mammalian Sec61 is associated with Sec62 and Sec63. *J Biol Chem*. 275:14550-14557.
1189 Misselwitz, B., O. Staeck, and T.A. Rapoport. 1998. J proteins catalytically activate Hsp70
1190 molecules to trap a wide range of peptide sequences. *Mol Cell*. 2:593-603.
1191 Nguyen, D., R. Stutz, S. Schorr, S. Lang, S. Pfeffer, H.H. Freeze, F. Forster, V. Helms, J. Dudek, and
1192 R. Zimmermann. 2018. Proteomics reveals signal peptide features determining the client
1193 specificity in human TRAP-dependent ER protein import. *Nat Commun*. 9:3765.
1194 Ojemalm, K., K.K. Halling, I. Nilsson, and G. von Heijne. 2012. Orientational preferences of
1195 neighboring helices can drive ER insertion of a marginally hydrophobic transmembrane
1196 helix. *Mol Cell*. 45:529-540.
1197 Plate, L., C.B. Cooley, J.J. Chen, R.J. Paxman, C.M. Gallagher, F. Madoux, J.C. Genereux, W.
1198 Dobbs, D. Garza, T.P. Spicer, L. Scampavia, S.J. Brown, H. Rosen, E.T. Powers, P. Walter,
1199 P. Hodder, R.L. Wiseman, and J.W. Kelly. 2016. Small molecule proteostasis regulators
1200 that reprogram the ER to reduce extracellular protein aggregation. *Elife*. 5.
1201 Plumb, R., Z.R. Zhang, S. Appathurai, and M. Mariappan. 2015. A functional link between the co-
1202 translational protein translocation pathway and the UPR. *Elife*. 4.
1203 Pobre, K.F.R., G.J. Poet, and L.M. Hendershot. 2019. The endoplasmic reticulum (ER) chaperone
1204 BiP is a master regulator of ER functions: Getting by with a little help from ERdj friends.
1205 *Journal of Biological Chemistry*. 294:2098-2108.
1206 Rapoport, T.A. 2007. Protein translocation across the eukaryotic endoplasmic reticulum and
1207 bacterial plasma membranes. *Nature*. 450:663-669.
1208 Rehm, A., P. Stern, H.L. Ploegh, and D. Tortorella. 2001. Signal peptide cleavage of a type I
1209 membrane protein, HCMV US11, is dependent on its membrane anchor. *Embo j*.
1210 20:1573-1582.

1211 Rutkowski, D.T., V.R. Lingappa, and R.S. Hegde. 2001. Substrate-specific regulation of the
1212 ribosome- translocon junction by N-terminal signal sequences. *Proc Natl Acad Sci U S A*.
1213 98:7823-7828.
1214 Rutkowski, D.T., C.M. Ott, J.R. Polansky, and V.R. Lingappa. 2003. Signal sequences initiate the
1215 pathway of maturation in the endoplasmic reticulum lumen. *J Biol Chem*. 278:30365-
1216 30372.
1217 Schorr, S., D. Nguyen, S. Hassdenteufel, N. Nagaraj, A. Cavalie, M. Greiner, P. Weissgerber, M.
1218 Loi, A.W. Paton, J.C. Paton, M. Molinari, F. Forster, J. Dudek, S. Lang, V. Helms, and R.
1219 Zimmermann. 2020a. Identification of signal peptide features for substrate specificity in
1220 human Sec62/Sec63-dependent ER protein import. *FEBS J*.
1221 Schorr, S., D. Nguyen, S. Haßdenteufel, N. Nagaraj, A. Cavalié, M. Greiner, P. Weissgerber, M.
1222 Loi, A.W. Paton, J.C. Paton, M. Molinari, F. Förster, J. Dudek, S. Lang, V. Helms, and R.
1223 Zimmermann. 2020b. Identification of signal peptide features for substrate specificity in
1224 human Sec62/Sec63-dependent ER protein import. *Febs j*. 287:4612-4640.
1225 Shan, S.O., and P. Walter. 2005. Co-translational protein targeting by the signal recognition
1226 particle. *FEBS Lett*. 579:921-926.
1227 Sharma, A., M. Mariappan, S. Appathurai, and R.S. Hegde. 2010. In vitro dissection of protein
1228 translocation into the mammalian endoplasmic reticulum. *Methods Mol Biol*. 619:339-
1229 363.
1230 Skach, W.R., L.B. Shi, M.C. Calayag, A. Frigeri, V.R. Lingappa, and A.S. Verkman. 1994. Biogenesis
1231 and Transmembrane Topology of the Chip28 Water Channel at the Endoplasmic-
1232 Reticulum. *Journal of Cell Biology*. 125:803-815.
1233 Snapp, E.L., N. McCaul, M. Quandte, Z. Cabartova, I. Bontjer, C. Kallgren, I. Nilsson, A. Land, G.
1234 von Heijne, R.W. Sanders, and I. Braakman. 2017. Structure and topology around the
1235 cleavage site regulate post-translational cleavage of the HIV-1 gp160 signal peptide.
1236 *Elife*. 6.
1237 Sun, S., and M. Mariappan. 2020. C-terminal tail length guides insertion and assembly of
1238 membrane proteins. *J Biol Chem*. 295:15498-15510.
1239 Sundaram, A., R. Plumb, S. Appathurai, and M. Mariappan. 2017. The Sec61 translocon limits
1240 IRE1α signaling during the unfolded protein response. *Elife*. 6.
1241 Vitale, M., A. Bakunts, A. Orsi, F. Lari, L. Tadè, A. Danieli, C. Rato, C. Valetti, R. Sitia, A. Raimondi,
1242 J.C. Christianson, and E. van Anken. 2019. Inadequate BiP availability defines
1243 endoplasmic reticulum stress. *eLife*. 8.
1244 Voigt, S., B. Jungnickel, E. Hartmann, and T.A. Rapoport. 1996. Signal sequence-dependent
1245 function of the TRAM protein during early phases of protein transport across the
1246 endoplasmic reticulum membrane. *J Cell Biol*. 134:25-35.
1247 Voorhees, R.M., and R.S. Hegde. 2015. Structures of the scanning and engaged states of the
1248 mammalian SRP-ribosome complex. *Elife*. 4.
1249 Voorhees, R.M., and R.S. Hegde. 2016a. Structure of the Sec61 channel opened by a signal
1250 sequence. *Science*. 351:88-91.
1251 Voorhees, R.M., and R.S. Hegde. 2016b. Toward a structural understanding of co-translational
1252 protein translocation. *Curr Opin Cell Biol*. 41:91-99.
1253 Walter, P., and G. Blobel. 1983. Preparation of microsomal membranes for cotranslational
1254 protein translocation. *Methods Enzymol*. 96:84-93.

1255 Wei, J., J.R. Gaut, and L.M. Hendershot. 1995. In vitro dissociation of BiP-peptide complexes
1256 requires a conformational change in BiP after ATP binding but does not require ATP
1257 hydrolysis. *J Biol Chem*. 270:26677-26682.
1258 White, S.H., and G. von Heijne. 2008. How translocons select transmembrane helices. *Annual*
1259 *Review of Biophysics*. 37:23-42.
1260 Willmund, F., M. del Alamo, S. Pechmann, T.T. Chen, V. Albanese, E.B. Dammer, J.M. Peng, and
1261 J. Frydman. 2013. The Cotranslational Function of Ribosome-Associated Hsp70 in
1262 Eukaryotic Protein Homeostasis. *Cell*. 152:196-209.
1263 Wu, X., C. Cabanos, and T.A. Rapoport. 2019. Structure of the post-translational protein
1264 translocation machinery of the ER membrane. *Nature*. 566:136-139.
1265 Yamamoto, Y., and T. Sakisaka. 2012. Molecular machinery for insertion of tail-anchored
1266 membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol*
1267 *Cell*. 48:387-397.
1268 Yanagitani, K., Y. Kimata, H. Kadokura, and K. Kohno. 2011. Translational pausing ensures
1269 membrane targeting and cytoplasmic splicing of XBP1u mRNA. *Science*. 331:586-589.
1270 Young, B.P., R.A. Craven, P.J. Reid, M. Willer, and C.J. Stirling. 2001. Sec63p and Kar2p are
1271 required for the translocation of SRP-dependent precursors into the yeast endoplasmic
1272 reticulum in vivo. *EMBO J*. 20:262-271.
1273 Zhang, Z.R., J.S. Bonifacino, and R.S. Hegde. 2013. Deubiquitinases sharpen substrate
1274 discrimination during membrane protein degradation from the ER. *Cell*. 154:609-622.
1275 Ziska, A., J. Tatzelt, J. Dudek, A.W. Paton, J.C. Paton, R. Zimmermann, and S. Hassdenteufel.
1276 2019. The signal peptide plus a cluster of positive charges in prion protein dictate
1277 chaperone-mediated Sec61 channel gating. *Biol Open*. 8.
1278

Figure 1. Sec63 is required for releasing marginally hydrophobic TMDs retained at the Sec61 translocon

(A) A diagram illustrating the translocon retained with the marginally hydrophobic C-terminal TMD of WRB-Venus, which is then translocated into the ER lumen and is glycosylated. The C-TMD with sufficient hydrophobicity is properly inserted into the ER membrane and thereby is not glycosylated. The red circle in the C-TMD indicates a charged residue.

(B) Cells expressing C-TMDs of substrates bearing 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 antibodies or anti-GFP antibodies for substrates. The percent of glycosylation is indicated below each panel.

(E) WT 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.

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

Figure 1

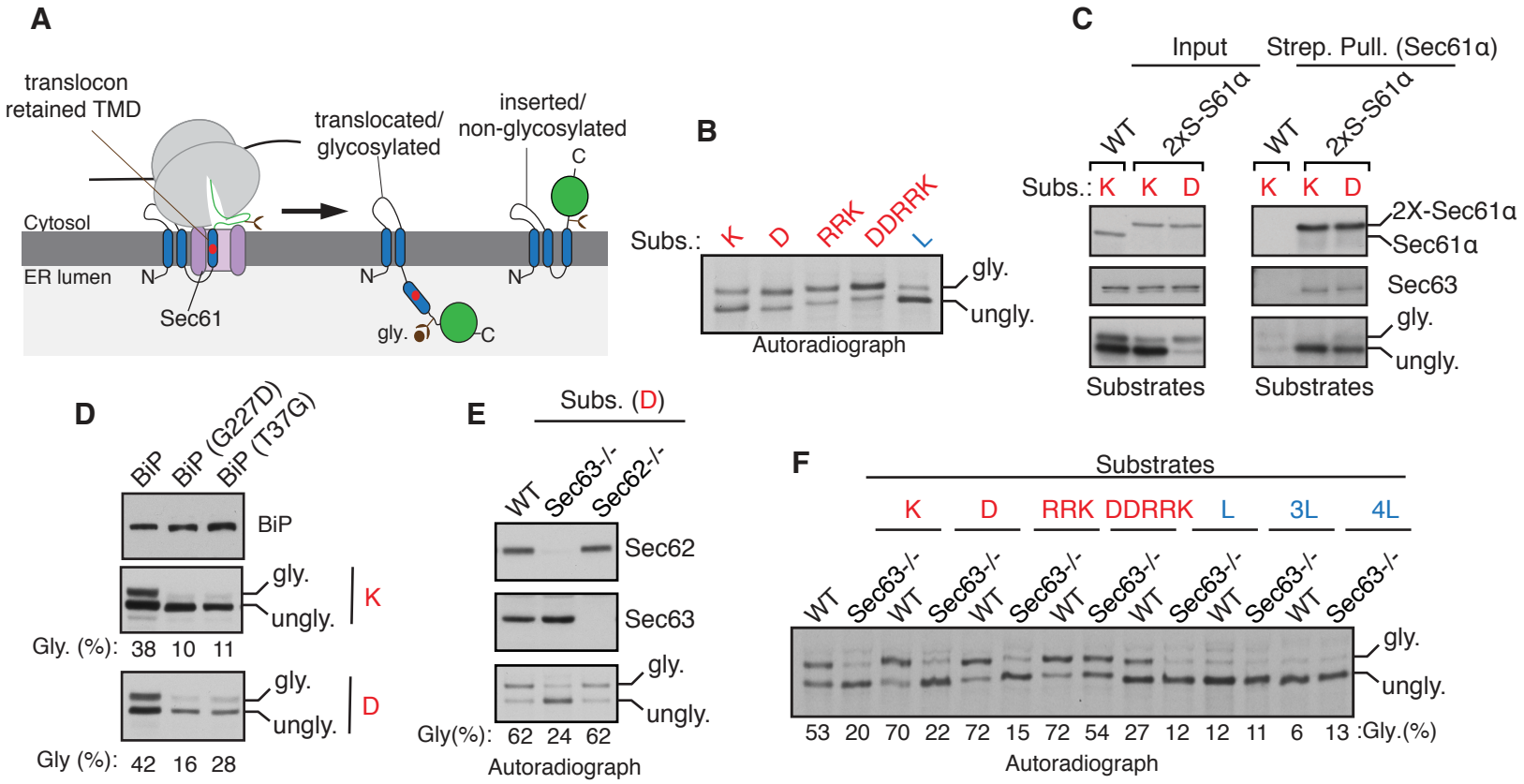


Figure 1. Sec63 is required for releasing marginally hydrophobic TMDs retained at the Sec61 translocon

Figure 2. Identification of translocon-retained endogenous substrates of Sec63

(A) WT 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.

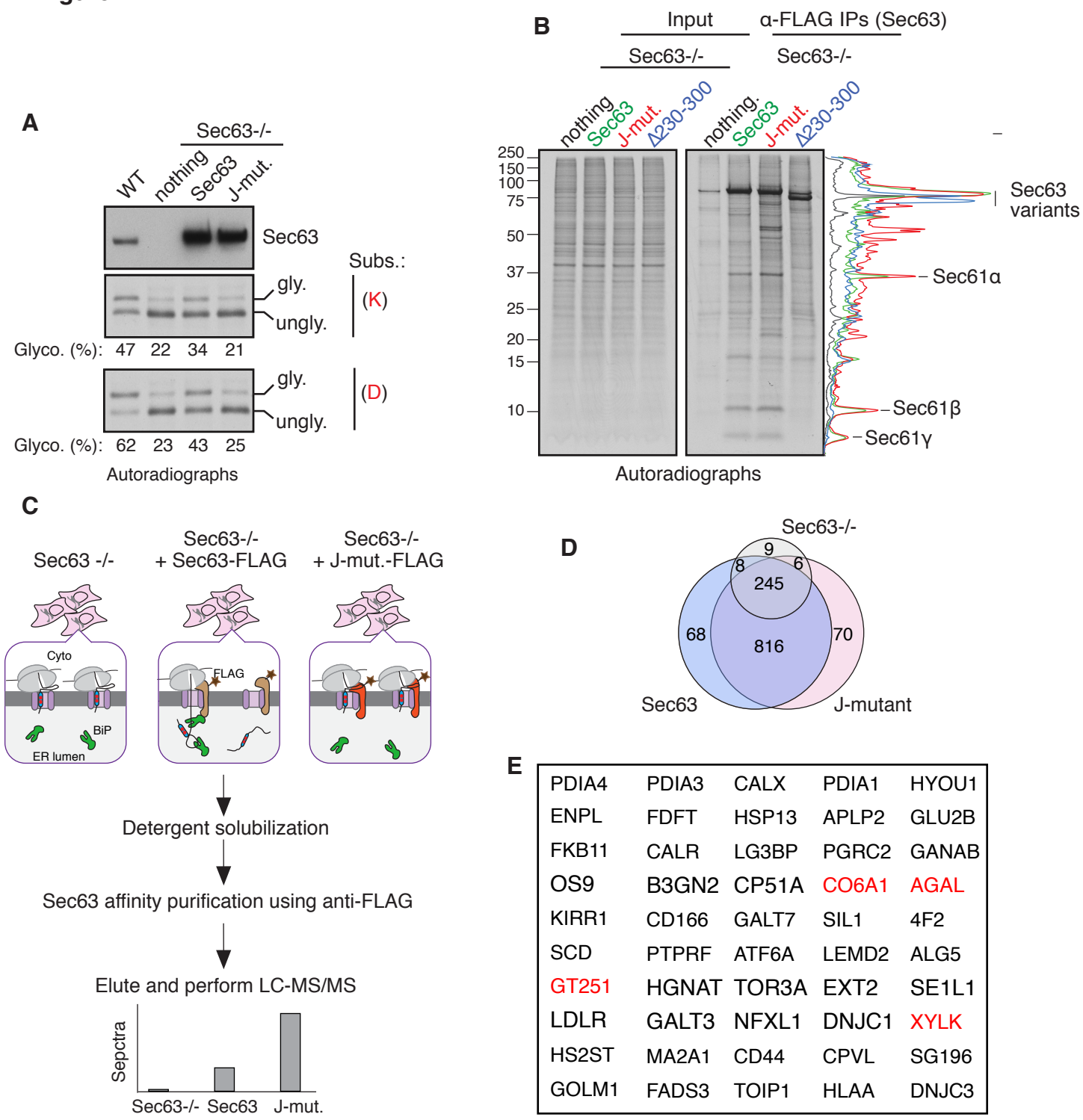
(B) The indicated Sec63-/- complemented cell lines were radiolabeled and analyzed by autoradiography after immunoprecipitation with anti-FLAG beads. The traces are densitometry profiles of lanes and shown at the right side of the autoradiograph, illustrating a selective enrichment of signals with the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

(C) A scheme describing affinity purification and identification of translocon-retained 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 top 50 secretory and membrane proteins from Table S2 that were enriched in Sec63 J-mutant. The red color labeled proteins are found also in the study by Schorr et al., 2020.

Figure 2



PDIA4	PDIA3	CALX	PDIA1	HYOU1
ENPL	FDFT	HSP13	APLP2	GLU2B
FKB11	CALR	LG3BP	PGRC2	GANAB
OS9	B3GN2	CP51A	CO6A1	AGAL
KIRR1	CD166	GALT7	SIL1	4F2
SCD	PTPRF	ATF6A	LEMD2	ALG5
GT251	HGNAT	TOR3A	EXT2	SE1L1
LDLR	GALT3	NFXL1	DNJC1	XYLK
HS2ST	MA2A1	CD44	CPVL	SG196
GOLM1	FADS3	TOIP1	HLAA	DNJC3

Figure 2. Identification of translocon-retained endogenous substrates of Sec63

Figure 3. Translocation of proteins with less hydrophobic signal sequences depend on Sec63

(A) WT HEK293 or Sec63-/- cells were transfected with AGAL-FLAG or CD44-FLAG. 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. “gly” indicates glycosylated forms. “ungly” denotes non-translocated/unglycosylated forms with uncleaved SSs. “degly” indicates Endo H digested de-glycosylated forms. Post-ER form of CD44 represents the Endo H resistant glycosylated form.

(B) WT HEK293 or Sec63-/- cells stably expressing the indicated constructs were transfected with AGAL-FLAG or CD44-FLAG. Cell lysates were directly analyzed by immunoblotting as in (A). # denotes FLAG-tagged Sec63 proteins.

(C) WT HEK293, Sec63-/-, or Sec62-/- cells were transfected with plasmids expressing AGAL-FLAG or CD44-FLAG, and the cell lysates were directly analyzed as in A.

(D) Top: Cartoons depicting the signal sequence (SS) of AGAL is either replaced with the SS of prolactin (Prl-AGAL) or mutated to generate 3L-AGAL. Blue color shade denotes the hydrophobic region (h-region) of the signal sequence. Hydrophobicity scores of h-regions were analyzed by grand average hydropathy (GRAVY). Bottom: The indicated cell lines were transfected with the mentioned constructs and analyzed by immunoblotting with an anti-FLAG antibody.

(E) Top: A cartoon depicting the SS of CD44 that is either replaced with the SS of prolactin (Prl-CD44) or mutated to generate 3L-CD44. Hydrophobicity scores were analyzed as in (D). The indicated cell lines were transfected and analyzed as in D.

Figure 3

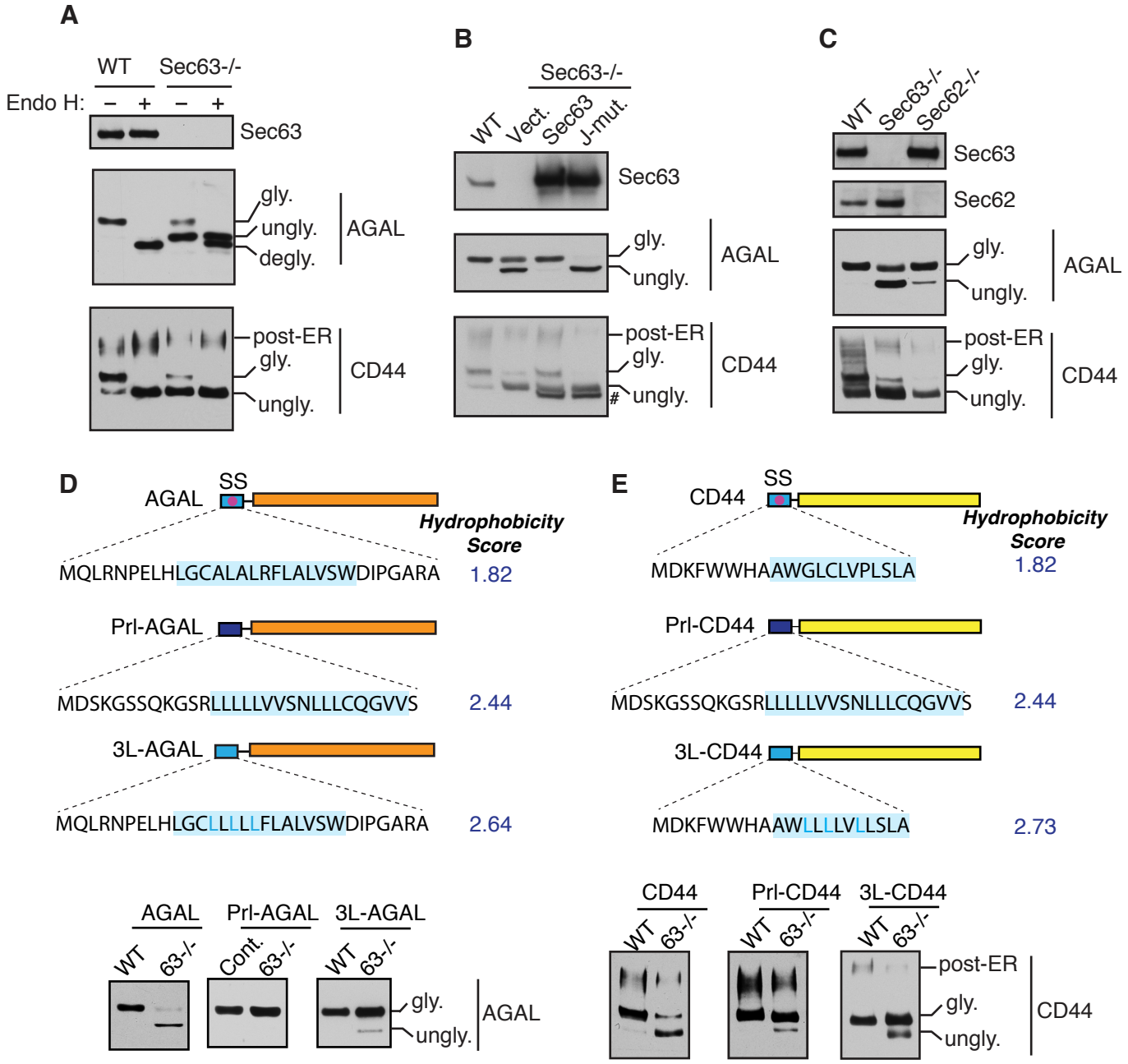


Figure 3. Translocation of proteins with less hydrophobic signal sequences depend on Sec63

Figure 4. Sec63 substrates are co-translationally translocated with an initial pause

(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. For the post-translational translocation, transcripts encoding AGAL-FLAG were translated and centrifuged to remove ribosomes. The supernatant was incubated with CRM and digested with PK and analyzed after immunoprecipitation with anti-FLAG antibody.

(C) AGAL or Prl-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 through 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, digested with PK, treated with RNase A to remove tRNA and analyzed by autoradiography. The percentage of protease protected fragments (PF) is shown below each panel. The star symbol indicates a distortion of the AGAL-130 nascent chain caused by the co-migration with abundant hemoglobin from RRL.

(E and F) RNCs of the indicated lengths of Prl-AGAL or 3L-AGAL were produced and analyzed as in D.

(G) Transcripts of AGAL-220 or 3L-AGAL-220 were translated in the presence of either rough microsomes (RM) derived from WT HEK293 cells or Sec63-/- cells and digested with PK before analyzing by autoradiography. Note that protease protected fragments were mostly disappeared in samples including 1% Triton X-100.

Figure 4

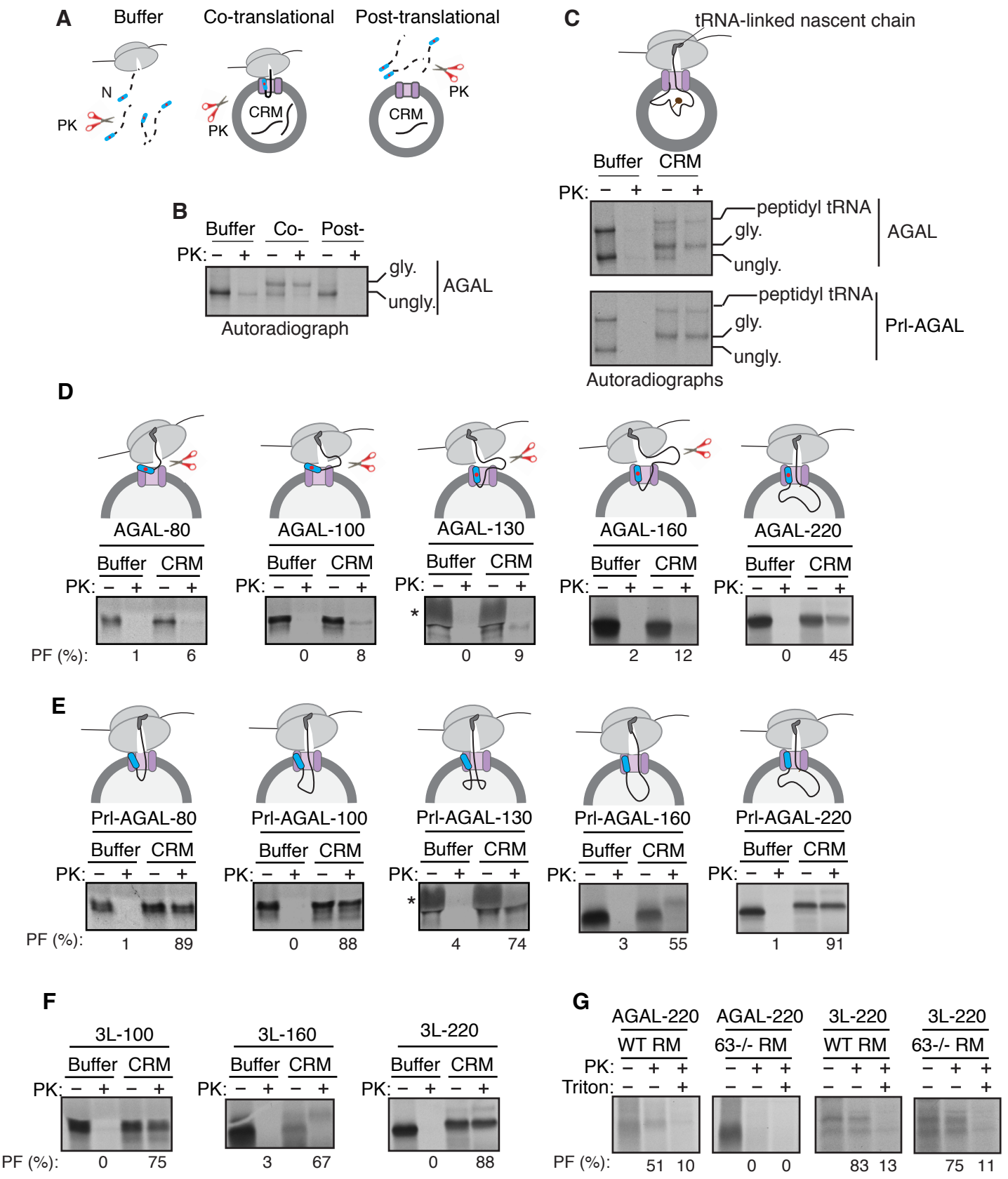


Figure 4. Sec63 substrates are co-translationally translocated with an initial pause

Figure 5. Sec63 mediates BiP binding to translocating nascent chains in the ER

(A) WT HEK293 cells or HEK293 cells that contain a 3xFLAG tag at the C-terminus of endogenous BiP were analyzed by immunoblotting with an anti-BiP antibody. Bag6 serves as a loading control.

(B) WT HEK293 and BiP-3xFLAG cell lines were immunostained with anti-FLAG for BiP and anti-PDI antibodies and imaged using a confocal microscope. Scale bars are 10 μm.

(C) Cells expressing AGAL-HA were radiolabeled for 2 minutes and chased for the indicated time points. The cell lysates were immunoprecipitated with anti-FLAG antibodies for BiP and analyzed by autoradiography.

(D) BiP-3xFLAG HEK293 or BiP-3xFLAG HEK293 Sec63-/- cells were transfected with either empty vector or AGAL-HA. Cells were radiolabeled for 2 minutes, lysed with or without ATP, and immunoprecipitated with anti-FLAG beads before analyzing by autoradiography.

(E) BiP-3xFLAG HEK293 or BiP-3xFLAG HEK293 Sec63-/- cells were transfected and radiolabeled as in D. Radiolabeled cells were lysed and immunoprecipitated with either anti-FLAG antibodies for BiP or anti-HA antibodies for AGAL and analyzed by autoradiography.

(F) A coomassie stained gel showing immunoprecipitated 3xFLAG tagged endogenous BiP from either BiP-3xFLAG HEK293 cells or BiP-3xFLAG HEK293 Sec63-/- cells. Ig represents immunoglobulins.

Figure 5

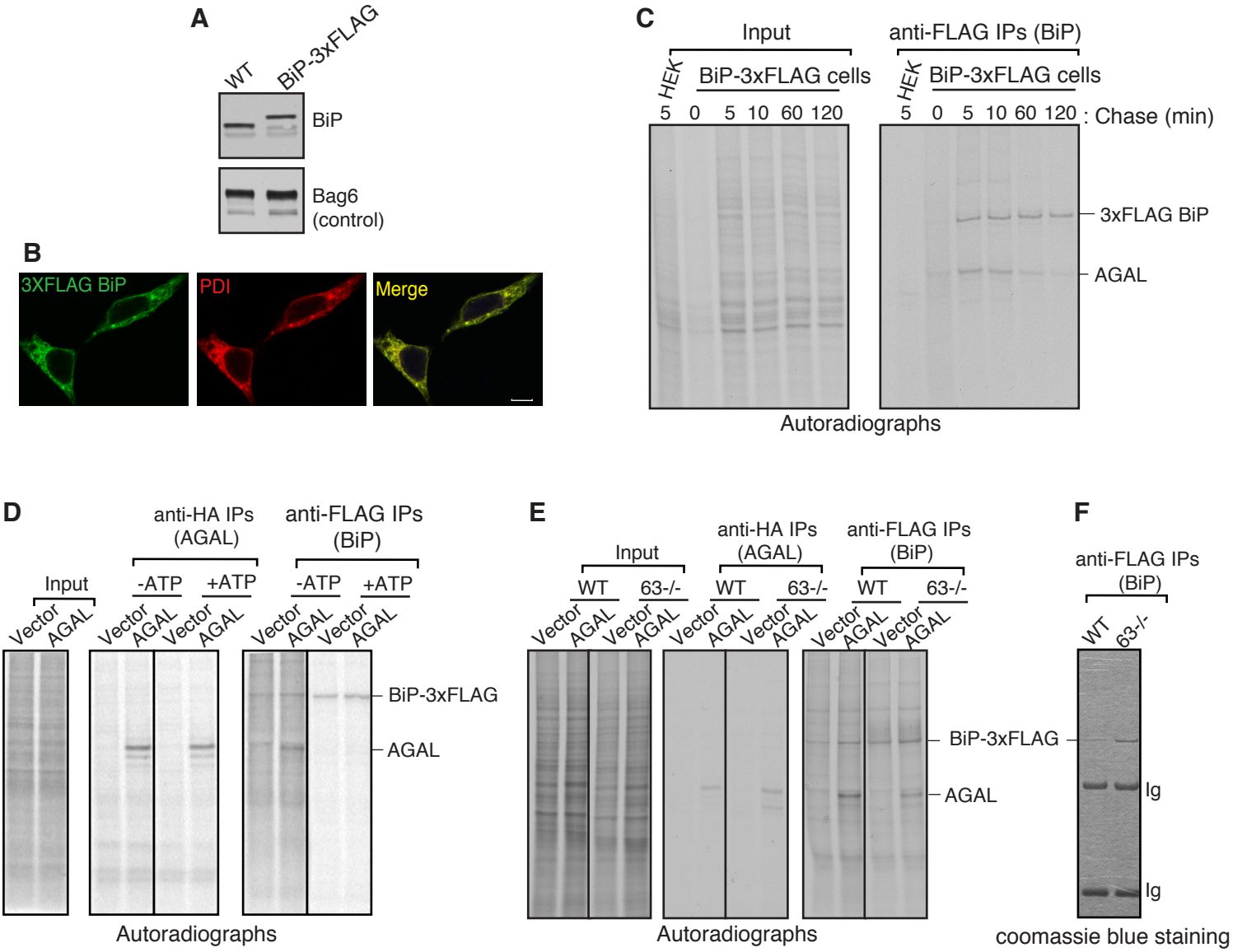


Figure 5. Sec63 mediates BiP binding to translocating nascent chains in the ER

Figure 6. Weak signal sequences sense BiP deficiency and prevent protein misfolding in the ER

(A) WT or Sec63-/- cells expressing AGAL or Prl-AGAL were radiolabeled for 30min and chased in the presence or absence of 20μM MG132 for the indicated time points. Both cells and medium were collected at each time point and analyzed by autoradiography after immunoprecipitation with an anti-HA antibody for AGAL. Secreted AGAL is given as percentage of total signals from both secreted and intracellular bands. “gly” indicates glycosylated forms. “ungly” denotes unglycosylated forms with uncleaved SSs.

(B) Cells expressing the indicated constructs were lysed in NP40 buffer, and the aliquots of total lysates and detergent soluble fractions were analyzed by immunoblotting. The detergent-insoluble fractions were lysed in SDS sample buffer and analyzed by immunoblotting. BiP, Sec63, and Bag6 were blotted from soluble fractions. The percent translocated/glycosylated protein values were calculated by insoluble glycosylated signal divided by the sum of glycosylated and unglycosylated signals from both soluble and insoluble fractions. The percent of non-translocated/unglycosylated form was calculated by insoluble unglycosylated signal divided by the sum of glycosylated and unglycosylated signals from both soluble and insoluble fractions.

(C) HEK293 cells were co-transfected with the indicated AGAL constructs and vector, BiP (G227D), or BiP (T37G). Cells were treated with 10μM MG132 for 6h before harvesting and analyzing as in B.

(D) HEK293 cells were treated with either control siRNA or BiP siRNA and analyzed by immunoblotting for the indicated antigens.

(E) BiP siRNA treated HEK293 cells were transfected with the AGAL constructs and harvested after treating the cells with 10μM MG132 for 6h. Cells were lysed and analyzed as in B.

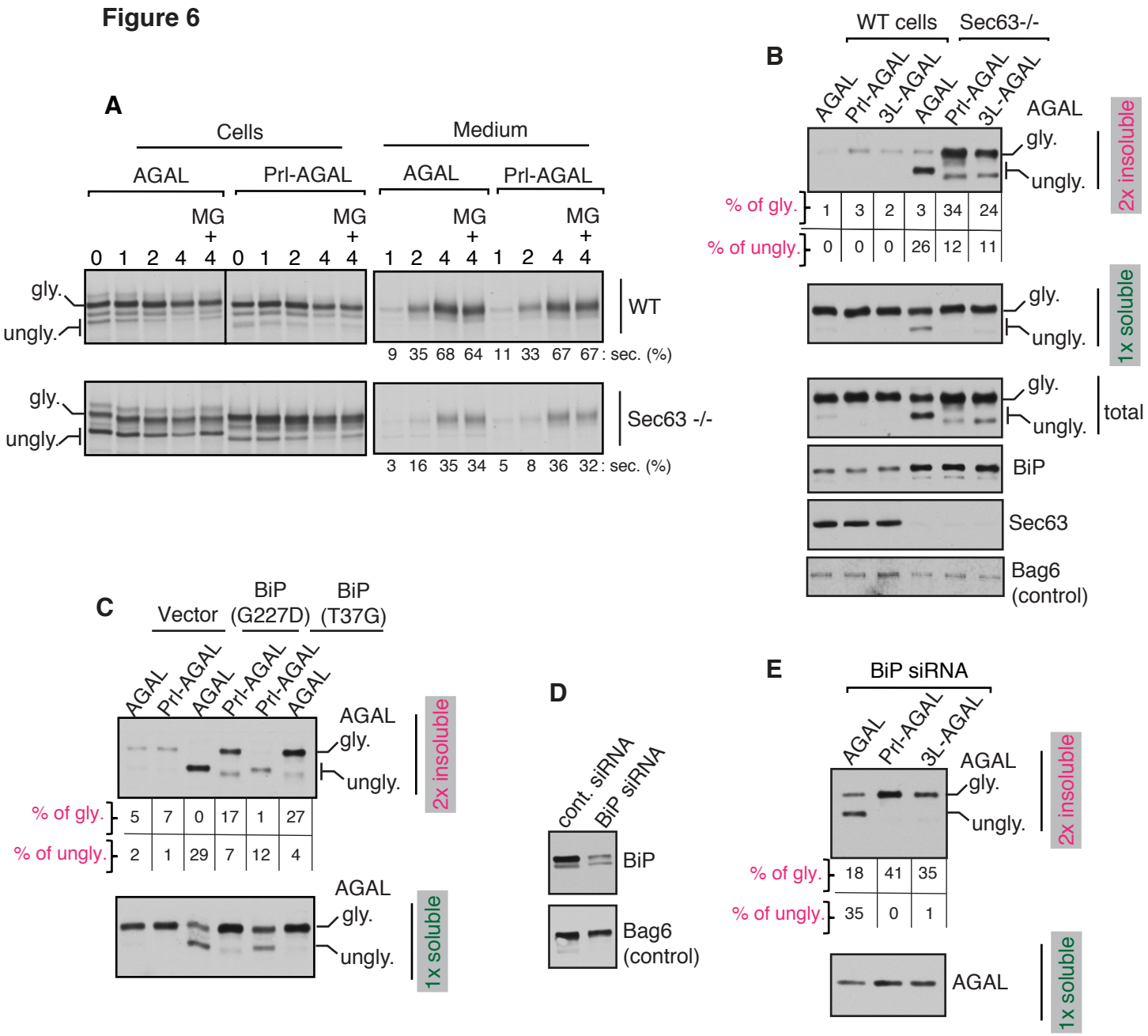


Figure 6. Weak signal sequences sense BiP deficiency and prevent protein misfolding in the ER

Figure 7. Model for signal sequence-dependent protein folding in the ER

The ribosome-associated nascent chain containing a weak signal sequence (SS) is efficiently recruited to the ER membrane but loosely engages the Sec61 translocon. This causes a pause in the initial translocation of ~160 amino acids in length that accumulates on the cytosolic side of the membrane. Sec63 is likely recruited to the site by sensing the accumulated nascent chains on the cytosolic side of the Sec61 translocon and activates ATPase activity of BiP through its luminal J-domain. Now, ADP-bound BiP can bind onto nascent chains and facilitate their translocation into the ER lumen, thus ensuring subsequent BiP-assisted folding of nascent chains in the ER. Under conditions of BiP deficiency, nascent chains with weak SSs still depend on BiP for translocation, and thus their BiP-assisted folding is guaranteed. However, if nascent chains fail to engage luminal BiP, their translocation is aborted, and they are mislocalized in the cytosol and eliminated by proteasomal degradation. Conversely, even under conditions of BiP deficiency, nascent chains with strong SSs are efficiently translocated in a BiP-independent manner and thus are at a high risk of misfolding and aggregating after translocation into the ER.

Figure 7

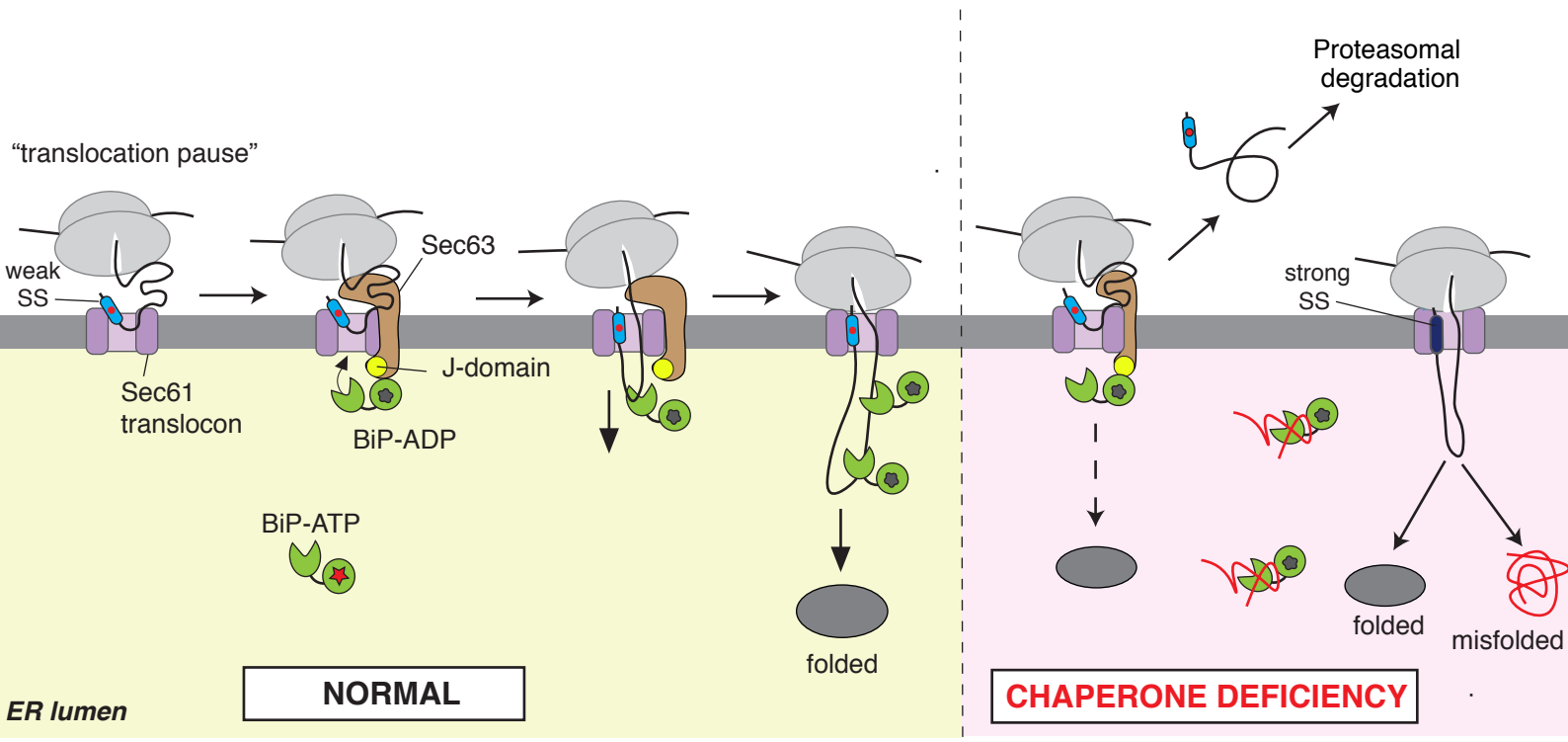


Figure 7: Model of signal sequence-dependent protein folding in the ER