McCaughey, J., & Stephens, D. J. (2019). ER-to-Golgi Transport: A Sizeable Problem. *Trends in Cell Biology*, 29(12), 940-953.
<https://doi.org/10.1016/j.tcb.2019.08.007>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.tcb.2019.08.007](https://doi.org/10.1016/j.tcb.2019.08.007)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://www.sciencedirect.com/science/article/pii/S0962892419301436#!>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

1 ER-to-Golgi transport: a sizeable problem

2

3 Janine McCaughey and David J. Stephens*

4

5 Cell Biology Laboratories, School of Biochemistry, University Walk, University of Bristol, BRISTOL,
6 BS8 1TD, UK

7 * Corresponding author:

8 Email: david.stephens@bristol.ac.uk

9 Tel: 00 44 117 331 2173

10 JM: 0000-0001-7709-2597 DJS: 0000-0001-5297-3240

11

12 Number of words: 3360

13 Keywords

14 COPII, vesicle, endoplasmic reticulum, Golgi, collagen, trafficking

15

16 **Abstract**

17 Metazoans require efficient and ordered secretion of extracellular matrix to coordinate cell and
18 tissue function. Many extracellular matrix proteins are atypically large, and their demand during
19 key stages of development presents a major challenge to the canonical secretion machinery. While
20 many of the molecular players in this pathway are known, little is understood about how they are
21 integrated in time and space. Recent advances in gene engineering and super-resolution
22 microscopy have underscored the spatio-temporal organization of the ER-Golgi interface. These
23 findings are challenging long held models of vesicular transport of large matrix proteins, such as
24 procollagen, and are implicating less well-defined carriers and direct interconnections between
25 organelles. Here, we discuss current models describing the dynamics and mechanisms of ER-Golgi
26 transport.

27

28 Glossary

29 **COPII:** a multiprotein complex that assembles in a GTP-dependent manner on the cytosolic face of
30 the endoplasmic reticulum to concentrate cargo and initiate transport carrier formation.

31 **ERES:** Endoplasmic reticulum exit site. Comprising the transitional ER, budding structures, and first
32 post-ER membranes of the ERGIC.

33 **ERGIC:** the ER-Golgi intermediate compartment that is the first post-ER compartment and plays a
34 key role in the models of large vesicle formation. It could act to maintain the physical separation of
35 ER and Golgi to prevent compartment mixing.

36 **Hsp47:** a chaperone that specifically enables the folding and assembly of procollagen and acts in
37 its ER export through interactions with both collagen and TANGO1. It can also bind to other
38 extracellular matrix proteins.

39 **KLHL12:** a ubiquitylation adaptor that was described to promote large carrier formation (see BOX
40 1).

41 **Large cargo:** we use this term to define those unusually large proteins that are packaged by the
42 COPII system but would be too large to fit within an 80 nm transport vesicle.

43 **Osteogenesis imperfecta (OI):** also known as brittle bone disease, OI is a disease of procollagen
44 biology that results in bones that break very easily. >90% of cases are caused by mutations in the
45 genes that encode type I procollagen.

46 **Procollagen:** the precursor of collagen, synthesized in the endoplasmic reticulum. There are 28
47 different types of collagen, this review focusses on the large collagens including fibrillar types I and
48 II.

49 **Sedlin:** a component of the TRAPP tethering complex that mediates fusion of post-ER membranes.
50 It binds to TANGO1 and mutations in humans cause X-linked spondyloepiphyseal dysplasia.

51 **TANGO1:** A cargo adaptor for procollagen, and likely other cargo, that spans the endoplasmic
52 reticulum membrane, physically lining cargo to the COPII coat.

53

54

55 **Oversized cargo: too big for COPII?**

56 Efficient extracellular matrix (ECM) formation is essential for normal development. The early
57 stages of metazoan embryo development, tissue organization and, for example, bone formation
58 place high demands on the secretory system. Secretory cargo like collagens, aggrecan, fibronectin,
59 fibrillin and laminins (Table 1) exit the endoplasmic reticulum (ER) and traffic to the Golgi
60 apparatus prior to reaching their destination, the extracellular space. Transport of proteins from
61 the ER to the Golgi is facilitated via coat complex type II (**COPII**, see Glossary) vesicles. These are
62 often considered to be spherical vesicles between 60-80 nm [\[1\]](#) that bud from specialized regions
63 of the ER membrane known as transitional ER and concentrate and package cargo proteins.
64 Together the transitional ER and newly formed carriers are considered to comprise ER-exit sites
65 (**ERES**) of which there are several hundred in a typical mammalian cell. Newly formed carriers
66 subsequently merge with either the Golgi apparatus or, in metazoans, the ER-Golgi intermediate
67 compartment (**ERGIC**) (BOX 1).

68 Many extracellular proteins, however, are too big to be incorporated into conventional ~80 nm
69 COPII carriers. Despite the obvious issues in size, ER to Golgi transport of these large matrix
70 proteins is COPII-dependent [\[2\]](#). Small molecule inhibitors, genetic depletion and knockout
71 experiments in both cells and animal models, as well as significant clinical data, have shown a
72 requirement for COPII proteins in the assembly of a functional ECM (Table 2, updated from [\[3\]](#)) [\[4-](#)
73 [7\]](#). Most data on the size of **large secretory cargo** proteins (Table 1) originates from electron
74 microscopy (EM) rotary shadow images from proteins extracted from fibroblast media (in the case
75 of fibrillin [\[8\]](#)) or extracellular matrix (in the case of **procollagens** [\[9\]](#)). Extracellular matrix proteins
76 are post-translationally modified and highly glycosylated, which makes it difficult to know how
77 flexible these molecules are upon exit from the ER. It is, nevertheless, assumed that fibrillar
78 collagens in particular possess a certain rigidity, which does not allow for a high flexibility [\[10\]](#). This
79 rigidity plus the large size of procollagens has led to ER-Golgi models of transport invoking
80 specialised, large COPII-dependent carriers [\[5, 11-17\]](#). However, recent data suggest that the local
81 ionic environment may lead to a greater degree of flexibility in these polymers such that they
82 should be considered as semi-flexible polymers, with a persistence length of ~100 nm, rather than
83 rigid rods [\[18\]](#). Here we discuss new findings in the field of large cargo trafficking and suggest
84 possibilities for transport by other means.

85 The case for large carriers

86 Fibrillar collagens are expressed in diverse species from sponges to humans and can thus be linked
87 with evolutionary steps leading to multicellularity [19]. The most abundant collagen in humans is a
88 fibrillar type I, which has been described as a rigid [10], rod-shaped trimer with a length of about
89 300 nm [9, 20]. Procollagen IV is the major network forming collagen of basement membranes
90 with an estimated overall length of 430 nm but with considerable flexibility arising from
91 interruptions in the triple helix [9]. Impaired collagen secretion and assembly leads to pathologies
92 like **osteogenesis imperfecta (OI)**, fibrosis, Marfan syndrome, Alport syndrome, and Ehlers-Danlos
93 syndrome [21, 22]. In order for procollagen to be secreted it needs to be assembled as homo- or
94 heterotrimers composed of three procollagen chains [23]. Triple helix formation is aided by the
95 heat shock protein **Hsp47**, also called Serpin-H1 [24], which stabilises the resulting folded trimers
96 [25-27]. Procollagen binds to Hsp47 at several locations within the cell from the ER to the *cis*-Golgi
97 [28]. Hsp47 binds to both monomeric [29] and triple helical procollagen [28, 30, 31] and contains
98 an RDEL sequence which acts as an ER-retention and retrieval signal and enables Hsp47 to be
99 recycled back from the *cis*-Golgi into the ER lumen [28]. Hsp47 does not seem to have a higher
100 affinity for specific types of procollagen, but binding to trimeric forms is favoured [30, 32, 33].
101 These trimers are then exported from the ER, via the ERGIC, to the Golgi apparatus [5, 28, 34].
102 Hsp47 can also bind to other small extracellular matrix proteins including decorin, fibromodulin,
103 and lumican [35] suggesting that it might play a role in their export from the ER.

104 Due to the presumed issue of cargo size and a lack of flexibility of fibrillar procollagen trimers, a
105 common hypothesis is that procollagen uses large COPII-dependent carriers as a means of
106 transport from the ER to Golgi [11, 12, 36-38]. COPII coats are in theory able to adopt more
107 flexible conformations that allow for a composition with up to 100 nm in diameter for spherical
108 shapes [39, 40]. Evidence for large COPII carriers was found in cells overexpressing **KLHL12** which
109 generates structures above 300 nm in diameter that co-label with Sec31A [37]. These structures
110 are few in number and COPII labelling in these cells does not display the usual fine distribution
111 throughout the cell as seen in many other cell lines [37]. In a later publication these structures co-
112 labelled with procollagen I and II, while large carriers positive for Sec31A colocalised with markers
113 for PEF1 and ALG-2 in the same cell system [12]. This was further supported by correlative light EM
114 (CLEM) showing procollagen I and Sec31A-positive structures with sizes >300 nm in KI6 cells, which
115 overexpress both procollagen I and FLAG-KLHL12 [11]. These carriers portrayed short-range
116 movement over a few microns when live cell imaging procollagen-GFP [11]. The use of large

117 carriers was further supported by *in vitro* budding assays showing large COPII structures
118 containing procollagen [11, 36].

119 Alternatively, COPII structures can assemble into tubular structures [41, 42] and other transport
120 intermediates including multi-budding structures [40, 43]. Vesicle expansion requires controlled
121 modulation of the Sar1-GTP-hydrolysis cycle. When incubated with the GTP-restricted form of Sar1
122 (Sar1-H79G), semi-intact cells produce long tubular structures extending from the ER [44]. Large
123 vesicles can tubulate *in vitro* upon impaired Sar1-GTP-hydrolysis and these are coated with Sec23-
124 Sec24 and Sec13-Sec31 [42]. Large vesicular-tubular structures might recruit fewer outer coat
125 components compared to spherical carriers, so there may be a reduced activity of GTP-hydrolysis
126 of Sar1 in tubular structure formation [45]. Similarly, loading of Sar1 with non-hydrolysable
127 analogues of GTP also leads to tubule formation of artificial liposomes [41, 45-47]. In mammalian
128 cells high level overexpression of secretory cargo can also enable the visualization of these tubules
129 [48]. It remains unclear whether cargo expression triggers tubule formation or provides enough
130 fluorophore for their visualization.

131 TANGO1 in ER-to-Golgi transport of collagens

132 The transport and Golgi organisation protein (**TANGO1**, encoded by the MIA3 gene) plays a key
133 role in ER to Golgi trafficking of large proteins and has drawn increasing attention in recent years.
134 TANGO1, an ER-resident transmembrane protein localising to ERES in mammalian cells [16], was
135 originally identified as a factor required for conventional secretion [49]. TANGO-related proteins
136 have also been implicated directly in the formation of large COPII-carriers that enable procollagen
137 transport from the ER [50-54]. TANGO1-KO mice [54] have major defects in bone formation and
138 chondrodysplasia most likely resulting from impaired secretion of collagens I, II, III, IV, VII and IX.
139 Secretion of Dumpy, a large (approximately 800 nm) protein in *Drosophila* [55], also requires
140 TANGO1 [52]. The biology of TANGO1 is further complicated by the diversity of the MIA gene
141 family (BOX 2 and Figure 1) which includes 3 members in humans, each of which undergoes
142 significant splicing. TANGO1 itself recruits other TANGO1-related proteins including cutaneous T-
143 cell lymphoma-associated antigen 5 (cTAGE5 or MIA2), TANGO1-like (TALI) and the short form of
144 TANGO1L (TANGO1S) to the ERES. These proteins also have defined roles in export of large cargo
145 molecules, especially collagens [53, 56-59]. There may be organism-specific differences in the role
146 of TANGO1 – possibly because MIA2/cTAGE5 and MIA2/TALI are not present in flies and/or
147 because flies do not express fibrillar collagens. Alternatively, loss of TANGO1 may be compensated

148 for in specific contexts through redundancy with other genes. MIA, MIA2, and MIA3 are each
149 encoded by a single gene in humans. The multiple splice forms of the *MIA* genes could potentially
150 provide the specificity to drive type-controlled recruitment of various procollagens and or large
151 cargo (Figure 1).

152 One current model proposes TANGO1 is central for large carrier formation [60]. Super-resolution
153 microscopy shows that TANGO1 assembles into rings at ERES that are positive for Sec31A and
154 contain procollagen [50, 51, 60]. These rings might act as a scaffold around which the COPII
155 machinery assembles and prevents premature coat formation [56]. The luminal SH3 domain [16]
156 of TANGO1 selectively binds large cargo proteins via Hsp47 [35, 61], and physically links these to
157 the COPII coat by direct binding to Sec23-Sec24. This interaction occurs at the Sec13-Sec31 binding
158 site on Sec23-Sec24 precluding outer coat recruitment and allowing the COPII vesicle to grow large
159 enough to incorporate bulky cargo proteins prior to GTP hydrolysis [56]. TANGO1-dependent
160 recruitment of **Sedlin**, which promotes Sar1 inactivation and release, could further help to stabilise
161 the inner COPII layer [15]. Sedlin is involved in post-ER tethering as part of the TRAPP complex.
162 This provides a means for TANGO1 to couple the fate of ER-derived carriers with their formation.

163 *In vitro* reconstitution provides a route to potentially defining large structures in more detail. Such
164 assays originally showed that TANGO1 is unique among cargo receptors as it is not incorporated
165 into the COPII vesicle itself [16]. Recent data from such experiments challenges this idea and
166 proposes a mechanism where TANGO1 and Sec12 together enter large COPII vesicles [62]. There is
167 abundant evidence that Sec12 remains at the ER and is not incorporated into (potentially only
168 small) COPII vesicles [1, 63, 64]. This can be reconciled if, in fact, any exit of Sec12 or other ER
169 resident proteins from the ER is managed by effective sorting of cargo from ER residents at the
170 point of exit and effective COPI-dependent recycling. Further, this could explain the absence of
171 obvious COPII labelling from procollagen-containing carriers emerging from the ER [65]; these
172 might represent membranes of the ERGIC that are not COPII-coated. One possibility is that Sec12
173 can leave the ER under specific circumstances such as autophagy. It is essential to consider fully
174 the way in which such *in vitro* experiments reflect cellular complexity.

175 COPII bud expansion requires additional membrane. While this membrane might come from the
176 ER itself, the current TANGO1-based model suggests that the ERGIC is the source. TANGO1 binds
177 to fusion machinery including SLY1, syntaxin 17 and 18 enabling recruitment of ERGIC membranes
178 in close apposition of COPII budding complex at ERES [13, 14]. It can also recruit the NRZ-tethering
179 complex composed of NBAS, RINT1 and ZW10 [60]. Mutations in NBAS in humans lead to

180 multisystem disorders [66] with impaired procollagen secretion resulting in an atypical form of OI
181 [67]. The role of TANGO1 in assembling this complex machine has been discussed in depth
182 elsewhere [60, 68, 69].

183 As well as being cargo receptors for collagens, MIA proteins, including TANGO1, play an important
184 role in ERES organisation [52, 70] and morphogenesis [50, 52] [70]. This role is consistent with its
185 original identification as a factor controlling ER-to-Golgi transport and Golgi organization [49].
186 Efficient organisation of ERES seems to rely on an abundance of TANGO1, since depletion of both
187 TANGO1L and TANGO1S results in dispersion of Sec16 from Sec31 [57, 70]. Individual depletion of
188 either TANGO1L or TANGO1S results in accumulation of procollagen type VII in the ER [57]. There
189 is also a suggestion that the membrane organization and cargo transport functions of TANGO1 are
190 separable [52]. One can imagine that different TANGO1-isoforms facilitate selective export of
191 various extracellular proteins at ERES. As with much of the other data, this is more consistent with
192 a general requirement for highly efficient COPII function to drive the export of procollagen from
193 the ER.

194 Alternative modes of ER-Golgi trafficking of large cargoes

195 Large COPII carriers have yet to be identified in cells expressing endogenous protein levels.
196 Primary fibroblasts do not contain any evident large COPII structures [11, 71]. Most of the
197 experiments resulting in large carriers were performed in cells overexpressing KLHL12 and or
198 procollagens in transformed cell lines [11, 12]. Furthermore, the micron-sized punctate structures
199 positive for Sec23 and procollagen in mouse osteoblasts are also positive for markers of autophagy
200 and ubiquitin [72]. These structures appear to be relatively immobile and do not traffic to the next
201 compartment, but instead are incorporated into lysosomes for degradation indicating a form of
202 noncanonical autophagy taking place at ERES to remove excess procollagen [72]. Indeed, collagen
203 can be secreted independent of the CUL3-KLHL12 complex and CUL3 depletion in human skin
204 fibroblasts blocks collagen synthesis at the level of translation, and upregulates sensors of the
205 unfolded protein response (UPR), PERK and IRE1 α [73]. The IRE1 α -XBP1s axis couples the
206 availability of nutrients, including lipids, to the COPII vesicle budding cycle via transcriptional
207 control of COPII subunit synthesis [74]. This could have implications for the formation of
208 lipoprotein-containing vesicles whose size on exit from the ER is less certain but is linked to COPII
209 directly because of disease-causing mutations in Sar1B ([75] and Table 2). These links between
210 metabolism and COPII-dependent trafficking likely impact on the role of autophagy in proteostasis

211 [\[76\]](#). Clear links have been shown between the COPII pathway and autophagy as well as
212 specifically in the handling of folded versus misfolded procollagen [\[72\]](#). One possibility is that the
213 role of ubiquitylation is more directly related to protein turnover than it is to biosynthetic
214 trafficking. There is clearly a complex integration of these pathways that we are yet to understand.
215 Therefore, it remains questionable whether any of these identified large COPII structures are
216 active cargo transporters from the ER to ERGIC or Golgi. Rather they may be acting as waystations
217 on a degradative pathway. Consistent with this, larger structures positive for markers of the ER,
218 suggests that they are not bona fide transport carriers [\[71\]](#). Our view is that these are more likely
219 to be accumulations of procollagen in “storage” or destined for degradation.

220 Despite the significant support in favour of procollagen trafficking via large COPII vesicles, there is
221 now evidence for an alternative pathway of procollagen trafficking in the absence of large carriers.
222 Early live cell imaging of procollagen-GFP showed small punctate structures tracking along long
223 range curvilinear tracks throughout the cell [\[2\]](#). Newer data using an engineered procollagen
224 reporter (with the tag inserted between N-propeptide and triple helical domain) shows that
225 procollagen can traffic to the Golgi in a COPII-dependent manner in the absence of large carriers
226 [\[71\]](#). These time-lapse movies show filling of the Golgi without the appearance of large carriers.
227 The Golgi appears to fill from the edge consistent with an intermediate compartment being
228 involved [\[71\]](#). These data suggest that procollagen could accumulate at ERES in a COPII-dependent
229 manner, followed by growth of the nascent carrier via e.g. fusion with ERGIC membranes, leading
230 to a post-ER, ERGIC element. This newly formed structure would then mature to become the first
231 bona fide cisterna of the Golgi [\[71\]](#) – similar to the cisternal maturation model where trafficking
232 between Golgi cisterna occurs without utilising vesicular intermediates [\[5, 77\]](#). The ERGIC is
233 biochemically separate from the underlying ER, requiring effective COPII-dependent sorting of
234 secretory cargo from non-residents, and COPI-dependent retrieval of and any ER residents
235 maintain compartment identity. If Sec12 does indeed transit, albeit transiently, to the ERGIC [\[62\]](#),
236 rapid recycling would prevent ongoing rounds of COPII assembly on post-ER compartments, whilst
237 retrieval of Hsp47 would prevent its accumulation in the Golgi. Importantly, formation of such an
238 intermediate would also prevent compartment mixing between the ER and the Golgi. The physical
239 proximity of most ERES to the juxtannuclear Golgi would make the “direct transfer” model
240 attractive in terms of efficiency, and potentially explain why microtubule-dependent translocation
241 of cargo over long distances is not essential for ER-to-Golgi membrane traffic. Thus, close physical
242 proximity and functional coupling of ERES and ERGIC are key features of the mammalian early

secretory pathway. This model is further supported by roles for TANGO1 in recruiting ERGIC membranes to enable expansion of carriers emerging ERES [13].

Concluding Remarks and Future Perspectives

The classical COPII pathway is considered by many to generate small 80 nm vesicles (Figure 2Ai, Key Figure) insufficient for procollagen transport. More semi-flexible procollagen polymers might be able to fit into carriers only slightly larger than 80 nm COPII vesicles (Figure 2Aii) and this has some experimental support [18], while other models propose the formation of larger carriers (Figure 2Aiii). In order to determine how large cargo proteins can be transported from the ER to the ERGIC and/or Golgi it is important to consider the structural organization of ERES and ER-ERGIC-Golgi dynamics [78]. In mammalian cells, the ER is approximately 300 – 500 nm away from ERGIC elements [79]. This means that cargo proteins with sizes of at least 300 nm will have little to no space to exist encapsulated in carriers without still being partially engulfed by ER or ERGIC. It has also been shown that *cis*-Golgi elements can reach out to budding vesicles facilitating a so called ‘hug and kiss’ mode to ‘collect’ vesicles from ERES [43, 80]. Due to the limited space in between the transitional ER and ERGIC and or Golgi elements it has also been proposed that procollagen might transfer to the next compartment via a direct tunnel-like pathway [5, 78] (Figure 2iv). Direct fusion and even tunnels between the ER and Golgi have been proposed recently [78]. We favour a model in which the ERGIC serves as the compartment boundary here preventing direct fusion of ER and Golgi and thus maintaining compartment identity.

Data also supports a model where juxta-Golgi ERES might be used preferentially for the transport of cargo [71]. One mode of operation could be that the formation of carriers is inextricably linked to their consumption. The high concentration of ERES, ERGIC, and Golgi in this region of most mammalian cells could facilitate this. Accumulation of cargo in small, peripheral ERES might result in the generation of small transport vesicles or could possibly, in the absence of sufficient ERGIC membranes, be unproductive for larger cargo.

A more flexible definition of cargo carriers in transit between the ER and Golgi seems the most likely explanation. The precise timing of COPII coat completion, initiation of uncoating and merger with the next compartment are not yet defined. Small changes in the dynamics of these processes could enable vesicles, larger tubular carriers, and even direct connections. It is highly likely that there will be subtle differences between cell types, tissues, and organisms. We favour, and would argue that the data support, a key role for the ERGIC. This could also explain why larger ERES and

274 tight functional coupling between COPII budding and the ERGIC appear necessary for procollagen
275 trafficking [81, 82]. This pathway also need not be exclusive to larger cargo. Indeed, a more direct
276 model is attractive in terms of efficiency. The exquisite sensitivity of procollagen secretion to
277 perturbation of the trafficking machinery might reflect a requirement for an optimised system that
278 is less of a necessity for small diffusible cargo.

279 Much remains to be defined concerning the export of all cargo from the ER (see “Outstanding
280 Questions”). The importance of understanding the morphology and spatial organization of the ER-
281 ERGIC-Golgi interface tests our abilities to image cells using light and electron microscopy
282 methods. It is likely that new developments in super-resolution light microscopy [83, 84] and
283 correlative light-electron microscopy will drive new understanding here. Lattice light sheet
284 imaging, serial block face scanning electron microscopy (SBF-SEM) [85], and focussed ion beam
285 milling SEM (FIB-SEM) [86] seem to hold particular promise in this area. While there remains much
286 that is unclear about the way in which procollagen and other cargoes traffic from the ER-to-Golgi,
287 their fundamental importance for metazoan development and ongoing health as well as key role
288 in major pathologies from cancer metastasis to fibrosis means that this will remain an active area
289 of study. Current advances in both super-resolution microscopy, spatially resolved proteomics,
290 and gene editing will greatly speed up the process of elucidating the mechanistic basis for
291 trafficking of large cargo proteins.

292

293 BOX 1: ER-to-Golgi transport via COPII vesicles

294 The core COPII machinery is well conserved between species [87]. The COPII machinery consists of
295 Sec12, Sec13 (each of which is encoded by a single gene) and other components for which
296 multiple isoforms exist (Sar1, Sec16, Sec23, Sec24 and Sec31; Figure 1). This additional complexity
297 likely enables efficient trafficking of diverse cargo. Sec16 dictates the spatial organisation of COPII
298 assembly on the ER membrane, the transitional ER. This, together with COPII-coated membranes
299 and ERGIC elements, forms an ERES [88-91]. COPII assembly is initiated by the ER transmembrane
300 protein and guanine nucleotide exchange factor Sec12, which binds Sec16 [92], and recruits and
301 activates Sar1-GDP [93, 94] (Figure 1A). Sar1-GTP then recruits the inner COPII components Sec23-
302 Sec24 (Figure 1B). The latter act as cargo receptors and lead to a concentration of cargo as well as
303 the recruitment of the outer COPII layer components Sec13-Sec31 [1, 95-97]. These protein
304 complexes together form the minimal machinery required to create COPII vesicles in vitro [98].
305 TFG (Trk-fused gene) interacts with Sec23 [99], as well as Sec16 [100], to further organise the
306 spatial features of COPII components by forming a matrix at the cytosolic side of the ER [81, 82,
307 100]. It plays a fundamental role in cells and mutations are known to cause hereditary spastic
308 paraplegia. Sec23-Sec24 together with Sec13-Sec31 induce GTP-hydrolysis of Sar1 [101, 102] and
309 drive scission of the COPII vesicle from the ER [46, 96] (Figure 1C) prior to uncoating of the vesicle
310 and merging with subsequent organelles (Figure 1D). The interplay between inner and outer COPII
311 coats is clearly fundamental to efficient ER export with the outer coat being required to induce
312 inner coat assembly [43].

313 Other factors modulate the COPII system, often through post-translational modification.
314 Ubiquitylation has been implicated in large vesicle formation. The CUL3-KLHL12 complex
315 monoubiquitylates Sec31A [37], which is presumed to stall the interaction of Sec31A with the
316 inner coat complex, allowing it therefore to grow in size prior to the outer coat attachment [37]. It
317 is possible that this is related to the mutually exclusive binding of Sec23 to either TANGO1 or to
318 Sec31. However, it seems that any site on Sec31 can be monoubiquitylated for this process to
319 function and so the role of ubiquitylated Sec31 remains unclear.

320 Transcriptional regulation of COPII expression, including of TANGO1, is also closely linked to the
321 secretion and assembly of ECM. CREB3L1 (which encodes the ER stress transducer OASIS)
322 regulates procollagen synthesis [103], and mutations in CREB3L1 result in OI [104]. OASIS controls
323 the expression of multiple genes involved in COPII-dependent trafficking including Sec23A and
324 Sec24D [105]. The related transcription factor CREB3L2, encoding BBF2H7, regulates Sec23A,

325 Sec23B and Sec24D expression. While it does not affect trafficking of small cargoes or laminin, loss
326 of CREB3L2/BBF2H7 leads to impaired collagen deposition resulting in craniofacial and skeletal
327 defects [106]. Furthermore, it is regulated in both a cargo-dependent manner and
328 developmentally [107] to enable the switch to procollagen type II expression by upregulating
329 expression of Sec23-Sec24, Sec13-Sec31, TANGO1, and other factors implicated in ER-to-Golgi
330 transport including Sedlin, and KLHL12.

331

BOX 2: The MIA gene family

MIA (originally defined as *melanoma inhibitory antigen*) encodes a protein of 131 amino acids. It includes a signal peptide and an SH3 domain, but lacks the transmembrane domain (TMD) found in other MIA2/MIA3 proteins and therefore has no means for retention in the ER possibly leading to secretion [108]. Expression of MIA is high in cartilage throughout development [109] and it can modulate integrin function [110]. MIA-knockout (KO) mice show defects in collagen organization, including fibre density [111], and increased cartilage regeneration [112]. Given it includes an SH3 domain, MIA might modulate matrix assembly, potentially even within the ER.

MIA2 encodes both cTAGE5 (792 aa) and TALI (1412 aa) that arise from alternative splicing. Both proteins can interact with TANGO1 and act in procollagen secretion [53]. cTAGE5 plays an essential role for localising Sec12 to ERES via direct binding but does not alter its exchange factor activity [58]. Knockdown of either TANGO1 or cTAGE5 leads to impaired trafficking of procollagen VII, showing that TANGO1 and cTAGE5 act in concert to facilitate export from the ER [53, 58]. TALI is not ubiquitously expressed (and is not evident in all vertebrates) therefore seems to have a more selective role [59]. It is present in liver and lung where fibrosis (overproduction of collagen-rich matrix) is common. TALI has a signal sequence, SH3 domain, and two predicted TMDs. The major transcript is MIA2-201 (predicted by APPRIS [113]), which consists of 653 aa and has a predicted signal sequence, as well as SH3 domain. TALI, together with TANGO1, interacts with apolipoprotein B in pre-chylomicrons and lipoproteins and drives ER exit of those proteins [59].

MIA3 encodes three different protein variants. The shortest is MIA3-202 (500 aa) which includes a signal sequence and SH3 domain but is devoid of predicted TMDs. MIA3-201 encodes a TANGO1 protein of 785 aa known as TANGO1S [57]. While this isoform lacks a strongly predicted signal sequence, it contains two TMDs and a coiled-coil domain but lacks the procollagen/Hsp47-binding SH3 domain of its longer relative, TANGO1L [54]. TANGO1L (1907 aa) annotated as MIA3-203 is the best described isoform of TANGO1 and is the one being referred to as such. Classical prediction algorithms fail to identify TMDs in TANGO1L, however, these have been demonstrated experimentally [16]. TANGO1L furthermore, has a signal sequence, SH3 domain, coiled coil domains and the PRD domain. This illustrates the need for careful experimental analysis rather than reliance on predictions alone.

362 **Acknowledgements**

363 We thank Nicola Stevenson for critical reading of the manuscript. Our lab is supported by grants
364 from the MRC (MR/P000177/1) and the BBSRC (BB/N000420/1). JM is funded by a postgraduate
365 scholarship from the University of Bristol. The funders had no role in the writing of this review.

366

367 **Conflict statement**

368 The authors declare no conflicts of interest.

369 **References**

- 370 1. Barlowe, C. et al. (1994) COPII: a membrane coat formed by Sec proteins that drive vesicle
371 budding from the endoplasmic reticulum. *Cell* 77 (6), 895-907.
- 372 2. Stephens, D.J. and Pepperkok, R. (2002) Imaging of procollagen transport reveals COPI-
373 dependent cargo sorting during ER-to-Golgi transport in mammalian cells. *J Cell Sci* 115 (Pt 6),
374 1149-60.
- 375 3. Unlu, G. et al. (2014) Trafficking mechanisms of extracellular matrix macromolecules:
376 insights from vertebrate development and human diseases. *Int J Biochem Cell Biol* 47, 57-67.
- 377 4. McCaughey, J. and Stephens, D.J. (2018) COPII-dependent ER export in animal cells:
378 adaptation and control for diverse cargo. *Histochem Cell Biol* 150 (2), 119-131.
- 379 5. Malhotra, V. and Ermann, P. (2015) The Pathway of Collagen Secretion. *Annu Rev Cell Dev*
380 *Biol* 31, 109-24.
- 381 6. Zanetti, G. et al. (2011) COPII and the regulation of protein sorting in mammals. *Nat Cell*
382 *Biol* 14 (1), 20-8.
- 383 7. Venditti, R. et al. (2014) Exiting the ER: what we know and what we don't. *Trends Cell Biol*
384 24 (1), 9-18.
- 385 8. Sakai, L.Y. et al. (1991) Purification and partial characterization of fibrillin, a cysteine-rich
386 structural component of connective tissue microfibrils. *J Biol Chem* 266 (22), 14763-70.
- 387 9. Bachinger, H.P. et al. (1982) Structural implications from an electronmicroscopic
388 comparison of procollagen V with procollagen I, pC-collagen I, procollagen IV, and a Drosophila
389 procollagen. *J Biol Chem* 257 (24), 14590-2.
- 390 10. Leblond, C.P. (1989) Synthesis and secretion of collagen by cells of connective tissue, bone,
391 and dentin. *Anat Rec* 224 (2), 123-38.
- 392 11. Gorur, A. et al. (2017) COPII-coated membranes function as transport carriers of
393 intracellular procollagen I. *J Cell Biol* 216 (6), 1745-1759.
- 394 12. McGourty, C.A. et al. (2016) Regulation of the CUL3 Ubiquitin Ligase by a Calcium-
395 Dependent Co-adaptor. *Cell* 167 (2), 525-538 e14.
- 396 13. Santos, A.J. et al. (2015) TANGO1 recruits ERGIC membranes to the endoplasmic reticulum
397 for procollagen export. *Elife* 4.
- 398 14. Nogueira, C. et al. (2014) SLY1 and Syntaxin 18 specify a distinct pathway for procollagen
399 VII export from the endoplasmic reticulum. *Elife* 3, e02784.
- 400 15. Venditti, R. et al. (2012) Sedlin controls the ER export of procollagen by regulating the Sar1
401 cycle. *Science* 337 (6102), 1668-72.
- 402 16. Saito, K. et al. (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites.
403 *Cell* 136 (5), 891-902.
- 404 17. Saito, K. and Katada, T. (2015) Mechanisms for exporting large-sized cargoes from the
405 endoplasmic reticulum. *Cell Mol Life Sci* 72 (19), 3709-20.
- 406 18. Rezaei, N. et al. (2018) Environmentally Controlled Curvature of Single Collagen Proteins.
407 *Biophys J* 115 (8), 1457-1469.
- 408 19. Ozbek, S. et al. (2010) The evolution of extracellular matrix. *Mol Biol Cell* 21 (24), 4300-5.
- 409 20. Lightfoot, S.J. et al. (1992) Type I procollagens containing substitutions of aspartate,
410 arginine, and cysteine for glycine in the pro alpha 1 (I) chain are cleaved slowly by N-proteinase,
411 but only the cysteine substitution introduces a kink in the molecule. *J Biol Chem* 267 (35), 25521-8.
- 412 21. Jobling, R. et al. (2014) The collagenopathies: review of clinical phenotypes and molecular
413 correlations. *Curr Rheumatol Rep* 16 (1), 394.
- 414 22. Kuivaniemi, H. et al. (1991) Mutations in collagen genes: causes of rare and some common
415 diseases in humans. *FASEB J* 5 (7), 2052-60.

- 416 23. Goldberg, B. et al. (1972) Precursors of collagen secreted by cultured human fibroblasts.
417 *Proc Natl Acad Sci U S A* 69 (12), 3655-9.
- 418 24. Nagata, K. et al. (1986) A major collagen-binding protein of chick embryo fibroblasts is a
419 novel heat shock protein. *J Cell Biol* 103 (1), 223-9.
- 420 25. Ishida, Y. et al. (2006) Type I collagen in Hsp47-null cells is aggregated in endoplasmic
421 reticulum and deficient in N-propeptide processing and fibrillogenesis. *Mol Biol Cell* 17 (5), 2346-
422 55.
- 423 26. Kawasaki, K. et al. (2015) Deletion of the collagen-specific molecular chaperone Hsp47
424 causes endoplasmic reticulum stress-mediated apoptosis of hepatic stellate cells. *J Biol Chem* 290
425 (6), 3639-46.
- 426 27. Marutani, T. et al. (2004) Accumulation of type IV collagen in dilated ER leads to apoptosis
427 in Hsp47-knockout mouse embryos via induction of CHOP. *J Cell Sci* 117 (Pt 24), 5913-22.
- 428 28. Satoh, M. et al. (1996) Intracellular interaction of collagen-specific stress protein HSP47
429 with newly synthesized procollagen. *J Cell Biol* 133 (2), 469-83.
- 430 29. Hosokawa, N. and Nagata, K. (2000) Procollagen binds to both prolyl 4-hydroxylase/protein
431 disulfide isomerase and HSP47 within the endoplasmic reticulum in the absence of ascorbate.
432 *FEBS Lett* 466 (1), 19-25.
- 433 30. Ito, S. and Nagata, K. (2017) Biology of Hsp47 (Serpine H1), a collagen-specific molecular
434 chaperone. *Semin Cell Dev Biol* 62, 142-151.
- 435 31. Ono, T. et al. (2012) Direct in vitro and in vivo evidence for interaction between Hsp47
436 protein and collagen triple helix. *J Biol Chem* 287 (9), 6810-8.
- 437 32. Tasab, M. et al. (2000) Hsp47: a molecular chaperone that interacts with and stabilizes
438 correctly-folded procollagen. *Embo J* 19 (10), 2204-11.
- 439 33. Koide, T. et al. (1999) Substrate recognition of collagen-specific molecular chaperone
440 HSP47. Structural requirements and binding regulation. *J Biol Chem* 274 (49), 34523-6.
- 441 34. Malhotra, V. et al. (2015) Procollagen export from the endoplasmic reticulum. *Biochem Soc*
442 *Trans* 43 (1), 104-7.
- 443 35. Ishikawa, Y. et al. (2018) The endoplasmic reticulum-resident collagen chaperone Hsp47
444 interacts with and promotes the secretion of decorin, fibromodulin, and lumican. *J Biol Chem* 293
445 (35), 13707-13716.
- 446 36. Yuan, L. et al. (2017) Cell-free Generation of COPII-coated Procollagen I Carriers. *Bio Protoc*
447 7 (22).
- 448 37. Jin, L. et al. (2012) Ubiquitin-dependent regulation of COPII coat size and function. *Nature*
449 482 (7386), 495-500.
- 450 38. Kim, S.D. et al. (2012) The [corrected] SEC23-SEC31 [corrected] interface plays critical role
451 for export of procollagen from the endoplasmic reticulum. *J Biol Chem* 287 (13), 10134-44.
- 452 39. Russell, C. and Stagg, S.M. (2010) New insights into the structural mechanisms of the COPII
453 coat. *Traffic* 11 (3), 303-10.
- 454 40. Stagg, S.M. et al. (2008) Structural basis for cargo regulation of COPII coat assembly. *Cell*
455 134 (3), 474-84.
- 456 41. Bacia, K. et al. (2011) Multibudded tubules formed by COPII on artificial liposomes. *Sci Rep*
457 1, 17.
- 458 42. Zanetti, G. et al. (2013) The structure of the COPII transport-vesicle coat assembled on
459 membranes. *Elife* 2, e00951.
- 460 43. Hutchings, J. and Zanetti, G. (2019) Coat flexibility in the secretory pathway: a role in
461 transport of bulky cargoes. *Curr Opin Cell Biol* 59, 104-111.
- 462 44. Aridor, M. et al. (2001) The Sar1 GTPase coordinates biosynthetic cargo selection with
463 endoplasmic reticulum export site assembly. *J Cell Biol* 152 (1), 213-29.

- 464 45. Saito, K. et al. (2017) Regulation of the Sar1 GTPase Cycle Is Necessary for Large Cargo
465 Secretion from the Endoplasmic Reticulum. *Front Cell Dev Biol* 5, 75.
- 466 46. Bielli, A. et al. (2005) Regulation of Sar1 NH2 terminus by GTP binding and hydrolysis
467 promotes membrane deformation to control COPII vesicle fission. *J Cell Biol* 171 (6), 919-24.
- 468 47. Lee, M.C. et al. (2005) Sar1p N-terminal helix initiates membrane curvature and completes
469 the fission of a COPII vesicle. *Cell* 122 (4), 605-17.
- 470 48. Simpson, J.C. et al. (2006) Biogenesis of tubular ER-to-Golgi transport intermediates. *Mol*
471 *Biol Cell* 17 (2), 723-37.
- 472 49. Bard, F. et al. (2006) Functional genomics reveals genes involved in protein secretion and
473 Golgi organization. *Nature* 439 (7076), 604-7.
- 474 50. Liu, M. et al. (2017) Tango1 spatially organizes ER exit sites to control ER export. *J Cell Biol*
475 216 (4), 1035-1049.
- 476 51. Raote, I. et al. (2017) TANGO1 assembles into rings around COPII coats at ER exit sites. *J*
477 *Cell Biol* 216 (4), 901-909.
- 478 52. Rios-Barrera, L.D. et al. (2017) Dual function for Tango1 in secretion of bulky cargo and in
479 ER-Golgi morphology. *Proc Natl Acad Sci U S A* 114 (48), E10389-E10398.
- 480 53. Saito, K. et al. (2011) cTAGE5 mediates collagen secretion through interaction with
481 TANGO1 at endoplasmic reticulum exit sites. *Mol Biol Cell* 22 (13), 2301-8.
- 482 54. Wilson, D.G. et al. (2011) Global defects in collagen secretion in a Mia3/TANGO1 knockout
483 mouse. *J Cell Biol* 193 (5), 935-51.
- 484 55. Wilkin, M.B. et al. (2000) Drosophila Dumpy is a gigantic extracellular protein required to
485 maintain tension at epidermal-cuticle attachment sites. *Current Biology* 10 (10), 559-567.
- 486 56. Ma, W. and Goldberg, J. (2016) TANGO1/cTAGE5 receptor as a polyvalent template for
487 assembly of large COPII coats. *Proc Natl Acad Sci U S A* 113 (36), 10061-6.
- 488 57. Maeda, M. et al. (2016) Distinct isoform-specific complexes of TANGO1 cooperatively
489 facilitate collagen secretion from the endoplasmic reticulum. *Mol Biol Cell* 27 (17), 2688-96.
- 490 58. Saito, K. et al. (2014) Concentration of Sec12 at ER exit sites via interaction with cTAGE5 is
491 required for collagen export. *J Cell Biol* 206 (6), 751-62.
- 492 59. Santos, A.J. et al. (2016) TANGO1 and Mia2/cTAGE5 (TALI) cooperate to export bulky pre-
493 chylomicrons/VLDLs from the endoplasmic reticulum. *J Cell Biol* 213 (3), 343-54.
- 494 60. Raote, I. et al. (2018) TANGO1 builds a machine for collagen export by recruiting and
495 spatially organizing COPII, tethers and membranes. *Elife* 7.
- 496 61. Ishikawa, Y. et al. (2016) Intracellular mechanisms of molecular recognition and sorting for
497 transport of large extracellular matrix molecules. *Proc Natl Acad Sci U S A* 113 (41), E6036-E6044.
- 498 62. Yuan, L. et al. (2018) TANGO1 and SEC12 are copackaged with procollagen I to facilitate the
499 generation of large COPII carriers. *Proc Natl Acad Sci* 115 (52), E12255.
- 500 63. Futai, E. et al. (2004) GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on
501 synthetic liposomes. *EMBO J* 23 (21), 4146-55.
- 502 64. Iwasaki, H. et al. (2017) Microscopy analysis of reconstituted COPII coat polymerization and
503 Sec16 dynamics. *J Cell Sci* 130 (17), 2893-2902.
- 504 65. Mironov, A.A. et al. (2003) ER-to-Golgi carriers arise through direct en bloc protrusion and
505 multistage maturation of specialized ER exit domains. *Dev Cell* 5 (4), 583-94.
- 506 66. Segarra, N.G. et al. (2015) NBAS mutations cause a multisystem disorder involving bone,
507 connective tissue, liver, immune system, and retina. *Am J Med Genet A* 167A (12), 2902-12.
- 508 67. Balasubramanian, M. et al. (2017) Compound heterozygous variants in NBAS as a cause of
509 atypical osteogenesis imperfecta. *Bone* 94, 65-74.
- 510 68. Saito, K. and Maeda, M. (2019) Not just a cargo receptor for large cargoes; an emerging
511 role of TANGO1 as an organizer of ER exit sites. *J Biochem*.
- 512 69. Peotter, J. et al. (2019) COPII-mediated trafficking at the ER/ERGIC interface. *Traffic*.

513 70. Maeda, M. et al. (2017) TANGO1 recruits Sec16 to coordinately organize ER exit sites for
514 efficient secretion. *J Cell Biol* 216 (6), 1731-1743.

515 71. McCaughey, J. et al. (2019) ER-to-Golgi trafficking of procollagen in the absence of large
516 carriers. *J Cell Biol* 218 (3), 929-948.

517 72. Omari, S. et al. (2018) Noncanonical autophagy at ER exit sites regulates procollagen
518 turnover. *Proc Natl Acad Sci U S A* 115 (43), E10099-E10108.

519 73. Kim, K. et al. (2018) Cullin3-RING ubiquitin ligases are intimately linked to the unfolded
520 protein response of the endoplasmic reticulum.

521 74. Liu, L. et al. (2019) Coupling of COPII vesicle trafficking to nutrient availability by the
522 IRE1alpha-XBP1s axis. *Proc Natl Acad Sci U S A* 116 (24), 11776-11785.

523 75. Jones, B. et al. (2003) Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid
524 absorption disorders. *Nat Genet* 34 (1), 29-31.

525 76. Wang, J. et al. (2017) Autophagosome formation: Where the secretory and autophagy
526 pathways meet. *Autophagy* 13 (5), 973-974.

527 77. Bonfanti, L. et al. (1998) Procollagen traverses the Golgi stack without leaving the lumen of
528 cisternae: evidence for cisternal maturation. *Cell* 95 (7), 993-1003.

529 78. Raote, I. and Malhotra, V. (2019) Protein transport by vesicles and tunnels. *J Cell Biol* 218
530 (3), 737-739.

531 79. Hanna, M.G. et al. (2018) Membrane Transport at an Organelle Interface in the Early
532 Secretory Pathway: Take Your Coat Off and Stay a While: Evolution of the metazoan early
533 secretory pathway. *Bioessays*, e1800004.

534 80. Kurokawa, K. et al. (2014) Contact of cis-Golgi with ER exit sites executes cargo capture and
535 delivery from the ER. *Nat Commun* 5, 3653.

536 81. McCaughey, J. et al. (2016) TFG Promotes Organization of Transitional ER and Efficient
537 Collagen Secretion. *Cell Rep* 15 (8), 1648-59.

538 82. Johnson, A. et al. (2015) TFG clusters COPII-coated transport carriers and promotes early
539 secretory pathway organization. *EMBO J* 34 (6), 811-27.

540 83. Guo, Y. et al. (2018) Visualizing Intracellular Organelle and Cytoskeletal Interactions at
541 Nanoscale Resolution on Millisecond Timescales. *Cell* 175 (5), 1430-1442 e17.

542 84. Schermelleh, L. et al. (2019) Super-resolution microscopy demystified. *Nat Cell Biol* 21 (1),
543 72-84.

544 85. Russell, M.R. et al. (2017) 3D correlative light and electron microscopy of cultured cells
545 using serial blockface scanning electron microscopy. *J Cell Sci* 130 (1), 278-291.

546 86. Xu, C.S. et al. (2017) Enhanced FIB-SEM systems for large-volume 3D imaging. *Elife* 6.

547 87. Schlacht, A. and Dacks, J.B. (2015) Unexpected ancient paralogs and an evolutionary model
548 for the COPII coat complex. *Genome Biol Evol* 7 (4), 1098-109.

549 88. Bannykh, S.I. and Balch, W.E. (1997) Membrane dynamics at the endoplasmic reticulum-
550 Golgi interface. *Journal of Cell Biology* 138 (1), 1-4.

551 89. Hughes, H. et al. (2009) Organisation of human ER-exit sites: requirements for the
552 localisation of Sec16 to transitional ER. *Journal of Cell Science* 122 (Pt 16), 2924-34.

553 90. Watson, P. et al. (2006) Sec16 defines endoplasmic reticulum exit sites and is required for
554 secretory cargo export in mammalian cells. *Traffic* 7 (12), 1678-87.

555 91. Connerly, P.L. et al. (2005) Sec16 is a determinant of transitional ER organization. *Curr Biol*
556 15 (16), 1439-47.

557 92. Montegna, E.A. et al. (2012) Sec12 binds to Sec16 at transitional ER sites. *PLoS One* 7 (2),
558 e31156.

559 93. Barlowe, C. and Schekman, R. (1993) SEC12 encodes a guanine-nucleotide-exchange factor
560 essential for transport vesicle budding from the ER. *Nature* 365 (6444), 347-9.

561 94. Nakano, A. and Muramatsu, M. (1989) A novel GTP-binding protein, Sar1p, is involved in
562 transport from the endoplasmic reticulum to the Golgi apparatus. *J Cell Biol* 109 (6 Pt 1), 2677-91.

563 95. Miller, E. et al. (2002) Cargo selection into COPII vesicles is driven by the Sec24p subunit.
564 *EMBO J* 21 (22), 6105-13.

565 96. Fromme, J.C. et al. (2007) The genetic basis of a craniofacial disease provides insight into
566 COPII coat assembly. *Dev Cell* 13 (5), 623-34.

567 97. Bi, X. et al. (2007) Insights into COPII coat nucleation from the structure of Sec23.Sar1
568 complexed with the active fragment of Sec31. *Dev Cell* 13 (5), 635-45.

569 98. Matsuoka, K. et al. (1998) COPII-coated vesicle formation reconstituted with purified coat
570 proteins and chemically defined liposomes. *Cell* 93 (2), 263-75.

571 99. Hanna, M.G.t. et al. (2017) TFG facilitates outer coat disassembly on COPII transport
572 carriers to promote tethering and fusion with ER-Golgi intermediate compartments. *Proc Natl*
573 *Acad Sci U S A* 114 (37), E7707-E7716.

574 100. Witte, K. et al. (2011) TFG-1 function in protein secretion and oncogenesis. *Nat Cell Biol* 13
575 (5), 550-8.

576 101. Antonny, B. et al. (2001) Dynamics of the COPII coat with GTP and stable analogues. *Nat*
577 *Cell Biol* 3 (6), 531-7.

578 102. Yoshihisa, T. et al. (1993) Requirement for a GTPase-activating protein in vesicle budding
579 from the endoplasmic reticulum. *Science* 259 (5100), 1466-8.

580 103. Murakami, T. et al. (2009) Signalling mediated by the endoplasmic reticulum stress
581 transducer OASIS is involved in bone formation. *Nat Cell Biol* 11 (10), 1205-11.

582 104. Symoens, S. et al. (2013) Deficiency for the ER-stress transducer OASIS causes severe
583 recessive osteogenesis imperfecta in humans. *Orphanet J Rare Dis* 8, 154.

584 105. Guillemyn, B. et al. (2019) A homozygous pathogenic missense variant broadens the
585 phenotypic and mutational spectrum of CREB3L1-related osteogenesis imperfecta. *Hum Mol*
586 *Genet* 28 (11), 1801-1809.

587 106. Melville, D.B. et al. (2011) The feelgood mutation in zebrafish dysregulates COPII-
588 dependent secretion of select extracellular matrix proteins in skeletal morphogenesis. *Dis Model*
589 *Mech* 4 (6), 763-76.

590 107. Ishikawa, T. et al. (2017) UPR transducer BBF2H7 allows export of type II collagen in a
591 cargo- and developmental stage-specific manner. *J Cell Biol* 216 (6), 1761-1774.

592 108. Blesch, A. et al. (1994) Cloning of a novel malignant melanoma-derived growth-regulatory
593 protein, MIA. *Cancer Res* 54 (21), 5695-701.

594 109. Bosserhoff, A.K. et al. (1997) Mouse CD-RAP/MIA gene: structure, chromosomal
595 localization, and expression in cartilage and chondrosarcoma. *Dev Dyn* 208 (4), 516-25.

596 110. Bauer, R. et al. (2006) Regulation of integrin activity by MIA. *J Biol Chem* 281 (17), 11669-
597 77.

598 111. Moser, M. et al. (2002) Ultrastructural cartilage abnormalities in MIA/CD-RAP-deficient
599 mice. *Mol Cell Biol* 22 (5), 1438-45.

600 112. Schmid, R. et al. (2010) Enhanced cartilage regeneration in MIA/CD-RAP deficient mice. *Cell*
601 *Death Dis* 1, e97.

602 113. Rodriguez, J.M. et al. (2018) APPRIS 2017: principal isoforms for multiple gene sets. *Nuc*
603 *Acids Res* 46 (D1), D213-D217.

604 114. Hofmann, H. et al. (1984) Localization of flexible sites in thread-like molecules from
605 electron micrographs. Comparison of interstitial, basement membrane and intima collagens. *J Mol*
606 *Biol* 172 (3), 325-43.

607 115. Morris, N.P. et al. (1986) The tissue form of type VII collagen is an antiparallel dimer. *J Biol*
608 *Chem* 261 (12), 5638-44.

609 116. Paulsson, M. et al. (1987) Extended and globular protein domains in cartilage
610 proteoglycans. *Biochem J* 245 (3), 763-72.

611 117. Erickson, H.P. et al. (1981) Fibronectin molecule visualized in electron microscopy: a long,
612 thin, flexible strand. *J Cell Biol* 91 (3 Pt 1), 673-78.

613 118. Engel, J. et al. (1981) Shapes, domain organizations and flexibility of laminin and
614 fibronectin, two multifunctional proteins of the extracellular matrix. *J Mol Biol* 150 (1), 97-120.

615 119. Boyadjiev, S.A. et al. (2006) Cranio-lenticulo-sutural dysplasia is caused by a SEC23A
616 mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nat Genet* 38, 1192-
617 1197.

618 120. Lang, M.R. et al. (2006) Secretory COPII coat component Sec23a is essential for craniofacial
619 chondrocyte maturation. *Nat Genet* 38 (10), 1198-203.

620 121. Cox, N.J. et al. (2018) Dynamic Glycosylation Governs the Vertebrate COPII Protein
621 Trafficking Pathway. *Biochemistry* 57 (1), 91-107.

622 122. Garbes, L. et al. (2015) Mutations in SEC24D, encoding a component of the COPII
623 machinery, cause a syndromic form of osteogenesis imperfecta. *Am J Hum Genet* 96 (3), 432-9.

624 123. Sarmah, S. et al. (2010) Sec24D-dependent transport of extracellular matrix proteins is
625 required for zebrafish skeletal morphogenesis. *PLoS One* 5 (4), e10367.

626 124. Ohisa, S. et al. (2010) sec24d encoding a component of COPII is essential for vertebra
627 formation, revealed by the analysis of the medaka mutant, vbi. *Dev Biol* 342 (1), 85-95.

628 125. Townley, A.K. et al. (2008) Efficient coupling of Sec23-Sec24 to Sec13-Sec31 drives COPII-
629 dependent collagen secretion and is essential for normal craniofacial development. *J Cell Sci* 121
630 (Pt 18), 3025-34.

631 126. Niu, X. et al. (2012) Sec13 safeguards the integrity of the endoplasmic reticulum and
632 organogenesis of the digestive system in zebrafish. *Dev Biol* 367 (2), 197-207.

633 127. Dannoura, A.H. et al. (1999) Anderson's disease: exclusion of apolipoprotein and
634 intracellular lipid transport genes. *Arterioscler Thromb Vasc Biol* 19 (10), 2494-508.

635 128. Beetz, C. et al. (2013) Inhibition of TFG function causes hereditary axon degeneration by
636 impairing endoplasmic reticulum structure. *Proc Natl Acad Sci U S A* 110 (13), 5091-6.

637

638 *The following tables are included here only to integrate the reference list.*

639 **Figure legends**

640 **Graphical abstract caption.**

641 **COPII-dependent transport modes of small and large cargo from the ER.** ER export of small cargo
642 proteins is facilitated by conventional COPII vesicles. Large cargo proteins, like collagens, destined
643 for transport to the ER-Golgi intermediate compartment (ERGIC) and further to the Golgi and
644 eventually extracellular space require alternative COPII-dependent pathways. Possible transport
645 modes include direct connections between ER and ERGIC, large COPII carrier formation and
646 medium sized COPII carriers with slightly more flexible procollagen. Transport intermediates can
647 originate from tubular and vesicular structures.

648

649 **Figure I, Box 1. Transport in the early secretory pathway.** Secretory cargo needs to be
650 transported from the ER via the ERGIC to the Golgi apparatus and subsequently the extracellular
651 space. The zoomed panel shows COPII assembly at an ER exit site initiated by Sec12 and Sar1 (A)
652 leading to the recruitment of inner layer (Sec23-Sec24) and outer layer (Sec13-Sec31) COPII
653 components with assistance of Sec16 and TFG that act to spatially organise the coat proteins (B).
654 After cargo concentration and vesicle growth, Sar1-GTP hydrolysis triggers the release from the ER
655 (C). Resulting COPII vesicles have a diameter of 60 – 80 nm. Shortly after budding the vesicle
656 undergoes uncoating (D), prior to fusion with the ERGIC/Golgi compartments.

657

658 **Figure 1: Motifs in proteins of the COPII machinery and TANGO1-related isoforms.** Motifs shown
659 are those predicted by smart.embl.de and or denoted on UniProt.org. The COPII machinery is
660 shown on the left and proteins of the MIA gene family on the right. Motifs shown include
661 transmembrane domains (TMD), proline-rich domains (PRD), Phox and Bem1 (PB1) domains,
662 coiled-coil dimerization (CCD) domains and Tether of ERGIC and ER (TEER) domains.

663

664 **Figure 2, Key Figure: Conventional transport of small cargo and possible transport modes for**
665 **large secretory cargo molecules from the ER to ERGIC.** A: Trafficking of cargo from ER to ERGIC in
666 a COPII-dependent manner. Small cargo proteins can utilise standard sized ~80 nm COPII vesicles
667 (i). Large procollagens are exported from the ER with the help of their chaperone Hsp47 acting as
668 an adaptor for TANGO1/cTAGE5. TANGO1/cTAGE5 act to enhance growth of COPII buds. Slightly
669 more flexible procollagen can be accommodated in slightly larger (~150 nm) COPII carriers (ii).

670 Large less flexible procollagen is transported via large (>350 nm) COPII carriers (iii). Enlarged
671 transport intermediates result in direct connections between ER and ERGIC to facilitate
672 procollagen trafficking (iv). Due to the spatial restriction between ER and ERGIC larger transport
673 intermediates may still remain bound to the ER or already started fusing with the ERGIC, resulting
674 in transient states of transport. B: Transport mechanisms of other large secretory cargo proteins
675 from the ER to ERGIC/Golgi remain to be identified.

676

677

678

679

680

681

682

683

684

685

686 *The tables that follow are pasted in line to ensure correct numbering of references. Do not use*
687 *these but instead the individual files.*

688 Table 1: Overview of large procollagens and other unconventional secretory cargo proteins.

Procollagen type	Length
Procollagen type I, fibrillar collagen	~330nm [9]

Procollagen type II, fibrillar collagen	~300nm [18]

Procollagen type III, fibrillar collagen	~300nm [114]

Procollagen type IV, networking collagen	~430nm [9]

Procollagen type V, fibrillar collagen	~330nm [9]

Procollagen type VII, networking collagen	~425nm [115]

Procollagen type XI, fibrillar collagen	>300 nm?

Aggrecan, proteoglycan	330 nm [116]

Fibrillin, glycoprotein	150 nm [8]

Firbonectin, glycoprotein	160nm [117]

Laminins, glycoproteins	110 - 165nm [118]

Table 2: Evidence for COPII-dependency and other factors of the early secretory pathway influencing unconventional cargo transport from the ER to the Golgi (expanded from [3]).

Compromised Component ¹	Affected cargo ²	Resulting effect on cargo transport	Organism	Tissue/cell type	Reference
Sec23A	COL1	Accumulation in ER	Human	Fibroblasts	[96]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[119-121]
Sec23B	COL1	Accumulation in ER	Human	Kidney cells	[121]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[120]
Sec24D	COL1	Accumulation in ER	Human	Fibroblasts	[122]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[123]
	COL2	Intracellular accumulation	Medaka	Chondrocytes	[124]
			Medaka	Notochord	
			Medaka	Myoseptum	
Sec13	COL1	Defective secretion and deposition	Human	Fibroblasts	[125]
			Human	Fibroblasts	
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[126]
Sar1B	Chylomicron	Retention in membrane-bound compartments	Human	Enterocytes	[127]
TFG	COL1	Decreased extracellular protein levels	Human	Fibroblasts	[81]
	<i>ER structural defects</i>	Hereditary spastic paraplegia	Human	Patients	[128]
TANGO1 /MIA3	COL7	Accumulation in ER	Human	Fibroblasts	[16]
	COL1		Mouse	Chondrocytes	[54]
	COL2	Intracellular accumulation	Mouse	Chondrocytes	
	COL3	Intracellular accumulation	Mouse	Endothelial cells	
	COL4	Intracellular accumulation	Mouse	Embryonic Fibroblasts	
	COL7	Intracellular accumulation	Mouse	Epithelial cells	
	COL9		Mouse	Epithelial cells	
	COL12		Human		
	(+TALI/MIA2) Chylomicron		Human		[59]
cTAGE5 /MIA2	COL7	Accumulation in ER	Human	A431 cells	[53]
Sedlin	COL1	Accumulation in ER	Human	Fibroblasts	[15]
	COL2	Accumulation in ER	Human	Chondrocytes	
CREB3L1/OASIS	COL1	Defective secretion and deposition	Human, mouse	Osteoblasts	[103-105]
CREB3L2/BBF2H7	COL2	Defective secretion and deposition	Zebrafish, medaka	Chondrocytes	[106, 107]

¹Compromised components are either depleted, mutated or knocked out.

693 ²Collagens are abbreviated as COL followed by the type number.

694

Highlights

- Export of cargo from the ER requires an adaptable system to accommodate cargo of diverse size and shape.
- In metazoans, the secretion of collagens has been widely studied due to their large size
- TANGO1 has emerged as a key player in assembling COPII-dependent machinery to drive export from the ER.
- The interplay between TANGO1 and the chaperone Hsp47 is central to understanding this process.
- New data and resulting models have challenged the “large carrier” model and have led to questions of whether complete coating of carriers with COPII is needed.

Outstanding Questions

While the in vitro characterisation of COPII function is comprehensive, key issues remain concerning its function in cells and tissues. Are emerging carriers ever fully coated? Does COPII have a more restricted role in cargo selection at the base of these larger structures? Does this occur preferentially at sites in close proximity to the Golgi?

Does procollagen selectively transfer from ERES to the Golgi via an ERGIC intermediate?

Do direct connections between the ER and ERGIC (or even the Golgi) exist?

Do these more direct pathways also facilitate bulk flow of small diffusible cargo?

Is the exquisite sensitivity of procollagen trafficking to defects in secretory membrane traffic more a function of secretory challenge than a specific molecular requirement for a specialized pathway?

What is the role of TANGO1 and other proteins from the MIA gene family in secretion more widely?

How does transcriptional, translational and post-translational control of ER-to-Golgi trafficking regulate cargo trafficking, notably in the context of procollagen and ECM?

Figure 1

A

B

McCaughey and Stephens Graphical Abstract

Table 1: Overview of large procollagens and other unconventional secretory cargo proteins.

Procollagen type	Length
Procollagen type I, fibrillar collagen	~330nm [4]

Procollagen type II, fibrillar collagen	~300nm [14]

Procollagen type III, fibrillar collagen	~300nm [112]

Procollagen type IV, networking collagen	~430nm [4]

Procollagen type V, fibrillar collagen	~330nm [4]

Procollagen type VII, networking collagen	~425nm [113]

Procollagen type XI, fibrillar collagen	>300 nm?

Aggrecan, proteoglycan	330 nm [114]

Fibrillin, glycoprotein	150 nm [3]

Firbonectin, glycoprotein	160nm [115]

Laminins, glycoproteins	110 - 165nm [116]

Table 2: Evidence for COPII-dependency and other factors of the early secretory pathway influencing unconventional cargo transport from the ER to the Golgi (expanded from [2]).

Compromised Component ¹	Affected cargo ²	Resulting effect on cargo transport	Organism	Tissue/cell type	Reference
Sec23A	COL1	Accumulation in ER	Human	Fibroblasts	[95]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[117-119]
Sec23B	COL1	Accumulation in ER	Human	Kidney cells	[119]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[118]
Sec24D	COL1	Accumulation in ER	Human	Fibroblasts	[120]
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[121]
	COL2	Intracellular accumulation	Medaka	Chondrocytes	[122]
			Medaka	Notochord	
			Medaka	Myoseptum	
Sec13	COL1	Defective secretion and deposition	Human	Fibroblasts	[123]
			Human	Fibroblasts	
	COL2	Accumulation in ER	Zebrafish	Chondrocytes	[124]
Sar1B	Chylomicron	Retention in membrane-bound compartments	Human	Enterocytes	[125]
TFG	COL1	Decreased extracellular protein levels	Human	Fibroblasts	[80]
	<i>ER structural defects</i>	Hereditary spastic paraplegia	Human	Patients	[126]
TANGO1 /Mia3	COL7	Accumulation in ER	Human	Fibroblasts	[12]
	COL1		Mouse	Chondrocytes	[54]
	COL2	Intracellular accumulation	Mouse	Chondrocytes	
	COL3	Intracellular accumulation	Mouse	Endothelial cells	
	COL4	Intracellular accumulation	Mouse	Embryonic Fibroblasts	
	COL7	Intracellular accumulation	Mouse	Epithelial cells	
	COL9		Mouse	Epithelial cells	
	COL12		Human		
(+TALI/Mia2)	Chylomicron		Human		[59]
cTAGE5 /Mia2	COL7	Accumulation in ER	Human	A431 cells	[53]
Sedlin	COL1	Accumulation in ER	Human	Fibroblasts	[11]
	COL2	Accumulation in ER	Human	Chondrocytes	
CREB3L1/OASIS	COL1	Defective secretion and deposition	Human, mouse	Osteoblasts	[101-103]
CREB3L2/BBF2H7	COL2	Defective secretion and deposition	Zebrafish, medaka	Chondrocytes	[104, 105]

¹Compromised components are either depleted, mutated or knocked out.

²Collagens are abbreviated as COL followed by the type number.