ER-to-Golgi transport: a sizeable problem

Janine McCaughey and David J. Stephens*

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

* Corresponding author:

Email: david.stephens@bristol.ac.uk

Tel: 00 44 117 331 2173

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

Number of words: 3360

Keywords

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

Abstract

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

Glossary

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

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

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

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

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

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

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

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

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

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

Oversized cargo: too big for COPII?

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

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

The case for large carriers

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

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

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

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

TANGO1 in ER-to-Golgi transport of collagens

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

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

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

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

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

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

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

Alternative modes of ER-Golgi trafficking of large cargoes

Large COPII carriers have yet to be identified in cells expressing endogenous protein levels. Primary fibroblasts do not contain any evident large COPII structures [11, 71]. Most of the experiments resulting in large carriers were performed in cells overexpressing KLHL12 and or procollagens in transformed cell lines [11, 12]. Furthermore, the micron-sized punctate structures positive for Sec23 and procollagen in mouse osteoblasts are also positive for markers of autophagy and ubiquitin [72]. These structures appear to be relatively immobile and do not traffic to the next compartment, but instead are incorporated into lysosomes for degradation indicating a form of noncanonical autophagy taking place at ERES to remove excess procollagen [72]. Indeed, collagen can be secreted independent of the CUL3-KLHL12 complex and CUL3 depletion in human skin fibroblasts blocks collagen synthesis at the level of translation, and upregulates sensors of the unfolded protein response (UPR), PERK and IRE1α [73]. The IRE1α-XBP1s axis couples the availability of nutrients, including lipids, to the COPII vesicle budding cycle via transcriptional control of COPII subunit synthesis [74]. This could have implications for the formation of lipoprotein-containing vesicles whose size on exit from the ER is less certain but is linked to COPII directly because of disease-causing mutations in Sar1B ([75] and Table 2). These links between metabolism and COPII-dependent trafficking likely impact on the role of autophagy in proteostasis

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

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

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

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

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

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

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

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

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

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.

Acknowledgements

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

Conflict statement

The authors declare no conflicts of interest.

References

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

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

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

70. Maeda, M. et al. (2017) TANGO1 recruits Sec16 to coordinately organize ER exit sites for efficient secretion. *J Cell Biol* 216 (6), 1731-1743.
71. McCaughey, J. et al. (2019) ER-to-Golgi trafficking of procollagen in the absence of large carriers. *J Cell Biol* 218 (3), 929-948.
72. Omari, S. et al. (2018) Noncanonical autophagy at ER exit sites regulates procollagen turnover. *Proc Natl Acad Sci U S A* 115 (43), E10099-E10108.
73. Kim, K. et al. (2018) Cullin3-RING ubiquitin ligases are intimately linked to the unfolded protein response of the endoplasmic reticulum.
74. Liu, L. et al. (2019) Coupling of COPII vesicle trafficking to nutrient availability by the IRE1alpha-XBP1s axis. *Proc Natl Acad Sci U S A* 116 (24), 11776-11785.
75. Jones, B. et al. (2003) Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nat Genet* 34 (1), 29-31.
76. Wang, J. et al. (2017) Autophagosome formation: Where the secretory and autophagy pathways meet. *Autophagy* 13 (5), 973-974.
77. Bonfanti, L. et al. (1998) Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell* 95 (7), 993-1003.
78. Raote, I. and Malhotra, V. (2019) Protein transport by vesicles and tunnels. *J Cell Biol* 218 (3), 737-739.
79. Hanna, M.G. et al. (2018) Membrane Transport at an Organelle Interface in the Early Secretory Pathway: Take Your Coat Off and Stay a While: Evolution of the metazoan early secretory pathway. *Bioessays*, e1800004.
80. Kurokawa, K. et al. (2014) Contact of cis-Golgi with ER exit sites executes cargo capture and delivery from the ER. *Nat Commun* 5, 3653.
81. McCaughey, J. et al. (2016) TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion. *Cell Rep* 15 (8), 1648-59.
82. Johnson, A. et al. (2015) TFG clusters COPII-coated transport carriers and promotes early secretory pathway organization. *EMBO J* 34 (6), 811-27.
83. Guo, Y. et al. (2018) Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution on Millisecond Timescales. *Cell* 175 (5), 1430-1442 e17.
84. Schermelleh, L. et al. (2019) Super-resolution microscopy demystified. *Nat Cell Biol* 21 (1), 72-84.
85. Russell, M.R. et al. (2017) 3D correlative light and electron microscopy of cultured cells using serial blockface scanning electron microscopy. *J Cell Sci* 130 (1), 278-291.
86. Xu, C.S. et al. (2017) Enhanced FIB-SEM systems for large-volume 3D imaging. *Elife* 6.
87. Schlacht, A. and Dacks, J.B. (2015) Unexpected ancient paralogs and an evolutionary model for the COPII coat complex. *Genome Biol Evol* 7 (4), 1098-109.
88. Bannykh, S.I. and Balch, W.E. (1997) Membrane dynamics at the endoplasmic reticulum-Golgi interface. *Journal of Cell Biology* 138 (1), 1-4.
89. Hughes, H. et al. (2009) Organisation of human ER-exit sites: requirements for the localisation of Sec16 to transitional ER. *Journal of Cell Science* 122 (Pt 16), 2924-34.
90. Watson, P. et al. (2006) Sec16 defines endoplasmic reticulum exit sites and is required for secretory cargo export in mammalian cells. *Traffic* 7 (12), 1678-87.
91. Connerly, P.L. et al. (2005) Sec16 is a determinant of transitional ER organization. *Curr Biol* 15 (16), 1439-47.
92. Montegna, E.A. et al. (2012) Sec12 binds to Sec16 at transitional ER sites. *PLoS One* 7 (2), e31156.
93. Barlowe, C. and Schekman, R. (1993) SEC12 encodes a guanine-nucleotide-exchange factor 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.*

Figure legends

Graphical abstract caption.

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

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

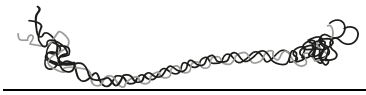
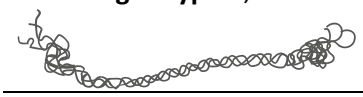
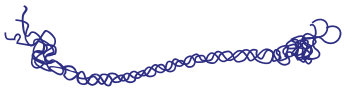
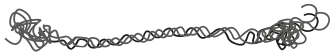
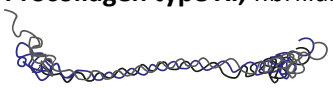
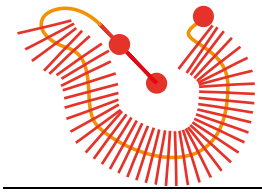
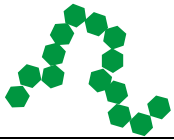
Figure 1: Motifs in proteins of the COPII machinery and TANGO1-related isoforms. Motifs shown are those predicted by smart.embl.de and or denoted on UniProt.org. The COPII machinery is shown on the left and proteins of the MIA gene family on the right. Motifs shown include transmembrane domains (TMD),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

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

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

The tables that follow are pasted in line to ensure correct numbering of references. Do not use 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]

689

690 Table 2: Evidence for COPII-dependency and other factors of the early secretory pathway influencing
691 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	
			Medaka	Notochord	[124]
			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]
	ER structural defects	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]

692 [1]Compromised components are either depleted, mutated or knocked out.

693 [2]Collagens 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 q

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?

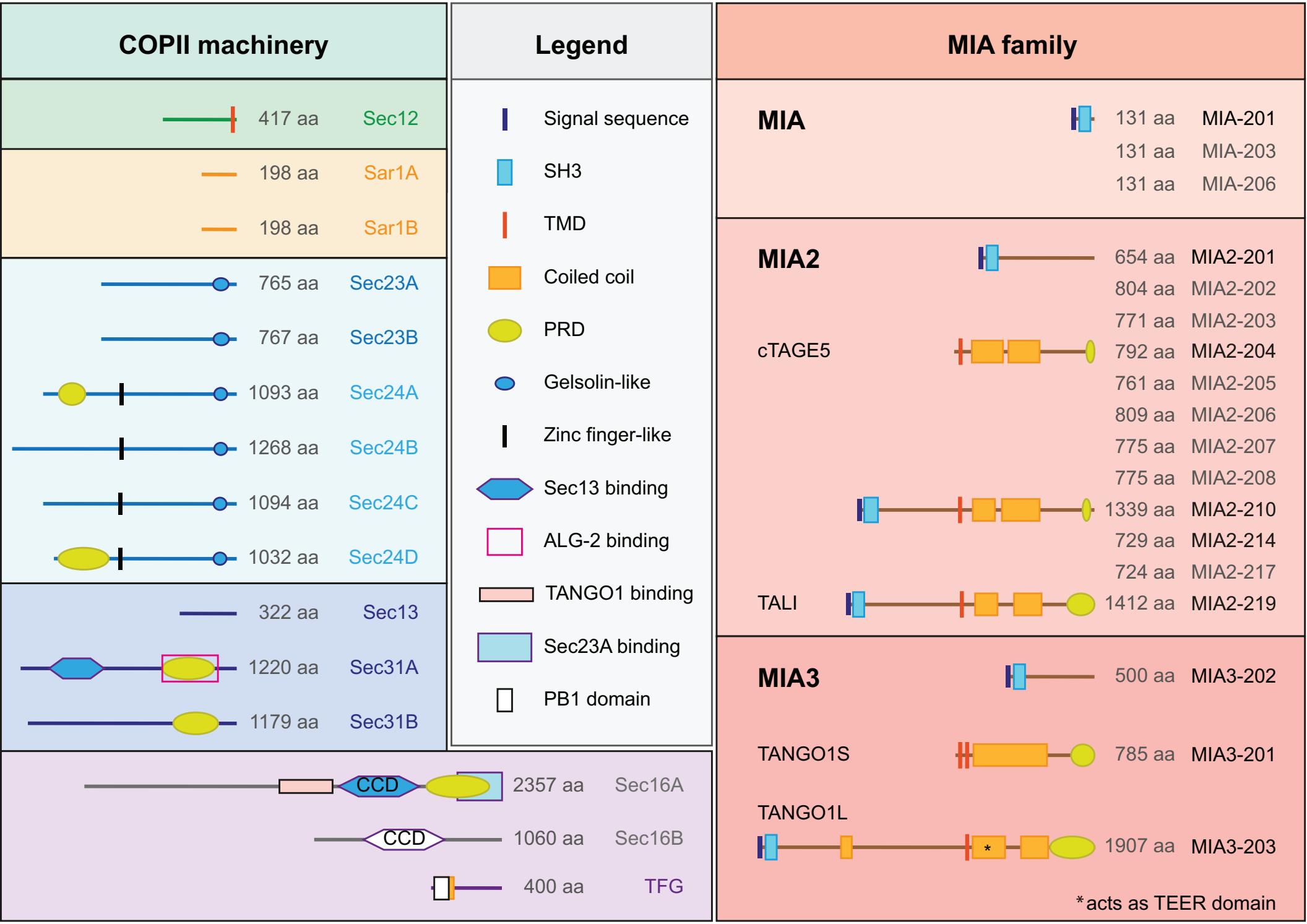
Do direct connections between the ER and ERGIC (or even the Golgi) exist?

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

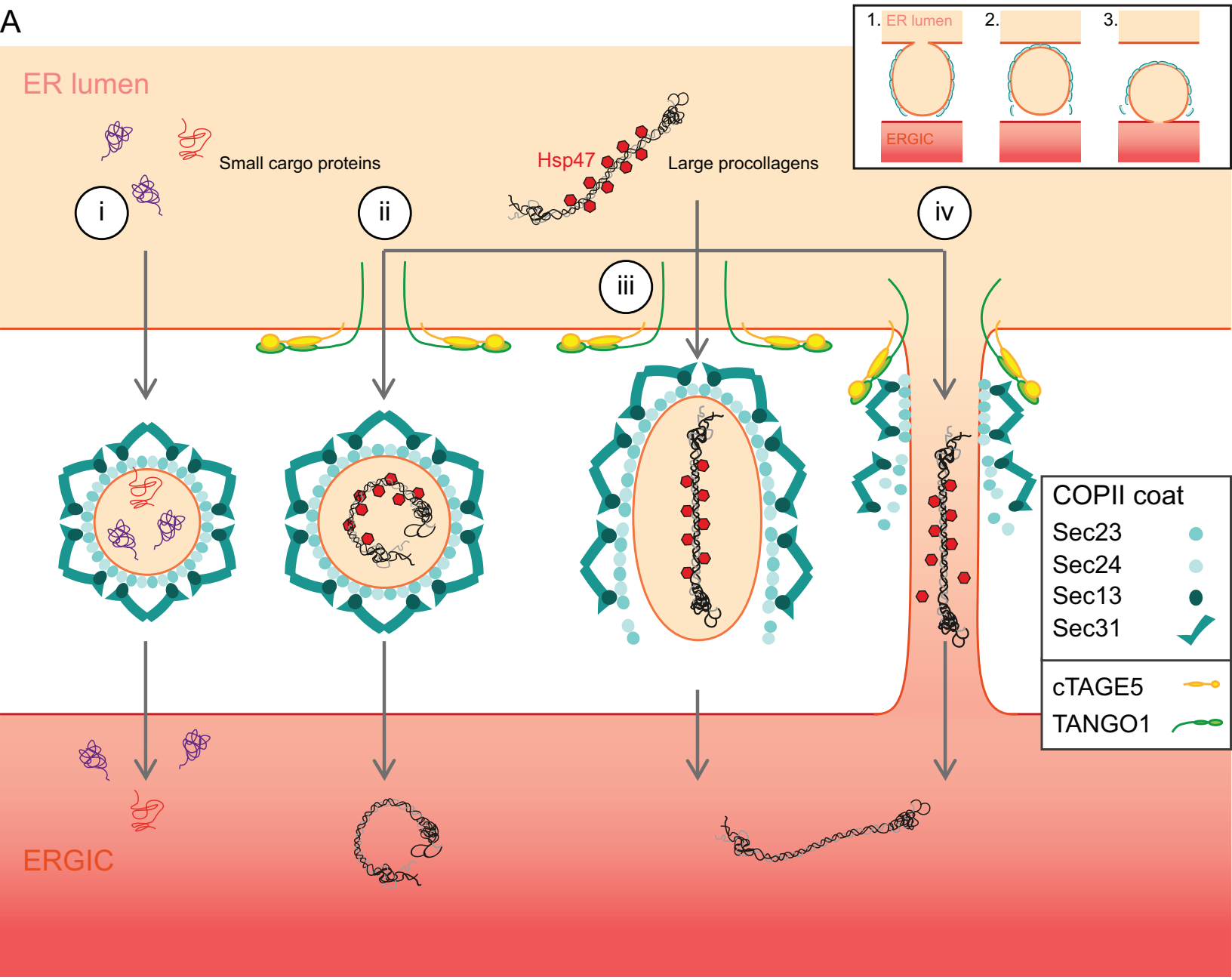
Is the exquisite sensitivity of procollagen trafficking to defects in secretory membrane traffic more a function of secretory challenge than a specific molecular requirement for a specialized pathway?

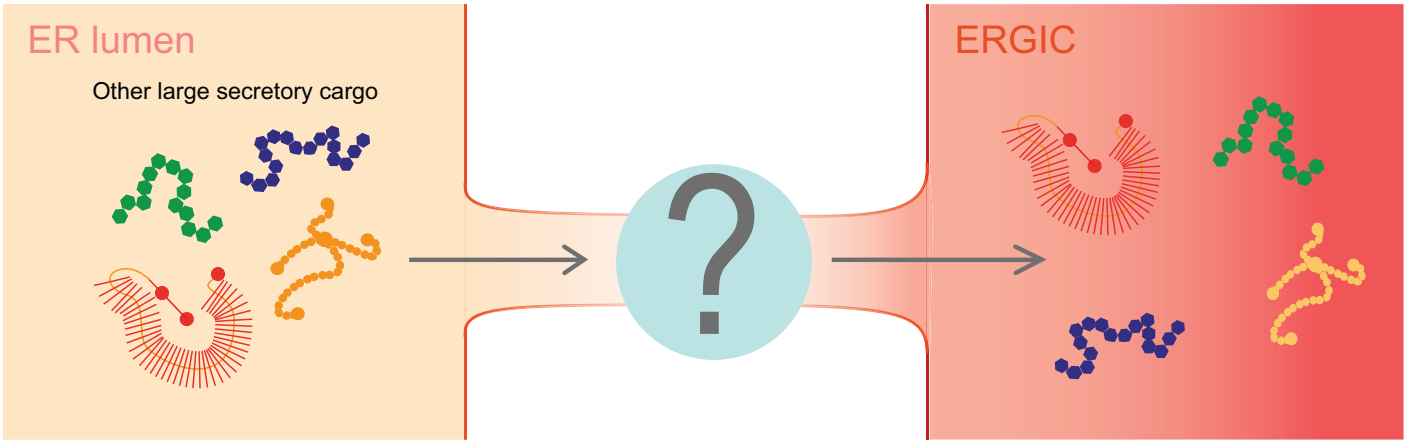
What is the role of TANGO1 and other proteins from the MIA gene family in secretion more widely?

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

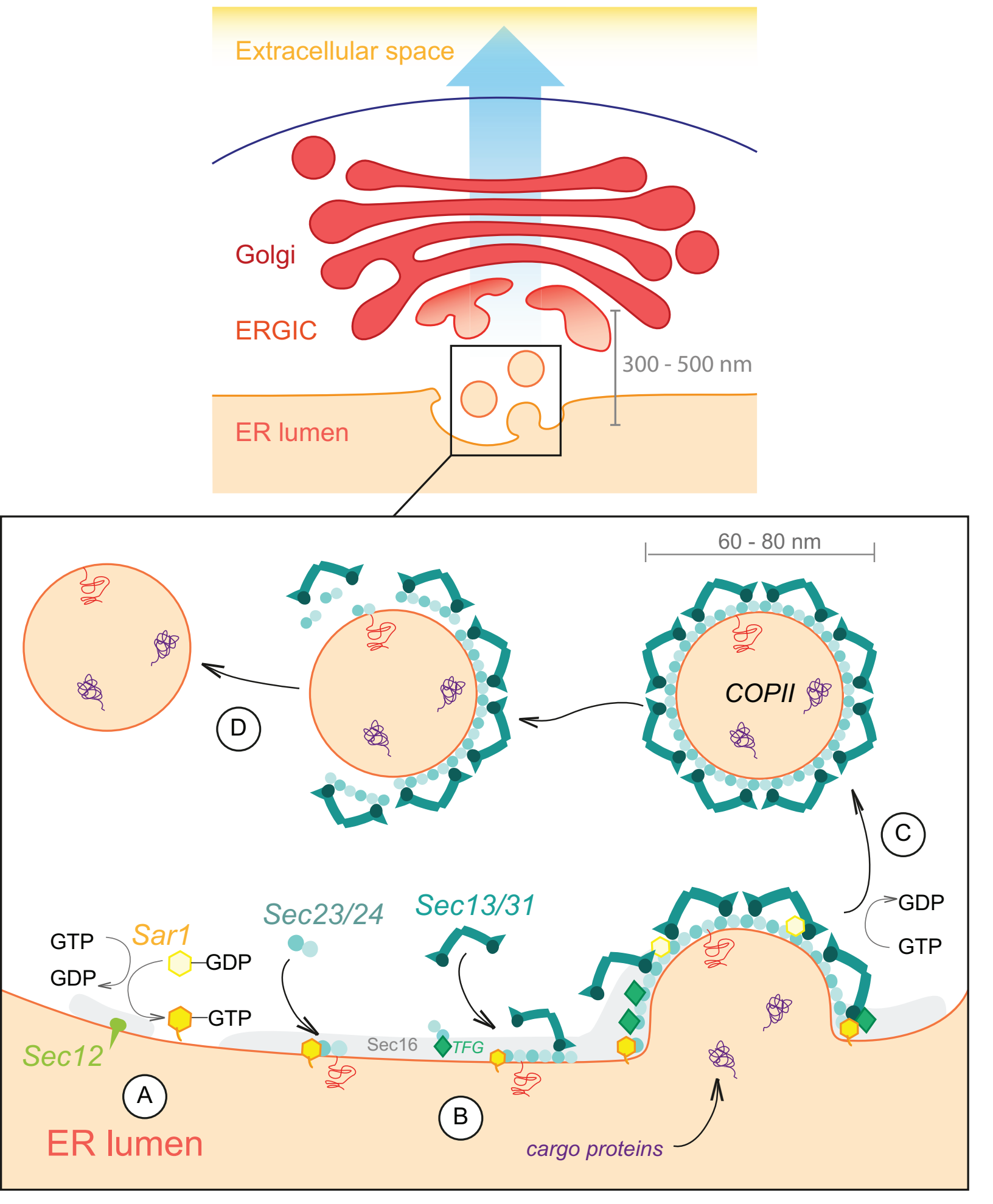

COPII machinery
417 aa Sec12
198 aa Sar1A
198 aa Sar1B
765 aa Sec23A
767 aa Sec23B
1093 aa Sec24A
1268 aa Sec24B
1094 aa Sec24C
1032 aa Sec24D
322 aa Sec13
1220 aa Sec31A
1179 aa Sec31B
CCD
2357 aa Sec16A
CCD
1060 aa Sec16B
400 aa TFG
Legend
Signal sequence
SH3
TMD
Coiled coil
PRD
Gelsolin-like
Zinc finger-like
Sec13 binding
ALG-2 binding
TANGO1 binding
Sec23A binding
PB1 domain
MIA family
MIA
131 aa MIA-201
131 aa MIA-203
131 aa MIA-206
MIA2
654 aa MIA2-201
804 aa MIA2-202
771 aa MIA2-203
cTAGE5
792 aa MIA2-204
761 aa MIA2-205
809 aa MIA2-206
775 aa MIA2-207
775 aa MIA2-208
1339 aa MIA2-210
729 aa MIA2-214
724 aa MIA2-217
TALI
1412 aa MIA2-219
MIA3
500 aa MIA3-202
TANGO1S
785 aa MIA3-201
TANGO1L
*
1907 aa MIA3-203
*acts as TEER domain

McCaughey and Stephens Figure 2 KEY FIGURE





McCaughey and Stephens Figure I



Graphical abstract

McCaughey and Stephens Graphical Abstract

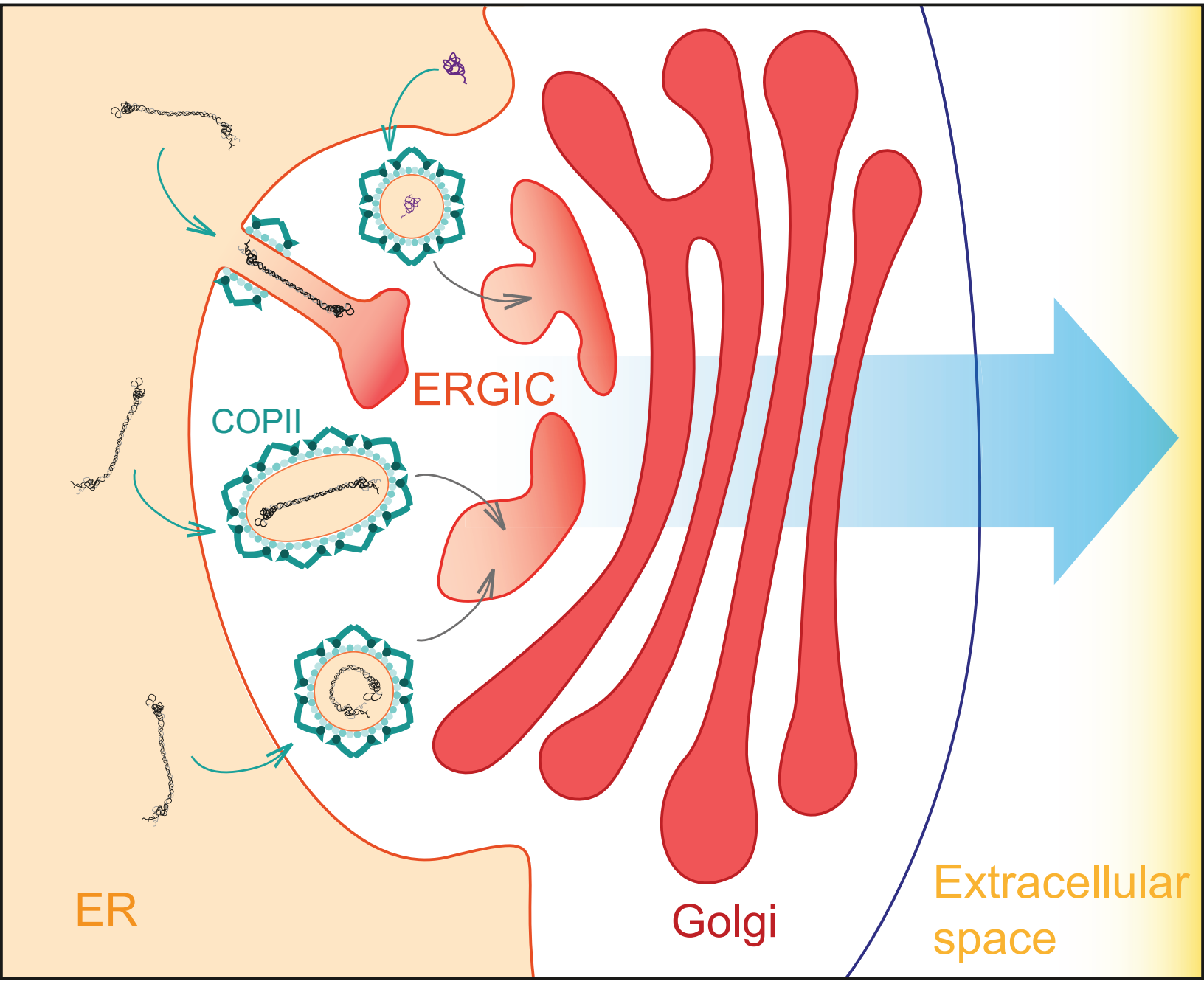


Table 1: Overview of large procollagens and other unconventional secretory cargo proteins.

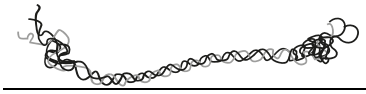
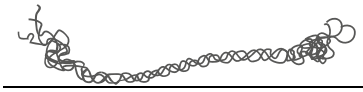
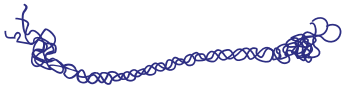
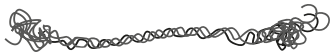
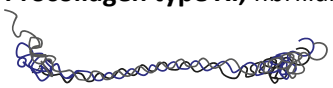
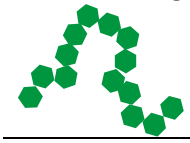
Procollagen type	Length
Procollagen type I, fibrillar collagen 	~330nm [4]
Procollagen type II, fibrillar collagen 	~300nm [14]
Procollagen type III, fibrillar collagen 	~300nm [112]
Procollagen type IV, networking collagen 	~430nm [4]
Procollagen type V, fibrillar collagen 	~330nm [4]
Procollagen type 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]
	ER structural defects	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]

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

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