MIA3/TANGO1 enables efficient secretion by constraining COPII vesicle budding

Authors: Janine McCaughey[1†], Judith M. Mantell[2], Chris R. Neal[2], Kate Heesom[3], and David J. Stephens[1*]

Affiliations

[1]Cell Biology Laboratories, School of Biochemistry, Faculty of Life Sciences, University Walk, University of Bristol, Bristol, BS8 1TD, UK.

[2]Wolfson Bioimaging Facility, Faculty of Life Sciences, University Walk, University of Bristol, Bristol, BS8 1TD, UK.

[3]Proteomics Facility, Faculty of Life Sciences, University Walk, University of Bristol, Bristol, BS8 1TD, UK.

*Correspondence to: david.stephens@bristol.ac.uk.

†Current address: Utrecht University, Kruytgebouw, Padualaan 8, 3584 CH Utrecht, The Netherlands

ORCID:
JM: 0000-0001-7709-2597
JMM: 0000-0002-9061-8933
CRN: 0000-0001-6604-280X
KH: 0000-0002-5418-5392
DJS: 0000-0001-5297-3240

Abstract

Complex machinery is required to drive secretory cargo export from the endoplasmic reticulum. In vertebrates, this includes transport and Golgi organization protein 1 (TANGO1), encoded by the Mia3 gene. Here, using genome engineering of human cells light microscopy, secretion assays, and proteomics, we show loss of Mia3/TANGO1 results in formation of numerous vesicles and a loss of early secretory pathway integrity. This restricts secretion not only of large proteins like procollagens but of all types of secretory cargo. Our data shows that Mia3/TANGO1 constrains the propensity of COPII to form vesicles promoting instead the formation of the ER-Golgi intermediate compartment. Thus, Mia3/TANGO1 facilities the secretion of complex and high volume cargoes from vertebrate cells.

Introduction

The first membrane trafficking step for secretion is driven by assembly of the COPII coat complex onto the endoplasmic reticulum (ER) membrane. In yeast and many other eukaryotes, this results in COPII vesicles that bud from the ER membrane (Bednarek et al., 1995). This process can be reconstituted *in vitro* using synthetic liposomes and a minimal COPII machinery of a small GTP binding protein Sar1p that, in its GTP-bound form, recruits an inner coat of Sec23p-Sec24p, and subsequently an outer coat of Sec13p-Sec31p (Matsuoka et al., 1998). Together these proteins are sufficient to generate 60-80 nm vesicles in an energy-dependent manner. Several other proteins support the COPII system including the guanine nucleotide exchange factor Sec12p that activates Sar1p, and Sec16p which potentiates vesicle formation (Espenshade et al., 1995). In metazoans, COPII proteins including Sec16 assemble at relatively stable sites on the ER membrane called transitional ER (Orci et al., 1994) from which COPII vesicles bud. In the most commonly accepted models, these then coalesce to form an ER-Golgi Intermediate Compartment (ERGIC, (Schweizer et al., 1988)). Collectively these structures form ER exit sites (ERES, (Hughes et al., 2009)). These sites are the location for cargo selection and bud formation. Despite the prevalence of this vesicular model in the literature, few studies have identified any significant number of 60-80nm secretory vesicles in vertebrate cells or tissues. While they have been detected (Martinez-Menarguez et al., 1999) they are not abundant.

In metazoans, many proteins have been identified which help orchestrate and regulate ERES membrane dynamics in diverse ways. TANGO1, encoded by Mia3, was identified in a genetic screen in *Drosophila* S2 cells, as a factor required for the secretion of a horseradish peroxidase reporter (Bard et al., 2006). Knockout of TANGO1 in *Drosophila* leads to defects in ER morphology, induction of ER stress, and defects in cargo secretion (Rios-Barrera et al., 2017). Analysis of TANGO1 function is complicated by the presence of multiple isoforms encoded by the Mia3 gene, the two principal ones being TANGO1S and TANGO1L indicating the short (785 amino acid) and long (1907 amino acid) forms (McCaughey and Stephens, 2019). The longer form contains an ER luminal SH3 domain that is reported to engage cargo, including procollagen by binding to the chaperone Hsp47 (Ishikawa et al., 2016). Notably, both evidence from RNAi experiments suggest that TANGO1L and TANGO1S function interchangeably in terms of procollagen secretion (Maeda et al., 2016). Thus, the roles and relevance of the two forms remains unclear.

TANGO1 recruits Sec16 (Maeda et al., 2017) and proteins encoded by Mia2 gene including cTAGE5 (cutaneous T cell lymphoma-associated antigen 5, (Ma and Goldberg, 2016; Saito et al., 2011)). Together cTAGE5 and TANGO1 form larger complexes with Sec12 (Maeda et al., 2016), the guanine nucleotide exchange factor that promotes assembly of COPII via GTP loading of Sar1. TANGO1 also integrates with SNARE proteins to recruit membranes from the ER-Golgi intermediate compartment (ERGIC) to ERES (Nogueira et al., 2014). This has been considered to provide additional membrane to promote bud expansion to facilitate encapsulation of large cargo such as fibrillar procollagens (Ma and Goldberg, 2016; Nogueira et al., 2014; Raote et al., 2018; Saito et al., 2009) and pre-chylomicrons (Santos et al., 2016) in 'mega-carriers'. This model has been extended following the identification of ring-like structures of TANGO1 that could be consistent with the "neck" of an emerging bud (Liu et al., 2017; Raote et al., 2017). TANGO1 is clearly a key component of the ER export machinery (Raote and Malhotra, 2021).

A Mia3 knockout mouse (Wilson et al., 2011) has been described that has substantial defects in bone formation leading to embryonic lethality. Biallelic mutations in TANGO1 have been described in humans that result in skipping of exon 8 and multiple defects including skeletal abnormalities, diabetes, hearing loss and mental retardation (Lekszas et al., 2020). This is reflected in TANGO1 mutant zebrafish models where multiple organs were found to be affected following loss of TANGO1 or its closely related orthologue, cTAGE5 (encoded by Mia2) (Clark and Link, 2021). Total loss of TANGO1 expression in humans is embryonically lethal showing an absence of bone mineralization (Guillemyn et al., 2021), reflecting the phenotype seen in Mia3$^{-/-}$ mice (Wilson et al., 2011).

Results

To investigate the relative contribution of TANGO1S and TANGO1L at the ER-Golgi interface we generated knockout human cell lines using CRISPR-Cas9. We designed guide RNAs against exon 2 to knockout TANGO1L and against exon 7 to knockout both TANGO1S and TANGO1L (Fig. 1A). The outcome from genome sequencing of these clones is shown in Fig. 1B. We were not able to design gRNAs to selectively target TANGO1S owing to the shared sequence with TANGO1L. Clonal mutant cell lines were validated using antibodies selective for TANGO1L or both TANGO1S and L by immunoblotting (Fig. 1C) and immunofluorescence (Fig. 1D). We obtained clones that lacked expression of TANGO1L only (denoted TANGO1 L-/S+), clones in which the transmembrane domain encoded by exon 7 as well as TANGO1S were absent but the remaining truncated TANGO1L protein was expressed (TANGO1 LΔ/S-), and clones in which both TANGO1L and TANGO1S were absent or the expression of the truncated TANGO1L drastically reduced (TANGO1 L-/S-). We present data from two clones of each. The trace amounts of protein persisting in the clones could be due to lines not being completely clonal or some residual expression within some cells (Fig. 1C). TANGO1 L-/S- clone 1 does not express detectable TANGO1S and only trace levels of TANGO1L and therefore is the closest to a complete TANGO1 knockout cell line. TANGO1 LΔ/S- clone 2 and TANGO1 L-/S- clone 1 grew very slowly (approximately 5-fold slower than wild-type (WT)).

We analyzed expression levels of COPII proteins and found that while Sec24A, Sec24C, Sec24D, Sec31A, and Sec12 were unaffected, TFG was upregulated in those knockout-cell lines where exon 7 of Mia3 was targeted resulting in a disrupted cytosolic domain. In all KO cell lines, expression of cTAGE5 (encoded by Mia2) was dramatically reduced and conversely. In most cell lines, immunofluorescence (Fig. 1D) confirmed a loss of TANGO1 expression, however the TANGO1 LΔ/S- lines had notable abnormal, large, TANGO1-positive structures. Labelling for GM130 in all KO cells showed obvious disruption of the Golgi consistent with fragmentation while retaining a broadly juxtanuclear location.

We used transmission electron microscopy to define the ER-Golgi interface. WT cells contained a typical connected juxtanuclear Golgi ribbon (Fig. 2A). TANGO1 L-/S+ cells show disruption to Golgi structure with more separated mini-Golgi stacks evident. TANGO1 LΔ/S- cells show a more severe version of this phenotype along with the presence of many small vesicular structures. TANGO1 L-/S- cells are packed with these small round vesicular structures along with other electron-lucent membranous structures that resemble degradative compartments. Scattered mini-Golgi elements are visible as well as occasional large electron dense structures consistent with enlarged ER with retained secretory cargo (Fig. 1 and (Fromme et al., 2007)). The ER-Golgi interface appears very different in these cells, with fewer pleiomorphic membranes between ER and Golgi structures.

We sought to confirm that this disruption was due to the loss of TANGO1 by reintroducing recombinant tagged proteins, TANGO1S-mScarlet-i (mSc, (Bindels et al., 2017)) or TANGO1L-HA (Raote et al., 2017)). Fig. 3 shows that TANGO1S-mSc expression reverses Golgi disruption. Furthermore, TANGO1L-HA also restores a compact juxtanuclear localization of the Golgi. Co-labeling these cells also showed that the pattern of TANGO1 labelling mirrored that of Sec16A consistent with a bona fide localization of the recombinant form.

We further analyzed the organization of the ER-Golgi interface using light microscopy (Fig. 4A and 4B, quantified in 4C-G). Fig. 4A shows localization of COPII proteins, Sec24C and Sec31A, in a characteristic punctate pattern in WT cells that is disrupted in TANGO1 knockout cells. In the most severe cases (TANGO1 L-/S-) the pattern of localization is diffuse with many more puncta detected. These changes in COPII protein distribution are also seen with Sec16A labelling (Fig. 4B). Notably COPII labelling remains clustered in a juxtanuclear position. A dramatic change in localization of ERGIC53 (Fig. 4B) is also seen with this classical marker of the ER-Golgi intermediate compartment becoming completely localized to the ER in TANGO1 LΔ/S- and L-/S- cells. GRASP65 and β-COP remain associated with the Golgi in all cells examined suggesting an impaired but still functional secretory pathway. Automated quantification of immunofluorescence data (Fig. 4C-G) showed an increase in the number of TFG, Sec16A, Sec24C, and Sec31A positive structures in most Mia3 knockout cell lines. This increase is consistent with the numerous vesicular structures, notably in TANGO1 L-/S- clone 1, and correlates closely with the impact on ERGIC53 distribution. Consistent with this, loss of peripheral β-COP labelling is consistent with loss of a functional ERGIC (Scales et al., 1997).

Since TANGO1 knockout cells are viable and adherent we explored the efficiency of membrane traffic in both targeted and unbiased assays. Using a biotin-controlled secretion system (Boncompain et al., 2012b) we monitored transfer of cargo (mannosidase II tagged with mCherry) from the ER (labelled with protein disulfide isomerase (PDI))

to the Golgi (labelled with GRASP65). At 30 minutes after addition of biotin to release mannII-mCh from the ER, around half of TANGO1 knockout cells retained mannII-mCh in the ER (Fig. 5A, B, quantified for all cell lines in C). We further explored this using E-cadherin-mCh. Here we saw a more dramatic phenotype where almost all TANGO1 LΔ/S- and L-/S- cells retained this reporter in the ER (Fig. 5D, E, quantified as green bars in Fig. 5F).

We analyzed the secreted proteome of WT and TANGO1 knockout cells. All TANGO1 knockout cells showed defects in secretion (in all 3 biological repeats of the experiment); TANGO1 L-/S- cells showed a >5-fold reduction in secretion of amyloid precursor protein, amyloid precursor-like protein 2, NPC2, clusterin, fibrillin-1, fibrillin-2, follistatin related protein 1, IGF binding protein 7, secreted frizzled protein 7, semaphorin 7A, SERPINE, thrombospondin-1, and TGF-beta. Of note this list includes multiple components of the extracellular matrix but of varying size including many small soluble proteins. Furthermore, we saw a significant increase in secretion of fibronectin. These data are not consistent with a selective defect in the secretion of large cargo.

We also analyzed the cell-derived matrix after removal of the cell monolayer by mass spectrometry (Supplementary Data 6). Proteins decreased at least 5-fold in abundance (and filtered for detection of >10 peptides) in TANGO1 L-/S- compared to WT cells in three independent repeats included collagen IV, protein-glutamine gamma-glutamyltransferase-2 (TGM2), fibronectin, and lysyl oxidase 1. Fibrillin-2, TGFβ, latent TGFβ binding protein 3, and perlecan were all reduced >5-fold in two of the three biological replicates. Interestingly we also see an increase in several proteins in all three repeats including the secreted signaling proteins Wnt5b (average 7.7-fold increase), semaphorin-3C (7.2-fold) and -3D (4.0-fold), and midkine (3.8-fold). Extracellular matrix (ECM) remodeling proteins ADAMTS1 (6.9-fold), and cell migration-inducing and hyaluronan-binding protein (3.0-fold) are also decreased in all three experiments. We also see increases in intracellular proteins including vimentin (5.2-fold) and interestingly, the plasma membrane integrin, ITGAV (2.6-fold). These are presumably retained on extraction of the cell layer but could importantly reflect key changes in cell architecture and adhesion. ITGAV is notable as a receptor for non-collagenous matrix ligands including fibronectin and vitronectin, notably increased in secretion in the soluble fraction. It is important to note that monitoring cell-derived matrix measures those proteins in the assembled matrix and does not necessarily reflect changes in secretion per se. Notably in both soluble proteomes and cell-derived matrix we see reductions in small soluble proteins including TGFβ along with many large glycoproteins of 350-500 kDa which likely require modified transport mechanisms compared to 60-80 nm vesicles. However, these are likely to be flexible polymers (Rezaei et al., 2018) that could be accommodated in a diversity of membrane-bound carriers, without requiring formation of large >500 nm megacarriers. We also see increases in the presence of several key matrix proteins in cell-derived matrix that suggest cellular adaptation to defective secretory pathway function.

TANGO1 has been defined previously as a factor that selectively controls secretion of procollagens through interaction with Hsp47 (Ishikawa et al., 2016). Indeed, the reported size of procollagen VII, the first defined cargo of TANGO1, underpins the model of a role for TANGO1 in large cargo secretion (Saito et al., 2009). This is further supported by analysis of Mia3 knockout mice which fail to form a skeleton (Wilson et al., 2011). Consistent with this, immunofluorescence shows some defects in assembly of a type I collagen matrix in TANGO1 knockout cells, most notably in TANGO1L-/S- cells (Figure 6A). This was further supported by immunoblotting that shows an increase in intracellular type I procollagen in TANGO1 knockout cells (Fig. 6B). We then used a biotin-controllable reporter to monitor procollagen transport in these cells (McCaughey et al., 2019). We were unable to derive stable cell lines from all TANGO1 knockout clones, notably those with the most severely disrupted Golgi morphology (LΔ/S-2 and L-/S- 1); we interpret this as an inability of these cells to manage overexpression of procollagen in the background of an impaired secretory pathway. Unlike WT cells, all TANGO1 knockout cells were unable to transport this procollagen reporter from ER-to-Golgi within 60 minutes (Fig. 6C-G).

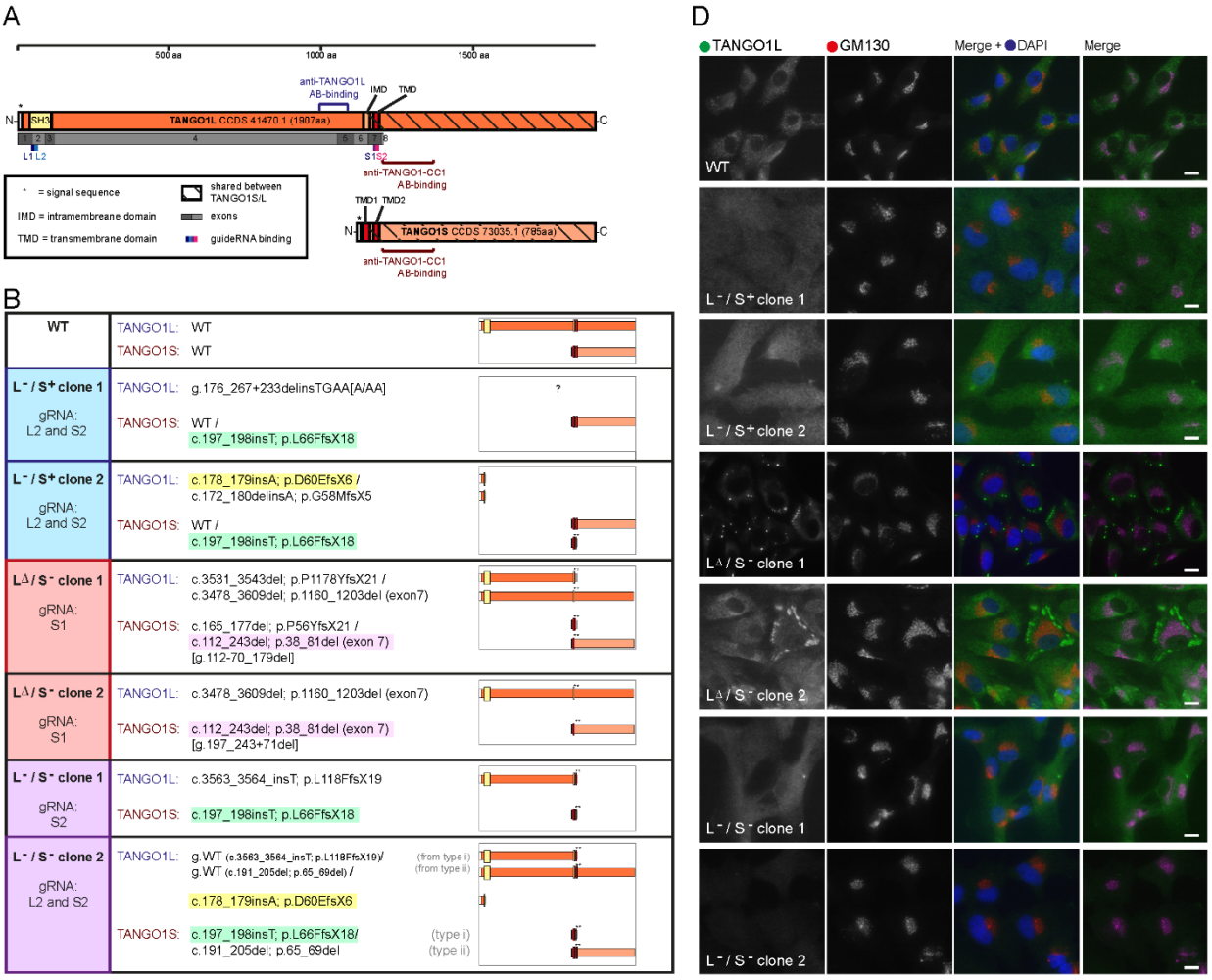

Fig. 1. (A) Schematic showing locations of gRNAs used and antibody epitopes in relation to the major isoforms encoded by Mia3, TANGO1L and TANGO1S. (B) Mapped genomic changes including a schematic of the predicted outcomes for encoded proteins. (C) Immunoblotting using TANGO1 antibodies that detect TANGO1L and TANGO1S (anti-TANGO1-CC1) with GAPDH used as a loading control. The same lysates were used to probe for Sec24C, TFG, Sec31, Sec24D. GAPDH or DIC74.1 was used as a loading control in each case, boxes indicate the individual gels that relate to those control blots. Lysates were also probed to detect TANGO1L, cTAGE5, Sec12, Sec24A, and Hsp47. (D) Immunofluorescence was used to confirm loss of TANGO1L expression (green in merge) and co-labelled to detect the Golgi using GM130 (red in merge). Bar = 10 µm.

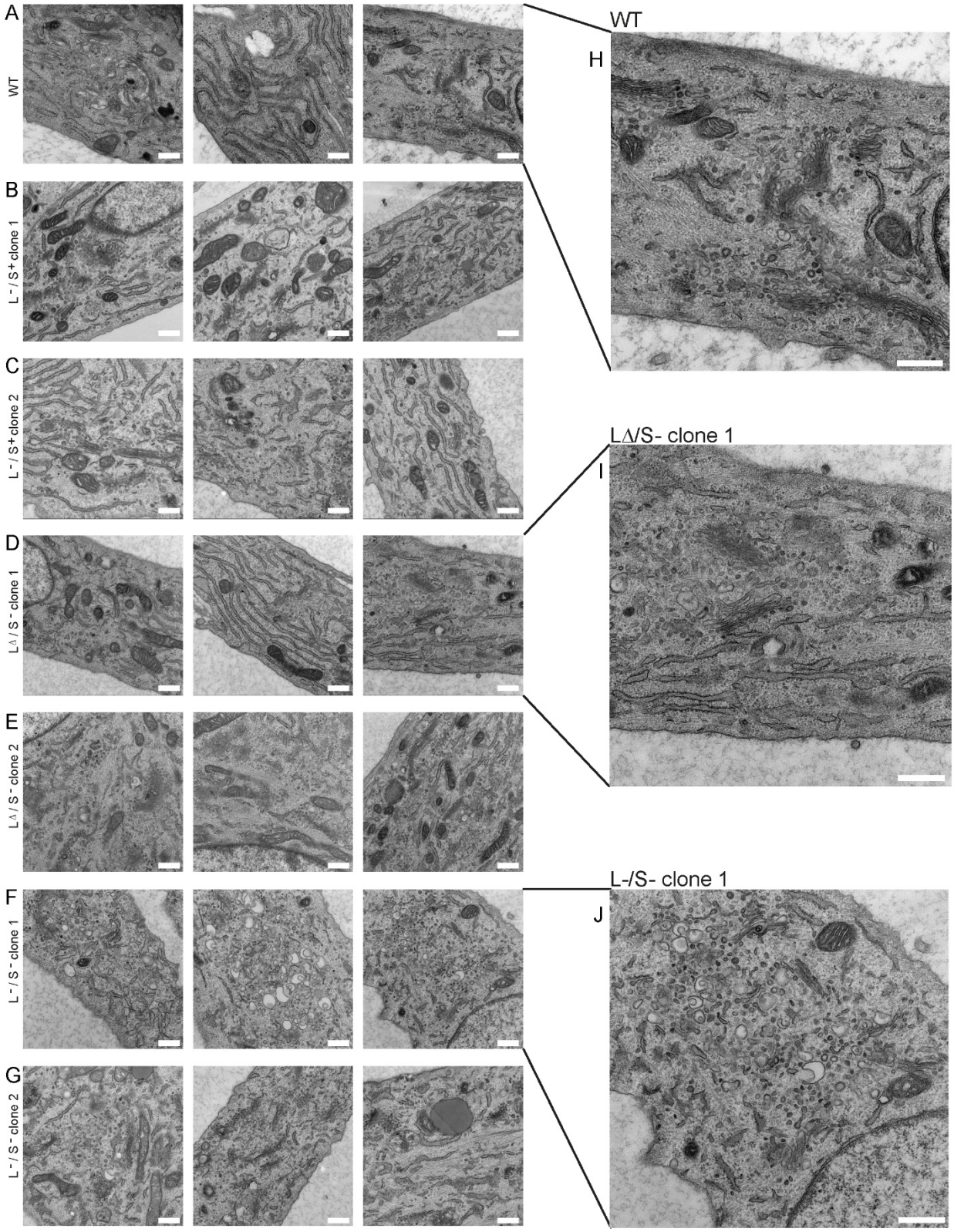

Fig. 2. (A-G) Transmission EM was used to identify Golgi membranes in TANGO1 KO cells compared to WT. Three examples of each are shown along with an enlarged version (H-J) of one from the WT, LΔ/S- clone 1 and L-/S- clone 1 cell lines. Bars = 500 nm.

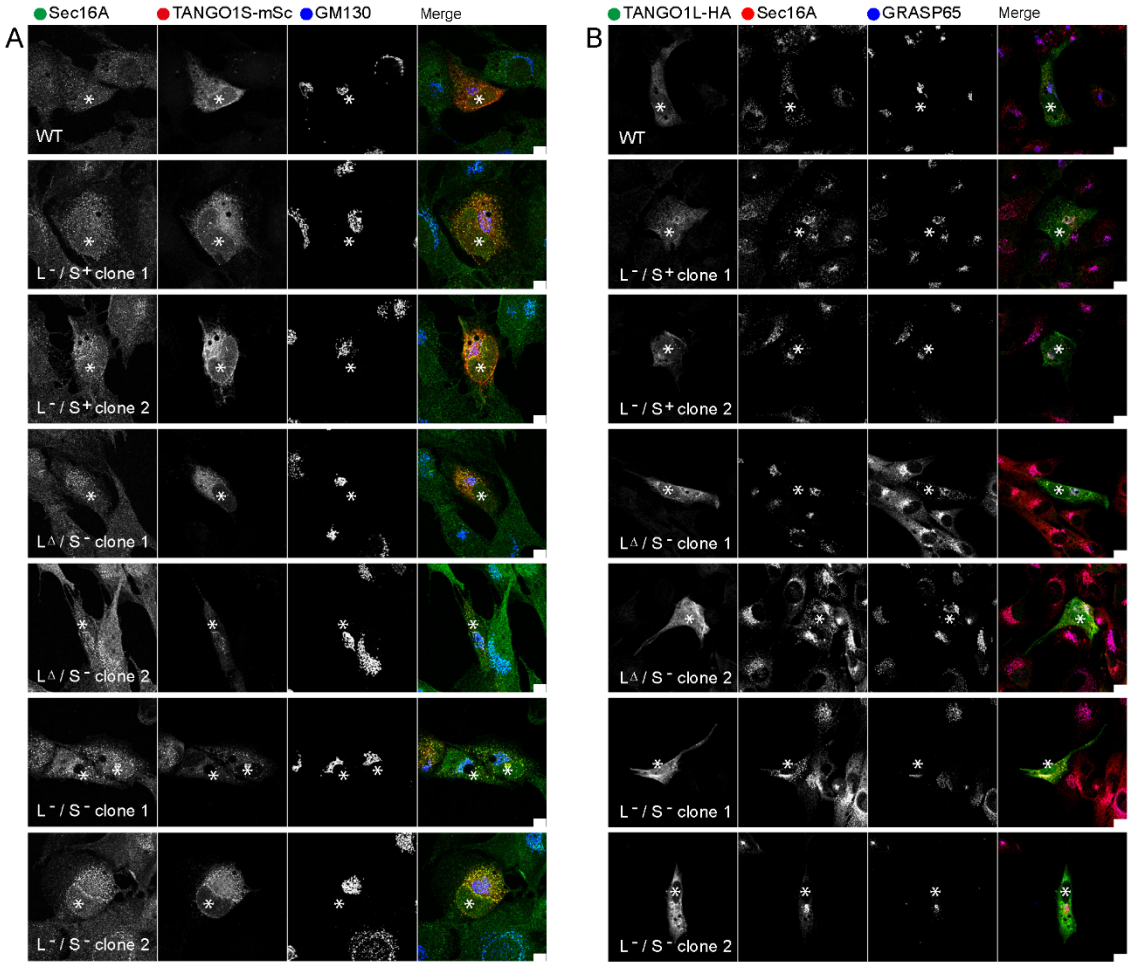

Fig. 3. Analysis of Sec16A (A) or Sec31A (B) localization and the Golgi apparatus (GM130 (A) and GRASP65 (B)) in cells expressing TANGO1S-mScarlet-i (A) or TANGO1L-HA (B). Bar = 10 µm. Asterisks highlight cells expressing the rescue constructs.

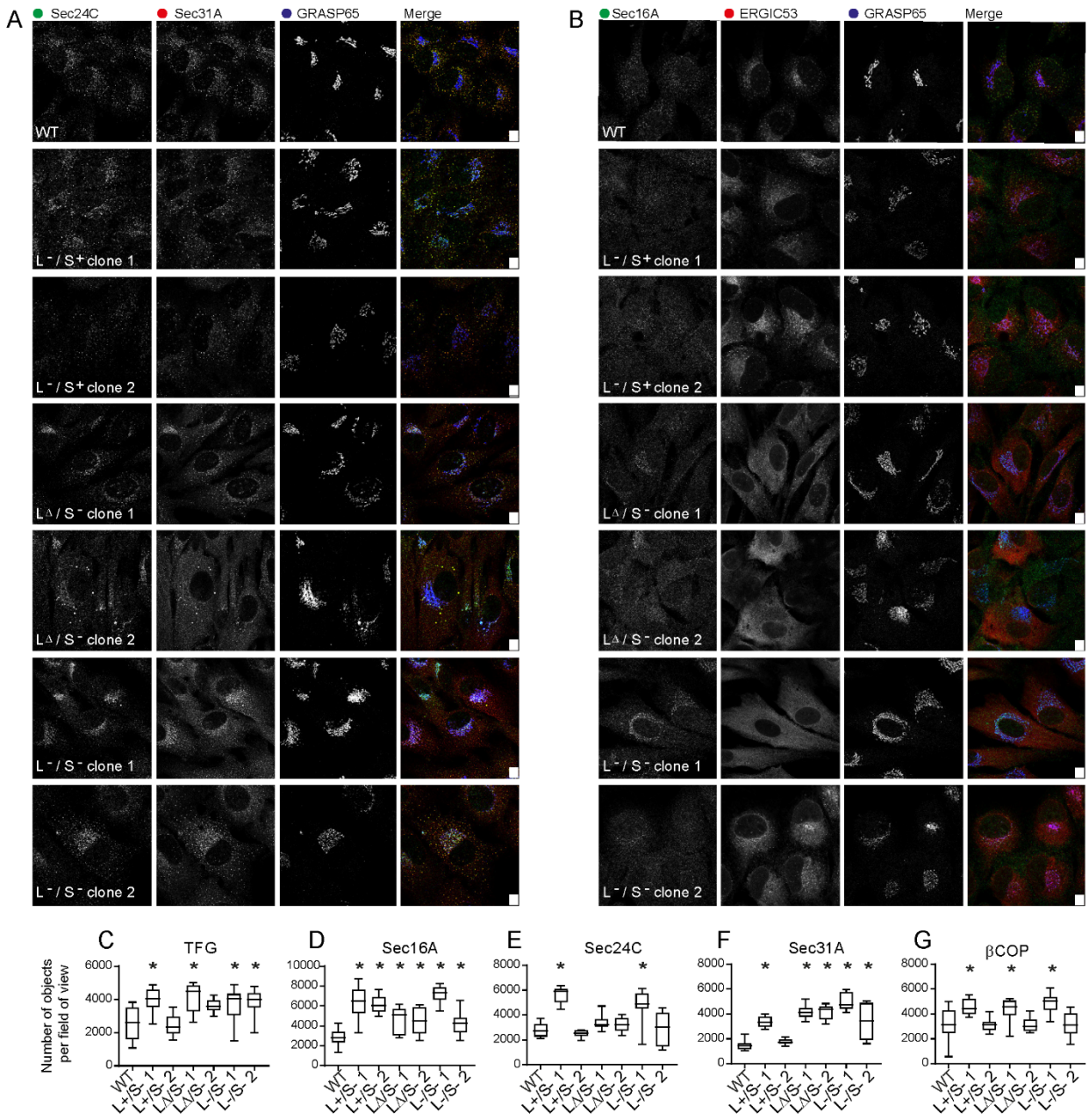

Fig. 4. Analysis of the distribution of COPII using immunofluorescence. (A) Sec24C, Sec31A, and GRASP65 labelling. (B) Sec16A, ERGIC53, and GRASP65 labelling. Bar = 10 µm. (C) Quantification of these data (number of objects per field of view) along with additional markers TFG and β-COP. Asterisks show p values <0.05 using Kruskal Wallis tests with Dunn's multiple comparison test for analysis of TFG and Sec24C where not all data were normally distributed, or from one-way ANOVA with Dunnett's multiple comparison test for Sec16A, Sec31A and β-COP where all the data were normally distributed.

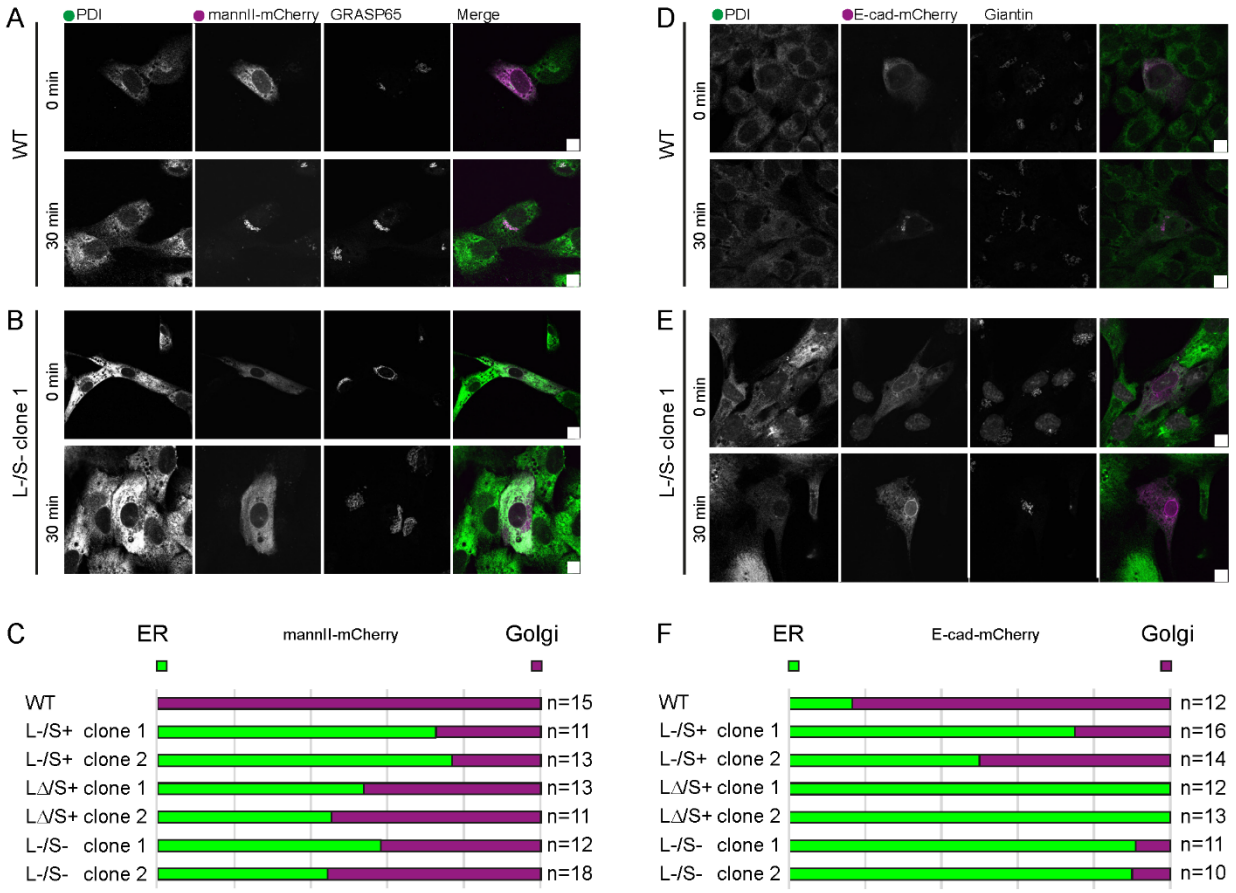

Fig. 5. Analysis of cargo trafficking using RUSH-engineered probes (A-C) mannII-mCherry and (D-F) E-cadherin-mCherry. PDI was used to define the ER and either (A-C) GRASP65 or (D-F) giantin to define the Golgi. Bar = 10 µm. (C) These data were then analyzed to determine the number of cells in which localization of the cargo protein to either the ER (green) or Golgi (magenta) was predominant at 30 minutes after the addition of biotin.

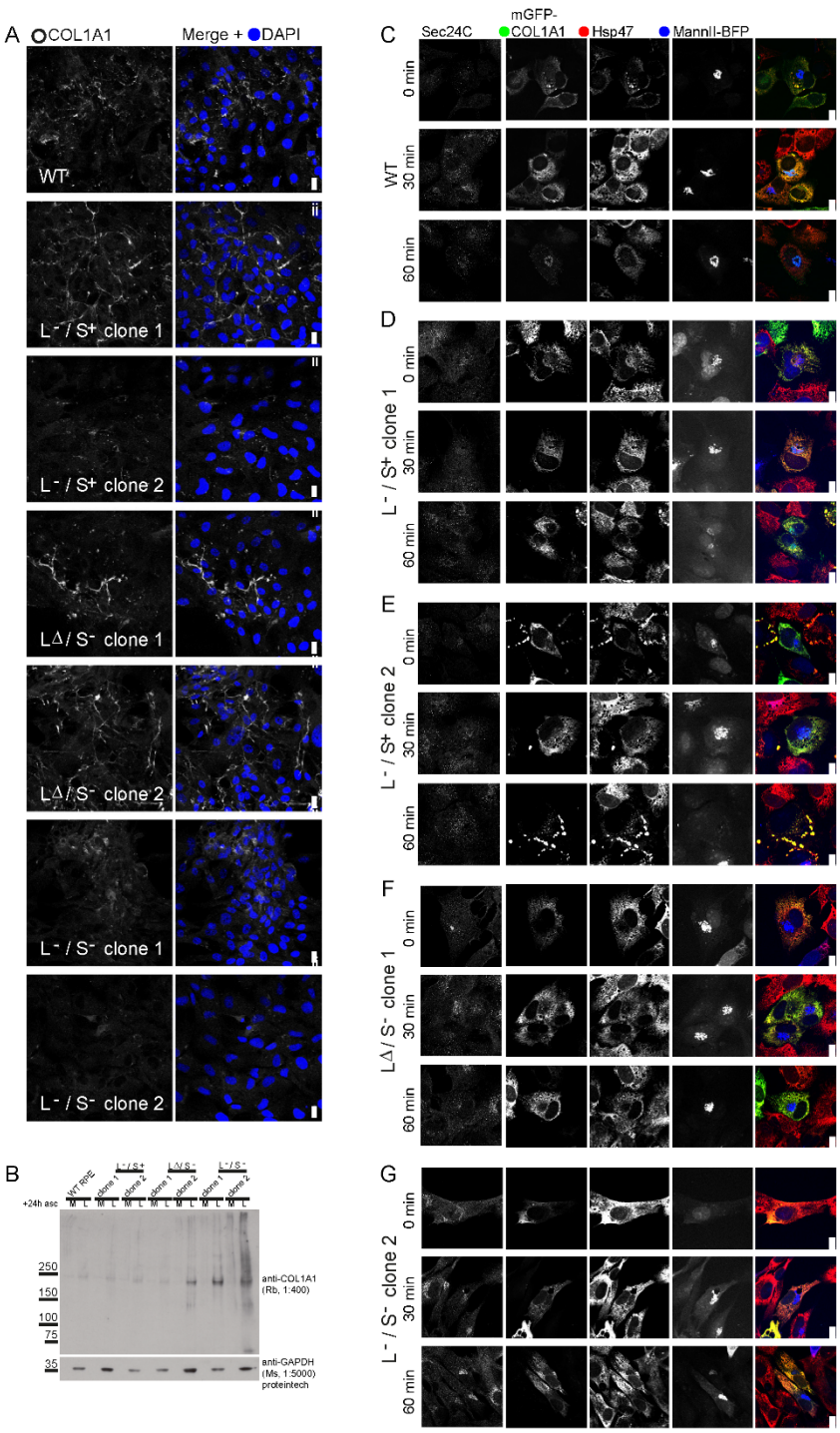

Fig. 6. (A) Analysis of collagen I localization. Bar = 10 µm. (B) Immunoblotting to detect type I procollagen in either media (M) or lysates (L) 24h after addition of ascorbate. GAPDH is included as a loading control. (C-G) Analysis of mGFP-COL1A1 trafficking using the RUSH system. Cells were co-labeled to detect Sec24C (not included in the merge image) and Hsp47. Transfected cells co-express a Golgi marker mannII-BFP with a separate ER-hook. Bar = 10 µm.

Discussion

We interpret these data as revealing that loss of TANGO1 expression results in dramatic morphological changes to the ER-Golgi interface coupled with significant changes in secretion. These primarily manifest as defects in extracellular matrix composition but do not correlate directly with changes in the secretion of procollagens over and above that of other proteins. TANGO1 is present in *Drosophila* which does not produce a fibrillar collagen matrix but does secrete some larger cargoes such as Dumpy (Rios-Barrera et al., 2017). In *Drosophila,* many cargoes are TANGO1-dependent (Liu et al., 2017) including type IV collagen, the sole collagen expressed. This work is also more consistent with a general role for TANGO1 in organization of ERES. Notably *C. elegans* do not express TANGO1 but TMEM131 (Zhang et al., 2020) and/or TMEM39 (Zhang et al., 2021) might play a similar role in collagen secretion in nematodes. Overall, our data support models where ECM proteins are the cohort of secretory cargo most sensitive to perturbation of early secretory pathway function. Thus, even in the absence of a collagen specific effect, the sensitivity of ECM assembly to loss of TANGO1 can still explain why bone and cartilage formation is the most dramatically affected process in Mia3 knockout *in vivo* (Wilson et al., 2011).

It is important to note that knockout of TANGO1L in our cells causes concomitant reduction in expression of cTAGE5 encoded by the Mia2 gene, likely due to interdependent stabilization of the cTAGE5-TANGO1 complex (Ma and Goldberg, 2016; Maeda et al., 2016). We consider this also likely in other systems. Thus, we cannot attribute specific phenotypes solely to Mia3 but instead must consider this as a holistic system that supports ER export. Our data notably reveal only limited defects following loss of TANGO1L only. Severe perturbation in cell organization and secretory function are only seen following loss of both isoforms. We consider that this further supports a primary role for the cytosolic domain of Mia3 isoforms. Additional roles for the luminal domain of TANGO1L in binding to Hsp47 and other cargo might further facilitate export of some cargo. However, we conclude that the core function of Mia3 is linked to its cytosolic domain facilitating endomembrane organization, limiting the formation of smaller vesicles and promoting formation and maintenance of an ERGIC compartment.

The loss of all punctate ERGIC53 labelling in TANGO1 L-/S- cells shows that the Mia2-Mia3 system has a fundamental role in maintaining the ERGIC as a steady-state compartment. This is consistent with the identification of regions including the first coiled coil of TANGO1 named TEER (Tether for ERGIC at the ER) that was reported to recruit ERGIC membranes to ERES to promote membrane expansion (Santos et al., 2015). Our data suggest that this domain has a key role in the biogenesis and/or maintenance of the ERGIC itself. Here, TANGO1 could limit COPII vesicle formation through delaying scission of COPII coated vesicles and facilitating formation of more amorphous tubule-vesicular structures that then bud to become independent from the underlying ER. Proline-rich domains (PRDs) of TANGO1 and cTAGE5 bind to multiple copies of Sec23 initially competing with Sec31 and thereby limiting recruitment of the outer layer of the COPII coat (Ma and Goldberg, 2016). This itself would limit GTPase stimulation by Sec13-Sec31 promoting bud growth. Such a model does not require formation of “megavesicles” and could result in more tubular ER export domains. This is consistent both with previous EM analyses suggesting *en bloc* protrusion of carriers (Mironov et al., 2003) and with the lack of obvious large carriers on live cell imaging of GFP-procollagen (McCaughey et al., 2019). TFG, notably upregulated in TANGO1 KO cells (Fig. 1), could also have a key role here in generating a local environment that promotes and stabilizes this process (Johnson et al., 2015; McCaughey et al., 2016; Witte et al., 2011).

In summary, our data show that TANGO1 plays an essential role in the organization of the ER-Golgi interface in mammalian cells and highlight the fundamental importance of endomembrane organization for effective secretion. TANGO1 promotes secretion by limiting the formation of COPII vesicles at ERES and facilitating the maturation of membrane-bound compartments from an ER to an ERGIC and finally Golgi identity. Cell-type specific tuning of Mia2 and Mia3 isoform expression could tip the relative balance of ER export carrier forms to adapt to cases of increased secretory load and adaptation for specific cargo types.

References:

Bard, F., L. Casano, A. Mallabiabarrena, E. Wallace, K. Saito, H. Kitayama, G. Guizzunti, Y. Hu, F. Wendler, R. Dasgupta, N. Perrimon, and V. Malhotra. 2006. Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature*. 439:604-607.

Bednarek, S.Y., M. Ravazzola, M. Hosobuchi, M. Amherdt, A. Perrelet, R. Schekman, and L. Orci. 1995. COPI- and COPII-coated vesicles bud directly from the endoplasmic reticulum in yeast. *Cell*. 83:1183-1196.

Bindels, D.S., L. Haarbosch, L. van Weeren, M. Postma, K.E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M.A. Hink, and T.W. Gadella, Jr. 2017. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods*. 14:53-56.

Boncompain, G., S. Divoux, N. Gareil, H. de Forges, A. Lescure, L. Latreche, V. Mercanti, F. Jollivet, G. Raposo, and F. Perez. 2012a. Synchronization of secretory protein traffic in populations of cells. *Nat Methods*. 9:493-498.

Boncompain, G., S. Divoux, N. Gareil, H. de Forges, A. Lescure, L. Latreche, V. Mercanti, F. Jollivet, G. Raposo, and F. Perez. 2012b. Synchronization of secretory protein traffic in populations of cells. *Nature methods*. 9:493-498.

Cheng, J.P., V.M. Betin, H. Weir, G.M. Shelmani, D.K. Moss, and J.D. Lane. 2010. Caspase cleavage of the Golgi stacking factor GRASP65 is required for Fas/CD95-mediated apoptosis. *Cell Death Dis*. 1:e82.

Clark, E.M., and B.A. Link. 2021. Complementary and divergent functions of zebrafish Tango1 and Ctage5 in tissue development and homeostasis. *Mol Biol Cell*:mbcE20110745.

Espenshade, P., R.E. Gimeno, E. Holzmacher, P. Teung, and C.A. Kaiser. 1995. Yeast SEC16 gene encodes a multidomain vesicle coat protein that interacts with Sec23p. *J Cell Biol*. 131:311-324.

Fromme, J.C., M. Ravazzola, S. Hamamoto, M. Al-Balwi, W. Eyaid, S.A. Boyadjiev, P. Cosson, R. Schekman, and L. Orci. 2007. The genetic basis of a craniofacial disease provides insight into COPII coat assembly. *Dev Cell*. 13:623-634.

Guillemyn, B., S. Nampoothiri, D. Syx, F. Malfait, and S. Symoens. 2021. Loss of TANGO1 leads to an absence of bone mineralization. *JMBR Plus*:e10451.

Hughes, H., A. Budnik, K. Schmidt, K.J. Palmer, J. Mantell, C. Noakes, A. Johnson, D.A. Carter, P. Verkade, P. Watson, and D.J. Stephens. 2009. Organisation of human ER-exit sites: requirements for the localisation of Sec16 to transitional ER. *J Cell Sci*. 122:2924-2934.

Ishikawa, Y., S. Ito, K. Nagata, L.Y. Sakai, and H.P. Bachinger. 2016. Intracellular mechanisms of molecular recognition and sorting for transport of large extracellular matrix molecules. *Proc Natl Acad Sci U S A*. 113:E6036-E6044.

Johnson, A., N. Bhattacharya, M. Hanna, J.G. Pennington, A.L. Schuh, L. Wang, M.S. Otegui, S.M. Stagg, and A. Audhya. 2015. TFG clusters COPII-coated transport carriers and promotes early secretory pathway organization. *EMBO J*. 34:811-827.

Lekszas, C., O. Foresti, I. Raote, D. Liedtke, E.M. Konig, I. Nanda, B. Vona, P. De Coster, R. Cauwels, V. Malhotra, and T. Haaf. 2020. Biallelic TANGO1 mutations cause a novel syndromal disease due to hampered cellular collagen secretion. *Elife*. 9.

Liu, M., Z. Feng, H. Ke, Y. Liu, T. Sun, J. Dai, W. Cui, and J.C. Pastor-Pareja. 2017. Tango1 spatially organizes ER exit sites to control ER export. *J Cell Biol*. 216:1035-1049.

Ma, W., and J. Goldberg. 2016. TANGO1/cTAGE5 receptor as a polyvalent template for assembly of large COPII coats. *Proc Natl Acad Sci U S A*. 113:10061-10066.

Maeda, M., T. Katada, and K. Saito. 2017. TANGO1 recruits Sec16 to coordinately organize ER exit sites for efficient secretion. *J Cell Biol*. 216:1731-1743.

Maeda, M., K. Saito, and T. Katada. 2016. Distinct isoform-specific complexes of TANGO1 cooperatively facilitate collagen secretion from the endoplasmic reticulum. *Mol Biol Cell*. 27:2688-2696.

Martinez-Menarguez, J.A., H.J. Geuze, J.W. Slot, and J. Klumperman. 1999. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell*. 98:81-90.

Matsuoka, K., L. Orci, M. Amherdt, S.Y. Bednarek, S. Hamamoto, R. Schekman, and T. Yeung. 1998. COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell*. 93:263-275.

McCaughey, J., V.J. Miller, N.L. Stevenson, A.K. Brown, A. Budnik, K.J. Heesom, D. Alibhai, and D.J. Stephens. 2016. TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion. *Cell Rep*. 15:1648-1659.

McCaughey, J., and D.J. Stephens. 2019. ER-to-Golgi Transport: A Sizeable Problem. *Trends Cell Biol*. 29:940-953.

McCaughey, J., N.L. Stevenson, S. Cross, and D.J. Stephens. 2019. ER-to-Golgi trafficking of procollagen in the absence of large carriers. *J Cell Biol*. 218:929-948.

Mironov, A.A., A.A. Mironov, Jr., G.V. Beznoussenko, A. Trucco, P. Lupetti, J.D. Smith, W.J. Geerts, A.J. Koster, K.N. Burger, M.E. Martone, T.J. Deerinck, M.H. Ellisman, and A. Luini. 2003. ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. *Dev Cell*. 5:583-594.

Nogueira, C., P. Erlmann, J. Villeneuve, A.J. Santos, E. Martinez-Alonso, J.A. Martinez-Menarguez, and V. Malhotra. 2014. SLY1 and Syntaxin 18 specify a distinct pathway for procollagen VII export from the endoplasmic reticulum. *Elife*. 3:e02784.

Notredame, C., D.G. Higgins, and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol*. 302:205-217.

Orci, L., A. Perrelet, M. Ravazzola, M. Amherdt, J.E. Rothman, and R. Schekman. 1994. Coatomer-rich endoplasmic reticulum. *Proc Natl Acad Sci U S A*. 91:11924-11928.

Palmer, K.J., H. Hughes, and D.J. Stephens. 2009. Specificity of cytoplasmic dynein subunits in discrete membrane-trafficking steps. *Mol Biol Cell*. 20:2885-2899.

Palmer, K.J., J.E. Konkel, and D.J. Stephens. 2005. PCTAIRE protein kinases interact directly with the COPII complex and modulate secretory cargo transport. *J Cell Sci*. 118:3839-3847.

Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D.J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Perez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A.F. Jarnuczak, T. Ternent, A. Brazma, and J.A. Vizcaino. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 47:D442-D450.

Raote, I., and V. Malhotra. 2021. Tunnels for Protein Export from the Endoplasmic Reticulum. *Annu Rev Biochem*.

Raote, I., M. Ortega-Bellido, A.J. Santos, O. Foresti, C. Zhang, M.F. Garcia-Parajo, F. Campelo, and V. Malhotra. 2018. TANGO1 builds a machine for collagen export by recruiting and spatially organizing COPII, tethers and membranes. *Elife*. 7.

Raote, I., M. Ortega Bellido, M. Pirozzi, C. Zhang, D. Melville, S. Parashuraman, T. Zimmermann, and V. Malhotra. 2017. TANGO1 assembles into rings around COPII coats at ER exit sites. *J Cell Biol*. 216:901-909.

Rezaei, N., A. Lyons, and N.R. Forde. 2018. Environmentally Controlled Curvature of Single Collagen Proteins. *Biophys J*. 115:1457-1469.

Rios-Barrera, L.D., S. Sigurbjornsdottir, M. Baer, and M. Leptin. 2017. Dual function for Tango1 in secretion of bulky cargo and in ER-Golgi morphology. *Proc Natl Acad Sci U S A*. 114:E10389-E10398.

Saito, K., M. Chen, F. Bard, S. Chen, H. Zhou, D. Woodley, R. Polischuk, R. Schekman, and V. Malhotra. 2009. TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell*. 136:891-902.

Saito, K., K. Yamashiro, Y. Ichikawa, P. Erlmann, K. Kontani, V. Malhotra, and T. Katada. 2011. cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites. *Mol Biol Cell*. 22:2301-2308.

Santos, A.J., C. Nogueira, M. Ortega-Bellido, and V. Malhotra. 2016. TANGO1 and Mia2/cTAGE5 (TALI) cooperate to export bulky pre-chylomicrons/VLDLs from the endoplasmic reticulum. *J Cell Biol*. 213:343-354.

Santos, A.J., I. Raote, M. Scarpa, N. Brouwers, and V. Malhotra. 2015. TANGO1 recruits ERGIC membranes to the endoplasmic reticulum for procollagen export. *Elife*. 4.

Satchwell, T.J., S. Pellegrin, P. Bianchi, B.R. Hawley, A. Gampel, K.E. Mordue, A. Budnik, E. Fermo, W. Barcellini, D.J. Stephens, E. van den Akker, and A.M. Toye. 2013. Characteristic phenotypes associated with congenital dyserythropoietic anemia (type II) manifest at different stages of erythropoiesis. *Haematologica*. 98:1788-1796.

Scales, S.J., R. Pepperkok, and T.E. Kreis. 1997. Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. *Cell*. 90:1137-1148.

Schweizer, A., J.A. Fransen, T. Bachi, L. Ginsel, and H.P. Hauri. 1988. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *J Cell Biol*. 107:1643-1653.

Townley, A.K., Y. Feng, K. Schmidt, D.A. Carter, R. Porter, P. Verkade, and D.J. Stephens. 2008. Efficient coupling of Sec23-Sec24 to Sec13-Sec31 drives COPII-dependent collagen secretion and is essential for normal craniofacial development. *J Cell Sci*. 121:3025-3034.

Weissman, J.T., H. Plutner, and W.E. Balch. 2001. The mammalian guanine nucleotide exchange factor mSec12 is essential for activation of the Sar1 GTPase directing endoplasmic reticulum export. *Traffic*. 2:465-475.

Wilson, D.G., K. Phamluong, L. Li, M. Sun, T.C. Cao, P.S. Liu, Z. Modrusan, W.N. Sandoval, L. Rangell, R.A. Carano, A.S. Peterson, and M.J. Solloway. 2011. Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse. *J Cell Biol*. 193:935-951.

Witte, K., A.L. Schuh, J. Hegermann, A. Sarkeshik, J.R. Mayers, K. Schwarze, J.R. Yates, 3rd, S. Eimer, and A. Audhya. 2011. TFG-1 function in protein secretion and oncogenesis. *Nat Cell Biol*. 13:550-558.

Zhang, Z., M. Bai, G.O. Barbosa, A. Chen, Y. Wei, S. Luo, X. Wang, B. Wang, T. Tsukui, H. Li, D. Sheppard, T.B. Kornberg, and D.K. Ma. 2020. Broadly conserved roles of TMEM131 family proteins in intracellular collagen assembly and secretory cargo trafficking. *Sci Adv*. 6:eaay7667.

Zhang, Z., S. Luo, G.O. Barbosa, M. Bai, T.B. Kornberg, and D.K. Ma. 2021. The conserved transmembrane protein TMEM-39 coordinates with COPII to promote collagen secretion and regulate ER stress response. *PLoS Genet*. 17:e1009317.

Acknowledgements

We thank Nicola Stevenson and other members of the Stephens lab for continued discussion and help with this work, the Wolfson Bioimaging Facility for support and advice, and Paul Martin, and Chrissy Hammond (University of Bristol, UK) and Karl Kadler, Martin Lowe, Qing-Jun Meng, Joe Swift, and Oliver Jensen (University of Manchester, UK) for helpful discussions.

Funding

This work was funded by a postgraduate research scholarship from the University of Bristol to JM with additional research grant support from UKRI-MRC and UKRI-BBSRC (MR/P000177/1 and BB/T001984/1). We gratefully acknowledge the support of the Wolfson Bioimaging Facility for this work, and UKRI-BBSRC for funding equipment used in this study (BB/L014181/1 and through BrisSynBio, a BBSRC/EPSRC-funded Synthetic Biology Research Centre (grant number: BB/L01386X/1).

Author contributions

Conceptualization, Data curation, Methodology, Formal Analysis, Validation, Visualization, Writing – original draft, Writing – review & editing: JM and DJS.

Investigation: JM (all experiments), CRN (EM embedding and sectioning), JMM (EM imaging and analysis), KH (proteomics).

Funding acquisition, Project administration, Resources, Supervision, and Writing original draft: DJS.

Competing Interests

The authors declare no competing interests. The funders had no role in the study design.

Reagent and Data Availability

All plasmids are available through Addgene: https://www.addgene.org/David_Stephens/. Proteomics data are available on PRIDE: Image data are available via our institutional repository at: . The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD024214 for cell-derived matrix proteomes and PXD024221 for soluble proteomes.

Materials and Methods

All chemicals and reagents were obtained from Sigma (now Merck, Watford, UK) if not stated otherwise.

Cell culture and generation of CRISPR-KO cell lines

A human telomerase reverse transcriptase-immortalised retinal pigment epithelium type 1 cell line (hTERT RPE-1, hereafter referred to as RPE-1; ATCC® CRL-4000) was used for all experiments, including the generation of stable cell lines. Cells were cultured at 37 °C and 5% CO2 in a humid environment. RPE-1 cells were grown in Dulbecco's modified Eagle medium (DMEM) F12 supplemented with 10 % decomplemented foetal bovine serum (FBS; Thermo Fisher Scientific). Cells were passaged every 3 - 4 days when a confluence of about 80 % was reached. For passaging cells were rinsed with phosphate buffered saline (PBS) following treatment with 0.05% trypsin-EDTA (Thermo Fisher Scientific) at 37 °C until cell detachment.

The CRISPR-KO cell lines were generated using the *TrueGuide Synthetic CRISPR gRNA* system (Thermo Fisher Scientific) with custom gRNA synthesis. GuideRNA sequences were obtained from CRISPR-KO library against MIA3 from Sigma-Aldrich. RPE-1 cells were transfected following the manufacturer's protocol using TrueCut™ Cas9 Protein v2 and 30 pmol gRNA duplex targeting the SH3-domain in TANGO1L (guide L1: cggtgaggctcttgaagatt or guide L2: ggattgtcgttttgtgaatt) and or the TMD in exon 7 present in both TANGO1S/L (guide S1: tgataaatacaggtttcca or guide S2: aacgaagcaattcccaaga). Cells were either transfected with only guides S1 or S2, or with both gRNAs S1+L1 or S2+L2. Clonal populations were obtained using single cell sorting and maintained in conditioned media (0.45 µm filtered one day old media supplemented with pen/strep from a confluent parent RPE-1 dish). Surviving clones were screened via immunoblotting targeting TANGO1L (2 ng per mL rabbit polyclonal anti-MIA3, Sigma-Aldrich Prestige, HPA056816-100UL) and a rabbit polyclonal targeting anti-TANGO1-CC1 (1211 – 1440aa/exon9-15 shared between both TANGO1S/L (a gift from Kota Saito, (Maeda et al., 2016))).

Stable GFP-COL1A1 expressing TANGO1 knockout cell lines were generated using irus containing the GFP-COL1A1 construct as described previously ((McCaughey et al., 2019)). In brief, the Lenti-XTM Packaging Single Shots (vesicular stomatitis glycoprotein pseudotyped version) system from Takara Bio Europe was used according to the manufacturer's instructions (631275). Growth medium was removed from an 80% confluent 6-cm dish of RPE-1, and 1 mL harvested virus supernatant supplemented with 8 µg·mL−1 polybrene (Santa Cruz Biotechnologies) was added to cells. After 1 h of incubation at 37°C and 5% CO2, 5 mL growth medium was added. Transfection medium was then replaced with fresh growth medium after 24 h. To select for transfected cells, cells were passaged in growth medium supplemented with 15 µg·mL−1 puromycin dihydrochloride (Santa Cruz Biotechnology) 72 h after transfection and sorted via fluorescence activated cell sorting according to the signal intensity of GFP. Media for GFP-COL-RPE was further supplemented with 5 µg·mL-1 puromycin maintain engineered cell lines.

Genotyping

Genomic DNA was obtained from clonal MIA3-KO populations using the PureLink genomic DNA extraction kit (Thermo Fisher Scientific) according to the manufacturer's protocol. Regions of interest were amplified by PCR using a One-Taq Hot Start DNA-polymerase (New England Biolabs) with standard buffer and 3% DMSO and the following primers targeting exon2 (L) and exon7 (S):

1kb product:

MIA3-S_F1: cagttatcaggagctatccg MIA3-S_R1: cggctgactggttatttctttagg	MIA3-L_F1: catgtgtggtagttggcacattgc MIA3-L_R1: atttcaacctctaatacgtatgcagc

5kb product:

MIA3-S_5kb_F: attaggaaggtcttgcc MIA3-S_5kb_R: catcttcctgaaagggc	MIA3-L_5kb_F: ctttgcccttctgctttattgg MIA3-L_5kb_R: ttgatctgaaaactatctgaaagcc

PCR programme:

	PCR of genomic DNA	PCR of cDNA
	1. 94°C; 2 min	1. 94°C; 0.5 min
35 cycles	2. 94°C; 1 min	2. 94°C; 0.5 min
	3. 50°C for 1kb product/ 45°C for 5kb product; 1 min	3. 60°C 1 min
	4. 68°C; 1 min	4. 68°C; 1 min
	5. 68°C; 10 min	5. 68°C; 6 min
	6. 4°C	6. 4°C

For sequence analysis of the mRNA, cDNA was generated using SuperScript III Reverse Transcription (Thermo Fisher Scientific) of extracted total mRNA using the RNeasy Plus MiniKit (Qiagen) using DTT as reducing agent. Amplification of cDNA was done as mentioned above using the following primers: MIA3-S-RT_F: ATGGCTGCGGCGCCTG, MIA3-S-RT_R: ACCTGCCACTGTGCCTTCTATCG and no supplementation of DMSO.

PCR products were ligated with the pGEM-T Easy vector system (Promega) according to the manufacturer's protocol and transformed into DH5α *E. coli* (New England Biolabs). The plasmid DNA was extracted from 3-9 positive colonies where possible after blue-white screening using a QIAprep® Spin Miniprep Kit (Qiagen) and sequenced using the standard T7 primer at Eurofins Genomics. Resulting sequences were compared using multiple sequence alignment via M-Coffee (Notredame et al., 2000) accessed at http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee and displayed with the help of version 3.21 of BOXSHADE, written by K. Hofmann and M. Baron https://embnet.vital-it.ch/software/BOX_form.html.

DNA constructs

Constructs were either generated for this study or acquired from Addgene (numbers indicated by #). The TANGO1L-HA construct was a gift from Vivek Malhotra (Saito et al., 2009). All restriction and modifying enzymes were purchased from New England Biolabs (Hitchin, UK). Str-KDEL-IRES-mannosidaseII-mScarlet-i (ManII-mSc; #117274) and procollagen-SBP-mGFP-COL1A1 (#110726) were described in (McCaughey et al., 2019). StrKDEL-IRES-mannosidase II-mTagBFP2 (ManII-BFP; #165461) was generated by using ManII-mSc as a template and replacing the mScarlet-i with a mTagBFP2 generated as a synthetic gene block by Integrated DNA Technologies via restriction digest using EcoRI and FseI and subsequent HiFi NEBuilder assembly (New England Biolabs).

Synthetic gene block from IDT for generating ManII-BFP:

TCCCACCGGTCGCCACCGGaattccATGAGCGAGCTGATTAAGGAGAACATGCACATGAAGCTGTACATGGAGGGCA
CCGTGGACAACCATCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATC
AAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCCTTCGACATCCTGGCTACTAGCTTCCTCTACGGCAGCAAGACCTT
CATCAACCACACCCAGGGCATCCCCGACTTCTTCAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCA
CATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAG
ATCAGAGGGGTGAACTTCACATCCAACGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTTCACCGAGAC
GCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAACGACATGGCCCTGAAGCTCGTGGGCGGGAGCCATCTGATCG
CAAACATCAAGACCACATATAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCTGGCGTCTACTATGTGGACTAC
AGACTGGAAAGAATCAAGGAGGCCAACAACGAGACCTACGTCGAGCAGCACGAGGTGGCAGTGGCCAGATACTGCGA
CCTCCCTAGCAAACTGGGGCACAAGCTTAATggccggcctTAAggcctcgagGGCC

All other constructs for RUSH experiments were gifts from Franck Perez (Institut Curie, Paris, France: Str-KDEL-IRES-ST-SBP-mCherry (# 65265), Str-KDEL-IRES-mannosidaseII-SBP-mCherry (#65253) and Str-KDEL_SBP-mCherry-Ecadherin (#65287) (Boncompain et al., 2012a).

A TANGO1S-mScarlet-i (TANGO1S-mSc; Addgene #165460) construct was generated using a two-step cloning. First, StrKDEL-IRES from ManII-mSc was replaced with the human coding sequence for TANGO1S (CCDS 73035.1) via restriction digest of ManII-mSc with EcoRV and AgeI and subsequent NEBuilder HiFi assembly with the synthetic gene block containing TANGO1S with sequence overlaps at both ends.

Synthetic gene block from IDT for generating TANGO1S-mSc (step 1):

aagcttggtaccgagctcggatcgatatcgcaAGCGCTgcaATGGACTCAGTACCTGCCACTGTGCC
TTCTATCGCCGCTACCCCGGGGGACCCGGAACTTGTGGGACCCTTGTCTGTGCTCTACGCAGCCTTCATA
GCCAAGCTGCTGGAGCTAGTTGCTACATTGCCTGATGATGTTCAGCCTGGGCCTGATTTTTATGGACTGC
CATGGAAACCTGTATTTATCACTGCCTTCTTGGGAATTGCTTCGTTTGCCATTTTCTTATGGAGAACTGT
CCTTGTTGTGAAGGATAGAGTATATCAAGTCACGGAACAGCAAATTTCTGAGAAGTTGAAGACTATCATG
AAAGAAAATACAGAACTTGTACAAAAATTGTCAAATTATGAACAGAAGATCAAGGAATCAAAGAAACATG
TTCAGGAAACCAGGAAACAAAATATGATTCTCTCTGATGAAGCAATTAAATATAAGGATAAAATCAAGAC
ACTTGAAAAAAATCAGGAAATTCTGGATGACACAGCTAAAAATCTTCGTGTTATGCTAGAATCTGAGAGA
GAACAGAATGTCAAGAATCAGGACTTGATATCAGAAAACAAGAAATCTATAGAGAAGTTAAAGGATGTTA
TTTCAATGAATGCCTCAGAATTTTCAGAGGTTCAGATTGCACTTAATGAAGCTAAGCTTAGTGAAGAGAA
GGTGAAGTCTGAATGCCATCGGGTTCAAGAAGAAAATGCTAGGCTTAAGAAGAAAAAAGAGCAGTTGCAG
CAGGAAATCGAAGACTGGAGTAAATTACATGCTGAGCTCAGTGAGCAAATCAAATCATTTGAGAAGTCTC
AGAAAGATTTGGAAGTAGCTCTTACTCACAAGGATGATAATATTAATGCTTTGACTAACTGCATTACACA
GTTGAATCTGTTAGAGTGTGAATCTGAATCTGAGGGTCAAAATAAAGGTGGAAATGATTCAGATGAATTA
GCAAATGGAGAAGTGGGAGGTGACCGGAATGAGAAGATGAAAAATCAAATTAAGCAGATGATGGATGTCT
CTCGGACACAGACTGCAATATCGGTAGTTGAAGAGGATCTAAAGCTTTTACAGCTTAAGCTAAGAGCCTC
CGTGTCCACTAAATGTAACCTGGAAGACCAGGTAAAGAAATTGGAAGATGACCGCAACTCACTACAAGCT
GCCAAAGCTGGACTGGAAGATGAATGCAAAACCTTGAGGCAGAAAGTGGAGATTCTGAATGAGCTCTATC
AGCAGAAGGAGATGGCTTTGCAAAAGAAACTGAGTCAAGAAGAGTATGAACGGCAAGAAAGAGAGCACAG
GCTGTCAGCTGCAGATGAAAAGGCAGTTTCGGCTGCAGAGGAAGTAAAAACTTACAAGCGGAGAATTGAA
GAAATGGAGGATGAATTACAGAAGACAGAGCGGTCATTTAAAAACCAGATCGCTACCCATGAGAAGAAAG
CTCATGAAAACTGGCTCAAAGCTCGTGCTGCAGAAAGAGCTATAGCTGAAGAGAAAAGGGAAGCTGCCAA
TTTGAGACACAAATTATTAGAATTAACACAAAAGATGGCAATGCTGCAAGAAGAACCTGTGATTGTAAAA
CCAATGCCAGGAAAACCAAATACACAAAACCCTCCACGGAGAGGTCCTCTGAGCCAGAATGGCTCTTTTG
GCCCATCCCCTGTGAGTGGTGGAGAATGCTCCCCTCCATTGACAGTGGAGCCACCCGTGAGACCTCTCTC
TGCTACTCTCAATCGAAGAGATATGCCTAGAAGTGAATTTGGATCAGTGGACGGGCCTCTACCTCATCCT
CGATGGTCAGCTGAGGCATCTGGGAAACCCTCTCCTTCTGATCCAGGATCTGGTACAGCTACCATGATGA
ACAGCAGCTCAAGAGGCTCTTCCCCTACCAGGGTACTCGATGAAGGCAAGGTTAATATGGCTCCAAAAGG
GCCCCCTCCTTTCCCAGGAGTCCCTCTCATGAGCACCCCCATGGGAGGCCCTGTACCACCACCCATTCGA
TATGGACCACCACCTCAGCTCTGCGGACCTTTTGGGCCTCGGCCACTTCCTCCACCCTTTGGCCCTGGTA
TGCGTCCACCACTAGGCTTAAGAGAATTTGCACCAGGCGTTCCACCAGGAAGACGGGACCTGCCTCTCCA
CCCTCGGGGATTTTTACCTGGACACGCACCATTTAGACCTTTAGGTTCACTTGGCCCAAGAGAGTACTTT
ATTCCTGGTACCCGATTACCACCCCCAACCCATGGTCCCCAGGAATACCCACCACCACCTGCTGTAAGAG
ACTTACTGCCGTCAGGCTCTAGAGATGAGCCTCCACCTGCCTCTCAGAGCACTAGCCAGGACTGTTCACA
GGCTTTAAAACAGAGCCCATAAgcagcaACCGGTccagtgtgctggaattaattcgctgtctgcgagg

Secondly, the ManII fragment of the resulting construct was replaced with a linker region via restriction digest using EcoNI and EcoRI and NEBuilder HiFi assembly reaction to allow for direct tagging of TANGO1S with mSc.

Gene block from IDT used for the generation of TANGO1S-mSc (step2, linker):

ATGGCTCCAAAAGGGCCCCCTCCTTTCCCAGGAGTCCCTCTCATGAGCACCCCCATGGGAGGCCCTG
TACCACCACCCATTCGATATGGACCACCACCTCAGCTCTGCGGACCTTTTGGGCCTCGGCCACTTCCTCC
ACCCTTTGGCCCTGGTATGCGTCCACCACTAGGCTTAAGAGAATTTGCACCAGGCGTTCCACCAGGAAGA
CGGGACCTGCCTCTCCACCCTCGGGGATTTTTACCTGGACACGCACCATTTAGACCTTTAGGTTCACTTG
GCCCAAGAGAGTACTTTATTCCTGGTACCCGATTACCACCCCCAACCCATGGTCCCCAGGAATACCCACC
ACCACCTGCTGTAAGAGACTTACTGCCGTCAGGCTCTAGAGATGAGCCTCCACCTGCCTCTCAGAGCACT
AGCCAGGACTGTTCACAGGCTTTAAAACAGAGCCCAgccgcagcagcgaattccATGGTGAGCAAGGGCG
AGGC

Plasmids were amplified in DH5α E. coli (New England Biolabs) and subsequent extraction of plasmid DNA was done using a MidiPrep kit (Thermo Fisher Scientific) and 50 mL cell suspension in Lysogeny broth (LB; Thermo Fisher Scientific) with 50 µg·mL-1 ampicillin (amp). For transformation 1 ng of plasmid DNA (or 2 µL ligation reaction) was added to 50 µl thawed chemically competent 5-alpha competent E. coli (New England Biolabs) on ice, mixed gently by flicking the tube and incubated on ice for 30 min. To seal membrane openings heat shock at

42 °C was performed for 30 sec and cells were incubated on ice for 5 min. Subsequently, 450 µL super optimal broth (SOC) outgrowth medium for cell recovery was added to the transformed cells and incubated at 37 °C and 220 rounds per min (rpm) for 30 min, prior to plating on LB plates containing necessary selective antibiotics and incubated overnight at 37 °C.

For plasmid amplification, resulting colonies were picked and grown in 50 mL LB with according antibiotics in suspension at 37 °C and 220 rpm overnight. Extracted plasmid DNA via PureLink kit (Thermo Fisher Scientific) performed according to the manufacturer's instructions, with elution in 100 µL sterile filtered MilliQ H2O, was used for subsequent transfection of human cells.

For screening of bacterial colonies for presence of the correct plasmid containing the insert of interest, colonies were picked and grown in 5 mL LB with according antibiotics in suspension as mentioned above, followed by plasmid extraction via a MiniPrep kit (Qiagen) according to the manufacturer's instructions (with an elution in 30 µL sterile filtered MilliQ H2O) and restriction digest with suitable restriction enzymes (New England Biolabs) of about 250 ng plasmid DNA for 3 h, using the corresponding protocol by New England Biolabs. DNA fragments were separated by size using gel electrophoresis of 1 – 1.5% agarose gels containing ethidium bromide running at 70 – 90 V for 40 – 50 min in Tris-acetate-EDTA (Ethylenediamine tetra acetic acid; TAE) buffer. Samples were subsequently compared on a transilluminator using UV light and positive colonies identified. Sequences were confirmed via MWG Eurofins tube sequencing services.

Cell transfection and RUSH experiments

Cells were seeded 1 – 2 days before transfection to ensure a confluence of about 60 – 80 %. A Lipofectamine2000 transfection solution was prepared according to the manufacturer's instructions containing a total of 0.8 (for rescue experiments with either TANGO1S-mSc or TANGO1L-HA) or 1 µg plasmid DNA per construct and 2.5 µL lipofectamine2000 (Thermo Fisher Scientific) in 200 µL OptiMEM (Thermo Fisher Scientific) per 35 mm well and was added drop wise onto the cells covered with 1 mL of fresh media. Transfected cells were incubated at culturing conditions for about 16 h prior to media change, followed by either fixation for immunofluorescence or time courses in presence or absence of culture medium supplemented with 40 µM biotin. For collagen trafficking experiments of GFP-COL1A1 cells, these were incubated first in presence of 50 µg·mL^{-1} ascorbate for 24 h and then in absence of ascorbate for 24 h prior to the trafficking experiment during which 500 µg·mL^{-1} ascorbate and 400 µM biotin was used to trigger transport from the ER to the Golgi.*Immunofluorescence*

For immunofluorescence cells were grown on 13 mm (thickness 1.5; VWR) autoclaved cover slips. Cells were washed with PBS and fixed with 4% paraformaldehyde for 15 min at RT following repeated rinsing with PBS. PFA-fixed cells were permeabilised with 0.1% (v/v) Triton-X100 for 10 min at RT and blocked with 3% bovine serum albumin (BSA) in PBS for 30-60 min. Immunolabelling with primary and secondary antibodies was performed at RT for 1 h in a humid environment and in the dark. Antibodies were diluted to the final working concentrations or dilutions as in blocking solution as follows: mouse monoclonal anti-β-COP (dilution 1:500, G61610), rabbit polyclonal anti- β'-COP (dilution 1:20, (Palmer et al., 2009)), 0.5 µg.ml^{-1} rabbit polyclonal anti-COL1A1 (NB600-408, Novus Biologicals), mouse monoclonal anti-ERGIC-53 (dilution 1:1000, clone G1/93, Alexis Biochemicals), rabbit polyclonal anti-giantin (dilution 1:2000, Poly19243, BioLegend), 0.25 µg.ml^{-1} mouse monoclonal anti-GM130 (610823, BD Bioscience), sheep polyclonal anti-GRASP65 (dilution 1:1500, a gift from Jon Lane, University of Bristol (Cheng et al., 2010)), rabbit polyclonal anti-HA-Tag (C29F4) (dilution 1:500, mAb #3724, Cell Signaling), 0.75 µg.ml^{-1} mouse monoclonal anti-Hsp47 (M16.10A1, ENZO), 0.005 µg.ml^{-1} mouse monoclonal anti-PDI (clone 2E6A11, 66422-1-Ig, Proteintech), rabbit polyclonal anti-Sec16A (dilution 1:500, KIAA0310, Bethyl Labs, Montgomery, TX), rabbit polyclonal anti-Sec24C (dilution 1:250 (Townley et al., 2008)), 0.25 µg.ml^{-1} mouse monoclonal anti-Sec31A (612350, BD Bioscience), 0.4 µg.ml^{-1} rabbit polyclonal anti-TANGO1L (HPA056816-100UL, Sigma Aldrich Prestige, 0.5 µg.ml^{-1} rabbit polyclonal anti-TFG (NBP2-24485, Novus Biologicals).

Samples were rinsed three times with PBS for 5 min after incubation with primary and secondary antibodies, respectively. As secondary antibodies 2.5 µg·mL^{-1} donkey anti-rabbit Alexa-Fluor-568-conjugated, donkey anti-mouse Alexa-Fluor-647-conjugated or donkey anti-sheep Alexa-Fluor-488-conjugated antibodies were used (Thermo Fisher Scientific).

Samples were washed with deionised water and mounted using ProLong Diamond Antifade (Thermo Fisher Scientific) with 4',6-diamidino-2-phenylindole (DAPI) for confocal imaging or without DAPI when transfected with

ManII-BFP. For imaging via widefield microscopy, MOWIOL 4-88 (Calbiochem, Merck-Millipore, UK) mounting media was used in combination with 1 µg·mL^{-1} DAPI in PBS (Thermo Fisher Scientific) for 3 min at RT prior to repeated washing and mounting.

Image acquisition

Images of cells transiently expressing GFP-COL and showing TANGO1L and GM130 were obtained through widefield microscopy using an Olympus IX-71 inverted microscope (Olympus, Southend, UK) as described previously (McCaughey et al., 2019).

All other immunofluorescence images were obtained with confocal microscopy using a Leica SP5II for or Leica SP8 system (TANGO1L-HA rescues only with SP8, Leica Microsystems, Milton Keynes, UK) as previously described (Saito et al., 2009). In brief, images were acquired at 400 Hz scan speed with bidirectional scanning set up, zoom factor three, three times frame averaging and, when necessary, line accumulation set to two. Fluorophores were excited with an argon laser (SP5II) or white light laser (SP8) at the required wavelengths (405, 488, 561 and 630 nm). Pixel size was chosen according to Nyquist sampling.

For immunofluorescence of extracellular COL1A1 cells were seeded near confluent on cover slips, grown for 3 days in total with the last 48 h in media with 50 µg.ml^{-1} ascorbate prior to fixation with PFA and followed by subsequent steps as described before, without permeabilization using Triton-X100. Images for extracellular collagen were acquired using a confocal SP5II system at zoom factor 1, with 3000 px, 400Hz speed, 3 times frame averaging in form of z-stacks containing three slices to capture the entire signal for collagen per field of view. Four fields of view per sample were chosen by viewing the DAPI channel only.

Immunoblots

For semiquantitative analysis of protein levels of the COPII machinery and MIA proteins in WT-RPE-1 and MIA3-KO clones, cells were seeded in 6-well plates and grown for 2-4 days until confluent. Cells were rinsed with ice cold PBS and lysed in 200 µL buffer containing 50 mM Tris-HCl, 150 mM NaCl, 1% (vol/vol) Triton X-100, and 1% (vol/vol) protease inhibitor cocktail (Calbiochem) at pH 7.4 on ice for 15 min and scraped using rubber policemen. Lysates were centrifuged at 13,500 rpm at 4°C for 10 min and the supernatant was denatured using LDS loading buffer and reducing agent (Thermo Fisher Scientific) at 95°C for 10 min and run under reducing conditions on a 3-8% Tris-Acetate precast gel for 135 min at 100 V in Tris-Acetate running buffer supplemented with antioxidant (Thermo Fisher Scientific). Transfer of protein bands onto a nitrocellulose membrane (GE Healthcare, Amersham, UK) was performed at 15 V overnight/ 300mA for 3-5 h or semi-dry at 90V for 1.5 h. The membrane was blocked using 5% (wt/vol) milk powder in tris buffered saline with tween20 (0.01% (vol/vol)) (TBST) for 30 min at RT and incubated with primary antibodies for 1.5 h at RT or overnight at 4°C. Primary antibody concentrations (where known) used for Wester-Blot (WB) analysis were as follows: 2.5 µg.ml-1 rabbit polyclonal anti-COL1A1 (NB600-408, Novus Biologicals), rabbit polyclonal anti-cTAGE5-CC1 (a gift from Kota Saito (Maeda et al., 2016)), mouse monoclonal anti-DIC74.1 (MAB1618, Merck), 0.33 µg.ml-1 mouse monoclonal anti-GAPDH (AM4300, Thermo Fisher Scientific) or 0.2 µg.ml-1 mouse monoclonal anti-GADPH clone 1E6D9 (60004-1-Ig, Proteintech), 0.75 µg.ml-1 mouse monoclonal anti-Hsp47 (M16.10A1, ENZO), rabbit polyclonal anti-Sec12 (a gift from the Balch Lab, (Weissman et al., 2001)), rabbit polyclonal anti-Sec24A (Satchwell et al., 2013), rabbit polyclonal anti-Sec24C (Townley et al., 2008), rabbit polyclonal anti-Sec24D (Palmer et al., 2005)), rabbit polyclonal anti-Sec31A (Townley et al., 2008), 2 µg.ml-1 rabbit polyclonal anti-TANGO1L (HPA056816-100UL, Sigma Aldrich Prestige), rabbit polyclonal anti-TANGO1-CC1 (a gift from Kota Saito, (Maeda et al., 2016)), 0.5 µg.ml-1 rabbit polyclonal anti-TFG (NBP2-24485, Novus Biologicals).

After repeated rinsing with TBST, the membrane was incubated for 1.5 h at RT with HRP-conjugated antibodies diluted in the blocking solution (1:5,000) against mouse (Jackson ImmunoResearch, AB_2340770) and rabbit (Jackson ImmunoResearch, AB_10015282), respectively. The wash step was repeated, and detection was performed using Promega enhanced chemiluminescence reaction reagents and autoradiography films (Hyperfilm MP, GE Healthcare) with 3 sec – 30 min exposure and subsequent development.

For analysis of type I collagen secretion cells were incubated in 1 mL serum-free culture medium supplemented with or without 50 µg·mL^{-1} ascorbate for 24 h and the media fractions collected prior to cell lysis as described above without scraping to obtain lysis fractions. The transfer onto the nitrocellulose membrane was performed overnight at 15 V.

Electron microscopy

WT and MIA3-KO cell lines were grown until confluent, prior to rinsing with serum-free culture medium and subsequent fixation in 2.5% glutaraldehyde in serum free medium for 30 mins at room temperature. Cells were osmicated using osmium ferrocyanide following standard procedures and removed from the culture surface with a rubber policeman. Cells were mixed and spun down into a BSA gel at room temperature. The encased cells were then dehydrated in an alcohol series and embedded in EPON, sections (70 nm) were stained with uranyl acetate and lead citrate and imaged in a Thermo Fisher Tecnai 12 BioTwin TEM operating at 120kV with images recorded on a Thermo Fisher CETA 4kx4k camera. At least 15 cells from each variant were imaged with every sectioned Golgi complex in each cell being imaged. To aid accurate quantification, section tilting was used to enable imaging of the Golgi membranes as close to perpendicular as possible in the electron beam.

Sample preparation and analysis of proteomes via mass spectrometry

Soluble secreted proteomes were obtained from a 6-well dish with a confluent layer of cells incubated in presence of 50 µg·mL−1 ascorbate for 24 h in 1 ml serum free media prior to sample collection. Media fractions were collected, and potential cell debris removed by centrifugation at 13500 rpm for 10 min at 4C. Samples were frozen prior to further processing for tandem-mass-tagging (TMT).

To obtain the proteome from the cell-derived matrix, cells were grown in 15 cm dishes for seven days in media supplemented with 50 µg·mL−1 ascorbate (media was refreshed every 72 hours). All clones were seeded to reach confluency on day 3. Upon sample collection cells were rinsed with ice-cold PBS and extracted in 8 mL 20 mM ammonium hydroxide and 0.5% triton-X-100 in PBS for 2 min on a shaker with subsequent repeated rinsing in ice-cold PBS and incubation with 10 µg·mL−1 DNAseI for 30 min at 37C. Samples were washed with deionised water twice and directly scraped using a rubber-policeman in 400 µL reducing agent containing LDS buffer (NuPAGE, Thermo Fisher Scientific) and boiled at 95C for 10 min prior to snap freezing in liquid nitrogen and storage at -80C until further processing for TMT.

TMT Labelling and High pH reversed-phase chromatography.

For the secretome analysis, each media sample was concentrated to approximately 100ul using a centrifugal filter unit with a 3kDa cut-off (Merck Millipore, Cork, Ireland), digested with trypsin (2.5µg trypsin; 37°C, overnight) and labelled with Tandem Mass Tag (TMT) ten plex reagents according to the manufacturer's protocol (Thermo Fisher Scientific, Loughborough, UK), and the labelled samples pooled.

For the cell-derived matrix analysis, each sample was loaded onto a 10% SDS-PAGE gel and electrophoresis performed until the dye front had moved approximately 1 cm into the separating gel. Each gel lane was then cut into a single slice and each slice subjected to in-gel tryptic digestion using a DigestPro automated digestion unit (Intavis Ltd.). The resulting peptides were quantified using a quantitative colorimetric peptide assay kit (Pierce/Thermo Scientific) and an equal amount of each labelled with Tandem Mass Tag (TMT) ten plex reagents according to the manufacturer's protocol (Thermo Fisher Scientific) and the labelled samples pooled.

For both secretome and cell-derived matrix analyses, the TMT-labelled pooled samples were desalted using a SepPak cartridge according to the manufacturer's instructions (Waters, Milford, Massachusetts, USA). Eluate from the SepPak cartridge was evaporated to dryness and resuspended in buffer A (20 mM ammonium hydroxide, pH 10) prior to fractionation by high pH reversed-phase chromatography using an Ultimate 3000 liquid chromatography system (Thermo Fisher Scientific). In brief, the sample was loaded onto an XBridge BEH C18 Column (130Å, 3.5 µm, 2.1 mm X 150 mm, Waters, UK) in buffer A and peptides eluted with an increasing gradient of buffer B (20 mM Ammonium Hydroxide in acetonitrile, pH 10) from 0-95% over 60 minutes. The resulting fractions were concatenated to generate a total of four fractions, which were evaporated to dryness and resuspended in 1% formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific).

Nano-LC Mass Spectrometry

High pH RP fractions were further fractionated using an Ultimate 3000 nano-LC system in line with an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). After washing with 0.5% (vol/vol) acetonitrile 0.1% (vol/vol) formic acid peptides were resolved on a 250 mm × 75 µm Acclaim PepMap C18 reverse phase

analytical column (Thermo Scientific) over a 150 min organic gradient, using 7 gradient segments (1-6% solvent B over 1min., 6-15% B over 58min., 15-32%B over 58min., 32-40%B over 5min., 40-90%B over 1min., held at 90%B for 6min and then reduced to 1%B over 1min.) with a flow rate of 300 nl min−1. Solvent A was 0.1% formic acid and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.0kV using a stainless-steel emitter with an internal diameter of 30 µm (Thermo Scientific) and a capillary temperature of 300°C.

All spectra were acquired using an Orbitrap Fusion Lumos mass spectrometer controlled by Xcalibur 3.0 software (Thermo Scientific) and operated in data-dependent acquisition mode using an SPS-MS3 workflow. FTMS1 spectra were collected at a resolution of 120 000, with an automatic gain control (AGC) target of 200 000 and a max injection time of 50ms. Precursors were filtered with an intensity threshold of 5000, according to charge state (to include charge states 2-7) and with monoisotopic peak determination set to Peptide. Previously interrogated precursors were excluded using a dynamic window (60s +/-10ppm). The MS2 precursors were isolated with a quadrupole isolation window of 0.7m/z. ITMS2 spectra were collected with an AGC target of 10 000, max injection time of 70ms and CID collision energy of 35%.

For FTMS3 analysis, the Orbitrap was operated at 50 000 resolution with an AGC target of 50 000 and a max injection time of 105ms. Precursors were fragmented by high energy collision dissociation (HCD) at a normalized collision energy of 60% to ensure maximal TMT reporter ion yield. Synchronous Precursor Selection (SPS) was enabled to include up to 10 MS2 fragment ions in the FTMS3 scan.

Proteomic Data Analysis

The raw data files were processed and quantified using Proteome Discoverer software v2.1 (Thermo Scientific) and searched against the UniProt Human database (downloaded August 2020: 167789 entries) using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10ppm, and MS/MS tolerance was set at 0.6Da. Search criteria included oxidation of methionine (+15.995Da), acetylation of the protein N-terminus (+42.011Da) and Methionine loss plus acetylation of the protein N-terminus (-89.03Da) as variable modifications and carbamidomethylation of cysteine (+57.021Da) and the addition of the TMT mass tag (+229.163Da) to peptide N-termini and lysine as fixed modifications. Searches were performed with full tryptic digestion and a maximum of 2 missed cleavages were allowed. The reverse database search option was enabled and all data was filtered to satisfy false discovery rate (FDR) of 5%.

Data analysis and statistics

Identification and analysis of objects in immunofluorescence images in Figure 4 was conducted in Volocity version 6.3 followed by statistical analysis using GraphPad Prism version 8. Data distribution was assessed for normality and then analysed using the Kruskal-Wallis test with Dunn's multiple comparison for non-parametric data or ordinary one-way ANOVA with Dunnett's multiple comparisons test for parametric data. Asterisks on plots indicate p-values <0.05. RUSH assays were scored by visual inspection of predominant localization of reporters.