ER-to-Golgi trafficking of procollagen in the absence of large carriers.

Janine McCaughey[1], Nicola L. Stevenson[1], Stephen Cross[2] and David J. Stephens[1,] *

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

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

* Correspondence: ORCID: 0000-0001-5297-3240, david.stephens@bristol.ac.uk

Other ORCID IDs:

JM: 0000-0001-7709-2597

NLS: 0000-0001-8967-7277

SC: 0000-0003-3565-0479

Number of characters: 33597

Running title: ER-to-Golgi traffic of procollagen.

Summary

Trafficking of procollagen is essential for normal cell function. Here, imaging of GFP-tagged type I procollagen reveals that it is transported from the endoplasmic reticulum to the Golgi, without the use of large carriers.

Abstract

Secretion and assembly of collagen is fundamental to the function of the extracellular matrix. Defects in the assembly of a collagen matrix lead to pathologies including fibrosis and osteogenesis imperfecta. Owing to the size of fibril-forming procollagen molecules it is assumed that they are transported from the endoplasmic reticulum to the Golgi in specialised large COPII-dependent carriers. Here, analysing endogenous procollagen and a new engineered GFP-tagged form, we show that transport to the Golgi occurs in the absence of large carriers. Large GFP-positive structures are observed occasionally but these are non-dynamic, are not COPII-positive, and label with markers of the ER. We propose a "short-loop" model of ER-to-Golgi traffic that, while consistent with models of ERGIC-dependent expansion of COPII carriers, does not invoke long-range trafficking of large vesicular structures. Our findings provide an important insight into the process of procollagen trafficking and reveal a short-loop pathway from the ER to the Golgi, without the use of large carriers.

Introduction

Collagen is the most abundant protein in the body. Fibrillar type I collagen plays a key role in bone, skin and tendon formation, providing tissues with the necessary structural support. Altered collagen secretion, processing, and assembly is linked to diseases including osteogenesis imperfecta (OI), fibrosis, chondrodysplasia, Ehlers-Danlos syndrome and many more (Forlino and Marini, 2016; Jobling et al., 2014). Type I collagen assembles from two type I α1 chains together with one type I α2 chain to form trimeric procollagen in the endoplasmic reticulum (ER) (Canty and Kadler, 2005; Goldberg et al., 1972) , with the α-helix from each chain forming a rigid 300 nm triple helix structure (Bachinger et al., 1982; Lightfoot et al., 1992). During procollagen biosynthesis proline hydroxylation stabilises the triple helical conformation (Blanck and Peterkofsky, 1975; Jimenez et al., 1973). This process requires the presence of ascorbic acid, which acts as a co-factor for prolyl-4-hydroxylase (Mussini et al., 1967). The collagen-specific chaperone *heat shock protein 47* (Hsp47) (Ito and Nagata, 2017; Satoh et al., 1996) is also required.

To be secreted efficiently, procollagen I must traffic from the ER to the Golgi via the ER-Golgi intermediate compartment (ERGIC) (Malhotra and Erlmann, 2015; Malhotra et al., 2015; Satoh et al., 1996). Conventional ER-to-Golgi transport is facilitated by coat protein complex type II (COPII) vesicles with a size of 60 - 90 nm in diameter. These vesicles are thus significantly smaller than the 300 nm length of procollagen. Nonetheless, COPII vesicles are essential for efficient collagen trafficking in cells (Stephens and Pepperkok, 2002; Townley et al., 2008; Townley et al., 2012) and in animal models, since perturbation of, or mutations in, key COPII components including Sec24D (Garbes et al., 2015; Moosa et al., 2015; Sarmah et al., 2010), Sec23A (Boyadjiev et al., 2006; Lang et al., 2006), and Sec13 (Schmidt et al., 2013; Townley et al., 2008; Townley et al., 2012) cause defects in collagen secretion. To accommodate all of these data the prevailing hypothesis for the mechanism of procollagen secretion proposes the formation of large COPII carriers (Gorur et al., 2017; Malhotra and Erlmann, 2015; McGourty et al., 2016; Nogueira et al., 2014; Saito et al., 2009; Saito and Katada, 2015; Santos et al., 2015; Venditti et al., 2012). Formation of these carriers is said to be facilitated by the ER transmembrane proteins TANGO1 (*transport and Golgi organisation 1*, also called Mia3*)* and cTAGE5 (a TANGO1-related protein) that form a dimer and localises to ER exit sites (ERES) in mammals (Saito et al., 2011). TANGO1 is considered to act as a tether to the ERGIC to expand the nascent procollagen-containing carrier during its formation.

Several publications have shown large structures reported to be ER-to-Golgi carriers of procollagen (Gorur et al., 2017; Jin et al., 2012; McGourty et al., 2016). These large structures are generally few in number and seen in systems where the ubiquitin ligase Cullin3 (CUL3) adaptor Kelch-like protein 12 (KLHL12) is overexpressed (Jin et al., 2012). CUL3 facilitates mono-ubiquitylation of Sec31, stalling the outer complex formation, leading to a delayed scission and enlargement of COPII vesicles (Jin et al., 2012). Subsequent work defined a mechanism for KLHL12-mediated ubiquitylation of Sec31A, with PEF1 and ALG2 shown to be

75 subunits of the CUL3-KLHL12 ubiquitin ligase (McGourty et al., 2016). Consistent with the previous finding, the large structures labelled for Sec31A, PEF1, and ALG-2 and, more significantly, with antibodies against procollagen (McGourty et al., 2016). Similar large, procollagen- and Sec31A-positive structures have been shown in KI6 cells that overexpress both type I procollagen and FLAG-KLHL12 (Gorur et al., 2017). Apparent encapsulation of procollagen into large Sec31A structures was shown by correlative light electron
80 microscopy (Gorur et al., 2017). Interestingly, smaller Sec31A puncta remained devoid of procollagen labelling. Short-range movement of these carriers, using live cell imaging of procollagen-GFP, did not appear to be Golgi-directed and only occurred over distances of a few microns. Further support for this model came from experiments showing COPII-dependent packaging of procollagen *in vitro* into large structures (Gorur et al., 2017; Yuan et al., 2017).

85 Pathways for transport of procollagen from the ER to the Golgi have to date relied extensively on a COL1A1-GFP construct with a C-terminal GFP-tag (Stephens and Pepperkok, 2002). Expression of this construct reveals the presence of small ascorbate-dependent, highly mobile long-range transport carriers containing procollagen-GFP. While generally effective there are caveats to its use. Many cells do not traffic this fusion protein effectively. This compromised ability to assemble and traffic is most likely because the folding of the
90 procollagen trimer proceeds from the C-terminus (Bourhis et al., 2012; Kadler et al., 1990).

Our aim was to test the prevailing hypothesis that large vesicular carriers mediate procollagen secretion. To achieve this, we generated a newly designed procollagen construct incorporating some key features to better control and define secretory cargo export from the ER. Surprisingly, live cell imaging using this construct revealed that procollagen traffics from the ER to the Golgi without using large vesicular carriers.
95 Instead the Golgi gradually fills with procollagen upon ascorbate addition, while the ER empties, suggesting the existence of a short trafficking loop connecting the ER and Golgi. Combined with work on endogenous procollagen and a variety of cell types we propose a new mode of procollagen trafficking via a short-loop pathway that facilitates local transfer of procollagen from the juxtanuclear ER to the Golgi.

Results

An ascorbate- and biotin-controllable procollagen reporter: mGFP-SBP-COL1A1

To visualise ER-to-Golgi trafficking of procollagen for study, we engineered monomeric GFP (mGFP) and a streptavidin-binding peptide (SBP)-tag into procollagen 1α1. The tag was introduced upstream of the naturally occurring N-terminal proteolytic cleavage site (located after the N-terminal pro-peptide) to generate mGFP-SBP-COL1A1 (abbreviated here to GFP-COL1A1, Fig. 1A). The inclusion of the SBP-tag allows synchronisation of trafficking using the Retention Using Selective Hooks (RUSH) system to control cargo exit from the ER using biotin (Boncompain et al., 2012). When expressed transiently in IMR-90 human lung fibroblasts (Fig. 1Bi), and in human telomerase immortalised retinal pigment epithelial cells (hTERT-RPE-1, Fig. 1Bii), GFP-COL1A1 colocalises with Hsp47 in the ER, but not with the *cis*-Golgi marker GM130 at steady state (Fig. 1Ci). Following the addition of ascorbate, GFP-COL1A1 became enriched in the Golgi area (Fig. 1Cii - iii).

As expected, transient transfection of GFP-COL1A1-positive cells led to highly variable expression levels. To reduce the variability a RPE-1 cell line was created that stably expresses GFP-COL1A1 (GFP-COL1A1-RPE). The population was sorted into four equally distributed populations with different expression levels according to signal intensity of GFP and the 25% of cells expressing the lowest level was chosen for future experiments. For further analysis of transport to the Golgi GFP-COL1A1-RPE cells were co-transfected with a bicistronic construct encoding an ER-hook (KDEL-tagged streptavidin to retain the SBP fusion protein in the ER), and a separate *trans*-Golgi marker that expressed the minimal Golgi targeting region of sialyltransferase fused to mCherry (ST-Cherry). The co-expression of the ER-hook here enabled the retention of SBP-tagged GFP-COL1A1 in the ER lumen until addition of biotin to trigger the dissociation of SBP from the streptavidin-bound KDEL “hook”.

Live cell imaging of synchronised procollagen transport using confocal microscopy was then used to monitor the dynamics of ER-to-Golgi transport. Prior to addition of ascorbate and biotin (asc/biotin) GFP-COL1A1 and the Golgi marker show no significant colocalization (Fig. 1D, t=0). After addition of asc/biotin GFP-COL1A1 structures of high signal intensity appeared near the Golgi, as well as in the cell periphery (Fig. 1D, t=15 and Video 1). The signal intensity of GFP-COL1A1 in the ER decreased gradually over time while the accumulation of GFP-COL1A1 in and around the *trans*-Golgi further increased (Video 1). After filling of the Golgi an overall decrease of GFP-COL1A1 signal intensity occurred over time. Line scans in Fig. 1D show clearly this accumulation of GFP-COL1A1 around the ST-Cherry labelled Golgi compartments (t=16 mins), before GFP-COL1A1 fills the *trans*-Golgi (here at 24 mins asc/biotin), and subsequently exits (t=40 mins and 51 mins in this example). These experiments are consistent with ascorbate-dependent transport of GFP-COL1A1 to the Golgi, followed by its secretion from the cell.

Surprisingly, in the above experiments we failed to observe the formation of any large, dynamic carriers prior to procollagen reaching the Golgi. Imaging at a higher frame rate (Fig. S1A and Video 2) also failed to reveal any large puncta that translocate towards the Golgi. In some experiments (more noticeably in transiently transfected cells) we did however detect small punctate structures that translocate to the Golgi along curvilinear tracks, consistent with previously described vesicular tubular structures (Stephens and Pepperkok, 2002). Fig. S1B and Video 2 show one such structure tracking towards to the Golgi over the nucleus (highlighted by arrowheads). Collectively these data suggest ER-to-Golgi transport in the absence of large carriers. We therefore sought to define this in more detail.

ER-to-Golgi trafficking of GFP-COL1A1 occurs without the use of large carriers

In approximately 50% of all live cells, we did observe between one and four large, round GFP-COL1A1 structures (calculated diameter >1 µm) that were similar in appearance to those previously reported. For example, Fig. 2A and 2B show GFP-COL1A1 distributed throughout the ER and in four larger (1 - 2 µm diameter), circular, GFP-positive structures with elevated signal intensity compared to the ER background (Fig. 2B, circles). However, these structures were relatively static and did not change in size or intensity despite transport of GFP-COL1A1 to the Golgi (Fig. 2A - D and Videos 3 - 4). They also did not appear to be the source or destination of the few more dynamic punctate structures observed (e.g. highlighted by arrows in Fig. 2A). In one example, a larger (~1.4 µm diameter) circular structure was observed moving over a distance of approximately 2 µm, while maintaining its size and the same slightly higher signal compared to the GFP-COL1A1 signal from the surrounding ER (Fig. 2C, circle and Video 4), but this structure did not interact with the Golgi. As seen in Figure 2C and 2D, again no notable vesicular transport towards the Golgi can be detected from the timepoint of accumulation around the Golgi to the end of the measurement after 26 mins asc/biotin in this cell (Video 4). Extensive analysis of GFP-COL1A1-RPE also failed to reveal any significant GFP-positive large carriers contributing to the accumulation of GFP-COL1A1 in the Golgi. The earliest stages of accumulation of GFP-COL1A1 proximal to the Golgi suggest an accumulation of GFP-COL1A1 at the ER-Golgi interface (e.g. Fig. D and Video 4 t=9’30” – t=11’00”), prior to entering the Golgi.

Large GFP-COL1A1 structures are positive for ER and Hsp47, not COPII

To further define the nature of the large immobile GFP-positive structures cells were fixed during live cell imaging experiments at time points when an accumulation of GFP-COL1A1 in or around the Golgi was observed (Videos 2 – 4 and 5i – iv). Fig. 3A shows labelling for the endogenous *cis/medial*-Golgi marker giantin and ST-Cherry in these cells; the separation between markers is consistent with the expected localisation of ST-Cherry as a *trans*-Golgi marker. We found that the large GFP-positive structures observed

were negative for giantin (Fig. 3Aii, circle), however they did colocalise with the collagen-specific chaperone Hsp47 (Fig. 3Bi, circle). Therefore, these structures are not bona fide Golgi elements. To determine whether these large Hsp47-positive structures were part of the ER or post-ER structures, we used an ER membrane marker, Cyt-ERM-mScarlet (ERM). This construct encodes the red fluorescent protein mScarlet-i (Bindels et al., 2017) fused to the ER-targeting sequence of cytochrome P450 (Costantini et al., 2012). Intriguingly, the large GFP-COL1A1 structures colocalised with both Hsp47 and ERM (Fig. 3Ci - ii, circles).

To investigate whether any of these structures were COPII carriers, cells were co-labelled for the COPII marker Sec31A (Fig. 3D). We analysed the size distribution of GFP-positive structures relative to their overlap with Sec31A (Fig. 3E). Large GFP-COL1A1 structures (>1 µm calculated diameter) were found in approximately one fifth of the cells analysed but these did not show any colocalization with Sec31A (Fig. 3Di, circle and Fig.3E); nor did 95% of small GFP-positive structures (<350 nm calculated diameter) (example shown in Fig. 3Di - ii, arrows). Any GFP-positive puncta that did show a higher percentage of overlap with Sec31A (example shown in Fig. 3Di – ii, arrowheads) were consistently smaller than those that were negative for Sec31A (Fig. 3E). Furthermore, the small GFP-positive puncta were negative for both ERM (Fig. 3Ci – ii, arrow and square) and Golgi markers (Fig. 3Ai – ii, arrows). Interestingly, while most small punctate GFP-positive structures were also negative for Hsp47 (Fig. 3Bi - ii, 3Ci, arrows), some were found to indeed colocalise with Hsp47, but not ERM (Fig. 3Cii, square).

In all cases, as shown in Fig. 2D and Fig. 3A, cells show a gradual accumulation of GFP-COL1A1 around the rim of the Golgi followed by subsequent filling as described above for Fig. 1, without the appearance of significant numbers of peripheral carriers of any size directed towards the Golgi. The only structures positive for both GFP-COL1A1 and COPII were small and did not stand out in size compared to other COPII puncta. Therefore, we sought to examine the transport and occurrence of large carriers of endogenous procollagen in fibroblasts secreting large quantities of type I procollagen.

Primary fibroblasts are devoid of large procollagen structures

Immunofluorescence of endogenous procollagen 1α1 in adult primary skin fibroblasts (NHDF-Ad) showed that it colocalises with the collagen specific chaperone Hsp47 in the ER (Fig. 4Ai and iv), consistent with our construct. We also observed partial localization to areas of the Golgi apparatus (Fig. 4Aii - iii, arrows), especially at the rim of the Golgi. This appearance is very similar to prior observations made using GFP-COL1A1 expressing cells (most evident in Fig.1 Ciii, 1D (t=40) and Fig.3Ai - ii, 3Bi – ii). However, in primary fibroblasts this observation is independent of ascorbate since colocalization with Golgi markers was observed before, as well as after, 30 mins of incubation with ascorbate. It is possible that sufficient ascorbate was present in the growth medium to facilitate some proline hydroxylation and procollagen export. Importantly, most procollagen-positive punctate structures were negative for the COPII marker Sec31A

(Fig. 4Bi - iv, arrows). Similarly, analysis of non-transformed, telomerase immortalised, human fibroblasts (BJ-5ta) showed procollagen accumulated in the Golgi area both in the presence and absence of ascorbate (Fig. 4Ci - iv). Some small procollagen puncta were evident, but again these did not colocalize with Sec31A (Fig. 4Ci - iv, arrows). It was however notable that no large structures labelled for either endogenous procollagen 1α1 or Sec31A were detected in these fixed samples in either cell line. Therefore, our data support a direct route of transport from the juxtanuclear ER to the Golgi without the use of large carriers for both endogenous and overexpressed type I procollagen.

GFP-COL1A1 transport to the Golgi is not dependent on an intact microtubule network

The absence of long-range carriers translocating from the cell periphery to the Golgi during live cell imaging led us to question the role of peripheral ER exit sites in procollagen transport. Long-range ER-to-Golgi transport in mammalian systems occurs by translocation along microtubules. Microtubules are not essential for trafficking but optimise the efficiency of transfer between organelles. An intact microtubule network is, however, required to maintain the juxtanuclear organization of the Golgi ribbon. Upon microtubule disassembly, such as in the presence of the drug nocodazole, the Golgi scatters into functional mini-stacks dispersed throughout the cytosol adjacent to ERES. Nocodazole can thus be used to disrupt long-range trafficking from peripheral ERES without affecting local ER-Golgi transport at the sites of the mini-stacks.

To determine the requirement for microtubules, and a juxtanuclear Golgi, in the transport of procollagen, cells were treated with nocodazole (NZ) for 60 - 120 mins prior to imaging. Fig. 5 shows images of GFP-COL1A1-RPE cells 60 mins after the addition of NZ followed by addition of asc/biotin/NZ. As expected, the Golgi apparatus redistributed to scattered structures of either circular shape or as short ribbons. In Fig. 5A and 5B GFP-COL1A1 is distributed throughout the ER, while two larger GFP-COL1A1 structures (diameter of approximately 1.5 – 2 µm) could be seen close to the nucleus with a signal intensity higher than that of GFP-COL1A1 in the ER background. In this example, GFP-COL1A1 structures are first visible, distributed throughout the cell, 19 mins after addition of asc/biotin/NZ (Video 6). Nearly all structures appeared at the edge of the scattered *trans*-Golgi elements. Gradual accumulation of GFP-COL1A1 at the edge of the *trans*-Golgi further increased for about 4 mins. No vesicular-tubular or circular transport carriers could be identified. An even distribution throughout the *trans-Golgi* could be observed between 26 and 35 mins after addition of asc/biotin/NZ (depending on the individual Golgi element) (Fig. 5A – B, Video 6). The overall signal intensity for GFP-COL1A1 subsequently decreased over time. Once the *trans*-Golgi had filled, small GFP-COL1A1 puncta were seen emerging from Golgi elements and moving towards the cell perimeter. Subsequent emptying of the Golgi, consistent with onward trafficking of GFP-COL1A1, was seen as it was in the absence of nocodazole (Fig. 5C, t=44’14”).

Figure 1: Controllable ER-to-Golgi trafficking of GFP-COL1A1 in the absence of large carriers.

The GFP-COL1A1 fusion construct is shown in **A**. It consists of a synthetic pro-sequence of human COL1A1 followed by a streptavidin-binding peptide (SBP), monomeric green fluorescent protein (mGFP), the N-propeptide cleavage site (indicated by the arrowhead) within the non-helical region and the triple-helical domain with the C-terminal nonhelical region followed by the C-propeptide of human COL1A1. The signal sequence is indicated by *. Widefield microscopy of cells transiently expressing GFP-COL1A1 (A) is shown in **B** and **C**. 16 hours after transfection, GFP-COL1A1 (green) colocalises with Hsp47 (red) in the ER in IMR-90 cells (Bi) and RPE-1 (Bii). Large images show whole cells, while the small panels show the enlargements with the corresponding channels in greyscale and the merge image including nuclear DAPI staining (blue). **C:** Shows same image set-up as in B with GFP-COL1A1 (green) and *cis*-Golgi marker GM130 (red) labelling in transiently GFP-COL1A1 expressing IMR-90 cells. Timepoints indicate length of incubation in the presence of 50 µg.ml^{-1} ascorbate prior to fixation. A concentration of GFP-COL1A1 in the Golgi can be observed after both 30 (Cii) and 60 mins (Ciii) in presence with ascorbate. Number of cells imaged n ≥ 10. **D:** Conf ocal images from live cell imaging of RPE-1 stably expressing GFP-COL1A1 (green; GFP-COL1A1-RPE) co-transfected with the *trans*-Golgi marker ST-Cherry (red; utilising the RUSH system). Acquisition at 1 frame every 30 seconds. Image layout as in C, followed by line graphs showing the signal intensity (y-axis) in arbitrary units for GFP-COL1A1 (green) and ST-Cherry (red) for the corresponding line with a 5-pixel width drawn through the Golgi in the enlarged overlays. X-axis shows the distance in µm. Image stills are derived from Video 1. Timepoints indicate mins of incubation in presence of 500 µg.ml^{-1} ascorbate and 400 µM biotin (asc/biotin). At timepoint 0 (t = 0 mins) GFP-COL1A1 and ST-Cherry show distinct localisation. Over time an accumulation of GFP-COL1A1 around the Golgi marker occurs and increases until t = 18 mins, with subsequent filling of the *trans*-Golgi (t = 24 mins). Later time points show the decrease in intensity of GFP-COL1A1 overall and in the Golgi area (t = 29, 40 and 51 mins), implying functional transport to the Golgi, followed by secretion from the cell. For each set of live imaging experiment 3 cells from the same dish were imaged simultaneously. A total of n = 4 sets was acquired. Scale bars = 10 µm and 1 µm (in enlargements).

McCaughey et al Figure 1

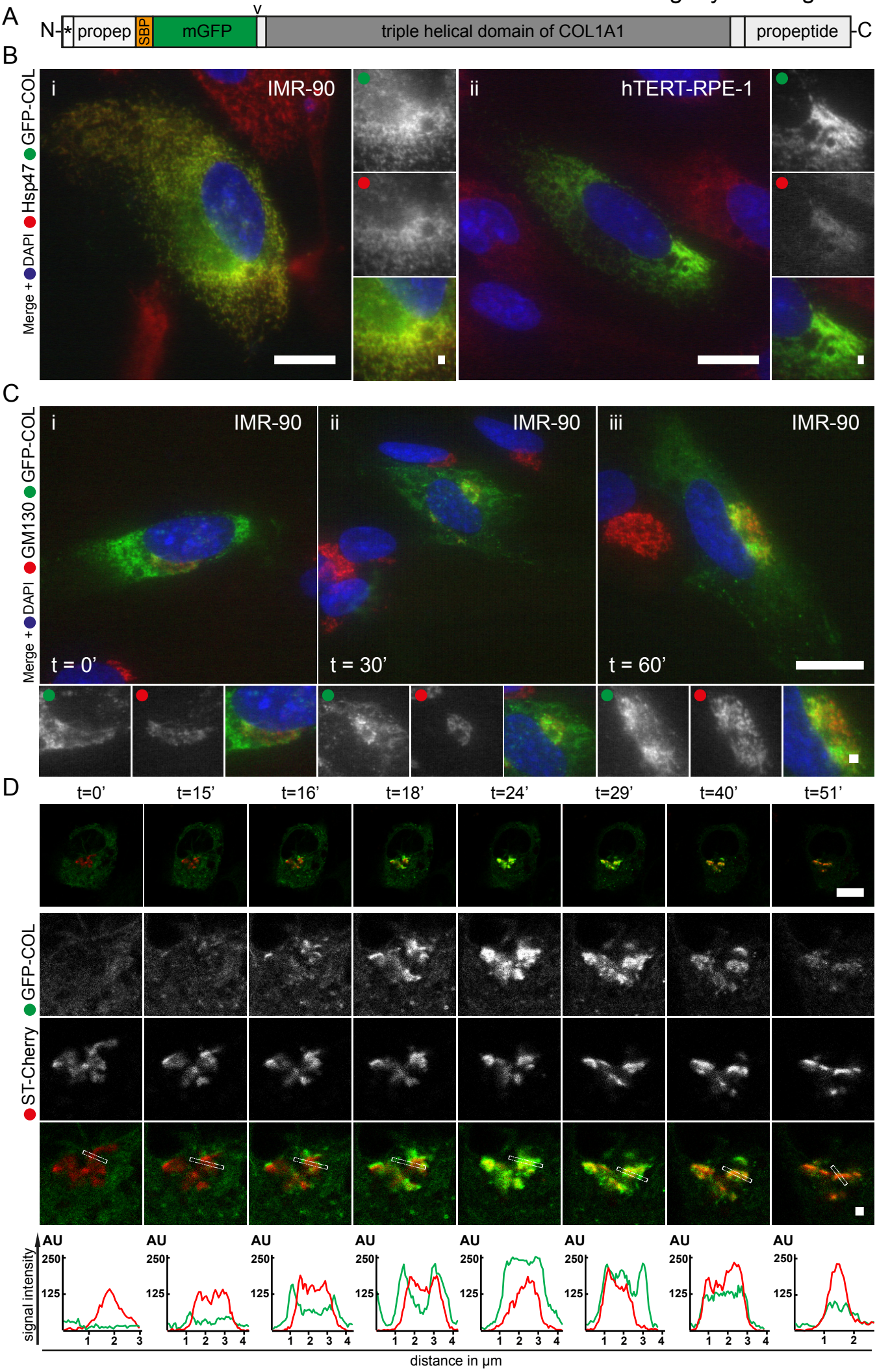


Figure 2: Transport of GFP-COL1A1 to the Golgi occurs without the use of large carriers.

Image stills from confocal live cell imaging of RPE-1 stably expressing GFP-COL1A1 (green; GFP-COL1A1-RPE) co-transfected with the *trans*-Golgi marker ST-Cherry (red). Timepoints indicate mins after addition of asc/biotin. **A:** Shows a whole cell imaged at 1 frame every 20 seconds (derived from Video 3). Probable post-Golgi structures are indicated by arrows. Corresponding enlargements of A are shown in **B** with channels for GFP-COL1A1 and ST-Cherry in greyscale, as well as the overlay image below. Large circular GFP-COL1A1-positive structures appear negative for the trans-Golgi (B, circles). Similar smaller structures are positive for GFP and the Golgi marker (asterisks) and follow the previously described phenotype of concentration of GFP-COL1A1 at the edge of the Golgi (t = 14 – 18 mins), followed by filling of the Golgi (t = 23 mins). **C:** Shows image stills of a whole cell taken at 1 frame every 30 seconds (derived from Video 4). Accumulation and filling of the Golgi with GFP-COL1A1 occurs within t = 10.5 mins. **D:** Corresponding enlargements of the Golgi area in C. Channels are displayed separately in greyscale with followed by a merge image. A total of n = 4 sets was acquired. For each set of live imaging experiment 3 cells from the same dish were imaged simultaneously. Scale bars = 10 µm and 1 µm (in enlargements).

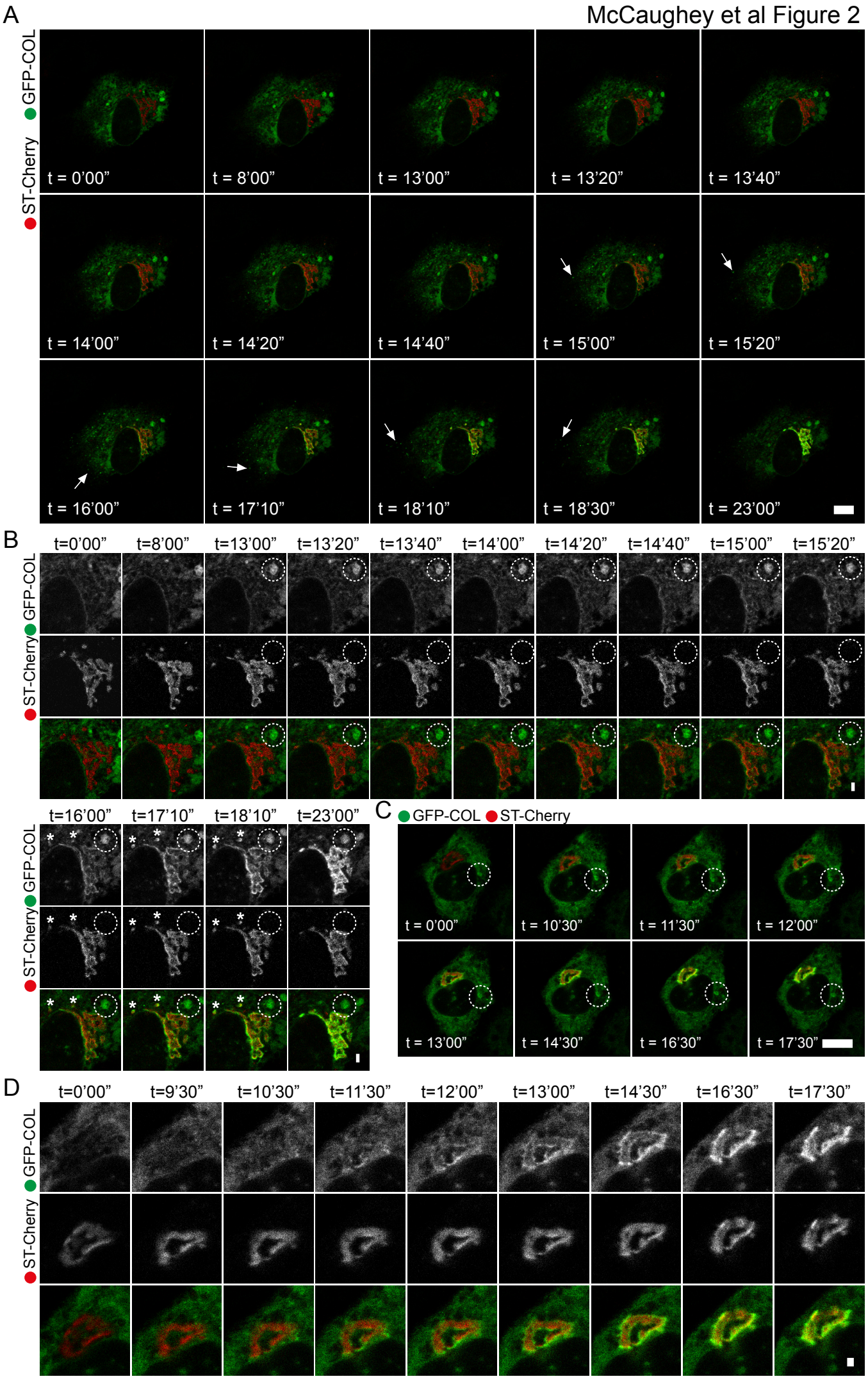

McCaughey et al Figure 2
A
GFP-COL
ST-Cherry
t = 0'00"
t = 8'00"
t = 13'00"
t = 13'20"
t = 13'40"
t = 14'00"
t = 14'20"
t = 14'40"
t = 15'00"
t = 15'20"
t = 16'00"
t = 17'10"
t = 18'10"
t = 18'30"
t = 23'00"
B
t=0'00"
t=8'00"
t=13'00"
t=13'20"
t=13'40"
t=14'00"
t=14'20"
t=14'40"
t=15'00"
t=15'20"
t=16'00"
t=17'10"
t=18'10"
t=23'00"
C
GFP-COL
ST-Cherry
t = 0'00"
t = 10'30"
t = 11'30"
t = 12'00"
t = 13'00"
t = 14'30"
t = 16'30"
t = 17'30"
D
t=0'00"
t=9'30"
t=10'30"
t=11'30"
t=12'00"
t=13'00"
t=14'30"
t=16'30"
t=17'30"

Figure 3: Large GFP-COL1A1 structures are positive for ER markers and Hsp47, but not COPII.

Confocal imaging of RPE-1 cells stably expressing GFP-COL1A1 (green; GFP-COL1A1-RPE) co-transfected with either ST-Cherry (red; A, B, D; 16 - 20 hours post-transfection) or the ER membrane marker ERM-mScarlet-i (red; C; 6 - 8 hours post-transfection). Cells were fixed after the given timepoints (mins after addition of asc/biotin (500 µg.ml^{-1} and 400 µM, respectively)) and labelled post-fixation with antibodies against further proteins of interest. Large panels show whole cells, while smaller panels show corresponding enlargements of areas of interest with the separate channels in greyscale, followed by the merge image including DAPI (acquired as separate channel and displayed in magenta in A, B and D or blue in C). Scale bars indicate 10 µm and 1 µm (enlargements). n ≥ 10.
A: Maximum projection images of z-stacks containing the Golgi apparatus. Duration of asc/biotin corresponds to duration of prior live imaging at approximately one image every 30 seconds until trafficking of GFP-COL1A1 to the Golgi was detectable by eye. Cells were labelled with the *cis/medial* Golgi marker Giantin (blue). Ai (after corresponding Video 5i) shows an accumulation of GFP-COL1A1 at the edge of the Golgi, without visible large GFP-positive structures. Small puncta can be observed in the cell periphery and close to the Golgi (Ai - ii, arrows). Large GFP-COL1A1-positive structures are negative for both the *trans*- and *cis/medial*-Golgi markers (Aii, circles). The corresponding video of live imaging prior to fixation of Aii is shown in Video 3 with image stills shown in Fig. 2A - B. **B:** Duration of asc/biotin corresponds to duration of prior live imaging at approximately one image every 30 seconds until trafficking of GFP-COL1A1 to the Golgi was detectable by eye (Video 5ii). GFP-COL1A1 within the ER at 57 mins asc/biotin colocalises with Hsp47 (blue) in the ER (Bi - ii). GFP-COL1A1 puncta in the cell periphery and close to the Golgi are negative for Hsp47 and the *trans*-Golgi marker ST-Cherry (arrows). Cells that show large (>1 µm diameter) GFP-COL1A1-positive structures appear negative for the *trans*-Golgi, but positive for Hsp47 (Bi, circles). **C:** Timepoint indicates incubation with 400 µg.ml^{-1} ascorbate prior to fixation. GFP-COL1A1 colocalises with the transiently expressed ERM-mScarlet-i (ERM; red) and Hsp47 (grayscale) in the ER, including large structures (Ci - ii, circles). Small GFP-positive puncta, likely post-Golgi carriers, show no colocalization with ERM or Hsp47 (Ci, arrows), while some small punctate GFP-COL1A1 structures are positive for Hsp47, but not ERM (Cii, square). **D:** Maximum intensity projection images of z-stacks of whole cells. Duration of asc/biotin corresponds to duration of prior live imaging at approximately one image every 30 seconds until trafficking of GFP-COL1A1 to the Golgi was detectable by eye (Video 5iii and iv, respectively). The red channel marked as "Golgi" shows the combined signal from ST-Cherry and antibody-labelling for giantin in the same channel to enable complete visualisation of the Golgi apparatus. Observed large GFP-positive structures do not colocalise with the COPII marker Sec31A (blue; Di, circles). Small GFP-COL1A1 puncta (<0.5 µm in diameter) close to the Golgi and in the cell periphery label for Sec31A (Di-ii, arrow heads). Most small, punctate, GFP-positive structures do not colocalise with Sec31A (Di - ii, arrows). **E:** Size distribution of GFP-positive objects relative to colocalization with the COPII marker Sec31A from confocal z-stacks with $\Delta z = 0.29$ µm and a sufficient number of slices to represent whole cells (as shown in D). The x-axis shows the calculated object diameter of GFP-positive objects in µm, while the y-axis shows the calculated percentage overlap of GFP-positive objects with Sec31A. Images for analysis were obtained after live cell microscopy as described in Fig. 1 and Fig. 2 and samples were fixed at the time when an accumulation of GFP-COL1A1 at the Golgi was visible by eye. In 11 of the 55 analysed cells no punctate or large GFP-COL1A1 objects were detected. These cells only show a GFP-COL1A1 accumulation around or in the Golgi. The remaining 44 cells show a total number of 1149 detect objects in the GFP-channel.

McCaughey et al Figure 3

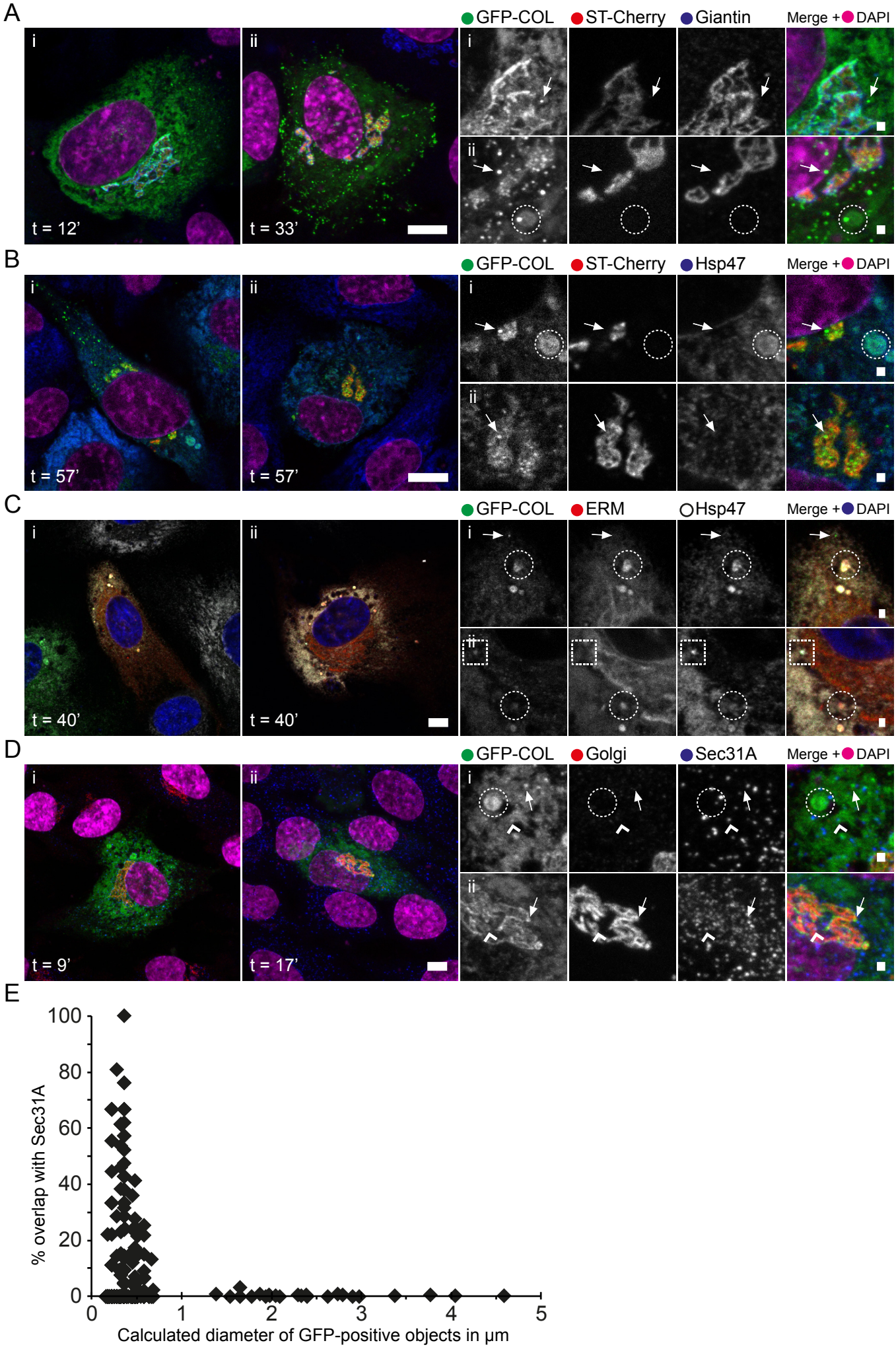


Figure 4: In primary fibroblasts, endogenous procollagen colocalises with Hsp47 in the ER.

Confocal images of cells after incubation in the presence of 500 µg.ml^{-1} ascorbate (A - B) or 50 µg.ml^{-1} ascorbate (C) for 0 (i - ii) and 30 mins (iii - iv), respectively. **A – B:** Maximum intensity projection images of z-stacks through primary adult skin fibroblasts NHDF-Ad. **C:** Shows single z-slice images of non-transformed foreskin fibroblasts (BJ-5ta). Cells were labelled using antibodies against COL1A1 (green) and either Hsp47 (A) or the COPII marker Sec31A (B – C) in red, as well as a *cis*-Golgi marker GRASP65 (blue). Panels show whole cells, as well as corresponding enlargements of the Golgi area on the right. Enlargements show the separate channels in greyscale as well as the overlay image including nuclear staining in magenta (DAPI, imaged as separate channel). Arrows highlight punctate COL1A1 structures with high signal intensity localising in close proximity to the Golgi. Scale bars = 10 µm for large panels and 1 µm for enlargements. NHDF-Ad show varying expression levels of Hsp47 and COL1A1 (Ai - iv). Localisation of COL1A1 to punctate structures in the Golgi area occurs at both 0 and 30 mins in the presence of ascorbate (Aii, iv and Bi - iv, arrows). BJ-5ta show lower levels of COL1A1 expression compared to NHDF-Ad, but accumulation of COL1A1 in the Golgi area can also be observed at both time points (Ci - iv, arrows). n ≥ 10 in each case.

McCaughey et al Figure 4

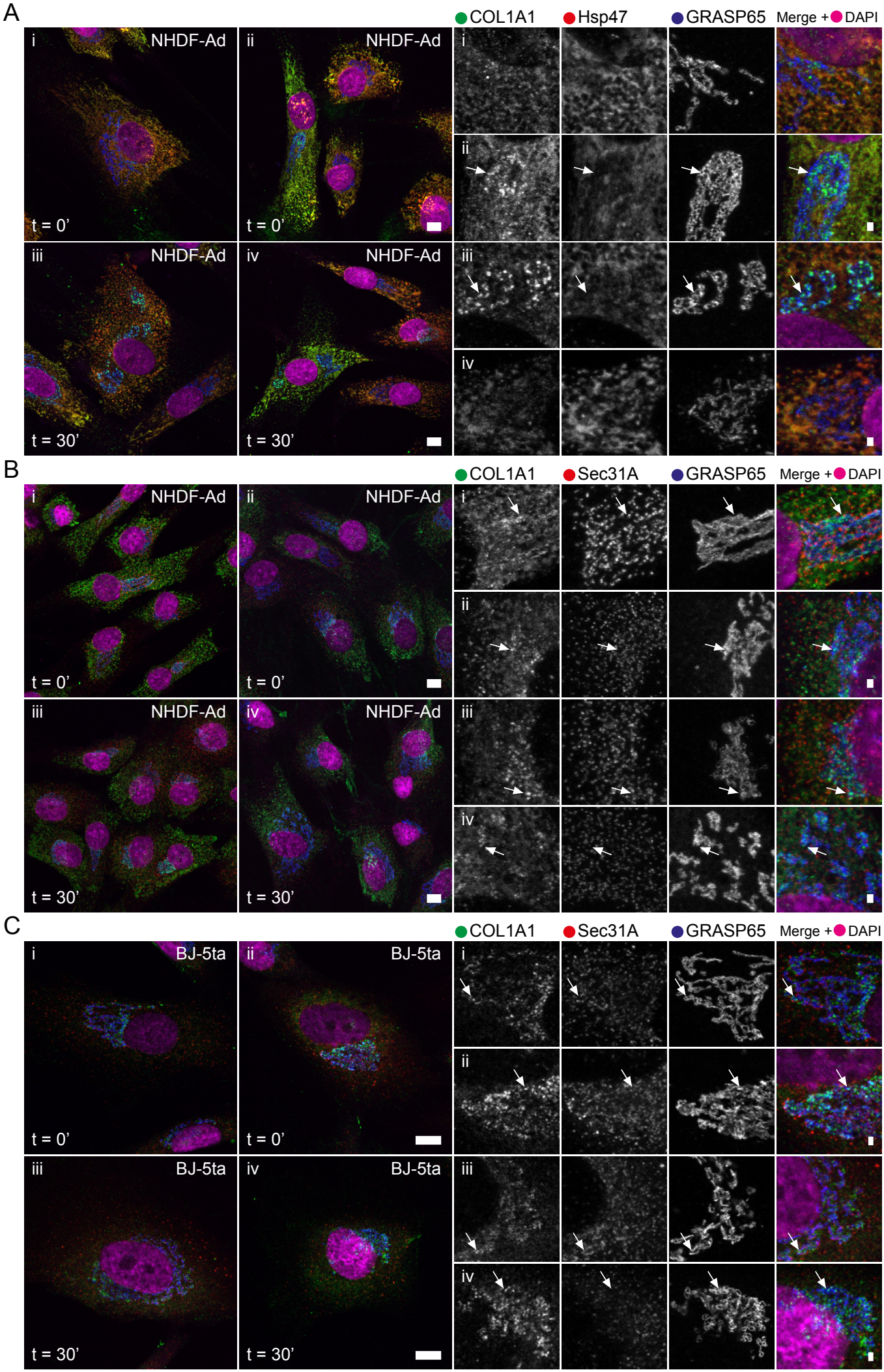


Figure 5: GFP-COL1A1 trafficking does not rely on an intact microtubule network

105 Still images from confocal live cell imaging of RPE-1 stably expressing GFP-COL1A1 (green; GFP-COL1A1-RPE) co-transfected with the *trans*-Golgi marker ST-Cherry (red) derived from Video 6. Acquisition at 1 frame every 26 seconds. Cells were incubated in presence of nocodazole for 60 mins prior to imaging. Time points indicate time in presence of asc/biotin (500 µg.ml^{-1} and 400 µM, respectively). **A**: Large panels show the entire cell, while enlargements in **B** highlight a zoomed in area of interest. Small panels show the green

110 channel and red channel in greyscale followed by the overlay. Scale bars indicate 10 µm, 1 µm (in enlargements). **C:** Example image of an enlarged overlay with a 5-pixel wide line drawn through the Golgi followed by corresponding line graphs of selected timepoints showing the signal intensity (y-axis) in arbitrary units for GFP-COL1A1 (green) and ST-Cherry (red) for the corresponding line. X-axis shows the distance in µm.

115

McCaughey et al Figure 5

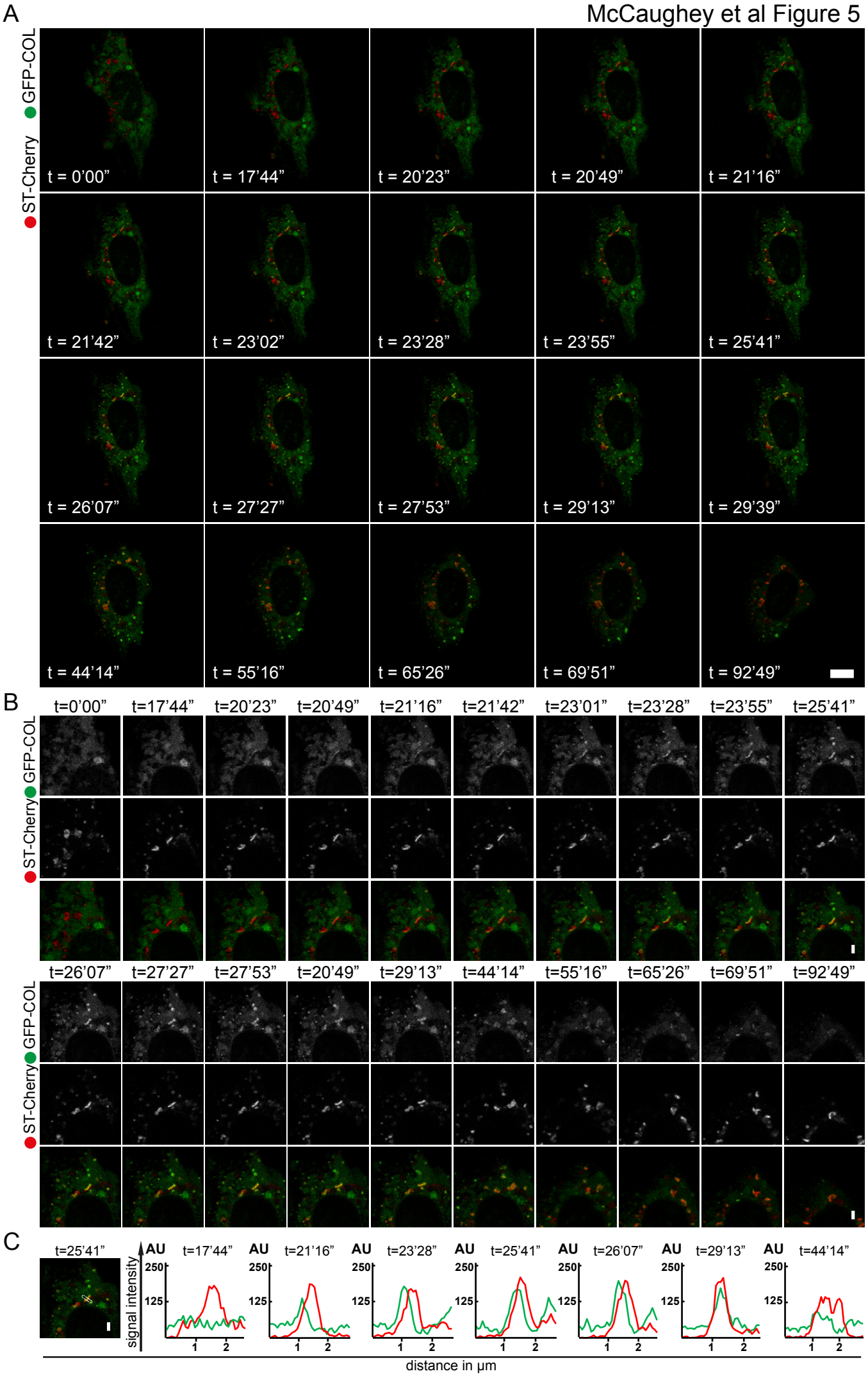
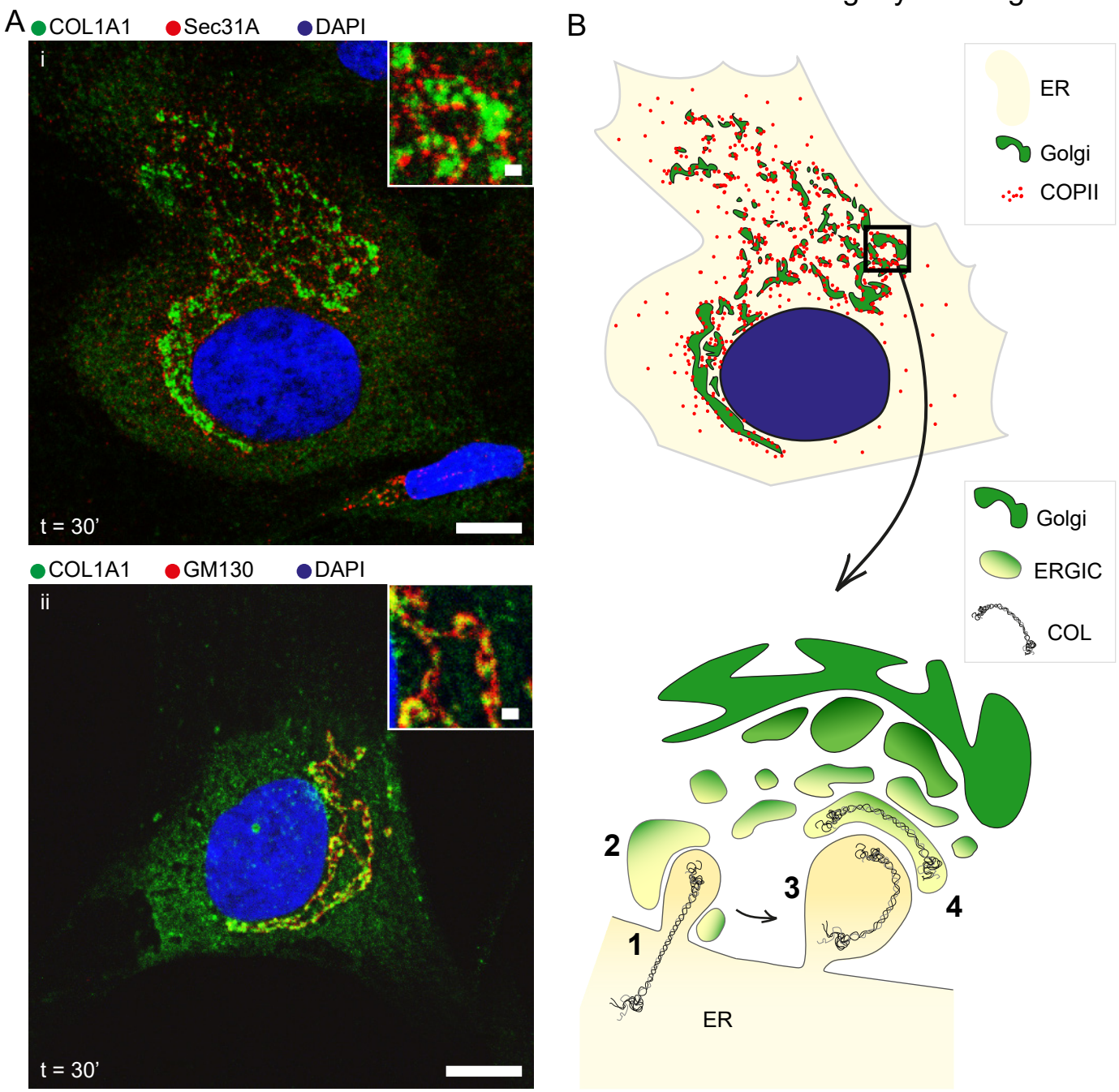


Figure 6: GFP-COL1A1 transport to the Golgi via a “short-loop” pathway

A: Immunofluorescence image of a primary skin fibroblast labelled for (i) endogenous COL1A1 (green) and Sec31A (red) and (ii) endogenous COL1A1 (green) and GM130 (red), both after 30 mins incubation in
120 presence with 50 µg.ml^{-1} ascorbate. B: Schematic of a cell with zoomed in region showing COL1A1 transport from ERES in close proximity to the Golgi. (1) COPII-dependent packaging of procollagen into nascent buds. (2) These buds are expanded by TANGO1-dependent fusion with the ERGIC. (3) Carrier expansion encapsulates procollagen. Questions remain as to the extent of expansion required and the flexibility of the trimer at this stage. (4) Scission of the carrier generates a compartment that almost immediately adopts the
125 identity of the ERGIC and progresses to become the first cisterna of the Golgi. In this model, the ERGIC acts as an intermediate to both expand the nascent ER-derived carrier and to ensure compartmentalization of ER and Golgi.

McCaughey et al Figure 6



Discussion

The currently accepted model for ER export of fibrillar procollagens, such as COL1A1, requires transport in large vesicles (Malhotra and Erlmann, 2015; Miller and Schekman, 2013; Saito and Katada, 2015; Stephens, 2012; Venditti et al., 2014). There are several reports of these structures in the literature from both cell-based and *in vitro* experiments (Gorur et al., 2017; Jin et al., 2012; McGourty et al., 2016; Yuan et al., 2017). Our data, using both stable expression of a newly engineered fluorescent procollagen reporter and endogenous labelling, do not support that model, at least for the long-range translocation of carriers from the peripheral ER to the Golgi. In contrast, our work shows that GFP-COL1A1 accumulates at sites juxtaposed to the Golgi without the formation of large carriers. Analysis of endogenous procollagen trafficking in fibroblasts confirmed this observation.

The construct we describe here is tagged at the N-terminus before the protease cleavage site. Very recently, similar constructs where the α2 chain of type I procollagen (COL1A2) was tagged at the N-terminus were described (Lu et al., 2018). In that work, the N-terminal propeptide was replaced with a fluorescent protein and the cleavage site removed. This was done intentionally and enabled the visualization of procollagen in the extracellular matrix. In our construct, we chose to retain the cleavage site such that extracellular procollagen itself does not become fluorescent. This enabled us to image intracellular trafficking more effectively. Furthermore, atypical retention of the N-propeptide due to impaired cleavage can lead to pathologies that overlap with those of osteogenesis imperfecta and Ehlers-Danlos Syndrome (Cabral et al., 2005; Malfait et al., 2013). Another recent development has been the engineering of a photoactivatable Dendra2 fluorescent protein into the endogenous locus of COL1A2 (Pickard et al., 2018). Here, Dendra2 was placed after the N-terminal cleavage site such that it is also retain in extracellular collagen fibrils. While this has advantages in terms of endogenous tagging and in analysis of fibril turnover, it does not enhance the ability to control ER export or to selectively analyse intracellular precursors versus extracellular pools. We also chose to tag COL1A1 rather than COL1A2 as it has the capacity to homotrimerize (Jimenez et al., 1977; Uitto, 1979), albeit as a minor fraction of total type I collagen, potentially lessening the impact of overexpression.

Our work supports the possibility of direct connections between the ER and Golgi to facilitate transfer of folded procollagen (Kurokawa et al., 2014; Malhotra and Erlmann, 2015). This idea has parallels with the cisternal maturation model of procollagen Golgi transit, where procollagen does not make use of vesicles while trafficking between Golgi stacks, but rather stays within the cisternae (Bonfanti et al., 1998). Direct connections are difficult to envisage given the well-characterized differences in composition between these organelles. A model that could reconcile this is the local formation of budding structures at ERES in close proximity to Golgi membranes (Fig. 6), as have been generated in *in vitro* budding assays (Gorur et al., 2017; Yuan et al., 2017). Direct connection of the ER to the ERGIC would both prevent compartment mixing of the

ER and Golgi and be consistent with data on the role of TANGO1 (Ma and Goldberg, 2016; Nogueira et al., 2014; Raote et al., 2018). One could view such a model as a maturation process where these nascent COPII-coated carriers form the ERGIC itself acquire compartment-specific markers and identity by direct fusion (Fig.6B). This is also consistent with the nocodazole experiments in which the Golgi is distributed adjacent to peripheral ERES.

This model raises questions about how large procollagen molecules could be encapsulated in COPII vesicles. Recent new data have challenged previous assumptions concerning the rigidity of triple helical collagen (Rezaei et al., 2018), showing instead that it can best be described as a semi-flexible polymer. Atomic force microscopy showed that the global curvature of collagen varies greatly with the salt concentration and pH. Importantly, this was seen with many collagen isotypes. It is therefore not inconceivable that the conditions at the point of ER exit are consistent with flexible procollagen that could be packaged into vesicular carriers closer to those classically described for COPII.

In many cases, small GFP-COL1A1 puncta were seen in our experiments that colocalise with Sec31A. These show a distinct size distribution from large, static, circular structures negative for the COPII marker. Therefore, our data are entirely consistent with COPII-dependent trafficking of procollagen from the ER via conventionally described ER exit sites and we do not dispute an absolute requirement for COPII in this process. Indeed, there is overwhelming support for this from *in vitro* (Gorur et al., 2017; Yuan et al., 2017), cell-based (Stephens and Pepperkok, 2002; Townley et al., 2008), and whole animal experiments (Garbes et al., 2015; Lang et al., 2006; Townley et al., 2008).

Our work also shows that trafficking and secretion of GFP-COL1A1 does not require an intact microtubule network. Transport of GFP-COL1A1 to Golgi stacks in our experiments appears very similar to that of GFP-tagged tumor necrosis factor (TNF-SBP-EGFP) (Fourriere et al., 2016), supporting the idea that this pathway might not be restricted to procollagen or other large cargo. We also think it unlikely that this short-loop pathway is itself mediated by the formation of large carriers, because large structures of the type described by others (Gorur et al., 2017; Jin et al., 2012; McGourty et al., 2016; Raote et al., 2017) are not evident in our experiments analysing either fixed or live cells. Notably, in those previously published experiments, COPII labelling appears very different to that one might expect from many other studies with few structures evident (Jin et al., 2012). Live imaging of both Sec31A-YFP and procollagen-CFP in KI6 cells that overexpress both procollagen and KLHL12, shows limited movement of an apparent Sec31A-positive and procollagen-positive structure over ~3 minutes (Gorur et al., 2017). This is not entirely consistent with long-range, vectorial ER-to-Golgi transport described previously (Stephens and Pepperkok, 2002) and the highly dynamic nature of the ER itself must be considered. Interestingly, some large FLAG-KLHL12 and Sec31A positive structures in KI6 cells colocalize with Hsp47 (Gorur et al., 2017). These structures would be consistent with our larger GFP-COL1A1 structures that are also positive for the collagen chaperone Hsp47.

In our experiments, some larger procollagen-containing structures are visible, but our data shows that these are most likely domains within the ER as they co-label for an ER membrane marker and Hsp47. Being contiguous with the ER, one might expect that these would label with markers of ER exit sites which are normally highly abundant in cells. Notably, we do not observe substantial labelling of these structures with Sec31A. The trafficking of Hsp47 itself is somewhat unclear. While it can bind to both monomeric and trimeric procollagen (Satoh et al., 1996) and has a KDEL retrieval sequence, both consistent with export from the ER, its localization at steady-state is almost entirely restricted to the ER (Duran et al., 2015; Kano et al., 2005; Salo et al., 2016). Current models favour a role for Hsp47 in maintaining the procollagen trimers in a non-aggregated state within the ER (Tasab et al., 2000) and possibly during further transit through the secretory pathway. *In vitro* data show that Hsp47 preferentially binds to the trimeric form of procollagen (Ishikawa et al., 2016; Ito and Nagata, 2017; Koide et al., 2006; Ono et al., 2012; Tasab et al., 2002) and accompanying procollagen after release from the ER (Oecal et al., 2016; Satoh et al., 1996). Interestingly, while most GFP-COL1A1 puncta in the cell periphery, as well as those close to the Golgi, were negative for Hsp47, some of these small punctate structures were positive for Hsp47 while also being negative for ER membrane. Therefore, these structures could be bona fide ER-to-Golgi transport carriers.

Mechanistically, a complex machinery, including KLHL12, is clearly involved in COPII-dependent trafficking from the ER. Our data, however, suggest that its role is not to generate large carriers in the cell periphery to direct this traffic. Rather, our findings indicate procollagen trafficking utilises what we term a “short-loop” pathway from the ER to the Golgi. This pathway could of course be used by other cargo. Advances in both light and electron microscopy provide opportunities to address this further. However, we conclude from our work that large carriers do not mediate long-range transport of procollagen from the peripheral ER to the Golgi.

Materials and methods

Unless stated otherwise, all reagents were purchased from Sigma-Aldrich (Poole, UK)

DNA constructs

All restriction and modifying enzymes were purchased from New England Biolabs (Hitchin, UK). ER-membrane marker pCytERM_mScarlet-i_N1 was a gift from Dorus Gadella (Addgene plasmid #85068) (Bindels et al., 2017). To generate the Str-KDEL-IRES-sialyltransferase-mCherry (ST-Cherry; Addgene #110727) construct, Str-KDEL_ST-SBP-mCherry (Addgene #65265), a gift from Franck Perez (Boncompain et al., 2012), was used as a template; the ST-SBP was exchanged for a ST sequence without the SBP-tag via restriction digest with AscI and SbfI.

The GFP-COL1A1 construct was designed using the NEB assembly tool (www.neb.com) to introduce a SBP-mGFP-tag between the N-propeptide and the corresponding cleavage site upstream of the triple helical domain of human procollagen 1α1. This was realised by using a two-step NEB HiFi Assembly reaction. First, the inserts procollagen-SBP, a synthetic construct composed of the sequence encoding the signal peptide and N-propeptide of human procollagen 1α1 (synthesised by MWG in pEX-A2), humanised monomeric GFP (mGFP; equivalent to NCBI accession number KP202880.1) and the genetic sequence encoding the triple helical domain and the C-terminus of human COL1A1 (NCBI accession number NM_000088.3) were amplified via PCR (primers listed below). mGFP and COL1A1 were subsequently purified via gel extraction. Prior to assembly, the remaining template vector procollagen-SBP-pEX-A2 was digested using DpnI. The lentiviral vector backbone pLVXPuro (Clontech/Takara Bio Europe, Saint-Germain-en-Laye, France) was linearized via overnight digest with EcoRI and subsequent heat inactivation. For the reaction the compounds to be combined were added in equimolar proportions of 0.05 pmol, except for procollagen-SBP with 0.18 pmol and incubated with the assembly master mix for 3 h at 50 °C, according to the NEB HiFi Assembly protocol. The transformation of NEB 5α competent *E. coli* was performed as described in the protocol. Sequencing of clones after the first assembly reaction showed successful insertion of mGFP and COL1A1 into pLVXPuro and an introduction of a unique PshAI restriction site between pLVXPuro and mGFP, which was used for a second assembly reaction. For this, 0.04 pmol procollagen-SBP, amplified using the primers listed below, and 0.02 pmol of the first assembly product (linearized using PshAI) were used for the assembly reaction. Resulting colonies yielded the desired construct procollagen-SBP-mGFP-COL1A1 in pLVXPuro (GFP-COL1A1; Addgene #110726).

Primers (synthesized by Sigma-Aldrich) used for amplification of the inserts: Forward primer for mGFP AGAGCCCATGTGCTGCTGCTGCATGGTGAGCAAGGGCGAG, reverse primer for mGFP GGGGCAGCAGCAGCACTTGTACAGCTCGTCCATGC, forward primer for COL1A1 GCTGTACAAGTGCTGCTGCTGCCCCCAGCTGTCTTATGGC, reverse primer for COL1A1

CGCGGTACCGTCGACTGCAGTTACAGGAAGCAGACAGG and final primers for procollagen-SBP amplification ACTCAGATCTCGACACCGGTCGCCACCATGTTCAGCTTTG (forward) and CCATGGTGGCGACCGGTGTCTTCATGGGCTCTCTCTGG (reverse).

Cell culture, cell lines and transfection

360 All cells were cultured at 37 °C and 5 % CO_2 in a humid environment. IMR-90 (ATCC Cat# CCL-186) were grown in MEM (Sigma-Aldrich) supplemented with L-glutamine and 10% decomplemented FBS (Gibco). DMEM (Gibco) containing 2 % M199 (Gibco) and 10% decomplemented FBS was used for cell culture of BJ-5ta hTERT-immortalized fibroblasts (ATCC Cat# CRL-4001). NHDF-Ad-Der primary fibroblasts (Lonza Cat# CC-2511) were grown in FBM medium supplemented with FGM-2 (Lonza). For generating the cell line stably
365 expressing the GFP-COL1A1 construct, as well as a cell line expressing just GFP as a control, hTERT RPE-1 cell line (ATCC Cat# CRL-4000) were used. These cells were grown in DMEM-F12 Ham (Sigma-Aldrich) supplemented with 10 % decomplemented FBS. Cells were not validated after purchase but were routinely screened annually (and confirmed negative) for mycoplasma contamination.

To make the stable GFP-COL1A1 cell line, virus containing the GFP-COL1A1 construct was generated using
370 the Lenti-XTM Packaging Single Shots (VSVG) system from Clontech/Takara according to the manufacturer's instructions (cat# 631275). Growth media was removed from an 80 % confluent 6 cm dish of hTERT-RPE1 and 1 ml harvested virus supernatant supplemented with 8 $\mu g.ml^{-1}$ polybrene was added to cells. After 1 hour of incubation at 37°C and 5% CO_2, 5 ml growth media was added. Transfection media was then replaced with fresh growth media after 24 hours. To select for transfected cells, cells were passaged in
375 growth media supplemented with 5 $\mu g.ml^{-1}$ puromycin dihydrochloride (Santa Cruz Biotechnology, Heidelberg, Germany) 72 hours post transfection. hTERT-RPE1 cells stably expressing GFP (GFP-RPE) were generated previously (Asante et al., 2014). Stable cell lines were maintained in growth media containing 3 $\mu g.ml^{-1}$ puromycin.

Cells were transfected at 80% confluence according to the manufacturer's protocol (Invitrogen/Thermo
380 Fisher, Paisley, UK), but using 2.5 µL Lipofectamine 2000 transfection solution (Invitrogen), 1 µg plasmid DNA and 250 µL OPTIMEM per 35 mm well, added drop wise onto the cells with 1 mL of fresh media. Transfection with ST-Cherry or GFP-COL1A1 was performed about 16 – 20 hours prior to fixation or live imaging, while transfection with ERM was performed 6-8 hours prior to the procollagen trafficking experiments.

Immunofluorescence

385 For immunofluorescence, cells were grown on 13 mm cover slips (0.17 mm thickness #1.5 (Thermo Fisher)) and fixed for 15 min at RT with 4% paraformaldehyde (Thermo Fisher) at RT. For fixation post live imaging, cells were grown in live cell dishes (MatTek Corp, Ashland, MA) and fixed by adding 8% PFA to an equal

amount of FluoroBrite DMEM imaging medium (Life Technologies, A18967-01). Cells were permeabilised with 0.1% (v/v) Triton-X100 (Sigma-Aldrich) for 10 mins at RT and blocked with 3% BSA (Sigma-Aldrich) for 30 min. Immunolabelling with primary and secondary antibodies was performed at RT for 1 hour in a humid environment and in the dark. Antibodies as follow were diluted in blocking solution to the final working concentrations or dilutions: 0.5 µg.ml^{-1} rabbit polyclonal anti-COL1A1 (NB600-408, Novus Biologicals), 1:2000 for rabbit polyclonal anti-giantin (Poly19243, Biolegend, London, UK), 0.25 µg.ml^{-1} mouse monoclonal anti-GM130 (610823, BD Biosciences, Wokingham, UK), 1:1500 sheep polyclonal anti-GRASP65 (gift from Jon Lane), 0.75 µg.ml^{-1} mouse monoclonal anti-Hsp47 (M16.10A1, ENZO, Exeter, UK) and 0.25 µg.ml^{-1} mouse monoclonal anti-Sec31A (612350, BD Biosciences).

Samples were rinsed three times with PBS for 5 mins 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. For labelling of endogenous procollagen donkey anti-rabbit Alexa-Fluor-488-conjugate was used.

Samples were washed with deionised water and mounted using ProLong Diamond Antifade with DAPI (Invitrogen) for confocal imaging or MOWIOL 4-88 (Calbiochem, Merck-Millipore, UK) mounting media for widefield imaging. Samples fixed after live imaging were incubated with DAPI (Invitrogen) for 3 mins at RT prior to repeated washing and storage in PBS at 4 °C in the dark until further imaging.

Procollagen trafficking experiments and image acquisition

GFP-COL1A1-RPE1 were incubated in medium containing 50 µg.ml^{-1} L-ascorbic acid-2-phosphate (Sigma-Aldrich; termed ascorbate throughout) for 24 hours - 48 hours prior to live image acquisition or fixation, to "flush out" overexpressed GFP-COL1A1 before a controlled accumulation in the ER for 24 h. Finally, to synchronise procollagen trafficking in GFP-COL1A1-RPE, medium was supplemented to contain 500 µg.ml^{-1} ascorbate during live imaging or prior to fixation. This higher concentration was found to more effectively synchronise export. Experiments performed with cells co-transfected with an ER-hook (ST-Cherry, to utilise the RUSH system) were conducted 16 - 20 hours post-transfection and the medium contained 500 µg.ml^{-1} ascorbate, as well as 400 µM biotin (Sigma-Aldrich), to enhance synchronisation and trigger GFP-COL1A1 release in a time-dependent manner. For analysis of the dependence on the microtubule network, cells were treated as mentioned above, but incubated in presence of 5 µM NZ for 60 – 120 mins prior to the initiation of the trafficking experiment via addition of asc/biotin/NZ to the cells. Images of cells transiently expressing GFP-COL1A1 were obtained through widefield microscopy using an Olympus IX-71 inverted microscope (Olympus, Southend, UK) combined with Exfo (Chandlers Ford, UK) Excite xenon lamp illumination, single pass excitation and emission filters combined with a multipass dichroic (Semrock, Rochester, NY) and images captured on an Orca-ER CCD (Hamamatsu, Welwyn Garden City, UK). The system was controlled using

Volocity (v. 5.4.1 Perkin-Elmer, Seer Green, UK). Chromatic shifts in images were registration corrected using TetraSpek fluorescent beads (Invitrogen/Thermo Fisher).

All other images were obtained using confocal microscopy using Leica SP5II for fixed or Leica SP8 for live samples (Leica Microsystems, Milton Keynes, UK). Z stacks of fixed samples labelled with Sec31A were acquired with $\Delta z = 0.29$ µm. For live cell imaging FluoroBrite DMEM (Thermo Fisher) was used as imaging medium. Live cell imaging was performed using a Leica SP8 confocal laser scanning microscope with 63x HC OL APO CS2 1.42 numerical aperture (NA) glycerol lens and an environmental chamber at 37°C with CO_2 enrichment and Leica LAS X software. Fluorophores were excited using ≤ 2% energy of the 65 mW Argon laser at 488 nm for the green and a 20mW solid state yellow laser at 561nm for the red channel, respectively. Time courses were acquired using the sequential scanning mode between lines, imaging speed set to 700, two times zoom and detection of the green and the red channel using 'hybrid' GaAsP detectors and corresponding notch filters. Each frame was acquired with a three times line average. One to three cells per sample were chosen that showed low to moderate expression of both GFP-COL1A1 and ST-Cherry (estimated by eye) and imaged using multi-position acquisition with 'Adaptive Focus Control' active for each cycle and each position to correct axial drift between frames. Minimisation of time between frames was undertaken to allow the highest temporal resolution possible with the given positions per sample resulting in intervals of 15 – 30 seconds between time points. Acquisition with a temporal resolution of 1 frame per second was accomplished by using a single position imaged at two times line averaging, imaging speed set to 1000 Hz and a zoom factor of 6.

All images shown were not altered from the raw data, unless mentioned. Movies and movie stills were enhanced in brightness and contrast for both channels, by using ImageJ autocorrection (except Fig. 1D). All movies were registration corrected using ImageJ's in-built 3D-correction for drifting, applied using the Golgi channel and smoothed using ImageJ's in-built *smooth* processing function.

Data analysis

Localisation analysis of images of GFP-COL1A1-RPE co-transfected with ST-Cherry and antibody labelled against Sec31A (post-fixation after live-imaging until an accumulation of GFP-COL1A1 in the Golgi was visible) was performed automatically. Spatial overlap between Sec31A and GFP-COL1A1 was measured using a custom plugin for ImageJ/Fiji (Schindelin et al., 2012; Schneider et al., 2012). First, the punctate Sec31A structures were detected using the Fiji plugin TrackMate (Tinevez et al., 2017), fit with a 2D ellipsoidal Gaussian distribution and false identifications isolated and removed according to a low-pass ellipticity filter. GFP-COL1A1, visible in a separate imaging channel, exhibited a mixed distribution of punctate and broader objects, which were identified separately then combined. Images were processed prior to object identification to enhance the respective structures being detected. For punctate GFP-COL1A1 structures,

images were convolved with a 3D Gaussian kernel to remove noise then processed with a rolling ball filter (Sternberg, 1983) to subtract non-punctate structures. Punctate GFP-COL1A1 structures were subsequently identified using the same approach as for Sec31A, but with an additional low-pass sigma (spot width) filter. For broader GFP-COL1A1 structures the raw image was also convolved with a 3D Gaussian kernel and rolling-ball filter, albeit with a larger radius. The images were then processed with a 2D median filter and thresholded using the maximum entropy approach (Kapur et al., 1985). Objects were identified as contiguous regions in the binarised image and filtered using a high-pass size filter. At this point punctate and broad GFP-COL1A1 structures were combined, with instances of spatial overlap resolved in favour of broad objects, unless the number of punctate objects per broad object exceeded a user-defined threshold of five. GFP-COL1A1 also accumulates around the Golgi. To remove these structures from the analysis, Golgi elements are identified in a separate fluorescence channel using a similar approach to the broader GFP-COL1A1 structures, except using the isodata thresholding approach (Ridler and Calvard, 1978). Any GFP-COL1A1 structures within 0.5 µm of a Golgi are removed. Finally, GFP-COL1A1 structures are filtered based on their mean intensity in the GFP-COL1A1 channel. Pixel-based overlap of Sec31A and GFP-COL1A1 is calculated along with the area of each object projected into the XY plane.

The code for data analysis is included as a .zip file in supplemental material. The plugin and the source code are also publicly-accessible on a new GitHub repository (https://github.com/SJCross/ModularImageAnalysis). This is linked to a service called Zenodo which provides a permanent DOI reference to that specific version (https://doi.org/10.5281/zenodo.1252337). The files are as follows: "Modular_Image_Analysis-v0.3.2.jar" is the plugin, which now includes all the third party libraries that were previous stored in the /jars folder; "ModularImageAnalysis-v0.3.2.zip" is the source code for the plugin itself; "Analysis.mia" is the .mia workflow file used for the final analysis; "Installation and usage.txt" describes how to install the plugin and run the Analysis.mia analysis file; "LICENSE.txt" is the license for the plugin itself; "dependencies.html" is an HTML formatted page listing all the dependencies used by the plugin and their associated licenses.

The analysis was performed with n=4 independent data sets that contained a total of 55 cells.

Author contributions

Conceptualization: DJS. Methodology: DJS, NLS, SC, JM. Software: SC. Validation: JM, NLS. Formal analysis: JM, NLS, SC. Investigation: JM, NLS. Resources: DJS, SC. Data curation: JM, NLS, SC, DJS. Writing – original draft: JM, SC, DJS. Writing – Reviewing and Editing: JM, NLS, DJS. Visualization: JM, DJS. Supervision: DJS. Project Administration: DJS. Funding Acquisition: DJS.

Acknowledgments

This work was funded by the Medical Research Council UK (grant no. MR/P000177/1) and the University of Bristol postgraduate research scholarship. We thank the Wolfson Foundation and University of Bristol for funding of the Wolfson Bioimaging Facility. This work also benefited from additional equipment funded by the BBSRC (through BrisSynBio, a BBSRC/EPSRC-funded Synthetic Biology Research Centre, grant L01386X) and an ALERT13 capital grant (BB/L014181/1). We thank other members of the lab for helpful discussions throughout the project. We thank Jon Lane, Franck Perez, and Dorus Gadella for sharing reagents with us and Ash Evans for creating the ST-Cherry construct.

Competing financial interests

The authors declare no competing financial interests.

References

Asante, D., N.L. Stevenson, and D.J. Stephens. 2014. Subunit composition of the human cytoplasmic dynein-2 complex. *J. Cell Sci.* 127:4774-4787.

Bachinger, H.P., K.J. Doege, J.P. Petschek, L.I. Fessler, and J.H. Fessler. 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*. 257:14590-14592.

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.

Blanck, T.J., and B. Peterkofsky. 1975. The stimulation of collagen secretion by ascorbate as a result of increased proline hydroxylation in chick embryo fibroblasts. *Arch. Biochem. Biophys.* 171:259-267.

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

Bonfanti, L., A.A. Mironov, J.A. Martínez-Menárguez, O. Martella, A. Fusella, M. Baldassarre, R. Buccione, H.J. Geuze, and A. Luini. 1998. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell*. 95:993-1003.

Bourhis, J.M., N. Mariano, Y. Zhao, K. Harlos, J.Y. Exposito, E.Y. Jones, C. Moali, N. Aghajari, and D.J. Hulmes. 2012. Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol*. 19:1031-1036.

Boyadjiev, S.A., J.C. Fromme, J. Ben, S.S. Chong, C. Nauta, D.J. Hur, G. Zhang, S. Hamamoto, R. Schekman, M. Ravazzola, L. Orci, and W. Eyaid. 2006. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nat Genet*. 38:1192-1197.

Cabral, W.A., E. Makareeva, A. Colige, A.D. Letocha, J.M. Ty, H.N. Yeowell, G. Pals, S. Leikin, and J.C. Marini. 2005. Mutations near amino end of alpha1(I) collagen cause combined osteogenesis imperfecta/Ehlers-Danlos syndrome by interference with N-propeptide processing. *J Biol Chem*. 280:19259-19269.

Canty, E.G., and K.E. Kadler. 2005. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci*. 118:1341-1353.

Costantini, L.M., M. Fossati, M. Francolini, and E.L. Snapp. 2012. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic*. 13:643-649.

Duran, I., L. Nevarez, A. Sarukhanov, S. Wu, K. Lee, P. Krejci, M. Weis, D. Eyre, D. Krakow, and D.H. Cohn. 2015. HSP47 and FKBP65 cooperate in the synthesis of type I procollagen. *Hum Mol Genet*. 24:1918-1928.

Forlino, A., and J.C. Marini. 2016. Osteogenesis imperfecta. *Lancet*. 387:1657-1671.

Fourriere, L., S. Divoux, M. Roceri, F. Perez, and G. Boncompain. 2016. Microtubule-independent secretion requires functional maturation of Golgi elements. *J Cell Sci*. 129:3238-3250.

Garbes, L., K. Kim, A. Riess, H. Hoyer-Kuhn, F. Beleggia, A. Bevot, M.J. Kim, Y.H. Huh, H.S. Kweon, R. Savarirayan, D. Amor, P.M. Kakadia, T. Lindig, K.O. Kagan, J. Becker, S.A. Boyadjiev, B. Wollnik, O. Semler, S.K. Bohlander, J. Kim, and C. Netzer. 2015. Mutations in SEC24D, encoding a component of the COPII machinery, cause a syndromic form of osteogenesis imperfecta. *Am J Hum Genet*. 96:432-439.

Goldberg, B., E.H. Epstein, Jr., and C.J. Sherr. 1972. Precursors of collagen secreted by cultured human fibroblasts. *Proc Natl Acad Sci U S A*. 69:3655-3659.

Gorur, A., L. Yuan, S.J. Kenny, S. Baba, K. Xu, and R. Schekman. 2017. COPII-coated membranes function as transport carriers of intracellular procollagen I. *J. Cell Biol.* 216:1745-1759.

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.

Ito, S., and K. Nagata. 2017. Biology of Hsp47 (Serpin H1), a collagen-specific molecular chaperone. *Semin Cell Dev Biol*. 62:142-151.

Jimenez, S., M. Harsch, and J. Rosenbloom. 1973. Hydroxyproline stabilizes the triple helix of chick tendon collagen. *Biochem. Biophys. Res. Commun.* 52:106-114.
Jimenez, S.A., R.I. Bashey, M. Benditt, and R. Yankowski. 1977. Identification of collagen alpha1(I) trimer in embryonic chick tendons and calvaria. *Biochem. Biophys. Res. Commun.* 78:1354-1361.
Jin, L., K.B. Pahuja, K.E. Wickliffe, A. Gorur, C. Baumgartel, R. Schekman, and M. Rape. 2012. Ubiquitin-dependent regulation of COPII coat size and function. *Nature*. 482:495-500.
Jobling, R., R. D'Souza, N. Baker, I. Lara-Corrales, R. Mendoza-Londono, L. Dupuis, R. Savarirayan, L. Ala-Kokko, and P. Kannu. 2014. The collagenopathies: review of clinical phenotypes and molecular correlations. *Curr Rheumatol Rep*. 16:394.
Kadler, K.E., Y. Hojima, and D.J. Prockop. 1990. Collagen fibrils in vitro grow from pointed tips in the C- to N-terminal direction. *Biochem J*. 268:339-343.
Kano, F., H. Kondo, A. Yamamoto, A.R. Tanaka, N. Hosokawa, K. Nagata, and M. Murata. 2005. The maintenance of the endoplasmic reticulum network is regulated by p47, a cofactor of p97, through phosphorylation by cdc2 kinase. *Genes Cells*. 10:333-344.
Kapur, J.N., P.K. Sahoo, and A.K.C. Wong. 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*. 29:273-285.
Koide, T., S. Asada, Y. Takahara, Y. Nishikawa, K. Nagata, and K. Kitagawa. 2006. Specific recognition of the collagen triple helix by chaperone HSP47: minimal structural requirement and spatial molecular orientation. *J Biol Chem*. 281:3432-3438.
Kurokawa, K., M. Okamoto, and A. Nakano. 2014. Contact of cis-Golgi with ER exit sites executes cargo capture and delivery from the ER. *Nat Commun*. 5:3653.
Lang, M.R., L.A. Lapierre, M. Frotscher, J.R. Goldenring, and E.W. Knapik. 2006. Secretory COPII coat component Sec23a is essential for craniofacial chondrocyte maturation. *Nat Genet*. 38:1198-1203.
Lightfoot, S.J., D.F. Holmes, A. Brass, M.E. Grant, P.H. Byers, and K.E. Kadler. 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*. 267:25521-25528.
Lu, Y., S.A. Kamel-El Sayed, K. Wang, L.M. Tiede-Lewis, M.A. Grillo, P.A. Veno, V. Dusevich, C.L. Phillips, L.F. Bonewald, and S.L. Dallas. 2018. Live Imaging of Type I Collagen Assembly Dynamics in Osteoblasts Stably Expressing GFP and mCherry-Tagged Collagen Constructs. *J. Bone Miner. Res.*
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.
Malfait, F., S. Symoens, N. Goemans, Y. Gyftodimou, E. Holmberg, V. Lopez-Gonzalez, G. Mortier, S. Nampoothiri, M.B. Petersen, and A. De Paepe. 2013. Helical mutations in type I collagen that affect the processing of the amino-propeptide result in an Osteogenesis Imperfecta/Ehlers-Danlos Syndrome overlap syndrome. *Orphanet J Rare Dis*. 8:78.
Malhotra, V., and P. Erlmann. 2015. The Pathway of Collagen Secretion. *Annu Rev Cell Dev Biol*. 31:109-124.
Malhotra, V., P. Erlmann, and C. Nogueira. 2015. Procollagen export from the endoplasmic reticulum. *Biochem Soc Trans*. 43:104-107.
McGourty, C.A., D. Akopian, C. Walsh, A. Gorur, A. Werner, R. Schekman, D. Bautista, and M. Rape. 2016. Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor. *Cell*. 167:525-538 e514.
Miller, E.A., and R. Schekman. 2013. COPII - a flexible vesicle formation system. *Curr. Opin. Cell Biol.* 25:420-427.
Moosa, S., B.H. Chung, J.Y. Tung, J. Altmuller, H. Thiele, P. Nurnberg, C. Netzer, G. Nishimura, and B. Wollnik. 2015. Mutations in SEC24D cause autosomal recessive osteogenesis imperfecta. *Clin Genet*.
Mussini, E., J.J. Hutton, Jr., and S. Udenfriend. 1967. Collagen proline hydroxylase in wound healing, granuloma formation, scurvy, and growth. *Science*. 157:927-929.
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.

Oecal, S., E. Socher, M. Uthoff, C. Ernst, F. Zaucke, H. Sticht, U. Baumann, and J.M. Gebauer. 2016. The pH-dependent Client Release from the Collagen-specific Chaperone HSP47 Is Triggered by a Tandem Histidine Pair. *J Biol Chem*. 291:12612-12626.

Ono, T., T. Miyazaki, Y. Ishida, M. Uehata, and K. Nagata. 2012. Direct in vitro and in vivo evidence for interaction between Hsp47 protein and collagen triple helix. *J Biol Chem*. 287:6810-6818.

Pickard, A., A. Adamson, Y. Lu, J. Chang, R. Garva, N. Hodson, and K. Kadler. 2018. Collagen assembly and turnover imaged with a CRISPR-Cas9 engineered Dendra2 tag. *bioRxiv*:331496.

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. *bioRxiv*:279380.

Ridler, T.W., and S. Calvard. 1978. Picture Thresholding Using an Iterative Selection Method. *IEEE Transactions on Systems, Man, and Cybernetics*. 8:630-632.

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., and T. Katada. 2015. Mechanisms for exporting large-sized cargoes from the endoplasmic reticulum. *Cell. Mol. Life Sci.* 72:3709-3720.

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.

Salo, V.T., I. Belevich, S. Li, L. Karhinen, H. Vihinen, C. Vigouroux, J. Magre, C. Thiele, M. Holtta-Vuori, E. Jokitalo, and E. Ikonen. 2016. Seipin regulates ER-lipid droplet contacts and cargo delivery. *Embo J*. 35:2699-2716.

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.

Sarmah, S., A. Barrallo-Gimeno, D.B. Melville, J. Topczewski, L. Solnica-Krezel, and E.W. Knapik. 2010. Sec24D-dependent transport of extracellular matrix proteins is required for zebrafish skeletal morphogenesis. *PLoS One*. 5:e10367.

Satoh, M., K. Hirayoshi, S. Yokota, N. Hosokawa, and K. Nagata. 1996. Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen. *J. Cell Biol.* 133:469-483.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 9:676-682.

Schmidt, K., F. Cavodeassi, Y. Feng, and D.J. Stephens. 2013. Early stages of retinal development depend on Sec13 function. *Biol Open*. 2:256-266.

Schneider, C.A., W.S. Rasband, and K.W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of Image Analysis. *Nature methods*. 9:671-675.

Stephens, D.J. 2012. Cell biology: Collagen secretion explained. *Nature*. 482:474-475.

Stephens, D.J., and R. Pepperkok. 2002. Imaging of procollagen transport reveals COPI-dependent cargo sorting during ER-to-Golgi transport in mammalian cells. *J Cell Sci*. 115:1149-1160.

Sternberg, S.R. 1983. Biomedical Image Processing. *Computer*. 16:22-34.

Tasab, M., M.R. Batten, and N.J. Bulleid. 2000. Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. *Embo J*. 19:2204-2211.

Tasab, M., L. Jenkinson, and N.J. Bulleid. 2002. Sequence-specific recognition of collagen triple helices by the collagen-specific molecular chaperone HSP47. *J Biol Chem*. 277:35007-35012.

Tinevez, J.Y., N. Perry, J. Schindelin, G.M. Hoopes, G.D. Reynolds, E. Laplantine, S.Y. Bednarek, S.L. Shorte, and K.W. Eliceiri. 2017. TrackMate: An open and extensible platform for single-particle tracking. *Methods*. 115:80-90.

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.

Townley, A.K., K. Schmidt, L. Hodgson, and D.J. Stephens. 2012. Epithelial organization and cyst lumen expansion require efficient Sec13-Sec31-driven secretion. *J. Cell Sci.* 125:673-684.

Uitto, J. 1979. Collagen polymorphism: isolation and partial characterization of alpha 1(I)-trimer molecules in normal human skin. *Arch. Biochem. Biophys.* 192:371-379.

Venditti, R., T. Scanu, M. Santoro, G. Di Tullio, A. Spaar, R. Gaibisso, G.V. Beznoussenko, A.A. Mironov, A. Mironov, Jr., L. Zelante, M.R. Piemontese, A. Notarangelo, V. Malhotra, B.M. Vertel, C. Wilson, and M.A. De Matteis. 2012. Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science*. 337:1668-1672.

Venditti, R., C. Wilson, and M.A. De Matteis. 2014. Exiting the ER: what we know and what we don't. *Trends Cell Biol*. 24:9-18.

Yuan, L., S. Baba, K. Bajaj, and R. Schekman. 2017. Cell-free Generation of COPII-coated Procollagen I Carriers. *Bio Protoc*. 7:e2450.