Author:
Mccaughey, Janine

Title:
The mechanism of procollagen trafficking through the early secretory pathway

The mechanism of procollagen trafficking through the early secretory pathway

Janine McCaughey

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Life Sciences

School of Biochemistry

September 2020

word count: 34605

Abstract

Cells and tissues are highly dependent on the secretion of extracellular matrix proteins like collagen to fulfil their roles in multicellular organisms and enable complex processes such as organ and bone formation during early development. Due to its unique structure, extensive post-translational modifications, requirement for fast abundant protein secretion during wound-healing processes and size of about 300 nm in length, type I procollagen (the precursor molecule of type I collagen) imposes high demand on the secretory pathway. Defects in collagen transport and secretion can lead to osteogenesis imperfecta, Ehlers-Danlos Syndrome, and fibrosis. Transport of proteins from the endoplasmic reticulum to the Golgi is facilitated via the COPII vesicle coat complex. These are typically considered to be spherical 60 – 80 nm diameter structures that bud from the endoplasmic reticulum membrane, encapsulate cargo proteins, and subsequently merge with the Golgi apparatus or, the endoplasmic reticulum-Golgi intermediate compartment. Despite having identified many factors involved in COPII organisation and regulation a major question in the field remains – how are large secretory cargo proteins like procollagen transported in a COPII-dependent manner through the early secretory pathway? Hypotheses include the use of large COPII *megacarriers* and direct tunnel-like connections. Here, I have used a new engineered GFP-tagged procollagen reporter, combined with experiments in primary fibroblasts, to characterise procollagen transport. In addition, fibroblasts from patients were investigated to help classify the pathogenicity of novel/rare mutations in genes involved in collagen processing and transport. GFP-procollagen was observed to be transported to the Golgi in absence of large carriers. Super resolution microscopy revealed small procollagen puncta overlapping, but not surrounded by, COPII. These data are consistent with models where large cargo proteins, including fibrillar collagens, are transported in a progressive manner in close proximity to Golgi membranes. This does not require large carriers, nor discrete cargo encapsulation by COPII.

‘Curiosity killed the cat, but knowledge brought it back‘

- Deviated from the British proverb

Acknowledgements

I would like to thank Prof Dr David J Stephens for permitting me to perform my master thesis in his lab, which resulted in the first step for making this PhD project possible.

I owe sincere gratitude to Prof Dr David J Stephens, for all his guidance, support and advice during this project both professionally and personally. I could not have wished for a better supervision.

I am grateful for the collaboration with Dr Meena Balasubramanian who made it possible to study endogenous procollagen in patient-derived cells with collagen related mutations. I am particularly grateful for Nicola Stevenson and Laura Vuolo, who encouraged, supported and counselled me in the laboratory and provided some of the cells to work with. I especially want to thank Stephen Cross for writing the Fiji plugin code for the data analysis without which this thesis could not have comprised the amount of data it has within this time frame. Further thanks go to Caroline Shak, Anantha Sundararaman, Borhan Uddin, Holly Baum, Dylan Bergen and Ash Evans for their helpful advice. Thanks to all the other researchers (Sam) of the C50 lab for their collegiality and support. I thoroughly enjoyed my time here with you!

I am very grateful for the opportunity to perform my experiments with the instruments provided by the Wolfson BioImaging Facility and in particular I want to thank Alan Leard, Dominick Alibhai, Katy Jepson and Judith Mantell for their training and advice. Great thanks also go to the proteomics and FACs facility of the University of Bristol, namely, Kate Heesom and Andrew Herman. I am also grateful for having received the postgraduate research UoB scholarship, without which none of this would have been possible.

A great amount of gratitude also goes to my partner Kyle Quinney and his dog Bumi, who made my time during the write-up as comfortable and entertaining as possible. Finally, I thank my friends, the staff squash club and the Bristol capoeira group for keeping me fit and healthy during these times. Thanks also goes to my parents and family for their support and encouragement.

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .....Janine McCaughey..... DATE:....17/09/2020....

Table of contents

List of Figures

List of Tables

Abbreviations

Abbreviation	description
aa	amino acids
ADAMTS	a disintegrin and metalloproteinase with thrombospondin motifs
amp	ampicillin
AREX	autoregulation of ER export
ascorbate	L-ascorbic acid-2-phosphate
BFA	brefeldin A
BFP	blue fluorescent protein
BiP	binding immunoglobulin protein
BMP	bone morphogenic protein
BSA	bovine serum albumin
CCS	Cole Carpenter syndrome
CDM	cell derived matrix
CLEM	correlative light electron microscopy
CLSD	cranio–lenticulo–sutural dysplasia
COL domain	collagen domain
$COL1^{mut}$	patient-derived cells from an individual with c.206_208delTGT, p.Leu69del in COL1A1
COMP	Cartilage Oligomeric Matrix Protein
COPI	coat protein complex type I
COPII	coat protein complex type II
CRISPR	clustered regularly interspaced short palindromic repeats
CRTAP	cartilage-associated protein
cTAGE5	cutaneous T-cell lymphoma-associated antigen 5
CTCF	corrected total cell fluorescence
C-terminus	carboxy-terminus
CUL3	ubiquitin ligase cullin3
DAPI	4′,6-diamidino-2-phenylindole
DIC	dynein intermediate chain
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
e.g.	for example
ECM	extracellular matrix
EDS	Ehlers-Danlos Syndrome
EDTA	ethylenediamine tetra acetic acid
EM	electron microscopy

ER	endoplasmic reticulum
ERAD	ER-associated degradation
ERES	ER exit site
ERGIC	ER to Golgi intermediate compartment
ERM	pCytERM_mScarlet-i_N1
FACS	fluorescence-activated cell sorting
FBS	foetal bovine serum
FDR	false discovery rate
FPKM	fragments per kilobase of exon per million reads mapped
fps	frame per second
FRAP	fluorescence recovery after photobleaching
FRET	fluorescent resonance energy transfer
FWHM	full width at half-maximum
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factor
GFP	green fluorescent protein
GFP-COL	procollagen-SBP-mGFP-COL1A1
GFP-COL-RPE	RPE-1 cell line stably expressing GFP-COL
GFP-RPE	RPE-1 cell line stably expressing mGFP
GRASP	Golgi reassembly stacking protein
GTP	guanosine triphosphate
h	hour or hours
HSP47	Heat shock protein 47
hTERT RPE-1	human telomerase immortalized retinal pigment epithelial -1 cells (referred to throughout as RPE-1)
IF	immunofluorescence
IRES	internal ribosome entry site
KD	knockdown
KLHL12	Kelch-like protein 12
KO	knockout
Lab	laboratory
LB	Lysogeny broth
ManII	α mannosidase II
ManII-BFP	StrKDEL-IRES-ManII-mTagBFP2
ManII-mSc	StrKDEL-IRES-ManII-mSc
mCherry	monomeric Cherry
MEM	minimum essential medium
mGFP	monomeric GFP
Mia	melanoma inhibitory antigen
min	minute or minutes

MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger ribonucleic acid
ms	milliseconds
MS	mass spectrometry
mSc	monomeric Scarlet-i
mScCOL1A2	procollagen-mScCOL1A2
MT	microtubule
NHDF-Ad	normal human dermal fibroblasts from adults
N-terminus	amino-terminus
NZ	nocodazole
O-GlcNAc	O-linked β-N-acetylglucosamine
OI	osteogenesis imperfecta
OMIM	Online Mendelian Inheritance in Man
P3H	prolyl-3-hydroxylase
P4H	prolyl-4-hydroxylase
P4HB	prolyl-4-hydroxylase subunit beta
$P4HB^{mut}$	patient-derived cells from an individual with c.1178A>G, p.Tyr393Cys in P4HB
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDI	protein disulphide isomerase
PFA	paraformaldehyde
PRD	proline-rich domain
pro-α1(I)	type I procollagen alpha 1 chain
pro-α2(I)	type I procollagen alpha 2 chain
RNA	ribonucleic acid
RNAseq	RNA sequencing
rpm	revolutions per minute
RT	room temperature
RUSH	retention using selective hooks
s	seconds
SBP	Streptavidin-binding peptide
SDS-PAGE	Sodium dodecyl sulphate – polyacrylamide gel electrophoresis
SNARE	soluble NSF (N-ethyl-maleimide-sensitive fusion protein) attachment protein receptor
SOC	Super Optimal broth
ST	sialyltransferase
ST-Cherry	Str-KDEL-IRES-sialyltransferase-mCherry
STED	stimulated emission depletion
Str-KDEL	Streptavidin-KDEL sequence
TAE	Tris-acetate-EDTA

TALI	TANGO1-like
TANGO1	transport and Golgi organisation 1
TBST	Tris-buffered saline-Tween
tER	transitional ER
TFG	Trk fused gene
TfR	transferrin receptor
TGFβ	transforming growth factor beta
TGN	*trans*-Golgi-network
TMD	transmembrane domain
TMT	Tandem mass tagging
TRAPP	transport protein particle
UPR	unfolded protein response
VSVG	vesicular stomatitis virus G
VTC	vesicular tubular cluster
VUS	Variants of Uncertain Significance
WB	Western Blot
WT	wildtype

Chapter 1: Introduction

1.1 Statement of contribution

Elements of the introduction, experimental methods and results chapters (especially figure legends and figures) are derived from texts and figures I originally wrote, composed and contributed to manuscripts for publications (listed in the appendix) (Balasubramanian et al., 2019; Balasubramanian et al., 2018; McCaughey and Stephens, 2018; McCaughey and Stephens, 2019; McCaughey et al., 2019).

1.2 The secretory pathway – an overview

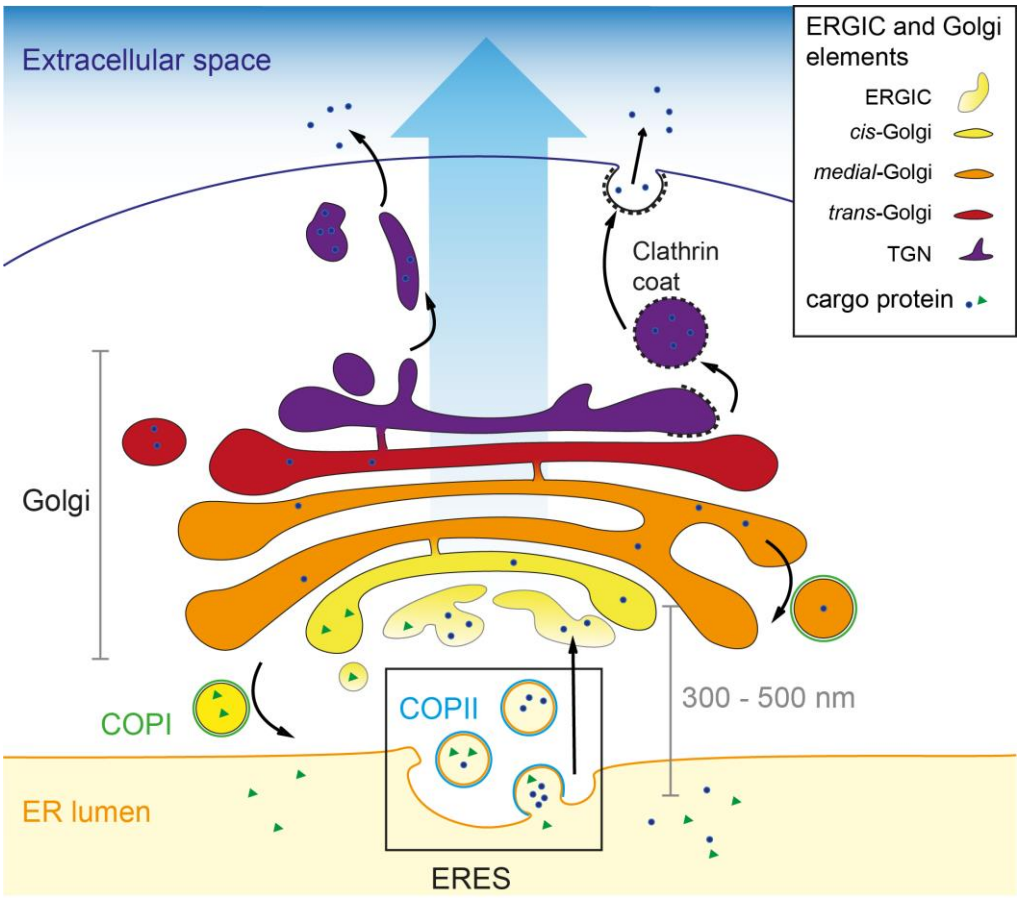


Fig. 1: A simplified overview of the secretory pathway in vertebrates.

Cargo proteins are packaged at ER exit sites (ERES) into coat complex type II (COPII) vesicles. From there secretory cargo needs to be transported via the ER–Golgi intermediate compartment (ERGIC) to the *cis*-, *medial*-, and *trans*-Golgi. From here clathrin coated vesicles, secretory granules and tubular structures can bud from the *trans*-Golgi network (TGN) to facilitate secretion into the extracellular space. Retrograde transport between Golgi elements, the ERGIC and the ER is facilitated by coat complex type I (COPI). This figure was adapted from one previously published (McCaughey and Stephens, 2019).

It is estimated that about a third of all proteins need to traverse the endoplasmic reticulum (ER) (Huh et al., 2003) for folding, post-translational modification, quality control and sorting (Bannykh et al., 1996; Mezzacasa and Helenius, 2002). From here, proteins destined for e.g. the extracellular space need to be transport to the Golgi apparatus (Golgi, 1989), or in vertebrates the ER-Golgi intermediate compartment (ERGIC, discussed in 1.2.2) prior to reaching the Golgi (Fig. 1) (Aridor et al., 1995; Bannykh and Balch, 1997; Hauri and Schweizer, 1992; Saraste and Kuismanen, 1992). The ER is approximately 300 – 500 nm away from ERGIC elements as shown by electron microscopy (EM) (Hanna et al., 2018). ER-to-ERGIC transport does not require the microtubule (MT) network, but the latter may play a role in stimulating this transport step (Scales et al., 1997; Watson et al., 2005). Transport from the ER to the ERGIC/Golgi is facilitated by coat protein complex type II (COPII) vesicles (Barlowe et al., 1994; Barlowe and Schekman, 1993; Yoshihisa et al., 1993), which will be elaborated in detail in 1.5.1. At ER exit sites (ERES) cargo proteins incorporated into COPII vesicles bud from the transitional ER (tER) (Barlowe et al., 1994; Kuehn and Schekman, 1997) and fuse with ERGIC membranes, releasing the cargo into the next compartment (Bannykh and Balch, 1997; Bannykh et al., 1996; Hauri and Schweizer, 1992; Saraste and Kuismanen, 1992). Therefore, retrograde flow of proteins and membranes being recycled back to earlier compartments is essential for homeostasis of membranous compartments and proteins involved in transport processes (Boncompain and Perez, 2013; Johannes and Popoff, 2008; Pavelka et al., 2008). Retrograde transport from the Golgi to the ERGIC and the ERGIC to the ER is mediated by coat complex type I (COPI) (Pepperkok et al., 1993; Scales et al., 1997; Stephens and Pepperkok, 2002). Secretory proteins can undergo further post-translational modifications, especially glycosylation, in the Golgi apparatus (1.2.3), which is a polarised organelle with the *cis*-Golgi facing ERES, followed by the *medial*-Golgi and the *trans*-Golgi network (TGN). Here, proteins are further sorted depending on their final intracellular or extracellular destination (Farquhar and Palade, 1998) including the Golgi, endosomes, lysosomes, mitochondria, peroxisomes, the plasma membrane or the extracellular space.

Another vesicular coat protein, clathrin (Pearse, 1975; Roth and Porter, 1964), plays a role at the TGN, where it can initiate the formation of clathrin coated vesicles destined for the extracellular space in addition to its role in endocytosis.

Assembly of vesicles, budding, fission, uncoating and heterotypic or homotypic fusion with other membranous structures is a common process amongst all vesicular structures. Rab-GTPases act in the recruitment and disassembly of coat proteins and facilitate association of some transport intermediates with molecular motors for MT- or actin-dependent directed transport (Boncompain and Perez, 2013). The Rabs are broadly considered to define organelle identity (e.g. Rab11 on late endosomes, Rab1 and Rab2 at the ER-Golgi interface). Uptake of proteins and nutrients from the

extracellular space is mediated by endocytosis assisted by e.g. the transferrin receptor (TfR).

Together, the secretory pathway provides an essential asset for controlling homeostasis, cell development and general cell function.

1.2.1 Quality control in the ER

Instead of export from the ER proteins that are not correctly assembled or misfolded need to be eliminated. Accumulation of misfolded or unassembled proteins in the ER triggers ER-stress and with it the unfolded protein response (UPR) (Morito and Nagata, 2015; Yoo et al., 2017), which in turn induces chaperone activity to promote accelerated refolding and maintains proteostasis. ER-stress is recognised by sensors like PERK and IRE1 and triggers UPR activation (Bertolotti et al., 2000; Pincus et al., 2010; Shen et al., 2005). PERK furthermore upregulates expression of chaperones and proteins involved in autophagy amongst others (Acosta-Alvear et al., 2007; Calfon et al., 2002; Shen et al., 2001; Yoshida et al., 2001). Proteins exhibiting incorrect conformation are cleared from the ER via ER-associated degradation (ERAD) or autophagy (Hartl and Hayer-Hartl, 2009; Ruggiano et al., 2014). ERAD facilitates ubiquitination of misfolded proteins in the ER and translocates them into the cytosol where they undergo degradation by the proteasome (Jensen et al., 1995; McCracken and Brodsky, 1996; Ward et al., 1995). Autophagy of the ER has been termed ERphagy and is linked to the activation of UPR (Baba et al., 1994; Bernales et al., 2006; Takeshige et al., 1992). Early autophagosomal structures assemble on ER membranes (Axe et al., 2008), including WIPI2, which subsequently recruits ATG16L amongst other proteins that promote elongation of the autophagosomal membranes and initiates LC3 lipidation required for autophagosome formation (Dooley et al., 2014). Subsequently, these autophagosomes merge with lysosomes (or endosomes) to become autolysosomes that readily degrade incorporated proteins. Upon correct assembly or folding, proteins can exit the ER.

1.2.2 The ER-Golgi intermediate compartment (ERGIC)

The identity and function of the ERGIC has been and remains a topic of extensive discussion in the field of protein transport.

The biochemical composition of ERGIC differs from its neighbours the ER and the Golgi (Schweizer et al., 1991). Functionally, ERGIC acts as cargo sorting intermediate (Appenzeller et al., 1999; Aridor et al., 1995; Ben-Tekaya et al., 2005; Klumperman et al., 1998; Martinez-Menarguez et al., 1999). The ERGIC is composed of vesicular tubular clusters (VTCs) or resembling pleomorphic structures with multiple buds at the edges (Appenzeller-Herzog and Hauri, 2006; Horstmann et al., 2002). Several studies using three-dimensional reconstitution of serial sections suggest that its membranes are not continuous with those of the ER or *cis*-Golgi (Appenzeller-Herzog and Hauri,

2006; Sesso et al., 1994). ERGIC-53 is considered a marker of the ERGIC and acts as a lectin-like cargo receptor (Aridor et al., 1995; Ben-Tekaya et al., 2005; Martinez-Menarguez et al., 1999). Some data are consistent with the ERGIC as a stable and non-moving compartment (Ben-Tekaya et al., 2005; Klumperman et al., 1998), while others suggest a more transient and dynamic role for ERGIC structures (Appenzeller-Herzog and Hauri, 2006; Hammond and Glick, 2000; Horstmann et al., 2002). The ERGIC might adapt its morphology and tendency to be dynamic depending on the need to traffic lower or higher amounts of cargo proteins.

1.2.3 The Golgi apparatus

Secretory proteins like collagens need to traverse the Golgi prior to secretion into the extracellular space. The Golgi has a characteristic structure of stacked cisternae. Golgi morphology differs between species, as exemplified in Fig. 2, and will therefore have an impact on ER, ERGIC and Golgi dynamics, including hetero- and homotypic fusion events and maturation. As one of the main protein sorting hubs in the cell, the Golgi apparatus needs to be able to recognise, capture, sort and modify cargo proteins, while maintaining its own morphology.

Golgi morphology and composition

In simpler, typically unicellular, eukaryotes there are usually one to a few discrete Golgi stacks per cell (Kurokawa and Nakano, 2019). Interestingly, yeast displays different forms of Golgi structure with several stacks in *Pichia pastoris* (Bevis et al., 2002), compared to unstacked Golgi elements in *Saccharomyces cerevisiae* (Okamoto et al., 2012) (Fig. 2).

Amongst eukaryotes the Golgi assembles into layers of cisternae that result in a Golgi stack. In vertebrates the Golgi stacks are interlinked with the help of Golgi reassembly stacking protein (GRASP)s (Feinstein and Linstedt, 2008; Puthenveedu et al., 2006; Rambourg et al., 1979) and can better be described as a single ribbon with close apposition to the nuclear envelope. The ribbon or stacks can be defined as *cis*-, *medial*- and *trans*-Golgi regions that contain specific proteins specialised for tethering vesicles, post-translational modifications and vesicle formation for both anterograde and retrograde trafficking (Boncompain and Perez, 2013).

The *cis*-side of the Golgi faces the ER from which cargo proteins are delivered from ERES. Anterograde transport (ER-to-Golgi transport is discussed in detail in 1.5) is facilitated by COPII (Barlowe et al., 1994), while retrograde transport from the ERGIC or Golgi and between Golgi stacks occurs in coordination with COPI-coated vesicles (Guo et al., 1994; Klumperman et al., 1998; Letourneur et al., 1994). COPI is also involved in anterograde transport from ERES to the ERGIC (Stephens and Pepperkok, 2001). *Cis*-Golgi elements form *de novo* from the ER into ERES and mature from *cis*- to *trans*-Golgi (Bannykh et al., 1998; Bevis et al., 2002; Losev et al., 2006; Mironov et al., 2003; Stephens and Pepperkok, 2002).

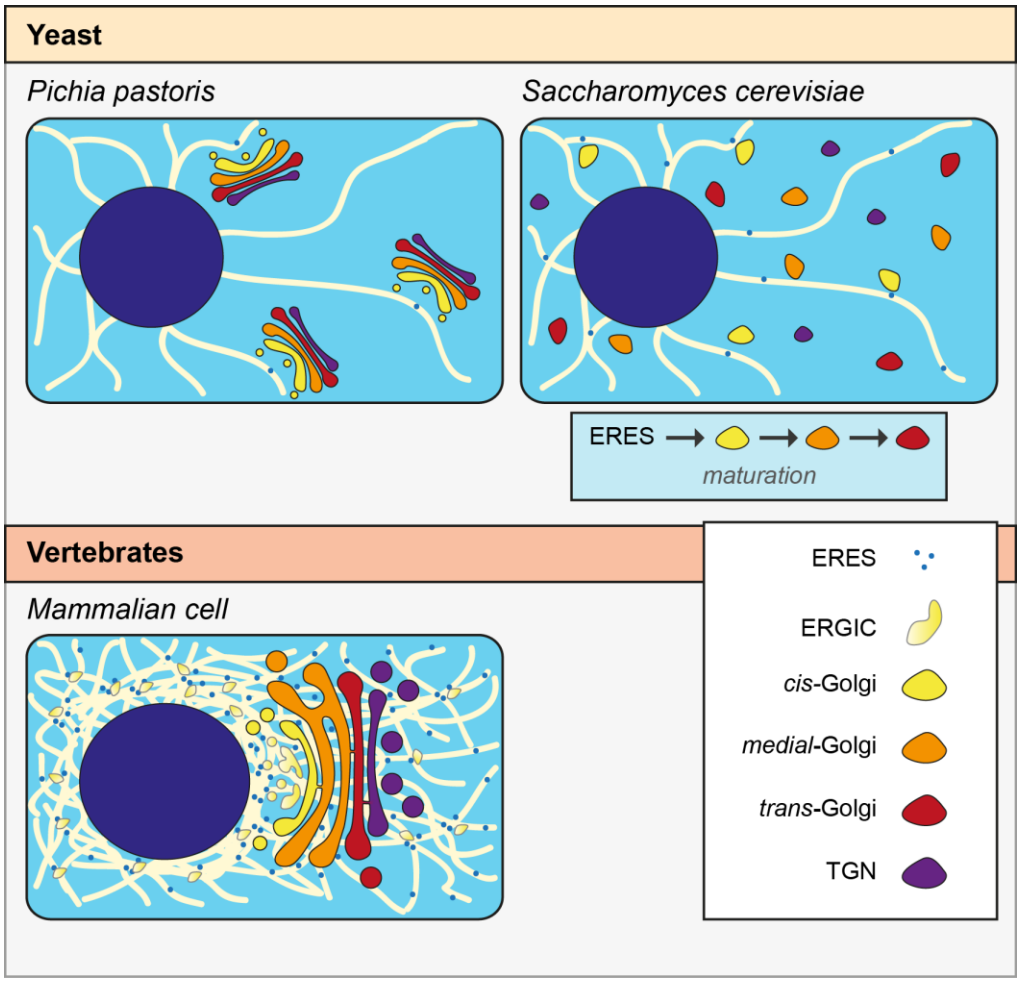


Fig. 2: Golgi structures in yeast and vertebrates.

The cytosol is displayed in light blue, the nucleus in dark blue and the ER in pale yellow. Golgi elements in Saccharomyces cerevisiae are formed *de novo* from the tER into ERES and mature from *cis*, via *medial* to *trans*-Golgi. This figure was created based on previous images published (Bevis et al., 2002; Suda and Nakano, 2012).

Different cisternae harbour different enzyme and protein compositions. As an example, the N-linked oligosaccharide modifying enzyme α mannosidase II (ManII) can be used as a marker for the *medial*-Golgi (Quinn et al., 1983), while sialyltransferase (ST)

resides in the *trans*-Golgi (Berger et al., 1993; Canty and Kadler, 2005; Griffiths et al., 1982; Roth and Berger, 1982).

Linking of Golgi stacks and tethering of transport vesicles: golgins

Golgins play an important role in tethering transport vesicles (e.g. Golgi matrix protein GM130 can tether vesicles during anterograde and retrograde transport (Wong and Munro, 2014)) and Golgi membranes (Munro, 2011). These Golgi-membrane anchored proteins contain long coiled-coil domains that reach out into the cytoplasm and are called golgins (Munro, 2011). Different golgins have been associated with facilitating transport of specific cargo proteins (Boncompain and Perez, 2013) and linking of Golgi cisterna is mainly dependent on GRASP55 and 65 (Barr et al., 1997; Grond et al., 2020; Puthenveedu et al., 2006; Xiang and Wang, 2010). GRASP65 binds GM130 and is localised to the *cis*- and *medial*-Golgi (Bachert and Linstedt, 2010; Barr et al., 1998). The involvement in vesicle tethering of giantin, the largest member of the golgin family (Linstedt and Hauri, 1993; Seelig et al., 1994), has yet to be discovered (Stevenson et al., 2020). This protein, however, has been shown to be required for the regulation of glycosylation (Kikukawa et al., 1990; Koreishi et al., 2013; Lan et al., 2016; Petrosyan et al., 2014; Stevenson et al., 2017), as well as the extracellular matrix (ECM) (Katayama et al., 2018; Kikukawa and Suzuki, 1992; Lan et al., 2016; Stevenson et al., 2020) e.g. impacting procollagen (the precursor of collagen) processing (Stevenson et al., 2020).

1.3 Procollagen – a large secretory cargo protein

1.3.1 Importance of understanding (pro)collagen transport

Collagens are one of the most abundantly expressed proteins in the animal kingdom and make up 30 – 35% of total proteins in the human body (Smith and Rennie, 2007). Collagen is most present in the ECM in form of centimetre-long fibrils organised into tissue specific arrangements like parallel bundles in tendon and wave-like orthogonal lattices in skin (Jansen et al., 2018). The extracellular scaffolds provided by collagen fibrils allow the attachment of various macromolecules, including cell surface receptors (Humphries et al., 2006; Wickstrom et al., 2011) and furthermore provides tissues with the necessary structural support and elasticity. Thus, the ECM plays an important role during development in tissue morphogenesis, differentiation and homeostasis, as well as cell migration, adhesion, wound healing, aging, and fibrosis (Frantz et al., 2010). An insufficient amount of collagen is associated with poor wound repair, osteoarthritis and tendinopathies, for which there are currently no effective treatment options (Butler et al., 2008; Ha et al., 2018). Development of therapeutics to address these conditions has been hampered by the limited understanding of the fundamental processes of collagen biosynthesis, processing and transport, maintenance, as well as repair.

The most abundant collagen is fibrillar type I collagen, which is mainly produced in fibroblasts and osteoblasts and assembles into fibrils together with type III (Fleischmajer et al., 1990) and type V collagen (Birk et al., 1988) in the ECM. Due to

the size of type I procollagen of potentially 300 nm in length (Bachinger et al., 1982; Lightfoot et al., 1992), this protein can also be used as a model to investigate the transport of large secretory proteins. Type I collagen plays a key role in the formation of skin, bone and tendons and will be the main focus from here on.

1.3.2 Structural differences within the collagen family

Structurally all collagens contain at least one collagen domain (COL domain) composed of Gly-X-Y triplet repeats (Ramshaw et al., 1998) (Fig. 3). The position of X is often occupied by a proline or hydroxyproline, whereas Y is often a lysine or hydroxylysine residue (Harwood et al., 1975; Morgan et al., 1970). These repetitive motifs result in an alpha-helical secondary structure. Three procollagen chains trimerize to form triple helical procollagen. In the case of type I procollagen, the COL domain is flanked by non-triple helical N- and C-terminal propeptides stabilised by inter- and intramolecular disulphide bonds (Prockop et al., 1979a; Prockop et al., 1979b).

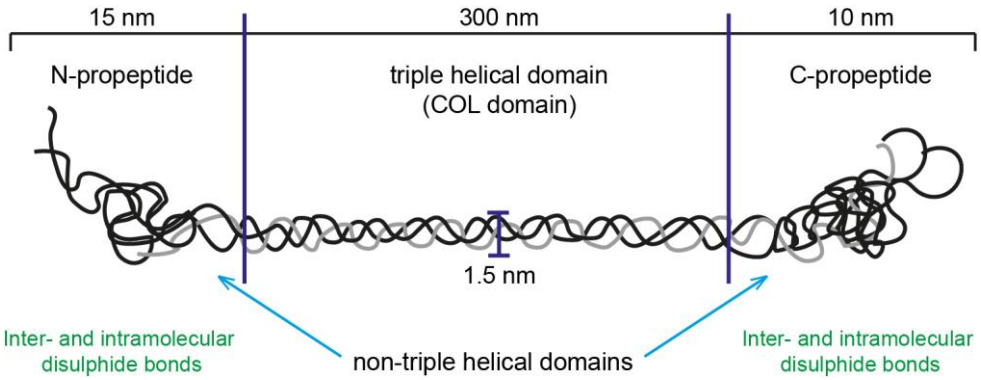


Fig. 3: Schematic of type I procollagen.

The triple helical COL domain (1.5 nm wide and about 300 nm in length) is flanked by the N- and C-terminal propeptides of 15 and 10 nm, respectively. The procollagen trimer is stabilised via inter- and intramolecular disulphide bonds and water bridges (Prockop et al., 1979a; Prockop et al., 1979b).

To date 28 different types of collagens encoded by 42 different collagen genes have been identified (Ricard-Blum, 2011). Collagens can be divided into different subfamilies according to their structure and function (Table 1). Fibrillar collagens (types I, II, III, V, XI, XXIV and XXVII) have a characteristic uninterrupted COL domain and are widely expressed in multicellular organisms ranging from sponges to humans (Ozbek et al., 2010). Type VII collagen has a unique structure containing fibronectin type II repeats and functions in connecting the epidermis to the dermis by forming anchoring fibrils (Keene et al., 1987).

Table 1: Different collagen subfamilies and their structure and function.

(Canty and Kadler, 2005; Ricard-Blum, 2011)

Subfamily of collagens	Collagen types	Structure	Function/ involvement
Fibrillar	I, II, III, V, XI, XXIV, XXVII	Non-interrupted COL domain	Formation of fibrillar collagens
Fibril-associated	IX, XII, XIV, XVI, XIX, XX, XXI, XXII	Interrupted COL domains	Associated with the surface of collagen fibrils
Network-forming	IV, VIII, X	Dimers and tetramers of trimers (VIII and X less than 600 aa)	Network formation
Membrane	XIII, XVII, XXIII, XXV	Transmembrane proteins, multiple COL domains in extracellular space	Tissue specific
Mulitplexins	XV, XVIII	Thrombospondin domain and multiple interrupted COL domains	Associated to tumour progression[1]
Other collagens	VI, VII, XXVI, XXVIII	1 – 2 COL domains, often rich in von Willebrand domains	

[1](Lakshmanachetty and Koster, 2016)

1.3.3 Procollagen folding and processing

Procollagen biosynthesis and trimer formation in the ER

Procollagen is synthesised in the ER and forms trimers composed of three procollagen chains (Goldberg et al., 1972). Trimerisation occurs prior to ER export. In the case of type I procollagen two α1 chains (pro-α1(I)) together with one α2 (pro-α2(I)) chain assemble in the ER lumen (Fig. 4A) to form trimeric procollagen (Canty and Kadler, 2005; Goldberg et al., 1972). This step is facilitated by several enzymes that act on different domains of the procollagen chains. The chaperone binding immunoglobulin protein (BiP) and the protein disulphide isomerase (PDI) bind the C-terminal propeptide (Chessler and Byers, 1993; Wilson et al., 1998). PDI and prolyl-4-hydroxylase (P4H) act to retain monomeric procollagen in the ER (Bottomley et al., 2001; Walmsley et al., 1999). P4H and other proteins like the collagen chaperone heat shock protein 47 (HSP47; (Ito and Nagata, 2017; Satoh et al., 1996; Tasab et al., 2000)) interact with the COL domain of the procollagen chains (Bachinger et al., 1978; Chessler and Byers, 1992; Vranka et al., 2004). The role of HSP47 will be addressed in more detail in 1.3.4.

P4H utilises the redox activity of ascorbic acid to maintain an active state and hydroxylates the monomeric procollagen chains to facilitate trimer stability (Berg and Prockop, 1973a; Berg and Prockop, 1973b; Mussini et al., 1967) (Fig. 4B). Procollagens undergo extensive port-translational modifications via O- and N-linked glycosylation. About a third of the hydroxylysines in type I collagen are, in addition, glycosylated prior to trimerisation (Pinnell et al., 1971).

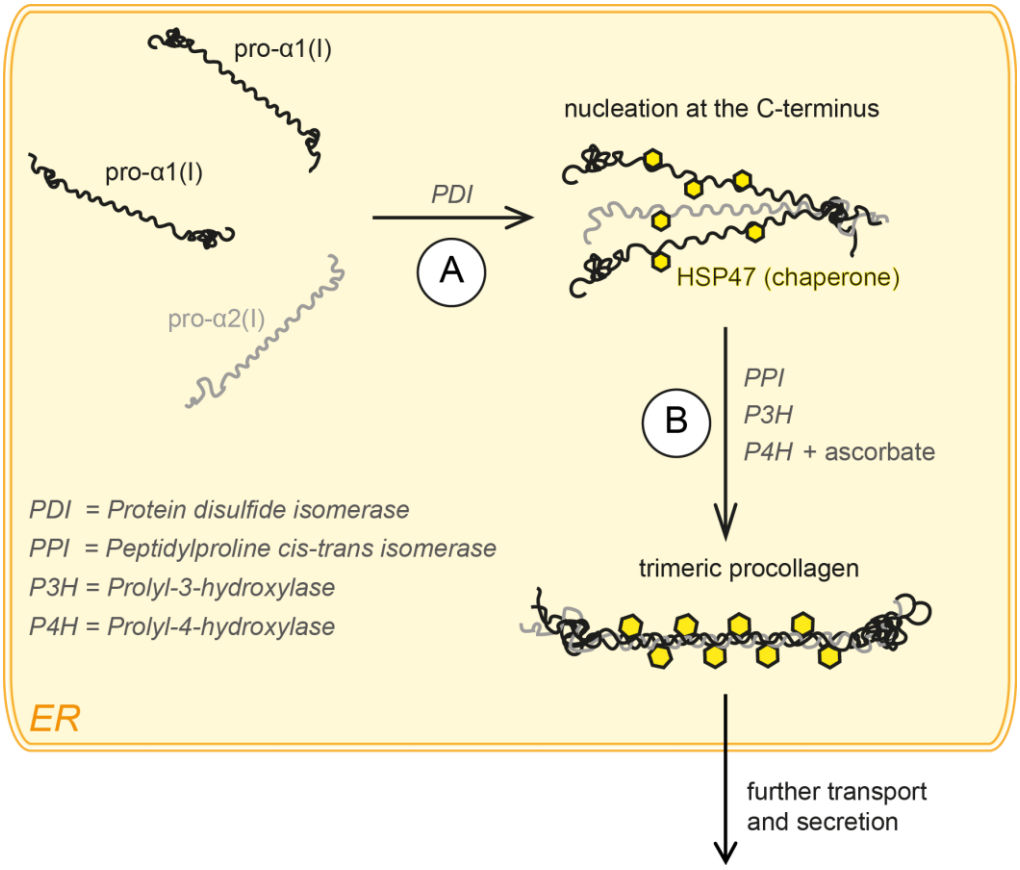


Fig. 4: Procollagen assembly and processing in the ER.

Procollagen alpha chains are synthesised in the ER. Disulphide bonds are formed within the N- and the C-propeptides of the same as well as different chains via PDI, initiating trimer formation from the C- to the N-terminus, assisted by the collagen chaperone HSP47 (A). The peptidyl proline *cis-trans* isomerase (PPI), as well as P3H and P4H (assisted by its co-factor ascorbate) work together to facilitate formation and stabilisation of the procollagen trimer (B).

Trimer formation itself occurs from the C-terminus towards the N-terminus in a zipper-like manner (Engel and Prockop, 1991). Upon correct assembly, folding and processing, procollagen trimers are subsequently transported to the next compartment. This transport step will be described in detail in a separate section 1.5.2 and 1.5.5. Once procollagen arrives in the Golgi apparatus it undergoes further post-translational processing.

Procollagen transit through the Golgi

Data on transport of procollagen through the Golgi apparatus laid the basis for the cisternal-maturation model with Golgi stacks maturing from *cis*- to *trans*, while stack-specific enzymes are recycled back via retrograde transport through either COPI (Lanoix et al., 1999; Love et al., 1998; Sonnichsen et al., 1996) or via inter-cisternal connections (Kweon et al., 2004; Trucco et al., 2004). Transit of procollagen through the Golgi leads to 300 – 400 nm distensions in cisternae (Leblond, 1989; Marchi and Leblond, 1984; Trelstad and Hayashi, 1979; Weinstock and Leblond, 1974) and procollagen does not appear to leave the Golgi stacks (Bonfanti et al., 1998) (Fig. 5A), while other large structures, such as engineered protein aggregates with ~400 nm diameters were seen to be transported between stacks by using megacarriers budding off from the cisternal rims (Volchuk et al., 2000). Inside the Golgi and the extracellular space, procollagen needs to undergo further processing to result in mature type I collagen.

Proteolytical cleavage of the N- and C-propeptides

Before assembly into fibres in the extracellular space, both N- and C-propeptides are cleaved (Leung et al., 1979)(Fig. 5). C-propeptide cleavage triggers fibril self-assembly of collagens by decreasing the critical concentration of collagen required for this step (Hulmes et al., 1989; Kadler et al., 1987; Miyahara et al., 1982). Although cleavage of the N-propeptide is not essential for fibrillogenesis (Bornstein et al., 2002; Hulmes et al., 1989; Romanic et al., 1992), it can alter fibril diameters by affecting accretion (Asgari et al., 2017; Linsenmayer et al., 1993). Cleavage of type I procollagen C-propeptide is facilitated by bone morphogenic protein 1 (BMP-1), a metalloproteinase of the tolloid family, and occurs most likely in the extracellular space (Kessler et al., 1996). N-proteinases of fibrillar procollagens are a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS)-2, -3 and -14 (Colige et al., 2005; Colige et al., 2002; Fernandes et al., 2001). There is evidence that cleavage of the N-propeptide can only occur on properly folded triple helical procollagen (Tanzawa et al., 1985). Since N-terminal processing is not impaired by brefeldin A (BFA)-treatment (which leads to a merging of the ER with *cis*- and *medial*-Golgi cisternae) removal of the N-propeptide can happen prior to reaching the *trans*-Golgi (Canty-Laird et al., 2012; Humphries et al., 2008). This is further supported by recent data showing a requirement for the golgin giantin for removal of the N-terminal propeptide (Stevenson et al., 2020). There is also evidence that meprin proteins α and β are involved in terminal propeptide cleavage of types I, II and III procollagen (Broder et al., 2013). Culture medium from fibroblasts was shown to be able to process procollagen into mature collagen, indicating that endopeptidases required for this enzymatic process can be secreted and remain active in the extracellular space (Kerwar et al., 1973; Layman and Ross, 1973).

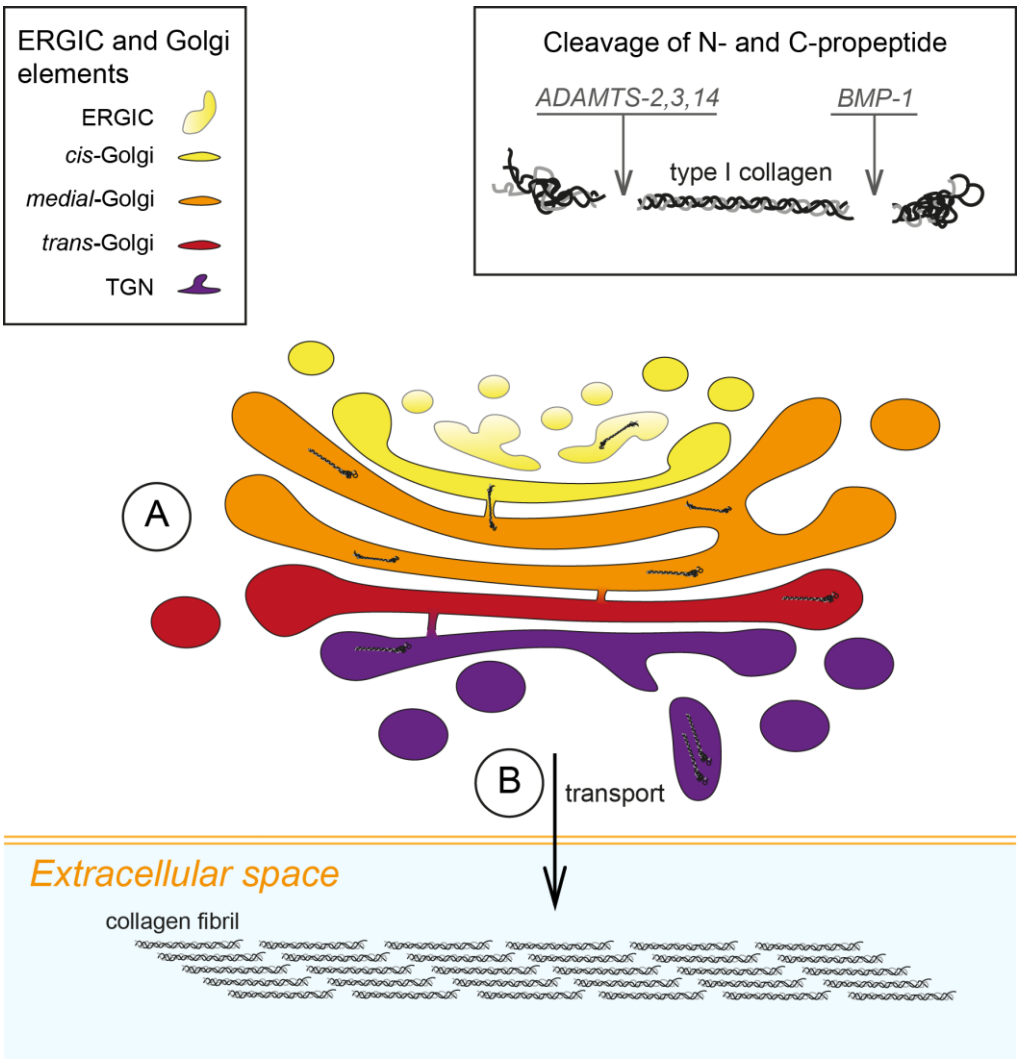


Fig. 5: Type I procollagen transit through the Golgi processing and fibrillogenesis.

The procollagen trimers transit through the Golgi without leaving the cisternae (A). The procollagen trimer is subsequently secreted from the cell into the extracellular space (B). Prior to fibrillogenesis, the N- and C-propeptide of the trimer are cleaved via endopeptidases ADAMTS-2,3,14 and BMP-1, resulting in mature type I collagen. Finally, collagen fibrils are assembled into the ECM in form of tissue-dependent aligned fibres. The staggered alignment of collagen fibrils results in the well-known banding pattern observed by EM (Birk and Trelstad, 1984; Kadler et al., 2008; Piez and Miller, 1974).

Post-Golgi transport of procollagen to the plasma membrane

After transit trough the *cis*- and *medial*-Golgi, cargo reaches the TGN (Griffiths and Simons, 1986; Keller and Simons, 1997; Orci et al., 1987) from which e.g. Rab6-positive carriers facilitate transport to the plasma membrane in a MT-dependent manner (Fourriere et al., 2019; Nakagawa et al., 2000; Teng et al., 2005). So far, the dependence of procollagen transport on Rab6 has only been shown for type X collagen (Fourriere et al., 2019), which is a small globular non-fibrillar collagen (Schmid and Linsenmayer, 1983). Post-Golgi carriers appear in a variety of morphologies from vesicular to tubular and can follow diverse routes to different destinations, including the

plasma membrane (Canty and Kadler, 2005; Ponnambalam and Baldwin, 2003). Procollagen bundles have been observed to leave the TGN in form of electron dense secretory vacuoles of 500 nm length by EM (Leblond, 1989; Marchi and Leblond, 1984; Trelstad, 1971; Trelstad and Coulombre, 1971; Trelstad and Hayashi, 1979; Weinstock and Leblond, 1974) (Fig. 5B). Using correlative light EM (CLEM) the small protein vesicular stomatitis virus G (VSVG) was also observed in the same post-Golgi carriers as type I procollagen, ranging from 300 – 1700 nm that detach from the TGN (Canty and Kadler, 2005; Polishchuk et al., 2003; Polishchuk et al., 2000).

Fibrillogenesis and assembly in the extracellular matrix (ECM)

Initial fibril formation may start at the stage of post-Golgi transport, as aligned procollagen bundles can be observed in the intracellular space and secreted into the cell culture medium (Fig. 5B) (Bruns et al., 1979; Goldberg, 1974; Hulmes, 1983). In the extracellular space collagen fibrils are assembled into aligned supramolecular insoluble fibre structures via inter- and intramolecular crosslinking facilitated by lysyl oxidases (Kagan and Trackman, 1991) and dependent on other extracellular scaffolds like fibronectin (Canty and Kadler, 2005; Velling et al., 2002; Wierzbicka-Patynowski and Schwarzbauer, 2003). Together with fibronectin, proteoglycans and cell adhesive glycoproteins collagen forms the ECM, which differs in its composition, alignment, and flexibility depending on the cell type, tissue, as well as the developmental stage (Kadler et al., 2008). Collagen is first deposited into the ECM by adult tissues during embryogenesis (Kalson et al., 2015) and extracellular collagen networks have to be maintained through life without renewal (Heinemeier et al., 2013). Thus, collagen synthesis, transport, and secretion as well as matrix remodelling and maintenance have to be well regulated.

1.3.4 Regulation of collagen synthesis, transport and secretion

Regulation of type I procollagen expression includes various pathways and factors including insulin-line growth factor-1, transforming growth factor and interleukins (Rossert et al., 2000). Increased levels of transforming growth factor beta (TGFβ) result in an increase in mRNA abundance for fibronectin and collagen (Ignotz et al., 1987). The N-propeptide of pro-α1(I) can also act as a negative regulator for pro-α1(I) synthesis during translational chain elongation or termination (Horlein et al., 1981; Oganesian et al., 2006; Paglia et al., 1979).

Post-transcriptional regulation of type I procollagen is also linked to the circadian rhythm. Cellular activities have been linked to and optimised to address the different necessities occurring throughout the day by the cell autonomous circadian clock which allows the tracking of day and night (or active and rest/inactive) cycles (Cermakian and Sassone-Corsi, 2000; Roenneberg and Merrow, 2005). Two distinct temporally different modes of procollagen transport and secretion have been reported in recent years: a circadian clock-regulated pathway for general ECM and tissue maintenance

that replaces old collagen with 10% of total collagen synthesised on a daily basis (Yeung et al., 2018), as well as a constitutive pathway with rapid secretion within minutes associated to a quick repair response and ongoing maintenance (Canty et al., 2004). During wound healing immune cells like macrophages, which are a part of the immune response regulated by the circadian clock (Man et al., 2016; Scheiermann et al., 2013), instruct mesenchymal cells like fibroblasts and thereby regulate matrix repair (Knipper et al., 2015; Mori et al., 2008; Morris et al., 2018).

The constitutive and the circadian clock-regulated pathways can be investigated *ex vivo* in tissues and cell culture through synchronisation of the circadian clock triggered by e.g. incubation of dexamethasone or serum shock (Yeung et al., 2018). Rhythmicity of collagen deposition is induced downstream of transcription (Yeung et al., 2018; Yeung et al., 2014).

Quality control of procollagen in the ER

Quality control in the ER also impacts on ER-to-Golgi transport. As a chaperone, HSP47 is required to prevent aggregation of procollagens in the ER and facilitates trimerisation and proper folding (Ishida et al., 2006; Kawasaki et al., 2015; Marutani et al., 2004). While also binding the monomeric procollagen (Hosokawa and Nagata, 2000), HSP47 appears to show a higher affinity for trimeric procollagens, independent of the specific procollagen type (Ito and Nagata, 2017; Koide et al., 1999; Koide et al., 2006; Ono et al., 2012; Satoh et al., 1996; Tasab et al., 2000). Inhibition of different transport steps in the Golgi showed that HSP47 can bind to procollagen both in the ER and the *cis*-Golgi (Satoh et al., 1996), from where it is recycled back to the ER via its C-terminal RDEL retention sequence (Satoh et al., 1996). RDEL, just like KDEL motifs, are recognised by KDEL receptors, which ensure that ER-resident proteins are not transported through the secretory pathway, but instead recycled back via COPI (Andres et al., 1990; Brauer et al., 2019). KO of HSP47 in mice impairs proper procollagen trimer formation and secretion and causes embryonic lethality (Nagai et al., 2000).

Besides quality controls like ERAD and UPR, an alternative mode for misfolded procollagen degradation has been identified. This process can occur at ERES, which are then marked by ubiquitin, recruit autophagy effectors and are subsequently incorporated by nearby lysosomes (Lu et al., 2018; Omari et al., 2018).

1.4 Collagen-related diseases

Efficient collagen assembly, processing and transport is essential for a functioning ECM, as defects lead to pathologies like osteogenesis imperfecta (OI), Ehlers-Danlos syndrome (EDS) and fibrosis (Forlino and Marini, 2016; Jobling et al., 2014).

1.4.1 Osteogenesis Imperfecta (OI)

In 1849 Willem Vrolik was the first to link various observations together and conclude that these belong to the same group of disorder giving it the name it is now known for: *osteogenesis imperfecta*, meaning imperfect bone formation (Schievink et al., 1994). Up to today different subtypes of OI with phenotypes of different severity have been classified. OI is an inherited bone fragility disorder usually diagnosed in childhood with an overall occurrence of 1 in 15,000 (Lim et al., 2017). OI reflects symptoms of low bone mass and increased bone fragility observed in patients with degenerative osteoporosis, which is often linked to old age (Balasubramanian et al., 2019; Rauch and Glorieux, 2004). Other phenotypes that can be linked to OI are hypermobility of skin and ligaments, hearing impairment, Wormian bones on radiographs of the skull, dentinogenesis imperfecta, blue sclera and muscle weakness (Marom et al., 2020; Rauch and Glorieux, 2004).

OI can be classified into at least 20 different subtypes based on the underlying affected genes. A commonly used classification relies on the system by Sillence, categorising OI into four major subtypes (Sillence et al., 1979). Clinical severity of OI can range from perinatal lethality to mild forms without fractures (Rauch and Glorieux, 2004; Shapiro et al., 1992). The severity of the common feature of bone fragility increases from type I (mild) < to type IV (moderately to severely deforming)< type III (severely deforming) < type II (perinatally lethal) (Rauch and Glorieux, 2004). So far mutations in 18 genes have been identified that cause OI. The most common mutations of OI patients are in the genes encoding for type I and II procollagen (COL1A1 and COL1A2) (Forlino and Marini, 2016; Marom et al., 2020; Pihlajaniemi et al., 1984). The remaining 15 - 25% of OI cases are caused by mutations in genes affecting type I collagen synthesis, processing, stability and transport, or functionality of osteoblasts, the cells responsible bone formation (Byers et al., 2006; Marom et al., 2020). The most severe cases of OI result from mutations disrupting the Gly-X-Y motif of the COL domain (Bonadio et al., 1985; Vogel et al., 1988) or can be associated with mutations at the C-propeptide, which is essential for chain recognition during trimerisation (Lees et al., 1997). Alterations of the COL domain also leads to a prolonged time in the ER, resulting in excessive post-translational modifications further affecting transport and correct fibril assembly (Bonadio et al., 1985). In addition, mutations in HSP47 also result in severe recessive OI (Christiansen et al., 2010). OI types V, VI and VII, which account for about 8% of children with OI, are not caused by mutations in COL1A1 or COL1A2 (Byers et al., 2006).

On a cellular level, mutations affecting type I collagen biosynthesis lead to distended ER, ER stress, a reduction in type I collagen secretion, defective signalling from the matrix to the cell and increased extracellular mineralisation when analysing osteoblasts (Marini et al., 2017; Marom et al., 2020). Other cells like fibroblasts also show impaired type I collagen synthesis and secretion subsequently affecting eye, lung and heart

valve tissue, explaining the non-skeletal phenotypes associated with OI (Dimori et al., 2020; Marom et al., 2020).

Diagnosis of OI is performed using multidisciplinary approaches involving genetic screening, and radiological analysis, it can, however, go unrecognised for a long time and only become apparent under special circumstances (Balasubramanian et al., 2019). The only treatment option for patients with OI (long-term and for treatment of children) are bisphosphonate injections, corrective surgery, and physical therapy. Bisphosphonate can increase bone mineral density by inhibiting bone resorption and osteoclast activity (Marom et al., 2020); however, bone quality is not improved. Novel treatment options using denosumab (Hoyer-Kuhn et al., 2014), sclerostin inhibitory antibodies (Glorieux et al., 2017) and inhibition of TGF-β (Grafe et al., 2014), have been started to be explored in order to also address the inherent bone fragility in combination with the bone mineral density (Marom et al., 2020). To date there is no cure for OI.

Cole Carpenter Syndrome

Cole Carpenter Syndrome (CCS; OMIM #112240 with OMIM being Online Mendelian Inheritance in Man (https://www.omim.org/)) is a rare subtype of OI with a distinct phenotype, but genetic heterogeneity. The European definition for a disease to be "rare" is an occurrence of up to 5 in 10,000. CCS is characterised by severe bone fragility, facial dysmorphism (overly large head with macrocephaly with wide open fontanelle, blue sclera, small nose with flat nasal bridge, and a broad face with ocular proptosis) with or without craniosynostosis (a birth defect resulting in bones of the skull joining together too early) and normal intelligence (Balasubramanian et al., 2018). Cole and Carpenter first reported two unrelated infant patients showing a phenotype similar to severe OI with bone deformities and multiple fractures, blue sclerae, craniosynostosis, hydrocephalus and distinct facial features in 1987 (Cole and Carpenter, 1987). A 12-year old patient with similar CCS phenotype was described three years later (Balasubramanian et al., 2015). This patient had a homozygous mutation in cartilage-associated protein (CRTAP), indicating CCS as a variant of recessive OI. CRTAP interacts with P3H1 to facilitate prolyl 3-hydroxylation of collagens (Morello et al., 2006). In 2015 a heterozygous missense mutation in P4H subunit B (P4HB) was linked to the original patients investigated by Cole and Carpenter by Rauch and colleagues, now known as CCS1 (CLCRP1 OMIM #11240)(Rauch et al., 2015). Another variation of CCS is CCS2 (CLCRP2 OMIM #616294) with a phenotype resembling both CCS and cranio–lenticulo–sutural dysplasia (CLSD; OMIM #607812) caused by a heterozygous mutation in the COPII component SEC24D. CCS2 was reported in a German boy and two foetuses from terminated pregnancies with severe bone fragility (Garbes et al., 2015).

1.4.2 Ehlers-Danlos Syndrome (EDS)

EDS is a heritable connective tissue disorder affecting primarily skin, joints, ligaments, blood vessels, and internal organs (De Paepe and Malfait, 2012). Characteristic phenotypes of EDS are hypermobility of joints, hyperflexible skin and dislocations.
Most cases of EDS result from mutations in genes encoding fibrillar collagens or enzymes responsible for their biosynthesis or post-translational processing (e.g. ADAMTS-2; leading to impairment of N-propeptide cleavage (Van Damme et al., 2016)). In addition to the 1997 Villefranche classification of EDS into six subtypes (Beighton et al., 1998), work in recent years has discovered that the biosynthesis of other ECM molecules, signalling pathways and intracellular transport can further affect EDS (De Paepe and Malfait, 2012). Since 2017 officially 13 different types of EDS have been recognised (Malfait et al., 2017).

1.4.3 Fibrosis

Fibrosis is characterised by the excessive deposition of unstructured ECM, resulting in abundant scarring tissue and stiffening, subsequently effecting the function of various organs like the lung, liver, kidney, heart or skin (Weiskirchen et al., 2019). Because of this drastic impact on vital organs, fibrosis is associated to approximately 45% all death (including cancer and cardiovascular diseases) (Wynn, 2007).

The mechanism of the fibrogenic response follows the same general route as the wound-healing response. Tissue injury, however, can cause excessive scarring and an increased number of matrix-depositing cells, or excessive activation of the former leading to a disorganised organ architecture (Pakshir and Hinz, 2018). An inflammatory response is activated upon cell death caused by injury and immune cells like macrophages migrate to the site of injury. The cytokines and chemokines released by the invading immune cells triggers the activation of mesenchymal cells, which enhance the immune signal and secrete large amounts of ECM.

To gain an insight into the mechanisms underlying procollagen transport from the ER to the Golgi, we must first look at what mechanisms take place for trafficking conventional smaller cargoes.

1.5 The early secretory pathway and procollagen transport

1.5.1 The coat protein complex type II (COPII)

The COPII coat enables the incorporation and transport of most cargo proteins from the ER to the ERGIC/Golgi. The human COPII machinery consists of Sec12, Sec13 and several isoforms of the other components Sar1(A/B), Sec16(A/B), Sec23(A/B), Sec24(A/B/C/D) and Sec31(A/B) as listed in Fig. 6. The core COPII components are well conserved amongst different species (Schlacht and Dacks, 2015). The multiple COPII isoforms most likely provide additional complexity to enable efficient trafficking of diverse cargo (Chung et al., 2016; Zeng et al., 2015).

COPII assembly

COPII assembly and regulation of ERES is dependent on many factors and pathways. To enable an easier overview, this first section concentrates on only a subset of proteins involved during this process and a more in-depth view follows.

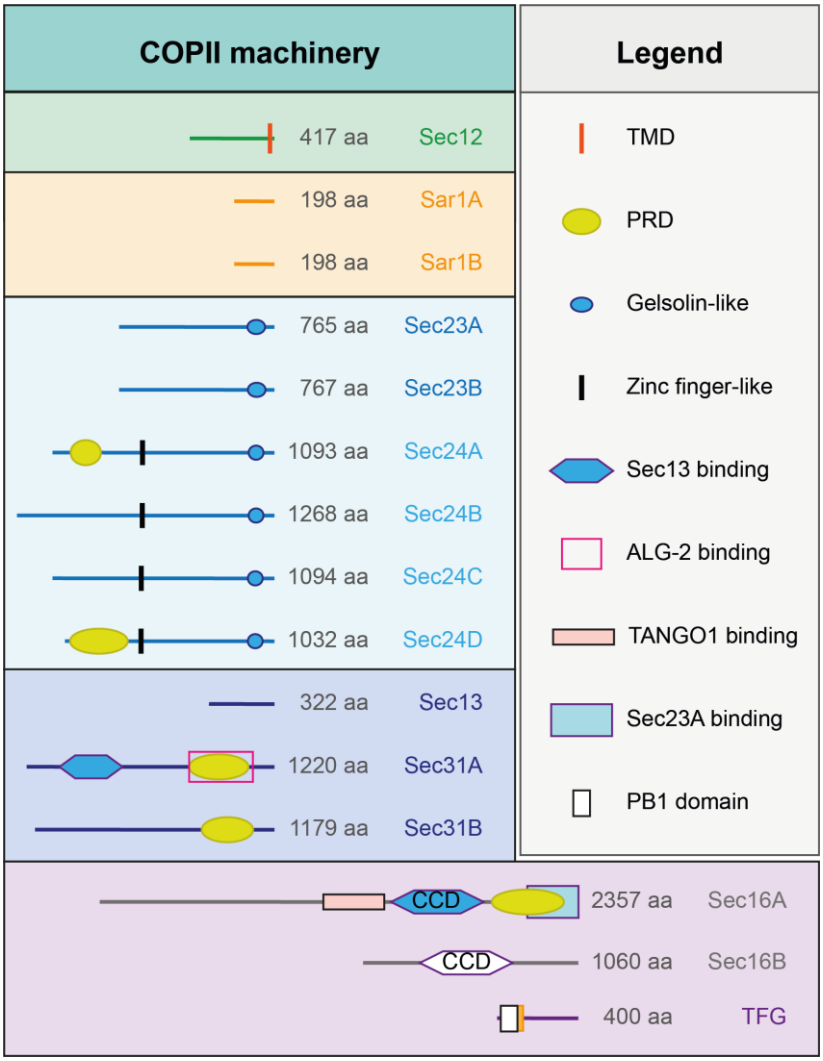


Fig. 6: Schematic domain structure of COPII and other ERES proteins.

Motifs shown are those predicted by smart.embl.de and/or denoted on UniProt.org. Motifs shown include transmembrane domains (TMDs), proline-rich domains (PRDs), Phox and Bem1 (PB1) domains and coiled-coil dimerization (CCD) domains. (McCaughey and Stephens, 2019)

In metazoans COPII assembly is initiated by the transmembrane protein Sec12 binding to Sec16 (Ivan et al., 2008; Montegna et al., 2012) at specific sites on the ER membrane, tER (Bannykh et al., 1996). The tER is marked by the scaffolding protein Sec16 (Connerly et al., 2005; Hughes et al., 2009; Watson et al., 2006). The tER, together with COPII structures that form at these sites and bud off, are referred to as ERES. Sec12 acts as a guanine nucleotide exchange factor (GEF) and concentrates and activates cytosolic Sar1 by catalysing guanosine diphosphate (GDP)-to-GTP exchange (Barlowe and Schekman, 1993; Montegna et al., 2012; Nakano and Muramatsu, 1989), which in turn results in a conformational change of Sar1 inserting

an amphipathic helix into the tER-membrane contributing to initial membrane curvature (Fig. 7A) (Farsad and De Camilli, 2003; Huang et al., 2001). Active Sar1-GTP then recruits the inner layer COPII components Sec23 and Sec24 which assemble in form of a bowtie structure (Bi et al., 2002; Matsuoka et al., 1998). Sec24 has multiple cargo binding sites and is the main player in cargo recruitment at ERES (Aridor et al., 1998; Miller et al., 2002; Miller et al., 2003). Subsequently, the inner coat drives recruitment of the outer coat proteins Sec13-Sec31 (Fig. 7B) (Barlowe et al., 1994; Bi et al., 2007).

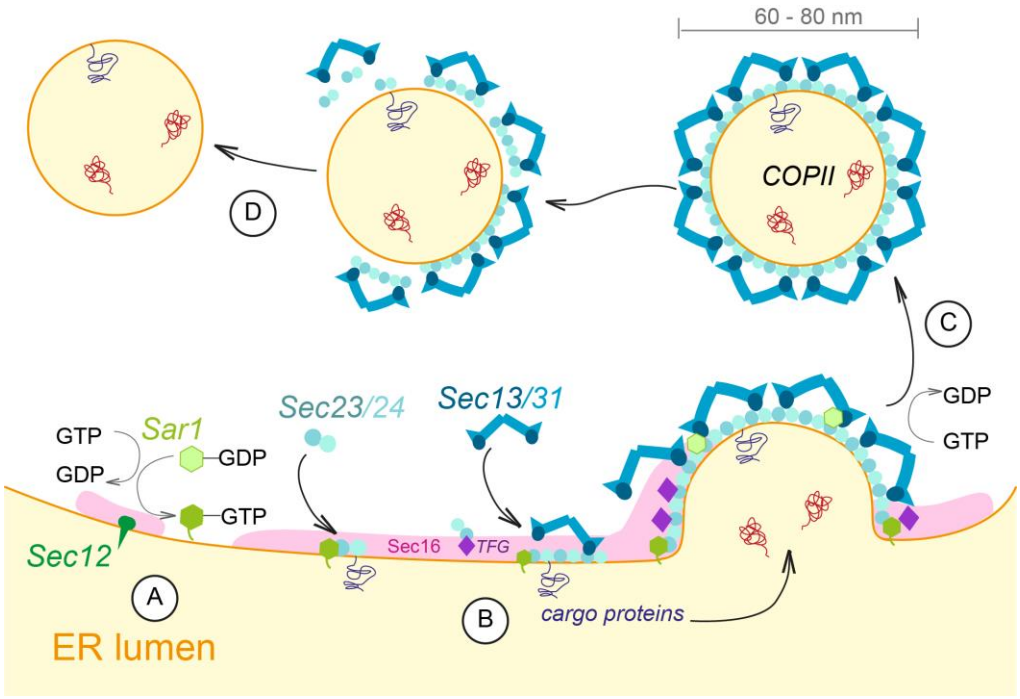


Fig. 7 Simplified overview of COPII assembly.

COPII assembles at an ERES initiated by Sec12 and Sar1 (A) leading to the recruitment of inner layer (Sec23–Sec24) and outer layer (Sec13–Sec31) COPII components with the assistance of Sec16 and Trk-fused gene (TFG), which act to spatially organise the coat proteins (B). After cargo concentration and vesicle growth, Sar1-GTP hydrolysis triggers release from the ER (C). The resulting COPII vesicles have a diameter of 60–80 nm. Shortly after budding, the vesicle undergoes uncoating (D) prior to fusion with the ERGIC/Golgi compartments. (McCaughey and Stephens, 2018; McCaughey and Stephens, 2019)

This step is assisted by ERES organising proteins Trk fused gene (TFG) (Johnson et al., 2015; McCaughey et al., 2016; Witte et al., 2011) and Sec16 (Espenshade et al., 1995; Miller and Schekman, 2013; Whittle and Schwartz, 2010). Sec13-Sec31 assemble in form of rod-shaped heterotetramers composed of two Sec13-Sec31 dimers (Stagg et al., 2008). The outer coat also further induces inner coat assembly (Hutchings and Zanetti, 2019). Increase of membrane curvature results in the formation of a COPII coated bud. Sar1-GTP hereby accumulates at the base of the budding structure and undergoes conformational changes induced by Sec23 activating the GTPase function of Sar1 (Nakano et al., 1994; Oka and Nakano, 1994; Yoshihisa et al., 1993). This step is further enhanced by Sec13 (Antonny et al., 2001; Bi et al., 2002).

This triggers the COPII vesicle with incorporated cargo proteins to undergo fission and release from the ER (Fig. 7C) (Bielli et al., 2005; Fromme et al., 2007; Yoshihisa et al., 1993). The resulting 60 – 80 nm COPII vesicle (Barlowe et al., 1994) needs to undergo uncoating prior to fusion with ERGIC membranes and transfer of cargo proteins (Fig. 7D) (Forster et al., 2006; Townley and Stephens, 2009; Zanetti et al., 2011). In addition, Sec24 has the ability to recruit and sort different soluble NSF (N-ethyl-maleimide-sensitive fusion protein) attachment protein receptors (abbreviated together as SNAREs) to allow for preparation of vesicles for subsequent homotypic fusion events (Adolf et al., 2016).

Some organisms lack Sec12 or Sec16 (Schlacht and Dacks, 2015) and alternative mechanisms for Sar1 activation might therefore exist. A possible alternative pathway for anterograde transport from the ER could be facilitated by COPI, as suggested from COPII-independent transport shown in Sar1-depleted cells (Cutrona et al., 2013).

While the minimal machinery required for *in vitro* formation of COPII vesicles in the presence of membranes consists of Sar1, Sec23-24 and Sec13-31 (Matsuoka et al., 1998), many more factors are required for this process *in vivo.*

1.5.2 Efficient ERES organisation and control of COPII-dependent budding

Cargo load can influence ERES organisation

There is evidence that Sec23 isoforms might be involved in differential cargo sorting, as seen from mouse genetics in Sec23A deficient mice (Zhu et al., 2015). Some proteins are sorted with the help of receptors like Erv14p that recognise transmembrane domains (TMDs) (Herzig et al., 2012). In combination with other sorting motifs and facilitated by coincidence detection, cargo receptors facilitate efficient recognition and ER-export (Pagant et al., 2015).

ERES structure appears to be tightly linked to its functionality and perturbations effecting transport processes e.g. triggered by the loss of the cargo receptor SURF4 (Erv29p in yeast) can have drastic effects on ERES organisation (like reduction in size) (Saegusa et al., 2018). Overall cargo load also determines abundance and size of ERES, in that e.g. a short-term increase in cargo induces the size ERES and abundance of COPII (Farhan et al., 2008; Forster et al., 2006; Guo and Linstedt, 2006). Other data suggest that the presence of larger cargoes may also have an increasing effect on COPII vesicle size (Gomez-Navarro et al., 2020). Recent work has identified a signalling pathway called autoregulation of ER export (AREX), which is activated by Sec24 recognising specific folded cargo including VSVG and type I procollagen (Subramanian et al., 2019). In turn, AREX up-regulates cargo export and down-regulates protein synthesis to reduce the amount of potentially active, folded proteins mislocalised in the ER, which might otherwise cause diseases like cancer (Choudhary et al., 2009; Subramanian et al., 2019). In this study, type I procollagen and VSVG were

used in temperature block experiments (shifting from 40 °C for 3 h to 32 °C) to accumulate both cargo proteins in an unfolded state in the ER, followed by temperature shift induced folding and cargo release (Subramanian et al., 2019).

Efficient ERES organisation

Efficient ERES organisation (especially in multicellular organisms and a requisite for procollagen export) requires further factors. Although many proteins involved in this process have been identified, the exact spatial and temporal organisation of ERES still requires extensive research. An overview including more factors involved in ERES regulation is shown in Fig. 8.

The ubiquitously expressed transmembrane protein transport and Golgi organisation 1 (TANGO1) binds and acts together with other proteins of the Mia (*melanoma inhibitory antigen*) family including cutaneous T-cell lymphoma-associated antigen 5 (cTAGE5) and TANGO1-like (TALI) at ERES (Saito et al., 2011). The different roles of alternative splice forms of TANGO1 (TANGO1S and L) will be discussed in a later section (1.5.4). TANGO1L is proposed to act as a cargo receptor for procollagen through either direct binding (type VII procollagen (Saito et al., 2009)) or indirect binding (via HSP47 to all kinds of procollagens and potentially other matrix proteins (Ishikawa et al., 2016)) via its ER-luminal SH3 domain (Fig. 8A). TANGO1L, in contrast to other cargo receptors remains ER-resident (Saito et al., 2009). Recent data suggests that Sec16 is recruited to ERES by TANGO1, since depletion of TANGO1S and TANGO1L together results in a localisation of Sec16A away from Sec31A also effecting transport of VSVG (Maeda et al., 2017; Maeda et al., 2016). This was not observed when depleting only cTAGE5 (Maeda et al., 2017).

The localisation of Sec12 to ERES (but not its activity as GEF) requires direct binding to cTAGE5 (Saito et al., 2014). TANGO1 furthermore interacts with SLY1, which is required for fusion of membranes (Nogueira et al., 2014). SLY1 interacts with the ER-specific SNARE Syntaxin 18 to enable fusion with ERGIC membranes (Nogueira et al., 2014) with the help of the NRZ tethering complex (Fig. 8B) (Raote et al., 2018). The latter may enable the formation of large COPII coated structures required for efficient packaging of large cargo proteins like procollagen (Maeda et al., 2016; Raote et al., 2018; Rios-Barrera et al., 2017; Saito et al., 2011; Santos et al., 2016; Santos et al., 2015).

TANGO1 recruits Sedlin, which is a component of the transport protein particle (TRAPP) tethering complex, to ERES, where it is involved in the GTP-cycle of Sar1 (Venditti et al., 2012). Some data shows that Sec16 can attenuate GTP-hydrolysis by Sar1, promoting delayed vesicle fission allowing cargo to be fully incorporated first (Kung et al., 2012; Yorimitsu and Sato, 2012). Efficient COPII assembly and vesicle growth is linked to the Sar1-GTP-cycling, as incubation of cells with a non-hydrolysable

GTP (Sar1-H79G) in semi-intact cells, results in tubular extrusions from the ER (Aridor et al., 2001) coated with COPII (Zanetti et al., 2013).

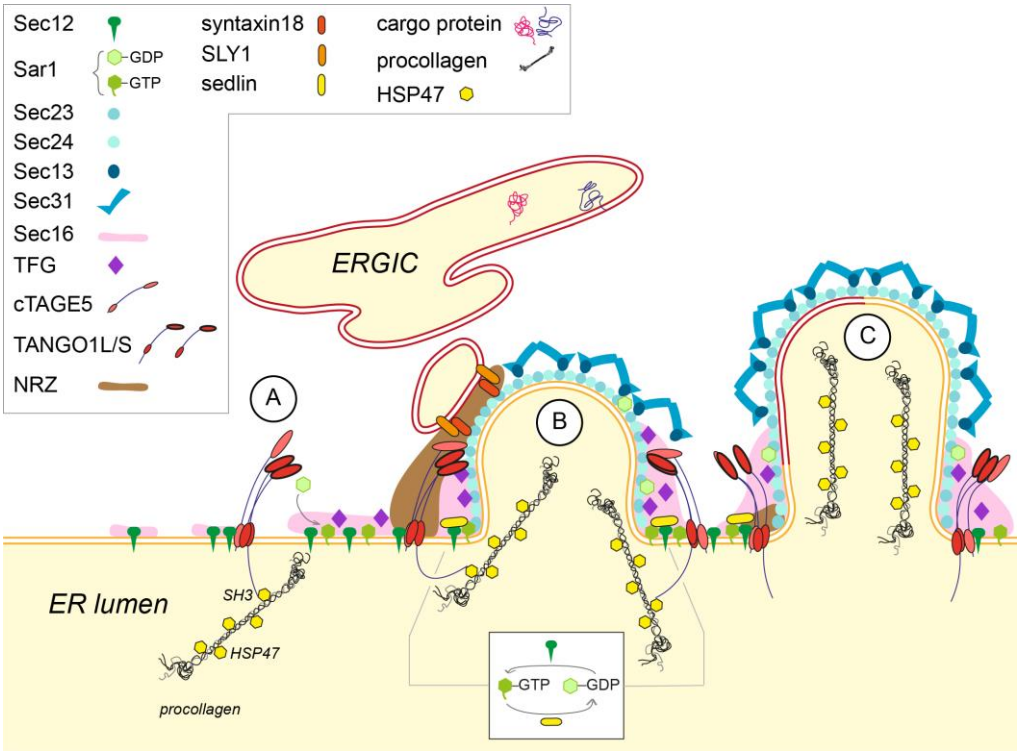


Fig. 8: The role of TANGO1S/L and cTAGE5 in efficient ERES organisation.

Sec12 interacts with cTAGE5 likely to concentrate it at ERES. TANGO1L binds HSP47 through its ER-luminal SH3-domain, directing procollagen to ERES (A). Efficient Sar1–GTP/GDP cycling is facilitated by Sedlin and Sec12. The COPII pre-budding complex is further regulated by TANGO1S/L together with cTAGE5 that bind to the COPII inner layer (B). The NRZ tethering complex is recruited by TANGO1-cTAGE5 and recruits ERGIC membranes to the forming COPII complex. TANGO1-cTAGE5 assemble the fusion machinery SLY1 and syntaxin18 that drive incorporation of ERGIC membranes into the vesicle to allow procollagen incorporation into the carrier (C). (McCaughey and Stephens, 2018)

Sec16 furthermore assists in recruiting other COPII components like Sec13, Sec23 and TFG (Espenshade et al., 1995; Hughes et al., 2009; Hughes and Stephens, 2010; Iwasaki et al., 2017; Whittle and Schwartz, 2010).

TFG localises to ERES and binds Sec23 (Hanna et al., 2017), as well as Sec16 (Witte et al., 2011) and links the ubiquitination pathway to COPII vesicles through binding of ALG-2 (Kanadome et al., 2017). The same interface that promotes binding of TANGO1, cTAGE5 and TFG on Sec23 is also required for binding of the outer COPII coat (Hanna et al., 2017; Saito et al., 2009; Saito et al., 2011). Thereby, TANGO1 can potentially compete with the outer layer preventing premature completion of coat formation (Ma and Goldberg, 2016).

TFG acts as a tether for partially coated vesicles to retain them in proximity to ERGIC elements (Hanna et al., 2018; Hanna et al., 2017) and to organise ERES into larger structures, which is potentially required for facilitating export of large cargoes like procollagen (Johnson et al., 2015; McCaughey et al., 2016). The close apposition of ERGIC with COPII by tethering via TFG might be further required for expansion of growing COPII buds via fusion with ERGIC membranes facilitated by TANGO1 (Santos et al., 2015). TFG depletion slows down secretion and can disable it completely for specific cargoes, while not inhibiting overall ER-to-Golgi transport (Johnson et al., 2015).

Ubiquitination plays a role in COPII regulation

Defects in collagen transport caused by Sec23A depletion could be rescued by the ubiquitin ligase Hrd1, which causes enrichment of type I collagen in the ECM and depletion of Hrd1 also leads to a decrease of collagen (Li et al., 2014). Furthermore, overexpression of the ubiquitin ligase Cullin3 (CUL3) adaptor Kelch-like protein 12 (KLHL12) leads to an enrichment of type I collagen in the ECM in fibroblasts (Jin et al., 2012). Depletion of CUL3 complex subunits PEF1 and ALG-2 in cells overexpressing type I procollagen causes its retention in the ER (McGourty et al., 2016). In addition, depletion of CUL3-KLHL12 leads to accumulation of type IV procollagen in the ER (Jin et al., 2012). Interestingly, ALG-2 promotes assembly and localisation of TFG to ERES (Kanadome et al., 2017). Furthermore, another PEF protein, peflin, seems to be a negative regulator for ALG-2 localisation to ERES, as well as for ALG-2/Sec31A interaction (Rayl et al., 2016). Data from this publication also showed that depletion of peflin leads to faster ER-Golgi transport of the VSVG (Rayl et al., 2016), indicating a general role of ALG-2 and the CUL3-KLHL12 complex in COPII regulation. The CUL3-KLHL12 complex is known to monoubiquitylate Sec31A (Jin et al., 2012) and since ALG-2 is regulated by calcium influx (Shibata et al., 2010; Shibata et al., 2007; Yamasaki et al., 2006), the CUL3 complex therefore plays an important role in calcium-dependent COPII regulation. Monoubiquitination of Sec31A increases the size of COPII vesicles (while de-ubiquitination reduces it) and cells overexpressing KLHL12 showed large structures positive for Sec31A and KLHL12 (Gorur et al., 2017; Jin et al., 2012; Kawaguchi et al., 2018; McGourty et al., 2016; Yuan et al., 2017).

1.5.3 Perturbations in the secretory pathway affecting procollagen transport

Diverse data from inhibitor, depletion and knockout (KO) experiments from both cells and animal models show that ER-to-Golgi transport of procollagen and other unconventional cargo proteins is COPII dependent (Stephens and Pepperkok, 2002). An overview is given in Table 2.

Mutations in Sec23A and Sec23B result in an accumulation of collagen in the ER

Inefficient COPII assembly caused by human mutations in Sec23A leads to defects in packaging type I and II procollagen, resulting in its accumulation in the ER. This phenotype is seen in patients with CLSD as well as in zebrafish chondrocytes (Boyadjiev et al., 2006; Boyadjiev et al., 2011; Fromme et al., 2007; Lang et al., 2006). Impaired procollagen transport is most likely caused by premature vesicle scission linked to the GTPase activity of Sar1 (Kim et al., 2012; Zanetti et al., 2011). Indeed, the GTPase cycle within the COPII mechanism is consistently linked to defects in ECM packaging. Collagen trafficking and bone formation in zebrafish and human cells are also impaired when Sec23A function is compromised by inhibition of O-GlcNAc(O-linked β-N-acetylglucosamine)-ylation of specific sites on Sec23A (Cox et al., 2018). Sec23B depletion in zebrafish, as well as human embryonic kidney cells, results in accumulation of collagen in the ER and similar phenotypes to those described for Sec23A depletion and mutation (Cox et al., 2018; Lang et al., 2006).

Sec24D is required for type II procollagen secretion and efficient export of type I procollagen

Sec24D seems to be specifically required for type II procollagen secretion. Sec24D mutations in zebrafish lead to a decrease of extracellular type II collagen and matrilin (a small ECM protein), while not affecting trafficking of other cargo proteins like β-1-integrin, fibronectin or cadherin (Sarmah et al., 2010). A Sec24D mutation in human fibroblasts from a patient displaying phenotypes of CLSD (as known from Sec23A mutations) and syndromic OI show retention of type I procollagen in the ER, as well as significantly thicker ER-tubules by EM (Garbes et al., 2015). These phenotypes also agree with observations in fish (*vbi* medaka mutant (Ohisa et al., 2010) and *bulldog* zebrafish mutant (Sarmah et al., 2010))

Sec13 is required for proteoglycan and collagen deposition into the ECM

Depletion of Sec13 causes concomitant loss of Sec31, as well as a clustering of Sec23-Sec24 complexes with accumulated cargo in close apposition to the nucleus (Townley et al., 2008). Interestingly, the transport marker tsO45-G-YFP remained unaffected, while type I procollagen secretion in human fibroblasts and proteoglycan deposition in zebrafish was impaired. These findings imply that proper interaction between the inner and outer COPII layers is essential for secretion of large cargo, while being highly dependent on the abundance of Sec13/31. (Townley et al., 2008). Skeletal abnormalities caused by Sec13 depletion mirror those of the mutant zebrafish *crusher* (Lang et al., 2006), as well as patients with CLSD.

Mutations in Sar1B lead to Chylomicron retention disease

Sar1B is considered to have an impact on COPII cage flexibility (Miller and Schekman, 2013; Zanetti et al., 2013) and mutations lead to chylomicron retention in human enterocytes causing Anderson's disease (Jones et al., 2003).

Table 2: Factors influencing unconventional cargo transport and secretion.

(Expanded from (Unlu et al., 2014). (McCaughey and Stephens, 2019))

Protein[1]	Cargo[2]	Defect[3]	In[4]	Tissue/cell type	Reference
Sec23A	COL1	Acc. in ER	H	Fibroblasts	(Fromme et al., 2007)
	COL2	Acc. in ER	Z	Chondrocytes	(Boyadjiev et al., 2006; Cox et al., 2018; Lang et al., 2006)
Sec23B	COL1	Acc. in ER	H	Kidney cells	(Cox et al., 2018)
	COL2	Acc. in ER	Z	Chondrocytes	(Lang et al., 2006)
Sec24D	COL1	Acc. in ER	H	Fibroblasts	(Garbes et al., 2015)
	COL2	Acc. in ER	Z	Chondrocytes	(Sarmah et al., 2010)
	COL2	Intracellular acc.	Mk	Chondrocytes	(Ohisa et al., 2010)
			Mk	Notochord	
			Mk	Myoseptum	
Sec13	COL1	Secretion and deposition	H	Fibroblasts	(Townley et al., 2008)
			H	Fibroblasts	(Niu et al., 2012)
	COL2	Acc. in ER	Z	Chondrocytes	
Sar1B	Chylo-microns and VLDL	Retention in membrane-bound compartments	H	Enterocytes	(Dannoura et al., 1999)
CREB3L1	COL1	Secretion and deposition	H, Ms	Osteoblasts	(Guillemyn et al., 2019; Murakami et al., 2009; Symoens et al., 2013)
CREB3L2	COL2	Secretion and deposition	Z, Mk	Chondrocytes	(Ishikawa et al., 2017; Melville et al., 2011)
TFG	COL1	Decreased in ECM	H	Fibroblasts	(McCaughey et al., 2016)
TANGO1	COL7	Acc. in ER	H	Fibroblasts	(Saito et al., 2009)
	COL1		Ms		(Wilson et al., 2011)
	COL2	Intracellular acc.	Ms	Chondrocytes	
	COL3	Intracellular acc.	Ms	Chondrocytes	
	COL4	Intracellular acc.	Ms	Endothelial cells	
	COL7	Intracellular acc.	Ms	Embryonic	
	COL9	Intracellular acc.	Ms	Fibroblasts	
	COL12	Intracellular acc.	H	Epithelial cells	(Santos et al., 2016)
(+TALI)	Chylo-micron	Intracellular acc.	H	Epithelial cells	
cTAGE5	COL7	acc. in ER	H	A431 cells	(Saito et al., 2011)
Sedlin	COL1	acc. in ER	H	Fibroblasts	(Venditti et al., 2012)
	COL2	acc. in ER	H	Chondrocytes	

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

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

[3]Transport defects, abbreviated: accumulation (acc.)

[4]In organisms, abbreviated: Human (H), Zebrafish (Z), Medaka (Mk), Mouse (Ms).

Chylomicrons seem to be transported from the ER to Golgi in a unique way by utilising so-called prechylomicron transport vesicles (Black, 2007). These bud from the ER in a L-FABP-dependent manner and trafficking and fusion with the *cis*-Golgi is COPII-dependent (Neeli et al., 2007; Siddiqi et al., 2003).

Further COPII-associated proteins that have an impact on large protein trafficking

Besides the core COPII machinery, many factors have been identified that play crucial roles in COPII regulation, as well as enabling transport of large secretory cargo proteins. Mutations in the transcription factor CREB3L2, which regulates Sec23A, Sec23B and Sec24D expression, does not affect trafficking of small cargoes or laminin in the zebrafish CREB3L2 mutant *feelgood*, but leads to impaired collagen deposition into the ECM resulting in craniofacial and skeletal defects (Melville et al., 2011). Depletion of TFG in human fibroblasts leads to a decrease in extracellular type I collagen compared to control samples (McCaughey et al., 2016) and KO of the golgin giantin causes osteochondrodysplasia in rats (Katayama et al., 2011).

This evidence supports the idea that procollagen and chylomicron transport from the ER to Golgi occurs in a COPII-dependent manner. Little is known about the other large secretory cargo proteins. However, TANGO1 and cTAGE5 are considered to play a key role in ER to Golgi trafficking of large proteins and have drawn more and more attention in recent years.

1.5.4 Mia proteins in trafficking of procollagen and other large secretory cargo

Besides the role of TANGO1 in ERES organisation it has been implied to facilitate ER-to-Golgi transport of procollagen. The role of TANGO1 as cargo receptor for procollagen was first identified through binding of type VII procollagen (Saito et al., 2009). Furthermore, knockdown (KD) of Sec12 or cTAGE5 results in impaired transport of type VII procollagen (Saito et al., 2011; Saito et al., 2014).

The Mia family

TANGO1 belongs to the Mia protein family shown in Fig. 9. In most cases the protein referred to as TANGO1 actually refers to TANGO1L (expressed by the Mia3 gene), which represents the biggest member of the Mia family resulting in the long form of TANGO1 with 1907 aa.

An alternative splice variant, the short form of TANGO1L, TANGO1S, has been recognised more recently (Maeda et al., 2016; Wilson et al., 2011). This protein is 785 aa long and shares most of the cytosolic domain with TANGO1L but lacks the ER luminal domain (including the binding site for HSP47/procollagen) (Wilson et al., 2011). While TANGO1S shares its second TMD with TANGO1L, the first one differs in its amino acid composition. While not predicted by algorithms, TANGO1L has two membrane-associated domains of which the first one is proposed to rather be an

intramembrane domain (which might act as a lipid diffusion barrier (Raote et al., 2020)), opposed to reaching into the cytosol and has been proven experimentally (Saito et al., 2009). Intriguingly, only the cytosolic proline-rich domain (PRD) domain of TANGO1 is required for type VII procollagen transport, since TANGO1S-cTAGE5 is able to facilitate this step in absence of TANGO1L with TANGO1S overexpressed (Maeda et al., 2016).

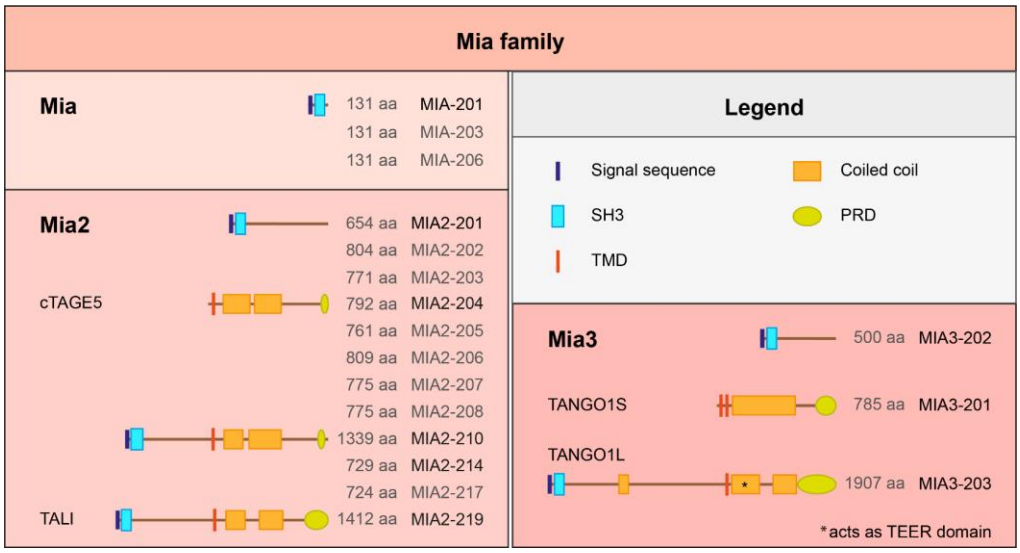


Fig. 9: Motifs in TANGO1-related isoforms.

Motifs shown are those predicted by smart.embl.de and or denoted on UniProt.org. Motifs shown include TMDs, proline-rich domains (PRDs) and Tether of ERGIC and ER (TEER) domains. (McCaughey and Stephens, 2019)

Analysis of Ensembl (www.ensembl.org) data shows that Mia, Mia2, and Mia3 are each encoded by a single gene in humans and multiple splice forms of the three Mia genes exist (Fig. 9). This results in similar TANGO-related peptides that could potentially provide the specificity to drive type-controlled recruitment of various procollagens and or large cargoes.

The smallest member of the Mia family is composed of 131 aa and due to lack of an ER-retention signal might be secreted (Blesch et al., 1994). Mia does, however, have a SH3 domain and might therefore influence matrix assembly, possibly at the stage of the ER. Increased Mia mRNA is associated with malignant lesions in melanocytes (Bosserhoff et al., 1996). Mia-KO mice show altered collagen fibre density, diameter and organisation (Moser et al., 2002) and an increase in the ability of regenerating cartilage (Schmid et al., 2010).

Both TALI and cTAGE5, amongst other proteins, are translated from alternative splice variants of mRNA encoded by the Mia2 gene. This gene can be traced back to a gene duplication in metazoans and its splice variants range from 654 to 1412 aa (TALI). TALI is common in tissues susceptible to fibrosis like lung and liver tissues and not

ubiquitously expressed (Santos et al., 2016). TALI interacts together with TANGO1 to bind and drive ER export of apolipoprotein B in pre-chylomicrons and lipoproteins (Santos et al., 2016).

The role of TANGO1S/L and cTAGE5 in procollagen transport

While the importance for TANGO1 in transport of type VII procollagen is undeniable, the role of TANGO1 in type I procollagen transport and other procollagens or large secretory cargoes has been less clear.

KD of TANGO1L in Saos2 cells leads to retention of type VII procollagen in ER, while secretion levels of type I procollagen remained unchanged (Saito et al., 2009). Overexpression of TANGO1L, however, did increase the speed of type II procollagen exit from the ER (Venditti et al., 2012). SLY1-KO impairs type VII procollagen secretion, while remaining redundant for general COPI and COPII function, including secretion of type I procollagen (Nogueira et al., 2014). Another large secretory protein (~800 nm long (Wilkin et al., 2000)) specific to *Drosophila* called Dumpy also requires TANGO1L for secretion (Rios-Barrera et al., 2017).

Interestingly, while transport of VSVG remained unaffected, mutations in the TRAPP complex (which interacts with TANGO1 through Sedlin) resembled phenotypes of type I and II procollagen retention in the ER resulting in spondyloepiphyseal dysplasia tarda in patients (Venditti et al., 2012).

Compared to these *in vitro* studies, a correlation between TANGO1L and other procollagens was more obvious in data obtained from TANGO1L-KO mice (Wilson et al., 2011). Here, both type IX procollagen and its binding partner Cartilage Oligomeric Matrix Protein (COMP) were retained in the ER (Wilson et al., 2011). TANGO1L-KO mice further suffered from dwarfism of the foetus, perinatal lethality caused by drastic impairment of osteogenesis and chondrodysplasia most likely caused by insufficient or impaired secretion of types I - IV and VII procollagens in various tissues (Wilson et al., 2011). This phenotype could not be reproduced in cells, but resembles those caused by HSP47-KO in fibroblasts leading to phenotypes common for OI (Christiansen et al., 2010; Ishida et al., 2006; Ishikawa et al., 2016).

HSP47 is known as an ER-chaperone for procollagens (Nagata et al., 1986), but in addition was shown to bind small leucine-rich proteins like fibromodulin, lumican and decorin, which interact with collagens in the extracellular space (Ishikawa et al., 2018). While type I procollagen transport is impaired in HSP47-KO cells, the secretion of fibronectin, COMP and fibrillin-1 remained unchanged (Ishikawa et al., 2016; Kawasaki et al., 2015)

TANGO1 was also shown to assemble in form of rings at ERES by super resolution microscopy which are positive for Sec31A and contain procollagen (Liu et al., 2017; Raote et al., 2018; Raote et al., 2017). Recently, the first patients with (biallelic) mutations in TANGO1 were described (Lekszas et al., 2020). These suffer from developmental defects affecting teeth, hearing, the skeleton and neurological functions, which can mostly be linked to impaired collagen secretion. Another recent study in zebrafish with cTAGE5 and or TANGO1L forms knocked out, showed tissue specific phenotypes relating to either Mia protein (Clark and Link, 2020). Here, cTAGE5-KO has a drastic impact on lipoprotein transport and increases cell stress in the intestine, neuronal system and muscles (but is dispensable for type II collagen secretion), while TANGO1L-KO shows overall a more severe phenotype including craniofacial defects, impaired collagen secretion, with only limited influence on lipoprotein transport, and cell stress in muscle cells (Clark and Link, 2020). KO of both TANGO1L and cTAGE5 forms in zebrafish leads to severe phenotypes comprising of those observed in single KO in addition to more severe craniofacial defects, multiple heart diseases and early lethality (Clark and Link, 2020). Of note, both the TANGO1L-KO mouse and fish still express TANGO1S.

Despite ample evidence for the importance of TANGO1 at ERES, the precise roles of TANGO1S/L and other Mia proteins in general (as opposed to procollagen-specific) export remains to be further investigated.

1.5.5 Proposed modes of procollagen trafficking from the ER to Golgi

Any understanding of the possible transport modes of large cargo proteins like procollagen from the ER to the ERGIC/Golgi requires consideration of the flexibility of both COPII and cargo proteins, as well as the dynamic nature of the endomembrane system.

The COPII coat can adopt multiple shapes and sizes

Data on COPII coat structures show that both the inner and outer coat interact in a flexible manner to create different shapes and sizes (Hutchings et al., 2018; Hutchings and Zanetti, 2019). *In vitro* studies show that the Sec13-31 complex can self-assemble into cage-like structures and is able to adopt various geometries allowing the formation of COPII cages of 60-120 nm diameter (Fig. 10A) (Stagg et al., 2006; Stagg et al., 2008). In addition, the COPII coat is able to assemble on membranes and induce tubulation as shown by cryo-electron tomography and sub-tomogram averaging (Fig. 10B) (Hutchings et al., 2018; Zanetti et al., 2013). In other cases, COPII has been shown to adopt a variety of shapes, including multi-budded structures resembling “trees” or “pearl-on-a-chain” like shapes (Fig. 10C) (Bacia et al., 2011; Gomez-Navarro et al., 2020). These structures are very evident in *in vitro* studies but have also been detected in cells.

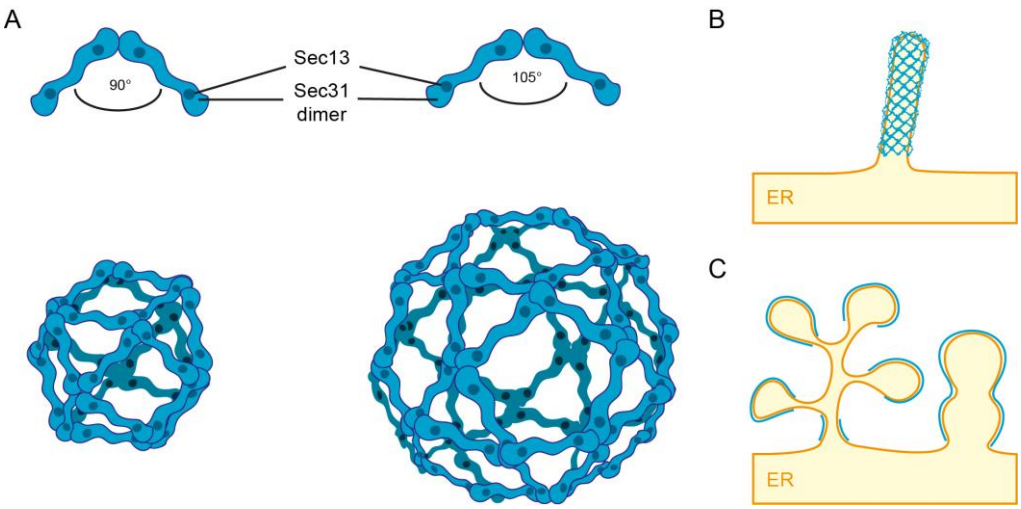


Fig. 10: The COPII coat can arrange into different shapes and sizes.

The angles between Sec13-Sec31 heterotetramers can vary and enable the formation of small cages of 60 nm, as well as large cage structures of 120 nm in diameter (A) (Stagg et al., 2006; Stagg et al., 2008), as well as tubular structures on membranes (B) (Zanetti et al., 2013). A multi-budded structure is shown in C (Bacia et al., 2011; Gomez-Navarro et al., 2020).

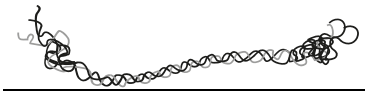
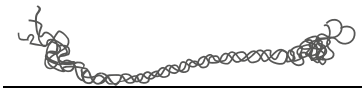
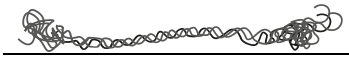
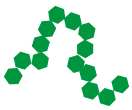
Cargo size and flexibility

Since studies on procollagen transport include other large types of procollagen apart from type I a visual overview over different large secretory cargo proteins is shown in Table 3. The length estimations originate from measurements of rotary shadow images acquired by EM (McCaughey and Stephens, 2019). Of note, despite displaying a higher length estimation, non-fibrillar procollagens such as type IV and VII have an increased flexibility compared to fibrillar procollagens (types I – III, V and XI), due to their interruptions of the triple helical COL domain (Bachinger et al., 1982). Interestingly, instead of expressing fibrillar collagens, *Drosophila* expresses only type IV collagen (known as Viking), a key component of most of the basal membrane underlying epithelial cells and surrounding organs in animals (Jayadev and Sherwood, 2017; Yurchenco, 2011). Basal membrane composition is conserved from fly to human containing laminin, nidogen and perlecan in addition to collagen IV (Hynes and Zhao, 2000). Since most extracellular matrix proteins are heavily glycosylated or otherwise post-translationally modified, it is difficult to estimate their flexibility when exiting the ER.

So far there has been little to no data available on the transport of other larger secretory cargo proteins like laminin, aggrecan, fibronectin or fibrillin (McCaughey and Stephens, 2019).

Table 3: Selection of large secretory cargo proteins.

Green hexagons in type VII procollagen indicate fibronectin type II repeats. (McCaughey and Stephens, 2019)

Secretory cargo protein	Length	Reference
type I procollagen, fibrillar collagen 	~330 nm	(Bachinger et al., 1982)
type II procollagen, fibrillar collagen 	~300 nm	(Rezaei et al., 2018)
type III procollagen, fibrillar collagen 	~300 nm	(Hofmann et al., 1984)
type IV procollagen, networking collagen 	~430 nm	(Bachinger et al., 1982)
type V procollagen, fibrillar collagen 	~330 nm	(Bachinger et al., 1982)
type VII procollagen, anchoring collagen 	~425 nm	(Morris et al., 1986)
type XI procollagen, fibrillar collagen 	>300 nm?	
aggrecan, proteoglycan 	330 nm	(Paulsson et al., 1987)
fibrillin, glycoprotein 	150 nm	(Sakai et al., 1991)
fibronectin, glycoprotein 	160 nm	(Erickson et al., 1981)
laminins, glycoproteins 	110 – 165 nm	(Engel et al., 1981)

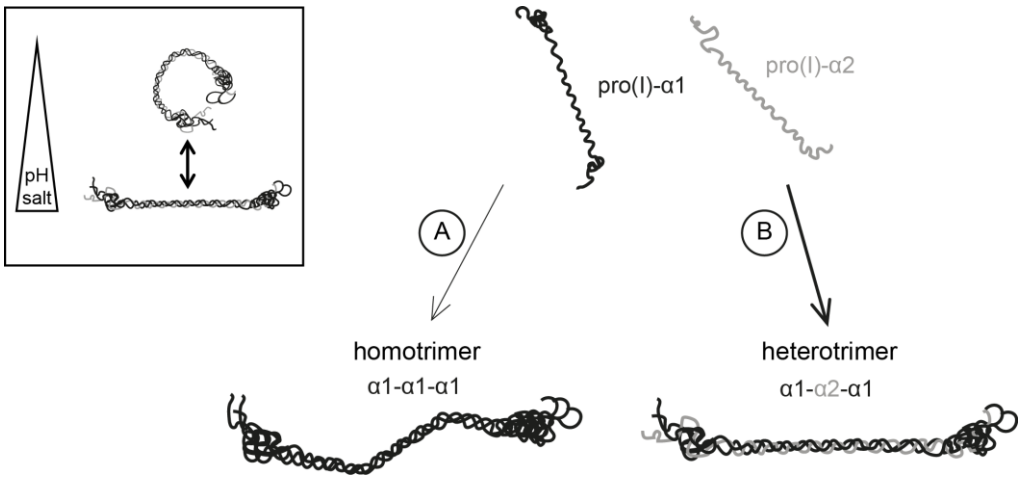


Fig. 11: Flexibility of type I procollagen trimers.

Flexibility of procollagen trimers is both dependent on the salt concentration and pH (Rezaei et al., 2018), as well as the composition, as homotrimers allow for local kinking (A), while heterotrimers retain the same amount of flexibility along the COL domain (B)(Chang et al., 2012).

Homotrimeric type I procollagen from mouse tendon composed of three α1 chains analysed by *in situ* modelling also suggests that the homotrimer is able to locally micro-unfold and promote kinking within the triple helical domain (Fig. 11A), enabling partial rotation and resulting in increased flexibility compared to the heterotrimer (Fig. 11B) (Chang et al., 2012). In addition, atomic force microscopy showed varying curvatures of collagen depending on the salt concentration and pH in many collagen isotypes indicating a higher flexibility for lower salt concentrations and pH (Fig. 11) (Rezaei et al., 2018).

Proposed modes of procollagen transport from the ER

Depending on the estimated flexibility and various experimental data, different modes of procollagen transport from the ER to the Golgi can be envisioned (Fig. 12C-D) opposed to transport of conventional cargo (Fig. 12A).

The structure of triple helical type I procollagen was originally reported to be rigid (Leblond, 1989) and approximately 300 nm in length (Bachinger et al., 1982; Lightfoot et al., 1992). This gave rise to the idea of the need for specialised COPII-dependent carriers, which can facilitate ER-Golgi transport of large cargo proteins (Fig. 12C) (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).

In fact, some studies show potential large carriers. Initially, overexpression of FLAG-KLHL12 was shown to induce Sec31A-positive structures with diameters >300 nm in cells (Jin et al., 2012). Overall COPII labelling in these cells showed few structures and not the usual small punctate distribution (Jin et al., 2012). In another study, these large

COPII *megacarriers* were observed to co-label for type I and II procollagen, while also positive for PEF1 and ALG-2 (McGourty et al., 2016). Further support for the model of large COPII procollagen containing carriers came from *in vitro* reconstitution assays (Gorur et al., 2017; Yuan et al., 2017), as well as CLEM (Gorur et al., 2017). In the latter, cells overexpressing both type I procollagen and KLHL12 were used and micron-sized procollagen and COPII positive structures were shown. Such structures showed minimal movement during live imaging using a C-terminally tagged procollagen (Gorur et al., 2017). Besides ubiquitination induced through KLHL12 overexpression, TANGO1 is also proposed to facilitate the formation of large vesicles through the interaction with the inner coat, as well as in concert with ERGIC-membrane recruitment (Reynolds et al., 2019; Santos et al., 2015) and Sedlin-dependent GTP-cycling to promote transport of procollagens from the ER (Liu et al., 2017; Raote et al., 2018; Raote et al., 2017; Rios-Barrera et al., 2017; Saito et al., 2011; Santos et al., 2015; Venditti et al., 2012; Wilson et al., 2011).

Evidence from more recent data, shows that depletion of CUL3 influences procollagen synthesis in human skin fibroblasts in addition to PERK and IRE1α, which function as sensors for UPR (Kim et al., 2018). Large structures seen in these studies have been observed to still be part of the ER and are destined for degradation, as visualised using fluorescence recovery after photobleaching (FRAP) experiments (Omari et al., 2020; Omari et al., 2018).This suggests that a pathway for procollagen transport independent on the CUL3-KLHL12 complex.

Some recent data proposes that TANGO1 and Sec12 are co-packaged with procollagen into large COPII transport vesicles (Yuan et al., 2018). Potentially, there might be an alternative pathway for this process. There is, however, abundant evidence that Sec12 remains at the ER and is not incorporated into (potentially only small) COPII vesicles (Barlowe et al., 1994; Futai et al., 2004; Iwasaki et al., 2017).

Interestingly, recent studies have challenged the idea of procollagen being a rigid structure and showing evidence in favour of a semi-flexible polymer instead (Rezaei et al., 2018). Therefore, it could be possible that procollagen is flexible at the stage of ER exit and could thus be packaged into more average sized COPII vesicles (Fig. 12B).

Alternatively, considering also the limited space between ER and ERGIC, it has been proposed that procollagen might travers to the ERGIC or Golgi via a direct, tunnel-like pathway (Fig. 12D) (Liu et al., 2017; Malhotra and Erlmann, 2015; Raote and Malhotra, 2019). This idea is based on transport of procollagen through the Golgi stack without leaving the cisternae (Bonfanti et al., 1998), but rather trafficking via maturation of the stack itself. Transient contact sites between the *cis*-Golgi and the ER have been reported, where cargo capture could be observed in form of *cis*-Golgi elements reaching out to ERES, resulting in a model termed ‘hug-and-kiss’ (Hutchings and

Zanetti, 2019; Kurokawa et al., 2014). This model needs to also carefully incorporate how compartment identity of the ER and ERGIC/Golgi is maintained.

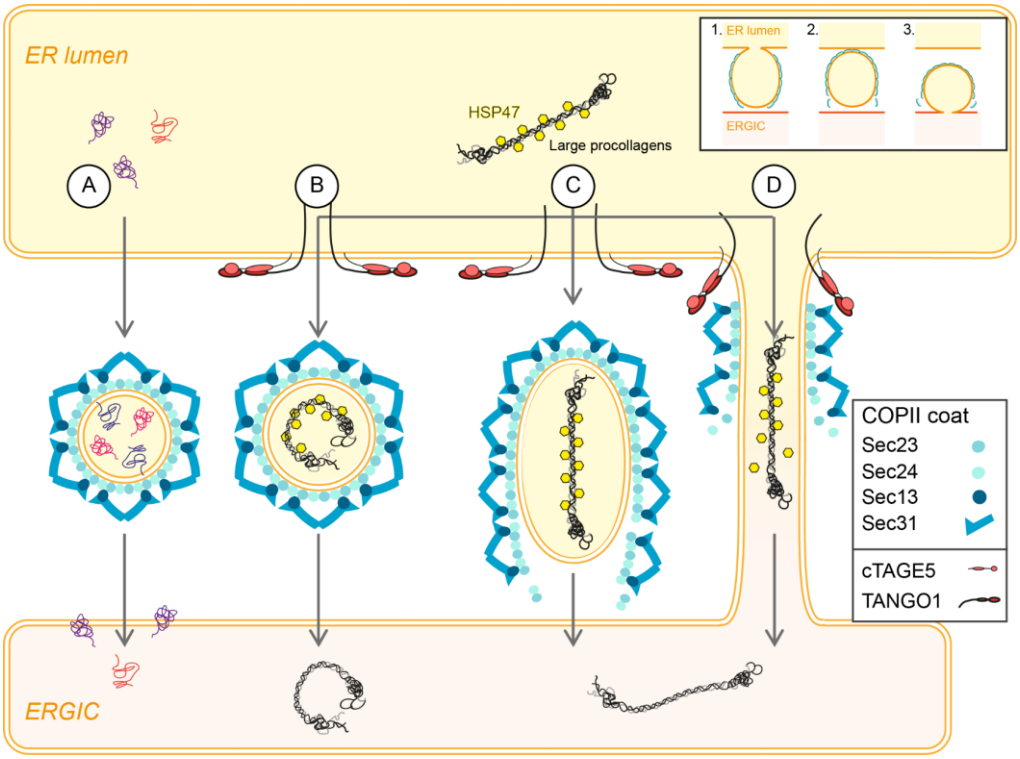


Fig. 12: Modes of ER to Golgi transport of procollagen.

Trimeric type I procollagen in the ER is stabilised by the collagen chaperone HSP47. Conventional small cargo proteins are transported to the ERGIC or Golgi via COPII vesicles with a diameter of about 80 nm (i). Model of flexible procollagen able to be incorporated into more average sized COPII carriers (ii) or rigid procollagen requiring large COPII carriers (iii). Transport to the next compartment facilitated via a hug and kiss like mechanism of the ERGIC/Golgi with the ER. The ER extends to allow passage of procollagen to the ERGIC via fusion/maturation. HSP47 is displayed here to accompany procollagen prior to dissociation from the trimer due to a shift in pH between compartments as currently favoured by literature and TANGO1 facilitating procollagen transport. (McCaughey and Stephens, 2019)

The intracellular conditions necessary for either of the hypothesised transport modes, as well as their specificity for certain large secretory proteins like collagens or small conventional cargo proteins remains to be elucidated. These transport modes might occur independently or in concert with each other, potentially in a situational and or cell type dependent manner.

1.6 Thesis aims

The aim of this project was to elucidate the mechanism behind procollagen transport from the ER to Golgi in mammalian cells. There are currently controversial ideas about the transport routes utilised by procollagens including large carriers and hug-and-kiss mechanisms. In order to investigate this, trafficking of type I procollagen was characterised in a spatial and time-resolved manner combining the use of a streptavidin-binding peptide (SBP) and monomeric green fluorescent protein (mGFP) -tagged pro-α1(I) (GFP-COL) and observations from primary fibroblasts. In addition, I aimed to use these tools and technologies to define a workflow to help classify newly identified collagenopathies from patient-derived cells.

The specific aims were:

1. Characterise the transport of type I procollagen using GFP-COL
2. Establish a cell biological workflow to help classify collagenopathies from patient-derived fibroblasts

Chapter 2: Experimental Methods

All experiments were performed at a room temperature (RT) of 20 °C if not stated otherwise. All reagents were purchased from Sigma-Aldrich (Poole, UK; now owned by Merck) unless noted otherwise. Some parts from this section are adopted from (listed in the appendix) (Balasubramanian et al., 2019; Balasubramanian et al., 2018; McCaughey et al., 2019) and were originally written by myself if not stated otherwise.

2.1 Generation of genetic constructs

Constructs were either homemade or acquired from Addgene (numbers indicated by #). All restriction and modifying enzymes were purchased from New England Biolabs (Hitchin, UK). ER-membrane marker pCytERM_mScarlet-i_N1 (ERM) was a gift from Dorus Gadella (#85068) (Bindels et al., 2017). The pro(I)-α2 fusion constructs were a gift from Sergey Leikin (GFP-COL1A2L, #120160; mApple-COL1A2L, #120161; mApple-COL1A2, #119840)(Omari et al., 2018). Cherry-LC3B was a gift from Jon Lane (University of Bristol, Bristol, UK; the GFP tag from GFP-LC3B (Betin et al., 2013), pLXG-3 SSF FV GFP-LC3B, was replaced by the Lane laboratory for an monomeric Cherry (mCherry) tag). Sar1-H79G was described previously (Aridor et al., 1995)

To allow expression of two different proteins using the same vector, bicistronic vector systems containing an internal ribosome entry site (IRES) element were used. The Str-KDEL-IRES-sialyltransferase-mCherry (ST-Cherry, #110727) construct was made by Ash Evans by using Str-KDEL-IRES-ST-SBP-mCherry (# 65265, a gift from Franck Perez (Institut Curie, Paris, France; (Boncompain et al., 2012)) as a template and exchanged the ST-SBP sequence for a ST only sequence via restriction digest using AscI and SbfI. Similarly, the Str-KDEL-IRES-mannosidaseII-mScarlet-i (ManII-mSc; #117274) was created using Str-KDEL-IRES-mannosidaseII-SBP-mCherry (a gift from Franck Perez, #65253 (Boncompain et al., 2012)) and monomeric Scarlet-i (mSc; a gift from Dorus Gadella, #85044 (Bindels et al., 2017)) were used as a template. The SBP-mCherry tag was replaced with mSc to generate ManII-mSc. To generate Str-KDEL-IRES-mScarlet-i-Sec23A (mScSec23A; #117273) mannosidaseII-SBP-mCherry was first replaced with mS and in a second step Sec23A originating from pLVX-Puro-mRuby-Sec23A (#36158, (Hughes and Stephens, 2010)) was amplified using polymerase chain reaction (PCR) and inserted after Str-KDEL-IRES-mSc. StrKDEL-IRES-mannosidase II-mTagBFP2 (ManII-BFP; with blue fluorescent protein (BFP)) was generated by using ManII-mSc as a template and replacing the mSc with a mTagBFP2 generated as a synthetic gene block (Table 4, Gene block 1) by Integrated DNA Technologies (IDT) via restriction digest using EcoRI and FseI and subsequent HiFi NEBuilder assembly (which is based on the Gibson assembly). To generate procollagen-mScCOL1A2 (mScCOL1A2) a synthesised gene block containing the sequence for human pro(I)-α2 and mSc (Table 4, Gene block 2) was inserted into XhoI and BamHI digested pLVXpuro using HiFi NEBuilder assembly. In a second step, this

intermediate plasmid was digested with XbaI and BamHI and assembled with block 2 (Table 4, Gene block 3) containing the first part of the COL1A2 coding sequence up until the PspXI restriction site. In a final step, the resulting intermediate vector was digested with SbfI and PspXI to insert the remaining COL1A2 from COL1a2-MykDDK (optained from OriGene's Collagen I (COL1A2) (NM_000089) Human Tagged ORF Clone – "COL1A2-MykDDK" (CAT#: RC208484)). The resulting construct has the mSc (flanked by linkers) inserted after the first 79 amino acids of human pro(I)-α2. The generated sequence for COL1a2 is identical with human COL1a2 (accession number BC054498.1) except for a single mutation C3235T, which does not affect the amino acid sequence. Furthermore, the sequence contains a natural variant (ancestral G, see SNP entry rs42524) from NM_000089.3 with C2116G, which corresponds to a proline to alanine change.

Table 4: Gene blocks (synthesised by Integrated DNA Technologies) and final sequences of newly generated constructs.

Gene block 1 used for ManII-BFP: TCCCACCGGTCGCCACCGGaattccATGAGCGAGCTGATTAAGGAGAACATGCACATGAAG CTGTACATGGAGGGCACCGTGGACAACCATCACTTCAAGTGCACATCCGAGGGCGAAGGCA AGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTT CGCCTTCGACATCCTGGCTACTAGCTTCCTCTACGGCAGCAAGACCTTCATCAACCACACC CAGGGCATCCCCGACTTCTTCAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCA CCACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTG CCTCATCTACAACGTCAAGATCAGAGGGGTGAACTTCACATCCAACGGCCCTGTGATGCAG AAGAAAACACTCGGCTGGGAGGCCTTCACCGAGACGCTGTACCCCGCTGACGGCGGCCTGG AAGGCAGAAACGACATGGCCCTGAAGCTCGTGGGCGGGAGCCATCTGATCGCAAACATCAA GACCACATATAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCTGGCGTCTACTATGTG GACTACAGACTGGAAAGAATCAAGGAGGCCAACAACGAGACCTACGTCGAGCAGCACGAGG TGGCAGTGGCCAGATACTGCGACCTCCCTAGCAAACTGGGGCACAAGCTTAATggccggcc tTAAggcctcgagGGCC
Sequence of ManII-BFP [flanked by the remaining vector backbone ranging between the AscI and FseI-sites.]: […]ATGAAGTTAAGTCGCCAGTTCACCGTGTTTGGCAGCGCGATCTTCTGCGTCGTAATCT TCTCACTCTACCTGATGCTGGACAGGGGTCACTTGGACTACCCTCGGGGCCCGCGCCAGGA GGGCTCCTTTCCGCAGGGCCAGCTTTCAATATTGCAAGAAAAGATTGACCATTTGGAGCGT TTGCTCGCTGAGAACAACGAGATTATCTCAAATATCAGAGACTCAGTCATCAACCTGAGCG AGTCTGTGGAGGACGGCCCGCGGGGGTCACCAGGCAACGCCAGCCAAGGCTCCATCCACCT CCACTCGCCACAGTTGGCCCTGCAGGCTGACCCCAGAGACTGTTTGGATCCCACCGGTCGC CACCGGaattccATGAGCGAGCTGATTAAGGAGAACATGCACATGAAGCTGTACATGGAGG GCACCGTGGACAACCATCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGG CACCCAGACCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCCTTCGCCTTCGACATC CTGGCTACTAGCTTCCTCTACGGCAGCAAGACCTTCATCAACCACACCCAGGGCATCCCCG

ACTTCTTCAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGA
CGGGGGCGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAAC
GTCAAGATCAGAGGGGTGAACTTCACATCCAACGGCCCTGTGATGCAGAAGAAAACACTCG
GCTGGGAGGCCTTCACCGAGACGCTGTACCCCGCTGACGGCGGCCTGGAAGGCAGAAACGA
CATGGCCCTGAAGCTCGTGGGCGGGAGCCATCTGATCGCAAACATCAAGACCACATATAGA
TCCAAGAAACCCGCTAAGAACCTCAAGATGCCTGGCGTCTACTATGTGGACTACAGACTGG
AAAGAATCAAGGAGGCCAACAACGAGACCTACGTCGAGCAGCACGAGGTGGCAGTGGCCAG
ATACTGCGACCTCCCTAGCAAACTGGGGCACAAGCTTAATGGGGCCCTCGAGGCCTTAA[…
](including stop codon)

Gene block 2 used for mScCOL1A2 containing the signal sequence-N-propeptide-mSc:

GGACTCAGATCtcgagATGCTCAGCTTTGTGGATACGCGGACTTTGTTGCTGCTTGCAGTA
ACCTTATGCCTAGCAACATGCCAATCTTTACAAGAGGAAACTGTAAGAAAGGGCCCAGCCG
GAGATAGAGGACCACGTGGAGAAAGGGGTCCACCAGGCCCCCCAGGCAGAGATGGTGAAGA
TGGTCCCACAGGCCCTCCTGGTCCACCTGGTCCTCCTGGCCCCCCTGGTCTCGGTGGGAAC
TTTGCTGCTtcagctgctgggcgcgccATGGTGAGCAAGGGCGAGGCAGTGATCAAGGAGT
TCATGCGGTTCAAGGTGCACATGGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGG
CGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGT
GGCCCCCTGCCCTTCTCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAGGGCCT
TCATCAAGCACCCCGCCGACATCCCCGACTACTATAAGCAGTCCTTCCCCGAGGGCTTCAA
GTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACCCAGGACACCTCC
CTGGAGGACGGCACCCTGATCTACAAGGTGAAGCTCCGCGGCACCAACTTCCCTCCTGACG
GCCCCGTAATGCAGAAGAAGACAATGGGCTGGGAAGCGTCCACCGAGCGGTTGTACCCCGA
GGACGGCGTGCTGAAGGGCGACATTAAGATGGCCCTGCGCCTGAAGGACGGCGGCCGCTAC
CTGGCGGACTTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGATGCCCGGCGCCTACA
ACGTCGACCGCAAGTTGGACATCACCTCCCACAACGAGGACTACACCGTGGTGGAACAGTA
CGAACGCTCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACaaagcagcagca
ttaattaaggcagcaggatcccgcgactctagataattctaccgggtaggggaggcgcttt

Gene block 3 used for mScCOL1A2 containing the first part of COL1A2:

TGTACaaagcagcagcattaattaagAACTTTGCTGCTCAGTATGATGGAAAAGGAGTTGG
ACTTGGCCCTGGACCAATGGGCTTAATGGGACCTAGAGGCCCACCTGGTGCAGCTGGAGCC
CCAGGCCCTCAAGGTTTCCAAGGACCTGCTGGTGAGCCTGGTGAACCTGGTCAAACTGGTC
CTGCAGGtgctcgtggtccagcgctcgagcgaTAAgGATCCCGCGACTCTAGATAATTCT

Sequence of mScCOL1A2 [flanked by the remaining pLVXPuro vector backbone] :

[…]**ATG**CTCAGCTTTGTGGATACGCGGACTTTGTTGCTGCTTGCAGTAACCTTATGCCTAG
CAACATGCCAATCTTTACAAGAGGAAACTGTAAGAAAGGGCCCAGCCGGAGATAGAGGACC
ACGTGGAGAAAGGGGTCCACCAGGCCCCCCAGGCAGAGATGGTGAAGATGGTCCCACAGGC
CCTCCTGGTCCACCTGGTCCTCCTGGCCCCCCTGGTCTCGGTGGGAACTTTGCTGCTtcag

ctgctgggcgcgccATGGTGAGCAAGGGCGAGGCAGTGATCAAGGAGTTCATGCGGTTCAA
GGTGCACATGGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGC
CGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCT
TCTCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAGGGCCTTCATCAAGCACCC
CGCCGACATCCCCGACTACTATAAGCAGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTG
ATGAACTTCGAGGACGGCGGCGCCGTGACCGTGACCCAGGACACCTCCCTGGAGGACGGCA
CCCTGATCTACAAGGTGAAGCTCCGCGGCACCAACTTCCCTCCTGACGGCCCCGTAATGCA
GAAGAAGACAATGGGCTGGGAAGCGTCCACCGAGCGGTTGTACCCCGAGGACGGCGTGCTG
AAGGGCGACATTAAGATGGCCCTGCGCCTGAAGGACGGCGGCCGCTACCTGGCGGACTTCA
AGACCACCTACAAGGCCAAGAAGCCCGTGCAGATGCCCGGCGCCTACAACGTCGACCGCAA
GTTGGACATCACCTCCCACAACGAGGACTACACCGTGGTGGAACAGTACGAACGCTCCGAG
GGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACaaagcagcagcattaattaagAACT
TTGCTGCTCAGTATGATGGAAAAGGAGTTGGACTTGGCCCTGGACCAATGGGCTTAATGGG
ACCTAGAGGCCCACCTGGTGCAGCTGGAGCCCCAGGCCCTCAAGGTTTCCAAGGACCTGCT
GGTGAGCCTGGTGAACCTGGTCAAACTGGTCCTGCAGGTGCTCGTGGTCCAGCTGGCCCTC
CTGGCAAGGCTGGTGAAGATGGTCACCCTGGAAAACCCGGACGACCTGGTGAGAGAGGAGT
TGTTGGACCACAGGGTGCTCGTGGTTTCCCTGGAACTCCTGGACTTCCTGGCTTCAAAGGC
ATTAGGGGACACAATGGTCTGGATGGATTGAAGGGACAGCCCGGTGCTCCTGGTGTGAAGG
GTGAACCTGGTGCCCCTGGTGAAAATGGAACTCCAGGTCAAACAGGAGCCCGTGGGCTTCC
TGGTGAGAGAGGACGTGTTGGTGCCCCTGGCCCAGCTGGTGCCCGTGGCAGTGATGGAAGT
GTGGGTCCCGTGGGTCCTGCTGGTCCCATTGGGTCTGCTGGCCCTCCAGGCTTCCCAGGTG
CCCCTGGCCCCAAGGGTGAAATTGGAGCTGTTGGTAACGCTGGTCCTGCTGGTCCCGCCGG
TCCCCGTGGTGAAGTGGGTCTTCCAGGCCTCTCCGGCCCCGTTGGACCTCCTGGTAATCCT
GGAGCAAACGGCCTTACTGGTGCCAAGGGTGCTGCTGGCCTTCCCGGCGTTGCTGGGGCTC
CCGGCCTCCCTGGACCCCGCGGTATTCCTGGCCCTGTTGGTGCTGCCGGTGCTACTGGTGC
CAGAGGACTTGTTGGTGAGCCTGGTCCAGCTGGCTCCAAAGGAGAGAGCGGTAACAAGGGT
GAGCCCGGCTCTGCTGGGCCCCAAGGTCCTCCTGGTCCCAGTGGTGAAGAAGGAAAGAGAG
GCCCTAATGGGGAAGCTGGATCTGCCGGCCCTCCAGGACCTCCTGGGCTGAGAGGTAGTCC
TGGTTCTCGTGGTCTTCCTGGAGCTGATGGCAGAGCTGGCGTCATGGGCCCTCCTGGTAGT
CGTGGTGCAAGTGGCCCTGCTGGAGTCCGAGGACCTAATGGAGATGCTGGTCGCCCTGGGG
AGCCTGGTCTCATGGGACCCAGAGGTCTTCCTGGTTCCCCTGGAAATATCGGCCCCGCTGG
AAAAGAAGGTCCTGTCGGCCTCCCTGGCATCGACGGCAGGCCTGGCCCAATTGGCCCAGCT
GGAGCAAGAGGAGAGCCTGGCAACATTGGATTCCCTGGACCCAAAGGCCCCACTGGTGATC
CTGGCAAAAACGGTGATAAAGGTCATGCTGGTCTTGCTGGTGCTCGGGGTGCTCCAGGTCC
TGATGGAAACAATGGTGCTCAGGGACCTCCTGGACCACAGGGTGTTCAAGGTGGAAAAGGT
GAACAGGGTCCCCCTGGTCCTCCAGGCTTCCAGGGTCTGCCTGGCCCCTCAGGTCCCGCTG
GTGAAGTTGGCAAACCAGGAGAAAGGGGTCTCCATGGTGAGTTTGGTCTCCCTGGTCCTGC
TGGTCCAAGAGGGGAACGCGGTCCCCCAGGTGAGAGTGGTGCTGCCGGTCCTACTGGTCCT
ATTGGAAGCCGAGGTCCTTCTGGACCCCCAGGGCCTGATGGAAACAAGGGTGAACCTGGTG
TGGTTGGTGCTGTGGGCACTGCTGGTCCATCTGGTCCTAGTGGACTCCCAGGAGAGAGGGG
TGCTGCTGGCATACCTGGAGGCAAGGGAGAAAAGGGTGAACCTGGTCTCAGAGGTGAAATT
GGTAACCCTGGCAGAGATGGTGCTCGTGGTGCTCCTGGTGCTGTAGGTGCCCCTGGTCCTG

CTGGAGCCACAGGTGACCGGGGCGAAGCTGGGGCTGCTGGTCCTGCTGGTCCTGCTGGTCC
TCGGGGAAGCCCTGGTGAACGTGGTGAGGTCGGTCCTGCTGGCCCCAATGGATTTGCTGGT
CCTGCTGGTGCTGCTGGTCAACCTGGTGCTAAAGGAGAAAGAGGAGCCAAAGGGCCTAAGG
GTGAAAACGGTGTTGTTGGTCCCACAGGCCCCGTTGGAGCTGCTGGCCCAGCTGGTCCAAA
TGGTCCCCCCGGTCCTGCTGGAAGTCGTGGTGATGGAGGCCCCCCTGGTATGACTGGTTTC
CCTGGTGCTGCTGGACGGACTGGTCCCCCAGGACCCTCTGGTATTTCTGGCCCTCCTGGTC
CCCCTGGTCCTGCTGGGAAAGAAGGGCTTCGTGGTCCTCGTGGTGACCAAGGTCCAGTTGG
CCGAACTGGAGAAGTAGGTGCAGTTGGTCCCCCTGGCTTCGCTGGTGAGAAGGGTCCCTCT
GGAGAGGCTGGTACTGCTGGACCTCCTGGCACTCCAGGTCCTCAGGGTCTTCTTGGTGCTC
CTGGTATTCTGGGTCTCCCTGGCTCGAGAGGTGAACGTGGTCTACCAGGTGTTGCTGGTGC
TGTGGGTGAACCTGGTCCTCTTGGCATTGCCGGCCCTCCTGGGGCCCGTGGTCCTCCTGGT
GCTGTGGGTAGTCCTGGAGTCAACGGTGCTCCTGGTGAAGCTGGTCGTGATGGCAACCCTG
GGAACGATGGTCCCCCAGGTCGCGATGGTCAACCCGGACACAAGGGAGAGCGCGGTTACCC
TGGCAATATTGGTCCCGTTGGTGCTGCAGGTGCACCTGGTCCTCATGGCCCCGTGGGTCCT
GCTGGCAAACATGGAAACCGTGGTGAAACTGGTCCTTCTGGTCCTGTTGGTCCTGCTGGTG
CTGTTGGCCCAAGAGGTCCTAGTGGCCCACAAGGCATTCGTGGCGATAAGGGAGAGCCCGG
TGAAAAGGGGCCCAGAGGTCTTCCTGGCTTAAAGGGACACAATGGATTGCAAGGTCTGCCT
GGTATCGCTGGTCACCATGGTGATCAAGGTGCTCCTGGCTCCGTGGGTCCTGCTGGTCCTA
GGGGCCCTGCTGGTCCTTCTGGCCCTGCTGGAAAAGATGGTCGCACTGGACATCCTGGTAC
AGTTGGACCTGCTGGCATTCGAGGCCCTCAGGGTCACCAAGGCCCTGCTGGCCCCCCTGGT
CCCCCTGGCCCTCCTGGACCTCCAGGTGTAAGCGGTGGTGGTTATGACTTTGGTTACGATG
GAGACTTCTACAGGGCTGACCAGCCTCGCTCAGCACCTTCTCTCAGACCCAAGGACTATGA
AGTTGATGCTACTCTGAAGTCTCTCAACAACCAGATTGAGACCCTTCTTACTCCTGAAGGC
TCTAGAAAGAACCCAGCTCGCACATGCCGTGACTTGAGACTCAGCCACCCAGAGTGGAGCA
GTGGTTACTACTGGATTGACCCTAACCAAGGATGCACTATGGATGCTATCAAAGTATACTG
TGATTTCTCTACTGGCGAAACCTGTATCCGGGCCCAACCTGAAAACATCCCAGCCAAGAAC
TGGTATAGGAGCTCCAAGGACAAGAAACACGTCTGGCTAGGAGAAACTATCAATGCTGGCA
GCCAGTTTGAATATAATGTAGAAGGAGTGACTTCCAAGGAAATGGCTACCCAACTTGCCTT
CATGCGCCTGCTGGCCAACTATGCCTCTCAGAACATCACCTACCACTGCAAGAACAGCATT
GCATACATGGATGAGGAGACTGGCAACCTGAAAAAGGCTGTCATTCTACAGGGCTCTAATG
ATGTTGAACTTGTTGCTGAGGGCAACAGCAGGTTCACTTACACTGTTCTTGTAGATGGCTG
CTCTAAAAAGACAAATGAATGGGGAAAGACAATCATTGAATACAAAACAAATAAGCCATCA
CGCCTGCCCTTCCTTGATATTGCACCTTTGGACATCGGTGGTGCTGACCAGGAATTCTTTG
TGGACATTGGCCCAGTCTGTTTCAAAACGCGTACGCGGCCGCtcgagcgaTAA[…]

The procollagen-SBP-mGFP-COL1A1 in pLVXPuro (GFP-COL; Addgene #110726) was made during a project undertaken prior to this thesis work as described in the supplementary information on Addgene. All genetic constructs used for transfection of human cells, that were newly generated for this work and not yet described in the literature (and including GFP-COL), are listed in Table 5.

Template plasmids were amplified in DH5α *E. coli* (New England Biolabs) and subsequent extraction of plasmid deoxyribonucleic acid (DNA) was done using a MidiPrep kit (Thermo Fisher Scientific) and 50 mL cell suspension in Lysogeny broth (LB; Thermo Fisher Scientific) with 50 $\mu g \cdot mL^{-1}$ ampicillin (amp).

Table 5: Overview of homemade genetic constructs used for cell transfection.

Genetic construct	Vector	kbp	Addgene number
procollagen-SBP-mGFP-COL1A1 (GFP-COL)	pLVXPuro	13.4	110726
procollagen-mScCOL1A2 (mScCOL1A2)	pLVXPuro	12.9	Not yet deposited
Str-KDEL-IRES-ManII-mSc (ManII-mSc)	pIRES_neo3	6.1	117274
Str-KDEL-IRES-ST-mCherry (ST-Cherry)	pIRES_neo3	5.9	110727
Str-KDEL-IRES-mScSec23A (mScSec23A)	pIRES_neo3	8.1	117273
Str-KDEL-IRES-ManII-mTagBFP2 (ManII-BFP)	pIRES_neo3	6.1	Not yet deposited

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 H_2O, 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 H_2O) 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.

2.2 Cell culture

All human cell types used were adherent and cultured at 37 °C and 5% CO_2 in a humid environment. Cells used in the experiments were verified to be negative for mycoplasma.

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 most experiments, including the generation of stable cell lines. RPE-1 cells were grown in 10 cm dishes [35 mm live cell dishes (MatTek)] 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.

Near primary lung fibroblasts IMR-90 (ATCC Cat# CCL-186) were grown in minimum essential medium (MEM) supplemented with L-glutamine and 10% decomplemented FBS (Thermo Fisher Scientific). DMEM containing 2% M199 (Thermo Fisher Scientific) and 10% decomplemented FBS was used for cell culture of BJ-5ta hTERT-immortalized foreskin fibroblasts (ATCC Cat# CRL-4001). Normal human dermal fibroblasts (primary adult skin fibroblasts NHDF-Ad-Der (Lonza Cat# CC-2511), from here on referred to as NHDF-Ad) were grown in FBM medium supplemented with fibroblast growth medium FGM-2 (Lonza). Fibroblasts were passaged every 4 -7 days when a confluence of about 70% was reached.

Primary control (C1 and C2; derived from unrelated adults) and patient skin fibroblasts coded COL^{mut} and $P4HB^{mut}$ with rare types of OI were obtained at the hospital of Sheffield and grown in Nutrient mix F10 supplemented with 12% FBS and were passaged (1:2) every 4 - 6 days when a confluence of 60 – 80% was reached.

For passaging old media was aspirated and cells were rinsed with phosphate buffered saline (PBS) in order to remove remaining media, which could subdue the following treatment with 0.05% trypsin-EDTA (Thermo Fisher Scientific). After a minimal amount of trypsin-EDTA was added to cover the cells, cells were incubated at 37 °C, to provide

the optimal working temperature for the protease, for about 5 – 15 min until cell detachment was visible using bright field light microscopy. The reaction was stopped by adding fresh media. After thorough suspension by pipetting up and down, the cell suspension was added to a new dish containing fresh media according to the growth rate of the individual cell lines used. RPE-1 cells were not passaged beyond 45 passages, while primary fibroblasts were not passaged beyond 20 and transformed fibroblasts beyond 30 passages.

Cell lines were frozen for longer storage by pelleting cell suspensions as mentioned above for passaging via centrifugation for 3 min at 1000 rpm and resuspending in freezing medium containing 1% dimethyl sulfoxide (DMSO), 40% FBS and 50% culture medium into cryovials. Cells were wrapped in tissue and stored at -80 °C for 3-4 days, prior to long term storage in liquid nitrogen. For growing cells from frozen, cryovials were thawed and added to a dish containing double the volume of warm culture medium without puromycin compared to normal culture conditions to dilute toxic effects of DMSO. Medium was changed on the subsequent day.

2.2.1 Stable cell lines and cell transfection

RPE-1 are easy to transfect and were used to generate stable cell lines via lentiviral integration of genetic constructs of interest, like GFP-COL to generate GFP-COL-RPE (performed by Nicola Stevenson) and GFP only as a control (GFP-RPE; (Asante et al., 2014)). Virus containing the GFP-COL1A1 construct was generated using the Lenti-XTM Packaging Single Shots (vesicular stomatitis glycoprotein pseudotyped version) system from Takara Bio Europe 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 $\mu g \cdot mL^{-1}$ polybrene was added to cells. After 1 h of incubation at 37°C and 5% CO_2, 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 5 $\mu g \cdot mL^{-1}$ puromycin dihydrochloride (Santa Cruz Biotechnology) 72 h after transfection. Media for GFP-COL-RPE and GFP-RPE was further supplemented with 3 – 5 $\mu g \cdot mL^{-1}$ puromycin maintain engineered cell lines.

To reduce the variability of GFP-COL expression levels the GFP-COL-RPE cell line was sorted into four equally distributed populations with different expression levels via fluorescence-activated cell sorting (FACS; flow cytometry performed by Andrew Hermann) according to signal intensity of GFP. The 25% of cells expressing the lowest level was chosen for future trafficking experiments.

For the generation of GFP-COL and mScCOL1A2 expressing RPE-1 (GFP-COL1A1/mScCOL1A2-RPE) GFP-COL-RPE were transduced using virus containing mScCOL1A2 as described above.

Human 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 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 8 – 20 h depending on the genetic construct used prior to media change, followed by fixation or live cell trafficking experiments.

2.3 Immunofluorescence

For immunofluorescence (IF), cells were grown on 13 mm (thickness 1.5; VWR) autoclaved cover slips in 12 or 6-well plates. For cell fixation, media was aspirated, cells were washed with PBS and fixed with either 4% paraformaldehyde (PFA) for 15 min at RT or ice-cold methanol at -20 °C for 4 min, following repeated rinsing with PBS. 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 (Thermo Fisher Scientific, A18967-01). 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 min. Treatment with Triton-X-100 was not performed, when investigating extracellular collagen via IF.

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 listed in Table 6 in blocking solution. Samples were rinsed three times with PBS for 5 min after incubation with primary and secondary antibodies, respectively.

As secondary antibodies 2.5 $\mu g \cdot mL^{-1}$ 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 for stimulated emission depletion (STED) microscopy. For imaging via widefield microscopy, MOWIOL 4-88 (Calbiochem, Merck-Millipore, UK) mounting media was used. Samples for widefield microscopy and post-fixation live cell samples were incubated with 1 $\mu g \cdot mL^{-1}$ DAPI in PBS (Thermo Fisher Scientific) for 3 min at RT prior to repeated washing and mounting or storage in PBS at 4 °C in the dark until further imaging.

Table 6: List of primary antibody dilutions or known working concentrations used for either immunofluorescence (IF) or Western Blot (WB) experiments.

Antibody description	Species	Dilution or concentration in $\mu g \cdot mL^{-1}$ for IF	Dilution or concentration in $\mu g \cdot mL^{-1}$ for WB	Company	Catalogue number or reference
alpha-tubuline clone B-5-1-2 (monoclonal)	mouse	1:1000	-	Sigma	T5168
ATG16L (polyclonal)	rabbit	2.50	-	MBL	PM040
COL1a1 (polyclonal)	rabbit	0.50	2.50	Novus Biologicals	NB600-408
COL1A1 3391 (monoclonal)	mouse	5.00	-	Merck	MAB3391
COL1a1 C-propeptide (monoclonal)	rabbit	10.00	-	QED	42024
COL1a1 C-propeptide (monoclonal)	rabbit	10.00	-	QED	42043
COL1a2 (polyclonal)	rabbit	-	1.25	Abcam	ab96723
dynein intermediate chain (DIC74.1; monoclonal)	mouse	-	1:1000	Merck	MAB1618
EEA1 Clone 14/EEA1 (RUO) (monoclonal)	mouse	1.25	-	BD Transduction Laboratories	610456
GAPDH (monoclonal)	mouse	-	0.33	Thermo Fisher Scientific	AM4300
GFP (monoclonal)	mouse	0.50	0.50	Covance	MMS-118P
Giantin (polyclonal)	rabbit	1:2000	-	BioLegend	Poly19243
GM130 (monoclonal)	mouse	0.25	-	BD Biosciences	610823
GRASP65 (polyclonal)	sheep	1:1500	-	gift from Jon Lane	(Cheng et al. 2010)
Hsp47 (monoclonal)	mouse	0.75	-	ENZO	M16.10A1
LF39 (N-propeptide COL1A1) (polyclonal)	rabbit	1:200	-	gift from Larry Fisher	(Fisher et al., 1995)
LF41 (C-propeptide COL1A1) (polyclonal)	rabbit	1:200	-	gift from Larry Fisher	(Fisher et al., 1995)
Rab11 (polyclonal)	rabbit	2.00	-	Sigma	R5903
RFP [6G6] (monoclonal)	mouse	-	1:1000	Chromotek	6g6-20
SBP, clone 20 (monoclonal)	mouse	-	1:1000	Merck	MAB10764
Sec12 (polyclonal)	rabbit	-	1:500	gift from Balch Lab	(Weissman et al., 2001)
Sec24A (polyclonal)	rabbit	-	1:50	Stephens Lab (homemade)	(Satchwell et al., 2013)
Sec24C (polyclonal)	rabbit	1:250	1:50	Stephens Lab (homemade)	(Townley et al., 2008)
Sec24D (polyclonal)	rabbit	1:100	1:50	Stephens Lab (homemade)	(Palmer et al., 2005)
Sec31A (monoclonal)	mouse	0.25	-	BD Biosciences	612350
SP1.D8 telopeptide COL1A1 (polyclonal)	rabbit	0.09	-	DSHB	deposited by Furthmayr, H.
TANGO1 (polyclonal)	rabbit	0.40	-	Sigma-Aldrich prestige	HPA056816-100UL
TFG (polyclonal)	rabbit	-	0.50	Novus Biologicals	NBP2-24485
transferrin receptor clone H68.4 (TfR; monoclonal)	mouse	1:500	-	Thermo Fisher Scientific	13-6800
WIPI2 clone 2A2 (monoclonal)	mouse	2.50	-	Bio-Rad,UK	MCA5780GA

2.4 Induction of procollagen trafficking

Induction of procollagen trafficking and secretion was performed by incubating cells for a stated amount of time in medium containing 50 µg·mL^{-1} L-ascorbate-2-phosphate (ascorbate).

To investigate intracellular trafficking in GFP-COL-RPE cells were first incubated with ascorbate as stated above for 24 h to flush out procollagen from the ER, followed by a controlled accumulation in the ER for 24 h via incubation in medium without addition of ascorbate and co-transfection with a bicistronic construct encoding an ER-hook (KDEL-tagged streptavidin to retain the SBP fusion protein in the ER), and a separate reporter protein. The co-expression of the ER hook here enabled the retention of SBP-tagged GFP-COL in the ER lumen until addition of biotin to trigger the dissociation of SBP from the streptavidin-bound KDEL "hook". Finally, to release procollagen in a synchronised manner, medium was supplemented to contain 500 µg·mL^{-1} ascorbate (as well as 400 µM biotin, when co-transfected with an ER-hook containing construct like ST-Cherry, ManII-BFP, ManII-mSc or mScSec23A). This system is based on the retention using selective hooks (RUSH) system (Boncompain et al., 2012).

For analysis of the dependence on the MT network, cells were treated as mentioned above, but incubated in presence of 5 µM nocodazole (NZ) for 60 – 120 min prior to the initiation of the trafficking experiment via addition of ascorbate/biotin/NZ to the cells.

2.5 Light microscopy

All images shown were not altered from the raw data, unless mentioned.

2.5.1 Confocal fixed cell and live cell imaging

Images of GFP-COL-RPE were obtained using confocal microscopy using Leica SP5II for fixed or Leica SP8 for live samples (Leica Microsystems, Milton Keynes, UK). Live cell imaging was performed using a Leica SP8 confocal laser scanning microscope with 63x HC OL APO CS2 1.42 numerical aperture 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 20 mW solid state yellow laser at 561 nm for the red channel, respectively. Time courses were acquired using the sequential scanning mode between lines, imaging speed set to 700 Hz, 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-COL and the chosen red reporter (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 10 – 30

seconds between time points for multiple positions. Acquisition with a temporal resolution of 1 frame per second (fps) was accomplished by using a single position imaged at two – three times line averaging, and 700 - 1000 Hz imaging speed and a zoom factor of 6, depending on the cell size.

Movies and movie stills were enhanced in brightness and contrast for both channels, by using ImageJ autocorrection. 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.

2.5.2 Super resolution microscopy

A gated STED SP8 X (Leica) system was used to acquire images for super resolution. GFP-COL-RPE samples were co-transfected with mScSec23A following the protocol for the induction of procollagen trafficking described above (2.4). Samples were fixed after 10 min incubation with 500 $\mu g \cdot mL^{-1}$ ascorbate and 400 µM biotin with 4% PFA in PBS and immunolabelled with an antibody targeting GFP (mouse) and Sec24C (rabbit). To enhance the signal of GFP-COL and mScSec23A in the samples, the secondaries were chosen for the same respective wavelength of GFP (using an anti-mouse Alexa-Fluor-488–conjugated secondary) and mSc (using a donkey anti-rabbit Alexa-Fluor-568–conjugated antibody). Immunofluorescence was performed as in 2.3. Samples were imaged at 400 Hz scan speed using unidirectional scanning. Signals were detected with gated hybrid SMD GaAsP detectors using a 100x HC PL APO CS2 oil immersion lens (numerical aperture 1.4; serial number 506378). A zoom of 3x to 5x with a pixel size of 23 – 24 nm was used, and fluorophores were excited at 488 and 568 nm via a white light laser using a sequential scanning setting starting with the red channel (excitation at 568 nm, with a 660-notch filter and the detection range of emission set to 578–634 nm with 0.3-6 ns gating), and followed by the green channel to avoid bleaching caused by the 592 nm STED laser (excitation at 488 nm using a 488-notch filter, and detection of emission at 500–545 nm with a 1.5–7.6 ns gating). For both channels a pinhole airy unit of 1 was selected and images were acquired with a frame averaging of three and line accumulation set to two. Laser intensities used for the depletion lasers were 1.424 W for the 592-nm depletion laser (green channel) and 1.323 W for the 660-nm depletion laser (red channel), respectively. The Software used was Leica LAS X 3.4.0.18371.

2.6 Analysis of imaging data

2.6.1 Data analysis of GFP-COL from images obtained by confocal microscopy

Localisation analysis of images of GFP-COL-RPE co-transfected with ST-Cherry and antibody labelled against Sec31A (post-fixation after live-imaging until an accumulation of GFP-COL in the Golgi was visible) was performed automatically. Spatial overlap between Sec31A and GFP-COL was measured using a custom plugin for ImageJ/Fiji (Schindelin et al., 2012; Schneider et al., 2012). The procedure the plugin uses is

described in the following paragraph. The analysis was performed with n=4 independent data sets containing a total of 55 cells. Both the plugin and the following description were created by Stephen Cross (McCaughey et al., 2019):

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-COL, 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-COL 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-COL structures were subsequently identified using the same approach as for Sec31A, but with an additional low-pass sigma (spot width) filter. For broader GFP-COL 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 a threshold was applied using the maximum entropy approach (Kapur et al., 1985). Objects were identified as contiguous regions in the binary image and filtered using a high-pass size filter. At this point punctate and broad GFP-COL 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-COL 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-COL structures, except using the isodata thresholding approach (Ridler and Calvard, 1978). Any GFP-COL structures within 0.5 µm of a Golgi are removed. Finally, GFP-COL structures are filtered based on their mean intensity in the GFP-COL channel. Pixel-based overlap of Sec31A and GFP-COL 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.

2.6.2 Estimation of object diameters from super resolution images

The following paragraph was written in my own words and published (McCaughey et al., 2019):

For estimating diameter sizes of larger GFP-positive objects and small puncta, as well as the closely or colocalizing objects labeling for the inner COPII layer (via Sec24C antibody labeling and mScSec23A in the same channel), whole cells were captured using the confocal mode of the STED, before zooming in on an area of interest (showing colocalization and or bigger GFP-positive structures). Using ImageJ's line tool, a line with a width of 5 pixels for smaller objects, and 10 pixels for larger objects, was drawn through objects of interest and line graph were generated for both channels. The resulting line graph were subsequently fitted to one to two Gaussian curves when possible using MAT LAB R2016a (MathWorks), and the full width at half-maximum (FWHM) was obtained and calculated to show the FWHM in nanometers. Where multiple peaks existed, the summed FWHM of both fitted Gaussian curves was displayed in the figure instead. For large objects where Gaussian fitting was not possible, the diameters measured manually via ImageJ were selected as estimated diameter. The estimated diameters represent the maximum object diameter of the measured objects. A total of 14 cells were investigated, and 20 objects of interest were measured.

2.6.3 Analysis of extracellular collagen levels from patient-derived fibroblasts

Prior to analysis of extracellular collagen levels, cells were grown on glass coverslips for 72 h and fixed after incubation in presence of 50 µg·mL^{-1} ascorbate for 0.5 h or 24 h, respectively. For IF analysis of extracellular type I collagen abundance in primary patient-derived fibroblasts, confocal images were compared using automated segmentation of collagen fluorescence intensity using Volocity (version 6.3, Perkin Elmer) for samples from P4HB^{mut} or via determining the corrected total cell fluorescence (CTCF) for samples from COL^{mut} using the following formula (Burgess et al., 2010):

$$CTCF = integrated\ density - (area\ of\ selected\ cell \cdot mean\ fluorescence\ of\ background)$$

Statistical analysis was performed using an unpaired Student's t-Test (p-values <0.05 indicate that means are significantly different (99 % confidence interval)).

2.7 RNA and Protein extraction and analysis

2.7.1 RNA-sequencing

The RNA-sequencing (RNA-seq) data from wild type (WT) RPE-1 cells was derived from previously published data (Stevenson et al., 2017). Raw RNA-seq data are available in the ArrayExpress database under accession no. E-MTAB-5618.

2.7.2 Western-Blot analysis

For semiquantitative analysis of protein levels of Sec24A and Sec24C and COL1A1 levels in WT-RPE-1, GFP-RPE, and GFP-COL-RPE cells, cells were seeded on 10 cm dishes and grown for 4 days. Cells were incubated in 2 mL serum-free culture medium supplemented with (Sec24C and Sec24D lysates, as well as COL1A1) or without 50 $\mu g \cdot mL^{-1}$ ascorbate for 24 h. The medium was collected, and the cells were lysed for 15 min in 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. Protein fractions of medium and lysate were centrifuged at 13,500 rpm at 4°C for 10 min. The cell pellet was discarded. The supernatant was denatured 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. Transfer of protein bands onto a nitrocellulose membrane was performed at 15 V overnight. 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. Primary antibody dilutions used for Wester-Blot (WB) analysis are listed in Table 6. 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 overnight exposure and subsequent development. Cells from the Sec31A, Sec12, and TFG blots were incubated in normal culture medium before cell lysis and scraping. The immunoblot for Sec31A was done as described above, while Western blots showing intracellular levels of Sec12 (a gift from Balch laboratory, Scripps Research Institute, La Jolla, CA) and TFG were run on a 4–12% Bis-Tris (NuPAGE, Thermo Fisher Scientific) precast gel at 200 V in 3-(N-morpholino)propanesulfonic acid (MOPS, Thermo Fisher Scientific) running buffer for 50 min instead.

For semiquantitative analysis of WB from patient derived cells the same protocol as described above was used with a few minor changes:

Cells were seeded confluent and grown for 24 h prior to incubation in serum free Ham's F10 medium supplemented with or without 50 $\mu g \cdot mL^{-1}$ ascorbate for 24 h. The medium was collected. The cell lysate was obtained as described above, with the exception that cells were not scraped. Samples were run on Tris-Acetate gels and transferred overnight as described above.

2.7.3 Immunoprecipitation experiments

For the GFP-trap, Chromotek GFP-Trap_A (GFP-Trap coupled to agarose bead particle size of ~90 µm; Code gta-20) were used. GFP–COL–RPE and GFP–RPE

(used as control) were seeded near confluent on 15-cm dishes and incubated in the presence of ascorbate (50 µg·mL^{-1}) for 24 h, followed by ascorbate starvation for 24 h and by 15-min incubation in presence or absence of 500 µg·mL^{-1} ascorbate before cell lysis. All following steps were performed at 4 °C. The lysis buffer contained 10 mM Tris-HCl (pH 7.4), 50 mM NaCl, 0.5 mM EDTA, 0.5% (vol/vol) IGEPAL and protease inhibitor cocktail (Calbiochem). For cell lysis, cells were rinsed with ice cold PBS and incubated in 0.5 mL lysis buffer for 15 min. Lysates were subsequently collected and incubated for 30 min while gently mixing. Beads were equilibrated with lysis buffer (without protease inhibitor cocktail and IGEPAL; referred to as dilution buffer) using 20 µL of bead slurry per sample. Lysates collected after centrifugation at 13,500 rpm for 10 min were incubated with the GFP-trap beads for 2 h. Samples were subsequently centrifuged at 2,700 g for 2 min to collect the beads with bound sample. After they were washed twice with dilution buffer containing protease inhibitor cocktail, samples were boiled after addition of 43 µL 2x LDS buffer containing a reducing agent (Thermo Fisher Scientific) at 95°C for 10 min, and beads were separated from the denatured protein samples by centrifugation as described above.

For pulldown experiments using streptavidin beads (Streptavidin agarose beads, Merck Millipore, catalogue number 16-126) the procedure followed is the same as described for using the GFP-trap, with the exception that proteins were eluted from the beads using 25 µL of dilution buffer containing protease inhibitor cocktail (Calbiochem) and 8 mM biotin and incubating the samples for 1 h on the rotator at 4 °C prior to adding loading buffer for sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) as described in 2.7.2.

2.7.4 Protein extraction of cell derived matrix (CDM) for tandem mass tagging (TMT) and proteomics

Cells were seeded near confluent and grown for 3 – 4 days in presence of 50 µg·mL^{-1} ascorbate. For extraction of the cells from the matrix, cells were washed twice with PBS and incubated in 20 mM NH_4OH in PBS for 5 min at RT with vigorous shaking on a rocker. Detached cells were rinsed three times with MiliQ H_2O and incubated with 10 µg·mL^{-1} DNAse I in water for 30 min at 37 ˚C. Following another wash with water, matrix proteins were covered with 5% acetic acid in PBS and incubated at 4 ˚C overnight for subsequent scraping in proteomics buffer (125 mM Tris-HCl, pH 6.8, 0.1% SDS, 10% glycerol, 1% DDT containing protease inhibitor cocktail). CDM protein extract and acetic acids fractions were combined, following an acetone precipitation at - 20 ˚C overnight and centrifugation at 15000xg for 10 min at 4 ˚C, followed by resuspension of the air-dried matrix protein pellet in proteomics buffer. Samples were then handed over to the proteomics facility (100 µg per sample) to perform proteomics via tandem-mass-tagging (TMT).

2.7.5 Proteomic analysis via mass spectrometry (MS)

Proteomics performed from GFP-trap experiments

The entire sample was loaded onto a 3–8% Tris-Acetate precast gel (NuPAGE, Thermo Fisher Scientific) and run for 135 min at 100 V in Tris-Acetate running buffer supplemented with antioxidant (all Thermo Fisher Scientific).

Gel slices were subsequently cut and digested for proteomics analysis via MS (performed by Kate Heesom). The following description was provided by Kate Heesom (McCaughey et al., 2019):

Each gel lane was cut into six slices, and each slice was subjected to reduction (10 mM dithiothreitol (DTT), 56°C for 30 min), alkylation (55 mM iodoacetamide, RT for 20 min), and in-gel tryptic digestion (1.25 µg trypsin per gel slice, 37°C overnight). The resulting peptides were fractionated using an Ultimate 3000 nano-LC system in line with an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific).

In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nanotrap column (Thermo Fisher 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 Fisher Scientific) over a 150-min organic gradient, using seven gradient segments (1–6% solvent B over 1 min, 6–15% B over 58 min, 15–32% B over 58 min, 32–40% B over 5 min, 40–90% B over 1 min, held at 90% B for 6 min, and then reduced to 1% B over 1 min) 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 nanoelectrospray ionization at 2.1 kV using a stainless-steel emitter with an internal diameter of 30 µm (Thermo Fisher Scientific) and a capillary temperature of 250°C. Tandem mass spectra were acquired using an LTQ-Orbitrap Velos mass spectrometer controlled by Xcalibur 2.1 software (Thermo Fisher Scientific) and operated in data-dependent acquisition mode. The Orbitrap was set to analyse the survey scans at 60,000 resolution (at mass to charge ratio [m/z] 400) in the mass range m/z 300–2,000 and the top 20 multiply charged ions in each duty cycle selected for tandem MS in the LTQ linear ion trap. Charge state filtering, where unassigned precursor ions were not selected for fragmentation, and dynamic exclusion (repeat count, 1; repeat duration, 30 s; exclusion list size, 500) were used. Fragmentation conditions in the LTQ were as follows: normalized collision energy, 40%; activation q, 0.25; activation time, 10 ms; and minimum ion selection intensity, 500 counts. The raw data files were processed and quantified using Proteome Discoverer software v1.4 (Thermo Fisher Scientific) and searched against the UniProt Human database (downloaded September 14, 2017; 140,000 sequences) plus the GFP sequence using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10 ppm, and tandem MS tolerance was set at 0.8 D. Search criteria included carbamidomethylation of cysteine (+57.0214) as a fixed

modification and oxidation of methionine and proline (+15.9949) as variable modifications. Searches were performed with full tryptic digestion, and a maximum of three missed cleavage sites were allowed. The reverse database search option was enabled, and all peptide data were filtered to satisfy false-discovery rate of 1%.

Proteomics via TMT of patient CDM

The following description was provided by Kate Heesom (Balasubramanian et al., 2019):

For TMT Labelling and high pH reversed-phase chromatography, the samples were labelled with Tandem Mass Tag (TMT) multiplex reagents according to the manufacturer's protocol (Thermo Fisher Scientific, Loughborough, UK) and the labelled samples pooled. The pooled sample was then 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 (130 Å, 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 min. The resulting fractions were evaporated to dryness and resuspended in 1% formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). High pH RP fractions were further fractionated using an Ultimate 3000 nano-LC system in line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific).

For the chromatography following TMT the same protocol as described above for GFP-trap samples was used except for a few changes (description provided by Kate Heesom (Balasubramanian et al., 2019)):

Peptides were ionized by nano-electrospray ionization at 2.0 kV using a stainless steel emitter with an internal diameter of 30 µm (Thermo Fisher Scientific) and a capillary temperature of 275°C. All spectra were acquired using an Orbitrap Fusion Tribrid mass spectrometer controlled by Xcalibur 2.0 software (Thermo Fisher 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 400 000 and a max injection time of 100 ms. Precursors were filtered with an intensity range from 5000 to 1E20, according to charge state (to include charge states 2-6) and with monoisotopic precursor selection. Previously interrogated precursors were excluded using a dynamic window (60 s +/- 10 ppm). The MS2 precursors were isolated with a quadrupole mass filter set to a width of 1.2 m/z. ITMS2 spectra were collected with an AGC target of 10 000, max injection time of 70 ms and CID collision

energy of 35%. For FTMS3 analysis, the Orbitrap was operated at 30 000 resolution with an AGC target of 50 000 and a max injection time of 105 ms. Precursors were fragmented by high energy collision dissociation at a normalized collision energy of 55% to ensure maximal TMT reporter ion yield. Synchronous Precursor Selection (SPS) was enabled to include up to 5 MS2 fragment ions in the FTMS3 scan.

The raw data files were processed and quantified using Proteome Discoverer software v2.1 (Thermo Fisher Scientific) and searched against the UniProt Human database (140000 entries) using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10 ppm, and MS/MS tolerance was set at 0.6 Da. Search criteria included the addition of the TMT mass tag (+229.163) to peptide N-termini and lysine as fixed modifications. Data was filtered to satisfy false discovery rate (FDR) of 5%.

2.8 Additional software

For creating the figures in this dissertation Adobe Photoshop and Illustrator CC 2015 were used. The thesis was written in Microsoft Word and tables generated using Microsoft Excel (both 365). The final PDF document was generated with the help of Adobe Acrobat 9 Pro.

Chapter 3: Transport of endogenous procollagen in fibroblasts

3.1 Statement of contribution

All work in this chapter is my own and has been published (listed in the appendix) (McCaughey et al., 2019).

3.2 Introduction

The large size of triple helical type I procollagen, as well as the limited flexibility arising from its triple helical domain, make procollagen a potentially difficult cargo to transport between organelles (Bachinger et al., 1982; Leblond, 1989; Lightfoot et al., 1992; Rezaei et al., 2018). In order to be secreted to the extracellular space, procollagen first needs to traffic from the ER to the Golgi. Due to its size, a current favoured hypothesis for ER-Golgi transport of procollagen involves the use of large COPII-*megacarriers* (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). These transport carriers are proposed to be up to micron size in diameter. This was reported in cells overexpressing KLHL12 and type I procollagen (Gorur et al., 2017). To investigate if this observation also corresponds to endogenous expression of procollagen in non-transformed cells antibodies against COL1A1 and COPII were used to label endogenous protein in different types of fibroblasts.

Fibroblasts are one of the main producers of type I collagen in the human body (Hass, 1940; Hass and McDonald, 1940; Wu et al., 2020). The fibroblasts presented here are primary skin fibroblasts NHDF-Ad, near primary lung fibroblasts IMR-90 and non-transformed neonatal foreskin fibroblasts BJ-5ta. Treatment with ascorbate enhances collagen secretion (Murad et al., 1981). To investigate the short-term effect of ascorbate on procollagen transport fibroblasts were incubated with control media and media supplemented with ascorbate for 30 min prior to fixation and image acquisition.

3.3 Results

NHDF-Ad were labelled with various antibodies recognising different epitopes in pro-α1(I) (Fig. 13).

IF of NHDF-Ad with both the COL1A1 (Novus) and the QED (clone 42043) antibody showed almost complete colocalization (Fig. 14A). COL1A1 was observed throughout the cell with an occasional increased signal intensity in parts of the Golgi region positive for GRASP65 (a *cis/medial*-Golgi marker).

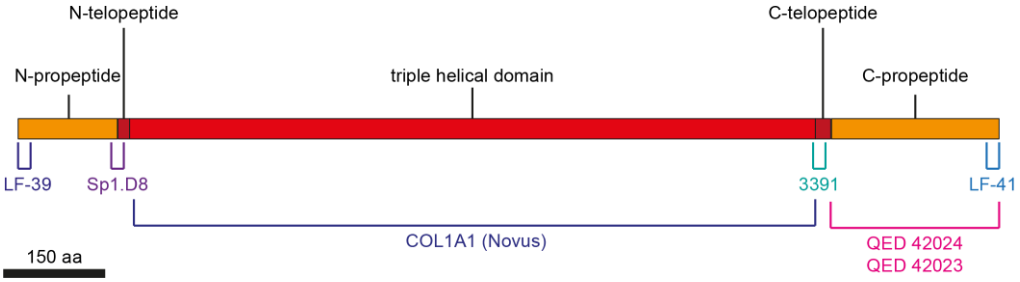


Fig. 13: Schematic of binding sites of COL1A1 antibodies.

Polyclonal antibodies LF-39 and LF-41 recognise the outmost regions of the N- and the C-propeptide (orange), respectively. The N- and C-telopeptides are recognised by monoclonal antibodies Sp1.D8 and 3391 (dark red), while the COL1A1 polyclonal antibody from Novus Biologicals (Novus) recognises the triple helical domain (red). Both monoclonals from QED recognise the C-propeptide of COL1A1.

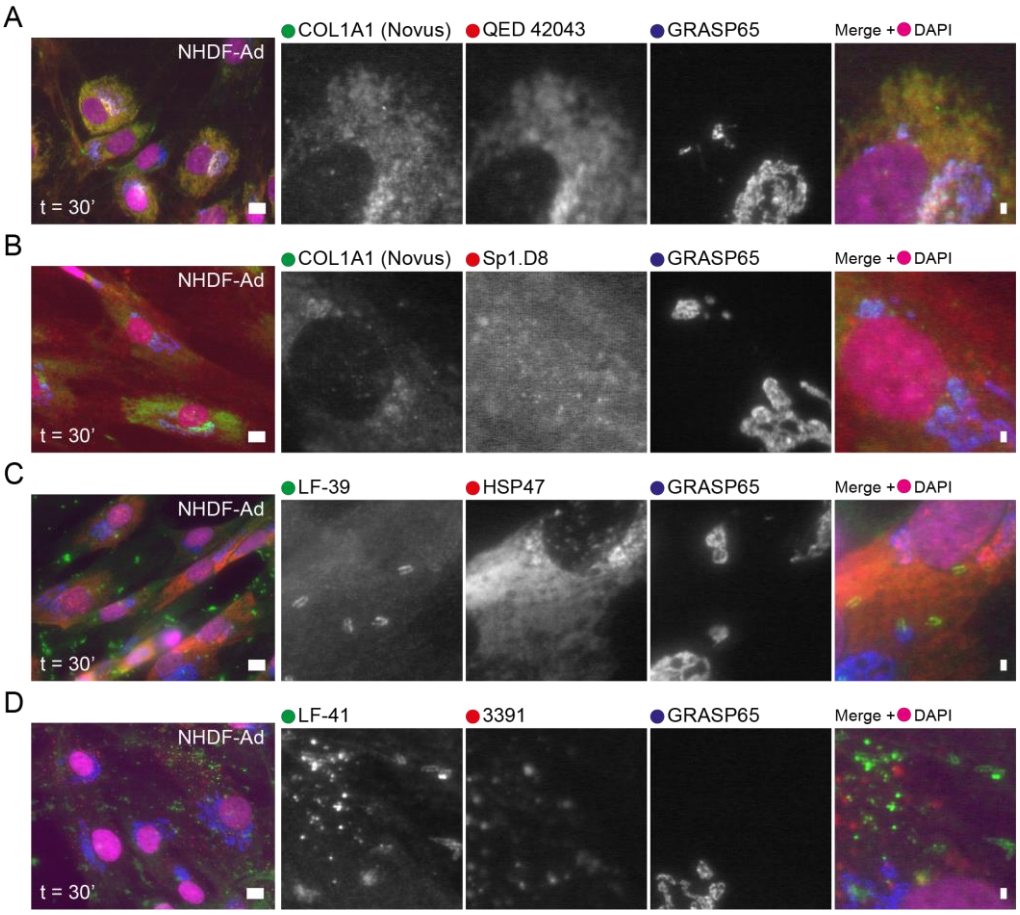


Fig. 14: Labelling of type I collagen in fibroblasts using various antibodies.

Widefield images of primary skin fibroblasts NHDF-Ad labelled with different antibodies against endogenous COL1A1 after 30 min incubation with 500 µg·mL^{-1} ascorbate before fixation with PFA; n>3 cells. Rabbit polyclonal antibodies are displayed in green, mouse monoclonal antibodies in red. All images show co-labelling with the *cis/medial*-Golgi marker GRASP65 in blue. Panels show the field of view, zoomed-in individual channels in grey scale, followed by the overlay image including nuclear DAPI labelling (imaged as separate channel in pseudo colour magenta). Scale bars, 10 µm; in enlargements, 1 µm. Endogenous labelling of COL1A1 was performed with COL1A1 (Novus) antibody in combination with COL1A1 C-peptide (PIP; QED 42043; A) or N-telopeptide COL1A1 (Sp1.D8) antibody (B). The N- and C-terminal propeptide of

COL1A1, labelled with LF-39 and LF-41, respectively, were co-labelled with either the collagen-specific chaperone HSP47 (C), or COL1A1 C-telopeptide marker 3391 (D). All images were equally brightness enhanced. (McCaughey et al., 2019).

The nuclear region remained devoid of signal for COL1A1. Labelling with the N-telopeptide marker Sp1.D8 resulted in a diffuse signal present throughout the entire cell with a high background signal (Fig. 14B) and only occasional puncta with high signal intensity that only sometimes co-label with COL1A1 (Novus). Antibodies directed against the N-(LF-39) and C-propeptides (LF-41) of COL1A1 were also tested. The signal for LF-39 observed by IF showed high intensity signals in the extracellular space and diffuse labelling of the cell body distinct from the signal obtained for the collagen chaperone HSP47 (Fig. 14C). Interestingly, LF41 (labelling the C-propeptide) and 3391 (labelling the C-telopeptide) both appeared in form of puncta both intra- and extracellular (Fig. 14D). These punctate structures do not appear to accumulate in the Golgi or colocalise with each other. Of note, none of the COL1A1 antibodies reveal obvious large circular structures of endogenous COL1A1.

To analyse intracellular trafficking of procollagen the COL1A1 (Novus) antibody was chosen for future experiments and IF labelling with this antibody is termed as “COL1A1” from here on. The localisation of endogenous COL1A1 in combination with markers for the Golgi, GRASP65, and COPII, or HSP47 was observed in presence and absence of incubation with ascorbate (Fig. 15).

Endogenous procollagen colocalises with HSP47 throughout the ER in NHDF-Ad (Fig. 15A). NHDF-Ad show varying expression levels of HSP47 and COL1A1. In addition, in some cells localisation of COL1A1 occurs in form of punctate structures in the Golgi area at both 0 and 30 min in the presence of ascorbate (Fig. 15A, images ii and iv, and B, images i–iv, arrows). These punctate structures do not colocalize with HSP47 or Sec31A. The COPII marker Sec31A can be observed as small punctate structures throughout the cell, clustering in close proximity to the Golgi labelling by GRASP65 (Fig. 15B, i-iii). None of the images in Fig. 15 show any obviously large circular structures (with diameters in micron size) positive for COL1A1, COPII or HSP47. Punctate structures positive for both COPII and COL1A1 at steady state are barely identifiable.

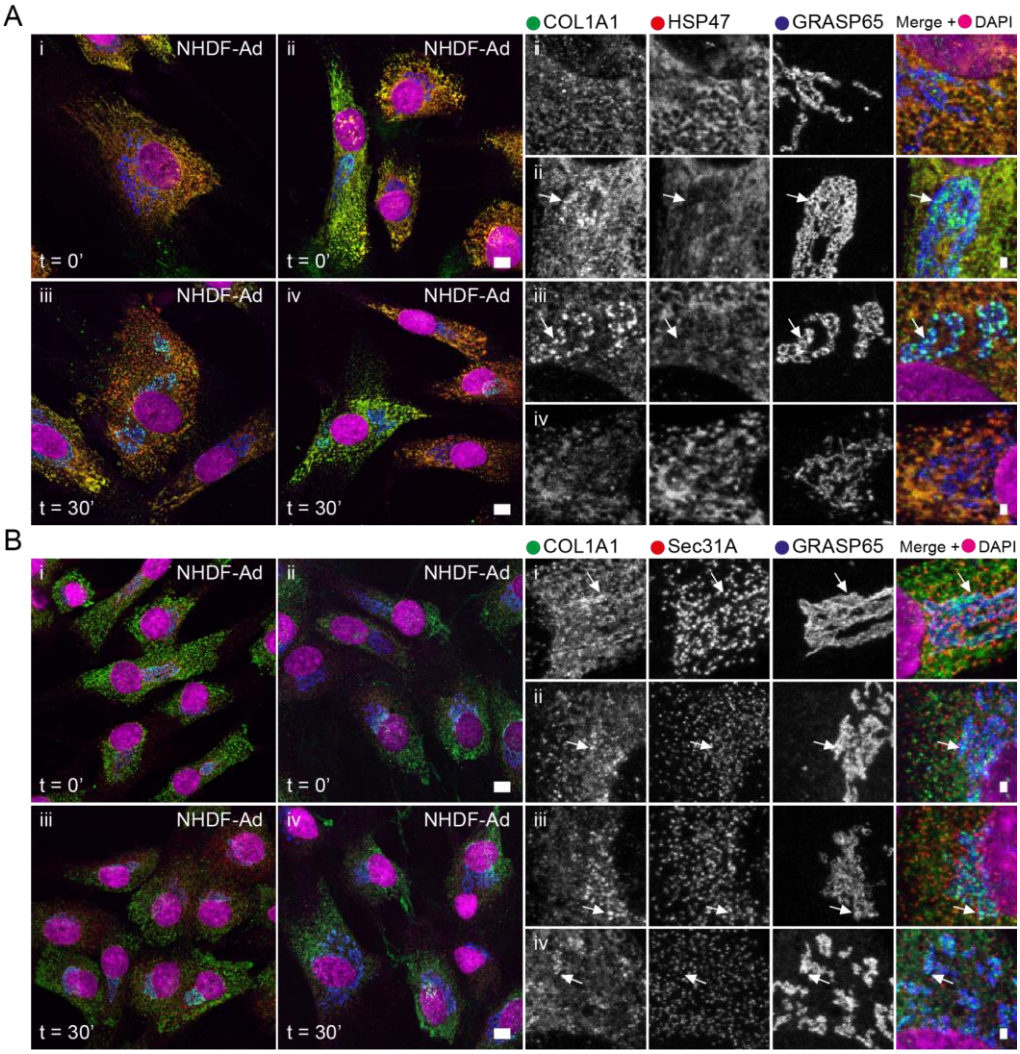


Fig. 15: Endogenous procollagen, HSP47 and COPII occurs as small puncta in primary skin fibroblasts.

Maximum intensity projection images of z stacks (z=3) of confocal images of primary skin fibroblasts NHDF-Ad after incubation in the presence of 500 µg·mL−1 ascorbate for 0 (images i and ii) and 30 min (images iii and iv), respectively. Cells were labelled using antibodies against COL1A1 (green) and either HSP47 (A) or the COPII marker Sec31A (B) in red, as well as a cis/*medial*-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 grey scale 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 localizing in close proximity to the Golgi, negative for HSP47 and COPII. Scale bars, 10 µm; in enlargements, 1 µm and n ≥ 10 cells in each case. All images were equally brightness enhanced. (McCaughey et al., 2019)

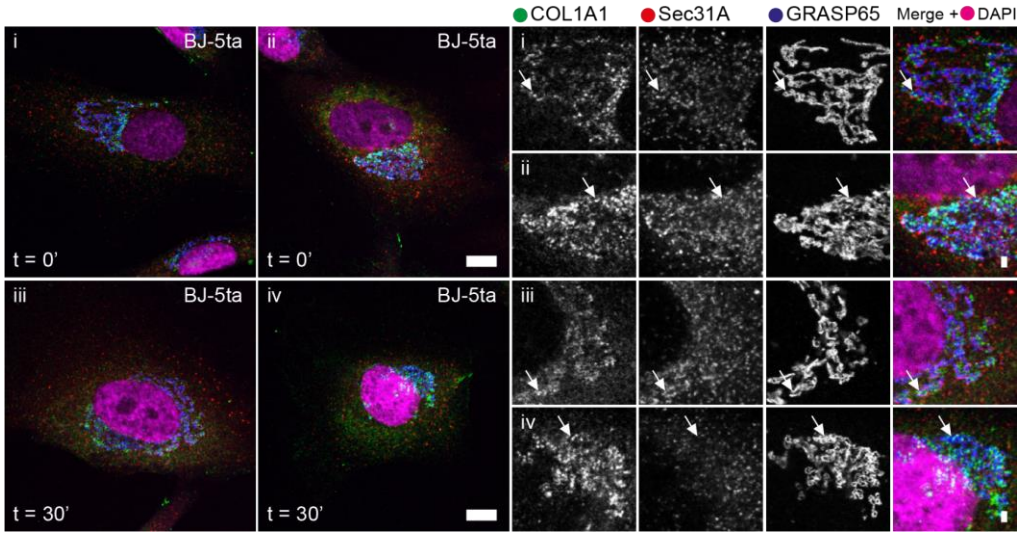


Fig. 16: Accumulation of endogenous procollagen in the Golgi area occurs unsynchronised in fibroblasts.

Confocal images of non-transformed foreskin fibroblasts (BJ-5ta) incubated with 50 µg·mL^{-1} ascorbate for 0 (images i and ii) and 30 min (images iii and iv), respectively. Cells were labelled using antibodies against COL1A1 (green) and COPII marker Sec31A in red, as well as a *cis/medial*-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 grey scale 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 localizing in close proximity to the Golgi. Scale bars, 10 µm; in enlargements, 1 µm and n ≥ 10 cells in each case. All images were equally brightness enhanced. (McCaughey et al., 2019).

Similar to NHDF-Ad, endogenous procollagen in a non-transformed human foreskin fibroblast cell line (BJ-5ta) localises to the ER and accumulates occasionally in the form of punctate structures in the Golgi area independent of addition of ascorbate to the media (Fig. 16, arrows). Immunofluorescent signals for both COPII and COL1A1 occur without the presence of obvious large (micron size) structures in these cells and procollagen puncta are again rarely positive for COPII.

To rule out that the lack of obvious large (e.g. micron sized) COL1A1-positive circular structures is antibody dependent, the same antibody as used in publications showing COPII- and COL1A1-positive large (micron sized) *megacarriers* (Gorur et al., 2017) was tested in IMR-90 and NHDF-Ad (Fig. 17 and Fig. 18).

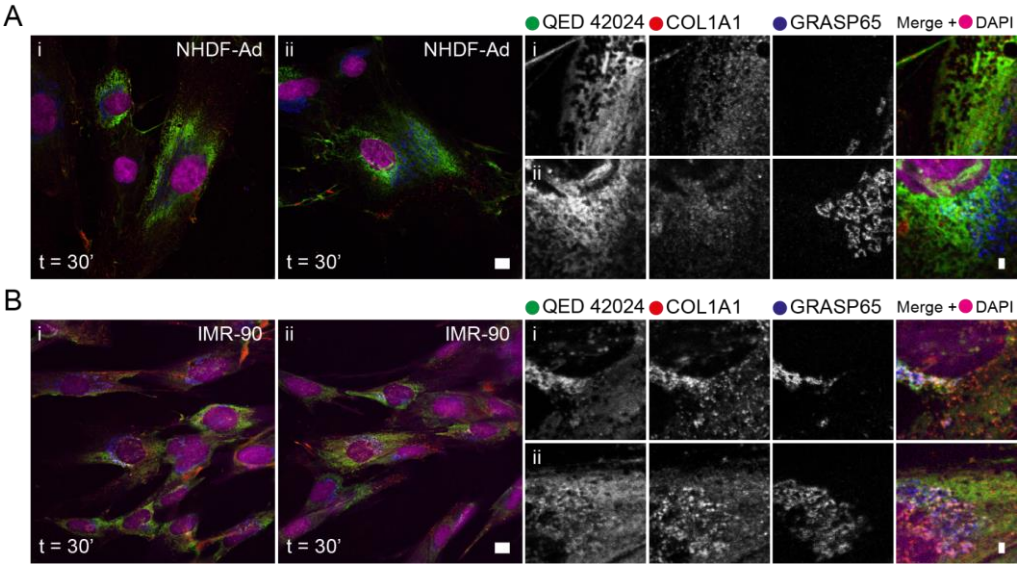


Fig. 17: Antibody testing of QED 42024 compared to COL1A1 (Novus) in skin and lung fibroblasts.

Confocal images of NHDF-Ad (A) and IMR-90 (B) labelled with a monoclonal antibody QED 42024, raised against the C-peptide (PIP) of COL1A1 in green, polyclonal antibody against COL1A1 (Novus) in red and *cis/medial*-Golgi marker GRASP65 in blue. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation. Panels show whole cells, as well as corresponding enlargements on the right. Enlargements show the separate channels in grey scale as well as the overlay image including nuclear staining in magenta (DAPI, imaged as separate channel). Scale bars, 10 µm; in enlargements, 1 µm and n > 10 cells in each case. All images were equally brightness enhanced. (McCaughey et al., 2019).

The observed localisation of procollagen in NHDF-Ad, as well as IMR-90 shows good colocalization between immunofluorescent signals for both QED 42024 and COL1A1 (Novus) (Fig. 17A-B). Again, no obvious large circular structures (micron size) positive for either antibody are observed in these cell types. To investigate the presence of COPII and procollagen positive structures cells were labelled with QED 42024 in combination with COPII markers Sec24C and Sec24D, respectively (Fig. 18).

Structures positive for both COL1A1 (here using QED 42024) and COPII (here with Sec24C/D) are rarely seen in fibroblasts expressing endogenous protein levels (Fig. 18A-B). When analysing a larger number of cells, occasional larger structures can be found in e.g. skin fibroblasts NHDF-Ad (Fig. 18A, ii, circle). These structures are, however, negative for COPII. Similarly, Fig. 18Bii shows above average sized structures positive for QED 42024, but not COPII (arrow), while there are some small punctate structures visible that are potentially positive for both COPII and COL1A1 (arrowhead). The latter do not appear to have an obviously larger size than that of other COPII structures observed.

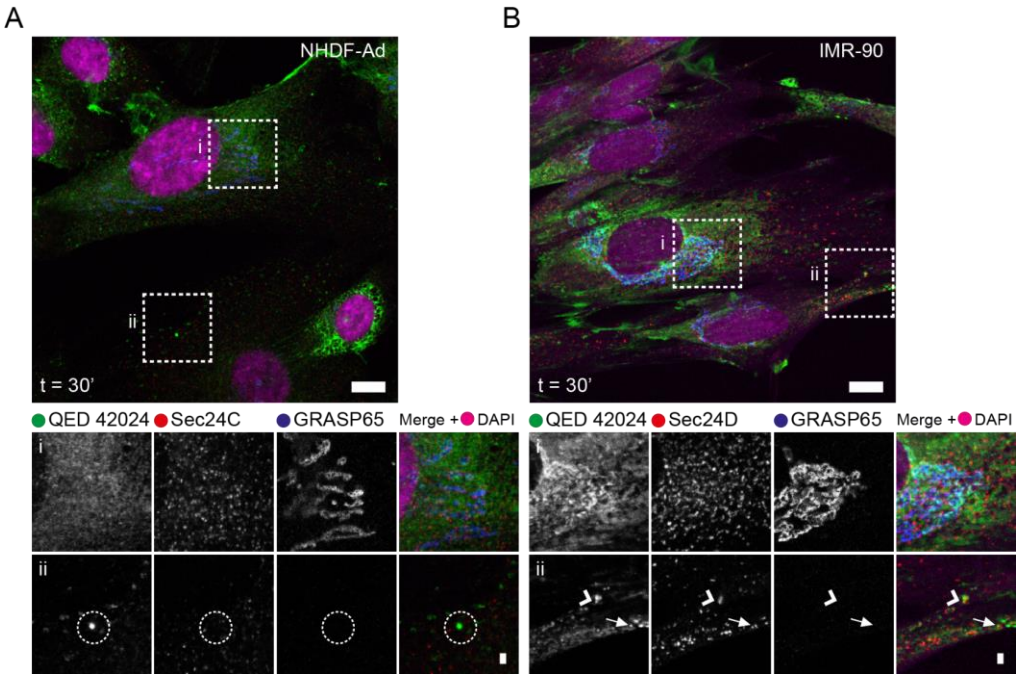


Fig. 18: A: Endogenous COPII and COL1A1 (via QED 42024-labelling) in fibroblasts can be visualised in form of small puncta.

Confocal images of NHDF-Ad (A) and IMR-90 (B) labelled with a monoclonal antibody QED 42024, raised against the C-peptide (PIP) of COL1A1a in green, a polyclonal against the inner layer COPII component Sec24C (A) or Sec24D (B) in red and *cis/medial*-Golgi marker GRASP65 in blue. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation. Panels show whole cells, as well as corresponding enlargements from boxed areas below. Enlargements show the separate channels in grey scale as well as the overlay image including nuclear staining in magenta (DAPI, imaged as separate channel). Circles indicate large structures positive for only QED 42024, arrows show punctate collagen structures negative for COPII and arrowheads indicate potential COPII- and QED 42024-positive structures. Scale bars, 10 µm; in enlargements, 1 µm and n > 60 cells in each case. All images were equally brightness enhanced. (McCaughey et al., 2019).

3.4 Discussion

The data demonstrates that different antibodies for type I collagen can result in a very different signal distribution of IF images. Only the COL1A1 (Novus) and the QED antibodies seem suitable to investigate intracellular procollagen. Other antibodies against COL1A1 tested here show high background signal with poor signal to noise ratio (Sp1.D8) and or non-specific binding (3391, LF-39 and LF-41) by IF. COL1A1 localises mainly to the ER in NHDF-Ad, BJ-5ta and IMR-90 fibroblasts, as seen by the overlap of signal for COL1A1 (Novus) with HSP47. Partial enrichment of COL1A1 in the Golgi area occurs in form of punctate structures both without and after supplementation of ascorbate to the media. Thus, ascorbate-dependent transport of endogenous COL1A1 in fibroblasts occurs unsynchronised. The abundance of COL1A1 especially in NHDF-Ad suggests a large pool of intracellular COL1A1 that might act as a reserve to allow fast response to e.g. signals initiating wound healing which might be linked to signalling of circadian clock genes (Cable et al., 2017; Hoyle

et al., 2017). None of the images acquired show immunofluorescent signals for endogenous COL1A1 in form of larger (e.g. micron sized) circular structures. One slightly larger than average sized structure positive for QED 42024 was observed in NHDF-Ad. This structure was, however, negative for both GRASP65 and COPII labelling Sec24C. Interestingly, most COL1A1 and COPII puncta observed in fixed cells do not colocalise in fibroblasts. Those structures identified positive for COPII marker Sec24D and COL1A1 (QED 42024) are no larger than other punctate structures positive for either COL1A1 or Sec24D.

Taken together these observations show that fibroblasts expressing endogenous protein levels do not show large (micron-sized) *megacarriers* positive for either COL1A1 and or COPII as previously reported by Gorur and colleagues in human fibrosarcoma (KI6) cells (Gorur et al., 2017) when labelling with QED 42024. Thus, the reported *megacarriers* (Gorur et al., 2017) might be a cell type dependent observation most likely caused by the stable overexpression of COL1A1 and KLHL12 in KI6 cells.

Furthermore, treatment with ascorbate does not allow for synchronised trafficking of procollagen in fibroblasts. Since the culture medium is prepared with FBS, it also cannot be ruled out that the medium itself contains small amounts of ascorbate. In addition, collagen secretion *in vivo* can be controlled by the circadian rhythm, which is not mimicked in cultured cells (Chang et al., 2020). To enhance synchronisation of collagen transport in cultured fibroblasts, the medium could be supplemented with agents like horse serum, dexamethasone, or cells could be temporarily temperature shifted to trigger initiation of a circadian cycle (Yeung et al., 2018; Yeung et al., 2014).

To enable synchronized transport of pro-a1(I) from the ER to the Golgi and visualise it in live cells, a different approach was chosen. A human type I procollagen was engineered to include a mGFP and SBP-tag to enable controllable transport through the secretory pathway.

Chapter 4: ER-to Golgi transport of procollagen in absence of large carriers

Statement of Contribution: If not stated otherwise, all work in this chapter is my own. The GFP-COL-RPE stable cell line was generated by Nicola Stevenson using the GFP-COL construct I generated during my Master's Thesis project. FACS sorting was performed in the flow cytometry facility at the University of Bristol by Andrew Hermann. The code for data analysis of COPII and GFP-COL overlap in relation to diameter was created by Steven Cross. Mass spectrometry and tandem-mass tagging for proteomics analysis was performed by the Proteomics facility at the University of Bristol led by Kate Heesom. RNA-seq was performed by The Genome Analysis Centre (Norwich, UK) with samples obtained by Dylan Bergen and is publicly available at ArrayExpress (https://www.ebi.ac.uk/arrayexpress/ accession number E-MTAB-5618). Most data in this chapter has been published and e.g. figure legends from the publication were written in my own words (listed in the appendix) (McCaughey et al., 2019).

4.1 Introduction

To visualise transport of type I procollagen through the early secretory pathway in a synchronised manner, human pro-α1(I) was tagged with SBP and mGFP (Fig. 19A). The tags were inserted after the N-terminal propeptide and upstream of the N-propeptide cleavage. Thus, following procollagen processing of the N-propeptide, the fluorescent reporter will be cleaved off together with the N-propeptide, resulting in extracellular type I collagen devoid of labelling. Therefore, this construct enables visualisation of type I procollagen transport from the ER to the Golgi apparatus, without additional signal from the extracellular space. The N-propeptide cleavage site was, furthermore, retained, since impaired cleavage can result in pathologies similar to those caused by OI and Ehlers-Danlos Syndrome (Cabral et al., 2005; Malfait et al., 2013).

An SBP-tag was included to enable use of the RUSH system (Boncompain et al., 2012). When cells expressing GFP-COL are in addition co-transfected with a streptavidin-KDEL construct, the latter acts as ER-hook. The streptavidin binds the SBP from GFP-COL and retains it in the ER. Release from the hook is triggered by addition of biotin. Biotin binds with a higher affinity to streptavidin and releases the SBP-containing GFP-COL from the hook, allowing it to freely traffic to the next compartment. To enhance collagen secretion ascorbate is added in addition to biotin for trafficking assays. The pro-α1(I) chain was chosen specifically for this, due to the stochiometric ratio of 2:1 of α1 to α2 chains per procollagen trimer. In addition, pro-α1(I), but not pro-α2(I), is capable of forming homotrimers (Sharma et al., 2017). Homotrimeric pro-α1(I) naturally occurs during embryonic development, wound healing and at steady state in small amounts in adult skin (Jimenez et al., 1977; Kay, 1986; Uitto, 1979). To further verify that the observations made also reflect the transport process of heterotrimeric

procollagen, a monomeric Scarlet-i (mSc; (Bindels et al., 2017)) tagged pro-α2(I) was tested in GFP-COL-RPE. These new genetically engineered type I procollagen chains were then used to test current models of ER-to-Golgi transport.

4.2 Results

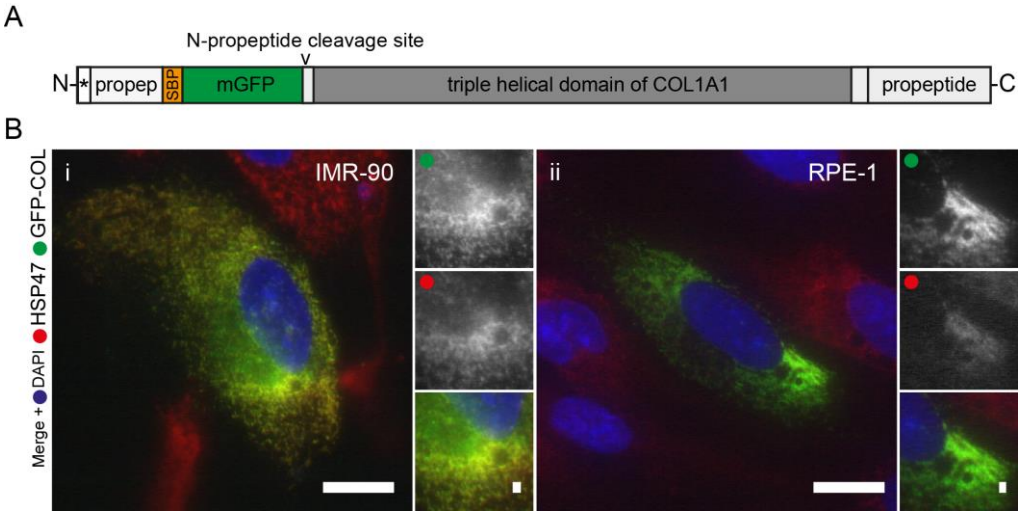


Fig. 19: GFP-COL transiently expressed in IMR-90 and RPE-1.

The GFP-COL fusion construct is shown in A. It consists of a synthetic pro-sequence of human pro-α1(I), followed by a SBP, 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 pro-α1(I). The signal sequence is indicated by an asterisk. Widefield microscopy of human lung fibroblasts IMR-90 (Bi) and human telomerase immortalised retinal pigment epithelial cells RPE-1 (Bii) transiently expressing GFP-COL (A). Images show cells 16 h after transfection with GFP-COL (green) colocalising with HSP47 (red) in the ER. Large images show whole cells, while small panels show enlargements with the corresponding channels in greyscale and the merge image including nuclear DAPI staining (blue). Number of cells imaged n ≥ 10 Scale bars indicate 10 and 1 µm. (McCaughey et al., 2019)

Fig. 19 shows GFP-COL (A) transiently expressed in human lung fibroblasts IMR90 (Bi) and human retinal pigment epithelium (RPE-1) cells (Bii). GFP-COL colocalises with the collagen chaperone HSP47 in the ER at steady state. To validate the capability of GFP-COL to traffic to the Golgi in an ascorbate-dependent manner, IMR-90 transiently expressing GFP-COL were incubated in presence of ascorbate for different time points prior to fixation (Fig. 20A), as well as imaged live (Fig. 20B). Both fixed cell and live cell experiments show that GFP-COL can accumulate in the Golgi as seen at 30 and 60 min post addition of ascorbate and GFP-positive puncta become apparent in the cell periphery (Fig. 20Aiii).

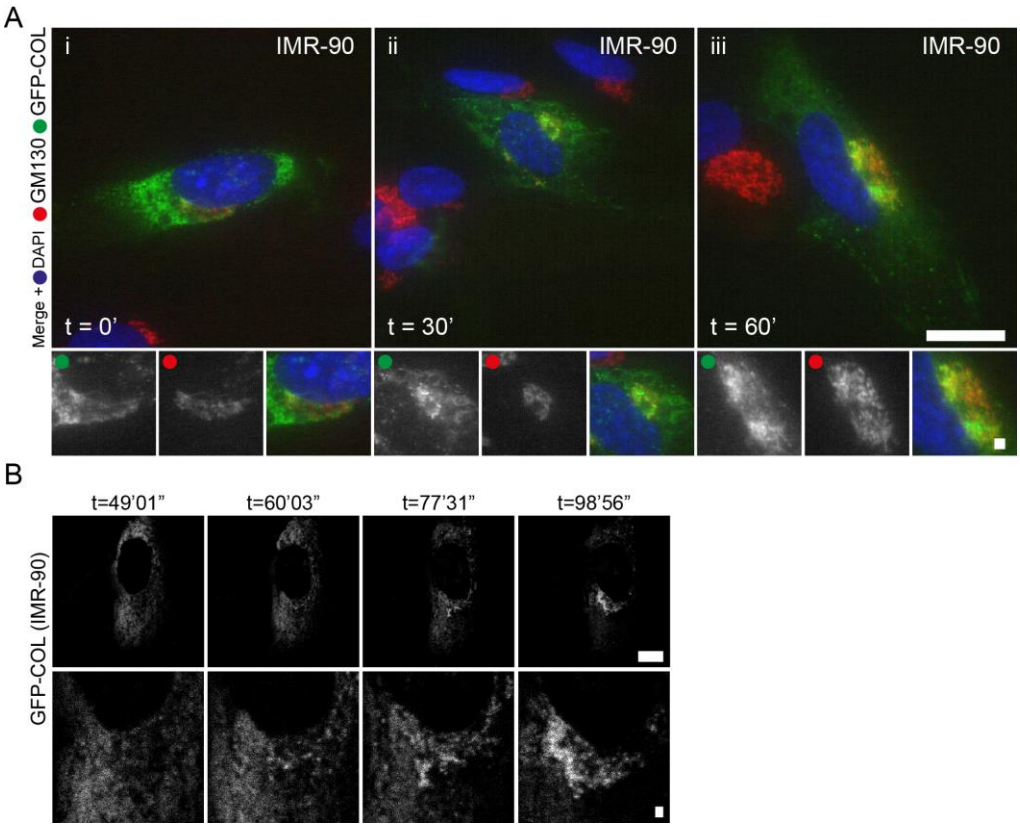


Fig. 20: GFP-COL transiently traffics ascorbate dependent when transiently expressed in fibroblasts.

Timepoints indicate incubation time in the presence of 50 µg·mL^{-1} ascorbate prior to fixation. Widefield microscopy of IMR-90 transiently expressing GFP-COL with GFP-COL (green) and *cis*-Golgi marker GM130 (red) and nuclear DAPI labelling (blue) of whole cells, followed by enlargements in corresponding greyscale images (A). A concentration of GFP-COL in the Golgi can be observed after both 30 (Aii) and 60 min (Aiii) in presence with ascorbate. Number of cells imaged n ≥ 10. Confocal image stills of live cell imaging of IMR-90 transiently expressing GFP-COL showing whole cells and enlargement of the Golgi area (B; video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V4.mp4). Number of cells imaged n ≥ 10 for fixed and n = 5 for live cells. Scale bars indicate 10 and 1 µm. (McCaughey et al., 2019)

To reduce the variability of expression levels and to exclude the investigation of artefacts caused by overexpression of GFP-COL, a RPE-1 cell line was created that stably expresses GFP-COL (GFP-COL-RPE) to allow synchronised transport of GFP-COL from the ER to the Golgi. RPE-1 cells were chosen, because these cells are telomerase-immortalised rather than transformed, near-diploid, have a proliferation rate of about ~14-18 h, and express different collagens including type I, as well as the COPII machinery including proteins like TFG and TANGO1 and are easy to transfect. Fig. 21 shows overall expression of different types of collagens in RPE-1 by RNA-seq. COL8A1 and COL4A1 and COL4A2 are the most abundantly expressed collagens in these cells.

Of interest, RPE-1 do express moderate amounts of COL1A1, however, only a very limited amount of COL1A2.

A

GENE	FPKM
COL1A1	12.67
COL1A2	0.02
COL2A1	0.00
COL3A1	0.13
COL4A1	109.02
COL4A2	177.11
COL4A3	6.01
COL4A4	13.51
COL4A5	16.75
COL4A6	0.79
COL5A1	0.00
COL5A2	44.87
COL5A3	0.00
COL6A1	1.48
COL6A2	0.00
COL6A3	0.98
COL6A5	0.00
COL6A6	0.00
COL7A1	5.51
COL8A1	336.36
COL8A2	2.74
COL9A1	0.00
COL9A2	0.13
COL9A3	0.02
COL10A1	0.04
COL11A1	19.89
COL11A2	0.53
COL12A1	72.74
COL13A1	24.27
COL14A1	0.04
COL15A1	0.02
COL16A1	2.71
COL17A1	0.21
COL18A1	8.65
COL19A1	0.00
COL20A1	0.00
COL21A1	0.10
COL22A1	0.00

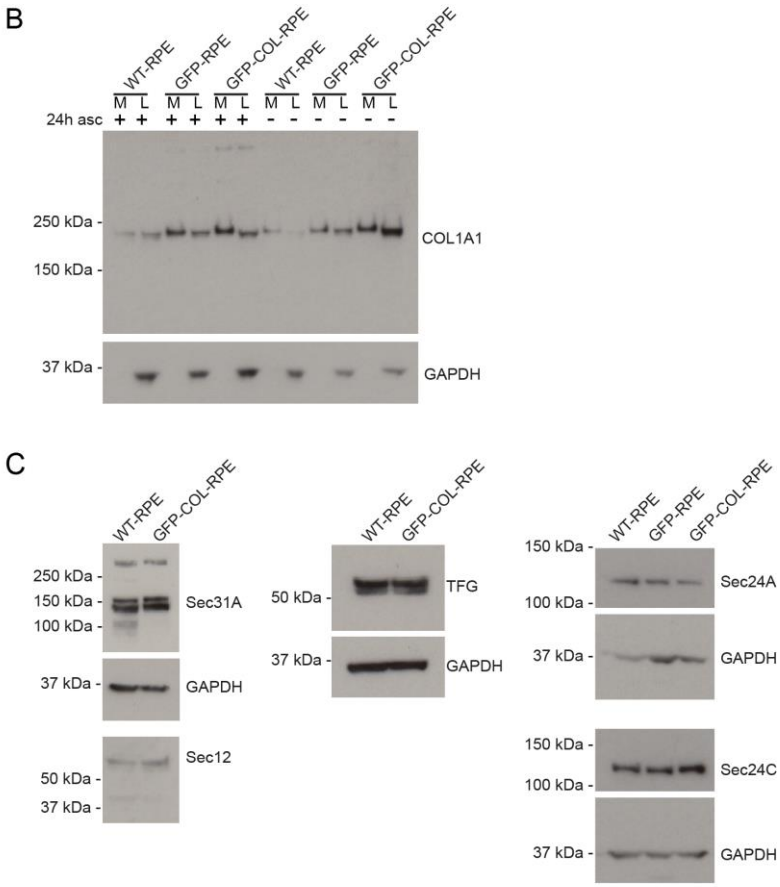


Fig. 21: Abundance of collagens and COPII machinery in wildtype and GFP-COL-RPE.

RNA-seq data from RPE-1 cells listing collagen isotypes detected and the number of fragments per kilobase of exon per million reads mapped (FPKM). Genes with an FPKM <0.1 are highlighted in grey (A). Immunoblots of cell lysates (L) and media fractions (M) of WT-RPE, GFP-RPE and GFP-COL-RPE cells incubated with (+) or without (-) 50 µg·mL^{-1} ascorbate for 24 h before collecting the samples. Antibodies blotted for are COL1A1 with GAPDH as loading control (B). Immunoblots of WT-RPE and GFP-COL-RPE cell lysates showing components of the COPII machinery Sec31A, Sec12, TFG, Sec24A, and Sec24C obtained after 4 days of growth on a 10-cm dish and in presence of ascorbate for 24 h in case of Sec24A and Sec24C and GAPDH as a loading control (C). (McCaughey et al., 2019)

GFP-COL-RPE cells were sorted by fluorescence-activated cell sorting (FACS) according to the intensity measurement of GFP resulting in the 25% of cells with lowest expression of GFP-COL. Abundance of type I procollagen observed on immunoblots with an antibody against the COL-domain of pro-α1(I) in WT-RPE, GFP-RPE and GFP-COL-RPE shows that GFP-COL-RPE do express slightly more pro-α1(I) compared to WT-RPE, but not much more compared to GFP-RPE (Fig. 21B). This shows variability

of collagen expression in different cell lines independent of GFP-COL expression. When comparing the lysate and media fractions one can observe a shift in the ratio of intracellular to extracellular collagen levels upon incubation in presence of ascorbate. In absence of ascorbate GFP-COL-RPE show that pro-α1(I) are higher in cell lysates than media fractions. This accumulation of intracellular pro-α1(I) is not as apparent in WT- or GFP-RPE. Incubation with ascorbate results in an obvious shift to more extracellular pro-α1(I) compared to intracellular levels in both GFP only and GFP-COL-RPE. This shows that secretion of type I collagen in RPE-1 is as expected ascorbate-dependent. In addition, stable expression of GFP-COL in RPE-1 does not influence the overall abundance of components of the COPII machinery like Sec12, TFG, Sec24A, Sec24C and Sec31A (Fig. 21C). Unless stated otherwise, GFP-COL-RPE were used for all further experiments.

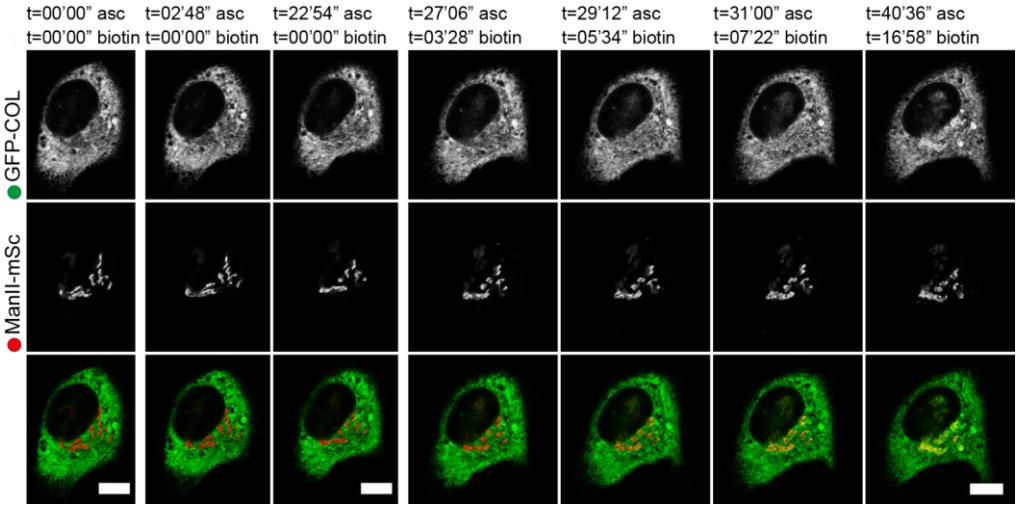


Fig. 22: GFP-COL can be trafficked using the RUSH system in a biotin controlled manner.

Image stills from confocal live-cell imaging imaged at one frame every 18 s (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V2.mp4). RPE-1 stably expressing GFP-COL (green; GFP-COL-RPE) 17 h after co-transfection with the *medial*-Golgi marker ManII-mSc (red; using the RUSH system). Time points indicate total time after addition of 500 $\mu g \cdot mL^{-1}$ ascorbate (asc) and 400 µM biotin. Panels show individual channels for GFP-COL and ManII-mSc in greyscale, followed by the overlay image below. No accumulation of GFP-COL in the Golgi region is visible after 27 min in presence of ascorbate but appears after ~5-7 min after addition of biotin. A total of n = 2 cells was acquired. Scale bar 10 µm. (McCaughey et al., 2019)

GFP-COL-RPE were further tested if trafficking using the RUSH system is biotin controllable. Cells co-transfected with a plasmid containing an ER-hook and a Golgi marker (the minimal localisation signal of ManII) tagged with mScarlet-i (ManII-mSc) after an IRES element were observed by live cell imaging. Fig. 22 shows that prior to addition of ascorbate and biotin, GFP-COL is distributed throughout the cell and the area marked by the Golgi reporter is devoid of signal for GFP-COL. Addition of

ascorbate alone for 23 min did not trigger accumulation of GFP-COL in the Golgi. When adding biotin to the same cell, however, rapid accumulation of GFP-COL in the Golgi occurs within 17 min post-addition. The time necessary for reaching apparent accumulation of GFP-COL in the Golgi area varies between cells from approximately 30 min to over 1 h. When using the RUSH system this time variation concentrates for most cells between 10 and 30 min but can occur in as little as 1 min and 23 seconds.

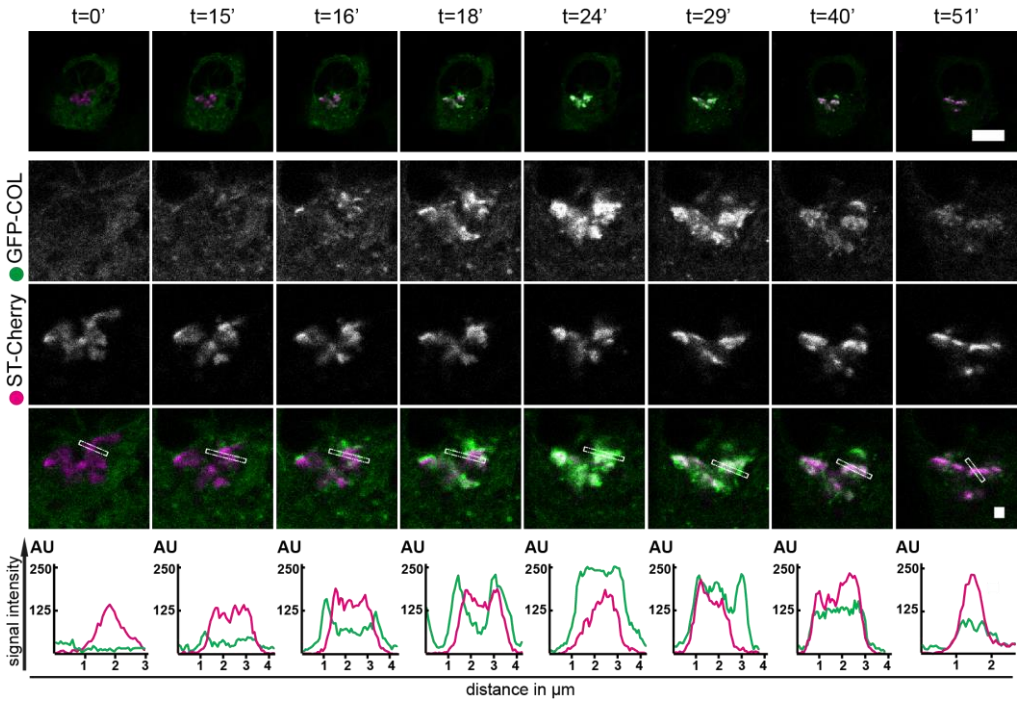


Fig. 23: Transport of GFP-COL to the Golgi occurs gradual in GFP-COL-RPE.

Confocal images from live-cell imaging of RPE-1 cells stably expressing GFP–COL (green; GFP–COL–RPE) co-transfected with the *trans*-Golgi marker ST-Cherry (magenta; using the RUSH system). Acquisition at one frame every 30 s (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V1.mp4). Images show whole cells with individual channels in greyscale and overlay images of enlargements of the Golgi area below, followed by line graphs showing the signal intensity (y axis) in arbitrary units for GFP–COL (green) and ST-Cherry (magenta) for the corresponding line with a 5-pixel width drawn through the Golgi in the enlarged overlays. The x axis shows the distance in micrometres. Time points indicate min of incubation in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin. For each set of live imaging experiments, three cells from the same dish were imaged simultaneously. A total of n = 4 sets was acquired. Scale bars, 10 µm; in enlargements, 1 µm. (McCaughey et al., 2019)

Fig. 23 shows that the transport of GFP-COL in GFP-COL-RPE from the ER to the Golgi apparatus occurs in a gradual manner from the edge of the Golgi. Here, GFP-COL-RPE were co-transfected with an ER-hook and separate *trans*-Golgi marker (minimal localisation signal for sialyltransferase) tagged with monomeric Cherry (ST-Cherry). Prior to addition of ascorbate and biotin GFP-COL localises in the ER with an even distribution of signal intensity distinct from the Golgi. After 15 min incubation with

ascorbate and biotin small punctate structures of GFP-COL with an increased signal intensity become visible throughout the cell and in proximity to the Golgi. Accumulation of GFP-COL in the Golgi area defined by ST-Cherry occurs in a gradual manner from the edge of the Golgi at 16 min over to filling of the Golgi at 24 min post-addition of ascorbate and biotin, as highlighted by the line graphs. During this process GFP-positive circular and very dynamic structures are emerging from the Golgi, as well as appearing in the cell periphery. At 29 to 51 min the GFP-COL signal gradually decreases in the Golgi apparatus. Over the time course of the video starting at about 24 min, signal intensity of GFP-COL in the ER gradually decreases, while the signal intensity of ST-Cherry does not. This implies that the decrease in signal intensity of GFP is not caused by effects of photobleaching. The dynamic structures slowly cease to be detectable with the overall intensity drop of GFP-COL after around 50 min (as seen in the corresponding video of Fig. 23).

About 5% of all cells imaged live show large (>500 nm diameter) vaguely circular GFP-positive structures (Fig. 24, circles). These structures appear to not colocalise with the Golgi labelling at any given time during the measurement and do not appear to be dynamic (Fig. 24A-D). Furthermore, these structures do not seem to interact with the trafficking of GFP-COL from the ER to the Golgi nor the dynamic punctate structures (Fig. 24A, arrows), nor change in signal intensity or shape throughout the time course of the experiment (Fig. 24B-C). Similar smaller structures that are Golgi-positive can be observed in some cells (Fig. 24B, asterisks). The localisation of GFP-COL and ST-Cherry over time for these structures is equivalent to the previously described accumulation at the edge of, with subsequent filling of the Golgi. Therefore, it can be assumed that these are smaller Golgi elements separate from the Golgi ribbon.

Interestingly, the process of GFP-COL concentrating in and around the Golgi area in RPE was observed in absence of any GFP-positive vesicular tubular clusters or carriers of any form, large or small translocating towards the Golgi, even when imaging at fast frame rates of 1 fps as seen in Fig. 25.

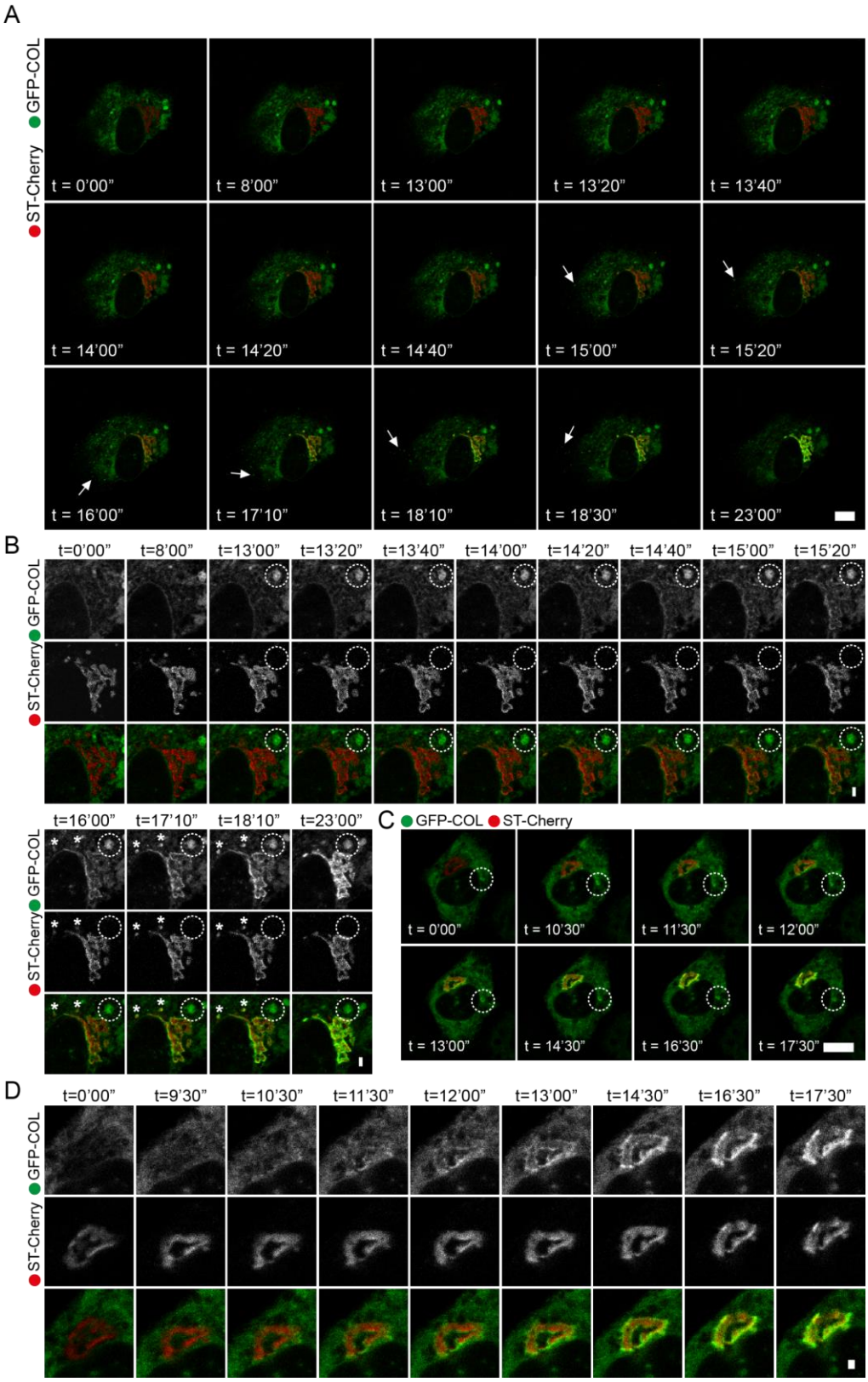


Fig. 24: Gradual accumulation of GFP-COL does not involve larger GFP-positive structures observed in some cells.

Image stills from confocal live-cell imaging of RPE-1 cells stably expressing GFP-COL (green; GFP-COL-RPE) co-transfected with the *trans*-Golgi marker ST-Cherry (red). Time points indicate min after incubation in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin. A whole cell imaged at one frame every 20 s (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/J

CB_201806035_V6.mp4) (A). Probable post-Golgi structures are indicated by arrows. Corresponding enlargements of A shown with channels for GFP-COL and ST-Cherry in greyscale, as well as the overlay image in B. Large circular GFP–COL-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-COL at the edge of the Golgi (t = 14–18 min), followed by filling of the Golgi (t = 23 min). C shows Image stills of a whole cell taken at one frame every 30 s (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V7.mp4). Accumulation and filling of the Golgi with GFP-COL occurs within t = 10.5 min. Corresponding enlargements of the Golgi area in C are displayed as separate channels in greyscale followed by the overlay in D. A total of n = 4 sets was acquired. For each set of live imaging experiments, three cells from the same dish were imaged simultaneously. Scale bars, 10 µm; in enlargements, 1 µm. (McCaughey et al., 2019).

Dynamic punctate structures can only be observed appearing in the cell periphery or emerging mostly from the Golgi during most videos (videos 1-3). The overall abundance of dynamic structures varies between cells as e.g. seen when comparing Fig. 24A and C. There are, however, few small GFP-positive punctate structures that appear to be translocating towards the Golgi from the cell periphery in the occasional RPE cell (video from Fig. 24 and Fig. 25) and GFP-COL transient expressing non-transformed foreskin fibroblasts BJ-5ta (Fig. 26B, arrows and corresponding video). Transport of GFP-COL in BJ-5ta appears very similar to ER-to-Golgi transport observed in GFP-COL-RPE. Localisation of GFP-COL prior to addition of ascorbate and biotin is distinct from the Golgi marker and GFP-COL accumulates at the edge of the Golgi at about 10 min ascorbate/biotin and subsequent filling of the Golgi at about 25 min (Fig. 26A-B). Filling of the Golgi occurs more localised to the nucleus facing side of the Golgi and small dynamic GFP-COL structures moving from the cell periphery to the Golgi appear more abundant and obvious in BJ-5ta compared to GFP-COL-RPE (Fig. 26A, arrows). These structures appear to be smaller than 500 nm in diameter. The limited number of such structures moving towards the Golgi by itself, however, does not seem to be sufficient as sole cause for the dramatic increase in signal for GFP-COL when accumulating in the Golgi area and are not detectable in most cells imaged.

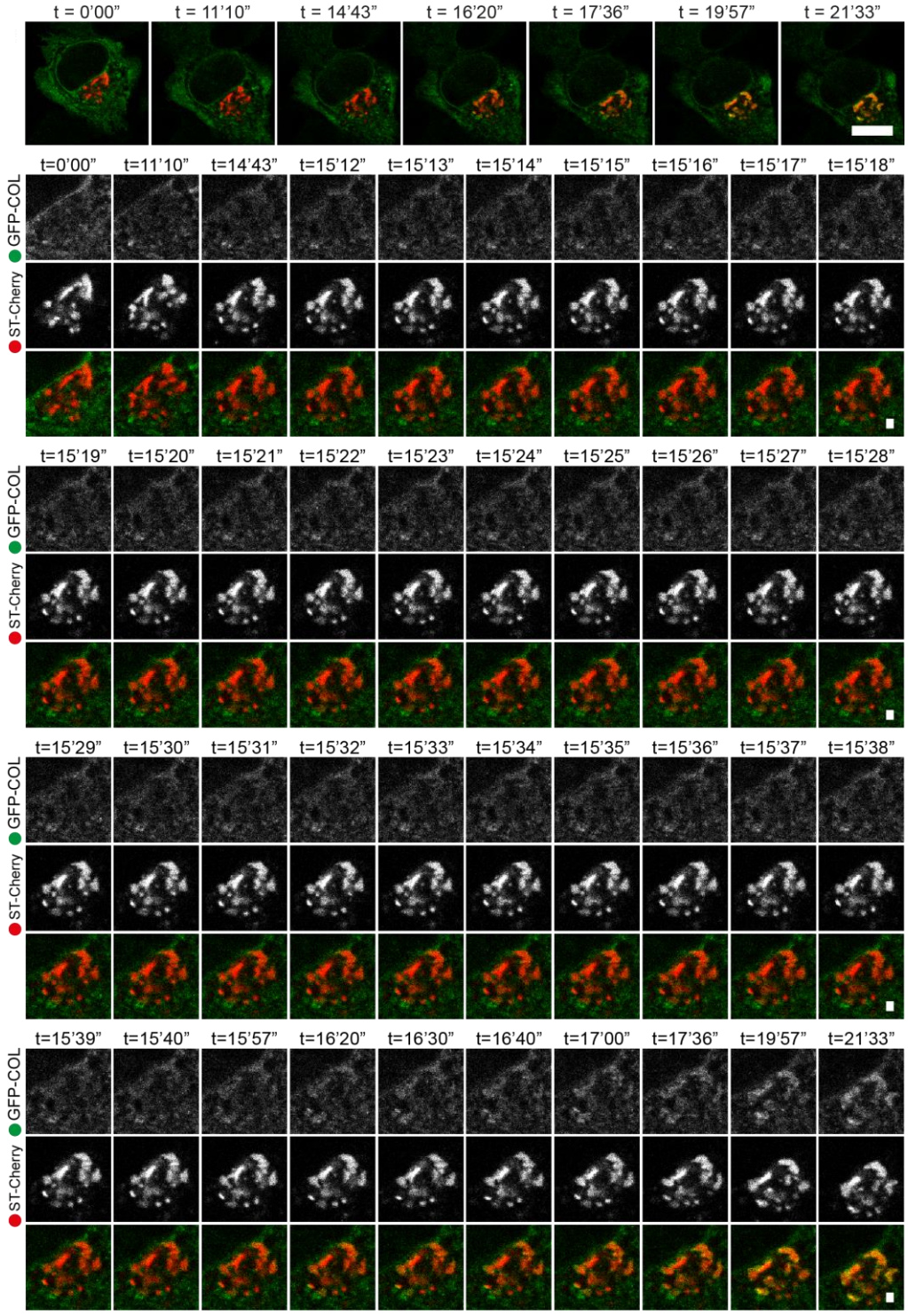


Fig. 25: Fast imaging of GFP-COL transport to the Golgi in absence of small or large vesicular carriers.

Image stills from confocal live-cell imaging of GFP-COL-RPE, with GFP-COL in green, 18.5 h after co-transfection with the *trans*-Golgi marker ST-Cherry (red; using the RUSH system) acquired at 1 fps (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V5.mp4). Time points indicate min after addition 500 µg·mL^{-1} ascorbate and 400 µM biotin. Shown are whole cells and corresponding enlargements with individual channels in greyscale, as well as the overlay image below. A total of n = 3 was acquired. Scale bars, 10 µm; in enlargements, 1 µm. (McCaughey et al., 2019).

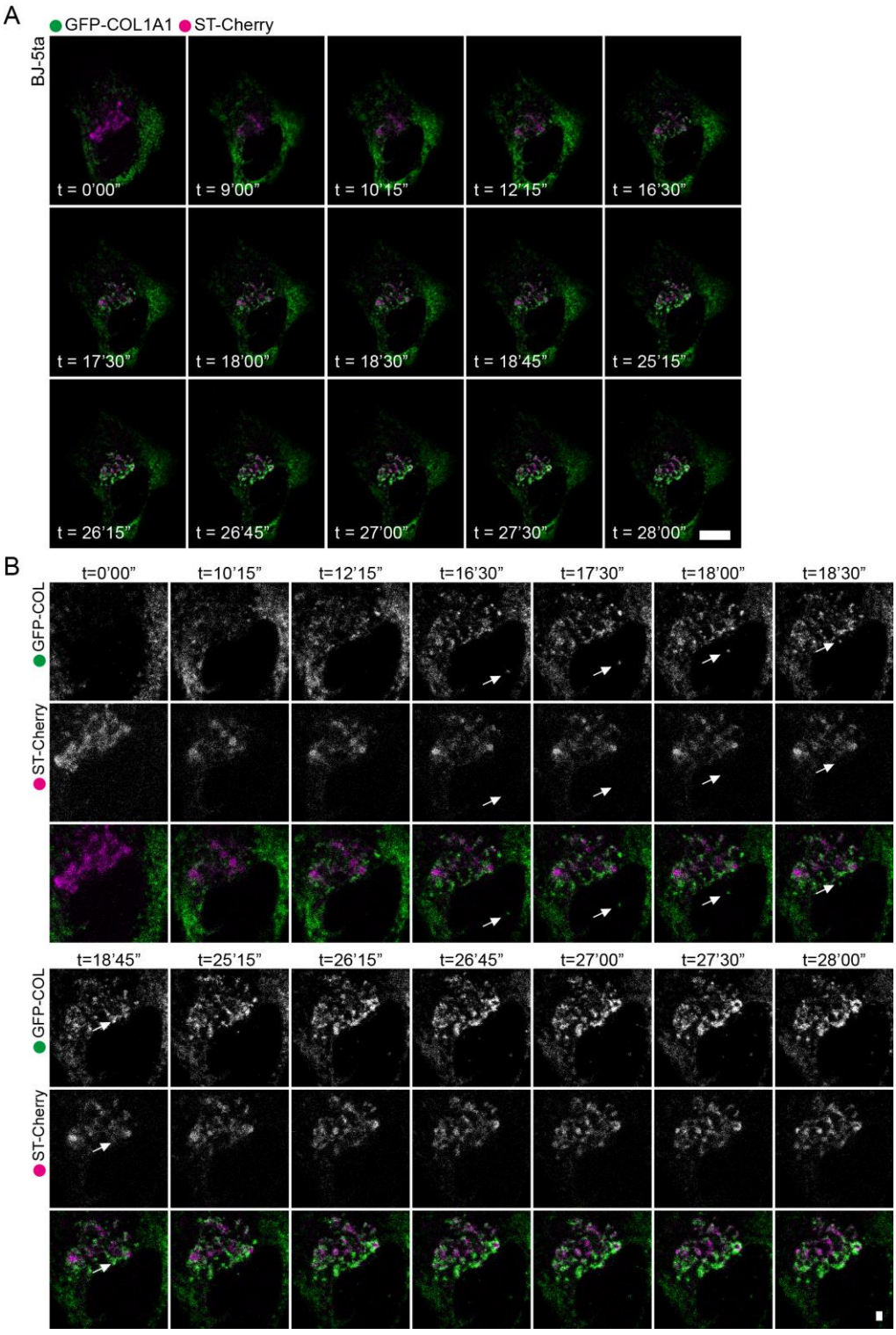


Fig. 26: Transport of transiently expressed GFP-COL to the Golgi in BJ-5ta reveals some punctate structures translocating towards the Golgi.

BJ-5ta transiently expressing GFP-COL (green) and the *trans*-Golgi marker ST-Cherry (magenta; using the RUSH system) 17 h after co-transfection. Time points indicate time after addition 500 µg·mL^{-1} ascorbate and 400 µM biotin. The cell was imaged at one frame every 15 s (video available at https://static-movie-usa.glencoesoftware.com/ source/10.1083/529/ e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V3.mp4). Panels show the whole cell (A) as overlay with corresponding enlargements of the Golgi area in greyscale for the separate channels GFP-COL and ST-Cherry, followed by the overlay image (B). An accumulation

and filling of the Golgi with GFP–COL becomes visible at ~10 min after addition of biotin. Arrows indicate a small punctate GFP-positive structure moving towards the Golgi from the cell periphery. Scale bars, 10 µm; in enlargements, 1 µm. (McCaughey et al., 2019).

To determine what the small dynamic punctate and the large not very mobile GFP-positive structures are, cells were fixed at time points when an accumulation of GFP-COL in or around the Golgi via live cell imaging was observed. The given time points of fixed samples in Fig. 27 - Fig. 32 are after live imaging and addition of ascorbate and biotin with subsequent fixation and labelling for components of interest.

Early stages of GFP-COL accumulation in the vicinity of the Golgi (Fig. 27i) suggests an accumulation of GFP-COL at the ER to Golgi interface, prior to entering the Golgi. This stage also does not show any visible large GFP-positive structures. Labelling for *cis/medial*-Golgi via antibodies targeting giantin usually appears to surround ST-Cherry labelling, verifying the assumed localisation of the *trans*-Golgi marker (Fig. 27). Small puncta can be observed in the cell periphery and close to the Golgi (Fig. 27i and ii, arrows). Large GFP-COL-positive structures (Fig. 27ii, circles) are negative for both the *trans*- and *cis*/*medial*-Golgi markers. Therefore, both large and small punctate GFP-positive structures appear negative for giantin and ST-Chery (Fig. 27, circles and arrows, respectively).

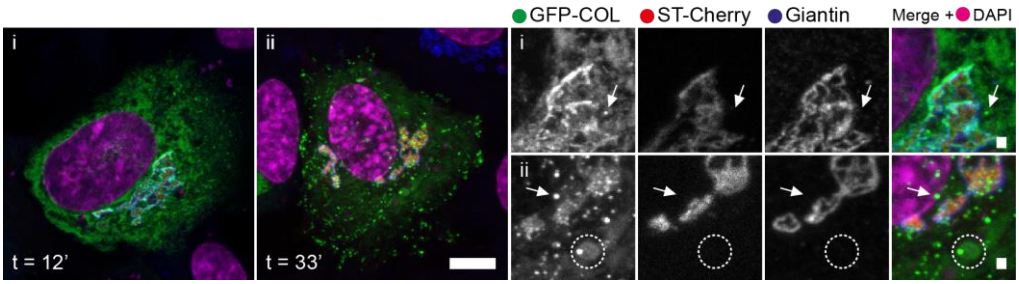


Fig. 27: Large circular GFP-COL structures are negative for Golgi markers.

Maximum projection confocal images of z stacks containing the Golgi apparatus of GFP-COL-RPE with GFP-COL (green) and 16 h post-transfection with ST-Cherry(red). Duration of 500 µg·mL^{-1} ascorbate and 400 µM biotin, respectively, corresponds to duration of prior live imaging at approximately one image every 30 s until trafficking of GFP-COL to the Golgi was detectable by eye. Cells were labelled with the *cis/medial*-Golgi marker giantin (blue). 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. Large GFP-COL-positive structures are highlighted by circles and small GFP-COL puncta negative for Golgi marker by arrows. Scale bars, 10 µm; in enlargements, 1 µm. n ≥ 10. (McCaughey et al., 2019).

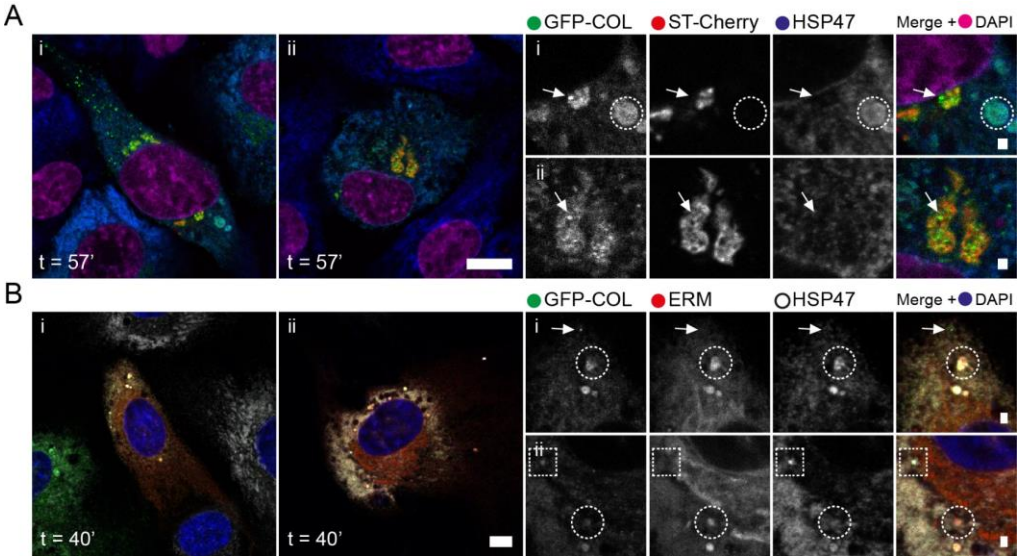


Fig. 28: Large GFP-COL structures are positive for ER-membrane and chaperone HSP47.

Confocal imaging of GFP-COL-RPE with GFP-COL in green co-transfected with either ST-Cherry 18 h after transfection (A) or the ER membrane marker ERM-mScarlet-i (ERM) 6–8 h after transfection (B) in red. Cells were fixed after the given time points of incubation with 500 µg·mL^{-1} ascorbate and 400 µM biotin and labelled for HSP47 (in blue in A or greyscale in B). 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 or blue in B). Arrows highlight small GFP-positive puncta negative for HSP47, Golgi or ERM, while circles indicate large GFP-positive structures. Puncta positive for both GFP-COL and HSP47, but not ERM are highlighted by a square. Scale bars, 10 µm; in enlargements, 1 µm. n ≥ 10. (McCaughey et al., 2019).

The large GFP-positive structures, however, are in fact positive for both endogenous collagen chaperone HSP47 and ER (defined by colocalization with a transiently expressed ER-membrane marker (ERM), cytochrome p450, tagged with mSc (Fig. 28A-B). Most small GFP-COL puncta are negative for HSP47, ERM and Golgi in proximity of the Golgi, as well as in the cell periphery (arrows). In some cells, a few small GFP-positive punctate structures could be identified that appeared positive for HSP47, but not ERM (Fig. 28B, square).

To investigate if the large GFP-positive structures are destined for degradation co-localisation with markers for early autophagosomal structures LC3B, WIPI2 and ATG16L were tested in GFP-COL-RPE (Fig. 29). Large GFP-positive structures were mutually exclusively co-labelled for either Cherry-LC3B or HSP47 (Fig. 29A, squares and circles, respectively). Fig. 29B shows a smaller punctate structure positive for ATG16L (arrow heads), while larger GFP-positive structures were mostly devoid of ATG16L (circles). Similarly, Large GFP-positive structures appear devoid of labelling for WIPI2 (Fig. 29Ci, circles), but some small punctate structures could be identified colocalising with this marker (Fig. 29Cii, arrow heads). Most small punctate structures

(Fig. 29C, arrow) and larger circular GFP-positive structures did not label for any of the autophagosomal markers.

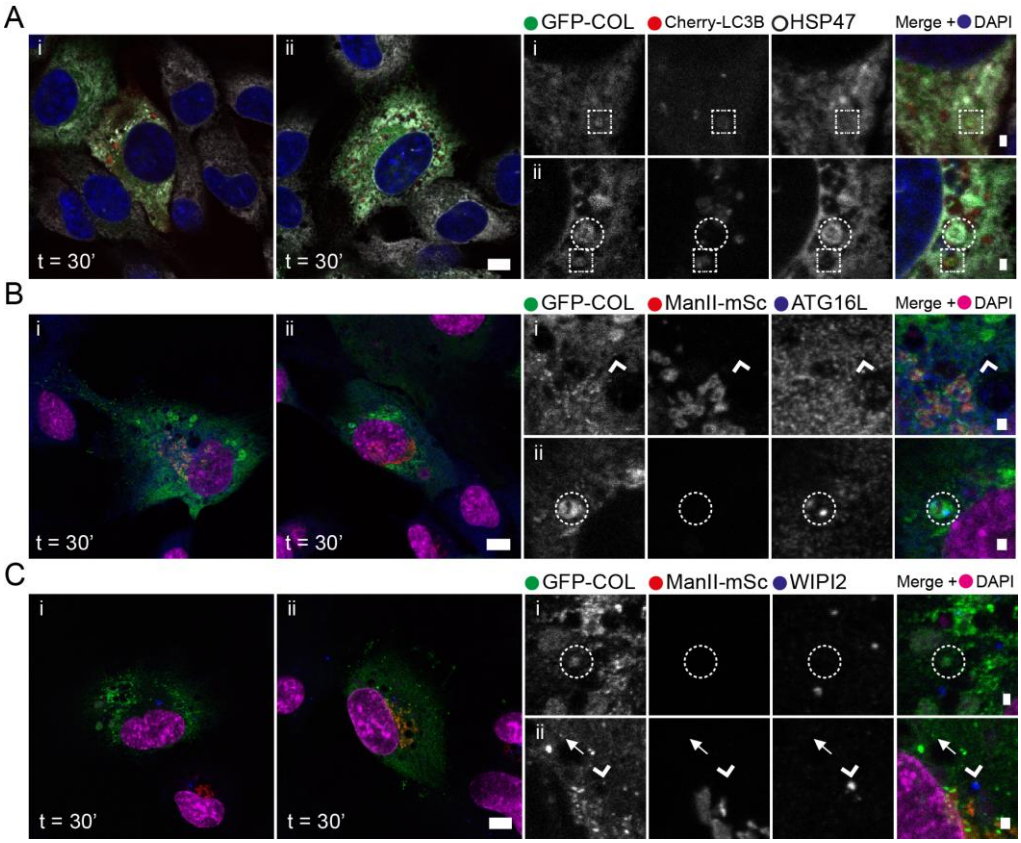


Fig. 29: Large GFP-COL structures are mostly negative for autophagosomal markers.

Confocal images of GFP-COL-RPE co-transfected with either the autophagosomal marker Cherry-LC3B (A) or ManII-mSc (B-C; red) 20 and 18 h after transfection, respectively, showing GFP-COL (green) and labelling for HSP47 is displayed in white, while autophagosomal markers ATG16L and WIPI2 are displayed in blue. Cells were fixed with PFA (except for C, which was fixed with methanol) after incubation with 500 µg·mL^{-1} ascorbate and 400 µM biotin for 30 min or just ascorbate when transfected with Cherry-LC3B. Panels show two examples of whole cells with corresponding enlargements showing the individual channels as greyscale, followed by the merge image including nuclear labelling with DAPI, imaged as separate channel in pseudo colour blue (A) or magenta (B-C). Circles indicate large GFP-positive structures not colocalizing with autophagosomal or endosomal markers. (A) Boxes indicate larger GFP-positive structures colocalizing with Cherry-LC3B. Arrowheads indicate small GFP-positive structures co-labelling for ATG16L or WIPI2 (B and C). Arrows highlight small GFP-positive puncta not co-labelling for autophagosomal markers. Scale bars, 10 µm; in enlargements, 1 µm and n ≥ 10. (McCaughey et al., 2019).

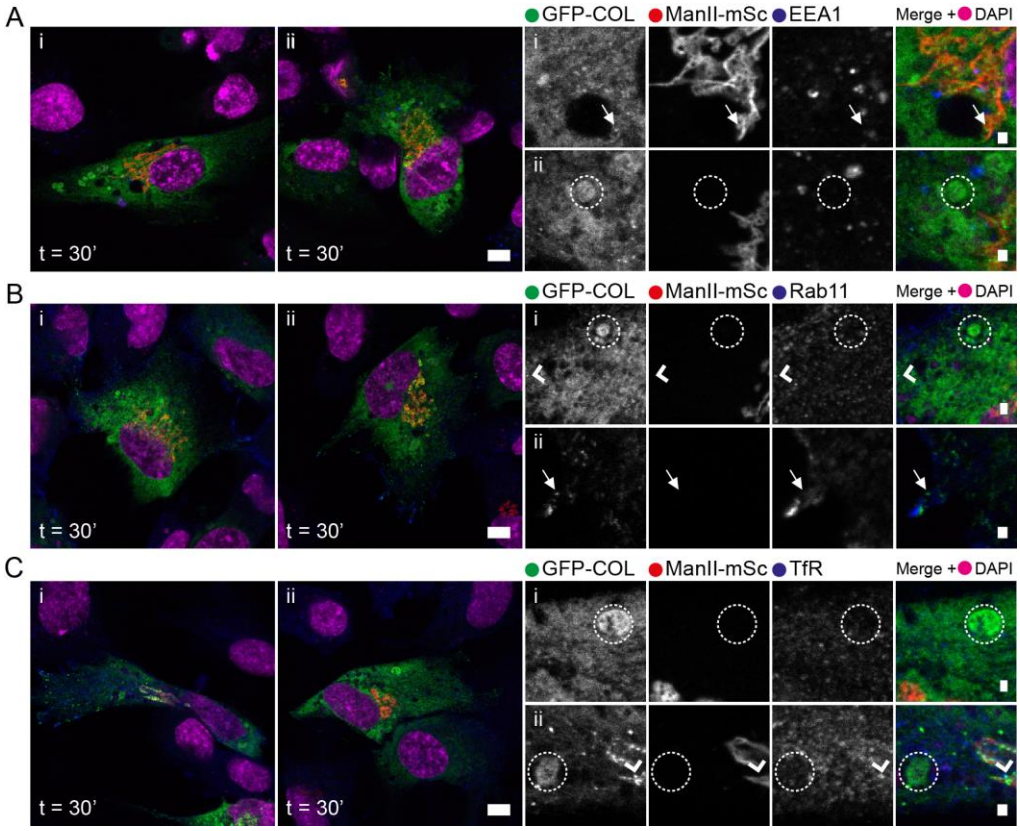


Fig. 30: Large GFP-COL structures are negative for endosomal markers.

Confocal images of GFP-COL-RPE 18 h after co-transfected with ManII-mSc (red) showing GFP-COL (green) and labelling endosomal markers EEA1 (A), Rab11 (B) or transferrin receptor (TfR; C) in blue, as well as nuclear labelling with DAPI, imaged as separate channel in pseudo colour magenta. Cells were fixed with PFA after incubation with 500 µg·mL^{-1} ascorbate and 400 µM biotin for 30 min. Panels show two examples of whole cells with corresponding enlargements showing the individual channels as greyscale, followed by the merge image. Circles indicate large GFP-positive structures not colocalizing with endosomal markers. Arrowheads indicate small GFP-positive structures co-labelling for Rab11 or TfR (B and C). Arrows highlight small GFP-positive puncta not co-labelling for endosomal markers. Scale bars, 10 µm; in enlargements, 1 µm and n ≥ 10. (McCaughey et al., 2019).

A similar result as observed for autophagosomal markers, was seen for endosomal markers (Fig. 30). Large GFP-positive structures, as well as most small punctate structures remained devoid of labelling for either EEA1, Rab11 or TfR (circles and arrows, respectively). Some small GFP-COL puncta did co-label with Rab11 or TfR (Fig. 30B-C, arrow heads).

KO and deletion experiments have previously shown that procollagen transport is COPII-dependent. This is also true for GFP-COL, since e.g. expression of a non-GTP-hydrolysable Sar1 (Sar1-H79G), which ultimately impairs COPII facilitated transport, in GFP-COL-RPE impairs transport to the Golgi apparatus (Fig. 31). Cells expressing Sar1-H79G can be identified by the dispersal of the Golgi markers ManII-mSc and giantin (Fig. 31ii and iii), while cells not expressing this construct, retain the Golgi ribbon

(Fig. 31i). No transport of GFP-COL can be observed in these cells, while cells not affected by Sar1-H79G show partial overlap with the Golgi markers (Fig. 31i).

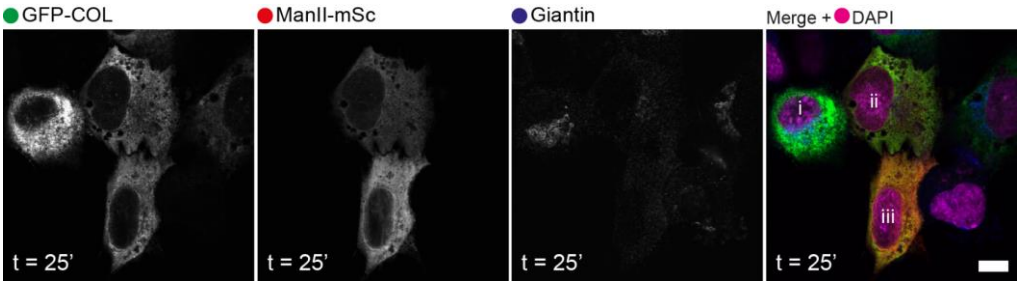


Fig. 31: GFP-COL trafficking is COPII dependent.

Example image of GFP-COL-RPE co-transfected with ManII-mSc and the GTP-locked form of Sar1 (Sar1-H79G), which results in a COPII block. Panels show the separate channels GFP-COL (green) and ManII-mSc (red), as well as antibody labelling with a *cis/medial*-Golgi marker giantin, in greyscale followed by the overlay image including nuclear labelling for DAPI (imaged as a separate channel in pseudo colour magenta). The time point indicates the incubation in presence of 500 $\mu g \cdot mL^{-1}$ ascorbate and 400 µM biotin before fixation with PFA and 7.5 h after transfection. Cells expressing Sar1-H79G show a scattered and disrupted Golgi apparatus, marked by ER labelling, compared with visible Golgi stacks in cells not affected by the COPII block (indicated by i). Scale bar, 10 µm and n ≥ 15 cells. (McCaughey et al., 2019).

To identify if any of the GFP-positive structures could be encapsulated by the COPII coat, cells were labelled for endogenous outer COPII coat marker Sec31A (Fig. 32). Interestingly, most small GFP-positive structures do not colocalise with Sec31A (Fig. 32A, arrows) and all large GFP-positive structures detected did not colocalise with Sec31A (Fig. 32A, circles). Some GFP-COL puncta were found to colocalise with the COPII marker (arrow heads). These puncta seem consistently smaller compared to those negative for Sec31A. Fig. 32B shows the analysis of images like A and gives a quantitative overview over the degree of overlap of GFP-positive circular objects with Sec31A in relation to their diameter. This data furthermore confirms that only smaller GFP-COL puncta (<700 µm diameter) show overlap with Sec31A, while large GFP-COL structures remain devoid for COPII labelling. Of note, in 11 of the 55 analysed cells obtained from n=4 independent experiments, despite showing accumulation of GFP-COL in the Golgi, no punctate or large GFP-COL objects were detected.

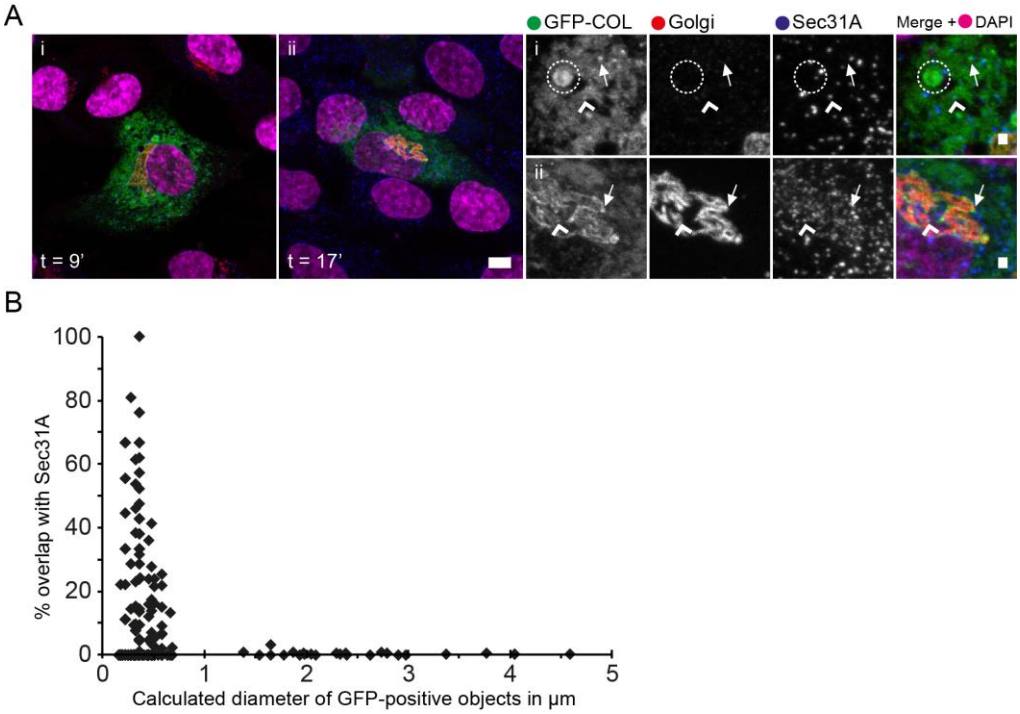


Fig. 32: Analysis of degree of overlap of GFP-positive circular objects depending on their diameter with Sec31A shows that large GFP-COL-positive structures are negative for Sec31A.

Maximum intensity projection confocal images of z stacks of whole cells (A). Duration in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin corresponds to duration of prior live imaging at approximately one image every 30 s until trafficking of GFP-COL to the Golgi was detectable by eye. 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 visualization of the Golgi apparatus. Labelling for Sec31A is shown in blue. Panels show two examples of whole cells with corresponding enlargements showing the individual channels as greyscale, followed by the merge image with nuclear DAPI labelling (imaged in a separate channel displayed in pseudo colour magenta). Large GFP-positive structures are highlighted by circles, while arrows indicate GFP-COL structures negative for Sec31A, and arrow heads indicate structures positive for Sec31A. Data obtained in (A) was used for the analysis of size distribution of GFP-positive objects relative to colocalization with the COPII marker Sec31A from confocal z stacks with Δz = 0.29 µm and a sufficient number of slices to represent whole cells (B). The x-axis shows the calculated object diameter of GFP-positive objects in micrometre, while the y-axis shows the calculated percentage overlap of GFP-positive objects with Sec31A. In 11 of the 55 analysed cells obtained from n=4 independent experiments, no punctate or large GFP-COL objects were detected. These cells only show a GFP-COL accumulation around or in the Golgi. The remaining 44 cells show a total number of 1,149 detected objects in the GFP channel. (McCaughey et al., 2019).

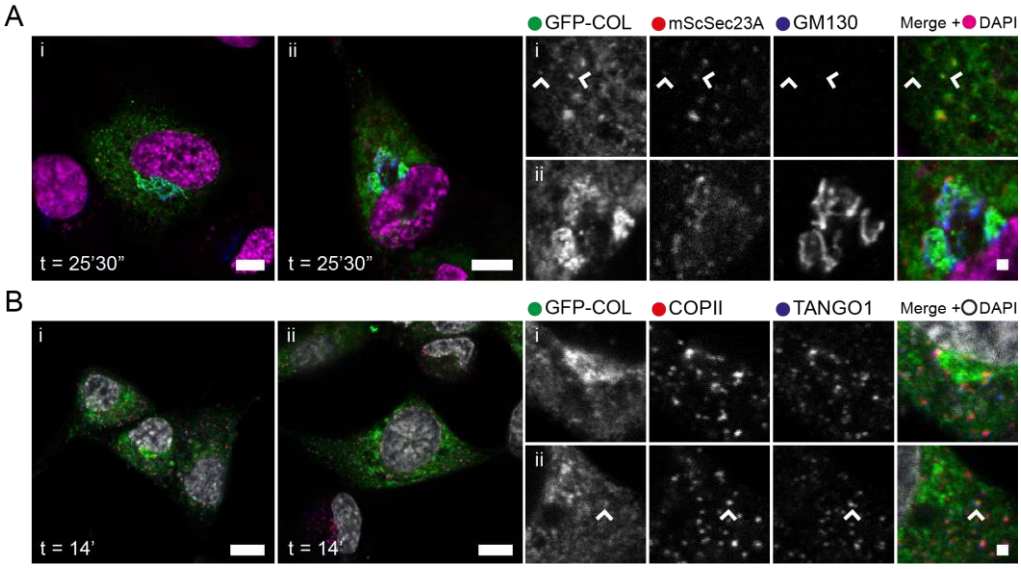


Fig. 33: GFP-COL can colocalise with TANGO1 and COPII in form of small puncta.

Confocal images of GFP-COL-RPE 20 h after co-transfected with mScSec23A (red) showing GFP-COL (green) and labelling for *cis*-Golgi marker GM130 (A) or TANGO1L (B) in blue, as well as nuclear labelling with DAPI, imaged as separate channel in pseudo colour magenta. Cells were fixed with PFA after incubation with 500 µg·mL^{-1} ascorbate and 400 µM biotin for the indicated timepoints. Panels show whole cells and enlargements with individual channels as greyscale, followed by the overlay image. Arrowheads highlight small GFP-positive puncta co-labelling for COPII markers and or TANGO1. COPII-labelling in B represents mScSec23A and Sec31A antibody labelling in the same channel to show whole COPII populations. Scale bars are 10 and 1 µm; n ≥ 10 cells.

To be able to better investigate the dynamic interplay between COPII and GFP-COL a new genetic construct containing the ER-hook was engineered to separately express COPII marker Sec23A N-terminally fused to mScarlet-i (mScSec23A). Co-expression of this construct in GFP-COL-RPE can be seen in Fig. 33. After 25 min incubation with ascorbate and biotin GFP-COL can be observed concentrating in the Golgi area marked by *cis*-Golgi marker GM130 (Fig. 33A). mScSec23A is found throughout the cell periphery and localising in proximity to labelling for the Golgi in form of punctate structures. Indicating that the new construct is functional as a COPII marker and for use in the RUSH system. Some of these mScSec23A puncta were found to colocalise with GFP-COL (Fig. 33Ai and Bii, arrow heads). The latter can also be found to colocalise with the ERES marker TANGO1 (Fig. 33Bii, arrow heads).

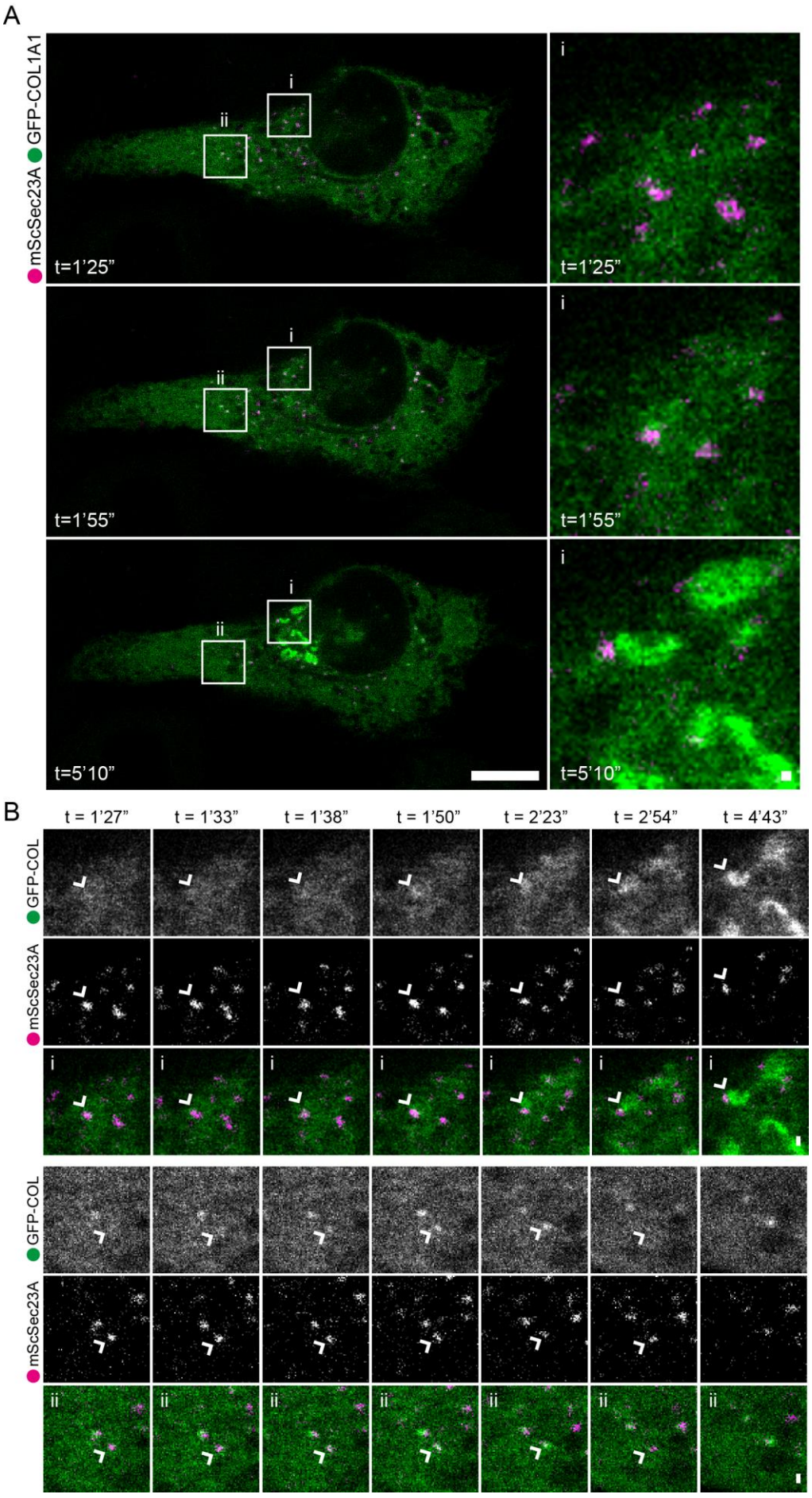


Fig. 34: GFP-COL accumulates in ERES prior to ER-Golgi trafficking.

Still images from confocal live-cell imaging GFP-COL-RPE) co-transfected with the inner layer COPII marker mScSec23A (magenta) (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V9.mp4). Acquisition at one frame every 2.79 s. Time points indicate time in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin 20 h after transfection. Large panels show the entire cell with selected enlargements of ERES in proximity to the Golgi (i) and the in the cell periphery (ii) with enlarged images of (i) to the right (A). The zoomed in areas of interest from A (also indicated by the square) with the separate channels in greyscale, followed by the overlay image are shown in (B). Arrowheads indicate a GFP-COL structure colocalising with mScSec23A. Scale bars, 10 µm; in enlargements, 0.25 µm; n = 9. (A) (McCaughey et al., 2019).

Live cell imaging of a RUSH assay using mScSec23A in GFP-COL-RPE reveals that GFP-COL accumulates in ERES marked by mScSec23A in quick succession after addition of ascorbate and biotin at 1 min 25 sec (Fig. 34). This accumulation can be observed in form of GFP-positive puncta appearing both at ERES in proximity to where GFP-COL accumulates in the Golgi after 5 min (Fig. 34A-Bi)., as well as in peripheral ERES (Fig. 34A-Bii). Throughout the video the COPII marker does not appear to be very dynamic, except for the general movement of the ER itself. Accumulation of GFP-COL in the Golgi area again occurs in a gradual manner and no large or small punctate structures can be identified moving from peripheral ERES towards the Golgi area. Interestingly, signal intensity for GFP-COL increases at peripheral ERES at e.g. 1 min 50 sec, prior to the signal decreasing again, while labelling for mScSec23A remains stable (Fig. 34Bii, 2 min 54 sec).

To more accurately analyse the diameters of GFP-COL and COPII-positive puncta, cells as in Fig. 34 were fixed at time points after incubation with ascorbate and biotin and the signal for both GFP and COPII was enhanced by labelling with antibodies against GFP and Sec24C measured in the same channel as GFP-COL and mScSec23A to allow super-resolution microscopy using stimulated emission depletion (STED; Fig. 35). Two major types of GFP- and COPII-positive structures could be observed by this method: slightly larger structures composed of clusters of COPII-positive objects with estimated maximum diameters using the full width at half-maximum (FWHM) values of about 220 - 860 nm for COPII and 350 – 830 nm for GFP-positive objects (Fig. 35Ci, vii and x), as well as small GFP- and COPII puncta with FWHM values of 130 – 320 nm and 100 – 220 nm, respectively (Fig. 35Cii-v and ix). Large GFP-positive objects with diameters of about 1 – 2.3 µm remained negative for COPII as observed before (Fig. 35Cvi and viii). Interestingly, COPII structures consistently displayed smaller estimated maximum diameters than the GFP-COL puncta they colocalise with (which also never exceeded 350 nm in diameter).

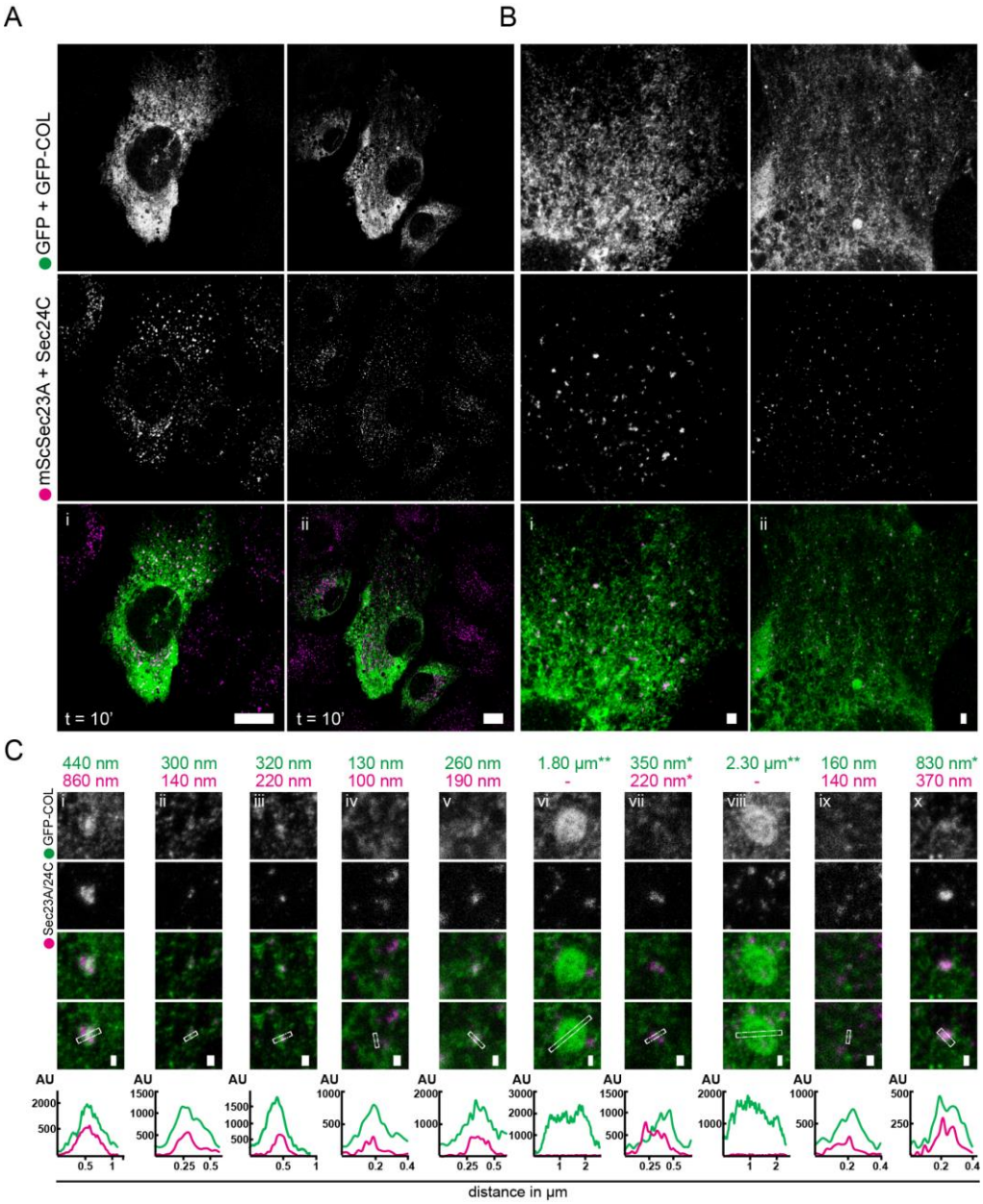


Fig. 35: GFP-COL puncta colocalising with COPII are <350 nm in diameter.

Confocal images of whole cells (A) and super-resolution gated STED images of a selected area (B) obtained using STED microscopy of GFP-COL-RPE co-transfected with the inner layer COPII marker mScSec23A fixed 10 min after incubation in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin. Cells were co-labelled with an antibody against GFP in the same channel as GFP-COL (green) and with an antibody against Sec24C in the same channel as mScSec23A (magenta). Enlargements of small GFP-positive puncta colocalizing with COPII and large GFP-positive structures extracted from images as displayed in B are shown in C. Panels show the different channels in greyscale, followed by the overlay image and the overlay containing the line with a width of 5 (images ii–v, vii, and ix) or 10 pixels (images i, vi, viii, and x) drawn through the object of interest to generate the corresponding line graphs, as shown below. Line graphs display the distance in micrometres on the x-axis and the signal intensity as an arbitrary unit (AU) on the y-axis. The FWHM values corresponding to the estimated maximum object diameters in nanometres or micrometres are displayed above the images for the corresponding line graphs fitted with a Gaussian curve (in green for GFP + GFP–COL and in magenta for mScSec23A + Sec24C). When two curves were used to fit the graphs, the sum of the FWHM was used as an

estimate (indicated by asterisk). When Gaussian fitting was not possible, the displayed estimated diameter value was measured via ImageJ (two asterisks). If no peak in the magenta channel could be identified, values were left blank (indicated by dash). Scale bar, 10 µm (A), 1 µm (B) and 0.25 µm (C); n = 14. (McCaughey et al., 2019).

In addition, the intensity maxima visible in the corresponding line graphs for both channels always appear slightly shifted from each other (Fig. 35Cii-v and ix). This implies that COPII object found to colocalise with GFP-COL puncta, are not evenly positioned around GFP-COL and show no evidence of forming a coat around such a structure.

In order to investigate if ER-to-Golgi transport of GFP-COL is dependent on an intact MT network, cells were treated with NZ. NZ treatment leads to the collapse of the MT network that is essential for long distance trafficking within cells. Fig. 36 shows the effect of treatment of RPE-1 with NZ for 0, 60 and 120 min. Prior to NZ treatment the MT network is intact and α-tubulin labelling can be observed in form of multiple tubules throughout the cytosol stretching from the perinuclear region towards the cell periphery (Fig. 36i-ii). After both 60 and 120 min of incubation with NZ α-tubulin labelling appears diffuse and devoid of tubules (Fig. 36iii-iv).

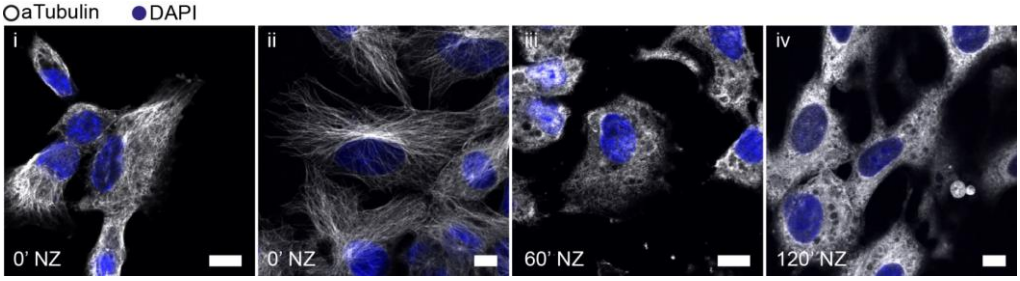


Fig. 36: The MT network collapses upon treatment with nocodazole (NZ).

Confocal images of GFP-COL-RPE 18.5 h after co-transfection with the *trans*-Golgi marker ST-Cherry (both not shown) and labelled with an antibody against α-tubulin (greyscale) and nuclear DAPI staining in blue. Time points indicate min after addition of NZ before fixation with methanol. The channel for α-tubulin was brightness and contrast enhanced using ImageJ's autocorrection. A total of n = 3 was acquired. Scale bar, 10 µm. (McCaughey et al., 2019).

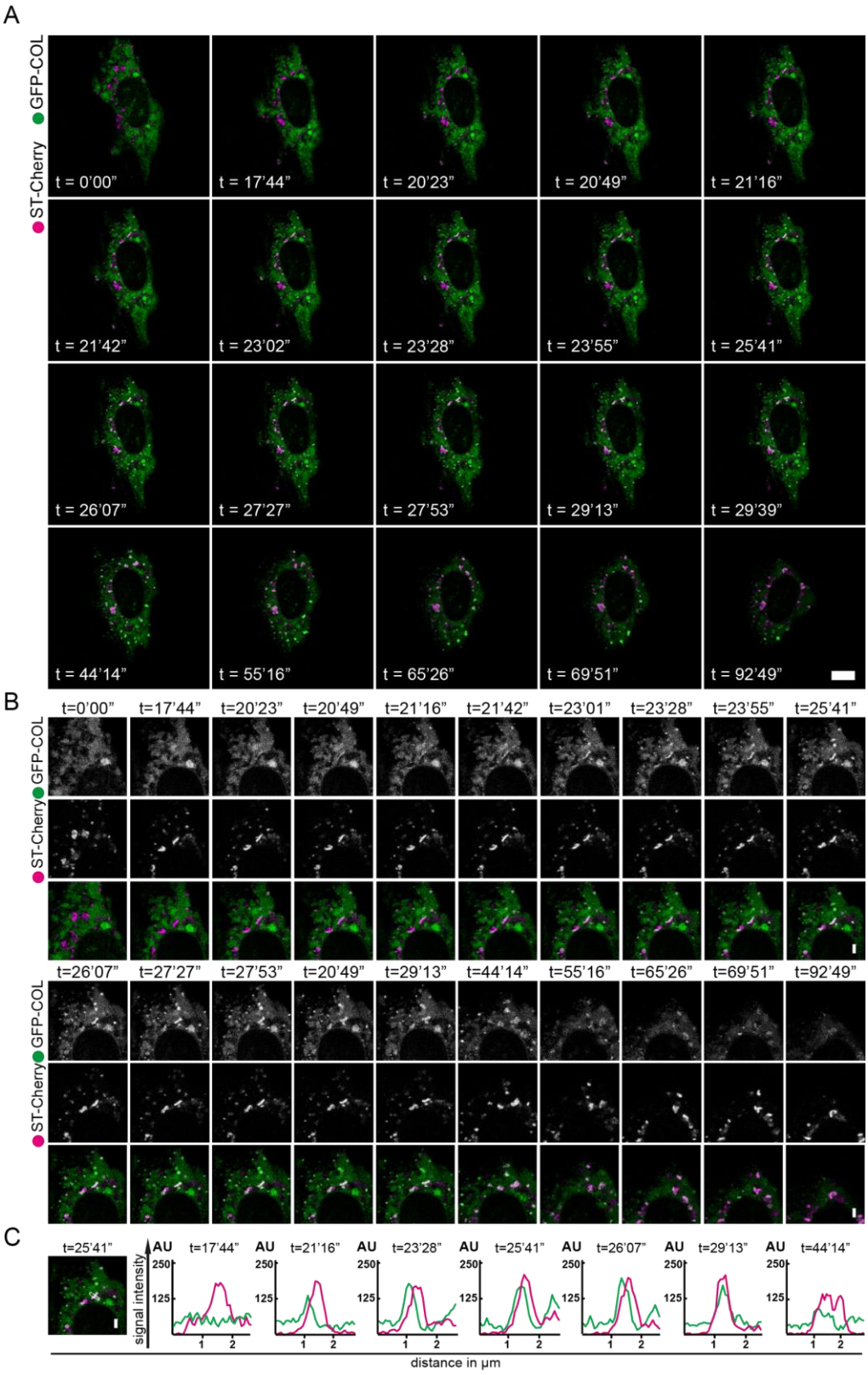


Fig. 37: ER-Golgi transport of GFP-COL does not rely on an intact MT network.

Still images from confocal live-cell imaging of GFP-COL-RPE with GFP-COL in green, co-transfected with the *trans*-Golgi marker ST-Cherry (magenta) (video available at https://static-movie-usa.glencoesoftware.com/source/10.1083/529/e5b860ff3ff151c33cea78fa6c5b27382c1799fb/JCB_201806035_V10.mp4). Acquisition at one frame every 26 s. Cells were incubated in presence of NZ for 60 min before imaging. Time points indicate time in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin. Large panels show the entire cell (A). Small panels show a zoomed

in area of interest with the green channel and magenta channel in greyscale followed by the overlay (B). Example image of an enlarged overlay with a 5-pixel-wide line drawn through the Golgi followed by corresponding line graphs of selected time points showing the signal intensity (y-axis) in arbitrary units (AU) for GFP-COL (green) and ST-Cherry (magenta) for the corresponding line (C). The x-axis shows the distance in micrometres. Scale bar, 10 µm (A) and 1 µm (B-C); n = 4. (McCaughey et al., 2019).

Fig. 37 shows image stills of a RUSH assay of NZ treated GFP-COL-RPE expressing ST-Cherry. Despite collapse of the MT network, GFP-COL is still able to accumulate around the scattered Golgi structures. Golgi labelling occurs in form of fragmented Golgi mini-stacks distributed throughout the cell. Upon incubation with ascorbate and biotin gradual accumulation of GFP-COL occurs in the Golgi fragments as seen before after additional 20 min in presence of NZ, ascorbate and biotin. Overall signal intensity of GFP-COL in the ER decreases over the time course of the video from 30 – 93 min post-release from the ER. Some GFP-positive circular structures, negative for Golgi, remain present throughout the video. Furthermore, an accumulation of GFP-COL clusters in the cell periphery can be observed with clusters merging over time to form bigger clusters. The amount of small GFP-COL puncta visible moving from the Golgi elements to the cell periphery is much lower in NZ treated cells compared to non-NZ treated ones. This implies that sorting within the Golgi is still possible in this state. Since the overall amount of GFP-COL decreases over time, it can be assumed that the pathway from the ER up to secretion from the cell is still intact.

GFP-trap experiments with subsequent proteomics analysis via mass spectrometry (MS) furthermore show that GFP-COL can bind and pull down endogenous pro-α2(I) in GFP-COL-RPE (Table 7). During this experiment pro-α2(I) could only be detected in pulldowns from GFP-COL-RPE lysates 15 min in presence of 500 µg·mL^{-1} ascorbate, but not in samples without addition of ascorbate, nor in control GFP-RPE cells with or without ascorbate. To test whether the observation of GFP-COL accumulation in the Golgi in a gradual manner in absence of large carriers also applies for the physiological type I procollagen composed of heterotrimers, different pro-α2(I) fusion proteins were tested.

GFP- or mApple-tags N-terminally fused to mouse pro-α2(I) downstream of a signal sequence and upstream of the COL-domain (replacing the propeptide; Fig. 38A and C), as well as replacing exons two and three but retaining the N-propeptide minor triple helix and N-terminal cleavage site upstream of the COL-domain (in case of mApple-COL1A2L; Fig. 38B) (Omari et al., 2018), were tested in RPE-1. These pro-α2(I) fusions localise to the ER and concentrate in the Golgi area marked by GM130 in RPE at steady state (Fig. 38).

Table 7: Proteomics data of collagens and GFP of a GFP-pulldown performed on control GFP-RPE and GFP-COL-RPE.

Cells were incubated without or in presence of 500 µg·mL^{-1} ascorbate for 15 min after 24 h ascorbate flush before cell lysis. Columns display the accession number, protein description, score, coverage, the number of proteins, and unique peptides and peptides, as well as the number of peptide spectrum matches (PSMs) and the peak area sorted by number of PSMs. Listed are only the proteins of interest: GFP and different forms of detected COL1A1, as well as COL1A2. N = 1. (McCaughey et al., 2019).

Sample	Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	Area
GFP-RPE									
0 min asc	tr	GFP	6405	72.38	2	39	39	2130	5.08E+10
	Q6LAN8	Collagen type I alpha 1 (Fragment)	18	2.71	5	2	2	6	5.49E+05
		Collagen alpha-2(I) chain				*Not detected*			
15 min asc	tr	GFP	6713	88.28	1	40	40	2192	4.14E+10
	Q6LAN8	Collagen type I alpha 1 (Fragment)	35	5.52	5	4	4	18	1.38E+07
		Collagen alpha-2(I) chain				*Not detected*			
Sample	**Accession**	**Description**	**Score**	**Coverage**	**# Proteins**	**# Unique Peptides**	**# Peptides**	**# PSMs**	**Area**
GFP-COL-RPE									
0 min asc	tr	GFP	3081	90.38	2	27	27	970	9.21E+09
	Q6LAN8	Collagen type I alpha 1 (Fragment)	722	14.5	6	10	10	382	3.52E+09
		Collagen alpha-2(I) chain				*Not detected*			
15 min asc	tr	GFP	1856	90.79	1	21	21	609	3.83E+09
	P02452	Collagen alpha-1(I) chain	714	21.31	6	10	18	397	2.78E+09
	D3DTX7	Collagen, type I, alpha 1, isoform	71	18.08	2	1	9	39	5.85E+06
	A0A087WTA8	Collagen alpha-2(I) chain	21	5.28	3	5	5	8	3.38E+06

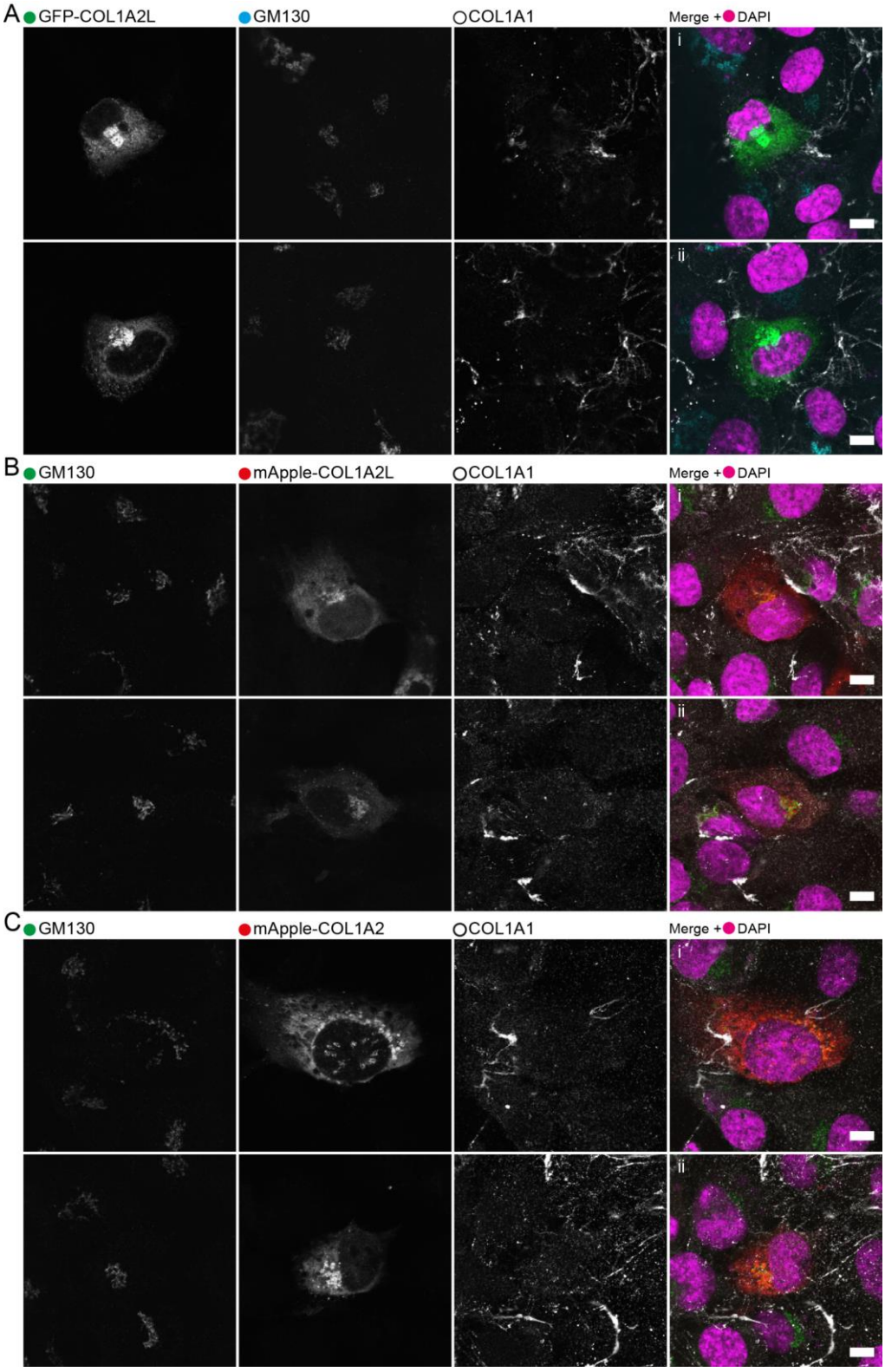


Fig. 38: Fluorescent protein fusions of proα2(I) localise to the ER and concentrate in the Golgi area at steady state in RPE.

Confocal images of RPE-1 cells transiently expressing different N-terminal mouse COL1A2 fusions with GFP (A) or mApple (B-C)(Omari et al., 2018). COL1A2L indicates that mApple replaced exons two and three retaining N-propeptide minor triple helix and N-terminal cleavage site, while other fusion proteins lack the N-terminal domain but have a signal sequence inserted prior to the tag. Panels show individual channels in greyscale, followed by merge images including nuclear labelling for DAPI (imaged in a separate channel and displayed in pseudo

colour magenta). Antibody labelling for COL1A1 is shown in greyscale, while GFP-fusions are shown in green and mApple constructs in red. *Cis*-Golgi marker GM130 is shown in cyan (A) or green (B-C). Scale bars are 10 µm; n ≥ 10 cells.

Endogenous pro-α1(I)/COL1A1 in these cells is mostly visible in the extracellular space and does not show obvious co-localisation with pro-α2(I) fusions in the cell.

To test fusion proteins of pro-α2(I) in GFP-COL during trafficking experiments a new mSc-tagged human pro-α2(I) (mScCOL1A2) was engineered. This procollagen chain had a mSc-tag inserted downstream of the N-propeptide and upstream of the triple helical COL-domain of human pro-α2(I). In theory, retention of the SBP-tagged GFP-COL in the ER should also retain mScCOL1A2 in the ER when bound to GFP-COL. Release from the ER by addition of ascorbate and biotin should trigger synchronised transport of both procollagen chains.

Binding of GFP-COL to mScCOL1A2 can be observed in streptavidin pulldowns from GFP-COL-RPE stably expressing mScCOL1A2 (GFP-COL1A1/mScCOL1A2-RPE) (Fig. 39). The streptavidin beads bind the SBP-tag from the GFP-COL and pulldown using streptavidin beads shows a similar pulldown efficiency compared to pulldowns using anti-GFP-beads (Fig. 39A). When blotting for SBP both the bound fractions from GFP- and the eluted fractions from streptavidin pulldowns show the expected bands when using lysates from GFP-COL-RPE. The band at about 60 kDa corresponds to the cleaved N-propeptide containing SBP-mGFP, while the band at ~180 kDa corresponds to full length procollagen-SBP-mGFP-COL and the ~120 kDa band to the full length procollagen-SBP-mGFP-COL without the C-terminus. No bands can be detected with either pulldown when blotting for SBP and using lysates from control cells only expressing GFP. To test whether GFP-COL can bind mScCOL1A2 a streptavidin pulldown was performed in cells stably expressing both GFP-COL and mScCOL1A2 (GFP-COL1A1/mScCOL1A2-RPE) and cells expressing GFP-COL and only mSc were used as controls (Fig. 39B-C). A band for COL1A2 can only be observed in cells expressing both procollagen fusion constructs in eluted fractions and not in control cells (Fig. 39B). A band with the same molecular weight of about ~160 kDa corresponding to full length mSc-tagged pro-α2(I) can be seen when blotting for RFP, which also recognises mSc, but is absent in control cells (Fig. 39C). GFP or RFP-traps could not be used for these experiments, since there was a cross-reaction of the RFP-beads also recognising GFP and vice versa.

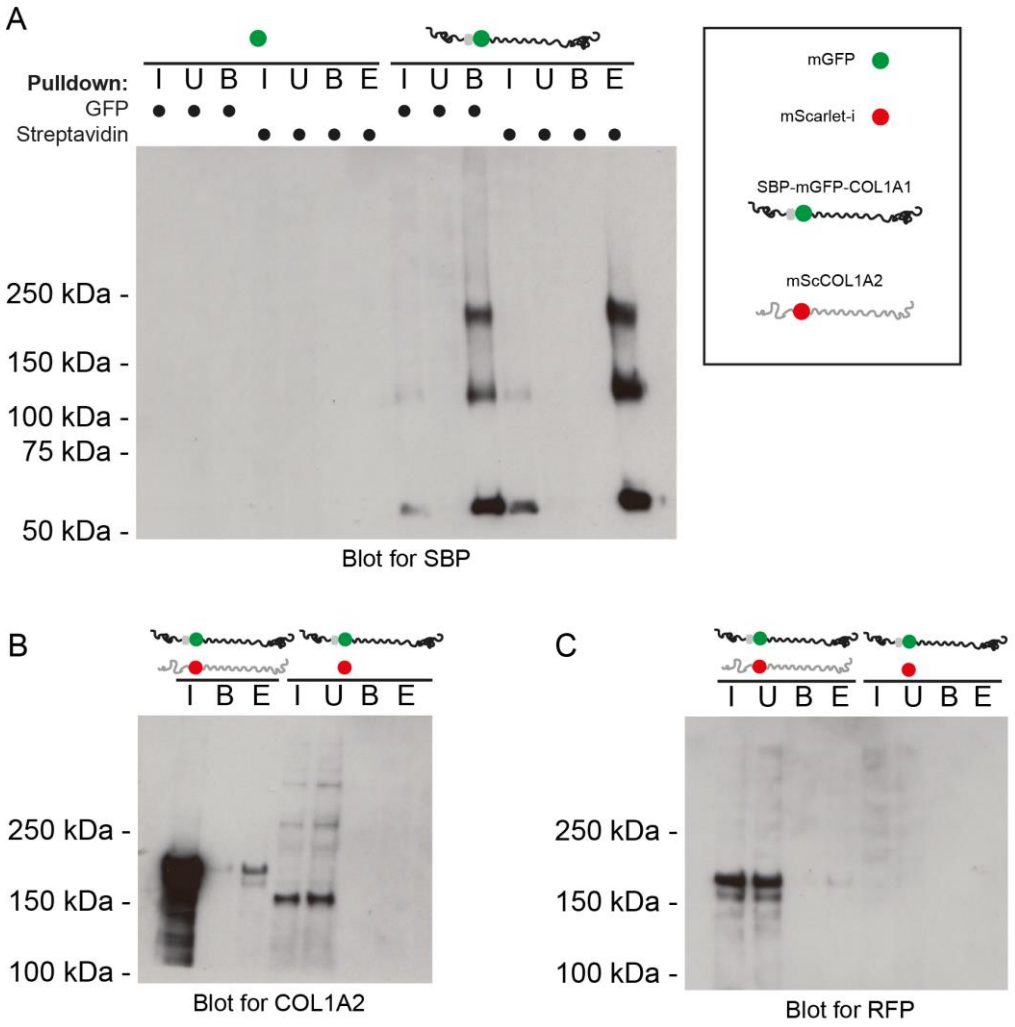


Fig. 39: GFP-COL can pulldown mScCOL1A2.

Immunoblots from GFP- (A) or streptavidin-pulldown experiments (A-C) blotted for SBP (A), COL1A2 (B) and RFP (which recognises mSc; C). Cell lysates used for the pulldown experiments are indicated in the legend and originated from GFP-RPE (only expressing GFP), GFP-COL-RPE (expressing SBP-mGFP-COL), GFP-COL1A1/mScCOL1A2-RPE expressing both GFP-COL and mScCOL1A2 stably, or GFP-COL-RPE stably expressing mScarlet-i. Lanes show input (I), unbound (U), bound (B) and for streptavidin pulldowns fractions eluted using 400 mM biotin (E). n = 1 (but could be reproduces using GFP- and RFP-pulldowns).

Subsequently, a RUSH assay was performed with GFP-COL-RPE transiently expressing mScCOL1A2 and co-transfected with an ER-hook and separate *medial*-Golgi marker mTag2BFP-tagged ManII (ManII-BFP) as seen in Fig. 40. At timepoint zero both GFP-COL and mScCOL1A2 colocalise partially in the ER. While the Golgi area marked by ManII-BFP is devoid of signal for GFP-COL, some mScCOL1A2 can already be detected in the Golgi prior to release of GFP-COL from the ER-hook (Fig. 40A).

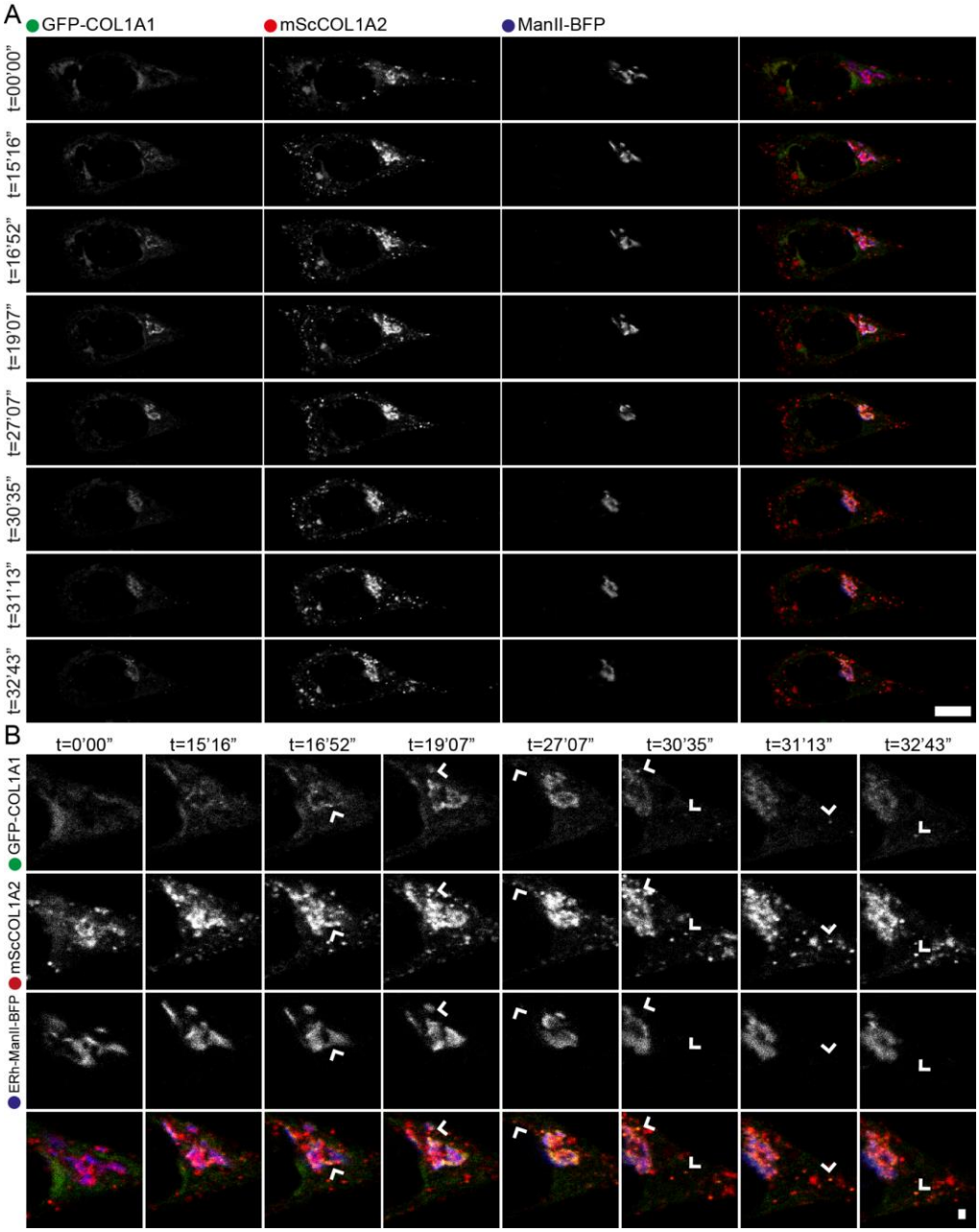


Fig. 40: GFP-COL and mScCOL1A2 colocalise in form of small puncta in RPE.

Confocal still images of GFP-COL-RPE (GFP-COL in green) transiently transfected with mScCOL1A2 (in red) and StrKDEL-IRES-ManII-BFP (ManII-BFP; in blue). Timepoints indicate time in presence of 500 µg·mL^{-1} ascorbate and 400 µM biotin. Whole cell shown with individual channels in greyscale, followed by the overlay image (A). Enlargements of the Golgi with individual channels in greyscale, followed by the overlay image (B). Arrow heads indicate small puncta positive for both GFP-COL and mScCOL1A2. Scale bars 10 µm (A) and 1 µm (B); n=3 cells.

As observed before, gradual accumulation of GFP-COL became visible at about 15 min and filling of the Golgi at about 20 min post-addition of ascorbate and biotin. Compared to timepoint zero, the signal intensity for mScCOL1A2 increases in the Golgi area at about 15 min post-ascorbate and biotin. Throughout the course of the experiment several small mScCOL1A2 structures are visible throughout the cell and several of

these puncta appear to colocalise with small GFP-COL puncta in the cell periphery and in proximity of the Golgi (Fig. 40B, arrow heads). Transport of mScCOL1A2 from the ER-to-Golgi was difficult to observe, due to a population of mScCOL1A2 already in the Golgi at the start of the video.

Upon fixation of GFP-COL1A1/mScCOL1A2-RPE at about 18 min ascorbate and biotin during RUSH assays (Fig. 41A) small puncta positive for both GFP-COL and mScCOL1A2 could be observed colocalising with COPII marker Sec31A (Fig. 41B, arrow heads). Some slightly larger structures positive for GFP-COL and mScCOL1A2 were negative for COPII (Fig. 41B, circle) and several small puncta positive for both procollagen fusions, but negative for Sec31A could be detected (Fig. 41B, arrows).

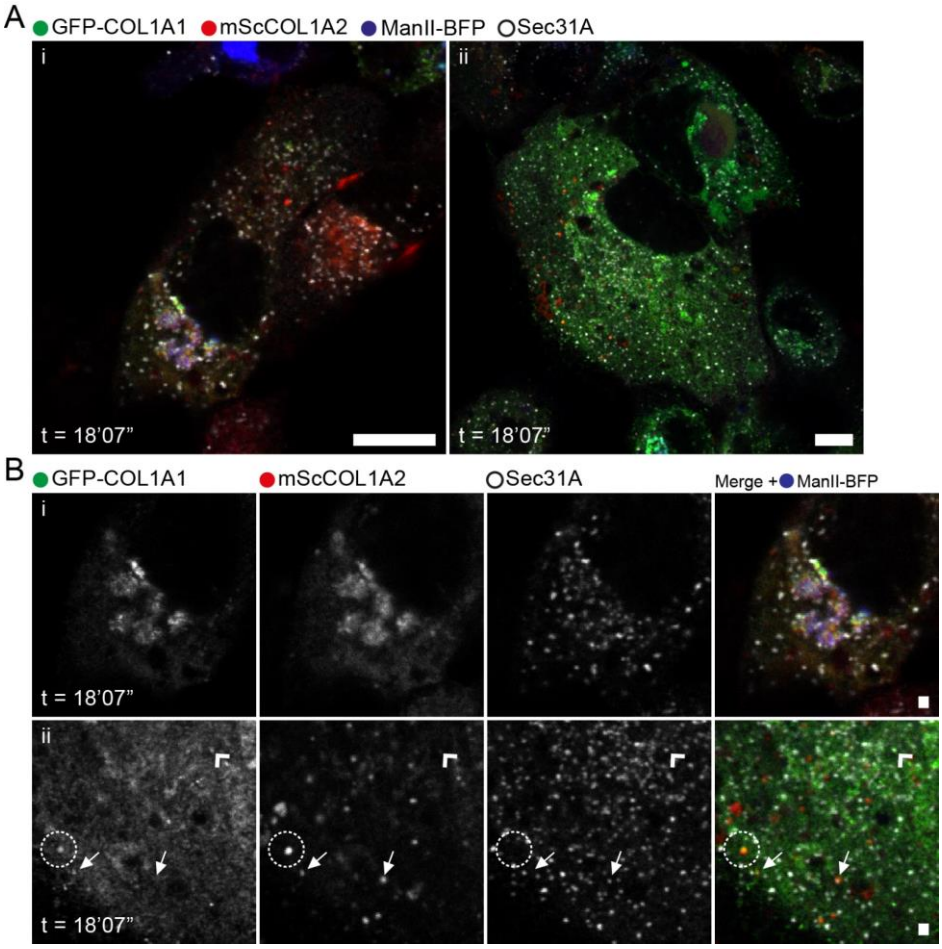


Fig. 41: GFP-COL1A1/mScCOL1A2-RPE show both collagens colocalise as small puncta, occasionally positive for Sec31A.

Confocal images of GFP-COL1A1/mScCOL1A2-RPE co-transfected with StrKDEL-IRES-ManII-BFP (ManII-BFP) 18 min and 7 s after incubation with 500 µg·mL^{-1} ascorbate and 400 µM biotin. Whole cells in (A) and enlargements with individual channels in greyscale, followed by the overlay image including ManII-BFP in blue (B). GFP-COL is displayed in green, mScCOL1A2 in red and COPII marker Sec31A in greyscale. Puncta positive for GFP-COL, mScCOL1A2 and COPII are indicated by arrowheads. Small punctate structures only positive for both procollagen chains are highlighted by arrows, while a slightly larger structure also negative for COPII is indicated by a circle. Scale bars are 10 µm (A) and 1 µm (B); n≥10 cells.

4.3 Discussion

The data demonstrate that GFP-COL is a suitable reporter for visualisation of pro-α1(I) transport from the ER to the Golgi. At steady state GFP-COL localises mostly to the ER and co-localises with HSP47. When transiently expressed in IMR-90 fibroblasts GFP-COL can be observed to traffic in an ascorbate-dependent manner to the Golgi. Notably, this transport process does not show any obvious large structures positive for GFP-COL moving towards the Golgi and only some smaller GFP-COL puncta can be seen potentially translocating from the cell periphery towards the Golgi area. Similarly, GFP-COL secretion in GFP-COL-RPE also occurs in an ascorbate-dependent manner as the ratio of COL1A1 abundance determined by immunoblots in the medium compared to that intracellular protein levels increases upon treatment with ascorbate. Since hydroxylation by P4H in the ER, which facilitates trimer stabilisation, is enhanced by the presence of ascorbate, these results indicate that GFP-COL is transported and secreted as a trimer. Despite low expression levels of endogenous pro-α2(I) in RPE-1 cells, GFP-pulldown experiments with subsequent proteomics analysis show that GFP-COL can bind its physiological binding partner. In addition, mScCOL1A2 can also be pulled down by GFP-COL. While it is likely that most of the observed GFP-COL forms homotrimers (Jimenez et al., 1977; Kay, 1986; Uitto, 1979), at least some population of GFP-COL trafficked consists of heterotrimers.

Using the RUSH system synchronised ER-to-Golgi transport of GFP-COL was achieved in a biotin and ascorbate-dependent manner. Here, accumulation of GFP-COL in the Golgi was observed after 2 min – 45 min in presence of ascorbate and biotin with the majority of cells showing filling of the Golgi at 15-30 min. This is consistent with temperature block experiments performed with type I procollagen which also resulted in an accumulation of pro(I)-α1 in the Golgi after about 16 min (Subramanian et al., 2019). Experiments with NZ treatment also show that GFP-COL transport to the Golgi occurs independent of and intact MT network and functional maturation of the Golgi (Fourriere et al., 2016). Transport of GFP-COL to Golgi stacks is very similar to observations made with smaller cargo molecules like transport of tumour necrosis factor TNF-SBP-EGFP (Fourriere et al., 2016). Juxta-Golgi accumulation of GFP-COL appears at a similar incubation time after addition of ascorbate and biotin to cells not pre-exposed to NZ, but GFP-COL abundance in the ER and especially in form of clusters in the cell periphery remains high for longer in NZ treated cells. This suggests that post-ER-to-Golgi transport is delayed if the MT network is impaired. Furthermore, this indicates that the observed pathway might not be restricted to large cargo proteins such as type I procollagen.

ER-to-Golgi transport in GFP-COL-RPE consistently appears to follow the same route: At the beginning of the time course GFP-COL localises to the ER and the Golgi area is devoid of signal for GFP, then GFP-COL begins to accumulate at the ER-Golgi interface (at ERES), prior to gradual filling of the Golgi. Small punctate GFP-COL

objects can be observed mostly moving from the Golgi to the cell periphery. After 45 min to 1 h post-addition of ascorbate and biotin the signal for GFP-COL in the Golgi decreases, indicating successful secretion from the cell. The loss of GFP-signal over time could be the outcome of successful cleavage of the N-propeptide-SBP-mGFP domain of GFP-COL and subsequent degradation, quenching the signal and or secretion into the extracellular space.

The SBP-mGFP-tag was introduced upstream of the naturally occurring N-terminal proteolytic cleavage site (located after the N-terminal pro-peptide). The GFP-COL construct is designed to retain the cleavage site such that extracellular procollagen itself does not become fluorescent. This enabled more effective visualisation of procollagen trafficking. Data from Nicola Stevenson of GFP-COL-RPE giantin-KO cells where the proteolytic cleavage of the N-propeptide-SBP-mGFP in GFP-COL-RPE is impaired, shows GFP-positive extracellular collagen fibrils (Stevenson et al., 2020). This implies that the construct does not impair fibril formation and can be transported effectively.

Only occasionally can a few small GFP-COL objects be observed moving from the cell periphery towards the Golgi. While this observation is more evident in BJ-5ta transiently expressing GFP-COL, the abundance of these structures does not seem sufficient to result in the drastic increase of GFP-COL signal in the Golgi observed. This implies that there are multiple possible routes of procollagen transport from the ER to the Golgi – transport via small structures from the cell periphery, and a more gradual dynamic transition without obvious visible carriers. Abundant KO and depletion experiments of COPII components shows that procollagen transport is COPII-dependent (Boyadjiev et al., 2006; Garbes et al., 2015; Lang et al., 2006; Moosa et al., 2015; Sarmah et al., 2010; Schmidt et al., 2013; Stephens and Pepperkok, 2002; Townley et al., 2008; Townley et al., 2012). Similarly, GFP-COL transport to the Golgi is impaired in cells expressing GTP-restricted Sar1-H79G, which blocks COPII (Aridor et al., 1995). While most GFP-COL puncta are negative for COPII, some do show colocalization with markers of the COPII coat. Furthermore, GFP-COL can be seen to accumulate at ERES marked by mScSec23A both in proximity to the Golgi and in the cell periphery as fast as 1.5 min after addition of ascorbate and biotin. Accumulation of GFP-COL in the Golgi is visible directly from those juxtaposed ERES. The signal for GFP-COL in peripheral ERES, however, subsides after about 1 min after first being visible. This suggests that ERES that are not in close proximity to the Golgi might not support an effective pathway for transport of large cargo proteins, such as type I procollagen. Furthermore, COPII puncta colocalising with GFP-COL analysed by super-resolution STED microscopy have diameters of <350 nm. Cells expressing mScCOL1A2 in addition to GFP-COL also show both procollagen chains colocalising with COPII in form of small (<350 nm) puncta. Furthermore, the relative localisation of COPII and GFP-COL objects at ERES, as well as the fact that measured diameters for GFP-COL

objects are consistently larger than those for COPII suggests that GFP-COL might not be fully encapsulated by the COPII coat, but rather partially coated. Therefore, a more dynamic and transitional pathway for COPII-dependent ER-to-Golgi transport of procollagen may be used, contrasting the idea of entirely COPII coated, and especially large, procollagen carriers (Gorur et al., 2017; Jin et al., 2012; 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).

To rule out that puncta positive for both GFP-COL and mScCOL1a2 consist of both proteins being transported in the same vesicular structure as monomers or in form of cleaved variants containing only GFP and mSc, fluorescent resonance energy transfer (FRET) assays were attempted. To achieve efficient FRET between GFP and mSc, both proteins need to be in the correct rotational alignment to each other and within ≤10 nm distance. The latter would only be the case, if both GFP and mSc are brought in sufficient proximity for FRET via heterodimerisation of GFP-COL with mScCOL1A2. However, FRET between to the two probes could not reproducibly be detected.

Interestingly, while most GFP-COL puncta in the cell periphery and close to the Golgi were negative for HSP47, some structures appeared positive for the collagen chaperone while also being negative for ER membrane. This agrees with the *in vitro* data of HSP47 preferentially binding 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).

Intriguingly, ER-to-Golgi transport of GFP-COL occurs in absence of large carriers even when imaged at fast rates like 1 fps. About 5% of the cells imaged show larger circular structures positive for GFP-COL. Live imaging shows that these structures do not change in signal intensity or shape during the time course of the experiments and do not seem to be very dynamic or involved the with gradual accumulation of GFP-COL in the Golgi. IF confirmed that these structures are negative for COPII, but instead co-label with markers for the chaperon HSP47 and ER membrane. Therefore, it can be assumed that the large GFP-COL objects with diameters of 0.5 – 2 µm are not bona fide carriers. These structures are likely a result of GFP-COL overexpression and subsequent formation of aggregates in the ER unable to participate in the process of trafficking and secretion. This agrees with the fact that large structures observed in literature occurred in cells overexpressing either KLHL12 or procollagen and KLHL12 (Gorur et al., 2017). These large GFP-COL structures may act as storage units or are artefacts from overexpression and may be destined for degradation. Large GFP-COL objects only occasionally colocalise with autophagosomal markers (mCherry-LC3B) and are negative for endosomal markers EEA1, Rab11 and TfR.

The absence of involvement of large structures or carriers stands in contrast to observations made in other cell lines, e.g. overexpressing KLHL12 and type I procollagen, where large (>400 nm diameter) objects positive for type I procollagen were shown to colocalise with COPII (Gorur et al., 2017; McGourty et al., 2016). These previously reported *"megacarriers"* are likely a result of overexpression in these cells and have recently been linked to a degradation pathway (Omari et al., 2018).

Together, the data suggest a more direct transport mode of procollagen from the juxtanuclear ER to the Golgi, without the use of large carriers (Kurokawa et al., 2014; Malhotra and Erlmann, 2015).

Chapter 5: Analysis of endogenous procollagen in fibroblasts from patients with rare variants of osteogenesis imperfecta

Statement of Contribution: The manuscripts resulting from work in this chapter have been published in collaboration with Meena Balasubramanian (listed in the appendix) (Balasubramanian et al., 2019; Balasubramanian et al., 2018). All patient-related work was undertaken by the collaborators in Sheffield, led by Meena Balasubramanian, while work on patient-derived skin fibroblasts is my own. Mass spectrometry and tandem-mass tagging for proteomics analysis was performed by the Proteomics facility at the University of Bristol led by Kate Heesom.

5.1 Introduction

Mutations in genes encoding for collagens or proteins essential for correct procollagen processing can cause collagenopathies (Cutrona et al., 2018; Jobling et al., 2014; Wong and Shoulders, 2019). Depending on the type and location of the mutation this can lead to more common or less often occurring phenotypes ranging from mild to severe. Yet unclassified variants, also referred to as Variants of Uncertain Significance (VUS) are to date not well studied (Balasubramanian et al., 2019). VUS are genetic mutations that have been identified, but not yet been confirmed as the cause of clinical phenotypes and thus the variant itself is not yet assigned to be pathogenic (Richards et al., 2015).

To analyse the pathogenicity, as well as the general phenotype of VUS/rare variants with OI on a cellular basis and clarify the clinical severity two different patient-derived cell types were analysed (Table 8). The first patient with a VUS has a mutation in the gene encoding for COL1A1, referred to as COL1^{mut} and displays a mild OI phenotype. The other patient has a mutation in subunit B of prolyl 4-hydroxylase (P4HB), referred to as P4HB^{mut}, with Cole Carpenter Syndrome (CCS) and severe OI phenotype (Balasubramanian et al., 2018).

The mutation from the COL1^{mut} patient causes a *de novo* in-frame deletion of a lysine residue at position 69 in the N-terminus of COL1A1. Since folding of procollagen trimers occurs directionally from C- towards the N-terminus, a mutation in the N-terminus usually does not affect trimer formation itself and does not alter the triple helical domain structurally. There is, however, evidence that, once in the extracellular space, the N-terminus of procollagen has a regulatory feedback function on procollagen synthesis (Oganesian et al., 2006). Therefore, a mutation in this area might affect overall collagen abundance and matrix composition.

Table 8: Genetic and clinical phenotypes of both patients with rare collagen-related variants.

Patient cell Identifier	gender	Age in years	Clinical phenotype	Mutation	Gene
COL1^{mut}	Male	7	failure to gain weight, talipes, fractures	c.206_208delTGT, p.Leu69del	COL1A1
P4HB^{mut}	Female	3	severe bone fragility, multiple fractures	c.1178A>G, p.Tyr393Cys	P4HB/ PDIA1

The P4HB gene encodes for PDI, also called P4HB. It is a multifunctional protein that can both catalyse the rearrangement of disulphide bonds at the cell surface and the ER-lumen, as well as act as a subunit of P4H. The P4H heterotetramer is composed of two P4HB subunits and two P4H alpha chains and is retained in the ER via a KDEL motif from P4HB (Helaakoski et al., 1989). P4H post-translationally hydroxylates proline residues in X-Pro-Gly motifs that occur predominantly in the COL-domain of procollagen chains. This enhances chain polarity and contributes towards stabilising procollagen trimers via water-bridges.

P4HB at high concentrations can also act as a chaperone in the ER-lumen inhibiting protein aggregation of misfolded proteins. Low concentrations of P4HB exhibit anti-chaperone activity by facilitating aggregation. (Lumb and Bulleid, 2002) Furthermore, P4HB is essential for cell viability in yeast (LaMantia et al., 1991). And to date there have been no mice or humans reported with a homozygous mutation in P4HB.

The heterozygous mutation from the P4HB^{mut} patient leads to a replacement of tyrosine with cysteine in position 393, which is localised in the second thioredoxin domain of the subunit. The tyrosine residue replaced here is highly conserved throughout evolution in eukaryotes and other members of the PDI protein family, indicating an important role in enzymatic functionality (Benham, 2012; Hatahet and Ruddock, 2009). P.Tyr393Cys has been reported previously and *in vitro* experiments confirmed impaired PDI activity in this variant (Rauch et al., 2015). Mutant PDI extracted from cells also indicated that this variant results in PDI associating more tightly with substrates via disulphide bridges than WT PDI (Rauch et al., 2015). This change in activity and binding affinity could influence collagen processing and secretion.

To help classify the pathogenicity of these rare/novel variants and identify the underlying cell biological processes resulting in the clinical phenotypes observed, patient-derived fibroblasts from both the COL1^{mut} and P4HB^{mut} patients were analysed.

5.2 Results

The 7-year old COL1^{mut} patient exhibits a short stature, a blueish tinge to his sclerae, a small jaw, failure to gain weight, hypermobility of small joints and dental enamel hypoplasia. He previously suffered from talipes and fractures (Fig. 42, arrows). The outer layer of the bone, the cortex, shows increased porosity and the COL1^{mut} patient has been diagnosed with osteopenia (low bone mass density). His skull is marked by Wormian bones, which are formed from additional ossification and used as a marker for diagnosing autosomal dominant genetic disorders like craniosynostosis and OI (Bellary et al., 2013). He furthermore has a delayed bone age of 1.5 years.

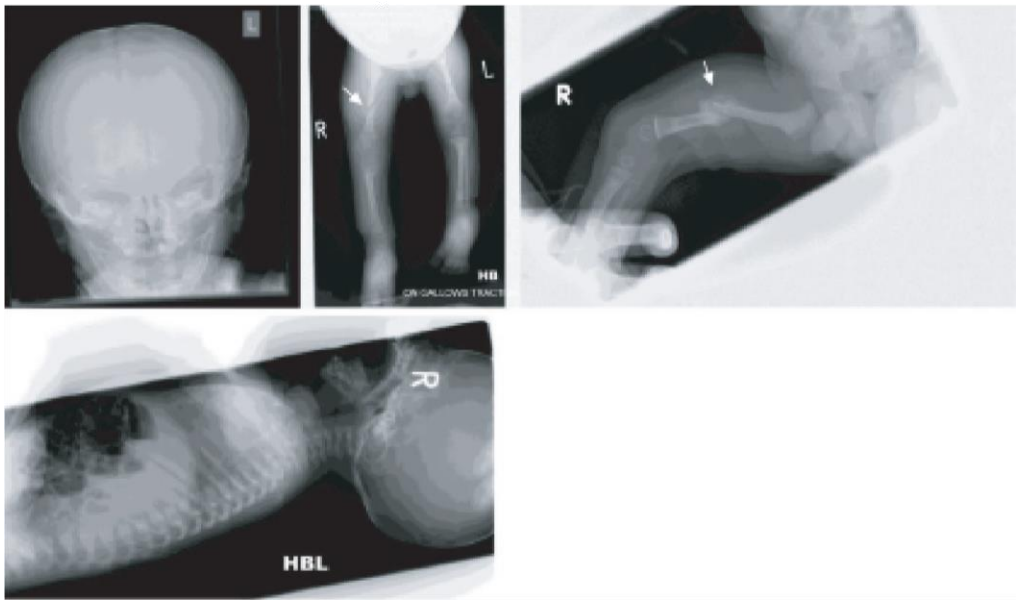


Fig. 42: X-ray images showing the skeletal phenotype of COL1^{mut} patient.

Radiographs show the skull, both legs, the right leg and the lateral spine. The COL1^{mut} patient shows osteopenia with Wormian bones. The vertebrae appear normal. Fractures are indicated by arrows. (Balasubramanian et al., 2019)

The 3-year old P4HB^{mut} patient also displays reduced bone density (osteopenia) and suffers from multiple fractures (Fig. 43, arrows). Fracture sites are often angulated with fracture adjacent bone displaced and bone axis orientation shifted) (Fig. 43B). There are compression fractures along the spine (Fig. 43D).The patient shows irregular sclerosis at the metaphysis of upper arm and leg bone (Fig. 43A - C), around healing fractures (Fig. 43E - F) and at the metaphysis of the shin bone in proximity to the knee (Fig. 43C, E - F). These symmetrical metadiaphyseal fractures are uncommon in OI. Radiographs taken at 31 months of age show forearm and leg bones with significant bowing deformity (Fig. 43F). Bisphosphonate therapy most likely caused an increase in bone density (bisphosphonate lines are visible on the forearm bones close to the wrist and the tarsal (midfoot) bones).

To investigate the collagen matrix, skin fibroblasts from control subjects and patients were analysed by IF.

Fig. 43: X-ray images showing skeletal phenotype of P4HBmut patient.

Radiographs show the right forearm (A), left forearm (B), right leg (C), lateral spine (D), apical and lateral right thighbone (E), as well as the lateral right forearm and right leg (F). The age at which the radiographs were taken is indicated at the top of the images. Arrows indicate fractures. (Balasubramanian et al., 2018)

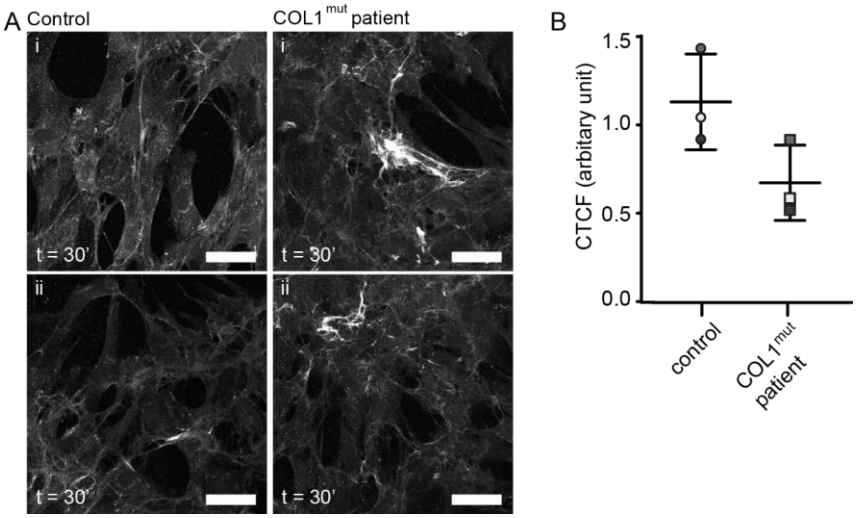


Fig. 44: Extracellular COL1A1 by IF from control and COL1^{mut} patient cells.

Maximum projection of confocal z-stacks consisting of 24 slices with Δz=0.08 µm. Experimental repeats n=3. A: Extracellular COL1A1 after cell growth for 48 – 72 h on coverslips and subsequent treatment with 50 µg·mL^{-1} ascorbate for 30 min prior fixation with PFA and antibody-labelling with COL1A1. Scale bar indicates 50 µm. B: Quantification of the mean CTCF for COL1A1 of randomly chosen fields of view from control sample (circles) and patient sample (squares) as in A. Data points show mean intensity from three independent experiments labelled with different grey values. Error bars show mean value and standard deviation. (Balasubramanian et al., 2019)

Fig. 44A shows extracellular COL1A1 after 30 min incubation with ascorbate prior to fixation and imaging. Extracellular COL1A1 labelling does not appear different in the patient cell matrix compared to the control. The only apparent difference observed are highly concentrated areas of collagen present in the COL1^{mut} patient sample but not in the control sample. Analysis of overall fluorescence intensity via calculating the CTCF revealed a consistent decrease in COL1A1 signal in patient samples compared to the control Fig. 44B.

The P4HB^{mut} and control samples show a similar collagen fibril morphology, as well as signal intensities. This was tested following incubation in presence of ascorbate for 30 min and 24 h (Fig. 45A - B). Quantification of signal intensity for COL1A1 of both mid- and high-intensity areas (Fig. 45C-D) further supported this. The COL1A1 matrix of P4HB^{mut} patient cells showed no statistically detectable difference compared to control samples (with a P value of 0.76 for the quantification of mid-intensity and 0.34 for high-intensity areas). There is, although statistically not significant, a trend towards an increase of type I collagen abundance in the P4HB^{mut} patient cells.

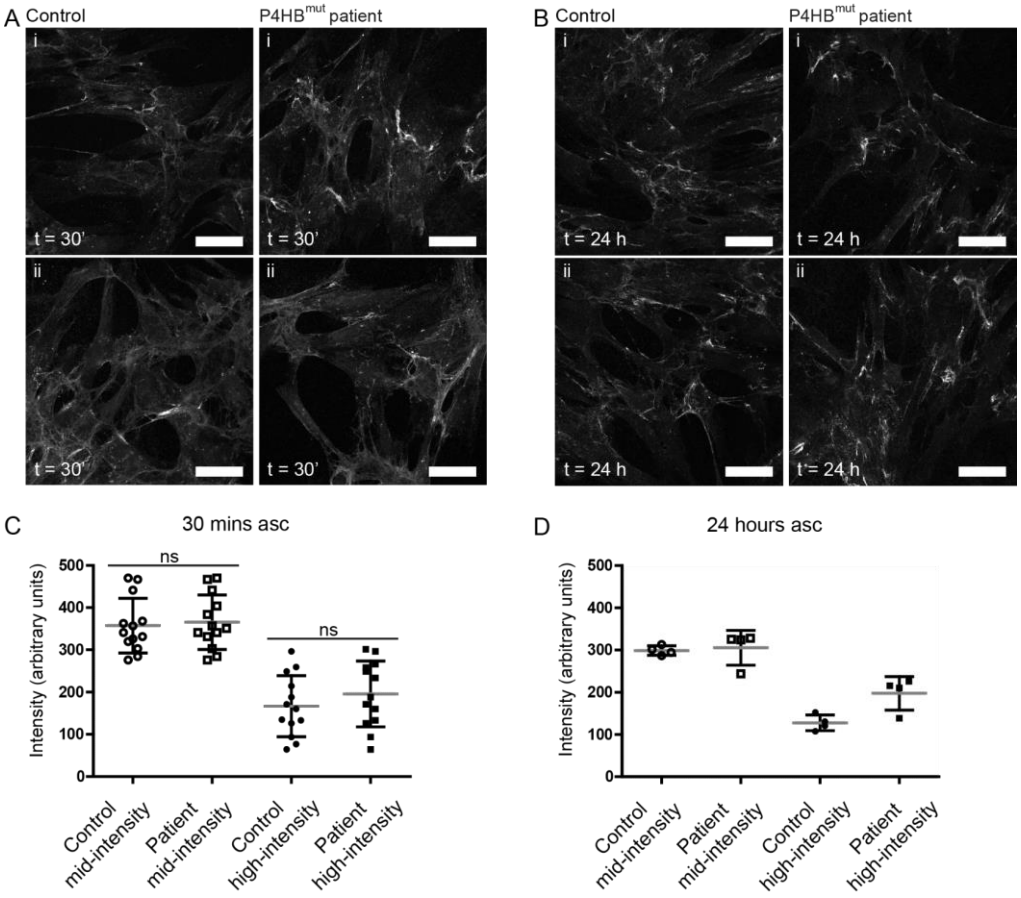


Fig. 45: Extracellular COL1A1 by IF from control and P4HB^{mut} patient cells.

Maximum projection of confocal z stacks consisting of 24 slices with Δz=0.08 µm. Extracellular COL1A1 after cell growth for 48 – 72 h on coverslips and subsequent treatment with 50 µg·mL^{-1} ascorbate for 30 min (A) or 24 h (B) prior fixation with PFA and antibody-labelling with COL1A1 (Novus Biologicals), C - D: Quantification of A and B for mid-intensity values (16%–62% or 40–159; empty shapes) and high-intensity values (63%–100% or 160–255; filled shapes) for control (circles) and patient (squares). Statistical analysis was performed using an unpaired Student's t-test (P values). Experimental repeats n=3. Scale bar indicates 50 µm. (Balasubramanian et al., 2018)

Immunoblot analysis of patient and control cells show type I collagen abundance in both lysate and media fractions in absence and presence of ascorbate for 24 h prior to sample collection and corresponding loading control (Fig. 46). All samples show a prominent band for COL1A1 at the expected molecular weight of about 150 kDa and two additional less intense bands at higher molecular weights corresponding to the processed forms of pro-α1(I).

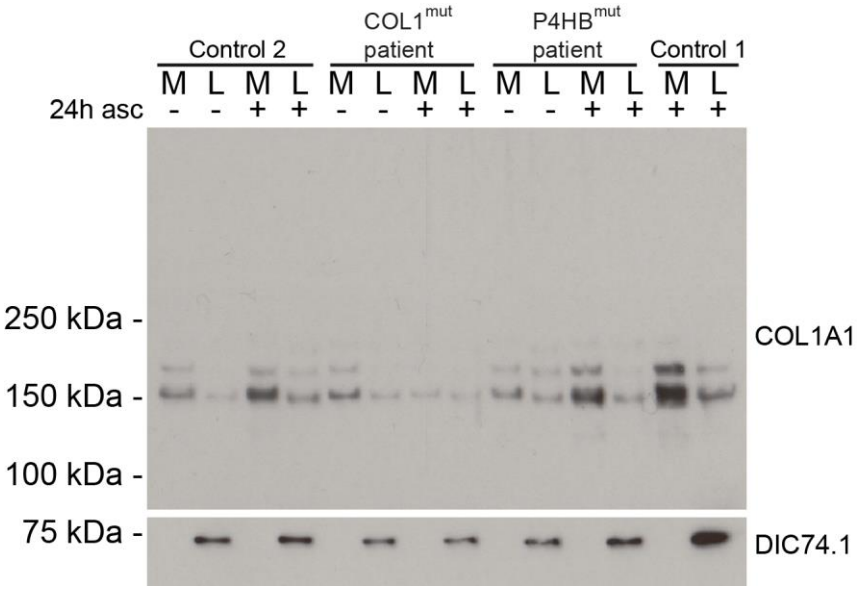


Fig. 46: COL1A1 levels in control and COL1^{mut} and PH4B^{mut} cell samples.

Cells were seeded confluent and grown in presence (+) or absence (-) of 50 µg·mL^{-1} ascorbic acid in serum-free media. For WBs analysis of lysate (L) and corresponding media (M) fractions of skin fibroblasts membranes were labelled for COL1A1 and dynein intermediate chain (DIC74.1) as loading control. Control samples 1 and 2 originate from two different control cell lines generated by the same procedure as the patient cells by the hospital of Sheffield. Marker bands indicate protein size in kDa. Experiment was performed with n=3 repeats. (Balasubramanian et al., 2019; Balasubramanian et al., 2018)

Both the control and P4HB^{mut} patient cells show an increase in type I collagen in the media fractions upon incubation with ascorbate, while collagen levels in the COL1^{mut} patient cells do not increase upon presence of ascorbate. Overall collagen abundance appears lower in COL1^{mut} patient cells compared to P4HB^{mut} and control cells.

Since the abundance of COL1A1 in the extracellular space was confirmed to be reduced in COL1^{mut} patient cells by both IF and immunoblot analysis, the CDM was extracted to further investigate matrix composition.

Fig. 47 shows the ratio of collagen isoforms detected using proteomics of COL1^{mut} patient and control extracellular matrices. COL1A1, COL1A2, COL3A1, as well as COL5A2 appear consistently reduced in patient cell matrix compared to control throughout all three repeats of the experiment. COL5A1, COL6A1, COL6A2, COL6A3, COL12A1 abundance vary between experiments and COL14A1 levels do not seem to differ much between control and patient sample. Other types of collagen were not detected in sufficient abundance in samples to be able to compare these quantitively.

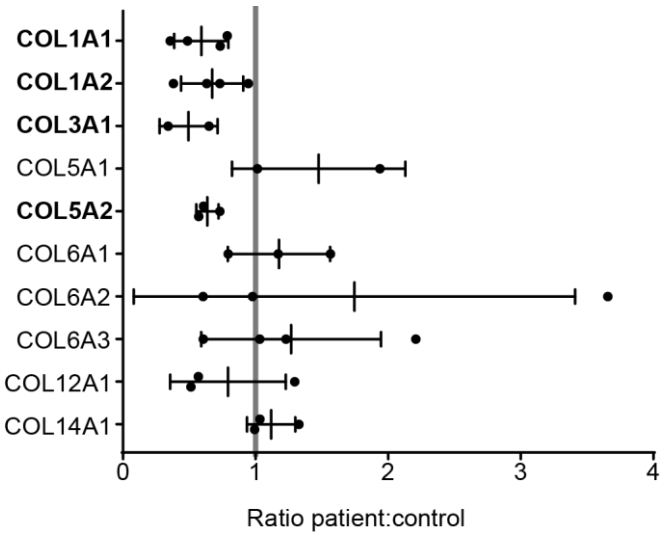


Fig. 47: MS of collagens in the cell-derived matrix (CDM) from control and COL1^{mut} patient cells.

Cells were grown for 5 days in 50 µg·mL^{-1} ascorbate supplemented media and the cell derived matrix was send for analysis via TMT-MS after extraction of cells with n=3. Error bars show mean ratio between patent and control samples and standard deviation. A ratio <1 indicates reduced collagen in the CDM from patient fibroblasts and was detected for collagens highlighted in bold. (Balasubramanian et al., 2019)

To investigate if there are any defects in intracellular procollagen localisation and or trafficking control and patient cells were labelled with antibodies recognising COL1A1 (Novus Biologicals) in combination with antibodies targeting either the Golgi apparatus via GM130 or GRASP65, the collagen chaperone HSP47 or COPII via Sec31A (BD Biosciences) (Fig. 48-Fig. 52). Intracellular procollagen localises in COL1^{mut}, P4HB^{mut} and control cells to the ER with occasional concentration in the Golgi area (Fig. 48, Fig. 51 and Fig. 52). Labelling intensity for HSP47 varies between cells in both control and patient fibroblasts (Fig. 49 and Fig. 51). There is no obvious accumulation of HSP47 visible in any of the imaged cells and the ER appears as expected with both tubular and sheet like structures. Similarly, labelling with Golgi markers shows that all samples display intact Golgi ribbons. COPII labelling investigating Sec31A also reveals normal localisation in form of small dispersed puncta throughout the cell with an accumulation at the ER-Golgi interface in both control and patient samples (Fig. 50 and Fig. 52). There is no difference visible by IF when comparing all three samples regarding the structures labelled for.

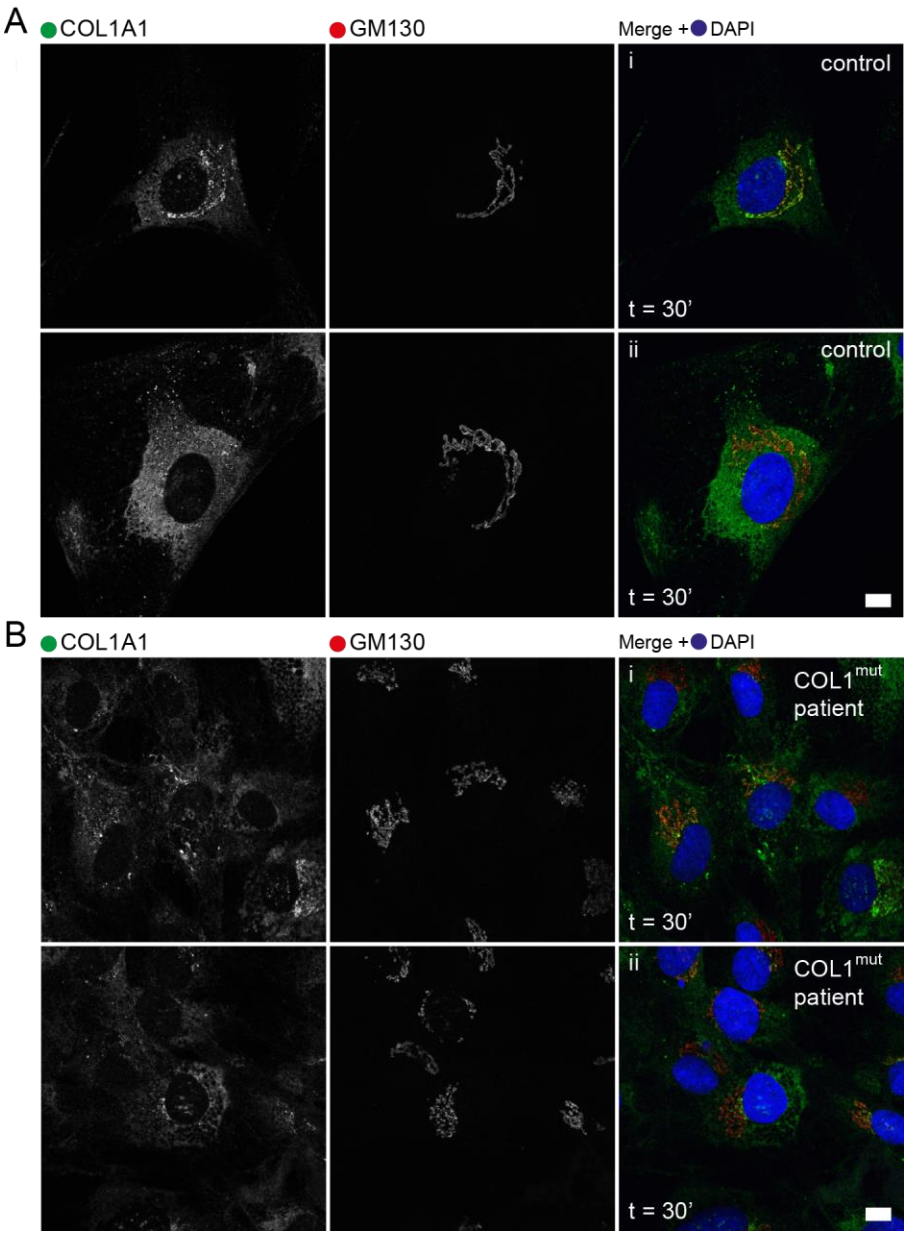


Fig. 48: Localisation of procollagen I and GM130 in control (A) and COL1^{mut} patient fibroblasts (B).

Confocal images of skin fibroblasts labelled with antibodies for COL1A1 (green, Novus Biologicals) and *cis*-Golgi GM130 (red, BD Biosciences) and nuclear DAPI staining in blue. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation with PFA. Number of experimental repeats was n=1 containing >50 cells per experiment. Due to no obvious differences observed and limited proliferative capacity of primary fibroblasts this experiment was considered not worth repeating. Images were equally brightness enhanced. Scale bars indicate 10 µm.

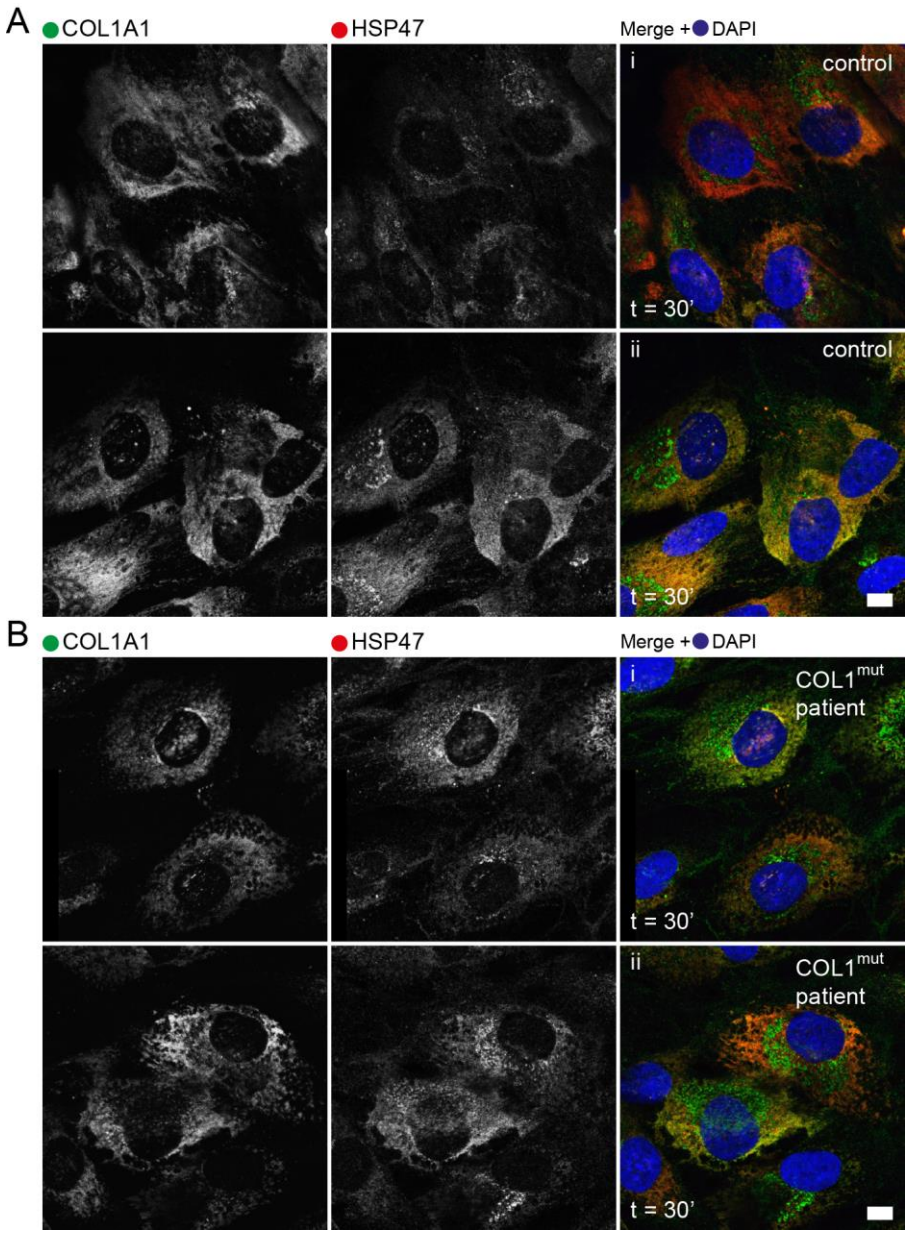


Fig. 49: Localisation of procollagen I and HSP47 in control (A) and COL1^{mut} patient fibroblasts (B).

Confocal images of skin fibroblasts labelled with antibodies for COL1A1 (green, Novus Biologicals) and HSP47 (red, Enzo) and nuclear DAPI staining in blue. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation with PFA. Number of experimental repeats was n=1 containing >50 cells per experiment. Due to no obvious differences observed and limited proliferative capacity of primary fibroblasts this experiment was considered not worth repeating. Images were equally brightness enhanced. Scale bars indicate 10 µm.

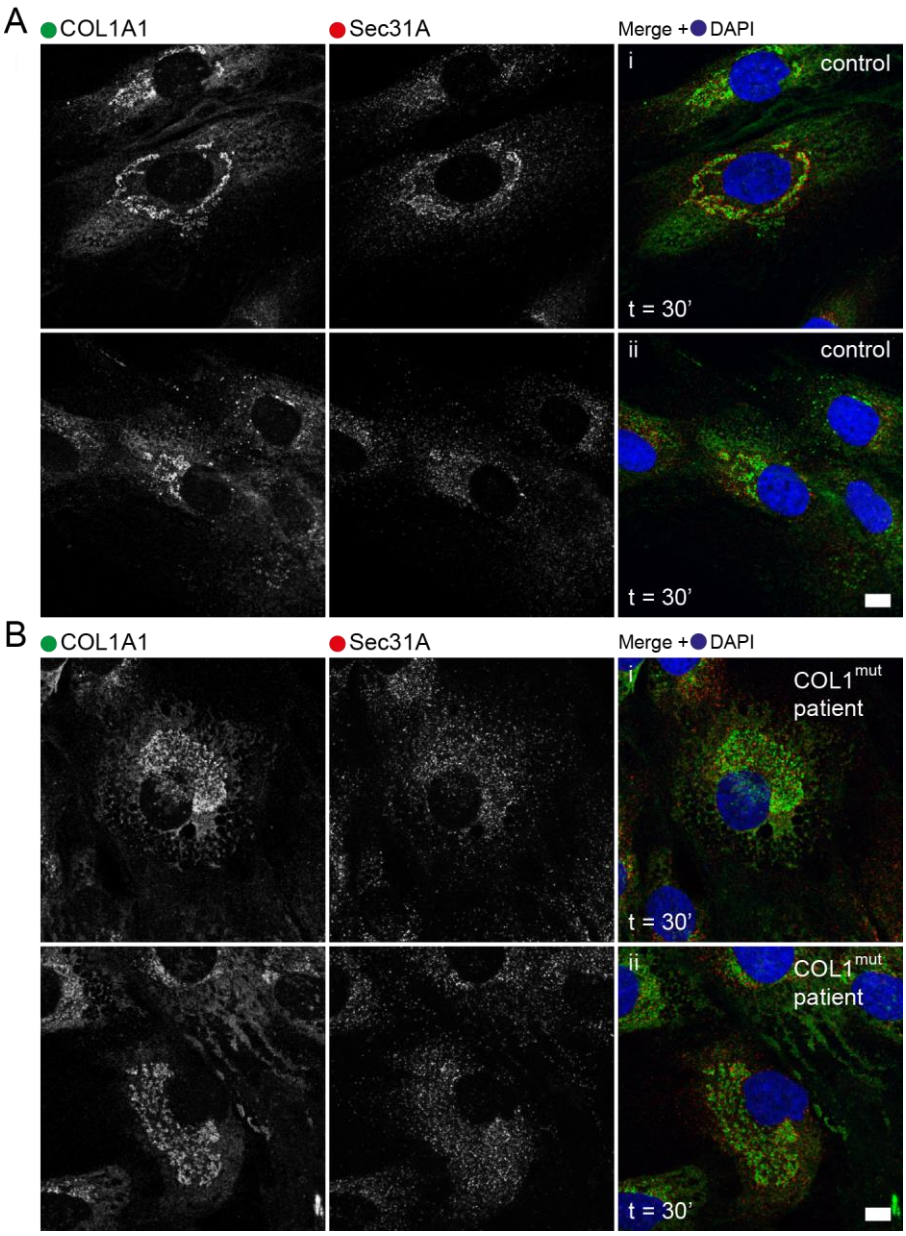


Fig. 50: Localisation of procollagen I and Sec31A in control (A) and COL1^{mut} patient fibroblasts (B).

Confocal images of skin fibroblasts labelled with antibodies for COL1A1 (green, Novus Biologicals) and COPII component Sec31A (red, BD Biosciences) and nuclear DAPI staining in blue. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation with PFA. Number of experimental repeats was n=2 containing >10 cells per experiment. Images were equally brightness enhanced. Scale bars indicate 10 µm.

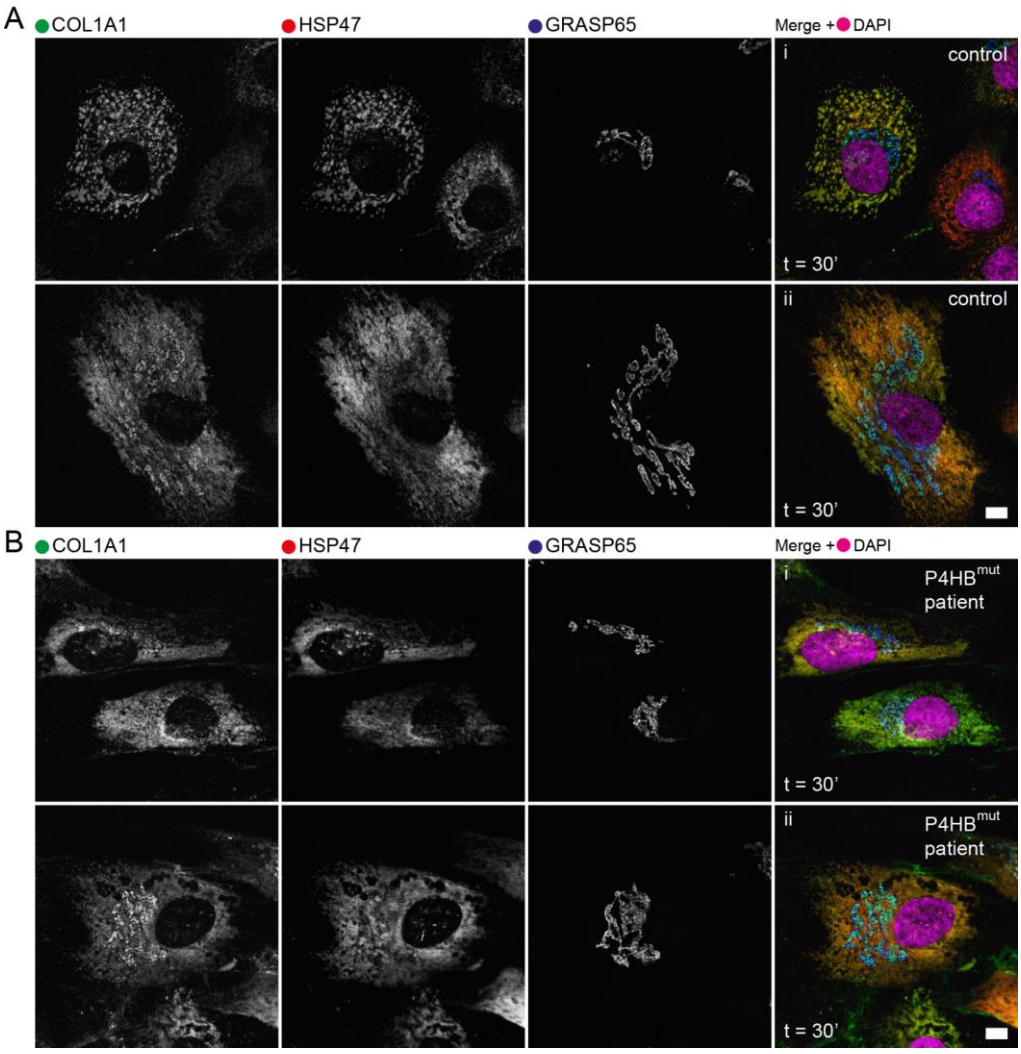


Fig. 51: Localisation of procollagen I, HSP47 and GRASP65 in control (A) and P4HB^{mut} patient fibroblasts (B).

Confocal images of skin fibroblasts labelled with antibodies for COL1A1 (green) and HSP47 (red) and Golgi marker GRASP65 in blue and nuclear DAPI staining (imaged in a separate channel) in pseudo-colour magenta. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation with PFA. Number of experimental repeats was n=2 containing >10 cells per experiment. Images were equally brightness enhanced. Scale bars indicate 10 µm.

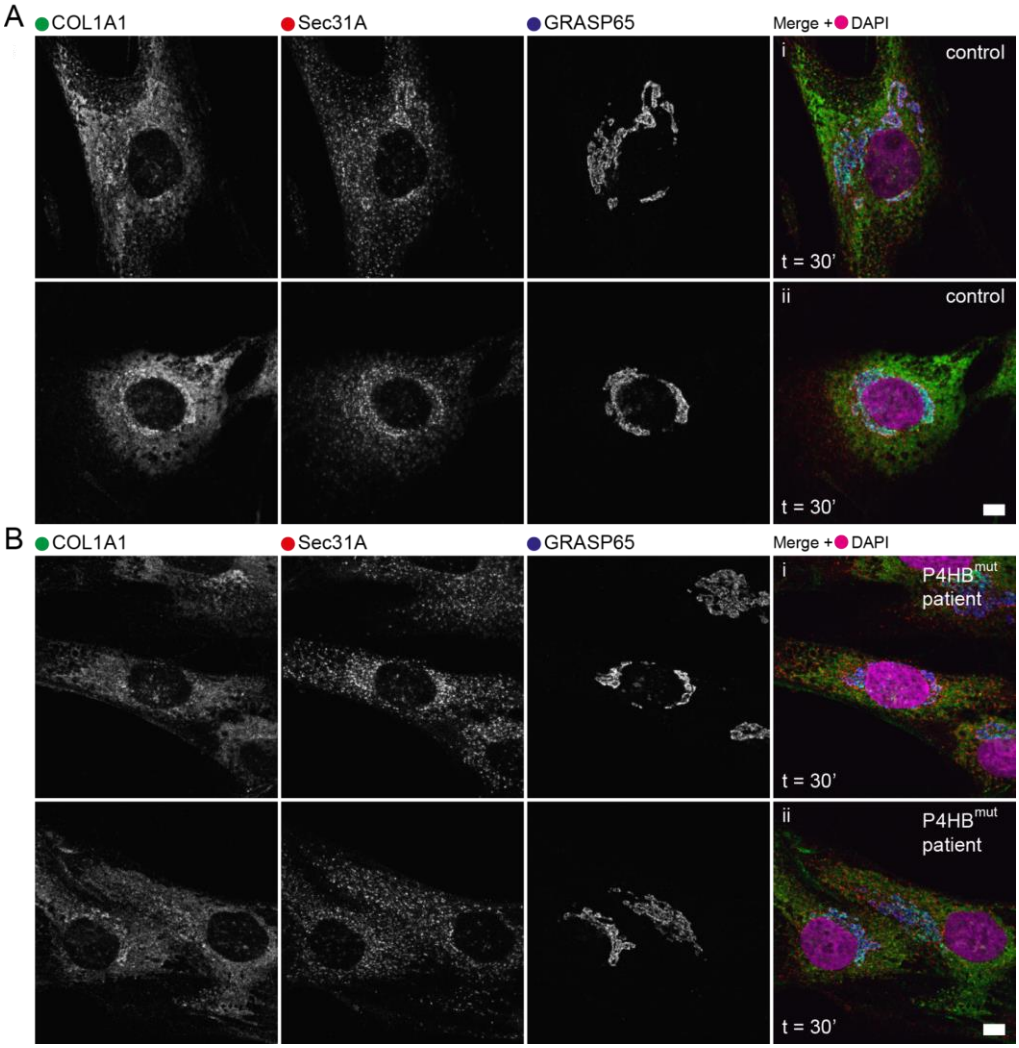


Fig. 52: Localisation of procollagen I, Sec31A and GRASP65 in control (A) and P4HB^{mut} patient fibroblasts (B).

Confocal images of skin fibroblasts labelled with antibodies for COL1A1 (green, Novus Biologicals) and COPII component Sec31A (red, BD Biosciences) Golgi marker GRASP65 in blue and nuclear DAPI staining (imaged in a separate channel) in pseudo-colour magenta. Cells were incubated with 50 µg·mL^{-1} ascorbate for 30 min prior to fixation with PFA. was n=2 containing >10 cells per experiment. Images were equally brightness enhanced. Scale bars indicate 10 µm.

5.3 Discussion

The data shows that the COL1^{mut} patient cells have a lower type I collagen abundance in the extracellular space and overall matrix composition is altered, while the intracellular localisation of procollagen, HSP47, ERES markers and the Golgi apparatus does not seem affected. Data acquired from the P4HB^{mut} patient cells did not reveal any significant differences between patient and control samples. There is nonetheless a trend towards an increase in type I collagen deposition into the extracellular space. Both patients suffer from low bone mass density (osteopenia),

resulting in mild OI (COL1^{mut}) and severe OI / CCS, with multiple fractures and skeletal deformities (P4HB^{mut}).

The reduced type I collagen abundance in the cells and the ECM of cells derived from the COL1^{mut} patient compared to controls are consistent with the skeletal phenotype of mild OI. Collagen fibrils were also observed to form unusual densely packed areas in the patient cell matrix by IF. This indicates defects in type I collagen deposition, also affecting overall matrix assembly. Proteomics data revealed a reduction in other extracellular matrix components including type III and V collagens. Both type III and V collagen interact with type I collagen in the extracellular space. Dermal collagen fibrils contain both type I and III collagen (Fleischmajer et al., 1990) and type V collagen can also be found in the same fibrils as type I in the avian corneal stroma (Birk et al., 1988). Type V collagen is an important regulator of fibrillogenesis of type I collagen containing fibrils (Birk, 2001; Birk et al., 1990) and mutations in type V collagen have been reported to cause EDS (Burrows et al., 1998; De Paepe et al., 1997; Johnston et al., 2017; Michalickova et al., 1998; Nicholls et al., 1996; Richards et al., 1998; Wenstrup et al., 1996). Furthermore, expression of type I and type III procollagen can be coregulated by the same signalling pathway (Walsh et al., 1987). The regulatory feedback function of the N-propeptide of type I procollagen (Bornstein et al., 2002; Card et al., 2010) might be altered upon the COL1^{mut} mutation, leading to a decrease in COL1A1 synthesis. While secretory protein abundance dictates matrix composition directly, ECM composition itself may also have a regulatory feedback function on synthesis of procollagen and other extracellular matrix components (Parker et al., 2014).

The collagen abundance and phenotype of the ECM in P4HB^{mut} patient cells was indistinguishable from that of control cells agreeing with previous findings of patient skin fibroblasts with the same mutation via EM (Rauch et al., 2015). Interestingly, IF analysis of intracellular COL1A1, HSP47, the Golgi apparatus and COPII marker Sec31A did not show any difference between either of the COL1^{mut} or P4HB^{mut} patients and control cells. This indicates that the ER morphology and procollagen trafficking in both P4HB^{mut} and COL1^{mut} cells appears normal. Delayed procollagen trafficking is often characterised by over-modification (Ishikawa and Bachinger, 2013). However, no over-modification or indications for a delay in procollagen trafficking were found here or in previous studies regarding other CCS patients (Amor et al., 2000; Cole and Carpenter, 1987; Rauch et al., 2015).

In a previous study investigating two other patients with the same mutations as P4HB^{mut} an increase in HSP47 was observed in skin fibroblasts by WB (Rauch et al., 2015). IF showed altered ER morphology following labelling for intracellular type I collagen and PDI with an accumulation of punctate structures positive for both markers localising in the ER (Rauch et al., 2015). These results indicate potential ER-stress. Overall

expression of PDI was reported to be normal in patient-derived cells compared to controls (Rauch et al., 2015).

This stands in contrast to the observed findings where no increase in HSP47 or change in ER morphology or type I collagen localisation was observed by IF. One should also consider the age difference between the different patients. The P4HB^{mut} patient cells were derived from a patient aged 3 years old, while the cells showing the ER-stress phenotype with altered ER morphology and increased HSP47 levels were derived from an 18-year old patient (Rauch et al., 2015). Furthermore, bisphosphonate treatment at the age of 7 months might also affect both intracellular and skeletal phenotypes of the P4HB^{mut} patient. This treatment seemed to have contributed to a decrease in the number of fractures suffered (Balasubramanian et al., 2018). Differences on a cellular basis concur with the fact that despite having the same underlying mutation in P4HB, different patients show different severity and clinical phenotypes (Balasubramanian et al., 2018). The specific skeletal phenotype, however, appears to be consistent for the CCS caused by this specific mutation in P4HB. All patients reported with the same mutation in P4HB display specific bone alterations characterised by multiple metaphyseal 'crumpling' fractures and metaphyseal sclerosis untypical for other types of OI, as well as bowing deformities (Balasubramanian et al., 2018; Cole and Carpenter, 1987). The first two phenotypes were not observed in CCS diagnosed patients tested negative for the P4HB^{mut} (Amor et al., 2000; Balasubramanian et al., 2015).

Since all clinically observed phenotypes in the patients can be directly associated to collagen and intracellular procollagen, ER and Golgi phenotypes were visually indistinguishable between control and patient cells and overall protein transport is likely unaffected in both patients.

There are indications that haploinsufficiency of PDI itself is not the cause for the CCS phenotypes *in vivo* as there are healthy individuals with heterozygous mutations in the P4HB locus. (Hahm et al., 2013; Jakobsson et al., 2008; Rauch et al., 2015; Wong et al., 2007). Since the P4HB^{mut} patient exhibits severe bone fragility, it seems that the mutation in P4HB affects matrix organisation and assembly rather than secretion and synthesis of type I procollagen itself. Despite PDI being ubiquitously expressed, the phenotype observed predominantly affects the skeleton and causes accumulation of cerebrospinal fluid in ventricles of the brain (Benham, 2012). The effect caused by P4HB^{mut} are likely tissue specific, since abnormalities are found in the skeletal, but e.g. not dermal tissues (Rauch et al., 2015). Thus, the severe OI-like bone fragility might be caused by cell stress in osteoblasts which perturbs bone formation, as seen in the OI mouse model (Lisse et al., 2008; Rauch et al., 2015). The mutation in P4HB might affect ECM assembly and organisation rather than affecting type I procollagen synthesis and transport (Balasubramanian et al., 2018).

Proteomics experiments as conducted for the COL1^{mut} patient cells would greatly contribute to assessing overall ECM composition in the P4HB^{mut} patient cells. To rule out a possible increase in HSP47 and ER-stress levels in the P4HB^{mut} patient cells, protein abundance via immunoblotting as done in previous studies could be performed. To investigate tissue specific changes in procollagen abundance and trafficking analysis of patient-derived osteoblasts would greatly complement the acquired data. The introduction of an additional cysteine residue in PDI close to the second catalytic centre, as in the P4HB^{mut}, might lead to decreased substrate-specificity and altered binding affinity to P4H alpha subunits and other non-specific proteins. To test this binding-assays with recombinantly expressed mutant PDI could be performed. Additional work elucidating the mechanistic processes occurring upon mutations in COL1A1 and P4HB will further contribute to understanding basic cell biological events, as well as enable development of new treatment.

Overall, both the clinical phenotype and the cellular analysis confirm that the rare COL1^{mut} and P4HB^{mut} variants are pathogenic. The cell-based work conducted shows that this type of analysis can help identify the underlying, mechanistic issues of rare genetic disorders affecting collagen secretion. In addition, clarifying pathogenicity of new variants helps with diagnosis and thus better suited treatment for affected patients. Variability in the severity and phenotype despite the same underlying mutation is a common characteristic of genetic disorders like OI. This makes it difficult to diagnose and treat patients, which is why a case by case exome sequencing and analysis using biochemical and cell biological approaches can greatly complement clinical assessment.

Chapter 6: Discussion

Data from both primary fibroblasts investigating endogenous type I procollagen and GFP-COL-RPE show procollagen present in form of small puncta but not in the quantity expected for efficient ER-to-Golgi transport. The larger (up to micron sized) GFP-COL structures occasionally observed were not involved in the transport process and were negative for COPII markers and were instead part of the ER. Most ERES are positioned in proximity of the Golgi apparatus (Klumperman et al., 1998), but ERES are also present in peripheral regions of the cell (Klumperman et al., 1998; Lotti et al., 1992). While ERGIC elements remain close to ERES all throughout the cell, the distance of peripheral ERES to the Golgi might result in peripheral sites utilising a different transport mode from those ERES in Golgi proximity. Indeed, transport of GFP-COL to the Golgi could only be visualised from ERES juxtaposed the Golgi and occurred in a gradual manner. This, together with super resolution microscopy indicating that procollagen is rather partially coated than entirely encapsulated by COPII suggests an alternative model (Fig. 53).

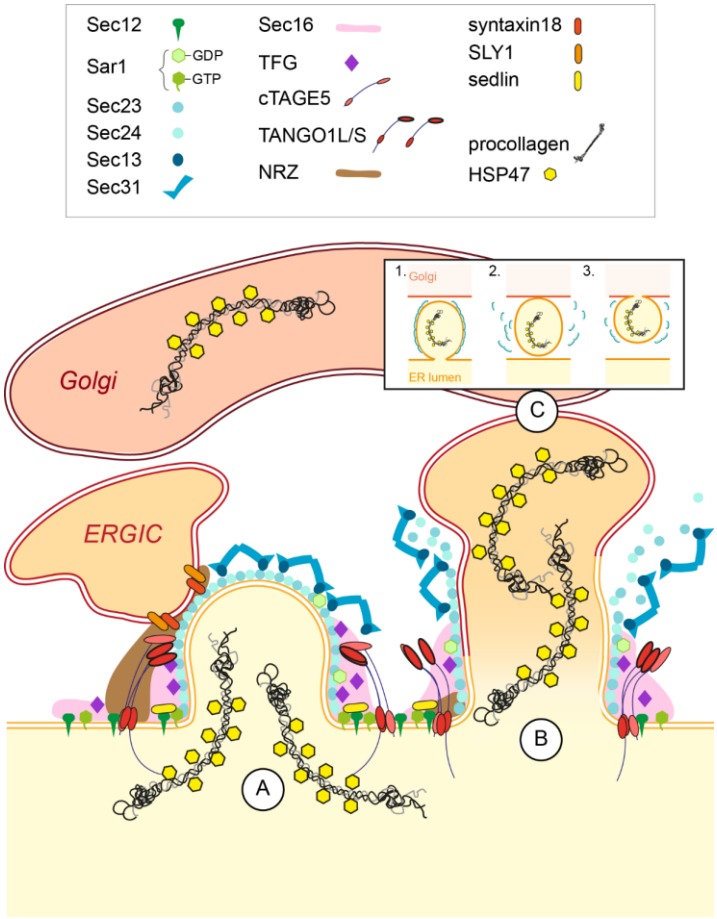


Fig. 53: Model for procollagen transport from ERES in proximity to the Golgi.

Procollagen is concentrated at ERES via HSP47 acting as an adaptor for recognition by TANGO1L (A). Via the NRZ tethering complex ERGIC membranes can fuse with pudding COPII structures. The enlarged budding complex with partially incorporated more flexible procollagen

undergoes rapid uncoating (B), allowing for transient merging with Golgi membranes/maturation to become the first post-ER Golgi cisterna (C).The close proximity with the Golgi does not allow for a permanent state of a large carrier encapsulating procollagen and rapid fusion with the help of SNARES takes place.

In this model, compartment identity can be retained by envisioning direct connections between ER and ERGIC, but not the Golgi itself. This is consistent with the role of TANGO1 facilitating COPII-bud expansion via fusion with ERGIC membranes (Ma and Goldberg, 2016; Nogueira et al., 2014; Raote et al., 2018). This step is enabled by interaction of TANGO1 with SNARE proteins Sly1 and syntaxin 18 at ERES (Nogueira et al., 2014). As the TANGO1-cTAGE5 complex not only binds the inner layer COPII components, but also HSP47 (Ishikawa et al., 2016) it can indirectly bind to collagen and other matrix proteins, linking bud-expansion using ERGIC-derived membrane, to cargo transport. Thus, a model incorporating recent findings could envision procollagen accumulating at ERES in a COPII-dependent manner with a partially coated bud incorporating procollagen. Procollagen transport is initiated by TANGO1-dependent bud-expansion, with a growing carrier containing, possibly slightly bent (Rezaei et al., 2018), procollagen (Fig. 53A). Possible COPII coated intermediates are most likely of transient nature and the limited space between the ER and the ERGIC does not allow for larger carriers to stably exist without being continuous with either the ER or ERGIC/Golgi. This carrier then matures to become part of the ERGIC, prior to fusion with or maturation to the first Golgi cisterna similar to the cisternal maturation model for inter-Golgi cisternae transport of procollagen (Bonfanti et al., 1998) (Fig. 53B-C).

Whether the growing COPII bud requires fusion with ERGIC membranes as an enhancement for efficient bud expansion to enable incorporation of large cargoes like procollagen is not known. So far there has been no experimental evidence for the involvement of this specific process in type I procollagen transport, as SLYI-KO only impairs type VII procollagen transport (Nogueira et al., 2014). Most data hints at Sar1-GTP cycling in combination with delayed binding of the outer COPII coat to drive COPII bud expansion (Aridor et al., 2001; McCaughey and Stephens, 2019). Due to the highly dynamic and transient system of ER-to-Golgi transport it is extremely difficult to biochemically test the proposed model. One could potentially make use of locally inducible depletion or inactivation of proteins like TANGO1 or SNARES via e.g. rapalog or light-inducible dimerization to trigger inactivation or mislocalisation (Nijenhuis et al., 2020; van Bergeijk et al., 2015).

Of note, some small dynamic GFP-COL puncta translocating to the Golgi from the cell periphery could be observed during live cell imaging. However, no such structures were found to be positive for mScSec23A and COPII puncta in general were not seen to be mobile. This agrees with previous live cell imaging of COPII consistently showing COPII structures to be immobile, while long-range transport of COPI-positive structures can

be observed (Presley et al., 2002; Scales et al., 1997; Shima et al., 1999; Stephens et al., 2000). This implies, that alternative transport pathways, in addition to the proposed model exist. Furthermore, the suggested model might not be limited to procollagen or large cargo proteins and might occur in concert with other transport mechanisms such as long range transport (which has also been observed with other recombinant GFP-tagged procollagen (Stephens and Pepperkok, 2002). To allow for better visualisation of small or dynamic GFP-COL puncta in the proximity of the Golgi where there is an increased background signal from ER-resident GFP-COL cells could be photobleached to improve the signal-to-noise ratio. Given the limitations of detection of at least 10 - 20 GFP molecules per voxel to detect a signal, transport of only few procollagen trimers might remain below the detection limit. Transport of several procollagen trimers per COPII bud seem like a more efficient solution for trafficking procollagen, especially with the high concentration of GFP-COL observed at ERES prior to filling of the Golgi. This would also allow cells to quickly and efficiently transport large amounts of procollagen necessary during wound healing. Large cargo loads of procollagen at ERES might also not require specific recognition via receptor proteins like TANGO1L to be transported (which agrees with findings in TANGO1L-KO fibroblasts) and may instead traffic to the next compartment via a bulk flow (Barlowe and Helenius, 2016). Experimental evidence for bulk flow is hard to obtain, since it relies on proving the non-involvement of interactors like cargo receptors (Barlowe and Helenius, 2016).

It also has to be taken into consideration that the imaging techniques used have limitations in both temporal (≤1 fps) and spatial resolution (90 – 200 nm for STED and confocal microscopy, respectively). Therefore, these techniques do not allow for the visualisation of small (<350 nm), potentially highly transient, COPII carriers packaging or encapsulating procollagen. One way to capture transport intermediates in such confined spaces with sufficient temporal and spatial resolution is to perform lattice light sheet microscopy, followed by CLEM using focused-ion-beam milling and scanning EM at different timepoints during RUSH assays of GFP-COL. Alternatively, fixed samples could be expanded in acrylamide gels (Chen et al., 2015) and analysed with super resolution light microscopy to investigate the three-dimensional shapes and connections of ER-Golgi transport intermediates by using a general membrane stain in combination with markers for ERES and procollagen.

Another caveat is that despite the fact that at steady state KDEL-bound proteins are observed to be solely ER-resident in RPE cells, it is possible that some population of proteins with a KDEL sequence, including the Str-KDEL hook used, does cycle between the ER and ERGIC/*cis*-Golgi. Of note, the absence of ascorbate prior to release from the hook further prevents abundant ER-Golgi transport of procollagen allowing for an additional control step. To rule out that the GFP-COL accumulating prior to release from the hook is not already in the ERGIC, an alternative membrane bound ER-hook could be used like e.g. a streptavidin-Sec61 construct. This model could also

be further tested using low temperature blocks that block ERGIC to Golgi transport, while allowing ER to ERGIC trafficking (Klumperman et al., 1998).

One major difficulty is to elucidate which of the impacts on the secretory pathway perturb procollagen transport directly or are just a side effect from reduced efficiency that only become visible when analysing more complex systems like procollagen trafficking. It seems likely that the capability of an ERES to transport procollagen is dependent on the efficient interplay of all proteins involved.

To better understand the role of key players at ERES and other factors including proteins of the Mia family, it is important to investigate the individual requirements for different factors such as TANGO1 isotypes. Some effort has been made to elucidate this; cTAGE5-KO has been shown to have a stronger impact on the secretion of lipoproteins compared to TANGO1L. To decipher the precise roles of TANGO1S and TANGO1L CRISPR-Cas9-KO could be performed generating cell lines lacking different isoforms of Mia proteins. Potential phenotypes observed in the ECM and in the early secretory pathway should then be investigated for differences. Subsequently, different isoforms of TANGO1S or TANGO1L should be stepwise reintroduced by recombinant expression to analyse rescue effects. This should not only give new insight into the exact role of both isoforms, but also allow a deeper understanding for the requirement of Mia proteins for general cargo secretion opposed to collagen specific effects. Different proteins of the Mia family could potentially act as different cargo receptors for various extracellular matrix proteins.

Analysis of samples from patient-derived cells with mutations in collagen-related conditions can greatly aid the understanding of the mechanisms underlying specific collagenopathies and possibly enable future patient-specific treatment. In addition, uncovering which mutations impact specific steps in procollagen transport will also contribute to the overall biochemical understanding of the early secretory pathway. It would be of great interest to engineer patient mutations into GFP-COL and to investigate the affect these have on trafficking, processing, and assembly into the extracellular matrix.

Future advances in super resolution techniques like expansion and focused-ion-beam scanning EM, as well as three-dimensional cell culture and work with organoids will greatly advance our current understanding of the molecular interactions required for procollagen transport. This will allow for the development of novel therapeutics and diagnostic tools to aid patients with currently incurable diseases like OI, EDS or fibrosis.

References

Acosta-Alvear, D., Y. Zhou, A. Blais, M. Tsikitis, N.H. Lents, C. Arias, C.J. Lennon, Y. Kluger, and B.D. Dynlacht. 2007. XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks. *Mol Cell*. 27:53-66.

Adolf, F., M. Rhiel, I. Reckmann, and F.T. Wieland. 2016. Sec24C/D-isoform-specific sorting of the preassembled ER-Golgi Q-SNARE complex. *Mol Biol Cell*. 27:2697-2707.

Amor, D.J., R. Savarirayan, A.S. Schneider, and A. Bankier. 2000. New case of Cole-Carpenter syndrome. *Am J Med Genet*. 92:273-277.

Andres, D.A., I.M. Dickerson, and J.E. Dixon. 1990. Variants of the carboxyl-terminal KDEL sequence direct intracellular retention. *J Biol Chem*. 265:5952-5955.

Antonny, B., D. Madden, S. Hamamoto, L. Orci, and R. Schekman. 2001. Dynamics of the COPII coat with GTP and stable analogues. *Nat Cell Biol*. 3:531-537.

Appenzeller-Herzog, C., and H.P. Hauri. 2006. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. *J Cell Sci*. 119:2173-2183.

Appenzeller, C., H. Andersson, F. Kappeler, and H.P. Hauri. 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat Cell Biol*. 1:330-334.

Aridor, M., S.I. Bannykh, T. Rowe, and W.E. Balch. 1995. Sequential coupling between COPII and COPI vesicle coats in endoplasmic reticulum to Golgi transport. *J Cell Biol*. 131:875-893.

Aridor, M., K.N. Fish, S. Bannykh, J. Weissman, T.H. Roberts, J. Lippincott-Schwartz, and W.E. Balch. 2001. The Sar1 GTPase coordinates biosynthetic cargo selection with endoplasmic reticulum export site assembly. *J Cell Biol*. 152:213-229.

Aridor, M., J. Weissman, S. Bannykh, C. Nuoffer, and W.E. Balch. 1998. Cargo selection by the COPII budding machinery during export from the ER. *J Cell Biol*. 141:61-70.

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

Asgari, M., N. Latifi, H.K. Heris, H. Vali, and L. Mongeau. 2017. In vitro fibrillogenesis of tropocollagen type III in collagen type I affects its relative fibrillar topology and mechanics. *Sci Rep*. 7:1392.

Axe, E.L., S.A. Walker, M. Manifava, P. Chandra, H.L. Roderick, A. Habermann, G. Griffiths, and N.T. Ktistakis. 2008. Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *J Cell Biol*. 182:685-701.

Baba, M., K. Takeshige, N. Baba, and Y. Ohsumi. 1994. Ultrastructural analysis of the autophagic process in yeast: detection of autophagosomes and their characterization. *J Cell Biol*. 124:903-913.

Bachert, C., and A.D. Linstedt. 2010. Dual anchoring of the GRASP membrane tether promotes trans pairing. *J Biol Chem*. 285:16294-16301.

Bachinger, H.P., P. Bruckner, R. Timpl, and J. Engel. 1978. The role of cis-trans isomerization of peptide bonds in the coil leads to and comes from triple helix conversion of collagen. *Eur J Biochem*. 90:605-613.

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.

Bacia, K., E. Futai, S. Prinz, A. Meister, S. Daum, D. Glatte, J.A. Briggs, and R. Schekman. 2011. Multibudded tubules formed by COPII on artificial liposomes. *Sci Rep*. 1:17.

Balasubramanian, M., E. Hobson, M. Skae, J. McCaughey, and D.J. Stephens. 2019. Developing pathways to clarify pathogenicity of unclassified variants in Osteogenesis Imperfecta genetic analysis. *Mol Genet Genomic Med*. 7:e912.

Balasubramanian, M., R. Padidela, R.C. Pollitt, N.J. Bishop, M.Z. Mughal, A.C. Offiah, B.E. Wagner, J. McCaughey, and D.J. Stephens. 2018. P4HB recurrent missense mutation causing Cole-Carpenter syndrome. *J Med Genet*. 55:158-165.

Balasubramanian, M., R.C. Pollitt, K.E. Chandler, M.Z. Mughal, M.J. Parker, A. Dalton, P. Arundel, A.C. Offiah, and N.J. Bishop. 2015. CRTAP mutation in a patient with Cole-Carpenter syndrome. *Am J Med Genet A*. 167A:587-591.

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

Bannykh, S.I., N. Nishimura, and W.E. Balch. 1998. Getting into the Golgi. *Trends in cell biology*. 8:21-25.

Bannykh, S.I., T. Rowe, and W.E. Balch. 1996. The organization of endoplasmic reticulum export complexes. *J Cell Biol*. 135:19-35.

Barlowe, C., and A. Helenius. 2016. Cargo Capture and Bulk Flow in the Early Secretory Pathway. *Annu Rev Cell Dev Biol*. 32:197-222.

Barlowe, C., L. Orci, T. Yeung, M. Hosobuchi, S. Hamamoto, N. Salama, M.F. Rexach, M. Ravazzola, M. Amherdt, and R. Schekman. 1994. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell*. 77:895-907.

Barlowe, C., and R. Schekman. 1993. SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. *Nature*. 365:347-349.

Barr, F.A., N. Nakamura, and G. Warren. 1998. Mapping the interaction between GRASP65 and GM130, components of a protein complex involved in the stacking of Golgi cisternae. *EMBO J*. 17:3258-3268.

Barr, F.A., M. Puype, J. Vandekerckhove, and G. Warren. 1997. GRASP65, a protein involved in the stacking of Golgi cisternae. *Cell*. 91:253-262.

Beighton, P., A. De Paepe, B. Steinmann, P. Tsipouras, and R.J. Wenstrup. 1998. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am J Med Genet*. 77:31-37.

Bellary, S.S., A. Steinberg, N. Mirzayan, M. Shirak, R.S. Tubbs, A.A. Cohen-Gadol, and M. Loukas. 2013. Wormian bones: a review. *Clin Anat*. 26:922-927.

Ben-Tekaya, H., K. Miura, R. Pepperkok, and H.P. Hauri. 2005. Live imaging of bidirectional traffic from the ERGIC. *J Cell Sci*. 118:357-367.

Benham, A.M. 2012. The protein disulfide isomerase family: key players in health and disease. *Antioxid Redox Signal*. 16:781-789.

Berg, R.A., and D.J. Prockop. 1973a. Purification of (14C) protocollagen and its hydroxylation by prolyl-hydroxylase. *Biochemistry*. 12:3395-3401.

Berg, R.A., and D.J. Prockop. 1973b. The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen. *Biochem Biophys Res Commun*. 52:115-120.

Berger, E.G., K. Grimm, T. Bachi, H. Bosshart, R. Kleene, and M. Watzele. 1993. Double immunofluorescent staining of alpha 2,6 sialyltransferase and beta 1,4 galactosyltransferase in monensin-treated cells: evidence for different Golgi compartments? *J Cell Biochem*. 52:275-288.

Bernales, S., K.L. McDonald, and P. Walter. 2006. Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. *PLoS Biol*. 4:e423.

Bertolotti, A., Y. Zhang, L.M. Hendershot, H.P. Harding, and D. Ron. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat Cell Biol*. 2:326-332.

Betin, V.M., B.K. Singleton, S.F. Parsons, D.J. Anstee, and J.D. Lane. 2013. Autophagy facilitates organelle clearance during differentiation of human erythroblasts: evidence for a role for ATG4 paralogs during autophagosome maturation. *Autophagy*. 9:881-893.

Bevis, B.J., A.T. Hammond, C.A. Reinke, and B.S. Glick. 2002. De novo formation of transitional ER sites and Golgi structures in Pichia pastoris. *Nat Cell Biol*. 4:750-756.

Bi, X., R.A. Corpina, and J. Goldberg. 2002. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature*. 419:271-277.

Bi, X., J.D. Mancias, and J. Goldberg. 2007. Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. *Dev Cell*. 13:635-645.

Bielli, A., C.J. Haney, G. Gabreski, S.C. Watkins, S.I. Bannykh, and M. Aridor. 2005. Regulation of Sar1 NH2 terminus by GTP binding and hydrolysis promotes membrane deformation to control COPII vesicle fission. *J Cell Biol*. 171:919-924.

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.

Birk, D.E. 2001. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron*. 32:223-237.

Birk, D.E., J.M. Fitch, J.P. Babiarz, K.J. Doane, and T.F. Linsenmayer. 1990. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *J Cell Sci*. 95 (Pt 4):649-657.

Birk, D.E., J.M. Fitch, J.P. Babiarz, and T.F. Linsenmayer. 1988. Collagen type I and type V are present in the same fibril in the avian corneal stroma. *J Cell Biol*. 106:999-1008.

Birk, D.E., and R.L. Trelstad. 1984. Extracellular compartments in matrix morphogenesis: collagen fibril, bundle, and lamellar formation by corneal fibroblasts. *J Cell Biol*. 99:2024-2033.

Black, D.D. 2007. Development and physiological regulation of intestinal lipid absorption. I. Development of intestinal lipid absorption: cellular events in chylomicron assembly and secretion. *Am J Physiol Gastrointest Liver Physiol*. 293:G519-524.

Blesch, A., A.K. Bosserhoff, R. Apfel, C. Behl, B. Hessdoerfer, A. Schmitt, P. Jachimczak, F. Lottspeich, R. Buettner, and U. Bogdahn. 1994. Cloning of a novel malignant melanoma-derived growth-regulatory protein, MIA. *Cancer Res*. 54:5695-5701.

Bonadio, J., K.A. Holbrook, R.E. Gelinas, J. Jacob, and P.H. Byers. 1985. Altered triple helical structure of type I procollagen in lethal perinatal osteogenesis imperfecta. *J Biol Chem*. 260:1734-1742.

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.

Boncompain, G., and F. Perez. 2013. The many routes of Golgi-dependent trafficking. *Histochem Cell Biol*. 140:251-260.

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.

Bornstein, P., V. Walsh, J. Tullis, E. Stainbrook, J.F. Bateman, and S.G. Hormuzdi. 2002. The globular domain of the proalpha 1(I) N-propeptide is not required for secretion, processing by procollagen N-proteinase, or fibrillogenesis of type I collagen in mice. *J Biol Chem*. 277:2605-2613.

Bosserhoff, A.K., R. Hein, U. Bogdahn, and R. Buettner. 1996. Structure and promoter analysis of the gene encoding the human melanoma-inhibiting protein MIA. *J Biol Chem*. 271:490-495.

Bottomley, M.J., M.R. Batten, R.A. Lumb, and N.J. Bulleid. 2001. Quality control in the endoplasmic reticulum: PDI mediates the ER retention of unassembled procollagen C-propeptides. *Curr Biol*. 11:1114-1118.

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.

Boyadjiev, S.A., S.D. Kim, A. Hata, C. Haldeman-Englert, E.H. Zackai, C. Naydenov, S. Hamamoto, R.W. Schekman, and J. Kim. 2011. Cranio-lenticulo-sutural dysplasia associated with defects in collagen secretion. *Clin Genet*. 80:169-176.

Brauer, P., J.L. Parker, A. Gerondopoulos, I. Zimmermann, M.A. Seeger, F.A. Barr, and S. Newstead. 2019. Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. *Science*. 363:1103-1107.

Broder, C., P. Arnold, S. Vadon-Le Goff, M.A. Konerding, K. Bahr, S. Muller, C.M. Overall, J.S. Bond, T. Koudelka, A. Tholey, D.J. Hulmes, C. Moali, and C. Becker-Pauly. 2013. Metalloproteases meprin alpha and meprin beta are C- and N-procollagen proteinases important for collagen assembly and tensile strength. *Proc Natl Acad Sci U S A*. 110:14219-14224.

Bruns, R.R., D.J. Hulmes, S.F. Therrien, and J. Gross. 1979. Procollagen segment-long-spacing crystallites: their role in collagen fibrillogenesis. *Proc Natl Acad Sci U S A*. 76:313-317.

Burgess, A., S. Vigneron, E. Brioudes, J.C. Labbe, T. Lorca, and A. Castro. 2010. Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance. *Proc Natl Acad Sci U S A*. 107:12564-12569.

Burrows, N.P., A.C. Nicholls, A.J. Richards, C. Luccarini, J.B. Harrison, J.R. Yates, and F.M. Pope. 1998. A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehlers-Danlos syndrome type II in two British families. *Am J Hum Genet*. 63:390-398.

Butler, D.L., N. Juncosa-Melvin, G.P. Boivin, M.T. Galloway, J.T. Shearn, C. Gooch, and H. Awad. 2008. Functional tissue engineering for tendon repair: A multidisciplinary strategy using mesenchymal stem cells, bioscaffolds, and mechanical stimulation. *J Orthop Res*. 26:1-9.

Byers, P.H., D. Krakow, M.E. Nunes, M. Pepin, and g. American college of medical. 2006. Genetic evaluation of suspected osteogenesis imperfecta (OI). *Genet Med*. 8:383-388.

Cable, E.J., K.G. Onishi, and B.J. Prendergast. 2017. Circadian rhythms accelerate wound healing in female Siberian hamsters. *Physiol Behav*. 171:165-174.

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.
Calfon, M., H. Zeng, F. Urano, J.H. Till, S.R. Hubbard, H.P. Harding, S.G. Clark, and D. Ron. 2002. IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature*. 415:92-96.
Canty-Laird, E.G., Y. Lu, and K.E. Kadler. 2012. Stepwise proteolytic activation of type I procollagen to collagen within the secretory pathway of tendon fibroblasts in situ. *Biochem J*. 441:707-717.
Canty, E.G., and K.E. Kadler. 2005. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci*. 118:1341-1353.
Canty, E.G., Y. Lu, R.S. Meadows, M.K. Shaw, D.F. Holmes, and K.E. Kadler. 2004. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *J Cell Biol*. 165:553-563.
Card, L., N. Henderson, Y. Zhang, P. Bornstein, and A.D. Bradshaw. 2010. Expression in SPARC-null mice of collagen type I lacking the globular domain of the alpha1(I) N-propeptide results in abdominal hernias and loss of dermal collagen. *Matrix Biol*. 29:559-564.
Cermakian, N., and P. Sassone-Corsi. 2000. Multilevel regulation of the circadian clock. *Nat Rev Mol Cell Biol*. 1:59-67.
Chang, J., R. Garva, A. Pickard, C.C. Yeung, V. Mallikarjun, J. Swift, D.F. Holmes, B. Calverley, Y. Lu, A. Adamson, H. Raymond-Hayling, O. Jensen, T. Shearer, Q.J. Meng, and K.E. Kadler. 2020. Circadian control of the secretory pathway maintains collagen homeostasis. *Nat Cell Biol*. 22:74-86.
Chang, S.W., S.J. Shefelbine, and M.J. Buehler. 2012. Structural and mechanical differences between collagen homo- and heterotrimers: relevance for the molecular origin of brittle bone disease. *Biophys J*. 102:640-648.
Chen, F., P.W. Tillberg, and E.S. Boyden. 2015. Optical imaging. Expansion microscopy. *Science*. 347:543-548.
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.
Chessler, S.D., and P.H. Byers. 1992. Defective folding and stable association with protein disulfide isomerase/prolyl hydroxylase of type I procollagen with a deletion in the pro alpha 2(I) chain that preserves the Gly-X-Y repeat pattern. *J Biol Chem*. 267:7751-7757.
Chessler, S.D., and P.H. Byers. 1993. BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta. *J Biol Chem*. 268:18226-18233.
Choudhary, C., J.V. Olsen, C. Brandts, J. Cox, P.N. Reddy, F.D. Bohmer, V. Gerke, D.E. Schmidt-Arras, W.E. Berdel, C. Muller-Tidow, M. Mann, and H. Serve. 2009. Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. *Mol Cell*. 36:326-339.
Christiansen, H.E., U. Schwarze, S.M. Pyott, A. AlSwaid, M. Al Balwi, S. Alrasheed, M.G. Pepin, M.A. Weis, D.R. Eyre, and P.H. Byers. 2010. Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. *Am J Hum Genet*. 86:389-398.
Chung, K.P., Y. Zeng, and L. Jiang. 2016. COPII Paralogs in Plants: Functional Redundancy or Diversity? *Trends Plant Sci*. 21:758-769.
Clark, E.M., and B.A. Link. 2020. Complementary and divergent roles for Ctage5 and Tango1 in zebrafish. *bioRxiv 2020.04.30.070664*.
Cole, D.E., and T.O. Carpenter. 1987. Bone fragility, craniosynostosis, ocular proptosis, hydrocephalus, and distinctive facial features: a newly recognized type of osteogenesis imperfecta. *J Pediatr*. 110:76-80.

Colige, A., F. Ruggiero, I. Vandenberghe, J. Dubail, F. Kesteloot, J. Van Beeumen, A. Beschin, L. Brys, C.M. Lapiere, and B. Nusgens. 2005. Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I-III and V. *J Biol Chem*. 280:34397-34408.

Colige, A., I. Vandenberghe, M. Thiry, C.A. Lambert, J. Van Beeumen, S.W. Li, D.J. Prockop, C.M. Lapiere, and B.V. Nusgens. 2002. Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3. *J Biol Chem*. 277:5756-5766.

Connerly, P.L., M. Esaki, E.A. Montegna, D.E. Strongin, S. Levi, J. Soderholm, and B.S. Glick. 2005. Sec16 is a determinant of transitional ER organization. *Curr Biol*. 15:1439-1447.

Cox, N.J., G. Unlu, B.J. Bisnett, T.R. Meister, B.M. Condon, P.M. Luo, T.J. Smith, M. Hanna, A. Chhetri, E.J. Soderblom, A. Audhya, E.W. Knapik, and M. Boyce. 2018. Dynamic Glycosylation Governs the Vertebrate COPII Protein Trafficking Pathway. *Biochemistry*. 57:91-107.

Cutrona, M.B., G.V. Beznoussenko, A. Fusella, O. Martella, P. Moral, and A.A. Mironov. 2013. Silencing of mammalian Sar1 isoforms reveals COPII-independent protein sorting and transport. *Traffic*. 14:691-708.

Cutrona, M.B., N.E. Morgan, and J.C. Simpson. 2018. Heritable Skeletal Disorders Arising from Defects in Processing and Transport of Type I Procollagen from the ER: Perspectives on Possible Therapeutic Approaches. *Handb Exp Pharmacol*. 245:191-225.

Dannoura, A.H., N. Berriot-Varoqueaux, P. Amati, V. Abadie, N. Verthier, J. Schmitz, J.R. Wetterau, M.E. Samson-Bouma, and L.P. Aggerbeck. 1999. Anderson's disease: exclusion of apolipoprotein and intracellular lipid transport genes. *Arterioscler Thromb Vasc Biol*. 19:2494-2508.

De Paepe, A., and F. Malfait. 2012. The Ehlers-Danlos syndrome, a disorder with many faces. *Clin Genet*. 82:1-11.

De Paepe, A., L. Nuytinck, I. Hausser, I. Anton-Lamprecht, and J.M. Naeyaert. 1997. Mutations in the COL5A1 gene are causal in the Ehlers-Danlos syndromes I and II. *Am J Hum Genet*. 60:547-554.

Dimori, M., M.E. Heard-Lipsmeyer, S.D. Byrum, S.G. Mackintosh, R.C. Kurten, J.L. Carroll, and R. Morello. 2020. Respiratory defects in the CrtapKO mouse model of osteogenesis imperfecta. *Am J Physiol Lung Cell Mol Physiol*. 318:L592-L605.

Dooley, H.C., M. Razi, H.E. Polson, S.E. Girardin, M.I. Wilson, and S.A. Tooze. 2014. WIPI2 links LC3 conjugation with PI3P, autophagosome formation, and pathogen clearance by recruiting Atg12-5-16L1. *Mol Cell*. 55:238-252.

Engel, J., E. Odermatt, A. Engel, J.A. Madri, H. Furthmayr, H. Rohde, and R. Timpl. 1981. Shapes, domain organizations and flexibility of laminin and fibronectin, two multifunctional proteins of the extracellular matrix. *J Mol Biol*. 150:97-120.

Engel, J., and D.J. Prockop. 1991. The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annu Rev Biophys Biophys Chem*. 20:137-152.

Erickson, H.P., N. Carrell, and J. McDonagh. 1981. Fibronectin molecule visualized in electron microscopy: a long, thin, flexible strand. *J Cell Biol*. 91:673-678.

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.

Farhan, H., M. Weiss, K. Tani, R.J. Kaufman, and H.P. Hauri. 2008. Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. *EMBO J*. 27:2043-2054.

Farquhar, M.G., and G.E. Palade. 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends in cell biology*. 8:2-10.
Farsad, K., and P. De Camilli. 2003. Mechanisms of membrane deformation. *Curr Opin Cell Biol*. 15:372-381.
Feinstein, T.N., and A.D. Linstedt. 2008. GRASP55 regulates Golgi ribbon formation. *Mol Biol Cell*. 19:2696-2707.
Fernandes, R.J., S. Hirohata, J.M. Engle, A. Colige, D.H. Cohn, D.R. Eyre, and S.S. Apte. 2001. Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. *J Biol Chem*. 276:31502-31509.
Fisher, L.W., J.T. Stubbs, 3rd, and M.F. Young. 1995. Antisera and cDNA probes to human and certain animal model bone matrix noncollagenous proteins. *Acta Orthop Scand Suppl*. 266:61-65.
Fleischmajer, R., E.D. MacDonald, J.S. Perlish, R.E. Burgeson, and L.W. Fisher. 1990. Dermal collagen fibrils are hybrids of type I and type III collagen molecules. *J Struct Biol*. 105:162-169.
Forlino, A., and J.C. Marini. 2016. Osteogenesis imperfecta. *Lancet*. 387:1657-1671.
Forster, R., M. Weiss, T. Zimmermann, E.G. Reynaud, F. Verissimo, D.J. Stephens, and R. Pepperkok. 2006. Secretory cargo regulates the turnover of COPII subunits at single ER exit sites. *Curr Biol*. 16:173-179.
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.
Fourriere, L., A. Kasri, N. Gareil, S. Bardin, H. Bousquet, D. Pereira, F. Perez, B. Goud, G. Boncompain, and S. Miserey-Lenkei. 2019. RAB6 and microtubules restrict protein secretion to focal adhesions. *J Cell Biol*. 218:2215-2231.
Frantz, C., K.M. Stewart, and V.M. Weaver. 2010. The extracellular matrix at a glance. *J Cell Sci*. 123:4195-4200.
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.
Futai, E., S. Hamamoto, L. Orci, and R. Schekman. 2004. GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on synthetic liposomes. *EMBO J*. 23:4146-4155.
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.
Glorieux, F.H., J.P. Devogelaer, M. Durigova, S. Goemaere, S. Hemsley, F. Jakob, U. Junker, J. Ruckle, L. Seefried, and P.J. Winkle. 2017. BPS804 Anti-Sclerostin Antibody in Adults With Moderate Osteogenesis Imperfecta: Results of a Randomized Phase 2a Trial. *J Bone Miner Res*. 32:1496-1504.
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.
Goldberg, W.M. 1974. Evidence of a sclerotized collagen from the skeleton of a gorgonian coral. *Comp Biochem Physiol B*. 49:525-529.
Golgi, C. 1989. On the structure of nerve cells. 1898. *J Microsc*. 155:3-7.
Gomez-Navarro, N., A. Melero, X.H. Li, J. Boulanger, W. Kukulski, and E.A. Miller. 2020. Cargo crowding contributes to sorting stringency in COPII vesicles. *J Cell Biol*. 219.
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.

Grafe, I., T. Yang, S. Alexander, E.P. Homan, C. Lietman, M.M. Jiang, T. Bertin, E. Munivez, Y. Chen, B. Dawson, Y. Ishikawa, M.A. Weis, T.K. Sampath, C. Ambrose, D. Eyre, H.P. Bachinger, and B. Lee. 2014. Excessive transforming growth factor-beta signaling is a common mechanism in osteogenesis imperfecta. *Nat Med*. 20:670-675.

Griffiths, G., R. Brands, B. Burke, D. Louvard, and G. Warren. 1982. Viral membrane proteins acquire galactose in trans Golgi cisternae during intracellular transport. *J Cell Biol*. 95:781-792.

Griffiths, G., and K. Simons. 1986. The trans Golgi network: sorting at the exit site of the Golgi complex. *Science*. 234:438-443.

Grond, R., T. Veenendaal, J.M. Duran, I. Raote, J.H. van Es, S. Corstjens, L. Delfgou, B. El Haddouti, V. Malhotra, and C. Rabouille. 2020. The function of GORASPs in Golgi apparatus organization in vivo. *J Cell Biol*. 219.

Guillemyn, B., H. Kayserili, L. Demuynck, P. Sips, A. De Paepe, D. Syx, P.J. Coucke, F. Malfait, and S. Symoens. 2019. A homozygous pathogenic missense variant broadens the phenotypic and mutational spectrum of CREB3L1-related osteogenesis imperfecta. *Hum Mol Genet*. 28:1801-1809.

Guo, Q., E. Vasile, and M. Krieger. 1994. Disruptions in Golgi structure and membrane traffic in a conditional lethal mammalian cell mutant are corrected by epsilon-COP. *J Cell Biol*. 125:1213-1224.

Guo, Y., and A.D. Linstedt. 2006. COPII-Golgi protein interactions regulate COPII coat assembly and Golgi size. *J Cell Biol*. 174:53-63.

Ha, C.W., Y.B. Park, S.H. Kim, and H.J. Lee. 2018. Intra-articular Mesenchymal Stem Cells in Osteoarthritis of the Knee: A Systematic Review of Clinical Outcomes and Evidence of Cartilage Repair. *Arthroscopy*. 35:277-288.

Hahm, E., J. Li, K. Kim, S. Huh, S. Rogelj, and J. Cho. 2013. Extracellular protein disulfide isomerase regulates ligand-binding activity of alphaMbeta2 integrin and neutrophil recruitment during vascular inflammation. *Blood*. 121:3789-3800, S3781-3715.

Hammond, A.T., and B.S. Glick. 2000. Dynamics of transitional endoplasmic reticulum sites in vertebrate cells. *Mol Biol Cell*. 11:3013-3030.

Hanna, M.G., J.L. Peotter, E.B. Frankel, and A. Audhya. 2018. Membrane Transport at an Organelle Interface in the Early Secretory Pathway: Take Your Coat Off and Stay a While: Evolution of the metazoan early secretory pathway. *Bioessays*. 40:e1800004.

Hanna, M.G.t., S. Block, E.B. Frankel, F. Hou, A. Johnson, L. Yuan, G. Knight, J.J. Moresco, J.R. Yates, 3rd, R. Ashton, R. Schekman, Y. Tong, and A. Audhya. 2017. TFG facilitates outer coat disassembly on COPII transport carriers to promote tethering and fusion with ER-Golgi intermediate compartments. *Proc Natl Acad Sci U S A*. 114:E7707-E7716.

Hartl, F.U., and M. Hayer-Hartl. 2009. Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol*. 16:574-581.

Harwood, R., M.E. Grant, and D.S. Jackson. 1975. Studies on the glycosylation of hydroxylysine residues during collagen biosynthesis and the subcellular localization of collagen galactosyltransferase and collagen glucosyltransferase in tendon and cartilage cells. *Biochem J*. 152:291-302.

Hass, G. 1940. Studies of collagen: II. Methods and results of implantation of collagen-forming cultures in granulation tissue. *Am J Pathol*. 16:549-560 543.

Hass, G., and F. McDonald. 1940. Studies of collagen: I. The production of collagen in vitro under variable experimental conditions. *Am J Pathol*. 16:525-548.

Hatahet, F., and L.W. Ruddock. 2009. Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation. *Antioxid Redox Signal*. 11:2807-2850.

Hauri, H.P., and A. Schweizer. 1992. The endoplasmic reticulum-Golgi intermediate compartment. *Curr Opin Cell Biol*. 4:600-608.
Heinemeier, K.M., P. Schjerling, J. Heinemeier, S.P. Magnusson, and M. Kjaer. 2013. Lack of tissue renewal in human adult Achilles tendon is revealed by nuclear bomb (14)C. *FASEB J*. 27:2074-2079.
Helaakoski, T., K. Vuori, R. Myllyla, K.I. Kivirikko, and T. Pihlajaniemi. 1989. Molecular cloning of the alpha-subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. *Proc Natl Acad Sci U S A*. 86:4392-4396.
Herzig, Y., H.J. Sharpe, Y. Elbaz, S. Munro, and M. Schuldiner. 2012. A systematic approach to pair secretory cargo receptors with their cargo suggests a mechanism for cargo selection by Erv14. *PLoS Biol*. 10:e1001329.
Hofmann, H., T. Voss, K. Kuhn, and J. Engel. 1984. Localization of flexible sites in thread-like molecules from electron micrographs. Comparison of interstitial, basement membrane and intima collagens. *J Mol Biol*. 172:325-343.
Horlein, D., J. McPherson, S.H. Goh, and P. Bornstein. 1981. Regulation of protein synthesis: translational control by procollagen-derived fragments. *Proc Natl Acad Sci U S A*. 78:6163-6167.
Horstmann, H., C.P. Ng, B.L. Tang, and W. Hong. 2002. Ultrastructural characterization of endoplasmic reticulum--Golgi transport containers (EGTC). *J Cell Sci*. 115:4263-4273.
Hosokawa, N., and K. Nagata. 2000. Procollagen binds to both prolyl 4-hydroxylase/protein disulfide isomerase and HSP47 within the endoplasmic reticulum in the absence of ascorbate. *FEBS Lett*. 466:19-25.
Hoyer-Kuhn, H., C. Netzer, F. Koerber, E. Schoenau, and O. Semler. 2014. Two years' experience with denosumab for children with osteogenesis imperfecta type VI. *Orphanet J Rare Dis*. 9:145.
Hoyle, N.P., E. Seinkmane, M. Putker, K.A. Feeney, T.P. Krogager, J.E. Chesham, L.K. Bray, J.M. Thomas, K. Dunn, J. Blaikley, and J.S. O'Neill. 2017. Circadian actin dynamics drive rhythmic fibroblast mobilization during wound healing. *Sci Transl Med*. 9.
Huang, M., J.T. Weissman, S. Beraud-Dufour, P. Luan, C. Wang, W. Chen, M. Aridor, I.A. Wilson, and W.E. Balch. 2001. Crystal structure of Sar1-GDP at 1.7 A resolution and the role of the NH2 terminus in ER export. *J Cell Biol*. 155:937-948.
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.
Hughes, H., and D.J. Stephens. 2010. Sec16A defines the site for vesicle budding from the endoplasmic reticulum on exit from mitosis. *J Cell Sci*. 123:4032-4038.
Huh, W.K., J.V. Falvo, L.C. Gerke, A.S. Carroll, R.W. Howson, J.S. Weissman, and E.K. O'Shea. 2003. Global analysis of protein localization in budding yeast. *Nature*. 425:686-691.
Hulmes, D.J. 1983. A possible mechanism for the regulation of collagen fibril diameter in vivo. *Coll Relat Res*. 3:317-321.
Hulmes, D.J., K.E. Kadler, A.P. Mould, Y. Hojima, D.F. Holmes, C. Cummings, J.A. Chapman, and D.J. Prockop. 1989. Pleomorphism in type I collagen fibrils produced by persistence of the procollagen N-propeptide. *J Mol Biol*. 210:337-345.
Humphries, J.D., A. Byron, and M.J. Humphries. 2006. Integrin ligands at a glance. *J Cell Sci*. 119:3901-3903.
Humphries, S.M., Y. Lu, E.G. Canty, and K.E. Kadler. 2008. Active negative control of collagen fibrillogenesis in vivo. Intracellular cleavage of the type I procollagen

propeptides in tendon fibroblasts without intracellular fibrils. *J Biol Chem*. 283:12129-12135.
Hutchings, J., V. Stancheva, E.A. Miller, and G. Zanetti. 2018. Subtomogram averaging of COPII assemblies reveals how coat organization dictates membrane shape. *Nat Commun*. 9:4154.
Hutchings, J., and G. Zanetti. 2019. Coat flexibility in the secretory pathway: a role in transport of bulky cargoes. *Curr Opin Cell Biol*. 59:104-111.
Hynes, R.O., and Q. Zhao. 2000. The evolution of cell adhesion. *J Cell Biol*. 150:F89-96.
Ignotz, R.A., T. Endo, and J. Massague. 1987. Regulation of fibronectin and type I collagen mRNA levels by transforming growth factor-beta. *J Biol Chem*. 262:6443-6446.
Ishida, Y., H. Kubota, A. Yamamoto, A. Kitamura, H.P. Bachinger, and K. Nagata. 2006. Type I collagen in Hsp47-null cells is aggregated in endoplasmic reticulum and deficient in N-propeptide processing and fibrillogenesis. *Mol Biol Cell*. 17:2346-2355.
Ishikawa, T., T. Toyama, Y. Nakamura, K. Tamada, H. Shimizu, S. Ninagawa, T. Okada, Y. Kamei, T. Ishikawa-Fujiwara, T. Todo, E. Aoyama, M. Takigawa, A. Harada, and K. Mori. 2017. UPR transducer BBF2H7 allows export of type II collagen in a cargo- and developmental stage-specific manner. *J Cell Biol*. 216:1761-1774.
Ishikawa, Y., and H.P. Bachinger. 2013. A molecular ensemble in the rER for procollagen maturation. *Biochim Biophys Acta*. 1833:2479-2491.
Ishikawa, Y., S. Ito, K. Nagata, L.Y. Sakai, and H.P. Bächinger. 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.
Ishikawa, Y., K. Rubin, H.P. Bachinger, and S. Kalamajski. 2018. The endoplasmic reticulum-resident collagen chaperone Hsp47 interacts with and promotes the secretion of decorin, fibromodulin, and lumican. *J Biol Chem*. 293:13707-13716.
Ito, S., and K. Nagata. 2017. Biology of Hsp47 (Serpin H1), a collagen-specific molecular chaperone. *Semin Cell Dev Biol*. 62:142-151.
Ivan, V., G. de Voer, D. Xanthakis, K.M. Spoorendonk, V. Kondylis, and C. Rabouille. 2008. Drosophila Sec16 mediates the biogenesis of tER sites upstream of Sar1 through an arginine-rich motif. *Mol Biol Cell*. 19:4352-4365.
Iwasaki, H., T. Yorimitsu, and K. Sato. 2017. Microscopy analysis of reconstituted COPII coat polymerization and Sec16 dynamics. *J Cell Sci*. 130:2893-2902.
Jakobsson, M., S.W. Scholz, P. Scheet, J.R. Gibbs, J.M. VanLiere, H.C. Fung, Z.A. Szpiech, J.H. Degnan, K. Wang, R. Guerreiro, J.M. Bras, J.C. Schymick, D.G. Hernandez, B.J. Traynor, J. Simon-Sanchez, M. Matarin, A. Britton, J. van de Leemput, I. Rafferty, M. Bucan, H.M. Cann, J.A. Hardy, N.A. Rosenberg, and A.B. Singleton. 2008. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*. 451:998-1003.
Jansen, K.A., A.J. Licup, A. Sharma, R. Rens, F.C. MacKintosh, and G.H. Koenderink. 2018. The Role of Network Architecture in Collagen Mechanics. *Biophys J*. 114:2665-2678.
Jayadev, R., and D.R. Sherwood. 2017. Basement membranes. *Curr Biol*. 27:R207-R211.
Jensen, T.J., M.A. Loo, S. Pind, D.B. Williams, A.L. Goldberg, and J.R. Riordan. 1995. Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. *Cell*. 83:129-135.
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.

Johannes, L., and V. Popoff. 2008. Tracing the retrograde route in protein trafficking. *Cell*. 135:1175-1187.

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.

Johnston, J.M., B.K. Connizzo, S.S. Shetye, K.A. Robinson, J. Huegel, A.B. Rodriguez, M. Sun, S.M. Adams, D.E. Birk, and L.J. Soslowsky. 2017. Collagen V haploinsufficiency in a murine model of classic Ehlers-Danlos syndrome is associated with deficient structural and mechanical healing in tendons. *J Orthop Res*. 35:2707-2715.

Jones, B., E.L. Jones, S.A. Bonney, H.N. Patel, A.R. Mensenkamp, S. Eichenbaum-Voline, M. Rudling, U. Myrdal, G. Annesi, S. Naik, N. Meadows, A. Quattrone, S.A. Islam, R.P. Naoumova, B. Angelin, R. Infante, E. Levy, C.C. Roy, P.S. Freemont, J. Scott, and C.C. Shoulders. 2003. Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nat Genet*. 34:29-31.

Kadler, K.E., A. Hill, and E.G. Canty-Laird. 2008. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol*. 20:495-501.

Kadler, K.E., Y. Hojima, and D.J. Prockop. 1987. Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. *J Biol Chem*. 262:15696-15701.

Kagan, H.M., and P.C. Trackman. 1991. Properties and function of lysyl oxidase. *Am J Respir Cell Mol Biol*. 5:206-210.

Kalson, N.S., Y. Lu, S.H. Taylor, T. Starborg, D.F. Holmes, and K.E. Kadler. 2015. A structure-based extracellular matrix expansion mechanism of fibrous tissue growth. *Elife*. 4.

Kanadome, T., H. Shibata, K. Kuwata, T. Takahara, and M. Maki. 2017. The calcium-binding protein ALG-2 promotes endoplasmic reticulum exit site localization and polymerization of Trk-fused gene (TFG) protein. *FEBS J*. 284:56-76.

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.

Katayama, K., M. Kuriki, T. Kamiya, Y. Tochigi, and H. Suzuki. 2018. Giantin is required for coordinated production of aggrecan, link protein and type XI collagen during chondrogenesis. *Biochem Biophys Res Commun*. 499:459-465.

Katayama, K., T. Sasaki, S. Goto, K. Ogasawara, H. Maru, K. Suzuki, and H. Suzuki. 2011. Insertional mutation in the Golgb1 gene is associated with osteochondrodysplasia and systemic edema in the OCD rat. *Bone*. 49:1027-1036.

Kawaguchi, K., A. Endo, T. Fukushima, Y. Madoka, T. Tanaka, and M. Komada. 2018. Ubiquitin-specific protease 8 deubiquitinates Sec31A and decreases large COPII carriers and collagen IV secretion. *Biochem Biophys Res Commun*. 499:635-641.

Kawasaki, K., R. Ushioda, S. Ito, K. Ikeda, Y. Masago, and K. Nagata. 2015. Deletion of the collagen-specific molecular chaperone Hsp47 causes endoplasmic reticulum stress-mediated apoptosis of hepatic stellate cells. *J Biol Chem*. 290:3639-3646.

Kay, E.P. 1986. Rabbit corneal endothelial cells modulated by polymorphonuclear leukocytes are fibroblasts. Comparison with keratocytes. *Invest Ophthalmol Vis Sci*. 27:891-897.

Keene, D.R., L.Y. Sakai, G.P. Lunstrum, N.P. Morris, and R.E. Burgeson. 1987. Type VII collagen forms an extended network of anchoring fibrils. *J Cell Biol*. 104:611-621.

Keller, P., and K. Simons. 1997. Post-Golgi biosynthetic trafficking. *J Cell Sci*. 110 (Pt 24):3001-3009.

Kerwar, S.S., G.J. Cardinale, L.D. Kohn, C.L. Spears, and F.L. Stassen. 1973. Cell-free synthesis of procollagen: L-929 fibroblasts as a cellular model for dermatosparaxis. *Proc Natl Acad Sci U S A*. 70:1378-1382.

Kessler, E., K. Takahara, L. Biniaminov, M. Brusel, and D.S. Greenspan. 1996. Bone morphogenetic protein-1: the type I procollagen C-proteinase. *Science*. 271:360-362.

Kikukawa, K., T. Kamei, K. Suzuki, and K. Maita. 1990. Electron microscopic observations and electrophoresis of the glycosaminoglycans in the epiphyseal cartilage of the congenital osteochondrodysplasia rat (ocd/ocd). *Matrix*. 10:378-387.

Kikukawa, K., and K. Suzuki. 1992. Histochemical and immunohistochemical distribution of glycosaminoglycans, type II collagen, and fibronectin in developing fetal cartilage of congenital osteochondrodysplasia rat (ocd/ocd). *Teratology*. 46:509-523.

Kim, K., S. Park, and J. Kim. 2018. Cullin3-RING ubiquitin ligases are intimately linked to the unfolded protein response of the endoplasmic reticulum.

Kim, S.D., K.B. Pahuja, M. Ravazzola, J. Yoon, S.A. Boyadjiev, S. Hammamoto, R. Schekman, L. Orci, and J. Kim. 2012. The [corrected] SEC23-SEC31 [corrected] interface plays critical role for export of procollagen from the endoplasmic reticulum. *J Biol Chem*. 287:10134-10144.

Klumperman, J., A. Schweizer, H. Clausen, B.L. Tang, W. Hong, V. Oorschot, and H.P. Hauri. 1998. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *J Cell Sci*. 111 (Pt 22):3411-3425.

Knipper, J.A., S. Willenborg, J. Brinckmann, W. Bloch, T. Maass, R. Wagener, T. Krieg, T. Sutherland, A. Munitz, M.E. Rothenberg, A. Niehoff, R. Richardson, M. Hammerschmidt, J.E. Allen, and S.A. Eming. 2015. Interleukin-4 Receptor alpha Signaling in Myeloid Cells Controls Collagen Fibril Assembly in Skin Repair. *Immunity*. 43:803-816.

Koide, T., S. Asada, and K. Nagata. 1999. Substrate recognition of collagen-specific molecular chaperone HSP47. Structural requirements and binding regulation. *J Biol Chem*. 274:34523-34526.

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.

Koreishi, M., T.J. Gniadek, S. Yu, J. Masuda, Y. Honjo, and A. Satoh. 2013. The golgin tether giantin regulates the secretory pathway by controlling stack organization within Golgi apparatus. *PLoS One*. 8:e59821.

Kuehn, M.J., and R. Schekman. 1997. COPII and secretory cargo capture into transport vesicles. *Curr Opin Cell Biol*. 9:477-483.

Kung, L.F., S. Pagant, E. Futai, J.G. D'Arcangelo, R. Buchanan, J.C. Dittmar, R.J. Reid, R. Rothstein, S. Hamamoto, E.L. Snapp, R. Schekman, and E.A. Miller. 2012.

Sec24p and Sec16p cooperate to regulate the GTP cycle of the COPII coat. *EMBO J*. 31:1014-1027.
Kurokawa, K., and A. Nakano. 2019. The ER exit sites are specialized ER zones for the transport of cargo proteins from the ER to the Golgi apparatus. *J Biochem*. 165:109-114.
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.
Kweon, H.S., G.V. Beznoussenko, M. Micaroni, R.S. Polishchuk, A. Trucco, O. Martella, D. Di Giandomenico, P. Marra, A. Fusella, A. Di Pentima, E.G. Berger, W.J. Geerts, A.J. Koster, K.N. Burger, A. Luini, and A.A. Mironov. 2004. Golgi enzymes are enriched in perforated zones of golgi cisternae but are depleted in COPI vesicles. *Mol Biol Cell*. 15:4710-4724.
Lakshmanachetty, S., and M.I. Koster. 2016. Emerging roles for collagen XV and XVIII in cancer progression. *Exp Dermatol*. 25:346-347.
LaMantia, M., T. Miura, H. Tachikawa, H.A. Kaplan, W.J. Lennarz, and T. Mizunaga. 1991. Glycosylation site binding protein and protein disulfide isomerase are identical and essential for cell viability in yeast. *Proc Natl Acad Sci U S A*. 88:4453-4457.
Lan, Y., N. Zhang, H. Liu, J. Xu, and R. Jiang. 2016. Golgb1 regulates protein glycosylation and is crucial for mammalian palate development. *Development*. 143:2344-2355.
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.
Lanoix, J., J. Ouwendijk, C.C. Lin, A. Stark, H.D. Love, J. Ostermann, and T. Nilsson. 1999. GTP hydrolysis by arf-1 mediates sorting and concentration of Golgi resident enzymes into functional COP I vesicles. *EMBO J*. 18:4935-4948.
Layman, D.L., and R. Ross. 1973. The production and secretion of procollagen peptidase by human fibroblasts in culture. *Arch Biochem Biophys*. 157:451-456.
Leblond, C.P. 1989. Synthesis and secretion of collagen by cells of connective tissue, bone, and dentin. *Anat Rec*. 224:123-138.
Lees, J.F., M. Tasab, and N.J. Bulleid. 1997. Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *EMBO J*. 16:908-916.
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.
Letourneur, F., E.C. Gaynor, S. Hennecke, C. Demolliere, R. Duden, S.D. Emr, H. Riezman, and P. Cosson. 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell*. 79:1199-1207.
Leung, M.K., L.I. Fessler, D.B. Greenberg, and J.H. Fessler. 1979. Separate amino and carboxyl procollagen peptidases in chick embryo tendon. *J Biol Chem*. 254:224-232.
Li, L., Y. Shen, Y. Ding, Y. Liu, D. Su, and X. Liang. 2014. Hrd1 participates in the regulation of collagen I synthesis in renal fibrosis. *Mol Cell Biochem*. 386:35-44.
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.

Lim, J., I. Grafe, S. Alexander, and B. Lee. 2017. Genetic causes and mechanisms of Osteogenesis Imperfecta. *Bone*. 102:40-49.

Linsenmayer, T.F., E. Gibney, F. Igoe, M.K. Gordon, J.M. Fitch, L.I. Fessler, and D.E. Birk. 1993. Type V collagen: molecular structure and fibrillar organization of the chicken alpha 1(V) NH2-terminal domain, a putative regulator of corneal fibrillogenesis. *J Cell Biol*. 121:1181-1189.

Linstedt, A.D., and H.P. Hauri. 1993. Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Mol Biol Cell*. 4:679-693.

Lisse, T.S., F. Thiele, H. Fuchs, W. Hans, G.K. Przemeck, K. Abe, B. Rathkolb, L. Quintanilla-Martinez, G. Hoelzlwimmer, M. Helfrich, E. Wolf, S.H. Ralston, and M. Hrabe de Angelis. 2008. ER stress-mediated apoptosis in a new mouse model of osteogenesis imperfecta. *PLoS Genet*. 4:e7.

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.

Losev, E., C.A. Reinke, J. Jellen, D.E. Strongin, B.J. Bevis, and B.S. Glick. 2006. Golgi maturation visualized in living yeast. *Nature*. 441:1002-1006.

Lotti, L.V., M.R. Torrisi, M.C. Pascale, and S. Bonatti. 1992. Immunocytochemical analysis of the transfer of vesicular stomatitis virus G glycoprotein from the intermediate compartment to the Golgi complex. *J Cell Biol*. 118:43-50.

Love, H.D., C.C. Lin, C.S. Short, and J. Ostermann. 1998. Isolation of functional Golgi-derived vesicles with a possible role in retrograde transport. *J Cell Biol*. 140:541-551.

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*.

Lumb, R.A., and N.J. Bulleid. 2002. Is protein disulfide isomerase a redox-dependent molecular chaperone? *EMBO J*. 21:6763-6770.

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.

Malfait, F., C. Francomano, P. Byers, J. Belmont, B. Berglund, J. Black, L. Bloom, J.M. Bowen, A.F. Brady, N.P. Burrows, M. Castori, H. Cohen, M. Colombi, S. Demirdas, J. De Backer, A. De Paepe, S. Fournel-Gigleux, M. Frank, N. Ghali, C. Giunta, R. Grahame, A. Hakim, X. Jeunemaitre, D. Johnson, B. Juul-Kristensen, I. Kapferer-Seebacher, H. Kazkaz, T. Kosho, M.E. Lavallee, H. Levy, R. Mendoza-Londono, M. Pepin, F.M. Pope, E. Reinstein, L. Robert, M. Rohrbach, L. Sanders, G.J. Sobey, T. Van Damme, A. Vandersteen, C. van Mourik, N. Voermans, N. Wheeldon, J. Zschocke, and B. Tinkle. 2017. The 2017 international classification of the Ehlers-Danlos syndromes. *Am J Med Genet C Semin Med Genet*. 175:8-26.

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.

Man, K., A. Loudon, and A. Chawla. 2016. Immunity around the clock. *Science*. 354:999-1003.
Marchi, F., and C.P. Leblond. 1984. Radioautographic characterization of successive compartments along the rough endoplasmic reticulum-Golgi pathway of collagen precursors in foot pad fibroblasts of [3H]proline-injected rats. *J Cell Biol*. 98:1705-1709.
Marini, J.C., A. Forlino, H.P. Bachinger, N.J. Bishop, P.H. Byers, A. Paepe, F. Fassier, N. Fratzl-Zelman, K.M. Kozloff, D. Krakow, K. Montpetit, and O. Semler. 2017. Osteogenesis imperfecta. *Nat Rev Dis Primers*. 3:17052.
Marom, R., B.M. Rabenhorst, and R. Morello. 2020. Osteogenesis imperfecta: an update on clinical features and therapies. *Eur J Endocrinol*. 183:R95-106.
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.
Marutani, T., A. Yamamoto, N. Nagai, H. Kubota, and K. Nagata. 2004. Accumulation of type IV collagen in dilated ER leads to apoptosis in Hsp47-knockout mouse embryos via induction of CHOP. *J Cell Sci*. 117:5913-5922.
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. 2018. COPII-dependent ER export in animal cells: adaptation and control for diverse cargo. *Histochem Cell Biol*. 150:119-131.
McCaughey, J., and D.J. Stephens. 2019. ER-to-Golgi Transport: A Sizeable Problem. *Trends in cell biology*. 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.
McCracken, A.A., and J.L. Brodsky. 1996. Assembly of ER-associated protein degradation in vitro: dependence on cytosol, calnexin, and ATP. *J Cell Biol*. 132:291-298.
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.
Melville, D.B., M. Montero-Balaguer, D.S. Levic, K. Bradley, J.R. Smith, A.K. Hatzopoulos, and E.W. Knapik. 2011. The feelgood mutation in zebrafish dysregulates COPII-dependent secretion of select extracellular matrix proteins in skeletal morphogenesis. *Dis Model Mech*. 4:763-776.
Mezzacasa, A., and A. Helenius. 2002. The transitional ER defines a boundary for quality control in the secretion of tsO45 VSV glycoprotein. *Traffic*. 3:833-849.
Michalickova, K., M. Susic, M.C. Willing, R.J. Wenstrup, and W.G. Cole. 1998. Mutations of the alpha2(V) chain of type V collagen impair matrix assembly and produce ehlers-danlos syndrome type I. *Hum Mol Genet*. 7:249-255.
Miller, E., B. Antonny, S. Hamamoto, and R. Schekman. 2002. Cargo selection into COPII vesicles is driven by the Sec24p subunit. *EMBO J*. 21:6105-6113.
Miller, E.A., T.H. Beilharz, P.N. Malkus, M.C. Lee, S. Hamamoto, L. Orci, and R. Schekman. 2003. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell*. 114:497-509.
Miller, E.A., and R. Schekman. 2013. COPII - a flexible vesicle formation system. *Curr Opin Cell Biol*. 25:420-427.

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.

Miyahara, M., F.K. Njieha, and D.J. Prockop. 1982. Formation of collagen fibrils in vitro by cleavage of procollagen with procollagen proteinases. *J Biol Chem*. 257:8442-8448.

Montegna, E.A., M. Bhave, Y. Liu, D. Bhattacharyya, and B.S. Glick. 2012. Sec12 binds to Sec16 at transitional ER sites. *PLoS One*. 7:e31156.

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*.

Morello, R., T.K. Bertin, Y. Chen, J. Hicks, L. Tonachini, M. Monticone, P. Castagnola, F. Rauch, F.H. Glorieux, J. Vranka, H.P. Bachinger, J.M. Pace, U. Schwarze, P.H. Byers, M. Weis, R.J. Fernandes, D.R. Eyre, Z. Yao, B.F. Boyce, and B. Lee. 2006. CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell*. 127:291-304.

Morgan, P.H., H.G. Jacobs, J.P. Segrest, and L.W. Cunningham. 1970. A comparative study of glycopeptides derived from selected vertebrate collagens. A possible role of the carbohydrate in fibril formation. *J Biol Chem*. 245:5042-5048.

Mori, R., T.J. Shaw, and P. Martin. 2008. Molecular mechanisms linking wound inflammation and fibrosis: knockdown of osteopontin leads to rapid repair and reduced scarring. *J Exp Med*. 205:43-51.

Morito, D., and K. Nagata. 2015. Pathogenic Hijacking of ER-Associated Degradation: Is ERAD Flexible? *Mol Cell*. 59:335-344.

Morris, J.L., S.J. Cross, Y. Lu, K.E. Kadler, Y. Lu, S.L. Dallas, and P. Martin. 2018. Live imaging of collagen deposition during skin development and repair in a collagen I - GFP fusion transgenic zebrafish line. *Dev Biol*. 441:4-11.

Morris, N.P., D.R. Keene, R.W. Glanville, H. Bentz, and R.E. Burgeson. 1986. The tissue form of type VII collagen is an antiparallel dimer. *J Biol Chem*. 261:5638-5644.

Moser, M., A.K. Bosserhoff, E.B. Hunziker, L. Sandell, R. Fassler, and R. Buettner. 2002. Ultrastructural cartilage abnormalities in MIA/CD-RAP-deficient mice. *Mol Cell Biol*. 22:1438-1445.

Munro, S. 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harb Perspect Biol*. 3.

Murad, S., D. Grove, K.A. Lindberg, G. Reynolds, A. Sivarajah, and S.R. Pinnell. 1981. Regulation of collagen synthesis by ascorbic acid. *Proc Natl Acad Sci U S A*. 78:2879-2882.

Murakami, T., A. Saito, S. Hino, S. Kondo, S. Kanemoto, K. Chihara, H. Sekiya, K. Tsumagari, K. Ochiai, K. Yoshinaga, M. Saitoh, R. Nishimura, T. Yoneda, I. Kou, T. Furuichi, S. Ikegawa, M. Ikawa, M. Okabe, A. Wanaka, and K. Imaizumi. 2009. Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. *Nat Cell Biol*. 11:1205-1211.

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.

Nagai, N., M. Hosokawa, S. Itohara, E. Adachi, T. Matsushita, N. Hosokawa, and K. Nagata. 2000. Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis. *J Cell Biol*. 150:1499-1506.

Nagata, K., S. Saga, and K.M. Yamada. 1986. A major collagen-binding protein of chick embryo fibroblasts is a novel heat shock protein. *J Cell Biol*. 103:223-229.

Nakagawa, T., M. Setou, D. Seog, K. Ogasawara, N. Dohmae, K. Takio, and N. Hirokawa. 2000. A novel motor, KIF13A, transports mannose-6-phosphate receptor to plasma membrane through direct interaction with AP-1 complex. *Cell*. 103:569-581.

Nakano, A., and M. Muramatsu. 1989. A novel GTP-binding protein, Sar1p, is involved in transport from the endoplasmic reticulum to the Golgi apparatus. *J Cell Biol*. 109:2677-2691.

Nakano, A., H. Otsuka, M. Yamagishi, E. Yamamoto, K. Kimura, S. Nishikawa, and T. Oka. 1994. Mutational analysis of the Sar1 protein, a small GTPase which is essential for vesicular transport from the endoplasmic reticulum. *J Biochem*. 116:243-247.

Neeli, I., S.A. Siddiqi, S. Siddiqi, J. Mahan, W.S. Lagakos, B. Binas, T. Gheyi, J. Storch, and C.M. Mansbach, 2nd. 2007. Liver fatty acid-binding protein initiates budding of pre-chylomicron transport vesicles from intestinal endoplasmic reticulum. *J Biol Chem*. 282:17974-17984.

Nicholls, A.C., J.E. Oliver, S. McCarron, J.B. Harrison, D.S. Greenspan, and F.M. Pope. 1996. An exon skipping mutation of a type V collagen gene (COL5A1) in Ehlers-Danlos syndrome. *J Med Genet*. 33:940-946.

Nijenhuis, W., M.M.P. van Grinsven, and L.C. Kapitein. 2020. An optimized toolbox for the optogenetic control of intracellular transport. *J Cell Biol*. 219.

Niu, X., C. Gao, L. Jan Lo, Y. Luo, C. Meng, J. Hong, W. Hong, and J. Peng. 2012. Sec13 safeguards the integrity of the endoplasmic reticulum and organogenesis of the digestive system in zebrafish. *Dev Biol*. 367:197-207.

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.

Oganesian, A., S. Au, J.A. Horst, L.C. Holzhausen, A.J. Macy, J.M. Pace, and P. Bornstein. 2006. The NH2-terminal propeptide of type I procollagen acts intracellularly to modulate cell function. *J Biol Chem*. 281:38507-38518.

Ohisa, S., K. Inohaya, Y. Takano, and A. Kudo. 2010. sec24d encoding a component of COPII is essential for vertebra formation, revealed by the analysis of the medaka mutant, vbi. *Dev Biol*. 342:85-95.

Oka, T., and A. Nakano. 1994. Inhibition of GTP hydrolysis by Sar1p causes accumulation of vesicles that are a functional intermediate of the ER-to-Golgi transport in yeast. *J Cell Biol*. 124:425-434.

Okamoto, M., K. Kurokawa, K. Matsuura-Tokita, C. Saito, R. Hirata, and A. Nakano. 2012. High-curvature domains of the ER are important for the organization of ER exit sites in Saccharomyces cerevisiae. *J Cell Sci*. 125:3412-3420.

Omari, S., E. Makareeva, L. Gorrell, M. Jarnik, J. Lippincott-Schwartz, and S. Leikin. 2020. Mechanisms of procollagen and HSP47 sorting during ER-to-Golgi trafficking. *Matrix Biol*.

Omari, S., E. Makareeva, A. Roberts-Pilgrim, L. Mirigian, M. Jarnik, C. Ott, J. Lippincott-Schwartz, and S. Leikin. 2018. Noncanonical autophagy at ER exit sites regulates procollagen turnover. *Proc Natl Acad Sci U S A*. 115:E10099-E10108.

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.

Orci, L., M. Ravazzola, M. Amherdt, A. Perrelet, S.K. Powell, D.L. Quinn, and H.P. Moore. 1987. The trans-most cisternae of the Golgi complex: a compartment for sorting of secretory and plasma membrane proteins. *Cell*. 51:1039-1051.

Ozbek, S., P.G. Balasubramanian, R. Chiquet-Ehrismann, R.P. Tucker, and J.C. Adams. 2010. The evolution of extracellular matrix. *Mol Biol Cell*. 21:4300-4305.

Pagant, S., A. Wu, S. Edwards, F. Diehl, and E.A. Miller. 2015. Sec24 is a coincidence detector that simultaneously binds two signals to drive ER export. *Curr Biol*. 25:403-412.

Paglia, L., J. Wilczek, L.D. de Leon, G.R. Martin, D. Horlein, and P. Muller. 1979. Inhibition of procollagen cell-free synthesis by amino-terminal extension peptides. *Biochemistry*. 18:5030-5034.

Pakshir, P., and B. Hinz. 2018. The big five in fibrosis: Macrophages, myofibroblasts, matrix, mechanics, and miscommunication. *Matrix Biol*. 68-69:81-93.

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.

Parker, M.W., D. Rossi, M. Peterson, K. Smith, K. Sikström, E.S. White, J.E. Connett, C.A. Henke, O. Larsson, and P.B. Bitterman. 2014. Fibrotic extracellular matrix activates a profibrotic positive feedback loop. *J Clin Invest*. 124:1622-1635.

Paulsson, M., M. Morgelin, H. Wiedemann, M. Beardmore-Gray, D. Dunham, T. Hardingham, D. Heinegard, R. Timpl, and J. Engel. 1987. Extended and globular protein domains in cartilage proteoglycans. *Biochem J*. 245:763-772.

Pavelka, M., J. Neumuller, and A. Ellinger. 2008. Retrograde traffic in the biosynthetic-secretory route. *Histochem Cell Biol*. 129:277-288.

Pearse, B.M. 1975. Coated vesicles from pig brain: purification and biochemical characterization. *J Mol Biol*. 97:93-98.

Pepperkok, R., J. Scheel, H. Horstmann, H.P. Hauri, G. Griffiths, and T.E. Kreis. 1993. Beta-COP is essential for biosynthetic membrane transport from the endoplasmic reticulum to the Golgi complex in vivo. *Cell*. 74:71-82.

Petrosyan, A., M.S. Holzapfel, D.E. Muirhead, and P.W. Cheng. 2014. Restoration of compact Golgi morphology in advanced prostate cancer enhances susceptibility to galectin-1-induced apoptosis by modifying mucin O-glycan synthesis. *Mol Cancer Res*. 12:1704-1716.

Piez, K.A., and A. Miller. 1974. The structure of collagen fibrils. *J Supramol Struct*. 2:121-137.

Pihlajaniemi, T., L.A. Dickson, F.M. Pope, V.R. Korhonen, A. Nicholls, D.J. Prockop, and J.C. Myers. 1984. Osteogenesis imperfecta: cloning of a pro-alpha 2(I) collagen gene with a frameshift mutation. *J Biol Chem*. 259:12941-12944.

Pincus, D., M.W. Chevalier, T. Aragon, E. van Anken, S.E. Vidal, H. El-Samad, and P. Walter. 2010. BiP binding to the ER-stress sensor Ire1 tunes the homeostatic behavior of the unfolded protein response. *PLoS Biol*. 8:e1000415.

Pinnell, S.R., R. Fox, and S.M. Krane. 1971. Human collagens: differences in glycosylated hydroxylysines in skin and bone. *Biochim Biophys Acta*. 229:119-122.

Polishchuk, E.V., A. Di Pentima, A. Luini, and R.S. Polishchuk. 2003. Mechanism of constitutive export from the golgi: bulk flow via the formation, protrusion, and en bloc cleavage of large trans-golgi network tubular domains. *Mol Biol Cell*. 14:4470-4485.

Polishchuk, R.S., E.V. Polishchuk, P. Marra, S. Alberti, R. Buccione, A. Luini, and A.A. Mironov. 2000. Correlative light-electron microscopy reveals the tubular-saccular ultrastructure of carriers operating between Golgi apparatus and plasma membrane. *J Cell Biol*. 148:45-58.

Ponnambalam, S., and S.A. Baldwin. 2003. Constitutive protein secretion from the trans-Golgi network to the plasma membrane. *Mol Membr Biol*. 20:129-139.

Presley, J.F., T.H. Ward, A.C. Pfeifer, E.D. Siggia, R.D. Phair, and J. Lippincott-Schwartz. 2002. Dissection of COPI and Arf1 dynamics in vivo and role in Golgi membrane transport. *Nature*. 417:187-193.

Prockop, D.J., K.I. Kivirikko, L. Tuderman, and N.A. Guzman. 1979a. The biosynthesis of collagen and its disorders (first of two parts). *N Engl J Med*. 301:13-23.

Prockop, D.J., K.I. Kivirikko, L. Tuderman, and N.A. Guzman. 1979b. The biosynthesis of collagen and its disorders (second of two parts). *N Engl J Med*. 301:77-85.

Puthenveedu, M.A., C. Bachert, S. Puri, F. Lanni, and A.D. Linstedt. 2006. GM130 and GRASP65-dependent lateral cisternal fusion allows uniform Golgi-enzyme distribution. *Nat Cell Biol*. 8:238-248.

Quinn, P., G. Griffiths, and G. Warren. 1983. Dissection of the Golgi complex. II. Density separation of specific Golgi functions in virally infected cells treated with monensin. *J Cell Biol*. 96:851-856.

Rambourg, A., Y. Clermont, and L. Hermo. 1979. Three-dimensional architecture of the golgi apparatus in Sertoli cells of the rat. *Am J Anat*. 154:455-476.

Ramshaw, J.A., N.K. Shah, and B. Brodsky. 1998. Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *J Struct Biol*. 122:86-91.

Raote, I., A.M. Ernst, F. Campelo, J.E. Rothman, F. Pincet, and V. Malhotra. 2020. TANGO1 membrane helices create a lipid diffusion barrier at curved membranes. *Elife*. 9.

Raote, I., and V. Malhotra. 2019. Protein transport by vesicles and tunnels. *J Cell Biol*. 218:737-739.

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.

Rauch, F., S. Fahiminiya, J. Majewski, J. Carrot-Zhang, S. Boudko, F. Glorieux, J.S. Mort, H.P. Bachinger, and P. Moffatt. 2015. Cole-Carpenter syndrome is caused by a heterozygous missense mutation in P4HB. *Am J Hum Genet*. 96:425-431.

Rauch, F., and F.H. Glorieux. 2004. Osteogenesis imperfecta. *Lancet*. 363:1377-1385.

Rayl, M., M. Truitt, A. Held, J. Sargeant, K. Thorsen, and J.C. Hay. 2016. Penta-EF-Hand Protein Peflin Is a Negative Regulator of ER-To-Golgi Transport. *PLoS One*. 11:e0157227.

Reynolds, H.M., L. Zhang, D.T. Tran, and K.G. Ten Hagen. 2019. Tango1 coordinates the formation of endoplasmic reticulum/Golgi docking sites to mediate secretory granule formation. *J Biol Chem*. 294:19498-19510.

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

Ricard-Blum, S. 2011. The collagen family. *Cold Spring Harb Perspect Biol*. 3:a004978.

Richards, A.J., S. Martin, A.C. Nicholls, J.B. Harrison, F.M. Pope, and N.P. Burrows. 1998. A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type II. *J Med Genet*. 35:846-848.

Richards, S., N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W.W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H.L. Rehm, and A.L.Q.A. Committee. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 17:405-424.

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

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.

Roenneberg, T., and M. Merrow. 2005. Circadian clocks: translation. *Curr Biol*. 15:R470-473.

Romanic, A.M., E. Adachi, Y. Hojima, J. Engel, and D.J. Prockop. 1992. Polymerization of pNcollagen I and copolymerization of pNcollagen I with collagen I. A kinetic, thermodynamic, and morphologic study. *J Biol Chem*. 267:22265-22271.

Rossert, J., C. Terraz, and S. Dupont. 2000. Regulation of type I collagen genes expression. *Nephrol Dial Transplant*. 15 Suppl 6:66-68.

Roth, J., and E.G. Berger. 1982. Immunocytochemical localization of galactosyltransferase in HeLa cells: codistribution with thiamine pyrophosphatase in trans-Golgi cisternae. *J Cell Biol*. 93:223-229.

Roth, T.F., and K.R. Porter. 1964. Yolk Protein Uptake in the Oocyte of the Mosquito Aedes Aegypti. L. *J Cell Biol*. 20:313-332.

Ruggiano, A., O. Foresti, and P. Carvalho. 2014. Quality control: ER-associated degradation: protein quality control and beyond. *J Cell Biol*. 204:869-879.

Saegusa, K., M. Sato, N. Morooka, T. Hara, and K. Sato. 2018. SFT-4/Surf4 control ER export of soluble cargo proteins and participate in ER exit site organization. *J Cell Biol*.

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.

Saito, K., K. Yamashiro, N. Shimazu, T. Tanabe, K. Kontani, and T. Katada. 2014. Concentration of Sec12 at ER exit sites via interaction with cTAGE5 is required for collagen export. *J Cell Biol*. 206:751-762.

Sakai, L.Y., D.R. Keene, R.W. Glanville, and H.P. Bachinger. 1991. Purification and partial characterization of fibrillin, a cysteine-rich structural component of connective tissue microfibrils. *J Biol Chem*. 266:14763-14770.

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.

Saraste, J., and E. Kuismanen. 1992. Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex. *Semin Cell Biol*. 3:343-355.

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.

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.

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.
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.
Scheiermann, C., Y. Kunisaki, and P.S. Frenette. 2013. Circadian control of the immune system. *Nat Rev Immunol*. 13:190-198.
Schievink, W.I., V.V. Michels, and D.G. Piepgras. 1994. Neurovascular manifestations of heritable connective tissue disorders. A review. *Stroke*. 25:889-903.
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.
Schlacht, A., and J.B. Dacks. 2015. Unexpected ancient paralogs and an evolutionary model for the COPII coat complex. *Genome Biol Evol*. 7:1098-1109.
Schmid, R., S. Schiffner, A. Opolka, S. Grassel, T. Schubert, M. Moser, and A.K. Bosserhoff. 2010. Enhanced cartilage regeneration in MIA/CD-RAP deficient mice. *Cell Death Dis*. 1:e97.
Schmid, T.M., and T.F. Linsenmayer. 1983. A short chain (pro)collagen from aged endochondral chondrocytes. Biochemical characterization. *J Biol Chem*. 258:9504-9509.
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.
Schweizer, A., K. Matter, C.M. Ketcham, and H.P. Hauri. 1991. The isolated ER-Golgi intermediate compartment exhibits properties that are different from ER and cis-Golgi. *J Cell Biol*. 113:45-54.
Seelig, H.P., P. Schranz, H. Schroter, C. Wiemann, and M. Renz. 1994. Macrogolgin--a new 376 kD Golgi complex outer membrane protein as target of antibodies in patients with rheumatic diseases and HIV infections. *J Autoimmun*. 7:67-91.
Sesso, A., F.P. de Faria, E.S. Iwamura, and H. Correa. 1994. A three-dimensional reconstruction study of the rough ER-Golgi interface in serial thin sections of the pancreatic acinar cell of the rat. *J Cell Sci*. 107 (Pt 3):517-528.
Shapiro, J.R., M.L. Stover, V.E. Burn, M.B. McKinstry, A.L. Burshell, S.D. Chipman, and D.W. Rowe. 1992. An osteopenic nonfracture syndrome with features of mild osteogenesis imperfecta associated with the substitution of a cysteine for glycine at triple helix position 43 in the pro alpha 1(I) chain of type I collagen. *J Clin Invest*. 89:567-573.
Sharma, U., L. Carrique, S. Vadon-Le Goff, N. Mariano, R.N. Georges, F. Delolme, P. Koivunen, J. Myllyharju, C. Moali, N. Aghajari, and D.J. Hulmes. 2017. Structural basis of homo- and heterotrimerization of collagen I. *Nat Commun*. 8:14671.
Shen, J., E.L. Snapp, J. Lippincott-Schwartz, and R. Prywes. 2005. Stable binding of ATF6 to BiP in the endoplasmic reticulum stress response. *Mol Cell Biol*. 25:921-932.
Shen, X., R.E. Ellis, K. Lee, C.Y. Liu, K. Yang, A. Solomon, H. Yoshida, R. Morimoto, D.M. Kurnit, K. Mori, and R.J. Kaufman. 2001. Complementary signaling pathways regulate the unfolded protein response and are required for C. elegans development. *Cell*. 107:893-903.
Shibata, H., T. Inuzuka, H. Yoshida, H. Sugiura, I. Wada, and M. Maki. 2010. The ALG-2 binding site in Sec31A influences the retention kinetics of Sec31A at the

endoplasmic reticulum exit sites as revealed by live-cell time-lapse imaging. *Biosci Biotechnol Biochem*. 74:1819-1826.

Shibata, H., H. Suzuki, H. Yoshida, and M. Maki. 2007. ALG-2 directly binds Sec31A and localizes at endoplasmic reticulum exit sites in a Ca2+-dependent manner. *Biochem Biophys Res Commun*. 353:756-763.

Shima, D.T., S.J. Scales, T.E. Kreis, and R. Pepperkok. 1999. Segregation of COPI-rich and anterograde-cargo-rich domains in endoplasmic-reticulum-to-Golgi transport complexes. *Curr Biol*. 9:821-824.

Siddiqi, S.A., F.S. Gorelick, J.T. Mahan, and C.M. Mansbach, 2nd. 2003. COPII proteins are required for Golgi fusion but not for endoplasmic reticulum budding of the pre-chylomicron transport vesicle. *J Cell Sci*. 116:415-427.

Sillence, D.O., A. Senn, and D.M. Danks. 1979. Genetic heterogeneity in osteogenesis imperfecta. *J Med Genet*. 16:101-116.

Smith, K., and M.J. Rennie. 2007. New approaches and recent results concerning human-tissue collagen synthesis. *Curr Opin Clin Nutr Metab Care*. 10:582-590.

Sonnichsen, B., R. Watson, H. Clausen, T. Misteli, and G. Warren. 1996. Sorting by COP I-coated vesicles under interphase and mitotic conditions. *J Cell Biol*. 134:1411-1425.

Stagg, S.M., C. Gurkan, D.M. Fowler, P. LaPointe, T.R. Foss, C.S. Potter, B. Carragher, and W.E. Balch. 2006. Structure of the Sec13/31 COPII coat cage. *Nature*. 439:234-238.

Stagg, S.M., P. LaPointe, A. Razvi, C. Gurkan, C.S. Potter, B. Carragher, and W.E. Balch. 2008. Structural basis for cargo regulation of COPII coat assembly. *Cell*. 134:474-484.

Stephens, D.J., N. Lin-Marq, A. Pagano, R. Pepperkok, and J.P. Paccaud. 2000. COPI-coated ER-to-Golgi transport complexes segregate from COPII in close proximity to ER exit sites. *J Cell Sci*. 113 (Pt 12):2177-2185.

Stephens, D.J., and R. Pepperkok. 2001. Illuminating the secretory pathway: when do we need vesicles? *J Cell Sci*. 114:1053-1059.

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.

Stevenson, N.L., D.J.M. Bergen, C.L. Hammond, and D.J. Stephens. 2020. Giantin is required for 1 intracellular N-terminal processing of type I procollagen. *bioRxiv 2020.05.25.115279*.

Stevenson, N.L., D.J.M. Bergen, R.E.H. Skinner, E. Kague, E. Martin-Silverstone, K.A. Robson Brown, C.L. Hammond, and D.J. Stephens. 2017. Giantin-knockout models reveal a feedback loop between Golgi function and glycosyltransferase expression. *J Cell Sci*. 130:4132-4143.

Subramanian, A., A. Capalbo, N.R. Iyengar, R. Rizzo, A. di Campli, R. Di Martino, M. Lo Monte, A.R. Beccari, A. Yerudkar, C. Del Vecchio, L. Glielmo, G. Turacchio, M. Pirozzi, S.G. Kim, P. Henklein, J. Cancino, S. Parashuraman, D. Diviani, F. Fanelli, M. Sallese, and A. Luini. 2019. Auto-regulation of Secretory Flux by Sensing and Responding to the Folded Cargo Protein Load in the Endoplasmic Reticulum. *Cell*. 176:1461-1476 e1423.

Suda, Y., and A. Nakano. 2012. The yeast Golgi apparatus. *Traffic*. 13:505-510.

Symoens, S., F. Malfait, S. D'Hondt, B. Callewaert, A. Dheedene, W. Steyaert, H.P. Bachinger, A. De Paepe, H. Kayserili, and P.J. Coucke. 2013. Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans. *Orphanet J Rare Dis*. 8:154.

Takeshige, K., M. Baba, S. Tsuboi, T. Noda, and Y. Ohsumi. 1992. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *J Cell Biol*. 119:301-311.

Tanzawa, K., J. Berger, and D.J. Prockop. 1985. Type I procollagen N-proteinase from whole chick embryos. Cleavage of a homotrimer of pro-alpha 1(I) chains and the requirement for procollagen with a triple-helical conformation. *J Biol Chem*. 260:1120-1126.

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.

Teng, J., T. Rai, Y. Tanaka, Y. Takei, T. Nakata, M. Hirasawa, A.B. Kulkarni, and N. Hirokawa. 2005. The KIF3 motor transports N-cadherin and organizes the developing neuroepithelium. *Nat Cell Biol*. 7:474-482.

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. *Journal of Cell Science*. 125:673-684.

Townley, A.K., and D.J. Stephens. 2009. Vesicle coating and uncoating: controlling the formation of large COPII-coated carriers. *F1000 Biol Rep*. 1:65.

Trelstad, R.L. 1971. Vacuoles in the embryonic chick corneal epithelium, an epithelium which produces collagen. *J Cell Biol*. 48:689-694.

Trelstad, R.L., and A.J. Coulombre. 1971. Morphogenesis of the collagenous stroma in the chick cornea. *J Cell Biol*. 50:840-858.

Trelstad, R.L., and K. Hayashi. 1979. Tendon collagen fibrillogenesis: intracellular subassemblies and cell surface changes associated with fibril growth. *Dev Biol*. 71:228-242.

Trucco, A., R.S. Polishchuk, O. Martella, A. Di Pentima, A. Fusella, D. Di Giandomenico, E. San Pietro, G.V. Beznoussenko, E.V. Polishchuk, M. Baldassarre, R. Buccione, W.J. Geerts, A.J. Koster, K.N. Burger, A.A. Mironov, and A. Luini. 2004. Secretory traffic triggers the formation of tubular continuities across Golgi sub-compartments. *Nat Cell Biol*. 6:1071-1081.

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.

Unlu, G., D.S. Levic, D.B. Melville, and E.W. Knapik. 2014. Trafficking mechanisms of extracellular matrix macromolecules: insights from vertebrate development and human diseases. *Int J Biochem Cell Biol*. 47:57-67.

van Bergeijk, P., M. Adrian, C.C. Hoogenraad, and L.C. Kapitein. 2015. Optogenetic control of organelle transport and positioning. *Nature*. 518:111-114.

Van Damme, T., A. Colige, D. Syx, C. Giunta, U. Lindert, M. Rohrbach, O. Aryani, Y. Alanay, P.O. Simsek-Kiper, H.Y. Kroes, K. Devriendt, M. Thiry, S. Symoens, A. De Paepe, and F. Malfait. 2016. Expanding the clinical and mutational spectrum of the Ehlers-Danlos syndrome, dermatosparaxis type. *Genet Med*. 18:882-891.

Velling, T., J. Risteli, K. Wennerberg, D.F. Mosher, and S. Johansson. 2002. Polymerization of type I and III collagens is dependent on fibronectin and enhanced by integrins alpha 11beta 1 and alpha 2beta 1. *J Biol Chem*. 277:37377-37381.

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.

Vogel, B.E., R. Doelz, K.E. Kadler, Y. Hojima, J. Engel, and D.J. Prockop. 1988. A substitution of cysteine for glycine 748 of the alpha 1 chain produces a kink at this site in the procollagen I molecule and an altered N-proteinase cleavage site over 225 nm away. *J Biol Chem*. 263:19249-19255.

Volchuk, A., M. Amherdt, M. Ravazzola, B. Brugger, V.M. Rivera, T. Clackson, A. Perrelet, T.H. Sollner, J.E. Rothman, and L. Orci. 2000. Megavesicles implicated in the rapid transport of intracisternal aggregates across the Golgi stack. *Cell*. 102:335-348.

Vranka, J.A., L.Y. Sakai, and H.P. Bachinger. 2004. Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J Biol Chem*. 279:23615-23621.

Walmsley, A.R., M.R. Batten, U. Lad, and N.J. Bulleid. 1999. Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. *J Biol Chem*. 274:14884-14892.

Walsh, M.J., N.S. LeLeiko, and K.M. Sterling, Jr. 1987. Regulation of types I, III, and IV procollagen mRNA synthesis in glucocorticoid-mediated intestinal development. *J Biol Chem*. 262:10814-10818.

Ward, C.L., S. Omura, and R.R. Kopito. 1995. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell*. 83:121-127.

Watson, P., R. Forster, K.J. Palmer, R. Pepperkok, and D.J. Stephens. 2005. Coupling of ER exit to microtubules through direct interaction of COPII with dynactin. *Nat Cell Biol*. 7:48-55.

Watson, P., A.K. Townley, P. Koka, K.J. Palmer, and D.J. Stephens. 2006. Sec16 defines endoplasmic reticulum exit sites and is required for secretory cargo export in mammalian cells. *Traffic*. 7:1678-1687.

Weinstock, M., and C.P. Leblond. 1974. Synthesis, migration, and release of precursor collagen by odontoblasts as visualized by radioautography after (3H)proline administration. *J Cell Biol*. 60:92-127.

Weiskirchen, R., S. Weiskirchen, and F. Tacke. 2019. Organ and tissue fibrosis: Molecular signals, cellular mechanisms and translational implications. *Mol Aspects Med*. 65:2-15.

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.

Wenstrup, R.J., G.T. Langland, M.C. Willing, V.N. D'Souza, and W.G. Cole. 1996. A splice-junction mutation in the region of COL5A1 that codes for the carboxyl propeptide of pro alpha 1(V) chains results in the gravis form of the Ehlers-Danlos syndrome (type I). *Hum Mol Genet*. 5:1733-1736.

Whittle, J.R., and T.U. Schwartz. 2010. Structure of the Sec13-Sec16 edge element, a template for assembly of the COPII vesicle coat. *J Cell Biol*. 190:347-361.

Wickstrom, S.A., K. Radovanac, and R. Fassler. 2011. Genetic analyses of integrin signaling. *Cold Spring Harb Perspect Biol*. 3.

Wierzbicka-Patynowski, I., and J.E. Schwarzbauer. 2003. The ins and outs of fibronectin matrix assembly. *J Cell Sci*. 116:3269-3276.

Wilkin, M.B., M.N. Becker, D. Mulvey, I. Phan, A. Chao, K. Cooper, H.J. Chung, I.D. Campbell, M. Baron, and R. MacIntyre. 2000. Drosophila Dumpy is a gigantic extracellular protein required to maintain tension at epidermal-cuticle attachment sites. *Current Biology*. 10:559-567.

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.

Wilson, V.J., M. Rattray, C.R. Thomas, B.H. Moreland, and D. Schulster. 1998. Effects of hypophysectomy and growth hormone administration on the mRNA levels of collagen I, III and insulin-like growth factor-I in rat skeletal muscle. *Growth Horm IGF Res*. 8:431-438.

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.

Wong, K.K., R.J. deLeeuw, N.S. Dosanjh, L.R. Kimm, Z. Cheng, D.E. Horsman, C. MacAulay, R.T. Ng, C.J. Brown, E.E. Eichler, and W.L. Lam. 2007. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet*. 80:91-104.

Wong, M., and S. Munro. 2014. Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. *Science*. 346:1256898.

Wong, M.Y., and M.D. Shoulders. 2019. Targeting defective proteostasis in the collagenopathies. *Curr Opin Chem Biol*. 50:80-88.

Wu, M., K. Cronin, and J.S. Crane. 2020. Biochemistry, Collagen Synthesis. *In* StatPearls, Treasure Island (FL).

Wynn, T.A. 2007. Common and unique mechanisms regulate fibrosis in various fibroproliferative diseases. *J Clin Invest*. 117:524-529.

Xiang, Y., and Y. Wang. 2010. GRASP55 and GRASP65 play complementary and essential roles in Golgi cisternal stacking. *J Cell Biol*. 188:237-251.

Yamasaki, A., K. Tani, A. Yamamoto, N. Kitamura, and M. Komada. 2006. The Ca2+-binding protein ALG-2 is recruited to endoplasmic reticulum exit sites by Sec31A and stabilizes the localization of Sec31A. *Mol Biol Cell*. 17:4876-4887.

Yeung, C.-Y.C., R. Garva, A. Pickard, J. Chang, D.F. Holmes, Y. Lu, V. Mallikarjun, J. Swift, A. Adamson, B. Calverley, Q.J. Meng, and K.E. Kadler. 2018. Circadian Clock Regulation of the Secretory Pathway. *bioRxiv*:304014.

Yeung, C.Y., N. Gossan, Y. Lu, A. Hughes, J.J. Hensman, M.L. Bayer, M. Kjaer, K.E. Kadler, and Q.J. Meng. 2014. Gremlin-2 is a BMP antagonist that is regulated by the circadian clock. *Sci Rep*. 4:5183.

Yoo, Y.S., H.G. Han, and Y.J. Jeon. 2017. Unfolded Protein Response of the Endoplasmic Reticulum in Tumor Progression and Immunogenicity. *Oxid Med Cell Longev*. 2017:2969271.

Yorimitsu, T., and K. Sato. 2012. Insights into structural and regulatory roles of Sec16 in COPII vesicle formation at ER exit sites. *Mol Biol Cell*. 23:2930-2942.

Yoshida, H., T. Matsui, A. Yamamoto, T. Okada, and K. Mori. 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell*. 107:881-891.

Yoshihisa, T., C. Barlowe, and R. Schekman. 1993. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic reticulum. *Science*. 259:1466-1468.

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

Yuan, L., S.J. Kenny, J. Hemmati, K. Xu, and R. Schekman. 2018. TANGO1 and SEC12 are copackaged with procollagen I to facilitate the generation of large COPII carriers. *Proc Natl Acad Sci*. 115:E12255.

Yurchenco, P.D. 2011. Basement membranes: cell scaffoldings and signaling platforms. *Cold Spring Harb Perspect Biol*. 3.

Zanetti, G., K.B. Pahuja, S. Studer, S. Shim, and R. Schekman. 2011. COPII and the regulation of protein sorting in mammals. *Nat Cell Biol*. 14:20-28.

Zanetti, G., S. Prinz, S. Daum, A. Meister, R. Schekman, K. Bacia, and J.A. Briggs. 2013. The structure of the COPII transport-vesicle coat assembled on membranes. *Elife*. 2:e00951.

Zeng, Y., K.P. Chung, B. Li, C.M. Lai, S.K. Lam, X. Wang, Y. Cui, C. Gao, M. Luo, K.B. Wong, R. Schekman, and L. Jiang. 2015. Unique COPII component AtSar1a/AtSec23a pair is required for the distinct function of protein ER export in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*. 112:14360-14365.

Zhu, M., J. Tao, M.P. Vasievich, W. Wei, G. Zhu, R.N. Khoriaty, and B. Zhang. 2015. Neural tube opening and abnormal extraembryonic membrane development in SEC23A deficient mice. *Sci Rep*. 5:15471.

Appendix

Published peer-reviewed journal articles and reviews relevant to this work:

1. McCaughey, J., and D.J. Stephens. 2018. COPII-dependent ER export in animal cells: adaptation and control for diverse cargo. Histochem Cell Biol. 150:119-131.

2. McCaughey, J., and D.J. Stephens. 2019. ER-to-Golgi Transport: A Sizeable Problem. Trends in cell biology. 29:940-953.

3. 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.

4. Balasubramanian, M., E. Hobson, M. Skae, **J. McCaughey**, and D.J. Stephens. 2019. Developing pathways to clarify pathogenicity of unclassified variants in Osteogenesis Imperfecta genetic analysis. Mol Genet Genomic Med. 7:e912.

5. Balasubramanian, M., R. Padidela, R.C. Pollitt, N.J. Bishop, M.Z. Mughal, A.C. Offiah, B.E. Wagner, **J. McCaughey**, and D.J. Stephens. 2018. P4HB recurrent missense mutation causing Cole-Carpenter syndrome. J Med Genet. 55:158-165.