TANGO1 assembles a machine for collagen folding and export

Anupama Ashok

TESI DOCTORAL UPF / 2018

Thesis supervisor

Dr.Vivek Malhotra

Cell and Developmental Biology Programme

Centre for Genomic Regulation (CRG)

My Parents
Vatsala and Ashok

Acknowledgements

Vivek, for giving me the opportunity to do a PhD in his lab. There are two things that I have learnt from you and will take with me moving forward. One, how to give a talk that is clear, concise and captivating. The other advice, which I think saves a lot of time, money and effort is to do the one key experiment to test the hypothesis.

Members of my thesis committee, Juan, Jerome, Salva and Pedro, thank you for your valuable feedback and support over the years.

My defense tribunal members, Carlos, Miguel, Juan, Jerome and Oriol, thanks for being a part of the panel.

Members of TANGO team, past and present, Patrik, Marghe Cristina, Antonio, Abira, Maria, Ishier and Ombry thank you all for your guidance and advice. It was great to have a team who struggled alongside me to work on an ambitious project with proteins as challenging as TANGO1 and Collagen.

Ombry, I wish I had you in the lab from when I started the PhD. Thank you for all your advice and input in the design of experiments, discussions, and writing of the thesis.

Ishier, the senior most member of the TANGO1 team, thank you for all your input and discussions during the course of my PhD.

Mari, you were the only other person there working alongside me in the lab all those late nights, weekends and holidays. I am profoundly grateful for all your advice and help over the last few years both in science and in my personal life.

Juan, I learnt so much from you, and over the years you have become one of my dearest friends. Thank you for all the discussions and feedback and for making sure I was doing ok every single day during the thesis writing process.

Amy, thank you for all the input, corrections and moral support during the thesis writing process and during my years in the lab. Thanks for giving me the non-TANGO person perspective on all things TANGO.

Natalie, my former bench mate, it was a lot of fun sharing a bench with you. You always knew how to lighten up the atmosphere. I will miss our dosa dinners. Thanks for all your advice and input over the years.

Gerard for being a wonderful colleague, you always made time for me to discuss experiments, guided me on all things Calcium and Catalan.
Carla, it was great to finally have another PhD student in the lab, I wish we had more years together, so that I can pass on my wisdom to you.
Julien, one of the kindest, most supportive colleagues I have ever known. Collaborating on Collagen XVII and your screen was a great learning experience.
Former members of the lab Cristi, Gabriel, David, Felix, Josse and Sandra.
Imma, you are absolutely amazing. The students of the CRG are lucky to have you. I cannot thank you enough for all your encouragement, support and advice through the years. You always had time for me and helped me through some really tough moments.
Gloria from HR, for her help and timely advice regarding all things administrative.
Rut, our secretary for all her help over the years.
Fellow PhD students Bogu and Krystal, you guys kept me optimistic and helped me power through the very stressful writing process.
Friends and colleagues in the department, who were always available despite their busy schedules to help me and have supported me over the years. The friends from cell culture room 590, thank you all for making the long hours in the cell culture entertaining.
I also would like to thank all my friends and colleagues from the CRG for their support and guidance.
The la Caixa fellowship for funding me for the last 4 years.

Colin,your kindness and support helped me persevere and made this work possible.
My grandfather Neelkant, his memory continues to be a source of inspiration.
My parents Ashok and Vatsala, I would not have made it through without your love and unwavering faith in me. You encouraged me to be positive and hopeful through the most disheartening moments. My brother Harish, I can never thank you enough for being there for me every single day. My cat Bundu, for all her unconditional love.

Abstract

COPII vesicles of 60-90nm diameter are known to export secretory cargoes from endoplasmic reticulum (ER). However, they cannot be employed to export bulky cargoes like the collagens that can reach up to 400 nm in length. Collagens, the most abundant secretory proteins, make up 25% of our dry body weight and required for building extracellular matrix, and to produce mineralized bones. The discovery TANGO1 has made the process of collagen export at the ER amenable to molecular analysis. I set out to identify its interactors through a proximity biotinylation coupled with mass spectrometry approach.

My results show that TANGO1 bridges the cytoplasmic export machinery with ER luminal folding machinery. It is noteworthy that several of the luminal interactors identified are exclusively dedicated to collagen folding and modification. This search also revealed the identity of protein called Torsin-1A, and my data show that it potentially functions in degrading unfolded collagens. I also asked whether TANGO1 is required to export transmembrane collagens. Transmembrane collagens are unique as they have a cytoplasmic domain that can in theory recruit COPII proteins to facilitate their export. I observed minimal dependency on TANGO1 for transmembrane collagen XVII export. Interestingly, TANGO1 binds both folded and misfolded Collagen XVII.

Altogether, my data suggest that TANGO1 functions predominantly to export soluble collagens its ability to bind both folded and unfolded collagens is used by the cells to eliminate unfolded collagens and to ensure only the full assembled and functional collagen are secreted.

Resum

Les vesícules COPII, que tenen un diàmetre de 60-90nm, transporten les proteïnes que seran secretades des del reticle endoplasmàtic (RE). No obstant, aquestes vesícules no poden ser usades per transportar proteïnes molt voluminoses, com els col·làgens que poden arribar a mesurar 400 nm de longitud. Els col·làgens, el tipus de proteïna secretada més abundant, suposen el 25% del pes sec del nostre cos i són necessaris per construir la matriu extracel·lular, així com per generar els ossos mineralitzats.

El descobriment de la proteïna TANGO1 ha permès l'anàlisi molecular del procés d'exportació dels col·làgens des del RE. Un dels objectius d'aquesta tesi era identificar quines proteïnes interaccionen amb TANGO1 usant un enfocament de biotinilació per proximitat conjugat amb una espectrometria de masses. Els meus resultats han mostrat que TANGO1 serveix de connexió entre la maquinària d'exportació citoplasmàtica i la maquinària de plegament de la llum del RE. És important destacar que una quantitat important de les proteïnes identificades que interaccionen amb TANGO1 al RE estan dedicades exclusivament al correcte plegament i modificació dels col·làgens. Aquesta investigació també ha revelat la identitat d'una proteïna anomenada Torsin-1A, i les meves dades han demostrat que la seva funció potencial és en la degradació de col·làgens que no han sigut correctament plegats. Una altra de les preguntes que he respost en aquesta tesi és si TANGO1 és necessari també per l'exportació de col·làgens amb un domini transmembrana.

Aquests col·làgens són únics ja que tenen un domini citoplasmàtic que en teoria pot reclutar directament les proteïnes COPII per facilitar la seva exportació del RE. Els meus resultats han mostrat una dependència mínima en TANGO1 per l'exportació del col·lagen amb domini transmembrana XVII. No obstant, és interessant remarcar que TANGO1 pot unir col·lagen XVII independentment de si està correctament o incorrectament plegat.

En conclusió, les meves dades suggereixen que TANGO1 funciona predominantment per exportar col·làgens solubles, i la seva habilitat per unir col·làgens tant si estan correctament o incorrectament plegats és usada per les

cèl·lules per eliminar aquells col·làgens mal plegats per assegurar que només els col·làgens totalment estructurats i funcionals seran secretats.

Preface

Our understanding of how proteins in the cell are transported from one compartment to the next and then to the extracellular space has vastly improved since it was first described by George Palade.

Proteins which are synthesized in the endoplasmic reticulum, are transported out of specialized delivery ports called ER exit sites and in carriers which are 60-90 nm in diameter called COP II vesicles. A combination of several techniques and years of work by several research groups have helped us gain a very clear mechanistic understanding of COP II mediated vesicular transport.

Yet, one of the biggest challenges in the field has been the export of collagens. Despite being one of the most abundant secretory cargoes that is essential for a variety of functions, nothing much was known about it export. This is partly due to its large size and complexity and the technical challenges involved in working with these proteins. The identification of TANGO1 and its requirement for collagen secretion has begun to help us gain mechanistic insights into how large cargoes are exported from the ER.

My aim when I began the thesis was to identify the interactome of TANGO1 which would then help us understand how it functions. My work reveals that TANGO1 works by bridging the cytoplasmic export machinery with the ER luminal folding machinery. It also points to a larger role for TANGO1 in the quality control of collagens.

This work was carried out in the Cell and Developmental Biology program of The Centre for Genomic Regulation (CRG) under the supervision of Dr.Vivek Malhotra.

Table of Contents

Introduction

1. Introduction

1.1 The Secretory pathway

The classical secretory pathway contains a series of organelles/compartments which participate in the synthesis, folding, modification and delivery of proteins to their correct destinations. The endoplasmic reticulum (ER) is the entry port for the journey of cargoes along the secretory pathway. Newly synthesized membrane proteins or signal sequence containing secretory proteins are translocated co- or post-translationally into the ER (Rapoport, 2007). The proteins then undergo folding and assembly, a process that is overseen by an extensive group of enzymes and chaperones. Some of the vital ER chaperones include Calnexin, Calreticulin, GRP94 and BiP, with BiP as the major player in the protein folding machinery of the ER. BiP functions in multiple aspects of protein folding regulation, including binding and retention of misfolded proteins (Hurtley, 1989), the unfolded protein response (UPR) and ER associated degradation (ERAD) (Hammond et al., 1994; Pincus et al., 2010). GRP94 also participates in the processes of folding and quality control, but unlike BiP, which chaperones most ER proteins, GRP94 has a relatively small set of substrates that includes collagens, integrins, and Toll-like receptors (Marzec et al., 2012) . An important feature of the folding process in the ER is protein modifications such as disulfide bonds and N-linked glycosylation. The chaperones calnexin and calreticulin bind the N-glycosylated proteins to ensure their proper folding. Most misfolded proteins retained by the quality control mechanisms of the ER are finally exported for degradation.

The fully folded proteins are then exported from the ER by coat protein complex II (COPII) coated vesicles, which are 60-90nm in diameter, from specially designed sites called ER exit sites (ERES). This process of COPII vesicle generation, starts with the protein Sec12, a guanine-nucleotide exchange factor (GEF) which activates the GTPase Sar1 (Barlowe and Schekman, 1993). The GTP bound- active Sar1 is targeted to the ER, where it recruits Sec23/Sec24 proteins to the form the pre-budding complex with the cargo. The outer coat proteins Sec13/31, binds Sec23/Sec24 and triggers the GTPase-activating

protein (GAP) activity of Sec23 for Sar1. The GTP hydrolysis of Sar1 is at present proposed to trigger membrane fission by which COPII coated vesicles separate from the ER (Figure 1) (Venditti et al., 2014).

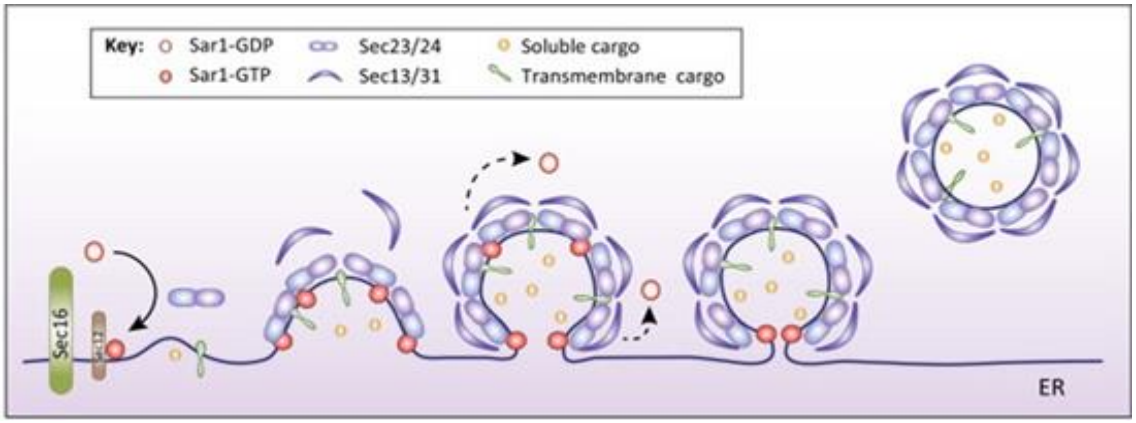

Figure 1. Biogenesis of a COP II carrier (Venditti et al 2014)
COPII vesicle generation, starts with the protein Sec12, which activates the GTPase Sar1.Inner coat proteins Sec23/Sec24 are recruited by Sar1 and form a complex with the cargo. Outer coat proteins Sec13/31, binds Sec23/Sec24 and promote the GTP hydrolysis of Sar1, resulting in vesicle fission from the ERES.

The COPII vesicles containing cargo are then transported to the Golgi complex via the ER-Golgi Intermediate compartment (ERGIC). The vesicles then undergo membrane fusion at the cis-Golgi. Depending on the cargo, the protein can undergo carbohydrate modifications or undergo proteolytic cleavage as they progress through the medial and trans compartments before finally being sorted to be transported onto extracellular space or other compartments.ER resident chaperones and enzymes that travel with the cargoes are transported back to the ER from the Golgi by coat protein complex I (COPI) coated vesicles, due to the presence of a KDEL sequence that binds the KDEL receptor (Barlowe and Miller, 2013). This process is essential to maintain the homeostasis in the compartments of the secretory pathway (Figure 2). The Golgi is a major hub of protein modification and a sorting station, I will not elaborate on its functions as it is not the focus of the thesis.

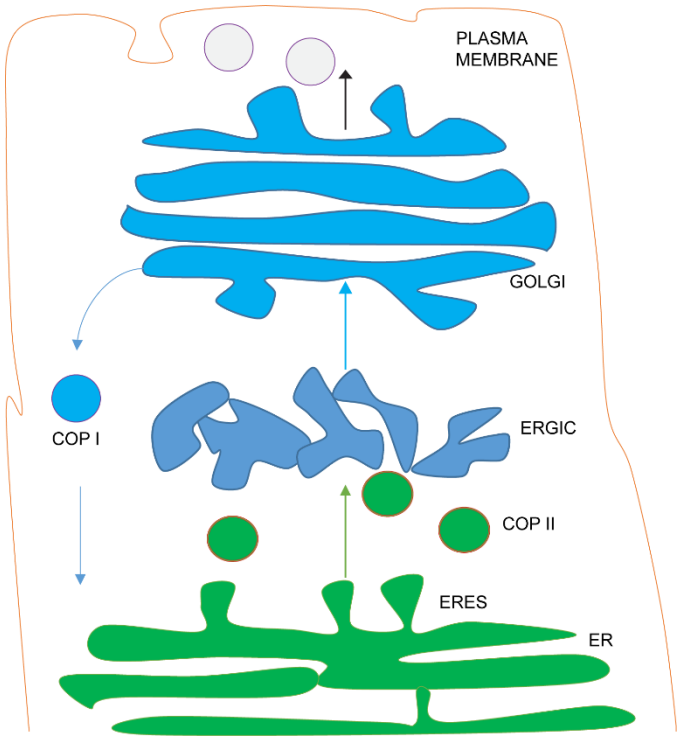

Figure 2. The classical secretory pathway
Signal sequence containing proteins are translocated into the ER, undergo folding, modification and are transported to the Golgi in COP II coated vesicles from ERES, passing through the ERGIC. ER resident proteins that travel to the Golgi are recycled back in COP I vesicles. Proteins in the Golgi undergo further modification and sorting and then onto the extracellular space or other compartments.

A combination of yeast genetics, in vitro reconstitution and imaging have greatly improved our understanding of vesicle mediated transport from the ER. But a longstanding challenge in the field has been the mechanism for the export of one of the most abundant secretory proteins: the collagens. Collagens are very large cargoes forming rigid structures which can be up to 425nm long (Sakai, 1986). After their synthesis, folding and modification in the ER they need to be exported to the next compartment. How do these large cargoes that cannot fit into standard

COPII vesicles exit the ER? The discovery of TANGO1, has enabled us to begin addressing some of these questions.

1.2 TANGO1

Transport and Golgi Organization 1 (TANGO1) was first identified in a Drosophila genome wide screen to identify genes affecting transport and Golgi organization (Bard et al., 2006). TANGO1 is a well conserved protein across metazoans and is required primarily for the export of collagens, chylomicrons and some components of the extracellular matrix (Saito et al., 2009; Santos et al., 2016; Wilson et al., 2011). TANGO1 is a transmembrane protein of 1907 amino acids, localized to the ER exit sites. The N terminus is in the ER lumen. The luminal portion of TANGO1 is composed of an SH3 like domain followed by an unstructured region, a coiled coil domain, and another large unstructured region. This is followed by a transmembrane domain and a partially membrane inserted region. Facing the cytoplasm are two coiled coil domains, culminating in a Proline rich Domain (PRD) at the C-terminus (Figure 3). The luminal SH3- like domain, functions as the cargo capturing module by binding collagens via its chaperone HSP47 (Ishikawa et al., 2016; Saito et al., 2009). The function of the luminal coiled coil domain is not known. The first coiled coil on the cytoplasmic side named is aptly titled the TEER domain – Tether of ERGIC at the ER for its function in recruiting ERGIC53 containing membranes (Santos et al., 2015). The second coiled coil domain CC2, binds another TANGO like protein called cTAGE5 (Saito et al., 2011). The carboxyl terminal Proline rich Domain (PRD) binds Sec23 and Sec16 (Ma and Goldberg, 2016; Maeda et al., 2017; Saito et al., 2009).

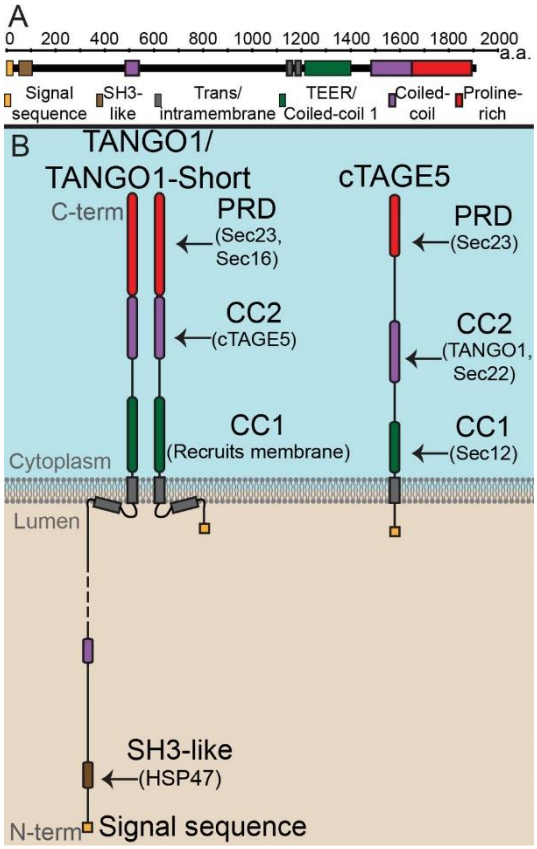

Figure 3. The TANGO1 family of proteins: domains and topology at the ERES (Raote et al 2018)

(A) TANGO1, full length depicting its various domains

(B) TANGO1, its isoform TANGO1-short and cTAGE5 are transmembrane proteins at the ERES. In the lumen of the ER, TANGO1, through its N-terminal SH3-like domains binds HSP47, which facilitates collagen binding. This is followed by a transmembrane domain and a membrane inserted region. There are two coiled coil domains on the cytoplasmic side. CC1 recruits ERGIC and CC2 binds cTAGE5.The C-terminal, Proline rich domain (PRD) binds Sec23A and Sec16. TANGO-1 short and cTAGE5, both lack the luminal domains of TANGO1, but have identical cytoplasmic domain architecture and similar binding partners. CC1 of cTAGE5 recruits Sec12.

An interesting feature of TANGO1 at the ERES is its organization into rings which serves to bring in more COPII coats to the site of export (Raote et al., 2017). A recent paper from the Malhotra lab has shed further light on how these rings are assembled through TANGO1's interactions with itself, COP II and with TANGO1-short, cTAGE5 in a cargo dependent manner. The recruitment of ERGIC by the TEER domain to expand the carrier for collagens is facilitated by the NRZ (NBAS/RINT1/ZW10) tether complex which is also essential for ring formation (Raote et al., 2018).

1.2.1 TANGO1- short

An isoform created due to TANGO1 splicing, it has a domain composition almost identical to TANGO1, with the exception that it lacks the entire luminal module and a slightly altered amino acid composition at the transmembrane region (Figure 3). Knockdown of TANGO1s and TANGO1 leads to decrease in cTAGE5 localization and Sec12 recruitment to ERES. Despite the lack of cargo binding domain, TANGO1s has been shown to be required for Collagen VII secretion and is a part of a protein complex of 700kDa consisting of Ctage5 and Sec12 (Maeda et al., 2016). Since it is identical to the cytoplasmic moiety of TANGO1, its domain architecture and binding partners are thought to be conserved.

1.2.2 cTAGE5

Cutaneous T-cell lymphoma–associated antigen 5 (cTAGE5), initially identified in an immunoscreening of meningioma cDNA library (Heckel et al., 1997) is also called meningioma-expressed antigen 6 (MEA6/MGEA6). It is overexpressed in several cancers (Comtesse et al.). A transmembrane protein of 804 amino acids localized to the ER exit sites, cTAGE5 has a signal sequence, followed by a single transmembrane region, and like TANGO1 it has two cytoplasmic coiled coil regions and a C-terminal Proline Rich Domain. cTAGE5 and TANGO1 form a very stable complex via binding of their respective Coiled coil 2 domains (Figure

3). Akin to TANGO1, the Proline rich domain of cTAGE5 also seems to interact with Sec23A and Sec24C (Saito et al., 2011).

Although knockdown of c TAGE5 does not affect levels of TANGO1, it did lead to an accumulation of collagen VII in the ER (Saito et al., 2011). In addition to this role, cTAGE5 is required for recruitment of Sec12 to ER exit sites, which functions as a guanine-nucleotide exchange factor for Sar1 GTPase. This cTAGE5 dependent recruitment of Sec12 is also crucial for efficient collagen VII export. A chimeric, protein composed of cTAGE5 and MIA2 known as TALI (TANGO1 like) is essential for chylomicron export (Pitman et al., 2011; Santos et al., 2016). cTAGE5 enables export of VLDL through regulation of COPII assembly (Santos et al., 2016; Wang et al., 2016). Recent findings also demonstrate a role for cTAGE5, in trafficking of proinsulin from the ER (Fan et al., 2017).

TANGO1, TANGO1- short, cTAGE5, COP II proteins and the NRZ complex assembled in a ring like structure could be the minimal export machinery required to export Collagens from the ER (Figure 4)

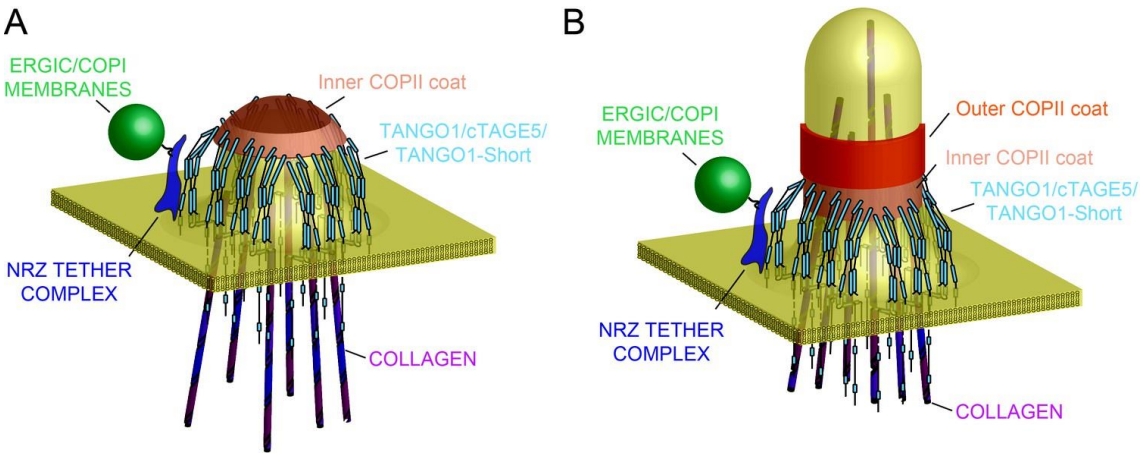

Figure 4. Model of TANGO1 ring assembly
(Raote et al 2018)

(A) TANGO1 family of proteins form rings at the ERES by virtue of their interactions with each other, the COP II coats and the NRRZ complex.

(B) A large carrier is generated by addition of the membranes which creates a vesicle large enough to fit collagens.

Collagen export involves the intricate co-ordination of molecular players, on the luminal and cytoplasmic sides of the ER and knowing who they are and what they do will give us the tools to address the questions that have been plaguing the field.

1.3 Extracellular Matrix (ECM)

ECM proteins have been key to evolution of multicellular organisms. Our understanding of ECM has come a long way since the early days when it was only thought of as the styrofoam packing material providing structural support. Its functions include cell adhesion, signaling, transduction of mechanical forces, and creation of growth factor and morphogen gradients all processes that are key during development and tissue morphogenesis. The core ECM proteins, termed the matrisome consist of 300 proteins, mainly fall into three categories: Proteoglycans, glycoproteins and collagens (Hynes and Naba, 2012; Mouw et al., 2014). A well-known member of the proteoglycan family is Perlecan, also known as Heparan sulfate proteoglycan 2. Glycoproteins are represented by the proteins fibronectin and laminin, which are indispensable for survival from the early stages of embryogenesis (Klinkenberg et al., 2014). Since the focus of the thesis is collagen export, it is essential to delve into the structure, folding and functional aspects of collagen in detail.

1.3.1 Collagens

Collagens are the most abundant secretory proteins comprising about 25% of our dry body weight (Myllyharju, 2004). The collagen family of proteins include 28 members and are characterized by the presence of a common structural motif, the triple helix. They can be homotrimers or heterotrimers, composed of 3 α chain polypeptides, which form triple helices by the presence of Gly-X-Y repeats, where X and Y are 4-hydroxyproline, respectively (Gelse, 2003). The triple helix structures give the molecules rod like rigidity that ranges from around 96% of the entire protein in Collagen I to under 10% in collagen XII (Ricard-Blum, 2011).

Rotary shadow electron microscopy studies have shown that they are rod like structures of up to 425 nm as in case of collagen VII (Sakai, 1986) . Secreted collagens not only supply structural support to tissues, they are also crucial for cell adhesion, migration, chondrocytes development into mineralized bones, wound healing, differentiation, and morphogenesis (Kadler et al., 2007).

Once they are secreted into the extracellular space, collagens undergo further modifications and form higher order supramolecular assemblies. Based on their organization in the extracellular space, collagens can be classified broadly into the following groups; fibril forming, FACITs (Fibril-Associated Collagens), hexagonal networks, beaded filaments, anchoring fibrils and transmembrane forms (Figure 5). (Table 1) (Kadler et al., 2007; Ricard-Blum, 2011).

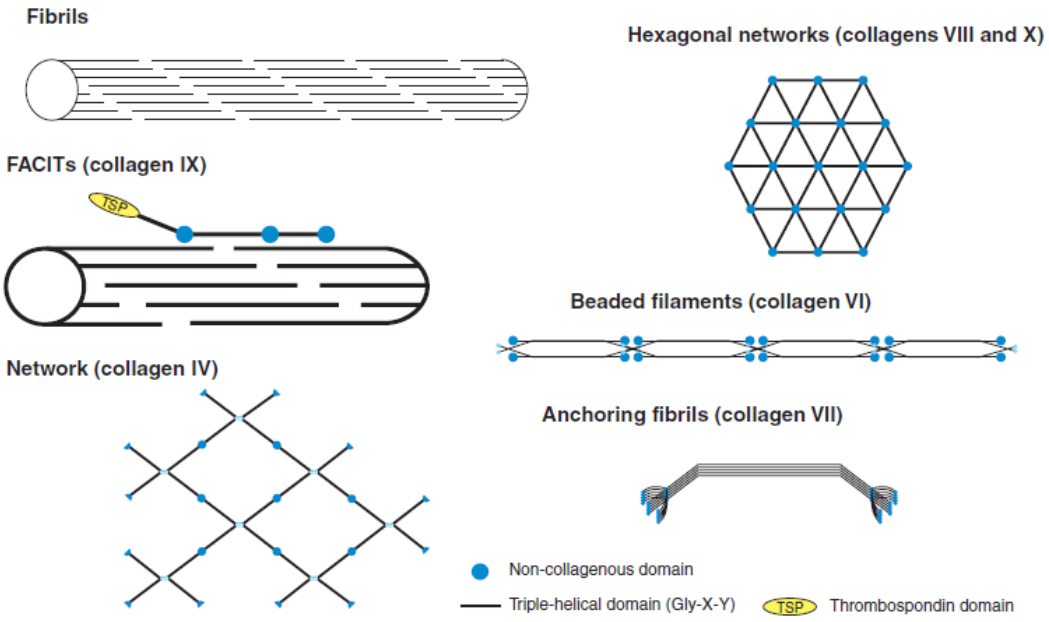

Figure 5: Supramolecular assemblies of Collagens (Ricard-Blum 2011)

Table 1. Classification of collagens based on their supramolecular assembly

Supramolecular assembly type	Collagens
Fibril-forming	I, II, III, V, XI, XXIV and XXVII
Fibril-associated Collagens With Interrupted Triple Helices (FACITs)	IX, XII,XIV, XVI, XIX, XX, XXI, XXII and XXVI
Hexagonal Networks	VIII and X
Type IV Collagens	IV
Beaded Filaments	VI
Anchoring Fibrils	VII
Transmembrane	XIII, XVII, XXIII and XXV

1.3.2 Collagen Folding and Modification Machinery

Despite being large molecules, the time taken from translation, through the steps of folding, modification to secretion ranges between 10 and 30 minutes (Bruckner et al., 1981; Morris et al.). This requires the folding and modification machinery in the ER to be highly efficient. A large number of chaperones and enzymes are dedicated specifically to collagens and required for the proper folding, hydroxylation, glycosylation and ultimately triple helix formation of collagens (Ishikawa and Bächinger, 2013). Before triple helix formation, the unfolded collagen chains have to undergo several modifications including hydroxylation of proline and lysine residues in the Gly-X-Y repeats and glycosylation. Three groups of enzymes, Prolyl 4-hydroxylases (P4Hs), Prolyl 3-hydroxylases (P3Hs) and Lysyl hydroxylases (LHs) catalyze the hydroxylation of unfolded collagen chains. The most abundant modification is the conversion of Proline in the Y position of the Gly-X-Y repeat to 4-hydroxyproline and is catalyzed by the Prolyl

4-hydroxylases. This confers thermal stability to the protein (Gorres and Raines, 2010). The proline in the X position is modified to 3-hydroxyproline by Prolyl 3-hydroxylases, this is a less frequent modification. Before lysine residues can be O-glycosylated, they need to be modified to hydroxylysines and this is carried out by Lysyl hydroxylases. This modification is important for intra and inter chain crosslinking (Ruotsalainen et al., 2006). All of hydroxylases have three isoforms with varied levels of expression, binding partners and distribution depending on the tissue. Hydroxylysines residues undergo further modification by addition of β-galactose and α-glucose. This is carried out by the enzymes, GLT25D1 (glycosyltransferase 25 domain containing 1), GLT25D2 (glycosyltransferase 25 domain containing 2) and Lysyl Hydroxylase 3(LH3) (Myllyharju, 2005).

ER chaperones BiP, calnexin, calreticulin, GRP94, PDI are required for the proper folding of the C-propeptides, which then initiates triple helix formation in the C to N direction in a zipper like manner. It is important at this stage for the large number of proline peptide bonds to be in the trans conformation. This is reaction is carried out by a family of proteins called Peptidyl prolyl cis–trans isomerases (PPIases), especially Cyclophilin B (CypB) which forms a complex with P3H1/CRTAP and FKBP65 (Ishikawa and Bächinger, 2013).

HSP47 is a collagen specific chaperone,fibroblasts from Hsp47 knockout mice, are unable to fold the collagen triple helix (Nagai et al., 2000).They bind preferentially to fully folded triple helical regions (Koide et al., 2006) and not to unfolded chains. HSP47 enables the selection and export of fully folded triple helical collagen by its association with the SH3-like domain of TANGO1 (Ishikawa et al., 2016). HSP47 remains bound to procollagen till it is exported into the Golgi (Nagata, 2003), where it dissociates due to change in pH.of Golgi. Once the collagens are transported through the Golgi, and secreted into the extracellular space, the pro-peptides are cleaved off by proteinases which initiates their assembly into the fibrils and networks (Figure 6).

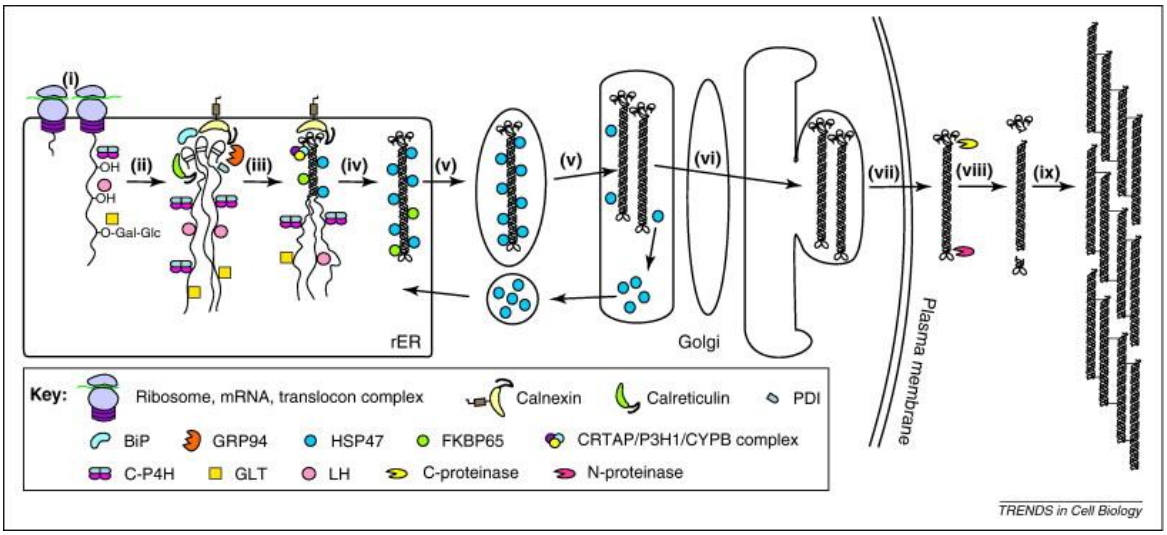

Figure 6. Collagen folding and biosynthesis
(Makareeva et al. 2011)

(i) collagen chains are translocated cotranslationally into the ER lumen and undergo hydroxylation and glycosylation

(ii) (ii) The chain association and folding of the C-propeptide domains of two proa1(I) and one pro-a2(I) chains is chaperoned by BiP, calnexin, calreticulin, GRP94 and PDI.

(iii) (iii) The triple helix formation proceeds in the C to N direction, assisted by CRTAP/P3H1/CYPB complex.

(iv) (iv) The triple helical structure is stabilized by binding of HSP47 and FKBP65

(v) (v)After folding is complete, the HSP47 bound collagen is transported to the golgi, where it dissociates and is recycled back to the ER.

(vi) (vi) Collagen is transported through the Golgi

(vii) (vii) Collagen is secreted into the extracellular space

(viii) .(viii)The C terminal and N terminal propeptides are cleaved (ix) Collagens assemble into fibers

1.3.3 Quality control of Collagen

In cells that produce large amounts of collagens such as fibroblasts, chondrocytes, osteoblasts around 15% of the newly synthesized collagen is quickly degraded (Bienkowski and Gotkin, 1995). In addition to this, collagen mutations, mutations in the chaperones will lead to an accumulation of unfolded or misfolded collagen in the Endoplasmic reticulum. This results in an altered ER state, stressing the cells, which if left unchecked, could lead to apoptosis. The unfolded protein response (UPR) is elicited by protein accumulation in the ER due exposure of hydrophobic residues, which sequester BiP and GRP94. This results in loss of BiP from UPR sensors IRE1, PERK and ATF6 and triggers a signaling cascade which decreases protein synthesis, increases chaperone levels and enables degradation of misfolded proteins via ER associated degradation pathways. All of the above steps are measures taken by the ER folding machinery to alleviate ER stress and maintain homeostasis (Walter and Ron, 2011).

How are collagens degraded intracellularly? Mutations in the C-propeptide region prevent chain assembly and triple helical formation. Such collagens bind BiP and are known to be degraded by ERAD pathway (Fitzgerald et al 1998.), whereas mutations in the triple helical regions do not have enough hydrophobic residues and therefore do not initiate UPR. The misfolded proteins go onto to form aggregates, which are ultimately cleared by autophagy and lysosomal degradation (Figure 7) (Ishida et al., 2009)

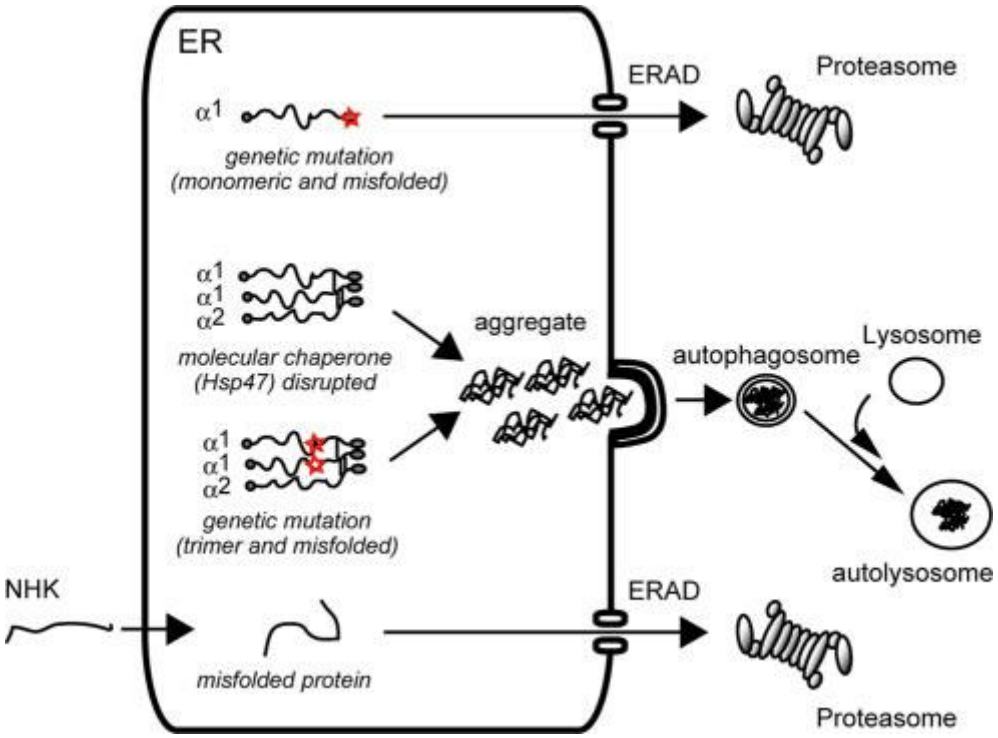

Figure 7. A model for collagen degradation in the ER.
(Ishida et al 2009)
Monomers of collagen are degraded by ERAD. Mutations in the triple helix or mutations in chaperones like HSP47 leads to aggregation which is proposed to be degraded by autophagy-lysosomal degradation.

.

Mutations in collagens or any of their folding machinery leads to the secretion of unfolded or misfolded collagens, which lead to collagenopathies where they disrupt the formation of the Extracellular Martix (ECM) undermining the stability of the structure leading to fragile skin and brittle bones. Due to the varied distribution of collagens across several tissues, collagenopathies affect several organs (Figure 8) Another facet of this problem is the secretion of excess collagen resulting in fibrosis (Ricard-Blum et al., 2018) It is therefore imperative to understand how the quality and quantity of collagens secreted is regulated. There are 28 types of Collagens but from the standpoint of my thesis, we will focus on two collagens: I and XVII.

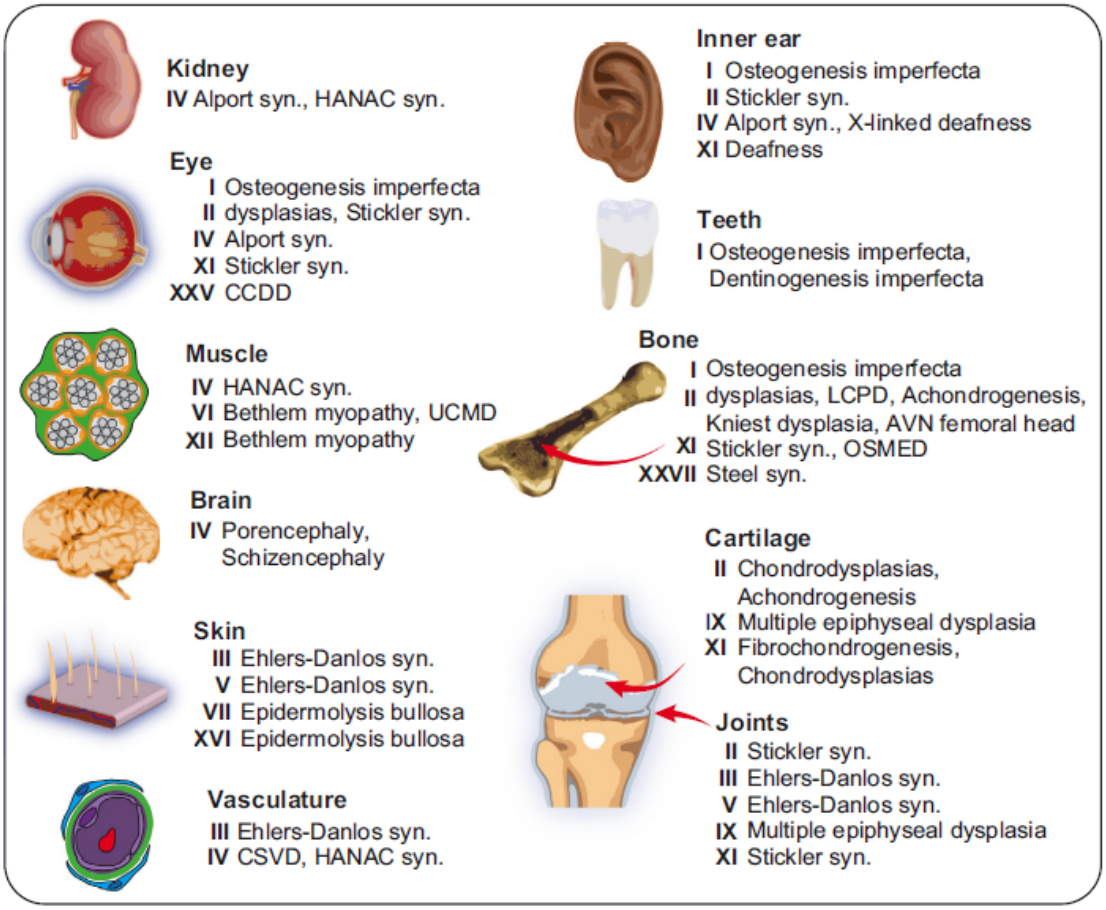

Figure 8. Pathologies resulting from mutations in the collagen triple helix
(Fidler et al 2018)
Mutations in the triple helix of various types of collagens results in a number of diseases and disorders affecting a wide variety of organs and tissues. The numbers in roamn numerals indicate collagen type. Abbreviations: AVN, avascular necrosis of femoral head; CCDD, congenital cranial dysinnervation disorder; CSVD, cerebral small-vessel disease; HANAC, hereditary angiopathy with nephropathy, aneurysms and muscle cramps; UCMD, Ullrich congenital muscular dystrophy; LCPD, Legg–Calvé–Perthes disease; OSMED, otospondylomegaepiphyseal dysplasia; syn, syndrome.

1.3.4 Collagen I

Collagen I is one of the most important and abundant constituents of the skin and bone extracellular matrix (Myllyharju, 2004). It is synthesized and secreted in copious amounts by dermal fibroblasts and osteoblasts. It is a heterotrimeric protein composed of two α 1 chains each of 1464 amino acids and one α 2 chain, a polypeptide of 1366 amino acids which associate at the carboxyl terminal propeptides to form a triple helix as it folds towards the amino terminus. The central large collagenous stretch is flanked by amino terminal and carboxyl terminal non-collagenous domains (Gelse, 2003) (Figure 9). The major modifications are hydroxylation and glycosylation which are essential for proper folding and triple helix formation. Upon secretion, the procollagen undergoes proteolytic cleavage at the amino and carboxyl terminal ends by two proteases a disintegrin and metalloproteinase with thrombospondin motifs 2 (ADAMTS2) and bone morphogenetic protein 1 (BMP1), respectively to give a 300nm long mature form (Prockop and Kivirikko, 1995).

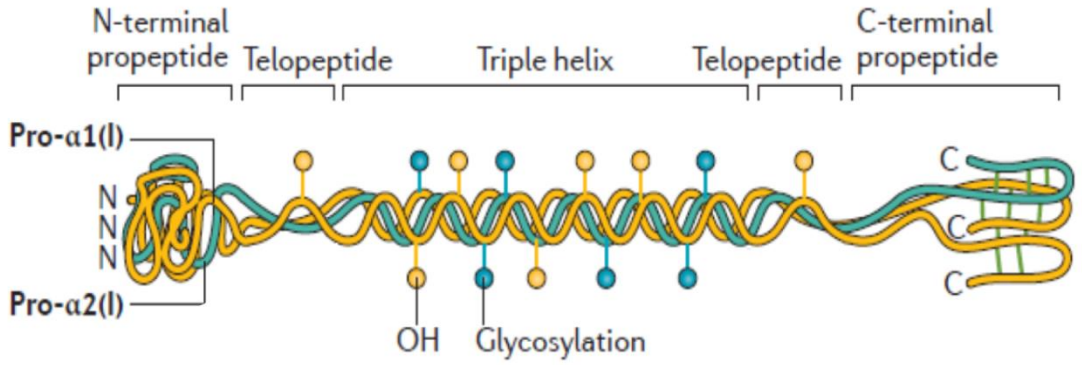

Figure 9. Structure of Collagen I
(Marini et al., 2017)
Two α1 and one α2 polypeptides associate at the C-terminal propeptide to fold into a triple helix from the C to N direction terminating in the folding of the N-propeptide. Hydroxylation and glycosylation of the polypeptides occur before triple helix formation.

Mutations in COL1A1 gene causes osteogenesis imperfecta (OI), a bone fragility disorder (Makareeva et al., 2008). Based on the clinical presentation it is

classified into 4 types. In addition to low bone mass and the propensity for fractures and deformities, other presentations include hearing loss, blue sclera, short stature, pulmonary issues, cardiac valve deformities and scoliosis (Marini et al., 2017). COL1A2 mutations results in a connective tissue disorder known as Ehlers–Danlos syndrome. It is characterized by hyper mobile joints, making them prone to dislocation, and elastic skin (Wirtz et al., 1987).

1.3.5 Collagen XVII

Collagen XVII belongs to the family of transmembrane collagens and at 1497 amino acids is their largest member. Primarily expressed in skin, mucous membranes, brain and the placenta(Fairlky et al., 1995; Seppänen et al., 2006).The most crucial function of Collagen XVII is in the skin where it is an important component of hemidesmosome complex which helps epidermal keratinocytes adhere to the basement membrane, by its interaction with Laminin 5 and α6β4 integrin (Borradori and Sonnenberg, 1999). Col17A1 gene mutations, cause Junctional Epidermolysis Bullosa, a skin fragility disorder(McGrath et al., 1995). It is also involved in skin blistering disorders wherein autoantibodies against Collagen XVII contribute to bullous pemphigoid,herpes gestationis, and cicatricial pemphigoid (Giudice et al., 1993). Structurally, Collagen XVII is a type II transmembrane glycoprotein, a homotrimer, composed of 3 α chain polypeptides of 180kDa each. At the N-terminus is a 466 amino acids long cytoplasmic domain, followed by 23 amino acid transmembrane region. The C-terminus is 1008 amino acids composed of 15 collagenous domains interrupted by non-collagenous regions and forms the extracellular domain (Figure 10) (Li et al., 1993). There is a coiled coil region adjacent to the transmembrane region called the NC16A region, this area is crucial for the start of the trimerization and triple helix formation which happens in the N to C direction which is unlike other soluble collagens where it occurs in the C to N direction(Areida et al., 2001). The phenomenon of constitutive shedding of transmembrane collagens from the cell surface was first observed in Collagen XVII where the 180kDa protein is cleaved at the ectodomain to yield a 120kDa protein. (Hirako et al., 1998; Schäcke et al., 1998). This cleavage is mediated by ADAM9 and ADAM10 (Franzke et al., 2009).

Figure 10. Structure of Collagen XVII
Collagen XVII is a transmembrane protein, with an N-terminal cytoplasmic domain followed by a transmembrane stretch. The NC16A domain, serves as a nucleator for triple helix formation. The large extracellular domain is composed of 15 collagenous domains.

Objectives

2. Objectives

TANGO1, initially discovered in a screen to identify proteins affecting transport and Golgi organization, was later shown to be required for the export of collagens, chylomicrons and other large cargoes. It is a large transmembrane protein at the ER exit sites, with sizeable luminal and cytoplasmic halves. Each of these halves are composed of subdomains with unique functions and interactors.

Over the course of the last few years we have uncovered new interactors and thereby been able to gain better mechanistic insights into the function of TANGO1. One of the main objectives of this thesis was to identify interactors of TANGO1, with a primary focus on its partners in the ER lumen as very little was known at the time of this study.

The other objective was to test if TANGO1 was required for the export of a transmembrane collagen. All collagens that were shown to need TANGO1 for their export were soluble cargoes without the ability to connect to cytoplasmic coat proteins to facilitate export. Transmembrane collagens like Collagen XVII have a cytoplasmic domain which in theory could bind COP II, do such proteins require TANGO1 for their trafficking?

Materials and methods

3. Materials and Methods

a. Cell culture

All cell lines were grown in Dulbecco's Modified Eagle's Medium (DMEM),4.5g/L glucose and ultraglutamine. Cells were maintained at 37C with 5%CO_2. Media was supplemented with 250um of ascorbic acid while performing collagen secretion assays.

Cell lines

The following cell lines were used in this study

- HeLa
- HEK293T
- HeLa ΔTANGO1, where TANGO1was deleted using the CRISPR/Cas9 system (Santos et al.,2015).
- U2OS
- U2OS ΔTANGO1+TANGO1s, where both TANGO1 and TANGO1-short were deleted using the CRISPR/Cas9 system. This cell line was generated by Julia Von Blume's lab at Max Planck Institute of Biochemistry, Martinsried.
- HeLa-Collagen XVII-GFP (see section on Collagen XVII)

All cell lines were regularly tested for mycoplasma.

b.siRNA oligos

siRNAs for TANGO1, Torsin-1A, Sec23IP, Non –Targeted Control were purchased from Eurofins MWG Operon:
TANGO1 siRNA- GAUAAG GUCUUCCGUGCUU
Torsin-1A siRNA- GGG CAG AAG CGG AGC CUU A
Sec23IP siRNA- GGA CAG GAA UAU UAA GAA A
For the control siRNA a pool of 4 siRNAs were used
NT1- UGG UUU ACA UGU CGA CUA ATT
NT2- UGG UUU ACA UGU UGU GUG ATT

NT3- UGG UUU ACA UGU UUU CUG ATT

NT4- UGG UUU ACA UGU UUU CCU ATT

c. siRNA mediated silencing

Cells for transfection were seeded at 4 x 10^5 cells per well of a 6 well plate in 2ml of culture medium. The cells were incubated at 37°C and 5% CO2 until the point of transfection. In 100ul of serum free media, 1ul of 100um siRNA was diluted to which 12 ul of Hiperfect transfection reagent was added.

The transfection mix was incubated at room temperature for 15 minutes for complex formation. The complexes were added drop-wise onto the cells and the plates were swirled for even distribution. The cells were incubated at the growth conditions for 24 hours after which the cells were trypsinized and the above siRNA transfection step repeated. The media was changed 24 hours post the second round of transfection and incubated for 24 hours.

Cells were lysed and tested for the silencing using specific primers or antibodies, at mRNA or protein level respectively.

d. RNA Extraction

RNA extraction was done using Qiagen RNA easy kit.the cells were washed with 1X PBS 3 times and. Cells were lysed using 350ul RLT buffer, and disrupted by pipetting after transferring to a microcentrifuge tube. 350ul of 70% ethanol was added to the lysate and mixed well. This was then transferred to the RNAesay spin column and spun down for 15 seconds at 8000g. Flow through was discarded. The column aws then washed with 700ul of Buffer RW1 and centrifuged at 8000 x g for 15 seconds and flow-through was discarded.

The column was then washed with 500ul Buffer RPE centrifuged at 8000 x g for 15 seconds, flow-through was discarded. The above step was repeated. The column was spun at full speed for 1 minute to remove any remaining buffer.

The column was placed in a fresh 1.5ml tube, 50ul of RNase free water was added to it and centrifuged for 1minute at 8000 x g to elute the RNA.

e. cDNA synthesis

cDNA was synthesized using SuperScript™ III Reverse Transcriptase Cat. No. 18080- 044. 1 ug of RNA was used for each reaction, to which 1ul of dNTP mix was added and the volume was made up to 13 ul with sterile water. The mixture was heated for 65 degree Celsius for 5 minutes and incubated on ice for 1 minute.

The cDNA synthesis mix was prepared for each sample, with the following composition:

4 µl 5X First-Strand Buffer

1 µl 0.1 M DTT

1 µl RNaseOUT (40 units/µl)

1 µl of SuperScript™ III RT (200 units/µl)

The cDNA synthesis mix was added to the RNA, dNTP mix from the first step and was incubated at 50°C for 60 minutes. The reaction was inactivated by heating at 70°C for 15 minutes.

f. PCR-polymerase chain reaction

All cloning experiments were performed using Thermo Scientific Phusion High-Fidelity DNA Polymerase (catalog no. F530S). All reactions and cycling parameters for the PCR were followed according to recommendations of the manufacturer.

g. Immunoprecipitation of FLAG tagged proteins

24-48 hours post transfection with the FLAG tagged construct of interest, the cells were lysed in FLAG lysis buffer (50 mM Tris HCl, pH 7.4, with 150 mM NaCl, 1

mM EDTA, and 1% TRITON X-100) with protease inhibhitors. Lysis was allowed to procced on ice for atleast 30 minutes. The supernatants were clarified by centrifuging the samples at 16,000 ×g for 15 minutes at 4 degrees.

For each lysate sample 40ul of resuspended ANTI-FLAG M2 affinity gel (Sigma Catalog Number A2220) was used. The beads were spun down at 5,000 x g for 30 seconds. They were washed with Tris buffered Saline (TBS) and equilibrated with lysis buffer. The beads were then collected and the lysate were added onto them.

The beads were allowed to incubate with the lysate at 4 degree celcius, with rotation for about 3 hours. The beads were centrifuged once at 5000xg to remove the supernatant.

3 washes, each of 10 minutes was performed using FLAG wash buffer (50 mM Tris HCl, pH 7.4, with 150 mM NaCl and 0.5% TRITON X-100). This was followed by a final wash with TBS. Elution was done by heating the beads with twice the volume of 2x SDS loading dye without the reducing agent.

h. APEX

Generation of TANGO1- APEX2 constructs
The pcDNA3.1 FLAG-APEX2 vector was used as the template for cloning. The APEX2-FLAG fusion tag was added after the PRD domain on the cytoplasmic end, whilst on the luminal side, it was added between the signal sequence and SH3-like domain of TANGO1.

Proximity labeling

(for more in depth protocol please refer to Hung et al., 2016)

HeLa and HEK293T cells were plated in a 10cm dish in complete medium for the following conditions: Control, TANGO1-Cyto-APEX2 and TANGO1-Lum-APEX2. 18 hours post seeding, the cells were transfected using Lipofectamine 3000 as per the manufactures protocol with the constructs mentioned above. 24 hours

post transfection the media was replaced with fresh complete media, with 500um of Biotin Phenol and incubated for 30 minutes at 37°C.

H_2O_2 was added to a final concentration of 1mM to the biotin phenol containing media and mixed well allowing the reaction to continue for 1 minute at room temperature.

After 1 minute the labeling solution was removed and the cells were rinsed with quenching solution (10 mM sodium ascorbate, 5 mM Trolox and 10 mM sodium azide in PBS) three times. The cells were washed one final time in the quencher solution collected and pelleted and frozen at -80°C.

Enrichment of biotinylated pool using streptavidin magnetic beads

The frozen pellets were lysed in 1ml of RIPA lysis buffer supplemented with protease inhibitors. The lysis was allowed to proceed at 4°C with rotation for 30 minutes. The lysates were then centrifuged at 16000xg for 20 minutes at 4°C.

The amount of protein in the lysates were quantified by BCA. For each sample 100ul of streptavidin magnetic bead was used, the beads were washed 3 times with RIPA buffer and were incubated with 1mg/ml of the lysate with rotation overnight at 4°C.

The supernatant post incubation was collected by pelleting the beads. Every experimental sample was washed in the following order: 2 times with RIPA, one wash each with 1M KCl, 0.1M Na2CO3, 2 M urea in10 mM Tris-HCl and final 2 washes with RIPA buffer again.

Mass Spectometry

Sample Digestion

Beads used in immunoprecipitation were resuspended in 6M urea, and attached proteins were reduced (1mM DTT, 1h, 37ºC), alkylated (2mM IAA, 30 min, RT), and directly digested with trypsin (1µg, 37ºC, 16h). Digested peptides were

desalted with C18 columns, evaporated to dryness, and resuspended into 10 µL of 0.1% formic acid in water.

Data acquisition

Digested samples were analyzed using a LTQ-Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an EasyLC (Thermo Fisher Scientific (Proxeon), Odense, Denmark). Peptides were separated by reverse-phase chromatography using a 25-cm column with an inner diameter of 75 µm, packed with 5 µm C18 particles (Nikkyo Technos Co., Ltd. Japan). Chromatographic gradients started at 93% buffer A and 7% buffer B and gradually increased to 65% buffer A / 35% buffer B in 60 min at a flow rate of 250 nl/min. The mass spectrometer was operated in positive ionization mode with nanospray voltage set at 2.2 kV and source temperature at 250°C. Ultramark 1621 for the FT mass analyzer was used for external calibration prior the analyses, and an internal calibration was performed using the background polysiloxane ion signal at m/z 445.1200. The acquisition was performed in data-dependent acquisition mode and full MS scans with 1 micro scans at resolution of 60,000 were used over a mass range of m/z 350-2000 with detection in the Orbitrap. Auto gain control (AGC) was set to 1E6, dynamic exclusion (60 seconds) and charge state filtering disqualifying singly charged peptides was activated. In each cycle of data-dependent acquisition analysis, following each survey scan, the top ten most intense ions with multiple charged ions above a threshold ion count of 5000 were selected for fragmentation at normalized collision energy of 35%. Fragment ion spectra produced via collision-induced dissociation (CID) were acquired in the Ion Trap, AGC was set to 5e4, isolation window of 2.0 m/z, activation time of 0.1ms and maximum injection time of 100 ms was used. All data were acquired with Xcalibur software v2.2.

Data analysis

Acquired spectra for the different LDL samples and tryptic peptides derived from hcVSMC cells were analyzed using the Proteome Discoverer software suite

(v2.0, Thermo Fisher Scientific) and the Mascot search engine (v2.5, Matrix Science). The data were searched against the Swiss-Prot human database (v.2017/10). At the MS1 level, a precursor ion mass tolerance of 7 ppm was used, and up to three missed cleavages were allowed. The fragment ion mass tolerance was set to 0.5 Da for MS2 spectra. Oxidation of methionine, and N-terminal protein acetylation were defined as variable modifications whereas carbamidomethylation on cysteines was set as a fixed modification. False discovery rate (FDR) in peptide identification was limited to a maximum of 5% by using a decoy database.

i.Collagen I Secretion Assay and western blotting

The cells were washed 3 times with IX PBS, and once with Optimem. Secretion assay was carried out in Serum free Optimem media supplemented with 0.5um cycloheximide to stop protein synthesis in addition to 250µm of Ascorbic acid. Media and cell lysate were collected at time points of zero hour, one hour and two hours.

Equal amounts of proteins were loaded onto an 8% gel for Collagen, TANGO1 and higher molecular weight proteins and 12% for tubulin, Torsin and loer molecular weight proteins with molecular weight markers. The gel was run in a Bio rad electrophoresis unit with 1X running Buffer, till the markers were well separated.

The gel was then transferred to a nitrocellulose membrane, using a transfer unit at 75volts for 3 hours for Collagens and TANGO1 and around 1 hour at 100 volts for regular proteins.The membrane was blocked in 5% BSA in TBST for one hour.

The appropriate antibodies were then incubated overnight. The blots were washed 3–5 times for 5 min with TBST. This was followed by incubation in the HRP-conjugated secondary antibody solution for 1 hr at room temperature. The blots were washed again 3–5 times for 5 min with TBST. The blots were developed using ECL (Luminata crescendo) and imaged using an AI600 imager.

j. Cell surface Biotinylation of Collagen XVII

Plasmids: All constructs used in this study were a kind gifts from Dr. Claus-Werner Franzke at the lab of Leena Bruckner-Tuderman ,Department of Dermatology, University Medical Center Freiburg.

U2OS and ΔTANGO1 U2OS cells were transfected with 1ug of FLAG-Collagen XVII construct using the Screenfect reagent, cells were maintained in DMEM + 10% FBS supplemented with 250um Ascorbic acid. 24 hours post transfection, cells were washed gently three times with ice cold PBS containing 0.1mM CaCl2 and 0.1mM MgCl2, pH7.4. For biotinylating cell surface proteins, Sulfo-NHS-LC-biotin (Thermo scientific) was used. An initial stock solution of 10mM was prepared in miliq water, 200ul of this solution was added to 1ml of PBS containing 0.1mM CaCl2 and 0.1mM MgCl2, to give a final concentration of 2mM. The cells were incubated with biotinylating reagent for 30 minutes at 4 degrees Celsius with gentle shaking. The reaction was quenched was aspirating the biotinylating reagent and washing the cells three times with 100mM glycine in PBS. The cells were washed one last time in PBS and lysed in 250ul of NP40- lysis buffer on ice for 30 minutes. The lysates were spun at 1600xg for 20 minutes and supernatant transferred to a fresh tube. 10 percent of the supernatant was saved and frozen at -20 to be later used as the input.

Immunoprecipitation of Biotinylated protein using Neutravidin agarose

40ul of neutravidin agarose per sample was washed with PBS followed by NP40 lysis buffer for equilibration. The supernatants, both the biotinylated and unbiotinylated controls were added to the beads and incubated overnight at 4 degrees Celsius with constant rotation.

The beads were spun down at 3000g for 5 minutes and the unbound samples were removed. The beads were then washed 3 times with the NP40 lysis buffer and a final wash with Tris Buffered Saline pH7.4.

The biotinylated proteins were eluted by heating the neutravidin agarose beads with 2X sample buffer at 70 degrees Celsius for 10 minutes.

k. Generation of a stable cell line expressing collagen XVII-GFP and dCas9-KRAB.

HeLa cells were transfected with Collagen XVII-GFP and after 24h, cells were selected for geneticin resistance (200 µg/ml) over the course of one week. Then, cells were analyzed by FACS and selected for high GFP expression and plated individual cells in a 96-well plate.

After expansion, cells were analyzed by immunofluorescence and flow cytometry to select clones expressing collagen XVII-GFP. Finally, to induce CRISPRi-mediated knockdown, selected clone were infected with the lentivirus vector pHR-SFFV-dCas9-BFP-KRAB (Gilbert et al., 2013) (addgene #46911).

l. CRISPRi/dCas9-mediated knockdown.

Cloning sgRNA into lentiviral vector.

Gene-specific CRISPR sgRNA oligonucleotide sequences are listed in table 6 Oligo were diluted to 100 µM in water and 1 µl of forward and reverse oligo were combined in a 50 µl reaction containing 200 mM potassium acetate, 60 mM HEPES, pH 7.4 and 4 mM magnesium acetate. Oligos were incubated for 5 min at 95°C and slow cooled (0.1 degrees/sec) for annealing. Annealed oligos were diluted 1/40 and 1 µl of insert were ligated into 10 ng of digested vector (pU6-sgRNA, addgene #60955 digested with BstXI and Blpl) using T4 DNA ligase in a final volume of 10 µl. Ligation was allowed to proceed for 1h. Ligated products were transformed into DH5α bacteria.

Pairs of oligonucleotides cloned in pU6-sgRNA. Three pairs of oligonucleotides were designed in order to generate three different sgRNA for each targeted genes. Sense sgRNA oligonucleotides were designed with 5’ TTG and 3’

GTTTAAGAGC overhangs (red) and antisens sgRNA oligonucleotides were designed with 5’ TTAGCTCTTAAAC and 3’ CAACAAG overhangs (blue)

Targeted genes		Oligonucleotide sequences (5'-3')
Gal4	oligo 1	TTGGAACGACTAGTTAGGCGTGTGTTTAAGAGC
	oligo 2	TTAGCTCTTAAACACACGCCTAACTAGTCGTTCCAACAAG
NSF	oligo 1	TTGGCCGGACGTGTCCGCGAAGAGTTTAAGAGC
	oligo 2	TTAGCTCTTAAACTCTTCGCGGACACGTCCGGCCAACAAG
cTAGE5	oligo 1	TTGGCTGTCGGGGCCACAATAAAGTTTAAGAGC
	oligo 2	TTAGCTCTTAAACTTTATTGTGGCCCCGACAGCCAACAAG
	oligo 3	TTGGACGGTTGAGGGGTAACCCCGTTTAAGAGC
	oligo 4	TTAGCTCTTAAACGGGGTTACCCCTCAACCGTCCAACAAG
	oligo 5	TTGGGATTCGGGTTCCGGACCGAGTTTAAGAGC
	oligo 6	TTAGCTCTTAAACTCGGTCCGGAACCCGAATCCCAACAAG
TANGO1L	oligo 1	TTGGGCAGCCAGACGAGCAGCCCGTTTAAGAGC
	oligo 2	TTAGCTCTTAAACGGGCTGCTCGTCTGGCTGCCCAACAAG
	oligo 3	TTGGGCTGGCCCGGCACCCGCCAGTTTAAGAGC
	oligo 4	TTAGCTCTTAAACTGGCGGGTGCCGGGCCAGCCCAACAAG
	oligo 5	TTGGGGTGACCACAACATGGCTGGTTTAAGAGC
	oligo 6	TTAGCTCTTAAACCAGCCATGTTGTGGTCACCCCAACAAG
TANGO1s	oligo 1	TTGGGGAAAGGCCCCAATTCAGGGTTTAAGAGC
	oligo 2	TTAGCTCTTAAACCCTGAATTGGGGCCTTTCCCCAACAAG
	oligo 3	TTGGCGGCATAAAGCGAAAGTGCGTTTAAGAGC
	oligo 4	TTAGCTCTTAAACGCACTTTCGCTTTATGCCGCCAACAAG
	oligo 5	TTGGGAGTCCGCCCCCTGAATTGGTTTAAGAGC
	oligo 6	TTAGCTCTTAAACCAATTCAGGGGGCGGACTCCCAACAAG

m. Lentivirus production and transduction.

Lentivirus was generated in human Embryo Kidney (HEK) 293 T cells using Mirus Trans-IT®-293 Transfection Reagent following the manusfacturers' recommendation. At 50% confluency, cells incubated in appropriate media lacking penicillin/streptomycin were transfected with 10µg of plasmid pHR-SFFV-dCas9-BFP-KRAB (addgene #46911) or 10µg of plasmid pU6-sgRNA (addgene #60955) carrying specific sgRNA along with lentiviral packaging plasmids pVSVg (3.5µg) and psPAX2 (6.5µg) (Addgene). Transfection was performed for each sgRNA using one 10-cm dish. After a 24 h transfection, the media was changed and after additional 24 h, the media was removed and filtered through a 0.45µm low protein binding membrane (VWR International, Radnor, PA). One ml of media was used immediately for infection of 1 million target cells plated in 6-well plate, supplemented with 8µg/ml polybrene (Sigma Aldrich, St. Louis, MO) and lacking penicillin/streptomycin. First, cells were infected for dCas9-KRAB-BFP expression. For this, 48 h after infection, cells were analyzed by FACS to select BFP positive cells. Then for CRISPRi/dCas9-mediated knockdown, 48h after infection for sgRNA expression, 2µg/ml puromycin (Sigma Aldrich, St. Louis, MO) was added to select transduced cells. Finally, cells were analyzed by flow cytometry after additional 5 days for collagen XVII expression.

n. Flow cytometry analysis for collagen XVII expression.

4 days after lentivirus infection for CRISPRi/dCas9-mediated knockdown, cells were incubated with trypsin for 15 min at 37°C to cleave the pool of Collagen XVII express at the cell surface. Then, cells were washed with PBS and incubated during 18h in complete medium allowing recovery of collagene XVII at the cell surface. Cells were harvested after incubation for 10min with 0.5 mM EDTA, washed and resuspended in optiMem containing 5% FCS. Cells were then incubated with an anti collagen XVII antibody. for 1h at 4°C followed by PBS wash

and incubated with a secondary antibody PE-conjugated for 1h at 4°C. Finally, cells were analyzed on a LSR Fortessa flow cytometer.

o.Immunoprecipitation Torsin HA- TANGO1

All Torsin constructs were kind gifts from the Christian Schlieker Lab at the Yale school of medicine. All constructs Torsin- WT, EQ and KA in mammalian expression vectors with a C-terminal HA tag.

HEK293T cells were seeded to yield 80% confluency the next day. Cells were transfected with Control –HA vector, HA C terminally tagged Torsin variants- Wildtype, EQ-Substrate trap mutant and KA-ATP binding deficient mutant. 24 hours post transfection, cells were lysed in TBS+Triton 1% lysis buffer for about an hour with rotation in the cold.. The supernatant was clarified by centrifugation at high speed.

HA agarose beads (26181 Pierce Anti-HA Agarose) 50ul per lysate sample, were washed with lysis buffer for equilibration, the supernatant from the lysates were then added to the beads and incubated overnight at 4 degrees with rotation. The beads were collected by centrifugation and washed three times with TBS+0.1% triton. This was followed by one final wash with TBS and then sample was eluted with 100 ul 2x sample buffer by heating at 70 degrees for 10 mins.

The eluates were separated on a 8% (TANGO1) and 12%(Torsin-HA) with loading of 5ul input and 30ul eluate of the IP.

p.Quantification using Image J

For immunofluorescence images, the fluorescence intensity for the Collagen I corresponding to the ER region of cell were measured across at least 5 fields and 30 cells for each experiment. The mean intensity was plotted in control versus test cells and the value of control being normalized to 1.

In three independent experiments, intensities of the collagen signal in the lysate and the supernatant was recorded by densitometry. The ratio of collagen in media by lysate was normalized to quantify secretion in control.

q.Statistical analysis

All results shown are mean ± standard error of the mean (SEM). Statistical testing was performed using Student's t-test (continuous data, two groups). The Student's t-test was performed after a one-way ANOVA for the comparison of more than 2 groups. *P<0.05; **P<0.01; ***P<0.001.

r. Immunofluorescence techniques

Cells were grown on coverslips and after the relevant incubation period post transfection with plasmids, or knockdown with siRNAs they were washed carefully 3 times with PBS and fixed with ice cold methanol (for staining ER exit sites) or 4% PFA for 15 minutes and then rised with PBS 3 more times. In case of fixation of PFA permeabilization was performed with 0.1% triton in PBS for 10 minutes. Cells were gently washed with PBS and blocked with blocking reagent (catalog no.) for 1 hour. Primary antibodies were diluted in blocking buffer and incubated at either 3 hours at room temperatute or overnight at 4 degree.

Primary antibodies were washed off with PBS and then incubated with secondary antibodies for 1 hour. Samples were also stained with DAPI before final washes and the coverslips were mounted on slides using mounting media (Calbiochem) and allowed to set overnight before using them for Confocal microscopy.

All imaging was performed using Leica TCS SPE with a 63x objective, and analyzed using Image J software.

s.Antibodies

The antibodies used in this thesis are listed in table

Antibody	Source	Catalog no.	raised in/ clonality	Dilution
BiP	Abcam	ab21685	rabbit	IF 1:1000
Calreticulin	R&D Systems	AF3898	goat, polyclonal	IF 1:100
Collagen I	abcam	ab138492	rabbit	WB 1:1000
Collagen XVII	Abcam	ab184996	rabbit	IF 1:200, WB 1:2000
cTAGE5	atlas antibodies	HPA000387	rabbit	WB 1:1000
FLAG	sigma	F1804	mouse	IF 1:500, WB 1:1000
HA	Roche	11867423001	rat	IF 1: 500 WB 1:1000
HSP47	Enzo	ADI-SPA-470-F	mouse	IF 1:500, WB 1:1000
Procollagen Type I C-Peptide	QED Bioscience Inc.	42024	mouse	IF 1:200
Sec23IP	atlas antibodies	HPA038403	rabbit	IF 1:500
Sec24C	Randy Schekman lab	-	rabbit	WB 1:500
Sec31 A	BD	612350	mouse	IF 1:500
Streptavidin HRP	Pierce	-	-	WB 1:10000
TANGO1	Sigma	HPA055922	rabbit	IF 1:500, WB 1:2000
Torsin-1A	ThermoScientific	MA5-15094	mouse	WB 1 :1000

Results

4. Results

4.1 Identification of TANGO1 interactors utilizing a proximity dependent biotinylation combined with mass spectrometry approach.

To better comprehend the process of TANGO1 based collagen export, it is crucial to have a comprehensive inventory of its interactors. Although traditional immunoprecipitation approaches were previously used to detect binding partners, they often have the limitation of not identifying fleeting yet significant interactors. The APEX2 catalyzed proximity biotinylation method offered us the possibility of catching transient interactions in addition to stable partners. The APEX2 is an engineered ascorbate peroxidase enzyme with a molecular weight of 27 kDa (Hung et al., 2016). When affixed to a protein of interest, in the presence of hydrogen peroxide and biotin phenol as a substrate, the enzyme catalyzes the production of a biotin phenoxyl radical that labels only vicinal endogenous proteins with biotin. The cells are then lysed and the biotinylated population is captured with streptavidin beads and identified by mass spectrometry.

As a transmembrane protein, TANGO1 contains a cytoplasmic and an ER luminal domain, separated by a membrane. By creating constructs where TANGO1 was tagged with the APEX2 enzyme at the cytoplasmic and luminal ends, we attempted to isolate domain specific interactors. An additional FLAG tag was added to both constructs to enable easy detection by immunofluorescence and immunoblotting. The APEX2-FLAG fusion tag was added after the PRD domain on the cytoplasmic end, whilst on the luminal side, it was added between the signal sequence and SH3-like domain of TANGO1. (Figure 11 A).

4.1.1 TANGO1 APEX2 fusion constructs localize to Endoplasmic Reticulum Exit sites and are functional.

It was important to evaluate if the TANGO1 tagged with APEX2 was targeted to the right location. HeLa cells were grown on coverslips, transfected with, TANGO1-Cyto-APEX2 and TANGO1-Lum-APEX2. After 24 hours of incubation they were fixed, permeabilized and stained with antibodies against FLAG and

Sec31 an ER exit site marker. In both cases, the constructs co-localized with Sec31A, indicating no aberrant effects of the APEX2 fusion with respect to the ability of the protein to be targeted to the right area (Figure 11.B).

Both constructs were targeted to the region of interest, but were they functional? Could they still bind to their partner cTAGE5? To test this functionality we performed a simple immunoprecipitation experiments to see if these constructs had the ability to associate with its stable binding partner cTAGE5. The constructs were transiently transfected into HEK293T cells, 24 hours post transfection cells were lysed and immunoprecipitated using FLAG beads. The lysates and eluates were separated on an SDS-PAGE, transferred onto a nitrocellulose membrane and immunoblotted against FLAG and cTAGE5 antibodies. There was clear recovery of cTAGE5 with TANGO1-Lum-APEX2 and TANGO1-Cyto-APEX2 and none in the control lane (Figure 11 C).

This confirms that addition of APEX2 tag at the luminal and cytoplasmic ends of TANGO1 did not alter its localization to ER exit sites or interfere with binding to its partners.

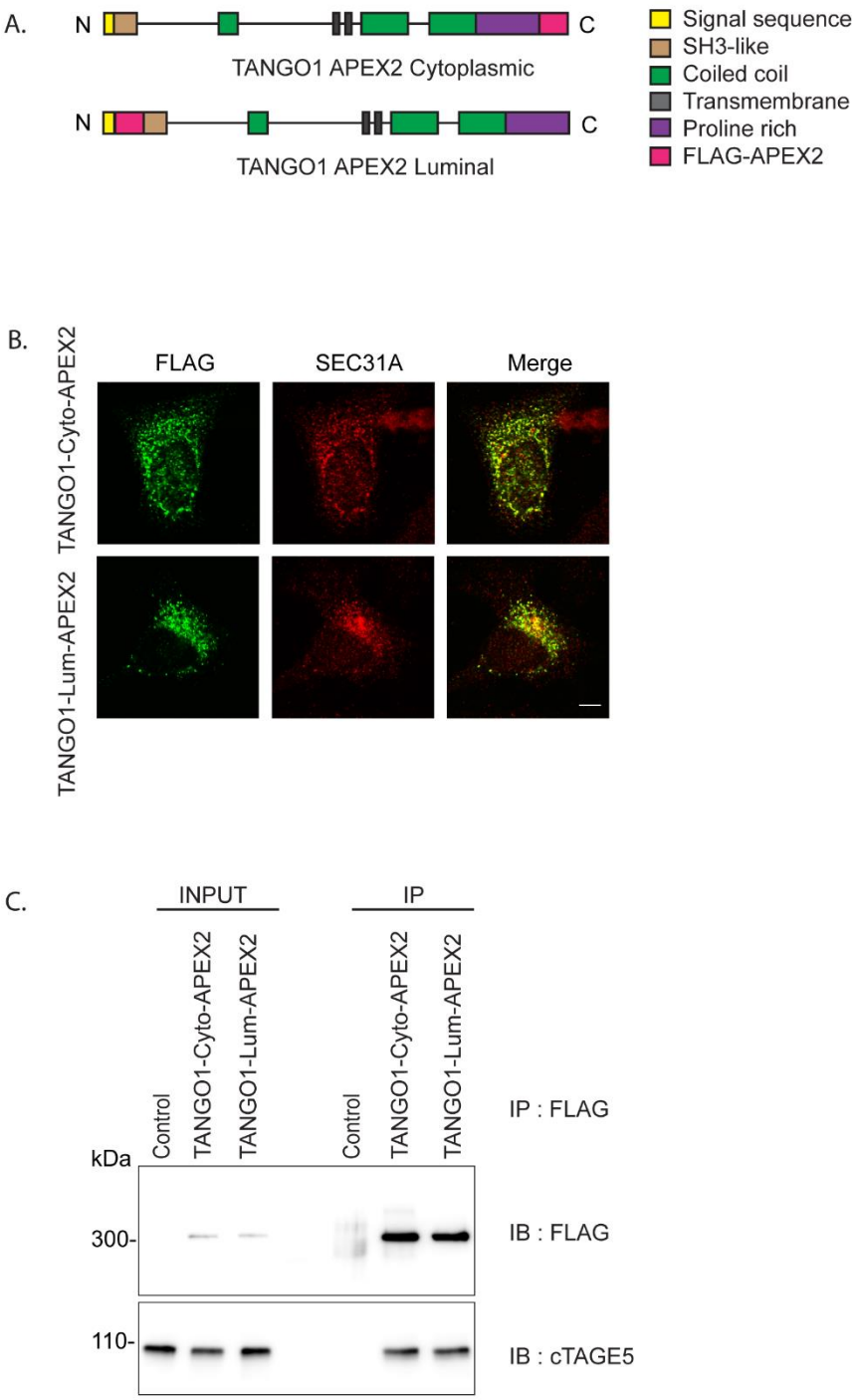

Figure 11.TANGO1-APEX2 fusion constructs localize to the ER exit sites and are functional

(A) Schematic depiction of APEX2 fusion to TANGO1 at cytoplasmic and Luminal ends.(B)TANGO1-Cyto-APEX2 and TANGO1-Lum-APEX2 were expressed in HeLa cells fixed, stained with anti-FLAG(green) and anti-Sec31(red) antibodies and visualized by confocal microscopy(Scale bar: 6um).

(C) Immunoprecipiation of Control FLAG tagged vector, TANGO1-Cyto-APEX2 and TANGO1-Lum-APEX2 in HEK293T cells. Lysates and Eluates from the immunoprecipitation were probed for FLAG and cTAGE5 antibodies.

4.1.2 TANGO1-Cyto-APEX2 and TANGO1-Lum-APEX2 biotinylates proteins in the Cytoplasm and ER.

The next step was to check if the APEX2 associated with TANGO1 had the capacity to biotinylate endogenous proteins. HeLa and HEK293T cell lines were transfected with TANGO1-Lum-APEX2 and TANGO1-Cyto-APEX2 together with a non-transfected control. 24 hours post transfection, the cells were incubated with 500uM Biotin Phenol in complete media for 30 minutes at 37 C degrees, H_2O_2 was added to the media at a final concentration of 1mM and the cells were incubated at room temperature for 1 minute after which the reaction was quenched and cells lysed. The lysates were separated by SDS-PAGE, transferred onto a nitrocellulose membrane and then immunoblotted with anti-Streptavidin-HRP. There were only three bands of endogenously biotinylated proteins in the control lane as opposed to several in TANGO1-APEX2 lanes (Figure 18). These results demonstrates the ability of the APEX2-tagged TANGO1 constructs to biotinylate a sizable pool of endogenous proteins.

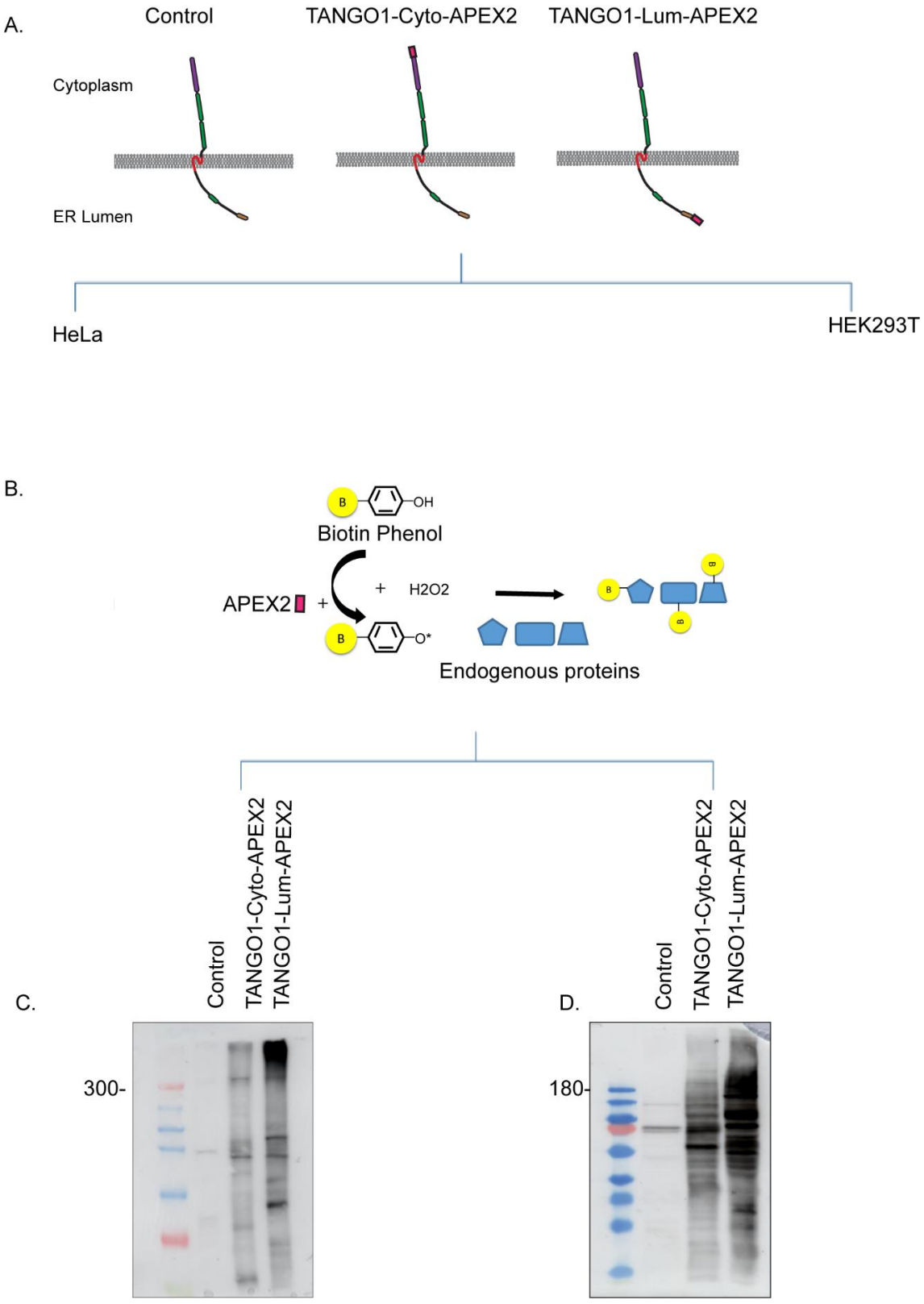

Figure 18.TANGO1- APEX2 fusion constructs biotinylate endogenous proteins in HeLa and HEK293T lines

(A) The constructs indicated were transfected into HeLa and HEK293T cell line

(B) The cells were incubated with biotin phenol in the presence of H_2O_2, the reaction generates biotin phenoxyl radicals that biotinylate endogenous proteins

(C) Immunoblotting of the biotinylated samples, from HeLa and HEK293T cells.

4.2 Identification TANGO1 interactome: Proteomic analysis by Mass spectrometry

4.2.1 Proteins identified by TANGO1-Cyto-APEX2

Data obtained from mass spectrometry was analyzed and filtered based on the criteria listed below to be considered potential hits

1. Zero peptides in the control sample.

2. At least >2 peptides in the experimental sample

Based on this filtering we obtained 41 proteins in HeLa cells and 235 proteins in HEK293T cells, this includes TANGO1 which was the protein with the highest enrichment. This is due to the fact that it is the bait protein and has the tendency to be autobiotinylated and therefore enriched during the immunoprecipitation step. The top 25 candidates from both cell lines have been listed in Appendix I and II. Interactors identified include, well known partners which have been shown to play a role in the collagen export pathway such as COP II components Sec23A, Sec24C and Sec31A, ER exit site protein Trk-fused gene (TFG), which has been shown to link ER exit sites to ERGIC, facilitating quick fusion of COPII carriers after uncoating with ERGIC (Hanna et al., 2017). Other intriguing hits include Annexin A11, which has been shown to stabilize Sec31 at ERES (Shibata et al., 2015) ,Sec23IP also known as p125,which is crucial for the organization of ER exit sites (Shimoi et al., 2005) and is essential for export of cargo from the ER (Ong et al., 2010). Casein kinase II subunit alpha (CK2), which is known to phosphorylate Sec31 and regulate COP II based transport (Koreishi et al., 2013) was also identified.p24 family proteins TMED7, TMED9 which function as cargo receptors are potential interactors as well. Myosin-9, a non-muscle myosin, which are known to participate in functions that require generation of force such as cytokinesis, cell migration, polarization, and adhesion, maintenance of cell shape, and signal transduction, has been identified with very high peptide abundance. Extended synaptotagmin-1, a protein known to tether ER to the Plasma membrane, it has also been shown to transfer lipids between membrane bilayers,

in a calcium dependent manner, is possibly an interactor of TANGO1 (Yu et al., 2016).

Actin modulating proteins Plastin-3, Profilin, and Transgelin 2 also feature in the top 25. We found proteins that are involved in vesicle motility like Kinectin. Cytoskeleton-associated protein 4 (CKAP4), also known as CLIMP-63 (cytoskeleton-linking membrane protein 63), is an ER transmembrane protein which binds microtubules on the cytoplasmic side and anchors ER to the cytoskeleton (Sandoz and van der Goot, 2015), was another curious hit.

The proteins biotinylated by TANGO1-Cyto-APEX2 in HEK293T cells, contained several well established interactors, suggesting right localization, better biotinylation and hits with higher confidence when compared to HeLa cells.

4.2.2 Proteins identified by TANGO1-Lum-APEX2

Excluding TANGO1, 113 and 202 proteins were biotinylated by the Luminal construct in Hela cells and HEK293T cells respectively. The top 50 hits based on peptide number and a brief description of their function have been listed in Appendix III A and IV A. The secreted cargoes identified were primarily ECM components- Collagens V, VI, VII and XII. Laminin chains α5, β1 and γ1 and peroxidasin homolog (Appendix III B and IV B). A large portion of machinery identified consists of chaperones and collagen modifying hydroxylases, glycosylation enzymes, oxidoreductases, calcium binding proteins all of which have a significant overlap among both cell lines. This allowed us to curate a list of 62 luminal interactors in both cell lines, giving us potential hits with high confidence. We further categorized them based on known functions into the following groups:

a. Enzymes, Chaperones and Foldases

In the group classified as Enzymes, Chaperones and Foldases, several proteins are known to be involved exclusively in the folding and modification of collagen (highlighted in bold).

They include the hydroxylases- Prolyl 3-hydroxylase 1, Prolyl 4-hydroxylase, Lysyl hydroxylase 2, Lysyl hydroxylase 3, Endoplasmic reticulum protein SC65, which carry out the most important modification of Collagens. Chaperones and co-chaperones of the DnaJ family, which promote folding and aid in degradation of misfolded proteins. Several protein disulfide isomerases Thioredoxin domain-containing protein 12, Thioredoxin domain-containing protein 5, Protein disulfide-isomerase TMX3, Protein disulfide-isomerase A5, they are oxidoreductases that isomerize disulfide bonds during folding.

Table 2 a: Chaperone, Enzymes and Folding complexes identified in both HeLa and HEK293T cell lines

Chaperone/Enzymes/Foldases		No. of peptides	
Protein names	Function	HeLa	HEK293T
Alpha-2-macroglobulin receptor-associated protein	Molecular chaperone for LDL receptor-related proteins	18	16
B-cell receptor-associated protein 31	Functions as a chaperone protein.	2	4
DnaJ homolog subfamily B member 11 protein 3	Serves as a co-chaperone for HSPA5.	8	13
DnaJ homolog subfamily B member 12	Acts as a co-chaperone with HSPA8/Hsc70; required to promote protein folding and trafficking, prevent aggregation of client proteins, and promote unfolded proteins to endoplasmic reticulum-associated degradation (ERAD) pathway	2	4
DnaJ homolog subfamily C member 10	Endoplasmic reticulum disulfide reductase involved both in the correct folding of proteins and degradation of misfolded proteins.	12	18
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	Catalytic subunit of the N-oligosaccharyl transferase (OST) complex which catalyzes the transfer of a high mannose oligosaccharide from a lipid-linked oligosaccharide donor to an asparagine residue within an Asn-X-Ser/Thr consensus motif in nascent polypeptide chains.	3	10

Chaperone/Enzymes/Foldases		No. of peptides	
Protein names	Function	HeLa	HEK293T
Endoplasmic reticulum protein SC65	Part of a complex composed of PLOD1, P3H3 and P3H4 that catalyzes hydroxylation of lysine residues in collagen alpha chains and is required for normal assembly and cross-linking of collagen fibrils.	5	3
Endoplasmic reticulum resident protein 44	Mediates thiol-dependent retention in the early secretory pathway.	11	18
ERO1-like protein alpha	Oxidoreductase involved in disulfide bond formation in the endoplasmic reticulum.	5	19
GDP-fucose protein O-fucosyltransferase 1	Catalyzes the reaction that attaches fucose	2	8
GPI transamidase component PIG-S	Component of the GPI transamidase complex.	3	3
GPI transamidase component PIG-T	Component of the GPI transamidase complex.	4	5
Inactive C-alpha-formylglycine-generating enzyme 2	Inhibits the activation of sulfatases by SUMF1.	2	7
LRP chaperone MESD (LDLR chaperone MESD)	Chaperone specifically assisting the folding of beta-propeller/EGF modules within the family of low-density lipoprotein receptors (LDLRs).	3	7
Malectin	Carbohydrate-binding protein with a strong ligand preference for Glc2-N-glycan.	4	7
Nucleotide exchange factor SIL1	Required for protein translocation and folding in the endoplasmic reticulum (ER)	5	3
Peptidyl-prolyl cis-trans isomerase FKBP9	PPIases accelerate the folding of proteins during protein synthesis.	9	2
Prenylcysteine oxidase 1	Involved in the degradation of prenylated proteins.	7	13
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 (Lysyl hydroxylase 2)	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens. T	16	11
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 (Lysyl hydroxylase 3)	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens.	19	18

Chaperone/Enzymes/Foldases		No. of peptides	
Protein names	Function	HeLa	HEK293T
Prolyl 3-hydroxylase 1	prolyl 3-hydroxylase activity catalyzing the post-translational formation of 3-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens, especially types IV and V.	19	12
Prolyl 4-hydroxylase subunit alpha-2 (4-PH alpha-2)	Catalyzes the post-translational formation of 4-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens and other proteins.	17	10
Protein canopy homolog 3	Toll-like receptor (TLR)-specific co-chaperone for HSP90B1.	3	4
Protein disulfide-isomerase A5	Protein disulfide isomerase	2	7
Protein disulfide-isomerase TMX3	Probable disulfide isomerase, which participates in the folding of proteins containing disulfide bonds.	3	4
Reticulocalbin-2	Binds calcium.	10	8
Selenoprotein F (15 kDa selenoprotein)	May be involved in redox reactions associated with the formation of disulfide bonds.	4	4
Thioredoxin domain-containing protein 12	Possesses significant protein thiol-disulfide oxidase activity.	5	5
Thioredoxin domain-containing protein 5	Possesses thioredoxin activity.	18	20
Vitamin K-dependent gamma-carboxylase	Mediates the vitamin K-dependent carboxylation of glutamate residues.	2	2
Xyloside xylosyltransferase 1	Alpha-1,3-xylosyltransferase.	2	4

b. Proteins involved in vesicular trafficking

The proteins with known or predicted functions in the trafficking process are listed in this category (Table 2b).The known interactor in this list is cTAGE5, a binding partner of TANGO1 and a key requirement for collagen secretion. ERGIC-1 and 3, seem to be interacting with TANGO1, not much is known about their function and they cycle between the ER and the Golgi. 3 members of the p24 family of proteins, TMED 4, 7 and 9 which are cargo receptors which were also hits in the cytoplasmic Apex2 were interactors on the luminal list as well.

Table 2 b: Proteins involved in trafficking identified in both HeLa and HEK293T cell lines

Protein Trafficking		No. of peptides	
Protein names	Function	HeLa	HEK293T
Endoplasmic reticulum export factor CTAGE5	Required for collagen VII (COL7A1) secretion by loading COL7A1 into transport carriers and recruiting SEC12 at the endoplasmic reticulum exit sites. {	11	4
Endoplasmic reticulum-Golgi intermediate compartment protein 1	Role in transport between endoplasmic reticulum and Golgi.	9	4
Endoplasmic reticulum-Golgi intermediate compartment protein 3	Role in transport between endoplasmic reticulum and Golgi.	2	2
Transmembrane emp24 domain-containing protein 4 (TMED4)	Involved in vesicular protein trafficking,	2	4
Transmembrane emp24 domain-containing protein 7 (TMED7)	Role in vesicular protein trafficking, mainly in the early secretory pathway.	4	5
Transmembrane emp24 domain-containing protein 9 (TMED9)	In COPI vesicle-mediated retrograde transport involved in the coatomer recruitment to membranes of the early secretory pathway.	4	9
Vesicular integral-membrane protein VIP36	Plays a role as an intracellular lectin in the early secretory pathway.	2	4
V-type proton ATPase subunit S1 (V-ATPase subunit S1)	. Involved in membrane trafficking and Ca(2+)-dependent membrane fusion.	2	3
Zinc transporter SLC39A7	Zinc transporter, that transports Zn(2+) from the endoplasmic reticulum/Golgi apparatus to the cytosol.	2	2

c. Proteins with Diverse Functions

Unique interactors, which function in multiple pathways are listed in this category (Table 2c). Some proteins stand out are that Nodal signaling agonists, NOMO1 and 2 identified with a very high enrichment, the unconventional AAA ATPase

Torsin-1A and its modulators LAP1 and LULL1 and Integrin β 1, a collagen receptor at the cell surface.

Table 2c: Proteins with Diverse functions identified in both HeLa and HEK293 cell lines

Diverse Functions		No. of peptides	
Protein names	Function	HeLa	HEK293T
Chloride channel CLIC-like protein 1	Chloride ion channel.	7	8
ER membrane protein complex subunit 1	Unknown	15	27
HLA class I histocompatibility antigen, Cw-17 alpha chain	Involved in the presentation of foreign antigens to the immune system.	2	3
Inhibitor of nuclear factor kappa-B kinase-interacting protein	Target of p53/TP53 with pro-apoptotic function.	17	13
Integrin beta-1	Receptor for collagen.	9	3
KDEL motif-containing protein 2	Unknown	4	10
Nicastrin	An endoprotease complex that catalyzes the intramembrane cleavage of integral membrane proteins such as Notch receptors and APP (amyloid-beta precursor protein.	5	3
Nodal modulator 1	Nodal signaling antagonist	26	33
Nodal modulator 2	Nodal signaling antagonist	26	35
Nucleobindin-2	Calcium-binding protein.	5	15
Reticulocalbin-2	Binds calcium.	10	8
Torsin-1A	Control of protein folding, processing, stability and localization as well as for the reduction of misfolded protein aggregates.	2	4
Torsin-1A-interacting protein 1	Required for nuclear membrane integrity. Induces TOR1A and TOR1B ATPase activity	2	10

Diverse Functions		No. of peptides	
Protein names	Function	HeLa	HEK293T
(LAP1)			
Torsin-1A-interacting protein 2 (Lumenal domain-like LAP1)	Required for endoplasmic reticulum integrity. Regulates the distribution of TOR1A between the endoplasmic reticulum and the nuclear envelope as well as induces TOR1A, TOR1B and TOR3A ATPase activity.	3	2
Transmembrane protein 109	Mediates cellular response to DNA damage by protecting against ultraviolet C-induced cell death.	2	2
Transmembrane protein 43	May have an important role in maintaining nuclear envelope structure by organizing protein complexes at the inner nuclear membrane.	7	13

d. Secretory cargoes

Proteins that are secreted were classified as cargoes (Table 2d)

Table 2d: Secreted cargoes identified in both HeLa and HEK293 cell lines

Secreted Cargoes		No. of peptides	
Protein names	Function	HeLa	HEK293T
ADP-dependent glucokinase	Catalyzes the phosphorylation of D-glucose to D-glucose 6-phosphate using ADP as the phosphate donor. GDP and CDP can replace ADP, but with reduced efficiency.	6	10
Laminin subunit beta-1	Mediates the attachment, migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components.	20	3
Mesencephalic astrocyte-derived neurotrophic factor	Selectively promotes the survival of dopaminergic neurons of the ventral mid-brain.	2	12

Secreted Cargoes		No. of peptides	
Protein names	Function	HeLa	HEK293T
Myeloid-derived growth factor (MYDGF)	Bone marrow-derived monocyte and paracrine-acting protein that promotes cardiac myocyte survival and adaptive angiogenesis.	5	4
Neudesin	Acts as a neurotrophic factor in postnatal mature neurons enhancing neuronal survival.	2	9

We also noticed that when we analyzed proteins based on the criteria that we have 0 peptides in the control we lost several previously identified interactors as these proteins bound with some affinity to the beads (Table 3); these include the collagen specific chaperone HSP47 (also known as Serpin H1), which has been shown to binding to triple helical collagen and to the SH3 domain of TANGO1 (Ishikawa et al., 2016). GRP-94, belongs to the HSP-90 family, is an ER luminal chaperone for several proteins including collagen (Marzec et al., 2012). Protein Disulfide isomerases- PDI, A3, A4, A6 and the general chaperone BiP were also found.

Table 3: Luminal interactors found in control sample but highly enriched in test samples

		No. of peptides			
		HeLa		HEK293T	
Protein names	Function	Control	TANGO1-Lum-APEX2	Control	TANGO 1-Lum-APEX2
Calumenin	Involved in regulation of vitamin K-dependent carboxylation	7	20	-	16
Coiled-coil domain-containing protein 47	endoplasmic reticulum degradation activity	-	16	6	25

		No. of peptides			
		HeLa		HEK293T	
Protein names	Function	Control	TANGO1-Lum-APEX2	Control	TANGO 1-Lum-APEX2
Endoplasmic reticulum chaperone BiP	Endoplasmic reticulum chaperone that plays a key role in protein folding and quality control in the endoplasmic reticulum lumen	25	55	12	53
Endoplasmic reticulum resident protein 29 (ERp29)	Plays an important role in the processing of secretory proteins within the endoplasmic reticulum (ER),	4	16	-	18
Endoplasmin (GRP-94)	Molecular chaperone that is in the processing and transport of secreted proteins.	23	69	15	64
Glucosidase 2 subunit beta	Regulatory subunit of glucosidase II	2	15	-	20
Hypoxia up-regulated protein 1	molecular chaperone	3	44	2	40
Neutral alpha-glucosidase AB	Catalytic subunit of glucosidase II	9	34	-	45
Protein disulfide-isomerase (PDI)	This protein catalyzes the formation, breakage and rearrangement of disulfide bonds	2	27		36
Protein disulfide-isomerase A3	Chaperone	13	30	11	43
Protein disulfide-isomerase A4	Chaperone	6	40		58
Protein disulfide-isomerase A6	chaperone that inhibits aggregation of misfolded proteins	5	14	4	24
Reticulocalbin-1	Regulates calcium-dependent activities in the endoplasmic reticulum lumen or post-ER compartment.	5	20		17

		No. of peptides			
		HeLa		HEK293T	
Protein names	Function	Control	TANGO1-Lum-APEX2	Control	TANGO 1-Lum-APEX2
Serpin H1 (47 kDa heat shock protein)	Binds specifically to collagen.	9	19	6	18

TANGO1 interactome identified by proximity biotinylation has several proteins in common with the Collagen Proteostasis network proteins

Several proteins identified as a part of the luminal TANGO1 interactome were chaperones, enzymes and foldases in the ER some of which are exclusive machinery for collagen folding. In a recent paper by DiChiara and colleagues, they created inducible fibrosarcoma cell lines, which express tagged collagen I, which was immunoprecipitated and quantitative mass spectrometry was performed. This yielded collagen I specific interactors, which they term as the collagen proteostasis network, it mainly consists of proteins involved in collagen folding, quality control and secretion (DiChiara et al., 2016). On comparing the two lists, we found a large overlap (Figure 13). This to us, presented the possibility that TANGO1 not only binds to fully folded collagen, it binds to collagens as it is in the process of folding.

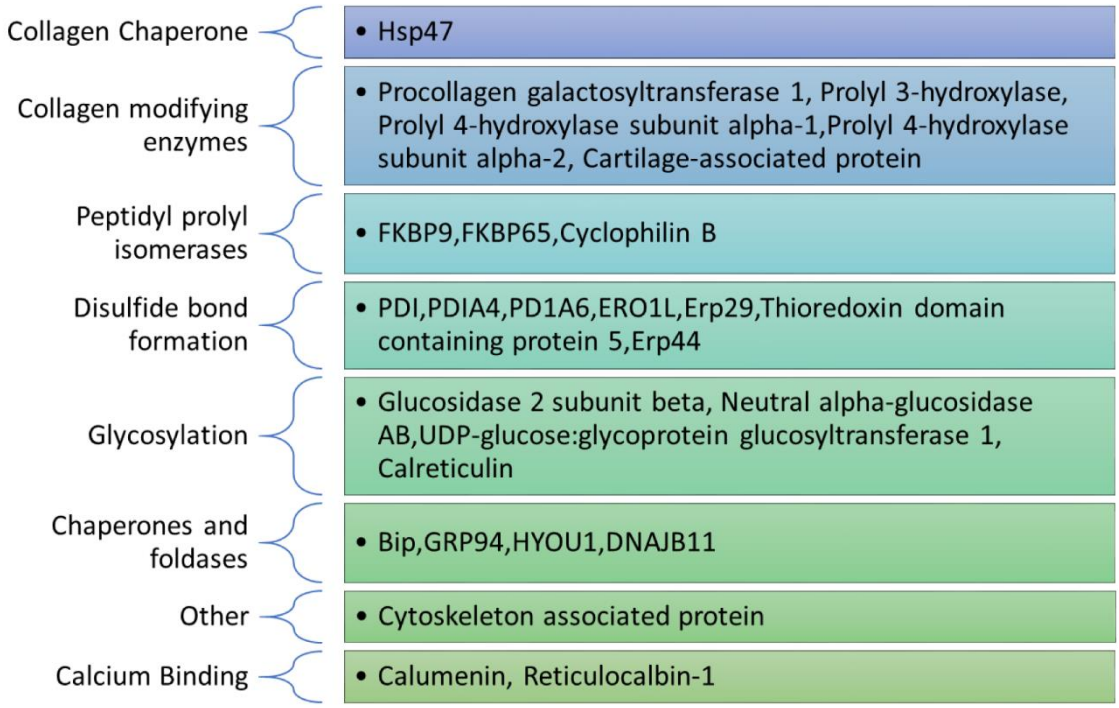

Figure 13. Proteins that interact with TANGO1 and Collagen I
A comparison of TANGO1 luminal interactors and Collagen I interactors showed overlap of a large number of proteins which are chaperones, modifying enzymes, peptidyl prolyl isomerases and calcium binding proteins.

4.3 A role for SEC23IP in the collagen export pathway

An interesting candidate amongst our cytoplasmic interactors was Sec23A-interacting protein (Sec23IP), also known as p125A, as the name suggests was first identified as an interactor of Sec23A (Tani et al., 1999). It was later shown to be mostly in complex with another COP II component Sec31 in the cytosol. After its recruitment to ER membranes, it binds Sec23A connecting the inner and outer COPII coats (Ong et al., 2010). Knockdown and overexpression studies have demonstrated the requirement of Sec23IP in organizing ER exit sites (ref). Knockdown studies have shown delayed cargo export of both secreted and transmembrane cargoes from the ER. It belongs to the Phospholipase A family, its C-terminal DDHD and SAM domains recognize and bind PI4P rich membranes and COP II assembly is modulated by PI4P (Klinkenberg et al., 2014) . The knockout of Sec23IP in mice show no major defects except for in spermatogenesis (Arimitsu et al., 2011). While its depletion in Xenopus has been linked to defects in neural crest cell migration, which could be due to defects in collagen secretion (McGary et al., 2010).

In humans a homozygous mutation in Sec23IP, which results in the deletion of the lipid binding module causes severe brain, cranial and skeletal development defects (Aridor, 2018; Reuter et al., 2017). The phenotype is quite similar to the craniofacial defects in humans caused by mutations in Sec23A which have defects in coupling (Boyadjiev et al., 2011, 2006) of the coat layers. All of the above hint towards a possible role in the collagen export pathway. So, when Sec23IP was as one of the possible interactors of TANGO1 on the cytoplasmic side, we were curious to test if it affected collagen export.

4.3.1 Loss of Sec23IP leads to intracellular accumulation of Collagen I

To test the role of Sec23IP in the export of collagens we used a knockdown approach in U2OS cells. Control and Sec23IP siRNAs were transfected over the course of two consecutive days in U2OS cells, 24 hours after the last transfection, cells were plated on coverslips and cultured for an additional 24 hours. The cells

were then washed, fixed, permeabilized and stained with antibodies against Sec23IP, to check the knockdown. To test for Collagen accumulation, we stained them with Collagen I and Calreticulin antibodies. The knockdown efficiency was very high, almost all cells were completely devoid of Sec23IP (Figure 14). Control U2OS cells showed very low levels of intracellular collagen indicating normal secretion while cells where Sec23IP was depleted had an increase in intracellular collagen levels.

Knockdown of Sec23IP, leads to intracellular accumulation of Collagen I. Sec23IP has been shown to affect transport of cargoes from the ER, what remains unexplored at the time of writing this thesis is if the collagen export an indirect effect of disruption of ERES or does Sec23IP bind TANGO1 directly and regulate its COPII binding and membrane addition facilitating efficient export of collagen.

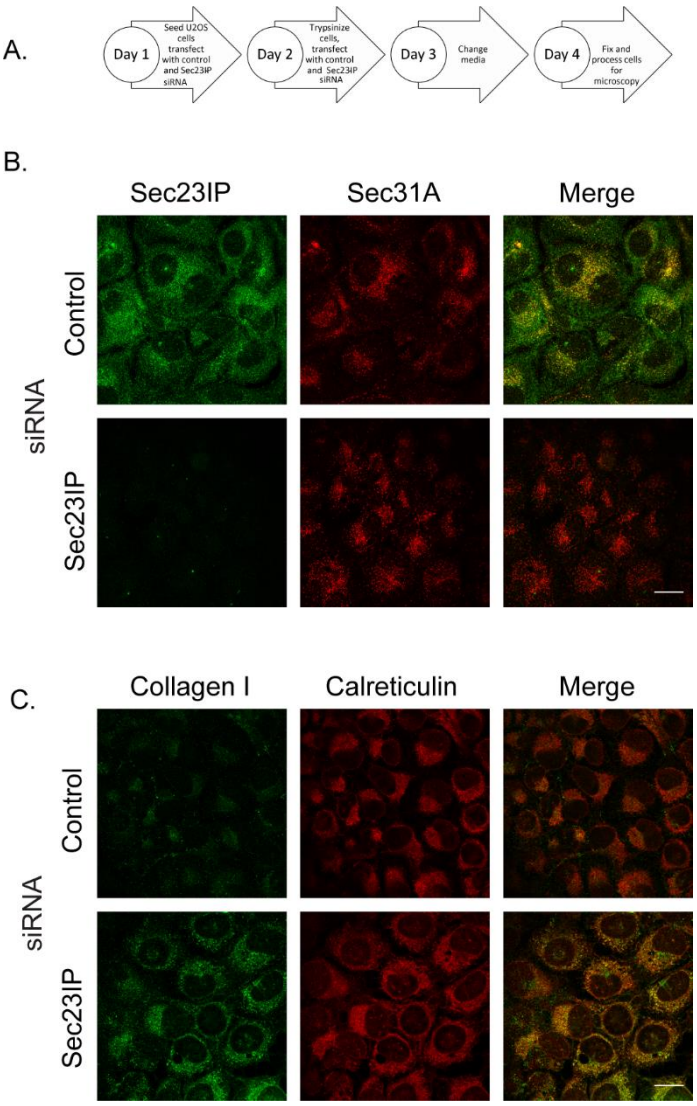

Figure 14. Sec23IP a protein at the ER exit site is required for Collagen I export

(A) Schematic of the experiment. U2OS cells were transfected with control and Sec23IP siRNAs for 2 consecutive days. 48 hours later cells were processed for microscopy.

(B) U2OS cells knocked down for Control or Sec23IP siRNA, were fixed, permeabilized and stained with antibodies against Sec23IP (green) and Sec31A(red) and visualized using confocal microscopy.

(C) U2OS cells knocked down for Control or Sec23IP siRNA, were fixed, permeabilized and stained with antibodies against Collagen I (green) and Calreticulin(red) and visualized using confocal microscopy. Scale bars 18 um

4.4 Torsin-1A

Amongst the several luminal interactors of TANGO1 that we identified, one protein, Torsin-1A though not very abundant looked to be an intriguing candidate. Torsin-1A belongs to the AAA+ (ATPases associated with cellular activities) family of proteins. It was first identified as the gene whose mutation is responsible for the movement disorder known as early onset torsion dystonia (Ozelius 1997). AAA+ proteins are known for using energy from the hydrolysis of ATP to orchestrate protein folding and unfolding, disassembly of large protein complexes, and fusion of membranes(White and Lauring, 2007).

They are expressed by all types of cells, with neuronal enrichment(Jungwirth et al., 2010) The five members of the Torsin family (Torsin 1A, Torsin 1B, Torsin2A, Torsin3A, and Torsin4A) (Rose et al., 2015). Torsin is the lone AAA ATPase protein found to date in the Endoplasmic reticulum and the nuclear envelope. It is widely conserved across metazoans, and its indispensability is supported by data from Torsin 1A knockout mice die within 48 hours post birth(Goodchild et al., 2005).

Torsin is an unconventional ATPase because it has no activity unless it is bound to its co-factor. Two transmembrane proteins, lamina-associated polypeptide 1 (LAP1) and luminal domain-like LAP1 (LULL1) function as co-factors in the nuclear envelope and endoplasmic reticulum respectively(Zhao et al., 2013).All Torsin containing organisms also possess the co-factors.

Structure wise it has the typical AAA+ architecture, an AAA domain with a conserved lysine, K108 in the Walker A motif for ATP binding and a glutamate E171 for ATP hydrolysis followed by an α-helical domain at the C-terminus(Hanson and Whiteheart, 2005). In vitro it forms hexameric assemblies with the co-factors.

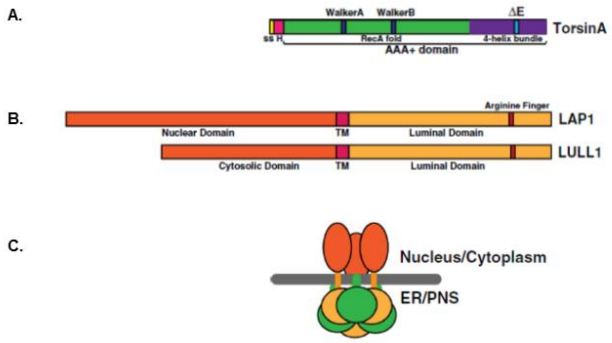

Figure 15. Domain organization of Torsin and its cofactors (Laudermilch and Schlieker 2016)

(A) Domain organization of Torsin-1A, ss: signal sequence; H: hydrophobic domain; ΔE: in-frame glutamate deletion that leads to DYT1 dystonia; TM: transmembrane domain.
(B) Schematic of the cofactors of Torsin-1A, LAP1 and LULL1
(C) Torsin can form a hexameric assembly with either cofactor.

Its multitudinous functions include in the ER include, protein quality control, trafficking of polytopic membrane proteins(Hewett et al., 2007; Nery et al., 2011; Torres et al., 2004)

Loss of Torsin has been shown to elevate the susceptibility of the cells to ER stress(Chen et al., 2010).

Our goal was to test if Torsin-1A had a role in the process of TANGO1 dependent collagen export.

4.4.1 Torsin-1A, an ER luminal protein intermittently co-localizes with Hsp47 and TANGO1

Torsin-1A is known to be located in the lumen of the endoplasmic reticulum. We wanted to test if it showed co-localization with TANGO1, its potential interactor and also with the collagen specific chaperone Hsp47.U2OS cells, were transfected with an HA tagged construct of Torsin-1A, 24 hours later the cells were fixed, permeabilized and stained with antibodies against HA, Hsp47, TANGO1 and visualized by confocal microscopy. The tagged Torsin construct was used due to the unavailability of a good antibody for immunofluorescence.

Torsin-1A(in red) had distinct reticular and nuclear envelope staining. It co-localized with Hsp47(in green) in most areas throughout the ER (Figure 16). TANGO1 appeared as punctate structures sometimes co-localizing with Torsin

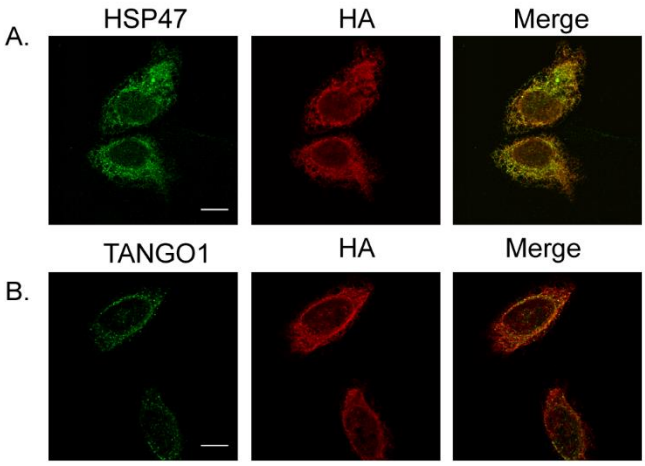

Figure 16. Torsin-1A is an ER luminal protein
U2OS cells were transfected with HA tagged Torsin-1A, cells were fixed, permeabilized and stained for Hsp47, HA, TANGO1 and visualized by confocal microscopy. Scale bar: 5um

4.4.2 Torsin-1A interacts with TANGO1

Torsin-1A was identified as an interactor through proximity biotinylation assay, which means that it was in the close vicinity of TANGO1. But does it interact with TANGO1? We decided to test if we could capture this interaction using an immunoprecipitation approach. Since Torsin is an AAA ATPase, we also wondered if this binding was ATP dependent and if TANGO1 interaction with Torsin was more stable in its ATP bound state. HA tagged Torsin-1A wild type and 2 mutants designated EQ, which it can bind but not hydrolyze ATP and KA a mutant that is deficient in binding ATP or an empty vector as control, were transfected into HEK293Tcells. 24 hours post transfection, cells were harvested, lysed and immunoprecipitated using HA agarose beads. The lysates and eluates were separated on an SDS-PAGE, transferred onto a nitrocellulose membrane and immunoblotted using antibodies against TANGO1 and HA. No TANGO1 was immunoprecipitated in the control sample, while all three variants, WT, EQ and KA were able to bind TANGO1 (Figure 17). This result is concurrent with the mass spectrometry data, Torsin-1A does bind to TANGO1.

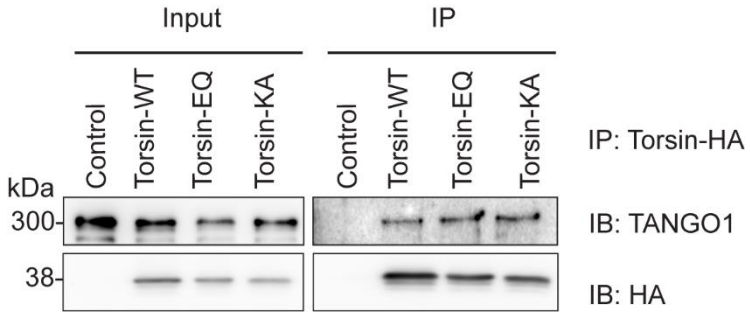

Figure 17. Torsin-1A binds TANGO1
HA taggedTorsin- WT,EQ and KA variants were expressed in HEK293T cells, immunoprecipitated and probed for HA and TANGO1.

4.4.3 Torsin-1A knockdown does not alter co-localization of TANGO1 with Sec31

We found that Torsin-1A interacts with TANGO1 and we wanted to test if it played a role in the process of TANGO1 mediated collagen export. The first step was to see if TANGO1 showed any change in localization in its absence.

Collagen I secreting U2OS cells were subjected to knockdown using control and Torsin-1A siRNA, the efficiency of knockdown was tested at the mRNA level using RT-PCR. The cells plated on coverslips were fixed, permeabilized and stained with antibodies against, Sec31 and TANGO1 and visualized by confocal microscopy.

Similar to the control condition, TANGO1 co-localized with Sec31 (Figure 18), Overall there seemed to be no altered localization of TANGO1 in the absence of Torsin.

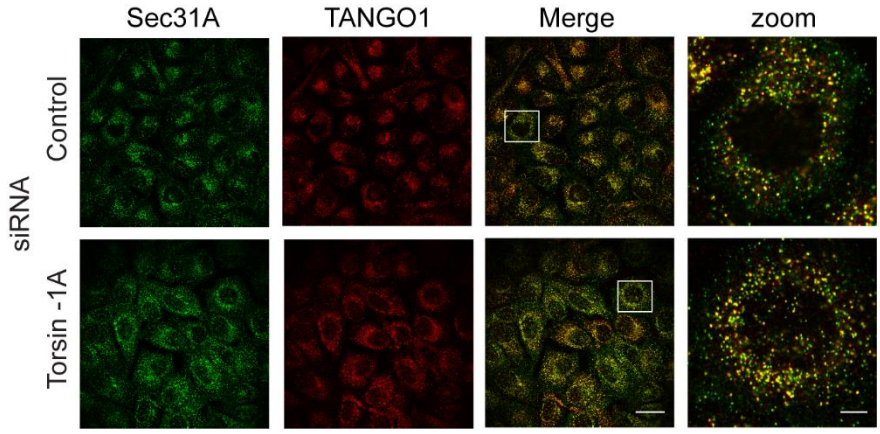

Figure 18. TANGO1 co localization with Sec31A remains unaltered in Torsin-1A knockdown
U2OS cells were transfected with control and Torsin-1A siRNAs. 48 hours after the last knockdown, cells were fixed, permeabilized and stained for TANGO1 and Sec31A and visualized by confocal microscopy. Scale bar: 5um

4.4.4 Depletion of Torsin-1A leads to a buildup of Collagen I in the Endoplasmic reticulum

Given that TANGO1 is required for collagen secretion(Saito et al., 2009; Wilson et al., 2011), and Torsin interacts with TANGO1, we wanted to test if Torsin had any role in the modulating the export of collagens.

U2OS cells, which produce Collagen I endogenously were transfected with Control and Torsin-1A siRNAs and 48 hours after the last knockdown, the cells were fixed and stained for Collagen I and ER resident chaperone BiP (Figure 19A). The knockdown levels of Torsin were checked at the mRNA level by RT-PCR (Figure 19 B). The control cells showed low levels of Collagen I inside the cells, while in cells lacking Torsin, there was a strong buildup of collagen I in the ER as evinced by co-localization with BiP (Figure 19 C). Upon quantification of the immunofluorescent signal from Collagen I in control and Torsin-1A knockdown cells, there was upto 4 fold increase of Collagen I in the cells depleted of Torsin-1A. (Figure 19 D)

.

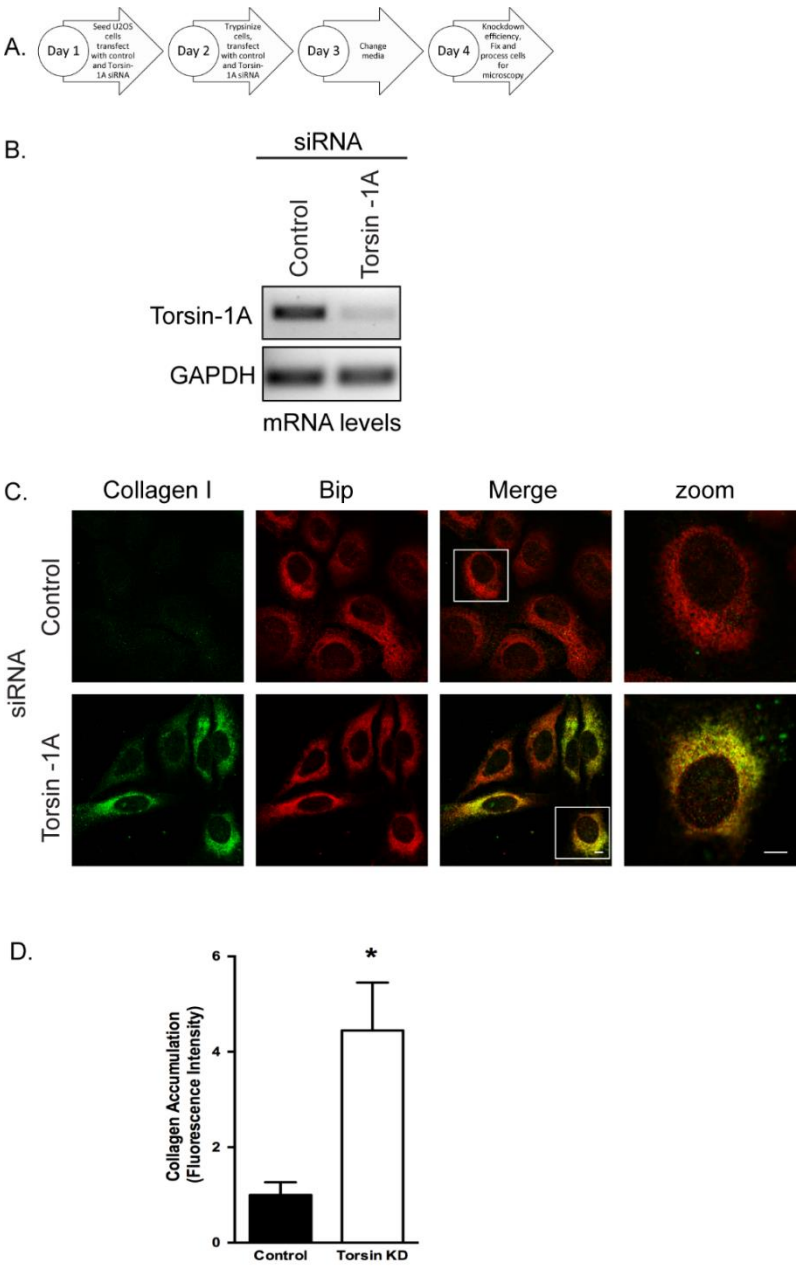

Figure 19. Knockdown of Torsin-1A results in a buildup of Collagen I in the ER

(A) Schematic of the experiment. U2OS cells were transfected with control and Torsin-1A siRNAs for 2 consecutive days. 48 hours later cells were tested for Torsin knockdown and processed for microscopy.

(B) mRNA levels of Torsin-1A and GAPDH, in control and Torsin-1A knockdown.

(C) U2OS cells knocked down for Control or Torsin-1A siRNA, were fixed, permeabilized and stained with antibodies against Collagen (green) and Bip(red) and visualized using confocal microscopy. Scale bars: 27μm and 5μm

(D) Quantification of fluorescence intensity of Collagen I. Error bars: standard error of the mean (SEM).

4.4.5 Lack of Torsin-1A shows an increase in intracellular Collagen I levels but no defect in secretion

The buildup of intracellular collagen I in cells devoid of Torsin-1A was very similar to the phenotype observed when TANGO1 is knocked down. Was lack of Torsin preventing the export of collagen? To test this we performed a secretion assay, with a chase over a period of 2 hours after cycloheximide treatment to prevent synthesis of new collagen.

U2OS cells, were transfected with Control and Torsin-1A siRNA. 48 hours after the last knockdown, the cells were washed with PBS and incubated with serum free media with the addition of cycloheximide and fortified with ascorbic acid. Lysate and media samples were collected at hours 0, 1 and 2. The samples were analyzed by immunoblotting against Collagen I, Tubulin and Torsin. There was significant amount of collagen in the lysate samples at hour 0 in the cells knocked down with Torsin when compared to the control at the same time point (Figure 20). Tubulin levels are equal in both samples. The media samples for both control and Torsin showed similar rates of secretion at 1 and 2 hour time points. There clearly seems to be an increase in collagen levels internally, without a block in secretion to the extracellular milleu.

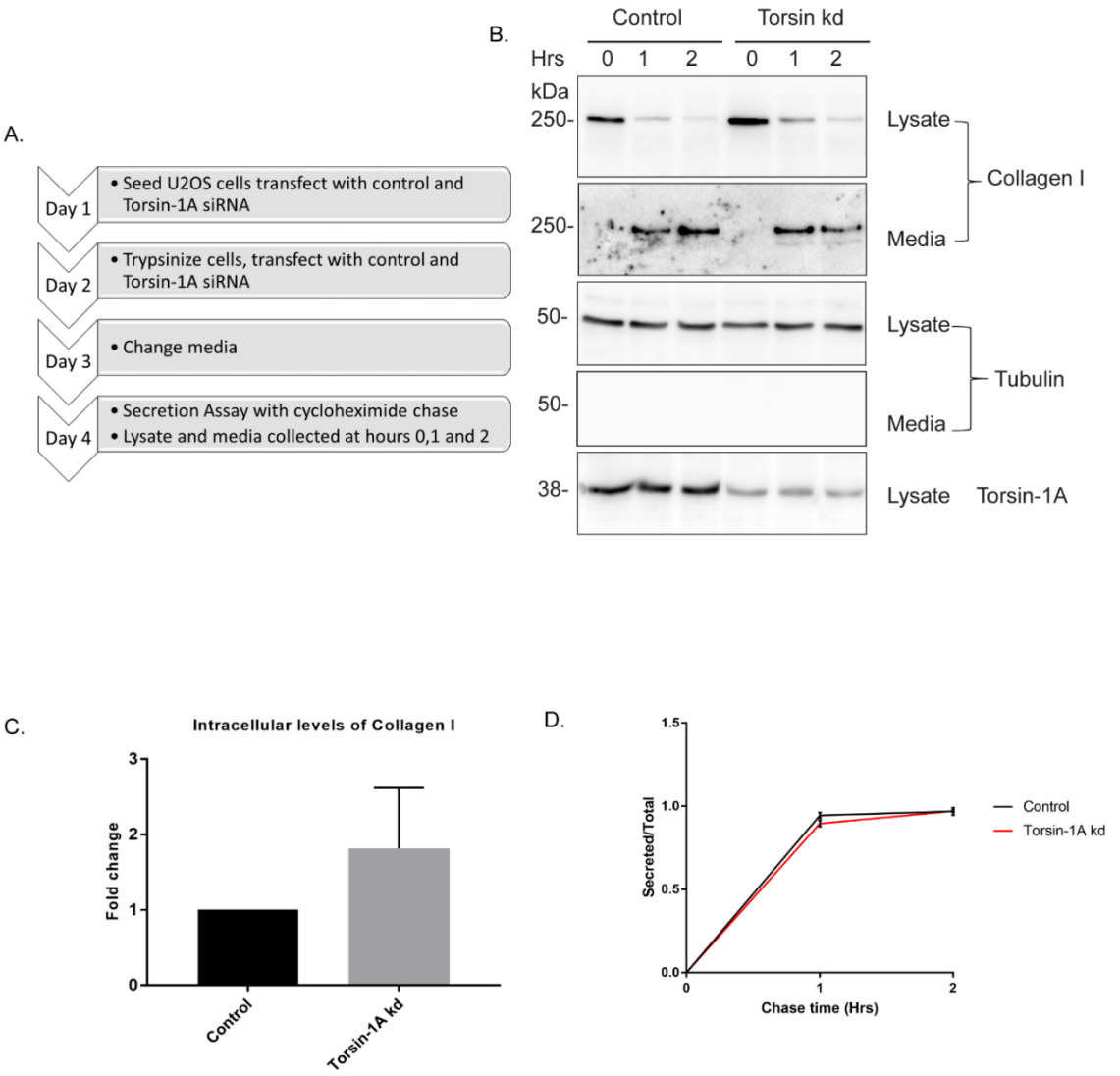

Figure 20.Torsin-1A does not affect collagen I secretion

(A) Schematic of the experiment. U2OS cells were knocked down with control or Torsin-1A siRNA, 48 hours later secretion assay was performed by incubating cells with media containing cycloheximide and ascorbic acid. Lysates and media were collected at 0, 1 and 2 hours.

(B) Lysate and media samples were immunoblotted using antibodies against Collagen I, Torsin-1A and Tubulin.

(C) Intracellular levels of Collagen I, normalized to tubulin at hour 0 from three independent experiments were quantified by densitometry.

(D) The rate of collagen secretion was measured from three independent experiments and plotted as a ratio of secreted by total. Error bars: standard error of the mean (SEM).

4.4.6 Loss of Torsin-1A affects degradation of unfolded collagen I

We wanted to test if the Torsin had a role in aiding the disposal of either excess or unfolded collagen. It is of note that in the absence of TANGO1, levels of Torsin increase by about 1.6 fold (Wilson et al., 2011) could this be to aid in the disposal of collagens?

We needed to establish an assay wherein we could clear out all the folded collagen from the ER, synthesize only unfolded collagen which would then be held back in the ER. This would then enable us to look at the rate of degradation of collagen from the ER.

For this I decided to use the drug 2,2-dipyridyl, which chelates iron and perturbs prolyl hydroxylases functioning thus preventing hydroxylation and leading to complete disruption of folding (Bonfanti and Mironov; Mironov et al., 2003) Cells were knocked down using Control and Torsin siRNAs, 48 hours after the last knockdown cells were treated with vehicle control and 0.3mM dipyridyl for 1 hour. One hour is the time needed for Collagen I synthesized in the ER to get to the Golgi. The newly synthesized Collagen is unable to be hydroxylated during the drug treatment and therefore becomes unfolded. At this point cells were shifted to serum free media containing Control vehicle and cycloheximide and dipyridyl with cycloheximide. The cells were incubated at 37 degrees and samples were collected at 0, 1.5 and 3 hours' time points.

The samples were analyzed by western blotting and immunoblotted with antibodies against Collagen I and tubulin. As seen before in the secretion assay of Torsin knockdown, in the untreated samples there was a marked increase in the intracellular levels of Collagen in the cells lacking Torsin 1a.

In control cells of the dipyridyl treated samples, there was a decrease in unfolded collagen levels at time points 1.5 and almost complete disappearance of intracellular pool without secretion. Whereas in the Torsin knockdown cells the pool of unfolded collagen remained quite stable even at 1.5 points. with some secretion into the media. These results indicate the possibility that lack of Torsin

impedes the ability of the ER to deal with the clearance of unfolded collagen and the excess collagen is instead secreted at hour 3.(Figure 21)

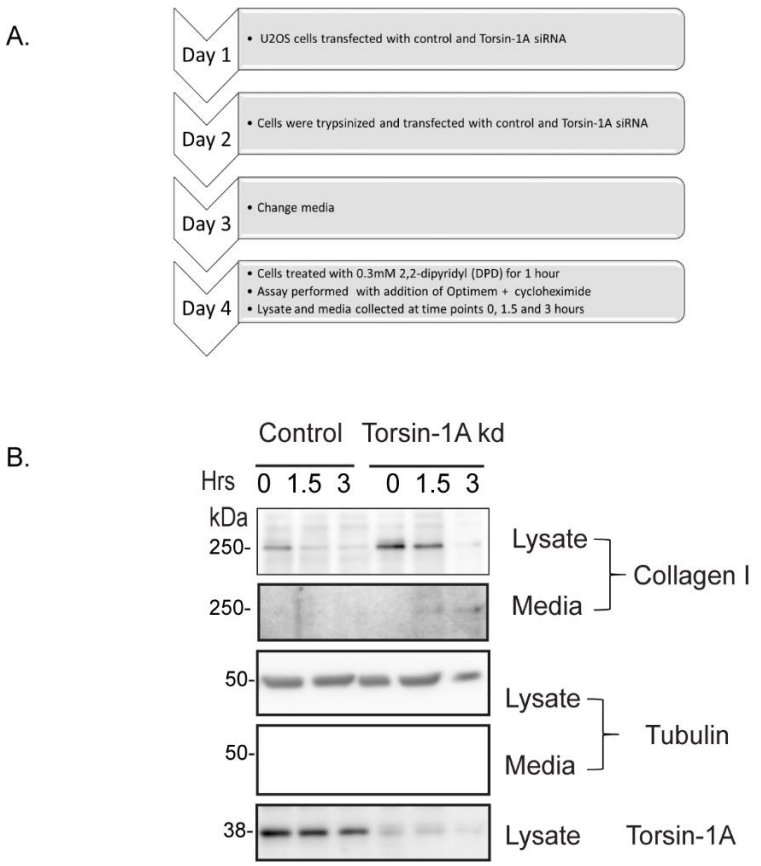

Figure 21. Degradation of unfolded Collagen I is slower in cells lacking Torsin-1A

(A) Schematic of the experiment U2OS cells were knocked down using Control and Torsin-1A siRNAs. 24 hours after the last knockdown cells were treated for an hour with 2,2-dipyridyl. This was then replaced with optimum with 0.5um cycloheximide like in a regular secretion assay and samples were collected over the course of 3 hours.

(B) Lysate and media samples were immunoblotted with antibodies against Collagen I, Tubulin and Torsin.

4.5 Depletion of TANGO1 does not affect trafficking of collagen XVII to the cell surface

TANGO1 had been shown to be required for the export of several collagens, I, II, III, IV,VII, and IX (Saito et al., 2009; Wilson et al., 2011). All of these collagens are soluble proteins which lack the ability to bind COPII coats on its own to initiate export. We wanted to know if TANGO1 was required for the export of transmembrane collagens. The seven transmembrane collagens identified so far are XIII, XVII, XXIII and XXV ectodysplasin A, the class A macrophage scavenger receptors, and the MARCO receptor (Franzke et al., 2003). Among these, Collagen XVII was a good candidate to test for TANGO1 dependency due to its large size and 15 collagenous domains.

4.5.1 Loss of TANGO1 and TANGO1 short has a mild effect on the export of Collagen XVII

TANGO1 short (TANGO1s) on its own had the ability to rescue soluble collagen VII secretion (Maeda et al., 2016), despite not having a cargo binding domain. Since Collagen XVII is membrane spanning and does not require TANGO1 for export, maybe TANGO1 short is sufficient on its own. The CRISPR in HeLa had depleted only TANGO1 (TANGO1L) and still had TANGO1 short (TANGO1s), we decided to use a U2OS cell line where both isoforms of TANGO1 had been removed. We first checked to see if Collagen I was blocked in these cells by immunofluorescence. Parental U2OS and cells where U2OS was depleted of both TANGO1L and TANGO1s were plated on coverslips, fixed, permeabilized and stained with antibodies against Collagen I, TANGO1 and Calreticulin. When visualized by confocal microscopy, as expected we observed that in comparison to the control cells, the cells lacking both forms of TANGO1 showed accumulations of Collagen in the ER colocalizing with Calreticulin (Figure 22).

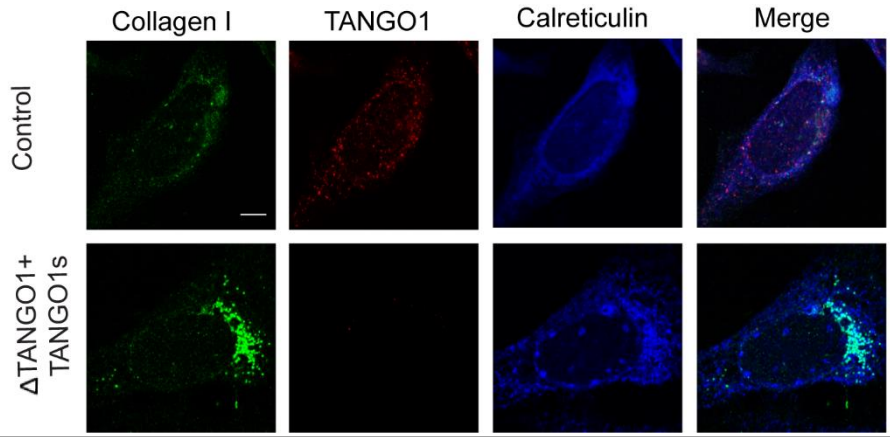

Figure 22. Depletion of TANGO1+TANGO1s shows Collagen retention in the ER
U2OS and U2OSΔTANGO1 were seeded on coverslips cultured for 24 hours.Cells were fixed, permeabilized and stained with antibodies against Collagen I, TANGO1, Calreticulin and visualized by confocal microscopy.Scale bar:9µm

We then went on to test if Collagen XVII was trafficked to the cell surface in this condition. The cell line was transfected with FLAG tagged Collagen XVII with parental U2OS as a control. Cells were fixed post transfection and not permeabilized in order to stain only the cell surface. The cell were stained with Collagen XVII antibody. Upon visualization using confocal microscopy, there were no apparent differences, Collagen XVII showed a cell surface staining in both cases (Figure 23).

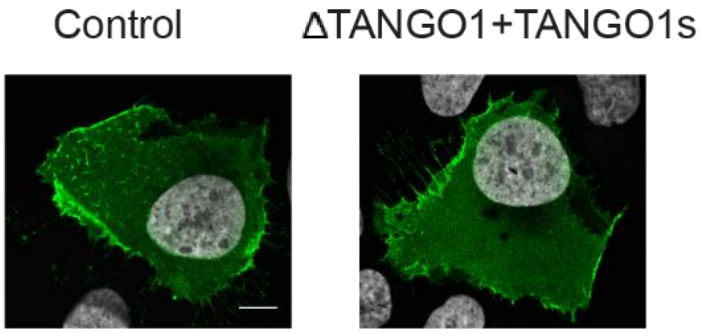

Figure 23. Trafficking of Collagen XVII to the cell surface is unaffected by deletion of TANGO1 and TANGO1-short
U2OS and U2OSΔTANGO1+TANGO1s were transfected with FLAG tagged collagen. Cells were fixed in methanol free 4% PFA to stain only the cell surface with Collagen XVII antibody (in green). Scale bar:9um

Since we did not see any major differences by immunofluorescence, in cell surface expression, even with the deletion of TANGO1 short, we decided to use a more quantitative approach by looking at cell surface biotinylation.

U2OS and ΔTANGO1+TANGO1s U2OS cells were transfected with FLAG-Collagen XVII, 24 hours post transfection, and a cell surface biotinylation assay was performed. Collagen XVII expression is slightly reduced in cells lacking TANGO1+TANGO1s compared to the parental. Looking at the biotinylated cell surface pool, localization of collagen XVII in the cells with no isoforms of TANGO1 showed anywhere between 20-30% reduction (Figure 24).

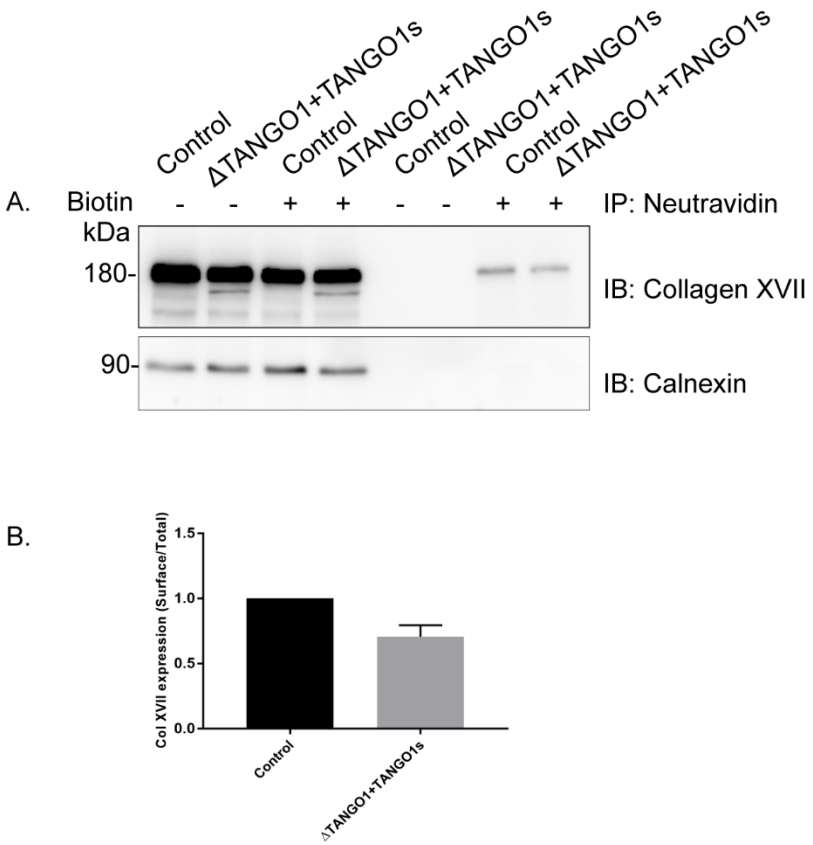

Figure 24. Cell surface biotinylation of Collagen XVII in ΔTANGO1+TANGO1s cells shows a small defect in trafficking

(A) Collagen XVII was expressed in control and cells depleted of TANGO1+TANGO1s, cell surface biotinylated with empty vehicle and biotin reagent, immunoprecipitated to enrich biotinylated pool and immunoblotted using antibodies against collagen XVII and calnexin as a control.

(B) The ratios of surface to total expression in the biotinylated pool of control versus TANGO1 depleted cells was plotted with control expression levels being normalized to 1. Error bars: standard error of the mean (SEM).

4.5.2 Knockdown of the TANGO1 family of proteins has a moderate effect on collagen XVII trafficking

Collagen XVII is primarily expressed in keratinocytes, while we could have used HaCaT cells which are immortalized keratinocytes for our assays. They are harder to transfect and knockdown. To ensure identical expression across experimental samples, we generated a HeLa cell expressing collagen XVII GFP (see materials and methods).

A Fluorescence activated cell sorting (FACS) method to quantitate cell surface expression of collagen XVII was our next approach to test the role of TANGO1, TANGO1s and cTAGE5 all of which are shown to be required for transport of soluble collagens. A CRISPRi/dCas9 mediated knockdown method was used to deplete the following genes GAL4, NSF, TANGO1, TANGO1s and cTAGE5 in HeLa cell line stably expressing Collagen XVII (Gilbert et al., 2013; Horlbeck et al., 2016). 4 days after lentivirus infection for CRISPRi/dCas9-mediated knockdown, cells were incubated with trypsin for 15 min at 37°C to cleave the pool of Collagen XVII expressed at the cell surface. Then, cells were washed with PBS and incubated for 18 hours in complete medium which allowed recovery of Collagen XVII cell surface expression. Cells were harvested after incubation for 10min with 0.5 mM EDTA, washed and resuspended in optiMem containing 5% FCS. Cells were then incubated with an anti-collagen XVII antibody for 1 hour at 4°C followed by PBS wash and incubated with a secondary antibody PE-conjugated for 1h at 4°C.The levels of Collagen XVII Expression, total indicated by levels of GFP and cell surface signal corresponding to the signal of PE was analysed using LSR Fortessa flow cytometer.

The ratio of cell surface expression to the total expression of collagen XVII was quantified. The results show that when compared to the control, the knockdown of NSF (Positive control), lead to a block in trafficking, resulting in only 20 % expression of Collagen XVII at the cell surface. While cTAGE5, TANGO1L and TANGO1s knockdown had almost no effect with cell surface expression levels (100%, 95% and 98% respectively). The only knockdown that had a significant effect was the double knockdown of TANGO1 full length and TANGO1 short

isoforms with 80% cell surface expression, indicating a 20% decrease in transport to cell surface. This data correlates well with the data obtained from the cell surface biotinylation experiments in the U2OS cells (Figure 25).

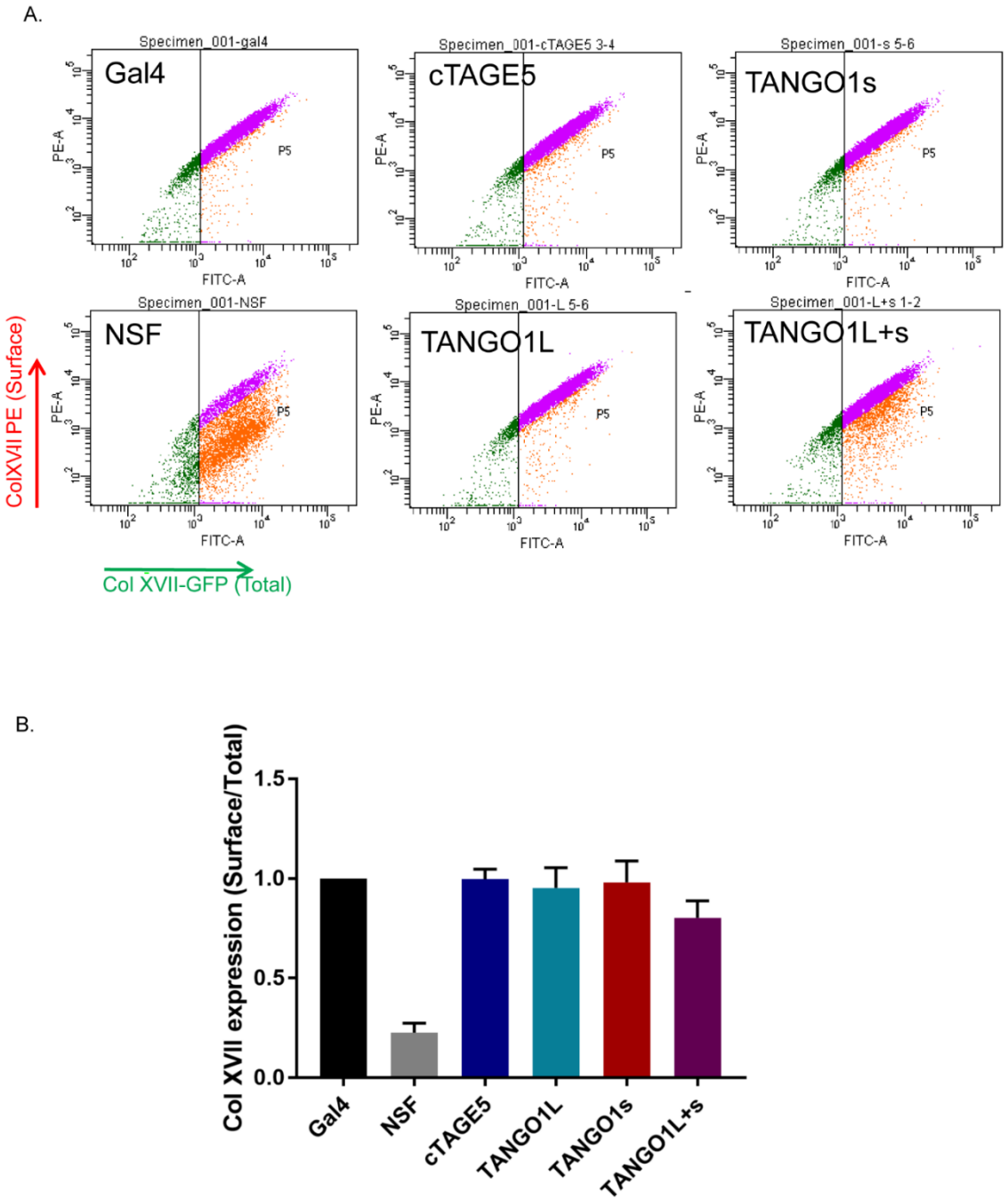

Figure 25. The TANGO1 family of proteins have minimal effect on trafficking of Collagen XVII to the cell surface

. (A) Clonal HeLa cell line stably expressing collagen XVII-GFP was infected with lentivirus to express sgRNA targeting the promoter of the indicated gene. Collagen XVII-GFP expressions (total and surface expression) were analyzed by flow cytometry.

(B) Quantification of collagen XVII expression. For each targeted gene, the ratio corresponding to the signal at the cell surface and the signal of the total expression was calculated. Each condition is normalized compared to cells infected with a sgRNA targeted Gal4 (Control condition). Error bars: standard error of the mean (SEM).

Three different approaches were used to test the role of TANGO1 in the export of collagen XVII and the sum total of results from these diverse methods clearly

indicate that TANGO1 KO has a mild effect on the export showing a 20 percent reduction but only in combination with TANGO1s depletion.

These results indicate a minimal role for TANGO1 in the export of Collagen XVII.

4.5.3 TANGO1 interacts with Collagen XVII through HSP47

Our results from immunofluorescence,cell surface biotinylation and FACS experiments demonstrated nominal requirement of TANGO1 for the trafficking of Collagen XVII.

On the basis of these data, we hypothesized that Collagen XVII probably is not a TANGO1 dependent cargo. To verify this stance, we tested if Collagen XVII was able to interact with TANGO1.

We used two variants of Collagen XVII, the full length wildtype and a mutant which cannot be glycosylated and hence is retained in the endoplasmic reticulum (ref), both tagged with FLAG. The wild type, the glycosylation defective mutant and an empty vector FLAG control were expressed in HEK293T cells.

The cells were then lysed and subject to immunoprecipitation using FLAG beads and eluted with flag peptide. The eluates were immunoblotted against antibodies for TANGO1, HSP47 and FLAG.

Binding with TANGO1 was seen in both wildtype and mutant forms, whereas only wildtype Collagen XVII showed binding with HSP47 (Figure 26). The reduced binding of the mutant to HSP47 could be explained by the fact that HSP47 only binds fully folded triple helices of collagen molecules (Koide et al., 2006). These results suggest that TANGO1 can interact with Collagen XVII, Hsp47 is also part of the complex and could mediate the binding. The collagen XVII mutant, binds TANGO1 and there seems to be no Hsp47 in this complex.This to us was an interesting result.

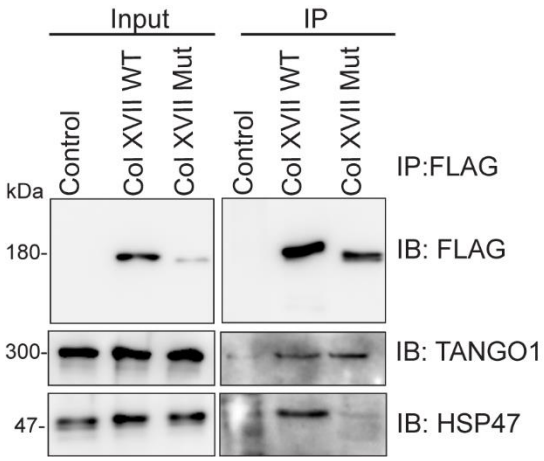

Figure 26. Collagen XVII wild type and mutant binds TANGO1
FLAG tagged Collagen XVII WT and Mutant variants were expressed in HEK293T cells, immunoprecipitated and probed for FLAG, HSP47 and TANGO1.

4.5.4 Collagen XVII binds Sec24C and TANGO1

Collagen XVII binds TANGO1 and yet is exported out of the ER even in its absence, this suggests that Collagen XVII is able to communicate with COPII components on its own. Being a transmembrane collagen with a sizeable cytoplasmic domain, Collagen XVII has the potential to bind COPII component Sec24 to mediate its export out of the ER. Upon analyzing the cytoplasmic domain of Collagen XVII for well-known motifs of COPII binding sites we found several potential sites containing diacidic, dihydrophobic motifs.. Sec24 has several isoforms (A, B, C and D) of which 24C has demonstrated propensity to bind to type II transmembrane proteins (Otsu et al., 2013). We tested if Sec24C could interact with Collagen XVII. FLAG tagged Collagen XVII was expressed in HEK239T cells and 24 hours post transfection, cells were lysed and immunoprecipitated using FLAG agarose beads and eluates were subject to electrophoresis and immunoblotted against FLAG, Sec24C and TANGO1 antibodies. Sec24C and TANGO1 were detected in the sample immunoprecipitated with lysates containing Collagen XVII and not control lysates

(Figure 27). The results clearly demonstrate that in the ER, Collagen XVII forms a complex with both COPII component Sec24C and TANGO1.

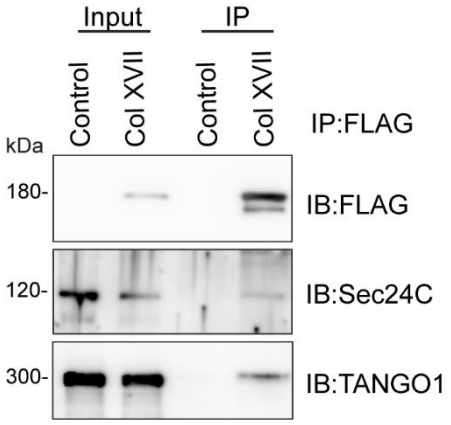

Figure 27. Collagen XVII binds Sec24C and TANGO1
FLAG tagged collagen XVII was expressed in HEK293T cells, immunoprecipitated and probed for FLAG, Sec24C and TANGO1.

4.6 Purification and Structural characterization of TANGO1

Our lab and other groups over the course of the last few years have provided valuable insights into the functions of TANGO1.The next step to deepen this understanding was to obtain structural information about the protein. TANGO1 is a transmembrane protein of 1907 amino acids, therefore not a great candidate for crystallography so Cryo-electron Microscopy (Cryo-EM) was going to be our method of obtaining a structure.

TANGO1 being transmembrane with a complex domain architecture posed several challenges in terms of choosing a method of protein expression. The bacterial expression systems were not an option and so we went with the mammalian expression system

A C-terminally FLAG tagged TANGO1 was transfected into HEK293T cells. 36 hours post transfection cells were washed and lysed in ice cold lysis buffer with rotation for about an hour. The supernatant was clarified by centrifugation at maximum speed. This supernatant was incubated with FLAG agarose beads, washed on a column and elution was done using FLAG peptide.

The lysate, flow through and eluates were subject to electrophoresis on an 8% SDS-PAGE gel and stained with Coomassie.

The purification with FLAG beads yielded protein which was free of most contaminants and we could visualize a strong band at 300kDa (Figure 28).

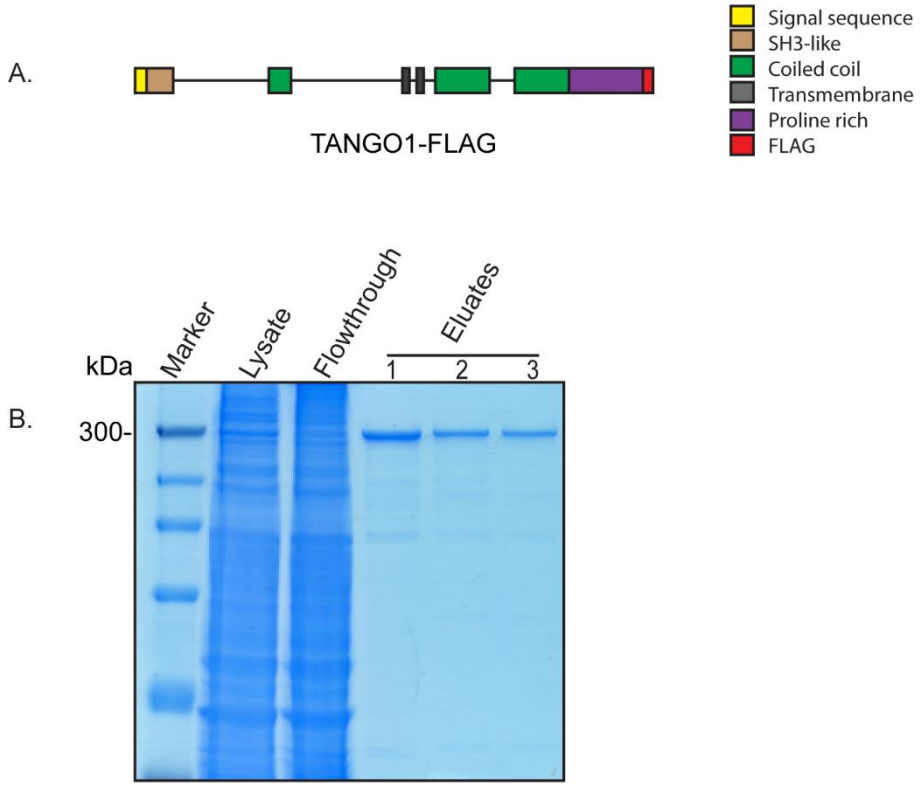

Figure 28: Purification of TANGO1

(A) Schematic of the TANGO1-FLAG construct used for purification

(B) Purified protein separated on an 8% SDS-PAGE and stained with coomassie blue with lanes indicating lysate, flowthrough and eluates.

Although we were able to obtain up to 1mg of protein, even after almost 8 months of optimizing detergents, buffers and purification conditions (data not shown) we had problems of aggregation and were unable to separate the protein from the aggregates despite using gel filtration columns. We did try to look at the structure initially by negative staining, but were unable to see anything of consequence due to issues of aggregation. The next plan was to try to express TANGO1 using the baculovirus expression system, but due to the limited time period of the thesis we decided to not continue these efforts.

Discussion

5. Discussion

5.1 The TANGO1 interactome

Since the discovery of TANGO1's role in export of collagens several interactors have been identified to date. Most of them have been identified through conventional immunoprecipitation approaches using either antibodies or tagged constructs. While this method has yielded valuable data, purifying protein complexes this way only allowed us to identify stable binding partners. The proximity biotinylation method of labelling proteome developed by the Ting lab presented us with an opportunity to identify proteins that function in concert with TANGO1. One of the major advantages was that proteins could be labelled while cells are alive and the protein is in its functional state, enabling us to capture the protein and its interactors in action. This allows us to label even transient interactors and proteins that are binders but are lost during the washing steps of the immunoprecipitation.

Although the data we obtained was spatially specific, with most hits being vicinal to TANGO1, the confidence of data could have been improved by having more replicates and better controls. Additional controls for TANGO1-Lum-APEX2 could have been an ER luminal protein, which would have allowed us to filter out proteins that could have been picked up out of their sheer abundance thus masking better hits. Use of additional cell lines especially those that are specialized for secreting certain cargoes would have enabled us to identify a richer subset of cargoes.

5.1.1 Cytoplasmic Interactors

The TANGO1-Cyto-APEX2 in HeLa was not as efficient as in the HEK293T cells, this could be due to the fact that protein was expressed in low levels and the biotinylation therefore was not as efficient.

TANGO1-Cyto-Apex2 identified several of the known and expected interactors at the ERES such as Sec23A, Sec24C, Sec31A, and TFG which clearly demonstrated enrichment of the right set of proteins. Some highly enriched interactors which we did not have the opportunity to test in our pathway, but could

potentially shed more light on the function of the cytoplasmic domain are as follows.

Myosin-9, this is interesting because there has been a study where they show that association of Myosin 9 /LARP6, is required for the synthesis and secretion of heterotrimeric collagen I (Cai et al., 2010) .How it fits into our pathway remains to be seen.

Annexin A11, it is one of the first annexins to be shown to have a role in the early secretory pathway. AnxA11 along with ALG-2 has been shown to be required for stabilizing Sec31A at ERES. Knockdown of Annexin A11 has been shown to accelerate export of tsO45 VSV-G from the ER but there have been conflicting reports stating otherwise (Shibata et al., 2015). Whether Annexin A11 interacts with the Cytoplasmic domain of TANGO1, and regulates the export of Collagen remains unknown and is worth exploring.

Among the cytoplasmic hits, we chose Sec23IP, an ERES modulator, to test its role in collagen secretion, we knocked down Sec23IP in U2OS cells and we looked for defects in collagen export and change in TANGO1 organization. As expected knockdown of Sec23IP lead to a disrupted ER exit site organization,.Sec23IP lacking cells showed a 2 fold increase in ER pool of collagen. Depletion of Sec23IP has previously been shown to affect transport from the ER (Ong et al., 2010), our data shows an block in collagen secretion, what remains to be examined is if this is an indirect effect of ERES disruption or if SEC23IP has a mode direct role by contributing to the addition of membranes at ERES.

Another interesting piece of information that adds credence to the possibility that Sec23IP is involved in the collagen export pathway is the discovery of a homozygous mutation in humans leading to a truncated form of Sec23IP resulting in osseous syndactyly, craniofacial and brain malformations (Reuter et al., 2017). Cranio-lenticulo-sutural dysplasia (CLSD), which is characterized by skeletal defects and craniofacial deformations is caused by a mutation in Sec23A (Boyadjiev et al., 2011, 2006) and the fibroblasts from these proteins had enlarged ER due to accumulation of Collagens.

Similar to our vitro data, we would predict that the Fibroblasts from patients with the homozygous mutations in Sec23IP will have massive accumulation of collagen in the ER.

5.1.2 Luminal interactors of TANGO1

TANGO1 is required for the secretion of a number of collagens including I, II, III, IV, VII, and IX (Saito et al., 2009; Wilson et al., 2011) and its luminal SH3 domain binds collagens through its interaction with the collagen specific chaperone Hsp47 (Ishikawa et al., 2016; Saito et al., 2009). The SH3 domain is followed by a large unstructured region and a coiled coil domain, the function of these domains and their binding partners remain undiscovered. To address this, we tagged the luminal domain with APEX2 to label proteins that might interact with not just the SH3 domain but also the rest of the luminal region.

The interactors that we found can be broadly categorized into Secretory cargoes, Enzymes, Chaperones and foldases, and proteins with diverse functions. The secretory cargoes were predominantly ECM proteins: Collagens, Laminins, Peroxidasin.

While requirement of TANGO1 for secretion of most collagens has been well documented but its role in Laminin export remains elusive. TANGO1 depletion in glial cells of the larval brain blocks the secretion of Laminin β and γ subunits (Petley-Ragan et al., 2016). In Drosophila fat body cells, the loss of collagen also leads to block of laminins in the ER (Maeda et al., 2016). This prevents us from understanding if Laminin is a direct cargo of TANGO1 or if its ER retention is a fallout of collagen block.

Laminin assembly and secretion is a complicated process, it is known in case of Laminin-1(α1β1γ1) that while the α chain could be secreted as a monomer, the β and γ chains were secreted only in the fully assembled form (Yurchenco et al., 1997). Testing the dependency on TANGO1 in such cases becomes technically challenging.

Little is known about ECM protein, peroxidasin, it could be a potential cargo of TANGO1. It is a large protein 1,479 amino acids long, with a complex domain architecture one leucine-rich repeat domain (LRR), four immunoglobulin-like motifs (Ig), a peroxidase domain (POX), and a C-terminal von Willebrand factor C domain (VWC) (Soudi et al., 2015). It is secreted by fibroblasts in the lung and skin. After TGF-β induced differentiation, myofibroblasts secrete peroxidasin which form fibrils in the ECM. Peroxidasin has also been shown to be highly expressed in mouse model of kidney fibrosis (Péterfi et al., 2009). Understanding how this protein is secreted could help us address issues of fibrosis which is caused by excess deposition of ECM.

We categorized Nodal modulator 1 and 2 (NOMO1/2) two very similar transmembrane proteins in the ER under the group of proteins with diverse functions. NOMOs are signaling factors belonging to the transforming growth factor-β (TGF-β) superfamily and are essential for vertebrate development. They are highly conserved proteins and function to antagonize Nodal signaling during mesendodermal patterning in zebrafish development (Haffner et al., 2004) .Altered nodal signaling, during gastrulation leads to abnormal movement of cells and developmental defects. ECM proteins play a crucial role during development, further studies are required to dissect if NOMO proteins regulate collagen secretion.

Thioredoxin domain containing protein 5 (TXNDC5), is one protein that was consistently identified as an interactor with both the cytoplasmic and luminal constructs in HeLa and HEK293T. This protein has the potential to be a stable binding partner. Also known as Endoplasmic reticulum resident protein 46 (ERp46), it belongs to the PDI family of proteins and is an oxidoreductase. It is not as well characterized as other ER chaperones, is known to associate with EDEM3 via its redox-active sites and activates its mannose-trimming activity. ER mannosidase I and ER degradation–enhancing α-mannosidase–like proteins(EDEMs) trim mannose from N-glycans is important for disposal of misfolded glycoproteins (Yu et al., 2018). ERp46 knockdown in β-cells led to a significant reduction in the insulin content, without affecting mRNA levels. It also resulted in ER stress (Alberti et al., 2009). Recently, it has been shown that

TXNDC5, is highly expressed in cardiac fibroblasts and promotes cardiac fibrosis by enhancing ECM production (Shih et al., 2018). TXNDC5 has also been shown to interact with Collagen I (DiChiara et al., 2016). All of the evidence clearly point to a role for Thioredoxin domain containing protein 5 in collagen folding and export, understanding how it functions in concert with TANGO1 could lead us to interesting breakthroughs.

5.1.3 TANGO1 luminal interactome shares machinery with Collagen Proteostasis proteins

The luminal interactome of TANGO1 had several proteins which are chaperones, enzymes, folding complexes. This isn't really surprising because the ER is home to protein folding machinery but what was interesting is that many of these proteins function uniquely in the folding and post translational machinery of collagens. When we compared our list of proteins with the collagen proteostasis machinery identified by DiChiara and colleagues we found several proteins in common (DiChiara et al., 2016).

These proteins are all required for the proper folding, modification of collagens. TANGO1 binds collagens via Hsp47, which binds to fully folded triple helical collagens, this suggests a mechanism to ensure selection of well modified folded collagens for export. This folding machinery being in the vicinity of TANGO1, some of them possibly direct interactors could suggest a larger role for TANGO1 in collagen folding. This presents a possibility that the luminal half of TANGO1 is not just required for the recruitment of the cargo, (in this case largely collagen) but also serves as an area for clustering collagen folding and modification machinery. TANGO1 could potentially function in the folding and quality control of collagen being exported from the ER choosing to send it for secretion versus degradation depending on the folding state.

The reason that collagens are retained in the ER in the absence of TANGO1 might not only be because of the lack of export machinery recruitment but also that they are improperly folded.

It is important to evaluate the state of collagen folding in the absence of TANGO1. Data from the mouse knockout of TANGO1 provides some evidence that in the absence of TANGO1 the collagens are abnormally glycosylated and form aggregates which also disrupts ECM formation once it is out of the cell (Wilson et al., 2011).

Another part that needs to be thoroughly analyzed if levels of TANGO1 is able to modulate the levels of rest of the collagen export machinery both on the luminal side and the cytoplasmic side. Data from my experiments using Collagen VII not included in this thesis, shows an upregulation of TANGO1 at mRNA and protein levels when collagen expression is upregulated.

Is TANGO1 the receptor for secretion and degradation? What happens to mutated collagens, those that are misfolded, hyper hydroxylated or glycosylated? Do they still bind TANGO1? My initial studies with Collagen VII mutants point to the direction of TANGO1 binding to both fully folded and mutant collagens (data not shown).

5.2 Torsin-1A, has a potential role in degradation of unfolded collagens

Torsin-1A, is an AAA ATPase located in the ER lumen. Its mutation results in the movement disorder Torsion Dystonia (Ozelius et al., 1997). Torsin is an unconventional ATPase, on its own is inactive and the ATPase activity is(Jokhi et al., 2013) induced by its partners in the ER and nuclear envelope, LULL1 and LAP1 respectively. Torsin has been implicated in a wide variety of processes including protein quality control, trafficking, nuclear envelope vesiculation and nuclear ribonucleoprotein (RNP) export (Jokhi et al., 2013). When we found Torsin-1A, LULL1 and LAP1 as potential interactors of TANGO1, it prompted us to investigate the role of this protein in our pathway. We hypothesized that it could be involved either in the disassembly of the TANGO1/Hsp47-Collagen complex or in the quality control of Collagen.

We showed that TANGO1 interacts with Torsin-1A, both the ATP bound form and the ATP deficient seemed to bind with a higher affinity to TANGO1 when compared to the wild type. Torsin-1A is in the ER lumen and biotinylated by the luminal TANGO1 APEX2, this suggests that the binding site on TANGO1 is in the luminal domains. Torsin binds TANGO1, do they function in the same pathway was our next question. We knocked down Torsin-1A in collagen secreting cells and asked if Collagen export was inhibited. Our immunofluorescence data showed a buildup of Collagen I in cells depleted of Torsin-1A. A secretion assay was performed to test if this build up was due to a defect in secretion. Concomitant with our immunofluorescence data, results of the secretion assay showed an increase in intracellular collagen at time 0, but there was no major defect in secretion as we were able to detect collagen in the media at hours 1 and 2.

The striking phenotype of the Torsin knockdown was the higher levels of intracellular collagen, without an apparent defect in secretion. This could either be due to increased synthesis of Collagen I or a decrease in degradation of collagens. Our RT-PCR data showed no change in the mRNA levels of ColA1.A(data not shown) reduction in degradation was the other possibility that needed to be tested. We know that a steady state around 15% of newly synthesized collagen I is degraded. It is known that collagens which are unfolded can be degraded by the ERAD pathway while triple helical collagens are disposed via the lysosomal pathway(Ishida et al., 2009) Torsin has been known to participate in ERAD mediated degradation of a few cargoes (Nery et al., 2011). Loss of Torsin has been shown to elevate the susceptibility of the cells to ER stress(Chen et al., 2010). To test if Torsin had a role in clearing of unfolded collagen from the ER, we decided to look at how fast unfolded collagens were cleared from the ER in the presence and absence of Torsin. The drug 2,2-dipyridyl an iron chelator that prevents functioning of the hydroxylases and therefore hydroxylation of collagens therefore resulting in its unfolding was used. In cells treated with 2,2 dipyridyl and depleted of Torsin there was a delay in degradation of unfolded collagen I. There was also some (Has et al., 2018)secretion at a later time point which I speculate could be a mechanism of

clearing out the overloaded ER. This result suggests a role for Torsin-1A in aiding the clearance of unfolded collagen I from the ER. We do not know if this activity of Torsin is dependent on the presence of TANGO1. Another key aspect to be tested is the role of the ATPase activity regulator LULL1. This study has identified Torsin-1A as a new interactor of TANGO1 in the lumen which has the potential to function not in the secretion of collagen I but in its degradation.

5.3 Collagen XVII requires both COPII component Sec24C and TANGO1 for its export

Collagen XVII, the largest of the transmembrane collagens, is a major component of the hemidesmosome complex. Mutations in Collagen XVII have been a leading cause of the skin fragility Epidermolysis Bullosa (Has et al., 2018). While a lot is known about the structural and functional aspects of Collagen XVII. Very little is known about how Collagen XVII is trafficked to the cell surface. Our aim was to test if Collagen XVII was transported in a TANGO1 dependent manner.

A TANGO1 knockout cell line was used to test the export of Collagen XVII to the cell surface, immunofluorescence studies showed no obvious defects in the trafficking to the cell surface. An isoform of TANGO1, lacking the cargo binding luminal domain TANGO1-short (TANGO1s) can facilitate collagen VII export (Maeda et al., 2016). Our Cell line lacking TANGO1, still contained TANGO1s and we hypothesized that this could possibly be why we didn't see a defect. A cell line depleted of both TANGO1 and TANGO1s also did not show an effect in trafficking of Collagen XVII to the cell surface by immunofluorescence. While a cell surface biotinylation approach in the same cell line showed a slight defect, around 20% in trafficking to the cell surface. The data from our FACS experiment where we tested TANGO1, cTAGE5, TANGO1s, TANGO1L+ TANGO1s, controls GAL4 and NSF, also showed minimal effect of the TANGO1 family proteins on Collagen XVII trafficking. The lack of a defect in export in the absence of the

TANGO1 family of proteins led us to investigate if Collagen XVII interacted with TANGO1 at all.

An immunoprecipitation experiment revealed that, both wildtype and mutant collagen XVII bind to TANGO1. This binding seems to be mediated by the collagen specific chaperone HSP47. There was reduced binding of HSP47 in the mutant pool, which is ideal as HSP47 is known to bind to triple helical collagens. But if the binding to TANGO1 is through Hsp47 it does not explain how the mutant pool is still bound to TANGO1. Could there be a different chaperone connecting TANGO1 to unfolded chains of collagen? This shift in binding could potentiate the switch between the secretion and degradation of collagens mediated by TANGO1.

Collagen XVII has a cytoplasmic domain, this could in theory bind COPII coats. and explain how they were being exported in the absence of TANGO1. Based on previously published literature, we looked for the motifs and sequences in the cytoplasmic domain that have been known to interact with COPII components. We found two di acidic motifs and several dihydrophobic motifs which could bind COPII. Sec24C is the isoform that has been shown to bind type II transmembrane proteins, so we tested if Collagen XVII could bind this protein. Immunoprecipitation of FLAG tagged Collagen XVII, revealed its binding to Sec24C and TANGO1. Collagen XVII, forms a complex with COP II component Sec24C and TANGO1. This combination could be essential for its export out of the ER.

The dual binding of Collagen XVII by both Sec24C and TANGO1 could be a mechanism to ensure selection of fully folded triple helical collagen. While Sec24 C may be involved in vesicle generation, TANGO1 binding via Hsp47 ensures selection of properly folded cargo.

A way to test this dual requirement hypothesis would be to test if a luminal only collagen XVII was entirely TANGO1 dependent for its export. Previous work has shown that it is possible to generate a fully folded luminal collagen XVII that is secreted out of the cells (Balding et al., 1997) . This clearly indicates the existence

of a receptor that can bind and facilitate the export of this luminal collagen. This receptor could be TANGO1 based on our immunoprecipitation data.

There have been reports of transmembrane protein, with the innate ability to bind COPII components and initiate their export from the ER also being cargoes of transmembrane receptors (Pagant et al., 2015). This mechanism allows for more efficient export of cargoes. For a protein as complex as collagens, which requires a gamut of the ER protein folding machinery, it becomes even more important that a mechanism like this exists. If future experimental results are consistent with our hypothesis, we can demonstrate that TANGO1 is not just a receptor that connects the export machinery to large cargo, but one that ensures that properly folded triple helical Collagens are allowed out of the ER.

Conclusions

6. Conclusions

In our quest to map the TANGO1 interactome, we found that on the cytoplasmic side, the potential partners are primarily components of ERES export machinery which was expected given what we already know so far. One such interactor, Sec23IP which is a modulator of ERES, in our study is required for the export of Collagen I from the ER.On the luminal side, we encountered several chaperones, known cargoes like collagens, potential cargoes like laminin, peroxidasin but the surprising find were the enzymes involved in modification and folding of collagens. We know that TANGO1 collects Collagens through HSP47, which is known to bind fully folded triple helical collagens, our finding of the collagen folding and modification machinery points to the possibility that TANGO1 could be involved in not just the collection and export but also the folding of collagens.

.One intriguing candidate Torsin-1A, an unconventional AAA ATPase was identified as an interactor of TANGO1 on the luminal side. We show that it binds TANGO1, and on Torsin-1A depletion, has no major effect on collagen export. There was an increase in the intracellular pool of Collagen in these cells, indicating reduced degradation. We found that cells lacking Torsin had a delay in degrading unfolded Collagen I.

We found that unlike other collagens the requirement for TANGO1 and its associates, TANGO1-short and cTAGE5 in the export of Collagen XVII is minimal at best. We have identified that Collagen XVII binds the COPII component, Sec24C which is likely enabling its export. We also demonstrate that TANGO1 can bind Collagen XVII. We are able to show that both wild type and mutant Collagen XVII, bind to TANGO1, with the wildtype co-precipitating HSP47 and the mutant binding negligible levels. This potentially is a very interesting result, and lays the groundwork into future studies where we can explore the possibility.TANGO1 could be the final checkpoint at the ERES for the quality control of collagens. TANGO1 could bind both folded and unfolded/ misfolded collagens, either directly or using a different set of chaperones as adaptors that and directs the collagens either for secretion or degradation, depending on its folding status.

7. References

Alberti, A., Karamessinis, P., Peroulis, M., Kypreou, K., Kavvadas, P., Pagakis, S., Politis, P.K., and Charonis, A. (2009). ERp46 is reduced by high glucose and regulates insulin content in pancreatic β-cells. Am. J. Physiol.-Endocrinol. Metab. *297*, E812–E821.

Areida, S.K., Reinhardt, D.P., Müller, P.K., Fietzek, P.P., Köwitz, J., Marinkovich, M.P., and Notbohm, H. (2001). Properties of the Collagen Type XVII Ectodomain: EVIDENCE FOR N- TO C-TERMINAL TRIPLE HELIX FOLDING. J. Biol. Chem. *276*, 1594–1601.

Aridor, M. (2018). COPII gets in shape: Lessons derived from morphological aspects of early secretion. Traffic.

Arimitsu, N., Kogure, T., Baba, T., Nakao, K., Hamamoto, H., Sekimizu, K., Yamamoto, A., Nakanishi, H., Taguchi, R., Tagaya, M., et al. (2011). p125/Sec23-interacting protein (Sec23ip) is required for spermiogenesis. FEBS Lett. *585*, 2171–2176.

Balding, S.D., Diaz, L.A., and Giudice, G.J. (1997). A Recombinant Form of the Human BP180 Ectodomain Forms a Collagen-like Homotrimeric Complex. Biochemistry *36*, 8821–8830.

Bard, F., Casano, L., Mallabiabarrena, A., Wallace, E., Saito, K., Kitayama, H., Guizzunti, G., Hu, Y., Wendler, F., DasGupta, R., et al. (2006). Functional genomics reveals genes involved in protein secretion and Golgi organization. Nature *439*, 604.

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

Barlowe, C.K., and Miller, E.A. (2013). Secretory Protein Biogenesis and Traffic in the Early Secretory Pathway. Genetics *193*, 383–410.

Bienkowski, R.S., and Gotkin, M.G. (1995). Control of Collagen Deposition in Mammalian Lung. Proc. Soc. Exp. Biol. Med. *209*, 118–140.

Bonfanti, L., and Mironov, A.A. Procollagen Traverses the Golgi Stack without Leaving the Lumen of Cisternae: Evidence for Cisternal Maturation. 11.

Borradori, L., and Sonnenberg, A. (1999). Structure and Function of Hemidesmosomes: More Than Simple Adhesion Complexes. J. Invest. Dermatol. *112*, 411–418.

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

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

Bruckner, P., Eikenberry, E.F., and Prockop, D.J. (1981). Formation of the Triple Helix of Type I Procollagen in cellulo. A Kinetic Model Based on cis-trans Isomerization of Peptide Bonds. Eur. J. Biochem. *118*, 607–613.

Cai, L., Fritz, D., Stefanovic, L., and Stefanovic, B. (2010). Nonmuscle Myosin-Dependent Synthesis of Type I Collagen. J. Mol. Biol. *401*, 564–578.

Chen, P., Burdette, A.J., Porter, J.C., Ricketts, J.C., Fox, S.A., Nery, F.C., Hewett, J.W., Berkowitz, L.A., Breakefield, X.O., Caldwell, K.A., et al. (2010). The early-onset torsion dystonia-associated protein, torsinA, is a homeostatic regulator of endoplasmic reticulum stress response. Hum. Mol. Genet. *19*, 3502–3515.

Comtesse, N., Niedermayer, I., Glass, B., Heckel, D., Maldener, E., Nastainczyk, W., Feiden, W., and Meese, E. MGEA6 is tumor-speci®c overexpressed and frequently recognized by patient-serum antibodies. 9.

DiChiara, A.S., Taylor, R.J., Wong, M.Y., Doan, N.-D., Rosario, A.M.D., and Shoulders, M.D. (2016). Mapping and Exploring the Collagen-I Proteostasis Network. ACS Chem. Biol. *11*, 1408–1421.

Fairlky, J.A., HEINTZ, P.W., NEUBURG, M., DIAZ, L.A., and GIUDICE, G.J. (1995). Expression pattern of the bullous pemphigoid-180 antigen in normal and neoplastic epithelia. Br. J. Dermatol. *133*, 385–391.

Fan, J., Wang, Y., Liu, L., Zhang, H., Zhang, F., Shi, L., Yu, M., Gao, F., and Xu, Z. (2017). *cTAGE5* deletion in pancreatic β cells impairs proinsulin trafficking and insulin biogenesis in mice. J. Cell Biol. *216*, 4153–4164.

Fitzgerald, J., Lamande, S.R., and Bateman, J.F. Proteasomal Degradation of Unassembled Mutant Type I Collagen Pro-␣1(I) Chains. 8.

Franzke, C.-W., Tasanen, K., Schumann, H., and Bruckner-Tuderman, L. (2003). Collagenous transmembrane proteins: collagen XVII as a prototype. Matrix Biol. *22*, 299–309.

Franzke, C.-W., Bruckner-Tuderman, L., and Blobel, C.P. (2009). Shedding of Collagen XVII/BP180 in Skin Depends on Both ADAM10 and ADAM9. J. Biol. Chem. *284*, 23386–23396.

Gelse, K. (2003). Collagens—structure, function, and biosynthesis. Adv. Drug Deliv. Rev. *55*, 1531–1546.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell *154*, 442–451.

Giudice, G.J., Emery, D.J., Zelickson, B.D., Anhalt, G.J., Liu, Z., and Diaz, L.A. (1993). Bullous pemphigoid and herpes gestationis autoantibodies recognize a common non-collagenous site on the BP180 ectodomain. J. Immunol. *151*, 5742.

Goodchild, R.E., Kim, C.E., and Dauer, W.T. (2005). Loss of the Dystonia-Associated Protein TorsinA Selectively Disrupts the Neuronal Nuclear Envelope. Neuron *48*, 923–932.

Gorres, K.L., and Raines, R.T. (2010). Prolyl 4-hydroxylase. Crit. Rev. Biochem. Mol. Biol. *45*, 106–124.

Haffner, C., Frauli, M., Topp, S., Irmler, M., Hofmann, K., Regula, J.T., Bally-Cuif, L., and Haass, C. (2004). Nicalin and its binding partner Nomo are novel Nodal signaling antagonists. EMBO J. *23*, 3041–3050.

Hammond, C., Braakman, I., and Helenius, A. (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. Proc. Natl. Acad. Sci. *91*, 913–917.

Hanna et al. (2017). Correction for Hanna et al., TFG facilitates outer coat disassembly on COPII transport carriers to promote tethering and fusion with ER–Golgi intermediate compartments. Proc. Natl. Acad. Sci. *114*, E8547–E8548.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. *6*, 519–529.

Has, C., Nyström, A., Saeidian, A.H., Bruckner-Tuderman, L., and Uitto, J. (2018). Epidermolysis bullosa: Molecular pathology of connective tissue components in the cutaneous basement membrane zone. Matrix Biol.

Heckel, D., Brass, N., Fischer, U., Blin, N., Steudel, I., Tureci, O., Fackler, O., Zang, K.D., and Meese, E. (1997). cDNA Cloning and Chromosomal mapping of a Predicted Coiled-Coil Proline-Rich Protein Immunogenic in Meningioma Patients. Hum. Mol. Genet. *6*, 2031–2041.

Hewett, J.W., Tannous, B., Niland, B.P., Nery, F.C., Zeng, J., Li, Y., and Breakefield, X.O. (2007). Mutant torsinA interferes with protein processing through the secretory pathway in DYT1 dystonia cells. Proc. Natl. Acad. Sci. *104*, 7271.

Hirako, Y., Usukura, J., Uematsu, J., Hashimoto, T., Kitajima, Y., and Owaribe, K. (1998). Cleavage of BP180, a 180-kDa Bullous Pemphigoid Antigen, Yields a 120-kDa Collagenous Extracellular Polypeptide. J. Biol. Chem. *273*, 9711–9717.

Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., et al. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. ELife *5*.

Hung, V., Udeshi, N.D., Lam, S.S., Loh, K.H., Cox, K.J., Pedram, K., Carr, S.A., and Ting, A.Y. (2016). Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat. Protoc. *11*, 456–475.

Hurtley, S.M. (1989). Interactions of misfolded influenza virus hemagglutinin with binding protein (BiP). J. Cell Biol. *108*, 2117–2126.

Hynes, R.O., and Naba, A. (2012). Overview of the Matrisome--An Inventory of Extracellular Matrix Constituents and Functions. Cold Spring Harb. Perspect. Biol. *4*, a004903–a004903.

Ishida, Y., Yamamoto, A., Kitamura, A., Lamandé, S.R., Yoshimori, T., Bateman, J.F., Kubota, H., and Nagata, K. (2009). Autophagic Elimination of Misfolded Procollagen Aggregates in the Endoplasmic Reticulum as a Means of Cell Protection. Mol. Biol. Cell *20*, 2744–2754.

Ishikawa, Y., and Bächinger, H.P. (2013). A molecular ensemble in the rER for procollagen maturation. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1833*, 2479–2491.

Ishikawa, Y., Ito, S., Nagata, K., Sakai, L.Y., and Bächinger, H.P. (2016). Intracellular mechanisms of molecular recognition and sorting for transport of large extracellular matrix molecules. Proc. Natl. Acad. Sci. *113*, E6036–E6044.

Jokhi, V., Ashley, J., Nunnari, J., Noma, A., Ito, N., Wakabayashi-Ito, N., Moore, M.J., and Budnik, V. (2013). Torsin Mediates Primary Envelopment of Large Ribonucleoprotein Granules at the Nuclear Envelope. Cell Rep. *3*, 988–995.

Jungwirth, M., Dear, M.L., Brown, P., Holbrook, K., and Goodchild, R. (2010). Relative tissue expression of homologous torsinB correlates with the neuronal specific importance of DYT1 dystonia-associated torsinA. Hum. Mol. Genet. *19*, 888–900.

Kadler, K.E., Baldock, C., Bella, J., and Boot-Handford, R.P. (2007). Collagens at a glance. J. Cell Sci. *120*, 1955–1958.

Klinkenberg, D., Long, K.R., Shome, K., Watkins, S.C., and Aridor, M. (2014). A cascade of ER exit site assembly that is regulated by p125A and lipid signals. J. Cell Sci. *127*, 1765–1778.

Koide, T., Nishikawa, Y., Asada, S., Yamazaki, C.M., Takahara, Y., Homma, D.L., Otaka, A., Ohtani, K., Wakamiya, N., Nagata, K., et al. (2006). Specific Recognition of the Collagen Triple Helix by Chaperone HSP47: II. THE HSP47-BINDING STRUCTURAL MOTIF IN COLLAGENS AND RELATED PROTEINS. J. Biol. Chem. *281*, 11177–11185.

Koreishi, M., Yu, S., Oda, M., Honjo, Y., and Satoh, A. (2013). CK2 Phosphorylates Sec31 and Regulates ER-To-Golgi Trafficking. PLoS ONE *8*, e54382.

Ma, W., and Goldberg, J. (2016). TANGO1/cTAGE5 receptor as a polyvalent template for assembly of large COPII coats. Proc. Natl. Acad. Sci. *113*, 10061–10066.

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

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

Makareeva, E., Mertz, E.L., Kuznetsova, N.V., Sutter, M.B., DeRidder, A.M., Cabral, W.A., Barnes, A.M., McBride, D.J., Marini, J.C., and Leikin, S. (2008). Structural Heterogeneity of Type I Collagen Triple Helix and Its Role in Osteogenesis Imperfecta. J. Biol. Chem. *283*, 4787–4798.

Marini, J.C., Forlino, A., Bächinger, H.P., Bishop, N.J., Byers, P.H., Paepe, A.D., Fassier, F., Fratzl-Zelman, N., Kozloff, K.M., Krakow, D., et al. (2017). Osteogenesis imperfecta. Nat. Rev. Dis. Primer *3*, 17052.

Marzec, M., Eletto, D., and Argon, Y. (2012). GRP94: An HSP90-like protein specialized for protein folding and quality control in the endoplasmic reticulum. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1823*, 774–787.

McGary, K.L., Park, T.J., Woods, J.O., Cha, H.J., Wallingford, J.B., and Marcotte, E.M. (2010). Systematic discovery of nonobvious human disease models through orthologous phenotypes. Proc. Natl. Acad. Sci. *107*, 6544–6549.

McGrath, J.A., Gatalica, B., Christiano, A.M., si, K., Owaribe, K., McMillan, J.R., Eady, R.A.J., and Uitto, J. (1995). Mutations in the 180–kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. Nat. Genet. *11*, 83.

Mironov, A.A., Mironov, A.A., Beznoussenko, G.V., Trucco, A., Lupetti, P., Smith, J.D., Geerts, W.J.., Koster, A.J., Burger, K.N.., Martone, M.E., et al. (2003). ER-to-Golgi Carriers Arise through Direct En Bloc Protrusion and Multistage Maturation of Specialized ER Exit Domains. Dev. Cell *5*, 583–594.

Morris, P., Fessler, I., and Fesslerj, J.H. and Secretion in Embryonic Chick Bone. 9.

Mouw, J.K., Ou, G., and Weaver, V.M. (2014). Extracellular matrix assembly: a multiscale deconstruction. Nat. Rev. Mol. Cell Biol. *15*, 771–785.

Myllyharju, J. (2004). Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet. *20*, 33–43.

Myllyharju, J. (2005). Intracellular Post-Translational Modifications of Collagens. In Collagen: Primer in Structure, Processing and Assembly, J. Brinckmann, H. Notbohm, and P.K. Müller, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 115–147.

Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hosokawa, N., and Nagata, K. (2000). Embryonic Lethality of Molecular Chaperone Hsp47 Knockout Mice Is Associated with Defects in Collagen Biosynthesis. J. Cell Biol. *150*, 1499.

Nagata, K. (2003). HSP47 as a collagen-specific molecular chaperone: function and expression in normal mouse development. Heat Shock Proteins Dev. Aging Evol. *14*, 275–282.

Nery, F.C., Armata, I.A., Farley, J.E., Cho, J.A., Yaqub, U., Chen, P., da Hora, C.C., Wang, Q., Tagaya, M., Klein, C., et al. (2011). TorsinA participates in endoplasmic reticulum-associated degradation. Nat. Commun. *2*, 393.

Ong, Y.S., Tang, B.L., Loo, L.S., and Hong, W. (2010). p125A exists as part of the mammalian Sec13/Sec31 COPII subcomplex to facilitate ER-Golgi transport. J. Cell Biol. *190*, 331–345.

Ozelius, L.J., Hewett, J.W., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C., de Leon, D., Brin, M.F., Raymond, D., Corey, D.P., et al. (1997). The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat. Genet. *17*, 40.

Pagant, S., Wu, A., Edwards, S., Diehl, F., and Miller, E.A. (2015). Sec24 Is a Coincidence Detector that Simultaneously Binds Two Signals to Drive ER Export. Curr. Biol. *25*, 403–412.

Péterfi, Z., Donkó, Á., Orient, A., Sum, A., Prókai, Á., Molnár, B., Veréb, Z., Rajnavölgyi, É., Kovács, K.J., Müller, V., et al. (2009). Peroxidasin Is Secreted and Incorporated into the Extracellular Matrix of Myofibroblasts and Fibrotic Kidney. Am. J. Pathol. *175*, 725–735.

Petley-Ragan, L.M., Ardiel, E.L., Rankin, C.H., and Auld, V.J. (2016). Accumulation of Laminin Monomers in Drosophila Glia Leads to Glial Endoplasmic Reticulum Stress and Disrupted Larval Locomotion. J. Neurosci. *36*, 1151–1164.

Pincus, D., Chevalier, M.W., Aragón, T., van Anken, E., Vidal, S.E., El-Samad, H., and Walter, P. (2010). BiP Binding to the ER-Stress Sensor Ire1 Tunes the Homeostatic Behavior of the Unfolded Protein Response. PLOS Biol. *8*, e1000415.

Pitman, J.L., Bonnet, D.J., Curtiss, L.K., and Gekakis, N. (2011). Reduced cholesterol and triglycerides in mice with a mutation in Mia2, a liver protein that localizes to ER exit sites. J. Lipid Res. *52*, 1775–1786.

Prockop, D.J., and Kivirikko, K.I. (1995). Collagens: Molecular Biology, Diseases, and Potentials for Therapy. Annu. Rev. Biochem. *64*, 403–434.

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

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

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663–669.

Reuter, M.S., Tawamie, H., Buchert, R., Hosny Gebril, O., Froukh, T., Thiel, C., Uebe, S., Ekici, A.B., Krumbiegel, M., Zweier, C., et al. (2017). Diagnostic Yield and Novel Candidate Genes by Exome Sequencing in 152 Consanguineous Families With Neurodevelopmental Disorders. JAMA Psychiatry *74*, 293.

Ricard-Blum, S. (2011). The Collagen Family. Cold Spring Harb. Perspect. Biol. *3*, a004978–a004978.

Ricard-Blum, S., Baffet, G., and Théret, N. (2018). Molecular and tissue alterations of collagens in fibrosis. Matrix Biol. *68–69*, 122–149.

Rose, A.E., Brown, R.S.H., and Schlieker, C. (2015). Torsins: not your typical AAA+ ATPases. Crit. Rev. Biochem. Mol. Biol. *50*, 532–549.

Ruotsalainen, H., Sipilä, L., Vapola, M., Sormunen, R., Salo, A.M., Uitto, L., Mercer, D.K., Robins, S.P., Risteli, M., Aszodi, A., et al. (2006). Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. J. Cell Sci. *119*, 625.

Saito, K., Chen, M., Bard, F., Chen, S., Zhou, H., Woodley, D., Polischuk, R., Schekman, R., and Malhotra, V. (2009). TANGO1 Facilitates Cargo Loading at Endoplasmic Reticulum Exit Sites. Cell *136*, 891–902.

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

Sakai, L.Y. (1986). Type VII collagen is a major structural component of anchoring fibrils. J. Cell Biol. *103*, 1577–1586.

Sandoz, P.A., and van der Goot, F.G. (2015). How many lives does CLIMP-63 have? Biochem. Soc. Trans. *43*, 222–228.

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

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

Schäcke, H., Schumann, H., Hammami-Hauasli, N., Raghunath, M., and Bruckner-Tuderman, L. (1998). Two Forms of Collagen XVII in Keratinocytes: A FULL-LENGTH TRANSMEMBRANE PROTEIN AND A SOLUBLE ECTODOMAIN. J. Biol. Chem. *273*, 25937–25943.

Seppänen, A., Autio-Harmainen, H., Alafuzoff, I., Särkioja, T., Veijola, J., Hurskainen, T., Bruckner-Tuderman, L., Tasanen, K., and Majamaa, K. (2006). Collagen XVII is expressed in human CNS neurons. Matrix Biol. *25*, 185–188.

Shibata, H., Kanadome, T., Sugiura, H., Yokoyama, T., Yamamuro, M., Moss, S.E., and Maki, M. (2015). A New Role for Annexin A11 in the Early Secretory Pathway via Stabilizing Sec31A Protein at the Endoplasmic Reticulum Exit Sites (ERES). J. Biol. Chem. *290*, 4981–4993.

Shih, Y.-C., Chen, C.-L., Zhang, Y., Mellor, R.L., Kanter, E.M., Fang, Y., Wang, H.-C., Hung, C.-T., Nong, J.-Y., Chen, H.-J., et al. (2018). Endoplasmic Reticulum Protein TXNDC5 Augments Myocardial Fibrosis by Facilitating Extracellular Matrix Protein Folding and Redox-Sensitive Cardiac Fibroblast ActivationNovelty and Significance. Circ. Res. *122*, 1052–1068.

Shimoi, W., Ezawa, I., Nakamoto, K., Uesaki, S., Gabreski, G., Aridor, M., Yamamoto, A., Nagahama, M., Tagaya, M., and Tani, K. (2005). p125 Is Localized in Endoplasmic Reticulum Exit Sites and Involved in Their Organization. J. Biol. Chem. *280*, 10141–10148.

Tani, K., Mizoguchi, T., Iwamatsu, A., Hatsuzawa, K., and Tagaya, M. (1999). p125 Is a Novel Mammalian Sec23p-interacting Protein with Structural Similarity to Phospholipid-modifying Proteins. J. Biol. Chem. *274*, 20505–20512.

Torres, G.E., Sweeney, A.L., Beaulieu, J.-M., Shashidharan, P., and Caron, M.G. (2004). Effect of torsinA on membrane proteins reveals a loss of function and a dominant-negative phenotype of the dystonia-associated ΔE-torsinA mutant. Proc. Natl. Acad. Sci. U. S. A. *101*, 15650.

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

Walter, P., and Ron, D. (2011). The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science *334*, 1081.

Wang, Y., Liu, L., Zhang, H., Fan, J., Zhang, F., Yu, M., Shi, L., Yang, L., Lam, S.M., Wang, H., et al. (2016). Mea6 controls VLDL transport through the coordinated regulation of COPII assembly. Cell Res. *26*, 787–804.

White, S.R., and Lauring, B. (2007). AAA+ ATPases: Achieving Diversity of Function with Conserved Machinery. Traffic *8*, 1657–1667.

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

Yu, H., Liu, Y., Gulbranson, D.R., Paine, A., Rathore, S.S., and Shen, J. (2016). Extended synaptotagmins are Ca^{2+} -dependent lipid transfer proteins at membrane contact sites. Proc. Natl. Acad. Sci. *113*, 4362–4367.

Yu, S., Ito, S., Wada, I., and Hosokawa, N. (2018). ER-resident protein 46 (ERp46) triggers the mannose-trimming activity of ER degradation–enhancing α-mannosidase–like protein 3 (EDEM3). J. Biol. Chem. *293*, 10663–10674.

Yurchenco, P.D., Quan, Y., Colognato, H., Mathus, T., Harrison, D., Yamada, Y., and O'Rear, J.J. (1997). The chain of laminin-1 is independently secreted and drives secretion of its - and -chain partners. Proc. Natl. Acad. Sci. *94*, 10189–10194.

Zhao, C., Brown, R.S.H., Chase, A.R., Eisele, M.R., and Schlieker, C. (2013). Regulation of Torsin ATPases by LAP1 and LULL1. Proc. Natl. Acad. Sci. *110*, E1545–E1554.

8. APPENDIX

8.1 Appendix I: Proteins labeled by TANGO1-Cyto-APEX2 in HeLa cells

Protein names	Function	No. of peptides
Transferrin receptor protein 1	Required for cellular iron uptake	8
Nucleobindin-2	Calcium-binding protein.	7
Thioredoxin domain-containing protein 5	Possesses thioredoxin activity.	7
Endoplasmic reticulum resident protein 44	Mediates thiol-dependent retention in the early secretory pathway,	6
Nucleolar GTP-binding protein 2	GTPase that associates with pre-60S ribosomal subunits in the nucleolus and is required for their nuclear export and maturation	5
Transmembrane emp24 domain-containing protein 7	Potential role in vesicular protein trafficking, mainly in the early secretory pathway.	4
Myoferlin	Calcium/phospholipid-binding protein that plays a role in the plasmalemma repair mechanism of endothelial cells that permits rapid resealing of membranes disrupted by mechanical stress.	4
Large neutral amino acids transporter small subunit 1	Transport of large neutral amino acids such as phenylalanine, tyrosine, leucine, arginine and tryptophan,	3
Myeloid-derived growth factor (MYDGF)	Bone marrow-derived monocyte and paracrine-acting protein that promotes cardiac myocyte survival and adaptive angiogenesis for cardiac protection	3
40S ribosomal protein S13	structural constituent of ribosome	3
Erythrocyte band 7 integral membrane protein	Regulates ion channel activity and transmembrane ion transport.	3
Eukaryotic translation initiation factor 3 subunit A (eIF3a)	RNA-binding component of the eukaryotic translation initiation factor 3 (eIF-3) complex, w	3

Protein names	Function	No. of peptides
Zinc finger CCCH domain-containing protein 13	Associated component of the WMM complex, a complex that mediates N6-methyladenosine (m6A) methylation of RNAs.	3
60S ribosomal protein L3-like	structural constituent of ribosome	2
Transmembrane emp24 domain-containing protein 9	Appears to be involved in vesicular protein trafficking,	2
Endoplasmic reticulum-Golgi intermediate compartment protein 1	Possible role in transport between endoplasmic reticulum and Golgi	2
Translocon-associated protein subunit gamma	TRAP proteins are part of a complex whose function is to bind calcium to the ER membrane and thereby regulate the retention of ER resident proteins.	2
SRSF protein kinase 2	Serine/arginine-rich protein-specific kinase	2
Transmembrane emp24 domain-containing protein 4	Involved in vesicular protein trafficking, mainly in the early secretory pathway. targeting.	2
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens.	2
Sideroflexin-1	Might be involved in the transport of a component required for iron utilization into or out of the mitochondria.	2
Dynein light chain 1, cytoplasmic	Cytoplasmic dynein 1 acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules.	2
Casein kinase II subunit alpha	Catalytic subunit of a constitutively active serine/threonine-protein kinase complex that phosphorylates a large number of substrates.	2
BAG family molecular chaperone regulator 3 (BAG-3)	Co-chaperone for HSP70 and HSC70 chaperone proteins.	2
Collagen alpha-1(V) chain	It is a minor connective tissue component of nearly ubiquitous distribution.	2

8.2 Appendix II: Proteins labeled by TANGO1-Cyto-APEX2 in HEK293T cells

Protein names	Function	No. of peptides
Myosin-9	Plays a role in cytokinesis, cell shape, and specialized functions such as secretion and capping.	48
Transitional endoplasmic reticulum ATPase (TER ATPase)	Involved in the formation of the transitional endoplasmic reticulum (tER).	27
Extended synaptotagmin-1 (E-Syt1)	Helps tether the endoplasmic reticulum to the cell membrane.	17
Fructose-bisphosphate aldolase A	scaffolding protein	14
Ubiquitin-associated protein 2-like (Protein NICE-4)	Plays an important role in the activity of long-term repopulating hematopoietic stem cells.	14
Src substrate cortactin (Amplaxin)	Plays a role in intracellular protein transport and endocytosis.	13
Transgelin-2	regulates the actin cytoskeleton dynamics	11
Kinectin (CG-1 antigen)	Receptor for kinesin thus involved in kinesin-driven vesicle motility.	11
Cytoskeleton-associated protein 4	Mediates the anchoring of the endoplasmic reticulum to microtubules.	10
Protein transport protein Sec31A	Component of the coat protein complex II (COPII) which promotes the formation of transport vesicles from the endoplasmic reticulum (ER).	10
Calnexin	Calcium-binding protein that interacts with newly synthesized glycoproteins in the endoplasmic reticulum.	9
Glycine--tRNA ligase	Catalyzes the ligation of glycine to the 3'-end of its cognate tRNA.	9
Heterogeneous nuclear ribonucleoprotein Q (hnRNP Q)	Heterogeneous nuclear ribonucleoprotein (hnRNP) implicated in mRNA processing mechanisms.	9

Protein names	Function	No. of peptides
Annexin A11	Required for midbody formation and completion of the terminal phase of cytokinesis.	9
Cold shock domain-containing protein E1	RNA-binding protein.	9
Protein disulfide-isomerase A4	Chaperone	8
Leucine-rich repeat-containing protein 59	Required for nuclear import of FGF1, but not that of FGF2.	8
Plastin-3 (T-plastin)	Actin-bundling protein found in intestinal microvilli, hair cell stereocilia, and fibroblast filopodia. May play a role in the regulation of bone development.	8
Membrane-associated progesterone receptor component 1 (mPR)	Component of a progesterone-binding protein complex	8
Eukaryotic translation initiation factor 4B (eIF-4B)	Required for the binding of mRNA to ribosomes.	8
Profilin-1	Binds to actin and affects the structure of the cytoskeleton	7
Ras-related protein Rab-7a	Key regulator in endo-lysosomal trafficking.	7
B-cell receptor-associated protein 31 (Bap31)	Functions as a chaperone protein.	7
Inosine-5'-monophosphate dehydrogenase 2 (IMP dehydrogenase 2)	Catalyzes the conversion of inosine 5'-phosphate (IMP) to xanthosine 5'-phosphate (XMP),	7
Protein transport protein Sec23A	Component of the coat protein complex II (COPII) which promotes the formation of transport vesicles from the endoplasmic reticulum (ER).	7

8.3 Appendix III A: Proteins labeled by TANGO1-Lum -APEX2 in HeLa cells

Protein names	Function	No. of peptides
Nodal modulator 2 (pM5 protein 2)	Antagonist of nodal signaling	26
Nodal modulator 1 (pM5 protein)	Antagonist of nodal signaling	26
Laminin subunit beta-1	Mediates the attachment, migration and organization of cells.	20
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens.	19
Prolyl 3-hydroxylase 1	Has prolyl 3-hydroxylase activity catalyzing the post-translational formation of 3-hydroxyproline in -Xaa-Pro-Gly- sequences in collagen	19
Thioredoxin domain-containing protein 5	Possesses thioredoxin activity	18
Alpha-2-macroglobulin receptor-associated protein (Alpha-2-MRAP)	Molecular chaperone for LDL receptor-related proteins	18
Inhibitor of nuclear factor kappa-B kinase-interacting protein (IKBKB-interacting protein)	Target of p53/TP53 with pro-apoptotic function.	17
Prolyl 4-hydroxylase subunit alpha-2 (4-PH alpha-2)	Catalyzes the post-translational formation of 4-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens and other proteins.	17
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 Lysyl hydroxylase 2	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens.	16
Coiled-coil domain-containing protein 47	endoplasmic reticulum-associated degradation activity	16
ER membrane protein complex subunit 1	protein folding in endoplasmic reticulum	15
Collagen alpha-1(V) chain	connective tissue component	12
Contactin-associated protein 1 (Caspr)	Required, with CNTNAP2, for radial and longitudinal organization of myelinated axons.	12

Protein names	Function	No. of peptides
DnaJ homolog subfamily C member 10	Endoplasmic reticulum disulfide reductase involved both in the correct folding of proteins and degradation of misfolded proteins.	12
Endoplasmic reticulum resident protein 44	Mediates thiol-dependent retention in the early secretory pathway.	11
Endoplasmic reticulum export factor CTAGE5	Required for collagen VII secretion by loading COL7A1 into transport carriers and recruiting PREB/SEC12 at the endoplasmic reticulum exit sites.	11
Reticulocalbin-2	Not known. Binds calcium.	10
Peptidyl-prolyl cis-trans isomerase FKBP9	Accelerates the folding of proteins during protein synthesis.	9
Integrin beta-1	receptors for collagen	9
Endoplasmic reticulum-Golgi intermediate compartment protein 1 (ERGIC-32)	Possible role in transport between endoplasmic reticulum and Golgi.	9
Transferrin receptor protein 1	Required for cellular iron uptake	9
Nucleolar GTP-binding protein 2	GTPase that associates with pre-60S ribosomal subunits in the nucleolus and is required for their nuclear export and maturation	9
Epidermal growth factor receptor	Receptor tyrosine kinase	9
DnaJ homolog subfamily B member 11	Serves as a co-chaperone for HSPA5. Binds directly to both unfolded proteins that are substrates for ERAD and nascent unfolded peptide chains.	8
Fibrillin-1	Structural component of microfibrils of the extracellular matrix,	8
Laminin subunit alpha-5	mediates the attachment, migration and organization of cells .	8
Reticulocalbin-3	Calcium binding	7
Prenylcysteine oxidase 1	Involved in the degradation of prenylated proteins.	7

Protein names	Function	No. of peptides
Transmembrane protein 43	May have an important role in maintaining nuclear envelope structure by organizing protein complexes at the inner nuclear membrane.	7
Chloride channel CLIC-like protein 1	Seems to act as a chloride ion channel.	7
Collagen alpha-1(XII) chain	Type XII collagen interacts with type I collagen-containing fibrils.	7
ADP-dependent glucokinase (ADP-GK)	Catalyzes the phosphorylation of D-glucose to D-glucose 6-phosphate.	6
Neural cell adhesion molecule L1	Neural cell adhesion molecule involved in the dynamics of cell adhesion and in the generation of transmembrane signals at tyrosine kinase receptors.	6
Collagen alpha-1(VII) chain	Stratified squamous epithelial basement membrane protein that forms anchoring fibrils which may contribute to epithelial basement membrane organization and adherence by interacting with extracellular matrix (ECM) proteins such as type IV collagen.	6
Myeloid-derived growth factor (MYDGF)	Bone marrow-derived monocyte and paracrine-acting protein that promotes cardiac myocyte survival a	5
ERO1-like protein alpha (ERO1-L)	Oxidoreductase involved in disulfide bond formation in the endoplasmic reticulum.	5
Nucleobindin-2	Calcium-binding protein.	5
Nicastrin	Essential subunit of the gamma-secretase complex, an endoprotease complex that catalyzes the intramembrane cleavage of integral membrane proteins such as Notch receptors and APP .	5
Thioredoxin domain-containing protein 12	Possesses significant protein thiol-disulfide oxidase activity.	5
Endoplasmic reticulum protein SC65	Part of a complex composed of PLOD1, P3H3 and P3H4 that catalyzes hydroxylation of	5

Protein names	Function	No. of peptides
	lysine residues in collagen alpha chains and is required for normal assembly and cross-linking of collagen fibrils	
Nucleotide exchange factor SIL1	Required for protein translocation and folding in the endoplasmic reticulum	5
Transmembrane emp24 domain-containing protein 9	involved in vesicular protein trafficking	4
Transmembrane emp24 domain-containing protein 7	Potential role in vesicular protein trafficking, mainly in the early secretory pathway.	4
Malectin	May play a role in the early steps of protein N-glycosylation.	4
GPI transamidase component PIG-T	Component of the GPI transamidase complex.	4
Translocation protein SEC63 homolog	Required for integral membrane and secreted preprotein translocation across the endoplasmic reticulum membrane.	4
Selenoprotein F (15 kDa selenoprotein)	contributes to the quality control of protein folding in the endoplasmic reticulum	4
KDEL motif-containing protein 2	glucosyltransferase	4
Prolyl 3-hydroxylase 2	Prolyl 3-hydroxylase that catalyzes the post-translational formation of 3-hydroxyproline on collagens	4

8.4 Appendix III B: Secreted Proteins labeled by TANGO1-Lum - APEX2 in HeLa cells

Protein names	Molecular function	No. of peptides

ADP-dependent glucokinase	ADP-specific glucokinase	6
Collagen alpha-1(V) chain	extracellular matrix constituent	12
Collagen alpha-1(VII) chain	extracellular matrix constituent	6
Collagen alpha-1(XII) chain	extracellular matrix constituent	7
Fibrillin-1	extracellular matrix constituent	8
Laminin subunit alpha-5	extracellular matrix constituent	8
Laminin subunit beta-1	extracellular matrix constituent	20
Myeloid-derived growth factor (MYDGF)	Angiogenesis	5

8.5 Appendix IV A: Proteins labeled by TANGO1-Lum -APEX2 in HEK293T cells

Protein names	**Function**	**No. of peptides**
Protein disulfide-isomerase A4	Chaperone	58
Neutral alpha-glucosidase AB	Catalytic subunit of glucosidase II that cleaves sequentially the 2 innermost alpha-1,3-linked glucose residues from the Glc(2)Man(9)GlcNAc(2) oligosaccharide precursor of immature glycoproteins	45
Cytoskeleton-associated protein 4	High-affinity epithelial cell surface receptor for APF.; Mediates the anchoring of the endoplasmic reticulum to microtubules.	41
UDP-glucose:glycoprotein glucosyltransferase 1	Recognizes glycoproteins with minor folding defects.	38
Protein disulfide-isomerase	This multifunctional protein catalyzes the formation, breakage and rearrangement of disulfide bonds.	36

Protein names	Function	No. of peptides
Nodal modulator 2 (pM5 protein 2)	Anatgonist of Nodal signaling	35
Nodal modulator 1 (pM5 protein)	Antagonist of nodal signaling	33
ER membrane protein complex subunit 1	Protein folding	27
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1(Lysyl hydroxylase 1)	Part of a complex composed of PLOD1, P3H3 and P3H4 that catalyzes hydroxylation of lysine residues in collagen	26
Calreticulin	Calcium-binding chaperone that promotes folding, oligomeric assembly and quality control	23
Thioredoxin domain-containing protein 5	Possesses thioredoxin activity.	20
Glucosidase 2 subunit beta	Regulatory subunit of glucosidase II	20
Aspartyl/asparaginyl beta-hydroxylase	specifically hydroxylates an Asp or Asn residue in certain epidermal growth factor-like (EGF) domains of a number of proteins.	20
Kinectin	Receptor for kinesin thus involved in kinesin-driven vesicle motility	20
ERO1-like protein alpha (ERO1-L)	Oxidoreductase involved in disulfide bond formation in the endoplasmic reticulum.	19
Desmoplakin	Major high molecular weight protein of desmosomes.	19
Endoplasmic reticulum resident protein 29 (ERp29)	Plays an important role in the processing of secretory proteins within the endoplasmic reticulum (ER),	18
Endoplasmic reticulum resident protein 44 (ER protein 44)	Mediates thiol-dependent retention in the early secretory pathway, f	18
Procollagen-lysine,2-oxoglutarate 5-	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens. T	18

Protein names	Function	No. of peptides
dioxygenase 3(Lysyl hydroxylase 3)		
Prolyl 4-hydroxylase subunit alpha-1 (4-PH alpha-1)	Catalyzes the post-translational formation of 4-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens and other proteins.	18
DnaJ homolog subfamily C member 10	Endoplasmic reticulum disulfide reductase involved both in the correct folding of proteins and degradation of misfolded proteins.	18
Reticulocalbin-1	May regulate calcium-dependent activities in the endoplasmic reticulum lumen or post-ER compartment.	17
Calnexin	Calcium-binding protein that interacts with newly synthesized glycoproteins in the endoplasmic reticulum.	17
Peptidyl-prolyl cis-trans isomerase FKBP10	PPIases accelerate the folding of proteins during protein synthesis.	17
Myosin-9	Cellular myosin that appears to play a role in cytokinesis, cell shape, and specialized functions such as secretion and capping.	17
Calumenin	Involved in regulation of vitamin K-dependent carboxylation	16
Alpha-2-macroglobulin receptor-associated protein	Molecular chaperone for LDL receptor-related proteins	16
Plectin	Interlinks intermediate filaments with microtubules and microfilaments and anchors intermediate filaments to desmosomes or hemidesmosomes.	16
Nucleobindin-2	Calcium-binding protein.	15
Nuclear pore membrane glycoprotein 210	Nucleoporin essential for nuclear pore assembly and fusion, nuclear pore spacing, as well as structural integrity.	15

Protein names	Function	No. of peptides
Peptidyl-prolyl cis-trans isomerase B (Cyclophilin B)	PPIase that catalyzes the cis-trans isomerization of proline imidic peptide bonds in oligopeptides and may therefore assist protein folding.	14
Procollagen galactosyltransferase 1	Beta-galactosyltransferase that transfers beta-galactose to hydroxylysine residues of type I collagen	14
DnaJ homolog subfamily B member 11	Serves as a co-chaperone for HSPA5. Binds directly to both unfolded proteins that are substrates for ERAD and nascent unfolded peptide chains,	13
Transmembrane protein 43 (Protein LUMA)	May have an important role in maintaining nuclear envelope structure by organizing protein complexes at the inner nuclear membrane.	13
Prenylcysteine oxidase 1	Involved in the degradation of prenylated proteins.	13
Inhibitor of nuclear factor kappa-B kinase-interacting protein	Target of p53/TP53 with pro-apoptotic function.	13
Mesencephalic astrocyte-derived neurotrophic factor	Up-regulated and secreted by the ER/SR in response to ER stress and hypoxia	12
Prolyl 3-hydroxylase 1	Has prolyl 3-hydroxylase activity catalyzing the post-translational formation of 3-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens,	12
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens.	11
Protein ERGIC-53	Mannose-specific lectin. LMAN1-MCFD2 complex forms a specific cargo receptor for the ER-to-Golgi transport of selected proteins. .	11
KDEL motif-containing protein 2	glucosyltransferase	10
Torsin-1A-interacting protein 1	Required for nuclear membrane integrity. Induces TOR1A and TOR1B ATPase activity	10

Protein names	Function	No. of peptides
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	Catalytic subunit of the N-oligosaccharyl transferase (OST) complex	10
ADP-dependent glucokinase	Catalyzes the phosphorylation of D-glucose to D-glucose 6-phosphate u	10
Prolyl 4-hydroxylase subunit alpha-2 (4-PH alpha-2)	Catalyzes the post-translational formation of 4-hydroxyproline in -Xaa-Pro-Gly- sequences in collagens and other proteins.	10
Neudesin	Acts as a neurotrophic factor in postnatal mature neurons enhancing neuronal survival.	9
Transmembrane emp24 domain-containing protein 9	Appears to be involved in vesicular protein trafficking, mainly in the early secretory pathway.	9
Mannosyl-oligosaccharide glucosidase	Cleaves the distal alpha 1,2-linked glucose residue from the Glc(3)Man(9)GlcNAc(2) oligosaccharide precursor in a highly specific manner.	9
Reticulocalbin-2	Not known. Binds calcium.	8
GDP-fucose protein O-fucosyltransferase 1	Catalyzes the reaction that attaches fucose through an O-glycosidic linkage	8

8.6 Appendix IV B: Secreted Proteins labeled by TANGO1-Lum - APEX2 in HEK293T cells

Protein names	Function	No. of peptides
ADP-dependent glucokinase	Catalyzes the phosphorylation of D-glucose to D-glucose 6-phosphate	10
Cartilage-associated protein	Necessary for efficient 3-hydroxylation of fibrillar collagen prolyl residues.	6

Protein names	Function	No. of peptides
Chitinase domain-containing protein 1	Saccharide- and LPS-binding protein with possible roles in pathogen sensing and endotoxin neutralization.	7
Coiled-coil domain-containing protein 134	Promotes proliferation and activation of CD8(+) T cells,	3
Collagen alpha-1(VI) chain	Collagen VI acts as a cell-binding protein.	7
Laminin subunit beta-1	mediates the attachment, migration and organization of cells	3
Laminin subunit gamma-1	mediates the attachment, migration and organization of cells	6
Mesencephalic astrocyte-derived neurotrophic factor	Inhibits cell proliferation and endoplasmic reticulum (ER) stress-induced cell death	12
Myeloid-derived growth factor (MYDGF)	Bone marrow-derived monocyte and paracrine-acting protein that promotes cardiac myocyte survival and adaptive angiogenesis	4
Neudesin	Acts as a neurotrophic factor in postnatal mature neurons enhancing neuronal survival.	9
Peroxidasin homolog	role in extracellular matrix formation	3