UNIVERSITÉ DE GENÈVE
Département de biologie cellulaire

FACULTÉ DES SCIENCES
Professeur Didier Picard

Département de physiologie
cellulaire et métabolisme

FACULTÉ DE MÉDECINE
Professeur Pierre Cosson

Role of TM9 proteins in intracellular transport and sorting of transmembrane domains

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève

pour obtenir le grade de Docteur ès science, mention biologie

par

Alexandre Vernay

de

Limoges (France)

Thèse N° 5069

Genève

Atelier de reprographie ReproMail

2017

UNIVERSITÉ DE GENÈVE
Département de biologie cellulaire

FACULTÉ DES SCIENCES
Professeur Didier Picard

Département de physiologie
cellulaire et métabolisme

FACULTÉ DE MÉDECINE
Professeur Pierre Cosson

Role of TM9 proteins in intracellular transport and sorting of transmembrane domains

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève

pour obtenir le grade de Docteur ès science, mention biologie

par

Alexandre Vernay

de

Limoges (France)

Thèse N° 5069

Genève

Atelier de reprographie ReproMail

2017

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès Sciences
Mention biologe

Thèse de *Monsieur Alexandre VERNAY*

intitulée :

"Role of TM9 Proteins in Intracellular Transport and Sorting of Transmembrane Domains"

La Faculté des sciences, sur le préavis de Monsieur P. COSSON, professeur ordinaire et directeur de thèse (Faculté de médecine, Département de physiologie cellulaire et métabolisme), Monsieur D. PICARD, professeur ordinaire et codirecteur de thèse (Département de biologie cellulaire), Monsieur F. LETOURNEUR, docteur (Dynamique des interactions membranaires normales et pathologiques, Université de Montpellier, France), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 26 avril 2017

Thèse - 5069 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres de mon jury, le Docteur François Letourneur et le Professeur Didier Picard pour avoir accepté d'évaluer ce travail.

Merci infiniment à Pierre de m'avoir accordé ta confiance, encadré, transmis (en partie !) ton savoir, ainsi que ta rigueur intellectuelle et scientifique. Ce travail n'aurait bien évidemment pas été possible sans ta précieuse aide.

Je remercie chaleureusement mes parrains de thèse, les Professeurs Bernhard Wehrle-Haller et Dominique Belin.

Merci à tous les membres du laboratoire, passés et présents, trop nombreux pour être tous cités ici, pour votre aide et votre bonne humeur. Plus particulièrement, merci tout d'abord à Anna pour ton aide tout au long de ma thèse. Merci à Jackie, avec qui j'ai partagé de nombreux ascenseurs émotionnels en sortie de réunion du lundi matin. Wanessa, pour ton aide, ta gentillesse et tes conseils. Hajer, Astrid, Marco, merci pour ces moments partagés autour du comptoir. Merci au mélomane du groupe Philippe et son Philippe-Show quotidien. Merci à Myriam, ma stagiaire épanouie. Bon courage aux 'petits' nouveaux Estelle, Tania et Otmane.
Last but not least, merci aux arrogants de post-docs que sont Ayman et Jade, ainsi que Romain ! Ayman d'abord, le 'grand frère' pour qui chaque manip est une œuvre d'art. Jade, la Jadastrophe des quartiers nord de Pertuis. Romainpédia 25% chelou, et avec qui 'on se comprend'. On a partagé de la science (un peu), du ski (un peu plus), des memes (beaucoup) et des fails (je n'ose pas dire combien). Je vous souhaite à tous beaucoup de bonheur et de réussite.

Merci à tous les membres du département PHYME, en particulier Tamara et Corine pour votre aide, Hervé pour l'organisation des sorties de département, Priscilla pour m'avoir permis de faire ces TP d'histologie, Lelio pour vos encouragements et votre bonne humeur, le petit Cyril pour son humour anti-grands. Merci à tous pour votre accueil, votre gentillesse et ces bons moments passés ensemble.

Une des plus grandes forces de Genève est indubitablement la possibilité de rencontrer des gens venant de partout dans le monde. Merci donc aux amis rencontrés ici, Dani et Doro en tête, toujours prêts à se réunir au Duke, au squash, en Italie ou en Ecosse. Piango, pour ta bonne humeur et ton rire communicatif, ainsi que pour les fameux Piango's training du mardi soir. Les malheureux de la coupe de Noel (Nico, Dani, Joanna, Ines, Jade, Romain et Anne-Laure), nager dans le lac en décembre, c'était un beau challenge et on a mérité notre médaille. Merci aussi aux participants du beach-volley du jeudi soir (Alfredo, Jordi, et ceux déjà nommés), on était mauvais mais on rigolait bien. Je ne peux bien évidemment pas citer tout le monde ici, et je m'en excuse, mais je remercie tous ceux que j'ai pu rencontrer pendant ces 4 ans et quelques mois, et qui ont participé de près ou de loin à cette aventure.

Merci aux grenoblois (même s'il n'y en a plus beaucoup sur place) d'être toujours là quand il le faut. Romain et Camille, Blaise et Caro, Julie et Kévin, le vieux BenJ, Bertrand mon fidèle acolyte de voyages, Omy, Françoise, Emilie et Ludo, VayVay la volaille, Camille et Sylvian, Neda, Alex, Laurie et Erwann, Martin, Clémence (qui ça ?), Marie et Jérome, Quentin, et les autres, merci pour ces retrouvailles qui ponctuent les années qui passent.

Merci à Anne-Laure d'être présente à mes côtés et de me soutenir.

Enfin, un immense merci à mes parents et à ma sœur, pour tout ce qu'ils m'apportent, et leur soutien sans faille.

La dernière ligne de ces remerciements revient à Valentina.

CONTENTS

List of figures

List of abbreviations

Arf	Adenosine diphosphate-Ribosylation Factor
ATF6	Activating Transcription Factor 6
ATP	Adenosine TriPhosphate
BiP	Binding immunoglobluin Protein
cDNA	complementary DeoxyriboNucleic Acid
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
COP	COat Protein
CRD	Carbohydrate Recognition Domain
DAF	Decay-Accelerating Factor
DNA	DeoxyriboNucleic Acid
ER	Endoplasmic Reticulum
ERES	Endoplasmic Reticulum Exit Site
ERGIC	Endoplasmic Reticulum to Golgi Intermediate Compartment
Erv	Endoplasmic Reticulum Vesicle protein
GAP	GTPase Activating Protein
GDP	Guanosine DiPhosphate
GEF	Guanine Exchange Factor
GOLD	GOLgi Dynamic
GPI	GlycosylPhosphatidylInositol
GTP	Guanosine TriPhosphate
Hsp	Heat-Shock Protein
Ig	ImmunoGlobulin
IP3	Inositol 1,4,5-triPhosphate

kDa	kiloDalton
mRNA	messenger RiboNucleic Acid
MTOC	MicroTubule Organization Center
NBD	Nucleotide-Binding Domain
NEF	Nucleotide Exchange Factor
NSF	N-ethylmaleimide-Sensitive Factor
OST	OligosaccharylTransferase
PDI	Protein Disulfide-Isomerase
PPiases	Peptidyl Propyl *cis/trans* Isomerase
Rer	Retention in Endoplasmic Reticulum
SBD	Substrate-Binding Domain
SCAP	Sterol regulatory element-binding protein Cleavage-Activating Protein
SNAP	Soluble N-ethylmaleimide-Sensitive Factor Attachment Protein
SNARE	Soluble N-ethylmaleimide-sensitive factor Attachment protein REceptors
SREBP	Sterol Regulatory Element-Binding Protein
SRP	Signal Recognition Particle
SRP-R	Signal Recognition Particle-Receptor
SRP-R	
TCR	T-Cell Receptor
TGN	*Trans*-Golgi Network
TM9SF	TransMembrane 9 SuperFamily
TMD	TransMembrane Domain
TRAM	TRanslocating chain-Associated Membrane protein

tRNA	Transfer RiboNucleic Acid
UPR	Unfolded Protein Response
Vps	Vacuolar Protein Sorting
VSV-G	Vesicular Stomatitis Virus-G protein
VTC	Vesiculo-Tubular Cluster

Résumé

Le transport membranaire est un processus par lequel des lipides et des protéines sont répartis au sein de la cellule. Ce processus est hautement régulé et est particulièrement important pour l'adressage d'une protéine au bon endroit, et au bon moment.

La voie de sécrétion forme un système endomembranaire complexe comprenant, entre autres, le réticulum endoplasmique et l'appareil de Golgi. Le transport et le triage des protéines dans la voie de sécrétion sont des processus essentiels qui déterminent la composition biochimique et les fonctions spécifiques de chaque organelle.

Au cours de ma thèse, j'ai étudié la famille des protéines TM9, d'abord chez *Dictyostelium discoideum* puis chez des cellules de mammifère. Nos résultats montrent que Phg1a, l'un des trois membres de la famille TM9 chez *Dictyostelium* transporte un ensemble de domaines transmembranaires (TMDs) hors du reticulum endoplasmique, et permet leur expression à la surface des cellules. Nous avons démontré que Phg1a s'associe de façon spécifique avec des TMDs riches en glycine, *via* des interactions intra-membranaires. Il existe quatre protéines TM9 chez les mammifères. TM9SF4, le plus proche orthologue humain de Phg1a, contrôle également le transport vers la surface de TMDs riches en glycine. Mes résultats plus récents ont montré que TM9SF4 interagit également avec des TMDs contenant un résidu chargé, et permet leur relocalisation du reticulum endoplasmique dans l'appareil de Golgi.

Une interprétation possible de mes résultats est que Phg1a/TM9SF4 pourrait être un élément du mécanisme assurant le tri des domaines transmembranaires dans l'appareil de Golgi.

Abstract

Membrane transport is the process by which lipids and proteins are distributed within the cell. This process is highly regulated and particularly important to target a protein to the correct place, and with the correct timing.

The secretion pathway forms a complex endomembrane system containing, among others, the endoplasmic reticulum and the Golgi complex. The protein transport and sorting in the secretion pathway are essential processes that determine the biochemical contents and specific functions of each organelle.

During my PhD, I studied the family of TM9 proteins, first in *Dictyostelium discoideum* and then in mammalian cells. Our results showed that Phg1a, one of the three members of TM9 proteins in *Dictyostelium,* transports proteins with a subset of transmembrane domains (TMDs) out of the endoplasmic reticulum (ER), and allows their expression at the cell surface. We demonstrated that Phg1a specifically associates with glycine-rich TMDs, through intra-membrane interactions. Four members of TM9 proteins exist in mammalian cells. TM9SF4, one of the closest human orthologs of Phg1a, also controls the surface transport of glycine-rich TMDs. Further studies showed that TM9SF4 interacts also with TMDs exhibiting charged residues, and allows its relocalization from the ER to the Golgi complex.

A possible interpretation of my results is that Phg1a/TM9SF4 may be a component of the mechanism ensuring the sorting of TMDs in the Golgi complex.

I. INTRODUCTION

I. Introduction

A. Transport along the early secretion pathway

Mammalian organisms are a complex assembly of many different cell types. Each of these cells is subdivided into distinct membrane compartments, restricting de facto each biological process to a specific environment. Compartmentalization of cells is particularly important to allow complex and multiple cellular functions to be achieved.

Intracellular organelles include for instance the nucleus, where the genome is confined; the endoplasmic reticulum (ER) where, among many functions, proteins are folded; the Golgi complex, which ensures the maturation of secreted proteins; mitochondria, where most of cellular ATP is produced. Finally, the plasma membrane delimits the cell and exposes different receptors or adhesion molecules to the surface. Any disorder in the cellular compartmentalization, or in the targeting of individual components of these organelles can have dramatic effects on the physiology or even the survival of the cell. Such disorders often also affect the physiology of the whole organism.

My work has mainly been focused on the complex endomembrane system devoted to the synthesis, the maturation and the transport of molecules, called the secretion pathway. As depicted in Figure 1, secreted proteins are transported from the ER to the Golgi complex for further modifications and maturation, before being secreted (for soluble proteins) or delivered to the plasma membrane (for membrane proteins).

Soluble and membrane proteins can also be distributed along the endomembranes of the secretion pathway.

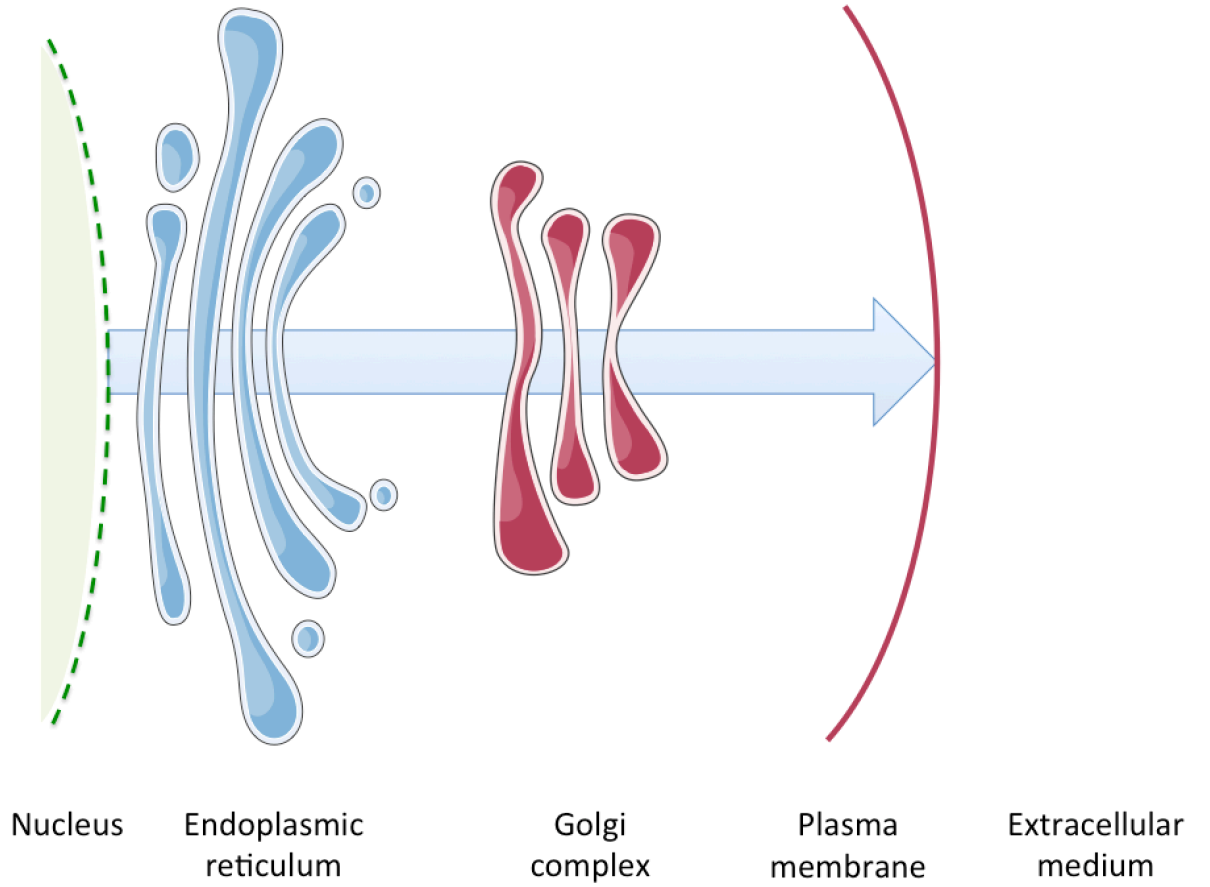

Figure 1: A simplified view of the secretion pathway. Proteins entering the secretion pathway are first synthetized in the Endoplasmic Reticulum (ER), reach then the Golgi complex where they are modified and mature to become fully active. Membrane proteins are then allowed to reach the plasma membrane, or soluble proteins can be secreted in the extracellular medium, or reach other subcompartments not depicted in this scheme, such as endosomes or lysosomes.

The aim of this manuscript is to describe the intracellular transport and the sorting of proteins in the early secretory pathway. This includes the translocation of proteins in the ER, transport at the ER to Golgi interface, and transport in the Golgi complex. Therefore, the first chapter of this manuscript describes the translocation of proteins in the early secretion pathway (A.). Proteins are first synthetized (1), and

inserted co-translationally in the endoplasmic reticulum (2), then transported to the Golgi complex (3). An unconventional secretion pathway has also been described (4).

Proteins are transported from one compartment to another by vesicular intermediates. The formation of vesicles is the subject of the second chapter (B.). There are different types of vesicles but they share common features, described in the first section (1). Specific molecular mechanisms involved in anterograde transport are detailed (2) as well as those involved in retrograde transport (3).

Proteins are not transported indiscriminately along the secretion pathway. On the contrary, transport is tightly linked to sorting events. The third chapter focuses on the general principles ensuring the sorting of proteins in the early secretion pathway (C.). Membrane sorting is based either on the exclusion of specific proteins from transport vesicles (1), or on the contrary on the concentration of proteins into transport vesicles (2). Finally, proteins in transport vesicles that are found at the same concentration to that found in the donor compartment are transported as part of the bulk flow.

Transmembrane domains (TMDs) are particularly relevant for my work, and the final chapter of this introduction is specifically devoted to the sorting of proteins containing these TMDs (for simplicity referred to as 'TMDs' in this chapter) in the early secretion pathway (D). I examine the structure of TMDs (1), the sorting motifs found in TMDs (2) and finally the mechanisms ensuring the sorting of TMDs (3).

1. Protein synthesis

Proteins are large macromolecules representing more than 50% of the dry mass of cells. They are composed mainly of a succession of amino acids, and they play many

different roles within the cells. These functions range from establishing the basic architecture of the cell (cytoskeleton) to ensuring the defence against pathogens (immunoglobulins), ensuring intercellular communication (through messengers such as hormones), enzymatic reactions (DNA replication), etc…

Proteins are usually soluble when they are found in the cytosol or in the lumen of organelles. Alternatively, integral membrane proteins can contain one or several transmembrane domains that are inserted into lipidic membranes. Irrespective of their final fate, the transcription of genes into messenger RNAs (mRNAs) is common to all proteins.

a) Transcription

Proteins are the product of the translation of mRNAs, the sequence of which are derived from the nucleotidic sequence of the corresponding genes. In eukaryotes, RNA polymerase II transcribes genes in the nucleus, and produces primary transcripts. This pre-mRNA undergoes post-transcriptional modifications, such as the addition of a 7-methylguanosine cap in 5' and a polyadenine tail in 3'. The role of the cap is to protect the mRNA from the degradation by ribonucleases and to facilitate its transport to the cytoplasm. The polyadenine tail also protects the mRNA from enzymatic degradation.

The non-coding elements, or introns, of the pre-mRNA are removed by the spliceosome, in a process called splicing. The remaining mature mRNA is then exported to the cytoplasm for its translation. The transcription process is summarized in Figure 2.

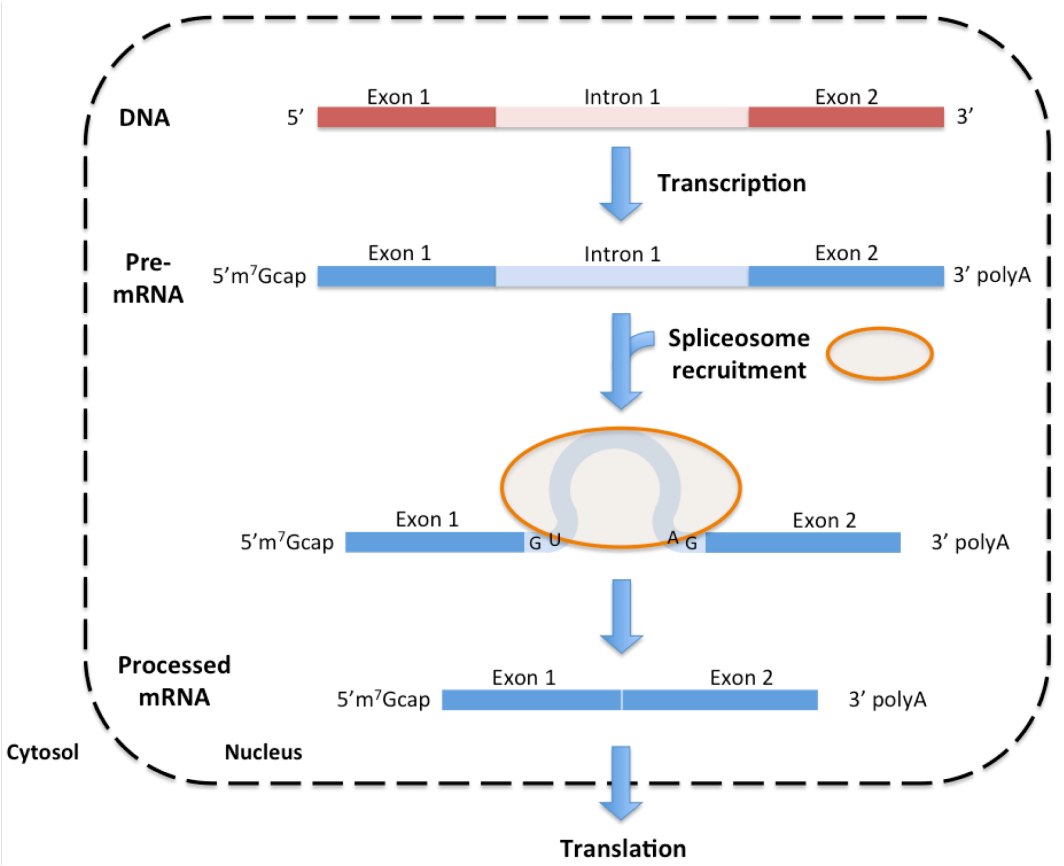

Figure 2: From DNA to RNA, the transcription process. In the nucleus, genes are organized in coding sequences called exons (dark red) and non-coding sequences, named introns (light red). The RNA Polymerase II transcribes protein-coding genes into a pre-messenger RNA (in blue), formed by the exons (dark blue) and the introns (light blue). The spliceosome complex, responsible for the removal of the introns, then processes this pre-messenger RNA. The final, fully processed mRNA is then exported in the cytosol for its translation.

b) Translation of cytosolic proteins

Cytosolic proteins are translated by free ribosomes in the cytosol. In eukaryotic cells, ribosomes are responsible for protein synthesis and are formed by four ribosomal RNAs and dozen of ribosomal proteins, organized in a small and a large subunit (Graifer & Karpova, 2015). The translation is divided into three steps: the initiation, the elongation and the termination.

The initiation phase corresponds to the binding of a ribosome, together with various proteins called Initiation Factors, to the G cap of the mRNA. The transcript is then scanned until an AUG initiation codon, base-pairing with the corresponding anticodon sequence of a methionyl initiator transfer RNA (tRNA), bound to the ribosome.

In the elongation phase, the ribosome adds aminoacids to the nascent polypeptide, strictly guided by the three-nucleotide genetic code. To achieve this process, Elongation Factors are needed for the selection of aminoacyl-tRNA, the formation of the peptide bond, and the translocation of the ribosome from one codon to the other.

Aminoacids are added to the nascent polypeptidic chains until the appearance of a stop codon, engaging Release Factors that terminate the translation, and release the ribosome from the mRNA. The newly synthetized protein terminates folding to reach its final three-dimensional shape, with the help of chaperones proteins. Post-translational modifications occurring during or after protein synthesis modulate the activity, the stability, or the localization of each protein. Both the ribosome and the mRNA template are recycled for the production of new polypeptides.

c) Translation of secreted and membrane proteins

Proteins destined to the secretory pathway are translated by ER-bound ribosomes, giving the ER its rough aspect (as opposed to the smooth ER, devoid of ribosomes, and responsible for the synthesis of fatty acids and phospholipids). Proteins

produced at the level of the rough ER can gain access to the secretion pathway, as detailed in the next section.

The simplest secreted protein usually bear a cleavable, N-terminal signal sequence or signal peptide composed by a short, hydrophobic stretch of 5-30 aminoacids, which is recognized by the cytosolic ribonucleoprotein named Signal Recognition Particle (SRP). SRP is a complex composed of 6 proteic subunits and a 7S RNA (Walter, Gilmore, Müller, & Blobel, 1982; Walter & Johnson, 1994). SRP binds the signal peptide of the nascent protein, stops the elongation and targets the ribosome to the rough ER, by binding the Signal Recognition Particle Receptor (SRP-R). This interaction guides the nascent protein to a protein translocator, or translocon, ensuring the passage of the nascent protein into the ER as translation resumes.

The translocon is a complex of proteins, assembled into a channel, allowing the passage of synthesized proteins into the lumen of the ER. Membrane proteins are, on the other hand, inserted into the membrane of the ER. The translocon was first described in yeast (Deshaies & Schekman, 1987; Johnson & van Waes, 1999; Rothblatt, Deshaies, Sanders, Daum, & Schekman, 1989) where the essential SEC61, SEC62 and SEC63 were identified. These three ER-resident transmembrane proteins associate into the SEC complex, making the channel of the translocon. Sec61 was identified as the core subunit of the channel (Stirling, Rothblatt, Hosobuchi, Deshaies, & Schekman, 1992) and is highly conserved in human (Sec61 α is the homologue of the yeast Sec61) and in bacteria (SecY in *E. coli*).

For soluble, secreted proteins, the passage through the translocon machinery is usually cotranslational. After the end of the translation, the release of the ribosome triggers the closure of the channel (Hamman, Hendershot, & Johnson, 1998). The translocation can also be post-translational, when fully synthetized proteins are

transported to the channel by Heat Shock Proteins Hsp40 and Hsp70 (Zimmermann, Eyrisch, Ahmad, & Helms, 2011).

Concerning the synthesis of integral proteins, TMDs need to be inserted into the ER membrane. The orientation and the number of TMDs determine the protein topology (Figure 3). Membrane proteins also use the translocon machinery, however the TMDs are not translocated into the lumen of the ER, but are rather captured in the translocon, and later inserted into the lipid bilayer of the ER. TMDs are thought to exit the translocon through its 'lateral gate'. The TRanslocating chain-Associated Membrane protein (TRAM), a component of the translocation channel creates a hydrophobic environment in the aqueous pore of the translocon, which facilitates the insertion of the TMD in the lipid bilayer.

There are several types of transmembrane proteins and their cotranslational insertion into the ER follows slightly different scenarios.

Type-I transmembrane proteins exhibit an N-terminal cleavable signal sequence, recognized by the SRP, as described above. After cleavage of the signal peptide, the synthesis of the protein resumes and it is inserted cotranslationally in the lumen of the ER, until the appearance of the TMD, which is recognized by the translocon and inserted in the ER membrane. The cytosolic C-terminal part of the nascent protein is synthetized last.

Type-II membrane proteins do not exhibit a cleavable signal sequence. Instead, the TMD is recognized as a non-cleavable signal sequence. It is recognized by the SRP and targets the nascent polypeptide to the translocon. After the release of the SRP, the C-terminal part of the protein is then synthetized through the translocon. This results in a protein with the N-terminal domain facing the cytosol and the C-terminal domain in the lumen of the ER.

Type-III membrane proteins, or polytopic membrane proteins, are characterized by multiple TMDs. Their synthesis and insertion in the ER membrane involves the same machinery as type I and type II membrane proteins, although the translocation events at the level of the translocon are more complex.

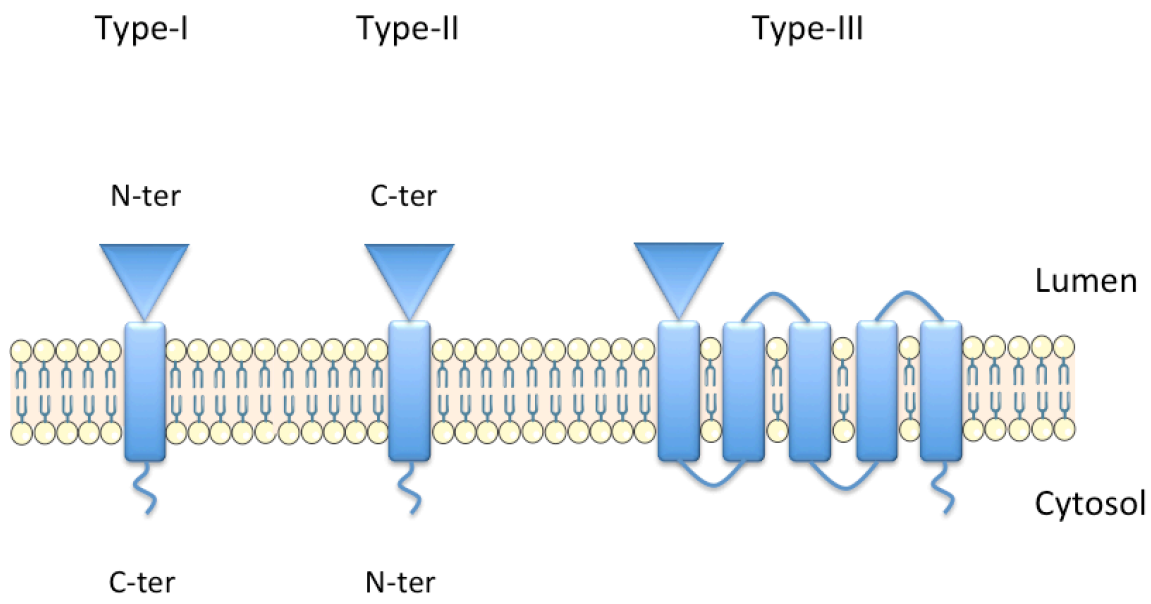

Figure 3: Organization of membrane proteins. Membrane proteins are formed by an N-terminal domain, one or several transmembrane domains crossing the lipid bilayer, and a C-terminal domain. Type-I membrane proteins have their N-terminal domain exposed in the ER lumen whereas the C-terminal part is in the cytosol. Type-II membrane proteins have their N-terminal part in the cytosol and the C-terminal part in the lumen. Finally, type-III membrane proteins cross a lipid bilayer several times, irrespective of the orientation.

An important feature that governs the orientation of TMDs across the lipid bilayer is the positive inside rule (Goder, Junne, & Spiess, 2004). Accordingly, over 85% of the cytosolically exposed domains of a membrane protein exhibit a net positive charge on the polypeptidic sequence flanking the TMD. By contrast, the luminal/extracellular domains are more likely to exhibit a net neutral or negative charge.

Several proteins are associated with the translocon (Figure 4). An important one is the Signal Peptidase which cleaves the signal peptide of most soluble and type-I membrane proteins, during the entry of the nascent polypeptide in the ER lumen. The Signal peptidase is a polytopic membrane protein, and belongs to the family of the Presenilin-Type Aspartic proteases. The cleavage of the signal peptide occurs inside the ER membrane (Weihofen, Binns, Lemberg, Ashman, & Martoglio, 2002).

Many proteins entering the lumen of the ER are N-glycosylated during the translocation process. The OligoSaccharyl Transferase (OST), a large 270 kiloDalton (kDa), multimeric and transmembrane protein complex, is responsible for this post-translational modification (Karamyshev et al., 2005). Briefly, a complex carbohydrate chain ($Glc_3Man_9GlcNAc_2$) is transferred from a lipidic dolichol pyrophosphate donor to the asparagine residue of a nascent polypeptide, in the context of a N-X-S/T acceptor site, where X can be any aminoacid (Kowarik et al., 2006). This ensures the proper folding and the stability of nascent proteins.

Proteins can also pass from the ER lumen to the cytosol in case of misfolding or misassembly, in a process called retrotranslocation. This mechanism is not discussed here.

Once synthesized, proteins destined to the secretion pathway are present in the ER, which is the first compartment of the secretion pathway

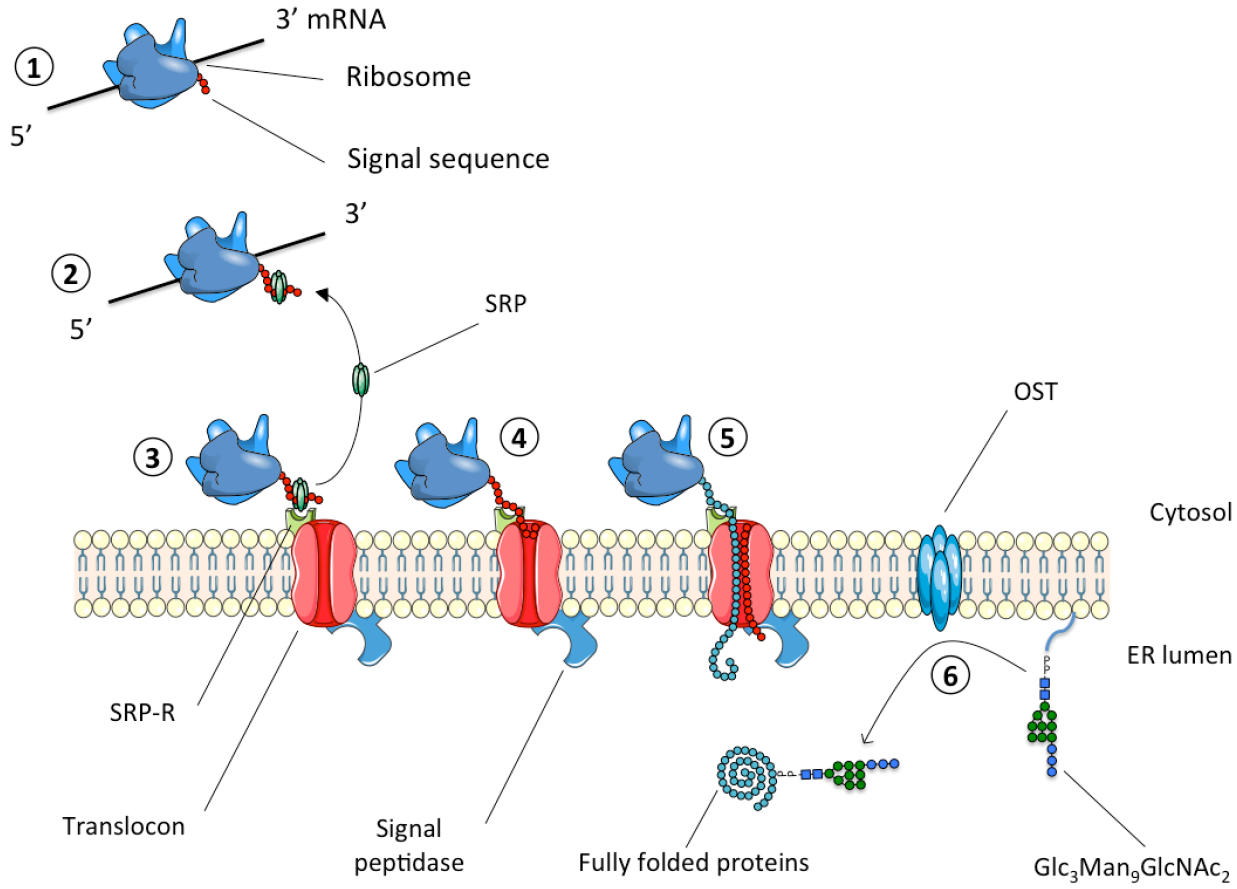

Figure 4: Translocon-associated proteins. The ribosome starts the translation of mRNA coding for secreted and membrane proteins in the cytosol (1). The signal sequence (red circles) is recognized by the Signal Recognition Particle (SRP, in green, 2), which drives the ribosome/mRNA/nascent polypeptide complex to the SRP-receptor (SRP-R), at the level of the ER, and more particularly close to the translocon machinery (red channel, 3). There, the translation is resumed (4) and the signal sequence is removed by the Signal Peptidase, in the lumen of the ER (5). During protein synthesis, the OligoSaccharylTransferase (OST) adds a pre-existing carbohydrate chain from a lipidic dolichol pyrophosphate to an asparagine residue of the nascent polypeptide (6).

2. The Endoplasmic Reticulum

The endoplasmic reticulum (ER) accounts for more than 50% of the total cellular membrane. The ER is composed of a highly organized network of tubules and cisternae, extending through the cytoplasm. The ER membrane is continuous with the outer membrane of the nucleus. Moreover, the ER interacts with microtubules, and this interaction maintains its shape and facilitates intracellular trafficking.

The main functions of the ER in mammalian cells are described in section 2.a. Proteins leave the ER at the level of Endoplasmic Reticulum Exit Sites described in the section 2.b, and reach the Endoplasmic Reticulum to Golgi Intermediate Compartment (ERGIC), described in section 2c.

a) Main functions of the endoplasmic reticulum

The ER is the site of many crucial reactions within the cells. As described in the first part of this chapter, the rough ER is the site of cotranslational insertion of soluble and transmembrane proteins destined to the exocytic pathway, and the place where they are glycosylated. The glycan moieties are also modified in the ER. It is also the place where nascent polypeptides are folded and assembled, with the help of many molecular chaperones. Enzymes including Peptidyl Propyl *cis/trans* Isomerases (PPIases) or members of the Hsp70 family (Heat-Shock Proteins 70) ensure the maintenance of a folding-competent state. Moreover, disulfide bond formation is catalysed by thiol:protein disulfide oxidoreductases and protein disulfide isomerases (PDIs) (Holtzman, 1997; Wilkinson & Gilbert, 2004).

The smooth ER is continuous with the rough ER but its membrane is devoid of ribosomes. It is the main site for synthesis of lipids notably phospholipids, glycolipids and cholesterol. In general, lipid biogenesis starts on the cytosolic face of the ER membrane, and the newly synthetized lipids are transferred into the luminal face of the ER membrane, with the help of flippases. Lipids are then transported along the secretion pathway.

Finally, the ER is the main calcium ion store of the cell, with a Ca^{2+} concentration 1'000 to 10'000 times higher than the cytosol. Calcium plays a key role in many intracellular signalling events and controls many cellular processes. A large part of intracellular signalling events thus involve the ER, and the control of calcium fluxes in and out of this organelle.

b) Endoplasmic Reticulum Exit Sites (ERES)

Once synthesized and correctly folded, proteins are allowed to progress through the secretion pathway. The exit out of the ER of correctly folded proteins takes place at specific sites called Endoplasmic Reticulum Exit Sites (ERES) or transitional ER (tER) (L. Orci et al., 1991; Palade, 1975). At these sites, the ER membrane is coated with COPII components (the cytosolic coat of vesicles responsible for anterograde transport, as described in the second part of this manuscript). In mammalian cells, these sites appear by fluorescence microscopy as largely immobile and long-lived punctae associated with the ER (Hammond & Glick, 2000; D. J. Stephens, Lin-Marq, Pagano, Pepperkok, & Paccaud, 2000).

The number and size of ERES can vary, depending on the cargo load that they have to deal with (Farhan, Weiss, Tani, Kaufman, & Hauri, 2008), underlining the plasticity of these structures. Secreted and membrane proteins are concentrated and then packaged into COPII-coated vesicles at the ERES and allowed to exit the ER. The next step of their journey to the plasma membrane is the passage in the Endoplasmic Reticulum to Golgi Intermediate Compartment (ERGIC).

c) The Endoplasmic Reticulum to Golgi Intermediate Compartment (ERGIC)

First described in 1984, the ERGIC, also known as the Vesiculo-Tubuluar Cluster (VTC) was initially proposed to be either a specialized sub-domain of the ER (Sitia & Meldolesi, 1992) or of the Golgi complex (Mellman & Simons, 1992), or an independent organelle within the cell. The ERGIC notably exhibits a different biochemical composition from both the ER and the Golgi (Schweizer et al., 1990). Its main roles are to concentrate the anterograde cargo proteins, and to control their quality and their sorting.

The ERGIC is identified by a high concentration of the mannose-binding receptor ERGIC-53. This protein is described later in this manuscript (See section C.2. c. 1). Both COPI and COPII coat components (described below) are found at the level of the ERGIC, suggesting that the ERGIC forms a platform for COPII-driven anterograde transport, as an intermediate to the route to the Golgi complex, and COPI-driven retrograde transport. But it is still under debate whether the ERGIC is a transient compartment rather than a stable compartment. Two models are today considered, the maturation model *versus* the stable compartment model (Figure 5).

- In the maturation model, the ERGIC is formed by the fusion of COPII-coated vesicles emanating from the ER. These transient clusters undergo homotypic fusion, and mature to eventually give rise to the *cis*-Golgi. However, this model is mostly based on the observation of the transport of an overexpressed viral protein, the thermosensitive mutant of the G protein from vesicular stomatis virus (tsO45-VSV-G) (D.J. Stephens & Pepperkok, 2001).

- In the stable compartment model, ERGIC receives lipids and proteins from the ER by COPII-coated vesicles forming an independent, stable organelle (C. Appenzeller-Herzog, 2006). Proteins are then thought to be transported to the Golgi complex possibly *via* COPI-mediated transport.

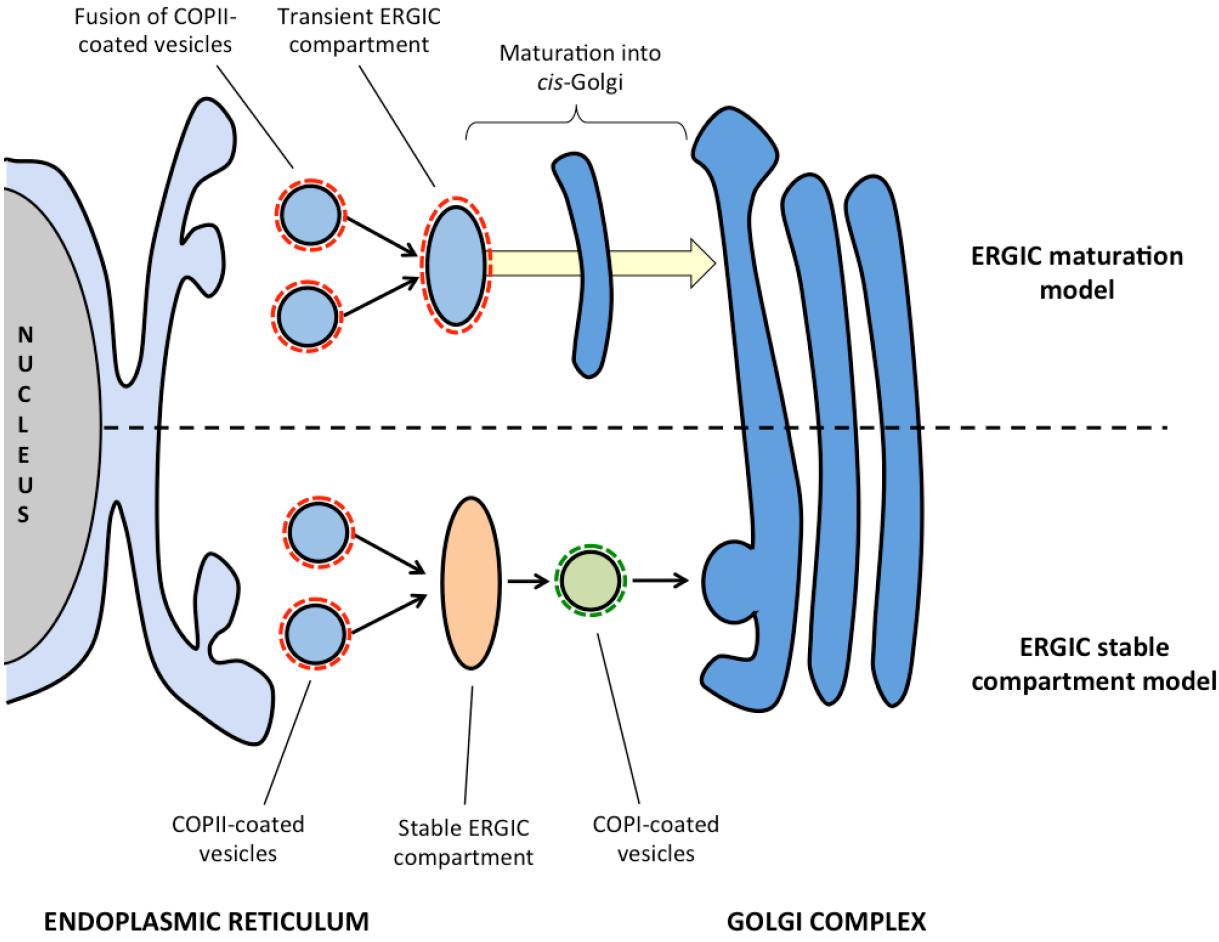

Figure 5: the ER-to-Golgi Intermediate Compartment. The ERGIC can be seen as a transient compartment (upper panel), in which COPII-coated vesicles (red coat) coming from the ER fuse between each other. This transient compartment then matures to eventually become the *cis*-Golgi itself. In the stable compartment model (lower panel), COPII-coated vesicles fuse and give birth to the ERGIC, as a stable compartment. From there, COPI-coated vesicles (green coat) transport lipids and proteins from the ERGIC to the *cis* face of the Golgi complex.

As mentioned before, the ERGIC is also the first sorting platform after ER exit. Escaped, ER-targeted proteins are recycled back *via* COPI-coated vesicles (Gaynor, Graham, & Emr, 1998; Letourneur et al., 1994; Pelham, 1994). On the contrary, secreted and membrane proteins are allowed to continue their journey to the Golgi complex.

In yeast, no ERGIC compartment has been formally described.

3. The Golgi complex

Proteins initially inserted in the ER can then be transported along the exocytic and endocytic pathways to reach many compartments in the cell, or to be secreted in the extracellular medium. Following exit from the ERGIC, the next compartment of the exocytic pathway is the Golgi complex.

The Golgi complex was first visualized by Camillo Golgi in 1898, and it was named after him (Bentivoglio, 1998). This organelle is at the crossroads of the ER, the cell surface, the endosomal and the lysosomal pathways. The Golgi complex is therefore a hub for collection and redistribution of proteins, as well as their modification and maturation, due notably to the presence of a large number of glycosylation enzymes.

In mammalian cells, the Golgi complex is a stack of cisternae, each delimited by a single membrane. It is attached to microtubules by motor proteins that ensure its localization close to the Microtubule Organizing Center (MTOC) near the cell nucleus. The flow of proteins leaving the ERGIC reaches the Golgi complex on its *cis* face, which can sometimes be closely apposed to the ER. Proteins are then transported through the *medial*-, and the *trans*-Golgi. They eventually leave the Golgi complex at the level of the

trans Golgi Network (TGN). A simplified scheme of the organization of the Golgi complex is proposed (Figure 6).

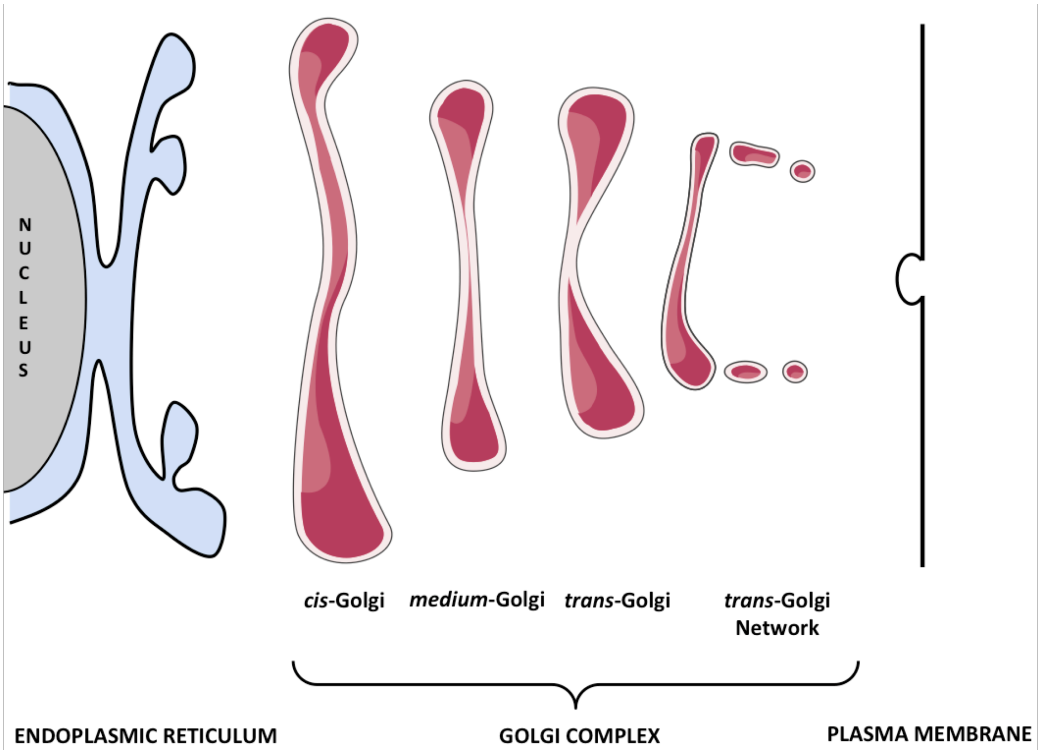

Figure 6: Organization of the Golgi complex. The flow of proteins and lipids exit the ER and reach the Golgi complex at its *cis* face. Transported molecules reach then the *medium*-Golgi and the *trans*-Golgi. Proteins and lipids exit the Golgi complex through the *Trans*-Golgi Network, a vast network of tubules and vesicles, where proteins are packaged and delivered to the plasma membrane, lysosomes or secretory granules for instance.

This high degree of compartmentalization is necessary to restrict enzymes in a specific place, since protein maturation is a regulated, multi-stepped process. Thus, enzymes responsible for the early modifications are found on the *cis* face, whereas the enzymes catalysing later modifications are found on the *trans*- face. The enzymes ensuring post-translational modifications, and notably glycosyltransferases and glycosidases are anchored in the membrane by a transmembrane domain.

The main functions of the Golgi complex are described in this section (a). Membrane transport within the Golgi complex is discussed next (b) and finally the events allowing proteins to exit the Golgi complex (c).

a) Main functions of the Golgi complex

The Golgi complex is a key organelle, found at the crossroad of the secretory, the endosomal and the lysosomal pathways. As it contains important sets of modification and glycosylation enzymes, it is a major site where secreted proteins undergo post-translational modifications. The Golgi complex is also a major hub for the packaging, the sorting and the shipping of functional proteins.

The functions of the Golgi complex are intimately linked to its structure. In mammalian cells, it is a membrane-bound organelle, composed of stacked, thin cisternae. As previously mentioned, the first acting enzymes are found on the *cis* face of the complex, whereas the late-acting enzymes are enriched on the *trans* side. For instance, the remodelling of the N-linked oligosaccharide starts with the trimming of the outermost mannose residues, and the addition of *N*-acetylglucosamine. The corresponding enzymes, the α1,3-1,6 mannosidase II and the β1,2 *N*-acetylglucosaminyltransferase II are found in the early cisternae of the Golgi complex. On the contrary, the addition of galactose residues or sialic acids (respectively by the β1,4 galactosyltransferase, the α1,6 sialyltransferase) occurs in the *trans*-Golgi or in the TGN (Kleene & Berger, 1993; Tommy Nilsson, Au, & Bergeron, 2009; Rabouille et al., 1995; Schoberer et al., 2010). The last modifications of proteins, such as addition of sialic acids or fucose residues typically take place in the TGN. Most of the glycosylation

enzymes in the Golgi complex are type II membrane proteins (Tommy Nilsson et al., 2009). On the contrary, the enzymes found in the TGN are often type I membrane proteins. Besides protein processing, lipids are also highly modified in the Golgi complex. This creates a lipid gradient where early Golgi compartments are depleted of modified lipids, and late compartments are enriched in these lipids, in particular sphingolipids (D'Angelo, Vicinanza, Di Campli, & De Matteis, 2008). Among other functions, the Golgi complex has been shown to be a minor calcium store, since it can pump Ca^{2+} and release it upon stimulation by inositol 1,4,5-trisphosphate (IP3) (Dolman & Tepikin, 2006; Missiaen, Dode, Vanoevelen, Raeymaekers, & Wuytack, 2007); to be a signalling platform, since the Ras/MAPK signalling pathway has been shown to be triggered from there, in response to growth factors such as epidermal growth factors (Chiu et al., 2002); to be involved in the Unfolded Protein Response (UPR) since the ATF6 (Activating Transcription Factor 6) needs to be processed in the Golgi before being functional (Ye et al., 2000), etc.

As previously mentioned, the Golgi complex is a key sorting platform. The membrane of the last *trans* stack, the TGN, exhibits regions coated with clathrin. Upon exit from the TGN, are then sent to their final destination, i.e. the plasma membrane, the endo/lysosomes, or secretion granules (De Matteis & Luini, 2008).

b) Anterograde transport in the Golgi complex

The Golgi complex is traversed constantly by a flow of proteins destined to be secreted, yet each Golgi cisterna maintains its biochemical identity (Glick & Luini, 2011). Two main models have been proposed to account for the anterograde transport

of proteins and lipids. The first is the 'vesicle-mediated transport' and the second is the 'cisternal maturation model'. It is important to note that the manner in which proteins and lipids are transported within the Golgi complex are still a subject hotly debated.

(1) Vesicle-mediated transport

In this model (see Figure 7, upper panel), Golgi cisternae are proposed to be stable compartments, each with its own biochemical composition. Secreted proteins are first modified in the *cis*-Golgi and are then transported to the next compartment by COPI-coated vesicles (C. L. Jackson, 2009; Palade, 1975; Pellett, Dietrich, Bewersdorf, Rothman, & Lavieu, 2013; Rothman, 1994; Rothman & Wieland, 1996). Each cisterna of the Golgi complex contains a specific set of enzymes that can modify secretory proteins. Resident Golgi enzymes are excluded from vesicles that transport the secretory cargo from one Golgi cisterna to the next. This model explains well how the biochemical identity of the stacks is maintained, but it doesn't explain the progression of larger cargoes that don't fit in COPI-coated vesicles. To account for this rare situation, it has been proposed that giant vesicles can form when necessary. It was originally proposed that COPI-coated vesicles were responsible for forward intra-Golgi transport (Rothman, 1994). Since evidence has accumulated that COPI-coated vesicles mainly transport proteins back from the Golgi to the ER (Duden, 2003), the role of COPI-coated vesicles in forward secretory transport is now more debated.

(2) Cisternal maturation

In the cisternal maturation model (see figure 7, lower panel), the *cis*-Golgi emerges from the homotypic fusion of vesicles derived from the ERGIC. This new compartment gradually matures and becomes successively a medium-Golgi, and trans-Golgi, and eventually a TGN cisterna. In the process, retrograde transport vesicles (presumably COPI-coated, see below) constantly recycle Golgi resident enzymes from the more mature compartment (e.g. the trans-Golgi cisterna) to the previous one (e.g. the medium-Golgi cisterna). The recycling rate, and therefore the identity of the different Golgi compartments, could be regulated by the activity of the Rab GTPases (Glick & Nakano, 2009). Direct evidence for the maturation of the cisternae has been obtained in yeast: taking advantage of the fact that the different Golgi cisternae are separated and can be visualized individually at the fluorescence level, the maturation of individual compartments was directly observed (Losev et al., 2006; Matsuura-Tokita, Takeuchi, Ichihara, Mikuriya, & Nakano, 2006). Besides, Golgi glycosylation enzymes have been demonstrated to be concentrated in COPI-coated carrier vesicles (Martinez-Menárguez et al., 2001). This latter result is however debated, since the enrichment of these enzymes in COPI-coated vesicles has not been observed by electron microscopy (L. Orci, Amherdt, Ravazzola, Perrelet, & Rothman, 2000).

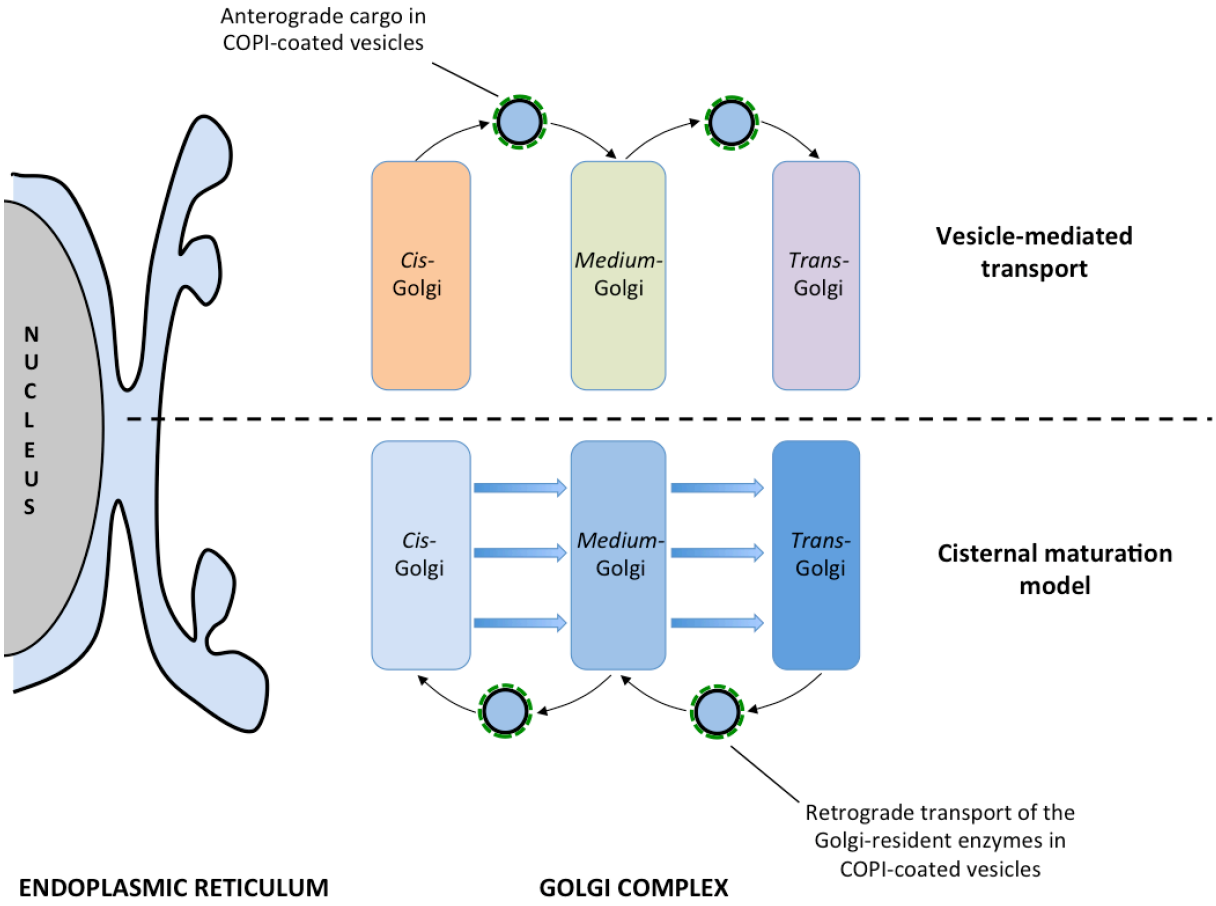

Figure 7: intra-Golgi trafficking. In the vesicle-mediated transport model (upper panel), incoming vesicles from the ER/ERGIC fuse with the *cis*-Golgi and proteins are modified there. Proteins are then transported to the *medium*- and then the *trans*-Golgi in COPI-mediated anterograde transport, for further modifications. Each Golgi stack maintains therefore its own biochemical content.
In the cisternae maturation transport model (lower panel), incoming vesicles reach the *cis*-Golgi, which gradually matures into *medium*- and eventually becomes the *trans*-Golgi. Golgi-resident enzymes are recycled back to the previous compartment in a COPI-mediated retrograde transport.

In summary, the mechanisms ensuring intra-Golgi trafficking are still the subject of an intense debate. The cisternal maturation transport is intellectually satisfying, but it is surprising that the Golgi enzymes were not seen to be concentrated in COPI-coated vesicles, to be recycled to the previous cistern. From a personal point of view, I would support the vesicle-mediated transport that corresponds more to the evidence available now.

Related to the subject of intra-Golgi transport, a recent study showed that the Golgi enzymes constantly and rapidly recycle through the ER, using an ER trapping assay that blocks recycled Golgi enzymes in the ER (Sengupta et al., 2015). These enzymes are sent back to the ER through tubule-shaped carriers, as it is described in the section B. 3. b) of this introduction. These results would be easy to integrate in the cisternae maturation transport, in which Golgi enzymes are constantly sent back to the previous cisternae. However, these results were recently contradicted (Villeneuve et al., 2017), where authors showed that the Golgi enzymes are not recycled back in the ER.

More studies are obviously needed to fully understand the protein trafficking within the Golgi complex.

c) Exit from the Golgi complex

Proteins having traversed the whole Golgi complex exit the most *trans* cisterna of the Golgi (the TGN) and can reach many intracellular destinations that are not the topic of this manuscript. Briefly, they can for instance reach the plasma membrane, concentrated into secretion granules or in constitutive secretion vesicles, or be transported to the endo/lysosomal pathway (De Matteis & Luini, 2008).

4. Unconventional secretion

Although a vast majority of secreted proteins reach their final destination through an ER/Golgi dependent pathway, a number of proteins are able to bypass this

classical secretion pathway and are delivered to their destination by a non-classical secretion or unconventional transport. These proteins, involved for instance in the inflammatory response or the cell survival, do not exhibit any signal sequence, and are delivered in a vesicular or non-vesicular manner. Both soluble and membrane proteins can be secreted by an unconventional secretion. At least four types of unconventional secretion have been described:

- Cytosolic proteins can be transported directly across the plasma membrane, as proposed for the fibroblast growth factor 2 (FGF2) (Nickel, 2011). It is speculated that FGF2 organizes as an hexamer which depends on phosphatidyl(4,5)-biphosphate ($PI(4,5)P_2$), and forms a ring-like structure opening a transient lipid pore (Steringer et al., 2012), and is then released in the extracellular medium.

- Direct export of proteins by an ATP-Binding Cassette (ABC) transporter, like the soluble yeast pheromone a-factor that is actively transported by the Ste6p ABC transporter (McGrath & Varshavsky, 1989; Michaelis, 1993).

- Export of proteins within vesicle-like particles, for instance derived from lysosomes or multivesicular bodies, as has been described for the secretion of the interleukin 1β (IL-1β) (C. Andrei et al., 1999; Cristina Andrei et al., 2004; MacKenzie et al., 2001; Qu, Franchi, Nunez, & Dubyak, 2007; Rubartelli, Cozzolino, Talio, & Sitia, 1990).

- Bypass of the Golgi complex of proteins, to directly reach the plasma membrane, as it as been described for instance for a mutant form of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) (Gee, Noh, Tang, Kim, & Lee, 2011; Yoo et al., 2002). Indeed, CFTR is transported to the surface in absence of the regulatory elements of the COPII-coated vesicle formation and is thought to reach the surface through the recycling endosome (Nickel & Rabouille, 2009).

B. Transport intermediates in the early secretion pathway

Along the early secretion pathway, proteins are transported through small, membrane-delimited carriers named vesicles. The vesicular transport transports the lipids and proteins between the different compartments of the secretory pathway in eukaryotic cells. Basically, proteins are packaged into a nascent, budding vesicle, which separates from the donor compartment in a process by membrane fission. The vesicle is then targeted to an acceptor compartment with which its membrane fuses.

Many proteins are involved in the formation of transport vesicles at various steps of the secretory pathway. However, the fundamental mechanisms leading to the formation of vesicles are shared between all transport vesicles. Therefore, the common processes leading to vesicle biogenesis will be discussed in a first part (1). Molecular mechanisms specifically involved in the anterograde transport will then be presented (2), before the ones implicated in retrograde transport (3).

1. Vesicular trafficking – Common mechanisms between all transport vesicles

As mentioned above, different proteins participate in vesicle formation, depending on the localization of the vesicle. However, the molecular mechanisms governing the initiation of vesicle formation, the vesicle budding, the detachment of vesicle from the donor compartment, the targeting and the fusion of the vesicle to its target compartment, are common between all the transport vesicles and are described in the following paragraphs.

a) Vesicle formation

Vesicle formation starts with the self-assembly of **co**at **p**roteins (COP), recruited from the cytosol to the membrane of the donor compartment. The polymerization of coat proteins drives the deformation of the membrane, from a flat patch to a spherical bud. Small guanine triphosphatases (GTPases) regulate the polymerization or the depolymerization of coat proteins from membranes, according to their activation state. Small GTPases are considered as binary switches, as the GTP-bound form drives the assembly of the coat proteins (active form), while the GDP-bound form induces the disassembly of the coat (inactive form). GTPases cycle between an active and inactive state thanks to the regulators GEFs (guanine exchange factors) and GAPs (GTPase activating proteins) (Nie, Hirsch, & Randazzo, 2003). GEFs catalyse the exchange of GDP with GTP, inducing a conformational change, making small GTPases active. GAPs catalyse the hydrolysis of the GTP into GDP, turning off the small GTPases activity. The small GTPase cycle is presented Figure 8, with the example of Sar1, the small GTPase triggering the formation of COPII-coated vesicles, as explained below.

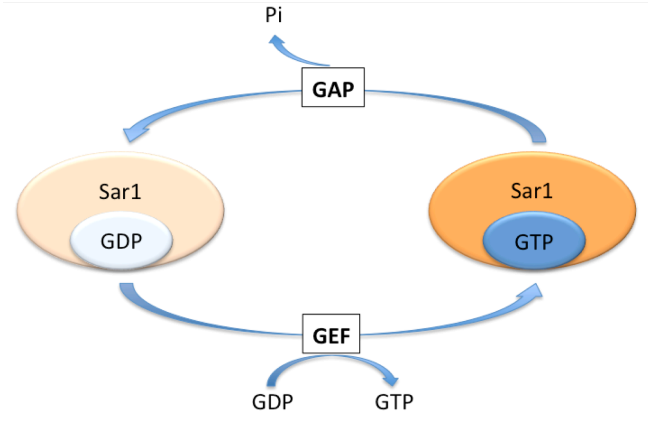

Figure 8: small GTPase cycle, the example of Sar 1. The Guanine Exchange Factor (GEF) exchanges GDP associated with Sar1 with cytosolic GTP. This turns the inactive Sar1 into its active state. On the contrary, the GTPase activating protein (GAP) stimulates the hydrolysis of the GTP by Sar1, bringing Sar1 back to its inactive state.

Once activated, small GTPases recruit COP components to the flat, pre-budding membrane, inducing a deformation of the membrane that ultimately gives birth to the vesicle.

A lot of efforts have been made to identify proteins involved in the scission of vesicles, i.e. the release of the vesicle from the donor compartment. However, no specific effectors have been identified so far. Indeed, vesicles were described to be able to bud in vitro simply in the presence of purified coat components (Matsuoka et al., 1998). It is therefore thought that the vesicle fission is driven by the polymerization of the coat itself (Juan S. Bonifacino & Glick, 2004; Kirchhausen, 2000). A more recent study showed that the vesicle fission was not dependent on GTP hydrolysis from the small GTPases (Adolf et al., 2013).

After budding of the vesicle, the coat of the newly formed vesicle is removed, most probably to facilitate the fusion with the acceptor compartment. The hydrolysis of GTP achieved by the small GTPases leads to the disassembly of the coat just after the scission, allowing the retrieval of the coat components for another round of budding. Recent studies suggest that the depolymerisation of the coat components (complete or not) may occur later than initially believed, at least until the initiation of the tethering with the target compartment (Trahey & Hay, 2010).

b) Vesicle targeting and fusion

Vesicle targeting and fusion is a highly regulated process. It involves notably tethering factors called **s**oluble **N**-ethylmaleimide-sensitive factor **a**ttachment protein **re**ceptors (SNAREs). Most of the 38 SNARE proteins found in the human genome are transmembrane proteins, containing a linker domain, one or two SNARE motifs (approximately 70 aminoacids in size) and are present at the surface of either vesicles or their target compartment. The very first SNAREs identified were Syntaxin1, SNAP25 and VAMP proteins, and were described in neurons in the late nineties. Most of the SNAREs are specific to cellular compartments.

Originally, they were classified as v-SNAREs (located at the surface of vesicles) and t-SNAREs (located at the surface of the targeted compartment). The interaction of a v-SNARE with a t-SNARE induces the so-called *trans*-SNARE complex, where four SNARE motifs form a coiled-coil motif, bringing two membranes closer and eventually contributing to their fusion. SNAREs are now classified based on their crystal structure (Fasshauer, Sutton, Brunger, & Jahn, 1998) and the observation that SNARE motifs contain either an arginine residue (R) or a glutamine residue (Q) at a critical position. Therefore, the four SNARE motifs leading to a functional SNARE complex are composed by one R-SNARE motif and three Q-SNARE motifs. With a few exceptions, most of the R-SNAREs are v-SNAREs and most of the Q-SNAREs are t-SNAREs.

The SNARE motifs drive the formation of the SNARE complex. Although a SNARE motif is unstructured in its monomeric form, it becomes highly stable when associated with other SNAREs motifs. In the current model, the assembly of SNAREs is in a 'zipper' manner, from the N-terminal to the C-terminal end, which brings the membrane from the vesicle and from the targeted compartment in close contact, initiating therefore their

fusion (Jahn & Scheller, 2006; Lin & Scheller, 1997; Otto, Hanson, & Jahn, 1997). After a series of intermediate steps where membranes merge in a stepwise manner (Figure 9), an aqueous pore is formed, which connects the lumen of the vesicle and of the targeted compartment. Briefly, the assembly of the SNARE complex exerts a mechanical force on the membranes, transmitted by the rigidity of the linker domain. This induces the bending and the deformation of the membrane in close contact, facilitating the formation of fusion stalks (Risselada & Grubmüller, 2012).

After fusion, assembled SNAREs form a stable complex in the resulting fused membrane, and must be mechanically separated. The disassembly is performed by NSF (**N**-ethylmaleimide-**s**ensitive **f**actor), together with the α-SNAP (**s**oluble **N**FS **a**ttachment **p**roteins) cofactor, which both bind the SNARE complex (Ryu et al., 2015; Söllner, Bennett, Whiteheart, Scheller, & Rothman, 1993; Zhao et al., 2015). When bound to ATP, NSF exhibits a split washer shape (Zhao et al., 2015) that becomes flat when bound to

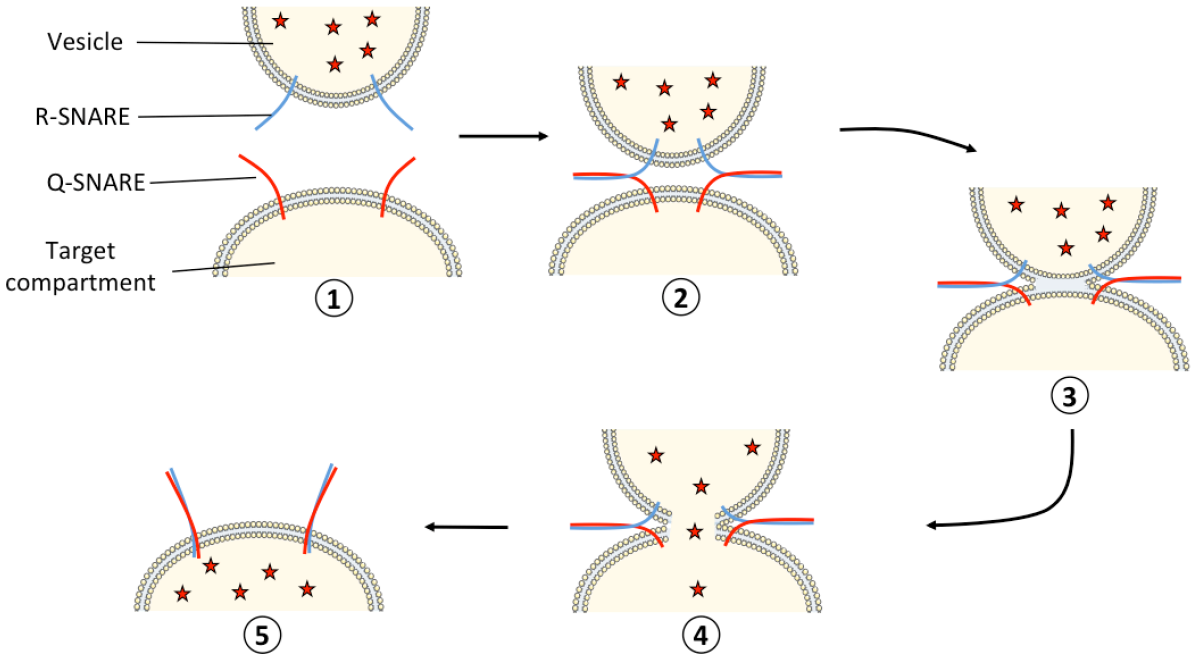

Figure 9: SNAREs-mediated vesicle fusion. (1) R-SNAREs are on the vesicle whereas Q-SNAREs are on the target compartment. (2) SNARE proteins interact with each other, in a 'zipper' manner, forming a *trans*-SNARE complex. (3) Membranes of the vesicle and the target compartment are close enough to initiate their fusion. (4) A fusion pore is formed, allowing the delivery of the vesicle content in the target compartment. (5) The vesicle completely fused with the target compartment.

ADP. This movement induces a conformational change of α-SNAP, forcing a shear motion, which disassembles the SNARE complex. All the proteins involved are then ready for a new cycle of fusion.

2. Anterograde transport – COPII-coated vesicles

Secretory or membrane proteins that are transported from the ERES to subsequent compartments travel in vesicles surrounded by a specific proteinaceous coat named COPII (for Coat Protein II). Soluble or membrane proteins as well as lipids transported by COPII-coated vesicles are usually referred to as clients or cargoes. Direct interactions of cargoes with COPII coat components, or indirect interactions mediated by cargo receptors or chaperones allow sorting of cargo proteins from ER resident proteins and at the same time drive the assembly of the COPII coat (Aridor, Bannykh, Rowe, & Balch, 1999; Venditti, Wilson, & De Matteis, 2014).

In the first part I described the proteins that form the COPII coat. The second part describes an alternative mechanism allowing ER exit of cargoes unable to enter COPII-coated vesicles.

a) COPII coat components

COPII-coated vesicles are responsible for the anterograde transport of proteins, from the ER to the Golgi complex. First described in yeast (C. Barlowe et al., 1994), COPII-coated vesicles incorporate ER soluble and membrane proteins into 60-80nm vesicles. The COPII coat is made of five proteins, recruited to the ER membrane (Figure

10). Polymerization of the COPII coat is initiated when the small GTPase Sar1 (secretion-associated RAS-related 1) is activated. This reaction is catalysed by the ER-membrane anchored Sec12 (Charles Barlowe, 2003). Sec12 is a GTP-Exchange Factor (GEF) and as such it allows the replacement of a GDP with a GTP in the nucleotide-binding site of Sar1. This induces a conformational change in Sar1, exposing its N-terminal amphipathic α-helix, which inserts into the ER membrane. Activated, ER-bound GTP-Sar1 imposes a membrane curvature and promotes the recruitment of the Sec23/Sec24 protein complex (Bi, Corpina, & Goldberg, 2002; Lee et al., 2005), therefore driving the formation of the inner coat. The Sec23/Sec24 complex exhibits a bowtie-shape, leading to the hypothesis that it could also contribute to the curvature imposed to the membrane. The Sec23/Sec24 complex is also implicated in the selection of the cargo to be transported, as discussed below. Sec24 binds and selects cargo proteins destined to be transported to the Golgi complex. Sec23 exhibits a Sar1-GAP activity, which catalyses

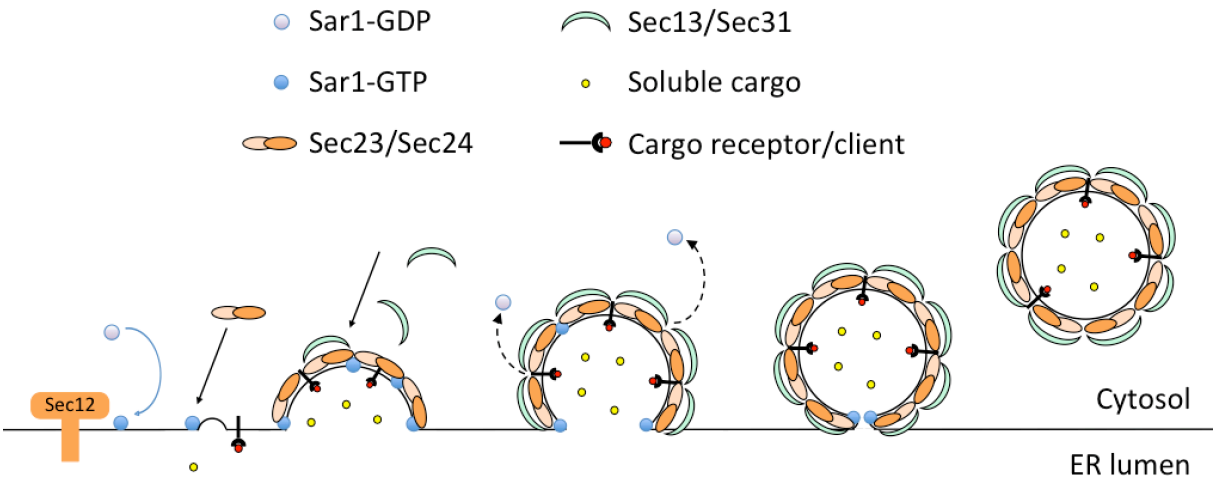

Figure 10: formation of COPII-coated vesicles. Sec12 promotes the exchange of the GDP of Sar1 with GTP. Sar1 becomes therefore activated and is inserted into the ER lipid bilayer. This provokes the curvature of the membrane and recruits the Sec23/Sec24 complex from the cytosol, forming the inner coat of COPII-coated vesicles. The Sec13/Sec31 polymerizes then upon the inner layer, stabilizing the complex and completing the membrane curvature. This is thought to provoke the release of the vesicle from the ER.

the hydrolysis of the GTP bound to Sec12. The Sec13/Sec31 complex polymerizes then and forms the outer layer of the COPII coat. The Sec13/Sec31 complex is formed by two Sec13 subunits and two Sec31 subunits. The outer coat interacts with the inner coat thanks to the proline-rich domain of Sec31, stabilizes the cargo-bound pre-budding Sec23/Sec24 complex, and completes the membrane curvature (Bi, Mancias, & Goldberg, 2007), leading to the release of the vesicle.

b) Non-canonical transport of large cargoes

Most of the secreted proteins are small enough to be packaged into 60-80nm COPII-coated vesicles that transport them from the ER to the Golgi complex. Some protein assemblies such as procollagen fibers are however too large (>300nm). *In vitro* studies consisting in the microinjection of a cDNA of a dominant-negative form of the small GTPase Sar1 were demonstrated to abort the transport of procollagen. This highlights the involvement of COPII coat components in this process (Mironov et al., 2003; David J. Stephens & Pepperkok, 2002). It is however not known if this effect is direct or not. Indeed, it is still to be demonstrated if COPII coat components are flexible enough to surround such large molecules, or if they participate in this process in another way.

3. Retrograde transport – COPI-coated vesicles

ER-resident proteins which mature in the Golgi complex, proteins involved in the vesicular transport or escaped proteins need to be sent back to the ER. This is mainly

achieved through COPI-coated vesicles, and the components that form these vesicles are firstly described. A COPI-independent retrograde transport also exists and is described in a second part.

a) COPI coat components

COPI-coated vesicles are responsible for retrograde transport, i.e. the transport from the Golgi complex to the ER. COPI-coated vesicles may also be involved in transport within the Golgi complex. These vesicles package soluble and membrane proteins into 75-nm diameter vesicles. The coatomer is a 600kDa protein complex which forms the coat of COPI vesicles. This complex is composed of seven subunits: α-, β-, β'-, γ-, δ-, ε-, and ζ-COP. Originally thought to be recruited en bloc, the coatomer is actually recruited in two layers. The inner coat made by γ-, δ-, ζ- and β-COP (or F-subcomplex) binds the cargo, and the outer coat is formed by the α-, β'- and ε-COP subunits (or B-subcomplex) (McMahon & Mills, 2004). The formation of the COPI-coated vesicles is summarized in Figure 11 and is initiated by the Arf1 (adenosine diphosphate-ribosylation factor 1) small GTPase. In the inactive form (GDP form), Arf1 can indirectly bind membranes, through an interaction with the p23/p24 proteins family (Beck, Adolf, Weimer, Bruegger, & Wieland, 2009). After activation by a GEF (GBF1 at the *cis*-Golgi for instance), Arf1-GTP undergoes a conformational change, which exposes a N-terminal amphipathic helix as well as a myristoyl-modified glycine residue, allowing the insertion in the lipid bilayer. Arf1-GTP recruits then the inner layer proteins, which mediate the incorporation of specific cargoes in the vesicles. The outer layer polymerizes then, contributing to the deformation and the fission of the vesicle. The exact mechanism of fission is still not totally understood. Hydrolysis of the GTP by ArfGAP provokes the

release of the coatomer from the vesicle. COPI components are then free to form a new vesicle.

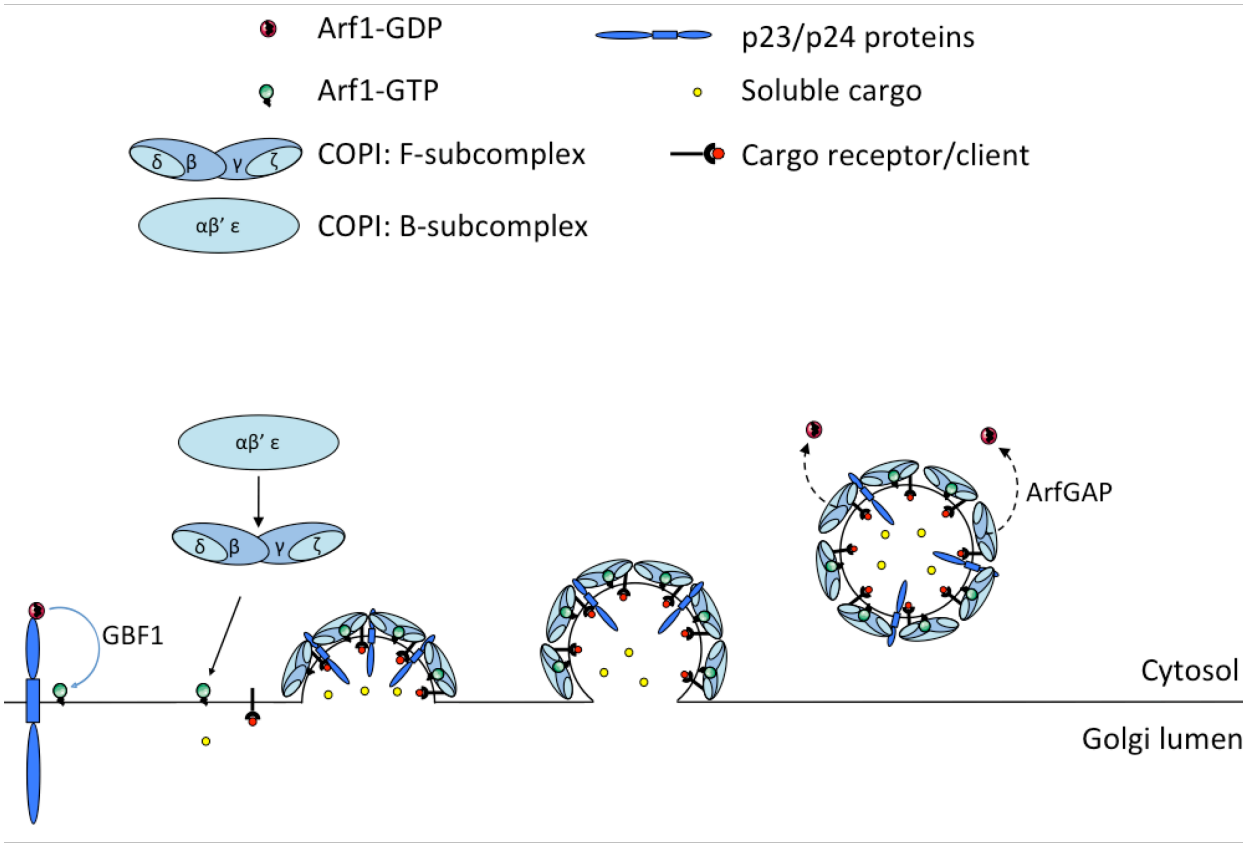

Figure 11: formation of COPI-coated vesicles. The formation of COPI-coated vesicles is initiated when Arf1 becomes activated. Inactivated (GDP-bound) Arf1 indirectly binds membranes through its interaction with p23/p24 proteins family and is inserted into the membrane bilayer once activated. Arf1-GTP recruits then the F-subcomplex of the COPI coat, formed by the γ-, δ-, ζ- and β-COP subunits. The B-subcomplex, formed by the α-, β'- and ε-COP subunits is then recruited, allowing the release of the vesicle from the donor compartment, for instance the Golgi complex. Arf1 is released from the vesicle after hydrolysis of the GTP by Arf1, upon stimulation by an ArfGAP.

As previously mentioned, COPI-coated vesicles are also thought to be involved in the anterograde transport, from the ERGIC to the Golgi complex, or within the Golgi complex. COPI-coated vesicles may perform bidirectional transport, moving proteins forward or backward, as suggested by Lelio Orci in 1997 (L. Orci et al., 1997). Two distinct populations of COPI-coated vesicles may be here responsible for the

anterograde and retrograde transport. This anterograde transport through COPI-coated vesicles may be explained by the action of different Arf isoforms (Volpicelli-Daley, 2005) or different isoforms of COPI components (Wegmann, Hess, Baier, Wieland, & Reinhard, 2004).

b) COPI-independent retrograde transport

Experiments performed in the late nineties brought to light the existence of an alternative route, which is independent from COPI-coated vesicles. Indeed, the inhibition of the COPI machinery by the microinjection of antibodies directed against COPI components or the expression of a dominant-negative form of Arf1, did not inhibit the retrograde transport of the Shiga toxin, or glycosyltransferases from the Golgi complex to the ER (Girod et al., 1999; Storrie, Pepperkok, & Nilsson, 2000). The transport of these proteins was then shown to be abrogated if an inactive form of the Ras-related proteins 6 (Rab6) GTPase was expressed. It has been hypothesized that this COPI-independent transport involves the tubulation of Golgi membrane from the *cis*-Golgi, without the recruitment of any coat.

It was hypothesized that the shape of the tubules confers a high surface for a low volume, limiting therefore the access for soluble cargoes, and favouring the retrieval of membrane proteins (Charles Barlowe & Helenius, 2016).

Although this COPI-independent retrograde pathway has been identified many years ago, little is known about the molecular mechanisms governing this process.

C. Sorting of proteins in the early secretion pathway: general principles

The preceding chapters describe the synthesis and the transport of proteins through the secretion pathway. Although the secretory pathway is constantly traversed by a flow of secreted proteins, constituents of the Endoplasmic Reticulum (e.g. the SRP receptor) or of the Golgi complex (e.g. specific glycosyltransferases or glycosidases) remain mostly in the organelle where their function is required. To achieve this specific localization, ER or Golgi proteins must indeed be separated from the secreted proteins, and from each other. This section is devoted to the description of the general mechanisms that ensure the proper intracellular transport, sorting and localization of individual proteins.

For this purpose, proteins exhibit within their primary, secondary, tertiary or quaternary structure targeting signals, also referred to as sorting motifs. These motifs can be compared to a postal address, allowing each protein to be delivered to the proper organelle. These sorting motifs are recognized by the transport machinery of the cell (e.g. COPI and COPII components) directly or indirectly through specific receptors. In this third chapter the molecular mechanisms which ensure the specificity of the protein transport in the secretion pathway are described. The aim of this section is not to make an extensive study of all the sorting processes, but to present a few examples illustrating the main sorting strategies.

Protein transport and sorting is achieved through three main modalities:

- Exclusion from transport vesicles to achieve retention in a compartment,
- Concentration in a transport vesicle through a direct interaction with COP-coat proteins or specific interaction with a cargo receptor,

- Bulk flow transport refers to situations where a protein is present in transport vesicles at a concentration identical to that found in the donor compartment.

These processes are highly controlled, and the localization of a protein to its final destination is not the result of a single sorting event but is rather caused by the additive effect of several sorting motifs and mechanisms.

In this part I first describe the mechanisms leading to the exclusion of proteins from transport vesicles. This process is exemplified here by the example of ER retention mediated by chaperones. The second part of this chapter describes the concentration of proteins into vesicles, first at the ER/Golgi interface, and then in the Golgi complex. The final part of this chapter deals with bulk flow transport.

The sorting mechanisms controlling the localization of TMDs themselves are the topic of the last chapter of this introduction (See Chapter D).

1. Exclusion from transport vesicles, the example of chaperone-mediated ER localization mechanisms

The first sorting and targeting event that proteins face is related to their folding status. If not completely folded, proteins are localized in the ER through multiple mechanisms involving several chaperones. Correctly folded proteins are allowed to progress through the secretion pathway whereas unfolded proteins are specifically recognized and excluded from COPII-coated vesicles. As a result, these proteins are

localized in the ER until they reach their final conformation. Retained unfolded proteins are eventually degraded.

Recognition of unfolded proteins is achieved by chaperones and relies on two different strategies: the recognition of unprocessed carbohydrate moieties by the Calnexin/Calreticulin system, and the binding of unfolded regions by BiP. The former is described in a first part and the latter in the second. These two strategies are complementary.

a) Calnexin/Calreticulin cycle

Newly synthetized proteins, once glycosylated, directly undergo the processing of their carbohydrate chains. This process occurs in the ER and is achieved by the glucosidase I and II, both of them removing the outermost glucose residues, leaving therefore a $Glc_1Man_9GlcNAc_2$ chain. Calnexin and Calreticulin, two lectins that reside in the ER, specifically recognize this structure. Calnexin is a transmembrane protein whereas Calreticulin is a luminal protein. The remaining glucose is then hydrolysed by the glucosidase II, disrupting therefore the interaction with the Calnexin/Calreticulin complex. If the protein is not correctly folded, it becomes the substrate for UDP-glucose:glycoprotein glycosyltransferase, a conformational sensor protein which adds back a glucose residue to the glycan moiety. This promotes the re-interaction with the Calnexin/Calreticulin complex, until the protein reaches its native form. These two chaperones also interact with the thiol-disulfide oxidoreductase ERp57 responsible for disulfide bonds formation. Permanently misfolded proteins see their oligosaccharidic moiety cleaved by ER α 1,2-mannosidase I, inducing their expulsion in the cytosol

followed by their degradation in the proteasome. On the other hand, glucose-free glycoproteins, which reach their native form, are allowed to continue their journey through the secretion pathway. This cycle between the Calnexin/Calreticulin complex, the glucosidase II and the UDP-glucose:glycosyltransferase participates in the conformational quality control which occurs in the ER (Caramelo & Parodi, 2008; Ellgaard & Helenius, 2003). Indeed, as a result, unfolded proteins are not packaged into COPII-coated vesicles and are therefore retained in the ER. A simplified view of this process is presented in Figure 12.

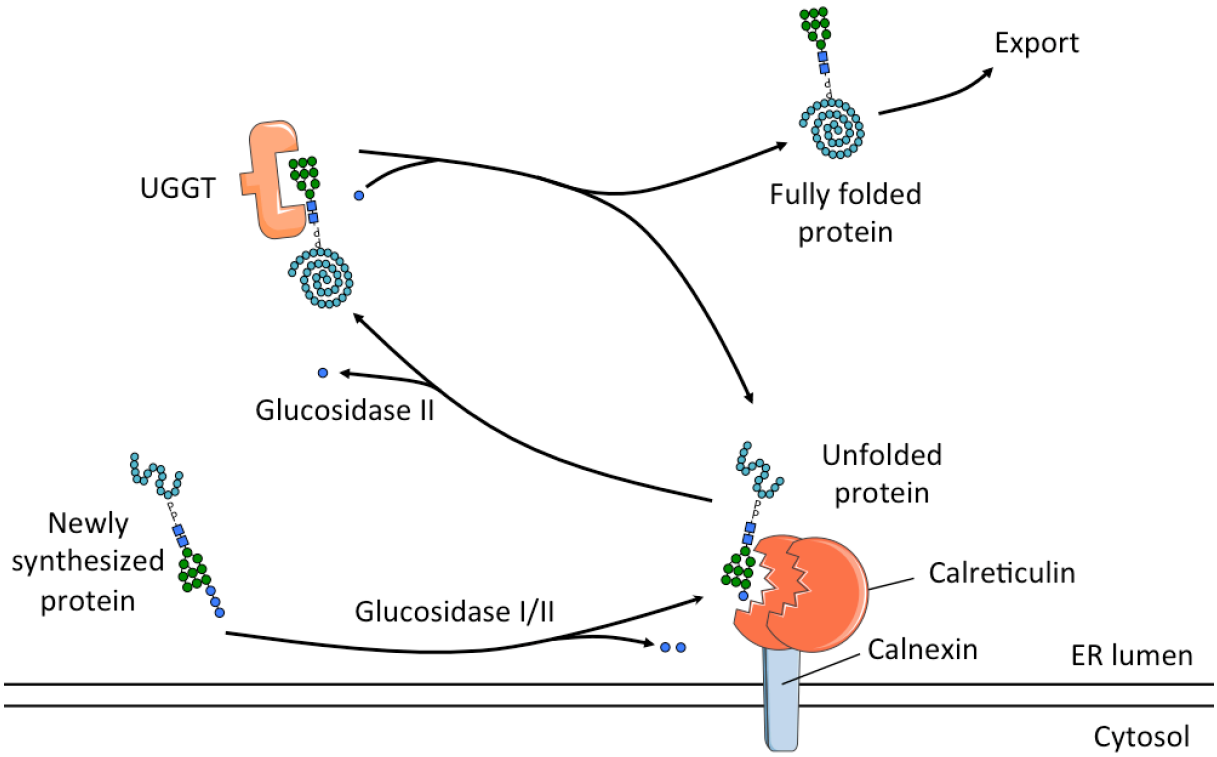

Figure 12: the Calnexin/calreticulin cycle. While they are synthetized, proteins are glycosylated. The carbohydrate chain is directly processed by the Glucosidase-I and –II, which remove the two outermost glucose residues (blue circles). The remaining $Glc_1Man_9GlcNAc_2$ chain is specifically recognized by the transmembrane Calnexin, in complex with the luminal Calreticulin. If the protein is correctly folded, the remaining glucose residue is removed by the Glucosidase-II, which disrupts the association with the Calnexin/Calreticulin, and promotes the association with ERGIC-53, as described later. This leads to the anterograde transport of the newly synthesized protein. On the contrary, if the protein is not correctly folded after hydrolysis of the last glucose residue, it becomes the substrate for the UDP-glucose:glycosyltransferase (UGGT). This enzyme adds a glucose residue to the carbohydrate chain, which allows another round of interaction with the Calnexin/Calreticulin, that promotes the correct folding of the protein. Permanently misfolded proteins are sent to the cytosol for degradation.

b) BiP

Since most of the newly synthetized proteins are glycoproteins, a lot of efforts have been made to decipher mechanisms leading to the correct folding of glycoproteins. However, it is also shown that mechanisms also exist to assist non-glycosylated proteins.

In this context, Binding immunoglobulin Protein (BiP), also denominated 78kDa glucose-regulated protein (GRP-78) binds to newly synthetized polypeptides, prevents them from aggregating. Binding of BiP excludes unfolded proteins from COPII-coated vesicles. BiP belongs to the family of Hsp70 proteins and is composed of two domains: a N-terminal nucleotide-binding domain (NBD) with an intrinsic low ATP hydrolysis activity and a C-terminal substrate-binding domain (SBD), which binds to nascent polypeptides (Otero, Lizák, & Hendershot, 2010; Yang, Nune, Zong, Zhou, & Liu, 2015). More precisely, the SBD of BiP is bipartite, with a binding pocket that is the acceptor site for the nascent polypeptide and a helical lid to close the binding cavity (Zhuravleva & Gierasch, 2015). BiP undergoes a complex allosteric regulation, depending on the nucleotide bound to the NBD. When bound to ATP, the protein is in an inactive state, with a low affinity for unfolded substrates, and the lid is open, allowing the interaction with its substrate. Here, the substrate can be brought to the SBD by several co-chaperones, which increase the ATPase activity of BiP (Behnke, Feige, & Hendershot, 2015). Upon ATP hydrolysis, a conformational change occurs, which greatly enhances the affinity of the SBD for its substrate. Besides, this change closes the lid on the bound substrate. This process prevents the substrate from aggregating or improperly folding. Finally, the exchange of the ADP for an ATP on the NBD releases the substrate from BiP, now available for another round of reaction.

The ADP/ATP cycle of BiP is regulated by Nucleotide Exchange factors (NEFs) and regulates the release of the substrate, while co-chaperones (Hsp40 family) increase the ATPase activity of the NBD. Some of these co-chaperones have also the ability to bind polypeptidic substrates and to transfer them to BiP.

The two mechanisms described above are part of the quality control of the ER. Substrates for the Calnexin/Calreticulin cycle or for BiP are recognized by these mechanisms, prevented to enter COPII-coated vesicles and are therefore retained in the ER. In this way, these quality control mechanisms participate in the protein sorting, ensuring the exclusion from COPII-coated vesicles of unfolded proteins. Other exclusion mechanisms exist but are not presented here.

2. Concentration of proteins into transport vesicles

Differential sorting can result from the exclusion or from the concentration of sorted proteins in vesicles. This leads to the selective transport of proteins from one compartment to another. In this section I first describe the mechanisms by which proteins are actively concentrated into COPII-coated vesicles, leading to a transport from the ER to the Golgi complex. The second part of this section deals with the concentration of protein into COPI-coated vesicles, associated with the transport from the Golgi complex to the ER. In the last part of this section, the mechanisms governing the concentration of proteins for localization in the Golgi complex are presented.

a) Mechanisms ensuring concentration of proteins into anterograde transport vesicles at the ER to Golgi interface

Proteins can be incorporated and concentrated into COPII-coated vesicles due to direct interactions between sorted proteins and COPII-coat components. This is the subject of the first part. The second part of this chapter is devoted to the indirect interaction between sorted proteins and COPII-coat components. This process is mediated by cargo receptors.

(1) Concentration of proteins into COPII-coated vesicles through direct interaction with coat proteins

It has been extensively shown that some cargoes (or clients) directly interact with the components of the COPII coat. This allows the concentration of the cargo at the ERES, where the budding of COPII-coated vesicles occurs. For instance, the di-acidic motif DXE (where X is any aminoacid) first identified on the cytosolic domain of the type I transmembrane protein vesicular stomatis virus glycoprotein (VSV-G), was shown to bind directly Sec24 (Miller et al., 2003; Mossessova, Bickford, & Goldberg, 2003). This allows the specific capture and packaging of transmembrane proteins containing this sorting motif into COPII-coated vesicles, moving them forward in the secretion pathway. Several isoforms of Sec24 have been described (Pagano et al., 1999), and it has been suggested that they may bind a variety of different motifs and cargoes. Moreover, the DXE motif was shown to restrict the access to vesicles ensuring Golgi-to-ER recycling (Fossati, Colombo, & Borgese, 2014). As a consequence, a motif for the anterograde

transport may promote the unidirectional transport through the secretion pathway, preventing their recycling by retrograde transport.

Other motifs have been described that bind other components of the COPII coat, such as the di-phenylalanine (FF) motif in the C-terminal part of p24, which interacts with Sec23 (Dominguez et al., 1998).

(2) Concentration of proteins into COPII-coated vesicles through a cargo receptor

Many proteins do not interact directly with COPII components but rather are taken in charge by cargo receptors, which then concentrate them into transport vesicles. In this case, the receptor is transported with its cargo, and releases it in the post-ER compartment. The empty receptor is then retrieved to the ER for another round of transport.

As described in the next paragraphs, different motifs are recognized by cargo receptors. Peptidic motifs can be found in the lumenal or cytosolic domains. Alternatively, post-translational modifications (such as glycosylation) can be recognized by the sorting machinery.

(a) ERGIC-53

ERGIC-53 is a type-I membrane protein of 53kDa, used as a marker for the ERGIC, but is also found in the ER and in the *cis*-Golgi. ERGIC-53 is a mannose-specific binding lectin, composed of a calcium- and pH-sensitive carbohydrate recognition domain

(CRD), a transmembrane domain and a short cytoplasmic domain (Velloso, Svensson, Schneider, Pettersson, & Lindqvist, 2002). The cytosolic domain of ERGIC-53 presents a KKFF (di-lysine, di-phenylalanine motifs) in position -4, -3, -2 and -1. As mentioned above, the di-phenylalanine is an ER-export motif whereas the di-lysine motif is an ER-localization motif (Y. C. Zhang, Zhou, Yang, & Xiong, 2009). Therefore, ERGIC-53 cycles between the ER, the ERGIC and the *cis*-Golgi.

The presence of two calcium ions on the CRD induces a conformational change, allowing the binding of high mannose moieties, which are present on mature cargo glycoproteins. Indeed, it was shown that ERGIC-53, when complexed with Ca^{2+}, binds high-mannose cargoes at the pH found in the ER. However, when the complex reaches the ERGIC, where the luminal pH is lower, the histidine 178 found in the CRD becomes protonated, inducing the release of the calcium ions. This leads to a conformational change in the mannose binding part of ERGIC-53, dissociating therefore the cargo from its receptor. ERGIC-53 is then recycled back to the ER, where the histidine 178 is deprotonated, allowing Ca2+ to bind the CRD. Another round of transport is then possible (Appenzeller, Andersson, Kappeler, & Hauri, 1999; Christian Appenzeller-Herzog et al., 2005; Christian Appenzeller-Herzog, Roche, Nufer, & Hauri, 2004). ERGIC-53 dimerizes after synthesis and then forms a homohexamer which represents the functional entity (Nufer, Kappeler, Guldbrandsen, & Hauri, 2003). The action of ERGIC-53 is described in Figure 13.

Various clients of ERGIC-53 have been identified, including the blood coagulation factors V and VIII (Nichols et al., 1998; B. Zhang et al., 2003), cathepsin C and Z (Appenzeller et al., 1999; Nyfeler, Zhang, Ginsburg, Kaufman, & Hauri, 2006; Vollenweider, Kappeler, Itin, & Hauri, 1998), the IgM biogenesis (Anelli et al., 2007a), among others.

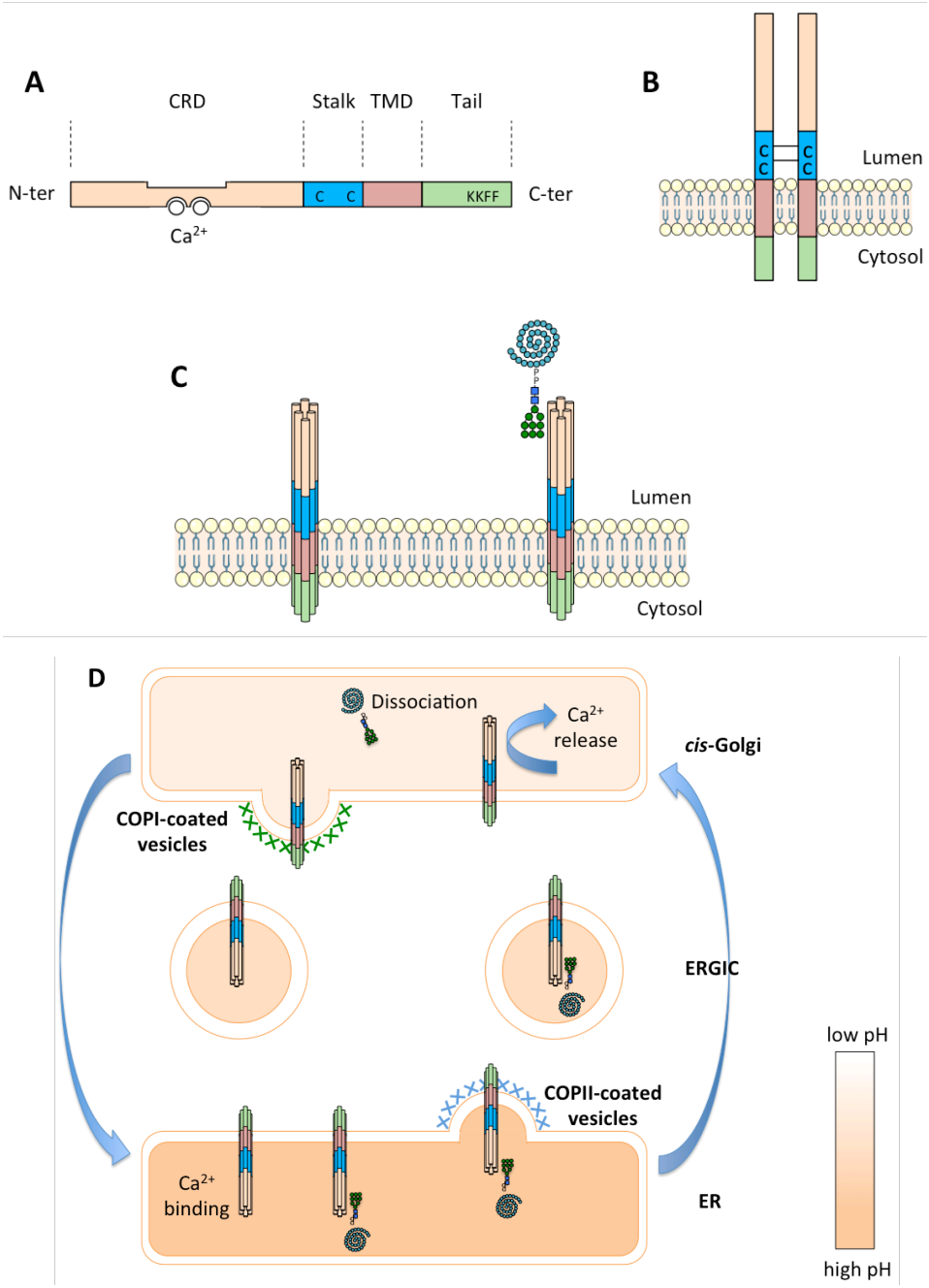

Figure 13: mode of action of ERGIC-53. (A) ERGIC-53 is a type-I membrane protein composed of a carbohydrate-recognition domain able to bind calcium ions, a stalk domain containing two cysteine residues, a transmembrane domain and a cytosolic domain exhibiting both anterograde and retrograde transport motif (a di-lysine motif and a di-phenylalanine motif). (B) Each ERGIC-53 monomer forms a homodimer, through the cysteine residues in the stalk region, forming disulfide bonds. (C) The functional ERGIC-53 protein is an homohexamer, formed by three ERGIC-53 dimers. At the high pH of the ER, the homohexamer is able to bind high mannose moieties, found on the glycosylated proteins. (D) Once bound to the carbohydrate moieties of correctly glycosylated protein, the ERGIC-53/cargo complex is then packaged into COPII-coated vesicles, thanks to the di-phenylalanine motif of ERGIC-53. Once in the ERGIC/Golgi complex, the lower pH allows the release of the calcium ions from the CRD of ERGIC-53, which promotes the dissociation of ERGIC-53 from its partner. This allows the anterograde transport of correctly glycosylated proteins. The di-lysine motif found on ERGIC-53 recruits then COPI-coat components, and recycles back ERGIC-53 in the ER, for another round of transport.

Because it only binds high-mannose moieties, ERGIC-53 acts as a secondary quality control mechanism. Indeed, it selects only correctly folded glycoproteins, acting in concert with the Calnexin/Calreticulin checkpoint.

(b) p24

The family of p24 proteins comprises, in human, four subfamilies: p24α, p24β, p24γ and p24δ. All p24 proteins are type-I membrane proteins found in COPI- and COPII-coated vesicles, but also in the ER, ERGIC and *cis*-Golgi membranes (Belden & Barlowe, 1996; Blum et al., 1996; Emery, Rojo, & Gruenberg, 2000; Langhans et al., 2008; Schimmöller et al., 1995; Sohn et al., 1996; Stamnes et al., 1995).

Early studies in yeast showed that deletion of one p24 protein was sufficient to impair the rapid anterograde transport of Gas1 (Belden & Barlowe, 2001a; Schimmöller et al., 1995), a GPI (glycosylphosphatidylinositol)-anchored protein. It was subsequently hypothesized that p24 proteins were implicated in the transport of GPI-anchored proteins. This result was later confirmed in mammalian cells when the knock-down of a member of the p24 family showed a delay in the transport of the decay-accelerating factor (DAF), a GPI-anchored protein (Takida, Maeda, & Kinoshita, 2008).

GPI-anchored proteins are soluble proteins attached to a membrane by a glycolipidic tail, covalently linked to its C-terminal extremity. In the ER, they are present in the lumen, and consequently cannot interact with cytosolic COPII-coat components. p24 proteins link COPII-coat components and GPI-anchored proteins.

As presented in Figure 14, p24 proteins are small transmembrane proteins exhibiting a luminal GOLD (for GOLgi Dynamics) domain thought to interact with

luminal GPI-anchored cargoes, a coiled-coiled domain that could be responsible for oligomer formation, a transmembrane domain and a short cytoplasmic tail exhibiting COPI- and COPII-coat binding motifs: a KKxx retrieval motif and a di-phenylalanine export motif (Anantharaman & Aravind, 2002; Ciufo & Boyd, 2000; Dominguez et al., 1998; Strating & Martens, 2009). Thus, p24 proteins cycle in the early secretory pathway.

Evidence suggests that p24 proteins function as a tetramer, where one member of each subfamily is represented (Dominguez et al., 1998; Füllekrug et al., 1999; Marzioch et al., 1999). However, little is known about the substrate specificity of p24 proteins.

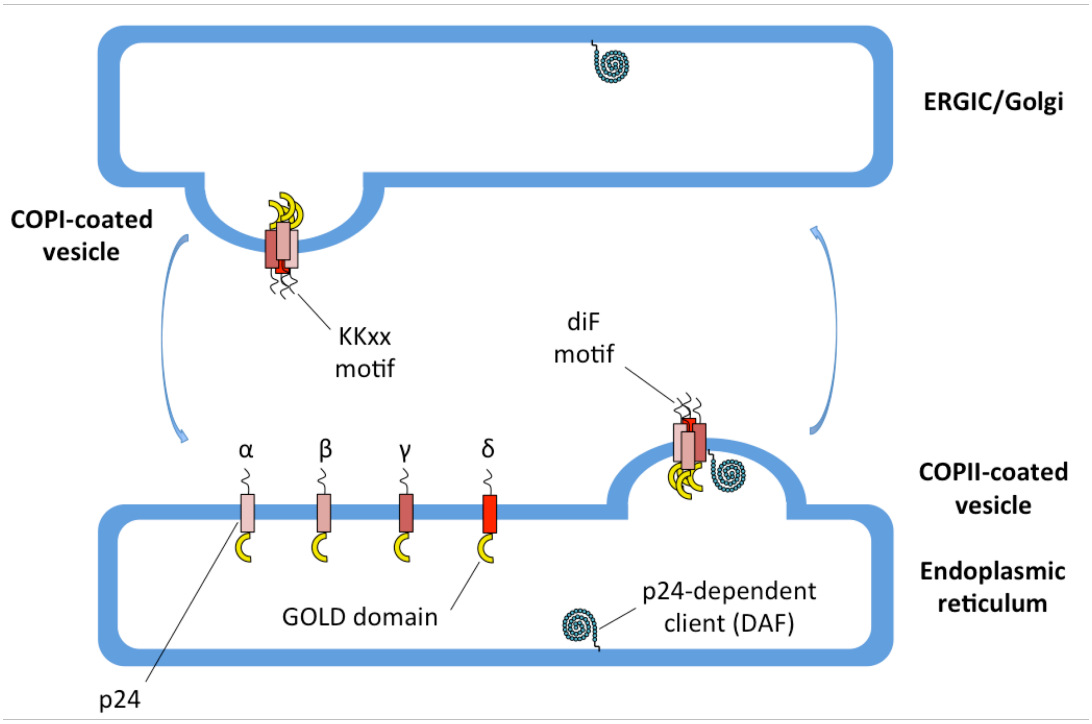

Figure 14: a simplified model for the action of p24 proteins. In the ER, p24 proteins form a multimer composed by the α, β, δ and γ subunits. Their luminal GOLD domain is thought to interact with GPI-anchored proteins (such as DAF) facing the lumen of the ER. The di-phenylalanine (diF) motif found in the cytosolic part of p24 proteins recruits the COPII-coat components, resulting in the exit out of the ER of p24 partners. Once in the ERGIC/Golgi complex, p24 proteins are recycled back in the ER, due to a C-terminal dilysine motif (KKxx), that recruits COPI-coat components. p24 proteins are then free for another round of anterograde transport.

(c) Erv29

Otte et al first identified the Erv29 protein in a proteomic analysis of COPII-coated vesicles (S. Otte et al., 2001). Erv29 is an integral membrane protein with 6 TMDs, with a molecular weight of 29kDa, and is located in both ER and Golgi membranes (Belden & Barlowe, 2001b). Later studies showed that Erv29 exhibits only 4 transmembrane domains (Foley, Sharpe, & Otte, 2007). Studies in yeast have shown that Erv29 null cells were not able to transport efficiently the soluble glycosylated pro-α-factor (gpαf) towards the cell surface. More precisely, these cells were not able to package gpαf in COPII-coated vesicles (Belden & Barlowe, 2001b). In addition, Erv29 was shown to interact directly and specifically with gpαf and to ensure its concentration in COPII-coated vesicles. Moreover, the export of gpαf depends on the level of expression of Erv29. Once in the Golgi complex, Erv29 and its substrate dissociate and Erv29 is recycled back to the ER thanks to the dilysine motif found at its C-terminus.

Later studies showed that an ILV sequence present in gpαf was responsible for its specific interaction with Erv29 (Stefan Otte & Barlowe, 2004). Moreover, the addition of this sequence on a soluble protein was sufficient to confer specific Erv29-dependent export. However, it is not known in detail how Erv29 is recruited by COPII-coat components.

Erv29 is a good example of how soluble proteins are actively recruited and packaged into COPII-coated vesicles, and a simplified view of this mechanism is proposed in Figure 15.

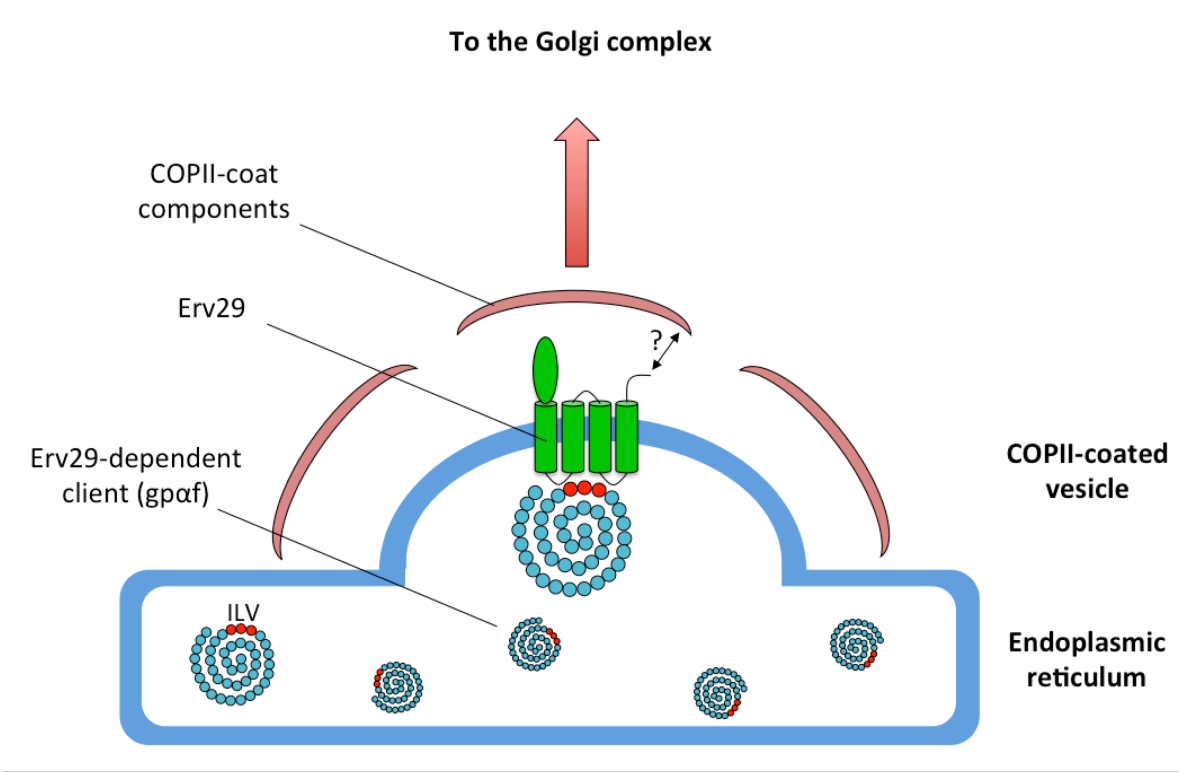

Figure 15: a simplified view of the action of Erv29. In the ER, soluble proteins exhibiting an ILV motif within their sequence (for instance gpαf) interact with the luminal domain of Erv29. This promotes the recruitment of COPII-coat components, by an unknown mechanism. This leads to the exit out of the ER of Erv29-dependent partners.

(d) SCAP

Regulating the activity of the cargo receptor itself is a way to control intracellular trafficking. This is nicely illustrated by the sterol regulatory element-binding proteins (SREBPs), for which the sorting out of the ER is achieved through the SCAP protein (SREBP Cleavage-Activating Protein).

SREBPs are transcription factors, inactive when membrane-bound. Once in the Golgi complex, they undergo a proteolytic cleavage that releases the active, soluble transcription factors, which in turn increases the gene transcription responsible for cholesterol synthesis (M. S. Brown & Goldstein, 1997; Horton, 2002). Once synthetized

in the ER, SREBPs are taken in charge by the eight-transmembrane domains protein SCAP, through the C-terminal WD-repeat domain of SCAP, a region rich in tryptophan and aspartic acid residues. SCAP drives in turn SREBPs into COPII-coated vesicles, because of the interaction of Sec24 with the MELADL hexapeptide found in a cytoplasmic loop of SCAP, between TMD 6 and 7 (Sun, Li, Goldstein, & Brown, 2005). However, the ER levels of cholesterol regulate the binding between SCAP and Sec24. Indeed, when ER cholesterol levels are high, SCAP undergoes a conformational change, in its sterol-sensing domain, that induces the binding to Insig, a membrane protein of the ER. In this context, the binding to Sec24, and the subsequent exit of SREBPs out of the ER, are aborted (Adams et al., 2004; Adams, Goldstein, & Brown, 2003; A. J. Brown, Sun, Feramisco, Brown, & Goldstein, 2002; Radhakrishnan, Sun, Kwon, Brown, & Goldstein, 2004; Sun et al., 2005). On the contrary, in a low cholesterol condition, SCAP directly binds the C-terminal domain of SREBPs and escorts them into COPII-coated vesicles, to be transported in the Golgi complex, where they are processed to release the soluble, active transcription factors, which in turn induces cholesterol synthesis (Y. Zhang, Motamed, Seemann, Brown, & Goldstein, 2013).

Thus, by inhibiting the activity of the cargo receptor of SREBPs, cells can limit the amount of active transcription factors produced in the Golgi complex, and consequently the cholesterol synthesis

b) Mechanisms ensuring concentration of proteins into retrograde transport vesicles at the Golgi-to-ER interface

Although anterograde secretory transport is tightly regulated, some proteins can inadequately escape the ER and reach the Golgi complex. In addition, receptor proteins involved in ER-to-Golgi transport need to be recycled back to the ER. Selective retrograde transport is necessary in these two situations.

As it is the case for the anterograde transport, different strategies exist: the direct interaction of the protein to be retrieved and the COPI-coat components or the capture of the client by a cargo receptor, thanks to specific motifs, which loads the complex into COPI-coated vesicles.

(1) Concentration of proteins into COPI-coated vesicles through direct interaction with coat proteins

Many transmembrane proteins are retrieved from the ERGIC or the *cis*-Golgi by the direct binding of a C-terminal dilysine motif (KKxx or KxKxx) with the COPI coat. Early studies showed that the presence of lysine residues in position -3/-4 or -3/-5 from the C-terminus was sufficient to localize a type-I transmembrane protein in the ER (M. R. Jackson, Nilsson, & Peterson, 1990). Later studies proved that the COPI coat was responsible for the retrieval of proteins bearing such di-lysine motifs (P. Cosson & Letourneur, 1994; Letourneur et al., 1994). Structural data showed that the lysine residues interact with the N-terminal WD-repeat domains of α–COP and β'-COP (L. P. Jackson et al., 2012; Ma & Goldberg, 2013). The efficiency with which this motif binds

the COPI coat in the cell and ensures ER localization is weakened by the length of the cytoplasmic domain (Vincent, Martin, & Compans, 1998).

Once bound to COPI-coat components, these transmembrane proteins are concentrated into COPI-coated vesicles and transported back to the ER.

Other motifs have been described that directly bind the COPI coat, as described for the δL motif, interacting with the δ-COPI subunit (P. Cosson, Lefkir, Démollière, & Letourneur, 1998). These motifs also act as efficient ER-targeting motifs in the cell.

(2) Concentration of proteins into COPI-coated vesicles by cargo receptors

As described for anterograde transport, a lot of proteins do not interact directly with COPI-coat components. They are recognized by specific receptors, which interact with sorting motifs, present with the polypeptidic sequence of the client. As seen before, these motifs are found in the soluble domains of the cargo. A few examples of cargo receptors implicated in retrograde transport is presented below.

(a) KDEL receptor

For the retrieval of soluble proteins, the most extensively studied couple of client/cargo receptor is probably the KDEL motif and its transmembrane KDEL receptor.

The KDEL sequence (HDEL in yeast) was first identified in 1987, at the C-terminal extremity of several soluble ER chaperones. The addition of this motif to secreted proteins led to their localization in the ER (Munro & Pelham, 1987). Biochemical studies

in yeast showed that carbohydrates of a chimeric HDEL-tagged invertase carried Golgi-specific modifications, proving therefore that the cargo and its receptor cycle between the ER and the Golgi complex (Pelham, 1988; Pelham, Hardwick, & Lewis, 1988; Semenza, Hardwick, Dean, & Pelham, 1990). Random mutagenesis allowed the identification of a mutant unable to retain the invertase-HDEL, and led to the identification of ERD2 (ER-retention defective 2), an integral membrane protein, as the HDEL receptor (Hardwick, Boothroyd, Rudner, & Pelham, 1992; Semenza et al., 1990).

Three KDEL receptors were identified in human: hERD2 (or KDEL receptor 1), ELP1 (ER luminal protein 1, or KDEL receptor 2), and KDEL receptor 3a and its isoform KDEL receptor 3b (Lewis & Pelham, 1990, 1992; Raykhel et al., 2007). KDEL receptors are seven transmembrane domain-proteins with the C terminal part of the three proteins placed in the cytosol. They are mainly localized in the Golgi complex (Scheel & Pelham, 1998; Townsley, Wilson, & Pelham, 1993). Extensive studies were performed to identify the binding site of the KDEL receptors. It is generally accepted that mutations in the cytosolic loops of the KDEL receptors result in the loss of the transport between the ER and the Golgi complex whereas mutations in the luminal loops impair the interaction with the KDEL-containing protein (Townsley et al., 1993). More precisely, four residues were shown to be critical for the binding of the KDEL motif (R5, D50, Y162 and N165) (Scheel & Pelham, 1998).

The association between the KDEL receptor and its substrate has been hypothesized to be dependent on the pH. Indeed, the association has been described to be favored in an acidic environment, and much weaker at a neutral or basic pH. This could explain the release of the cargo from the KDEL receptor in the ER, the pH in this organelle being higher than that in the Golgi complex (Wilson, Lewis, & Pelham, 1993; Wu et al., 2000, 2001).

KDEL receptors have also been described to stimulate retrograde transport in COPI-coated vesicles (Aoe et al., 1997; Aoe, Lee, van Donselaar, Peters, & Hsu, 1998). Indeed, when bound to its ligand, the KDEL receptor was shown to interact with Arf GAP1, promoting therefore the recruitment of the coatomer. In this way, the KDEL receptor is also able to modulate the retrograde transport through COPI-coated vesicles. This process is presented in Figure 16.

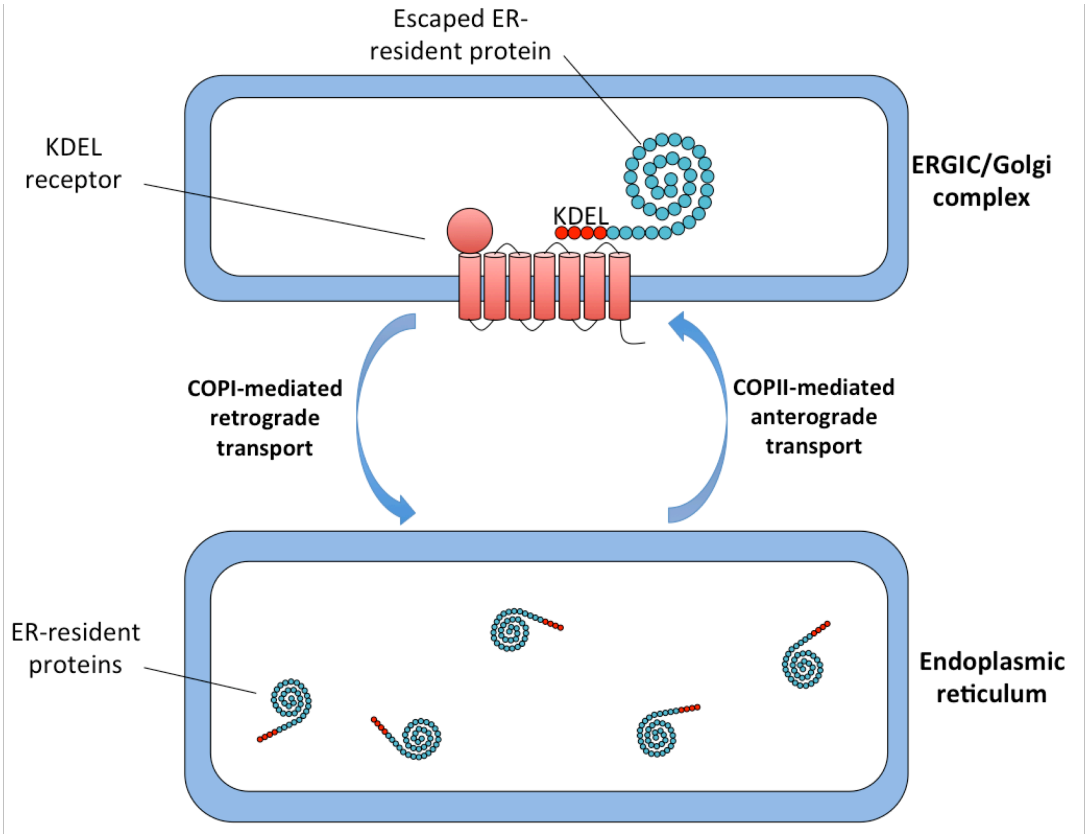

Figure 16: Action of the KDEL-receptor. In the ERGIC/Golgi complex, escaped, ER-resident soluble proteins exhibiting a KDEL motif are taken in charge by the KDEL-receptor. The low pH of the ERGIC/Golgi complex allows the association of the KDEL-receptor with the KDEL-bearing protein. This interaction favors the recruitment of COPI-coat components that transports the KDEL-receptor and its partner back to the ER. Here, the higher pH promotes the dissociation of the KDEL-receptor from its partner. This results in the ER localization of proteins bearing a KDEL-motif.

(b) ERp44

Some cargo receptors can act as chaperones. It is the case for instance for the ERp44 protein. ERp44 is a Protein Disulfide-Isomerase (PDI) family member mainly localized in the ERGIC (Anelli et al., 2007a), firstly described because it interacts with the unassembled IgM subunits (Anelli et al., 2003, 2007a).

IgM are penta- or hexamers, composed by a μ heavy and a light chain of immunoglobulins. The constant domain 1 (CH1) of the μ heavy chain, after synthesis, is first bound by BiP and is retained in the ER, until the assembly with the light (L) chain, into μ_2L_2 complex. Once correctly glycosylated, each IgM is taken in charge by ERGIC-53 for anterograde transport. Here, ERGIC-53 acts as a platform for the assembly of the final penta- or hexamer through disulfide bonds formation (Anelli et al., 2007b; Sitia et al., 1990). Indeed, the hexameric ERGIC-53 may concentrate μ_2L_2 subunits, facilitating the formation of disulfide bond. On the contrary, a conserved cysteine residue (Cys29) of ERp44 has been shown to interact with unpolymerized μ_2L_2 subunits, at the permissive pH of the ERGIC/*cis*-Golgi. This causes a conformational change in ERp44, exposing the KDEL-like sequence RDEL, found in its C-terminal tail (Anelli et al., 2003). Once in the ER, the restrictive pH favours the dissociation of ERp44 and the μ_2L_2 subunit (Anelli & van Anken, 2013), which is then available for ERGIC-53, and for another round of oligomerization. The action of ERp44 is illustrated in Figure 17.

By binding and retrieving unassembled, free IgM subunits, ERp44 prevents the secretion of misfolded IgM, participating therefore in the quality control along the secretion pathway.

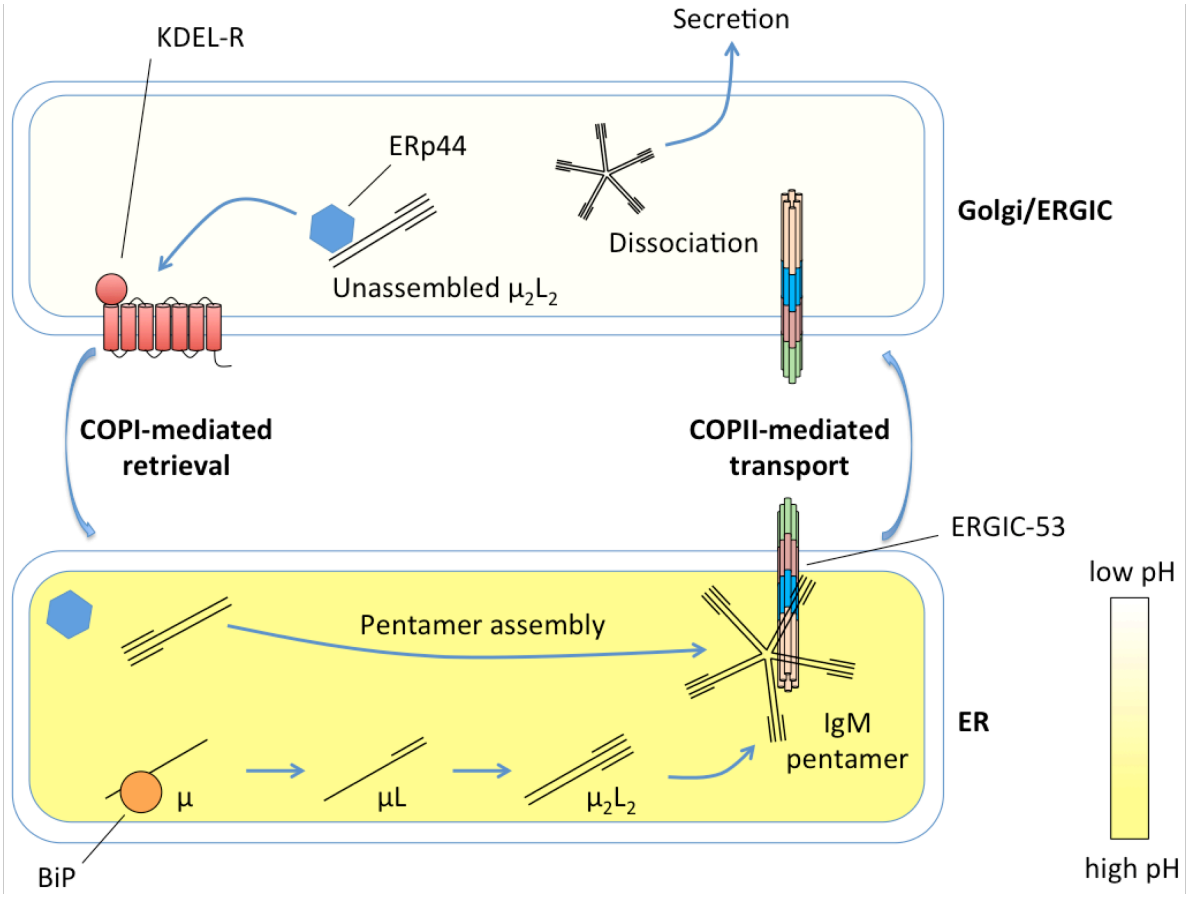

Figure 17: a simplified model proposed for the action of ERp44. Once synthesized and glycosylated in the lumen of the ER, IgM monomers (formed by two μ heavy chains and two L light chains μ_2L_2) interacts with ERGIC-53. Here, ERGIC-53 acts as a platform for the formation of the final IgM pentamer, and the complex is subsequently transported to the ERGIC/Golgi complex in a COPII-mediated manner. The pH of the ERGIC/Golgi complex compartment allows the dissociation of ERGIC-53 and the IgM pentamer. On the contrary, escaped, unassembled μ_2L_2 monomers that have been transported out of the ER interact with ERp44, thanks to a conserved cysteine residue. This interaction provokes a conformational change in ERp44 that exposes a KDEL-like sequence in the C-terminal part of ERp44. The KDEL-receptor interacts with this motif and transports the ERp44/IgM monomer back to the ER, in a COPI-mediated manner. Here, the higher pH allows the dissociation of the IgM monomer from ERp44. The IgM monomer is therefore free for another round of interaction with ERGIC-53, for the assembly of the final IgM pentamer.

(c) Erv41/Erv46

Proteomic analysis of purified COPII-coated vesicles revealed that the Erv41-Erv46 complex was enriched in these vesicles (S. Otte et al., 2001). Both proteins of 41 and 46 kDa are membrane proteins with 2 TMDs, short cytosolic domains but large luminal segments. A C-terminal KKxx retrieval motif is found in Erv46 and ensures the

ER localization of the complex. On the contrary, it was demonstrated that the isoleucine and leucine residues found at positions -4 and -3 of Erv41 C-terminus were critical for COPII-component binding, whereas a weak di-phenylalanine motif (FY) is exhibited by Erv46. These two motifs ensure the packaging of Erv41-Erv46 in COPII-coated vesicles. As a result, the complex cycles between the ER, the ERGIC and the *cis*-Golgi (Lelio Orci, Ravazzola, Mack, Barlowe, & Otte, 2003; S. Otte et al., 2001).

Lack of Erv41 leads to a reduce ER localization of soluble proteins (Shibuya, Margulis, Christiano, Walther, & Barlowe, 2015). These soluble proteins are normally localized in the ER but do not present any KDEL motif. Therefore, Erv41-Erv46 complex is believed to act as a retrograde cargo receptor, for a new class of ER localized proteins. However, the motif recognized by Erv41-Erv46 is still to be determined.

As we saw, many strategies exist to ensure the concentration of the correct proteins into COPI- or COPII-coated vesicles. The final localization of a protein is not the result of a single sorting event but the result of several. Moreover, the strength of each sorting motif can be weakened by single aminoacid modifications.

c) Mechanisms ensuring sorting of proteins in the Golgi complex

The Golgi complex is constantly traversed by a flow of lipids and proteins, but needs to maintain its specific biochemical composition. Since it is still hotly debated whether the Golgi localization of proteins is achieved through retention mechanisms (in the intra-Golgi vesicle-mediated transport model) or through a constant active transport

(in the intra-Golgi cisternal maturation transport model), the sorting mechanisms in the Golgi complex are only poorly understood.

Most of the knowledge about the localization of Golgi resident proteins comes from the study of glycosyltransferases. The role of TMDs in the Golgi localization of proteins is described in the next chapter.

Glycosyltransferases enzymes are responsible for the maturation of sugars coupled to proteins and lipids. Golgi glycosyltransferases are usually type-II membrane proteins, with the catalytic domain exposed in the lumen of the Golgi cisternae, a single TMD and a short, 5-20 amino acid cytosolic tail. Glycosyltransferases are believed to be localized in the Golgi by multiple mechanisms (Fenteany & Colley, 2005). These enzymes exhibit generally both anterograde and retrograde sorting motifs. This results in a dynamic process that localizes them in particular compartments of the Golgi complex at steady-state, through continual cycles of anterograde and retrograde transport (Okamoto, Yoko-o, Miyakawa, & Jigami, 2008), but as previously mentioned, this is still hotly debated.

It has been shown that several yeast glycosyltransferases cycle within the Golgi complex (and between the ER and the Golgi complex, but this is under intense debate since recent studies contradicted this result, as previously described) in a manner that requires a functional coatomer (Todorow, Spang, Carmack, Yates, & Schekman, 2000). The Vacuolar Protein Sorting 74 (Vps74) has been shown to be particularly important in this process (Tu, Tai, Chen, & Banfield, 2008). Indeed, the loss of Vps74 in yeast results in the mislocalization of the Golgi glycosyltransferase Kre2, but not Rer1, to the vacuole. Although no canonical dilysine motif has been found in the cytosolic tail of most of the

glycosyltransferases, a cytosolic motif F/L-L/I/V-X-X-R/K (where X can be any amino acid) was necessary to bind the peripheral Golgi membrane protein Vps74 (Tu et al., 2008), in a PIP_4-dependent manner (Wood et al., 2009). Moreover, Vps74 has the ability to directly bind to COPI component (Tu, Chen, & Banfield, 2012; Tu et al., 2008). In the current model, Vps74 concentrates glycosyltransferases into COPI-coated vesicles, facilitating therefore the recycling to the ER or to the previous Golgi sub-compartment (Tu et al., 2012). This leads to the dynamic Golgi localization of these glycosyltransferases. Vps74 is the only protein that localizes proteins in the Golgi complex though their cytosolic domain identified yet, and a simplified model of the action of Vps74 is proposed in Figure 18. However, several glycosyltransferases do not rely on Vps74 for their Golgi localization, underlining the fact that other mechanisms exist and are still to be discovered.

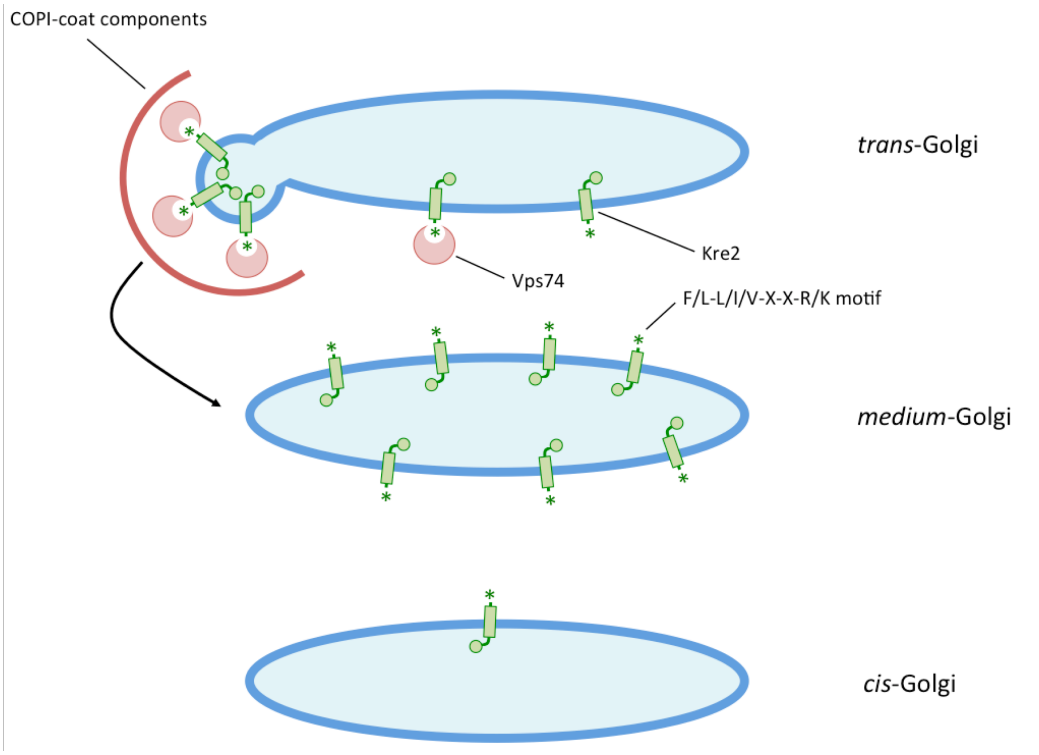

Figure 18: Golgi localization of Golgi-resident proteins through Vps74. At the level of the Golgi complex the cytosolic Vps74 protein interacts with the F/L-L/I/V-X-X-R/K motif, present in the cytosolic domain of the Kre2 protein. This in turn promotes the recruitment of the COPI-coat components to package the Kre2/Vps74 complex into COPI-coated vesicles, and their retrieval to the previous cisterna of the Golgi complex.

3. Bulk flow transport of proteins

The bulk flow transport corresponds to the transport of a protein where its concentration in the transport vesicle is the same as that found in the donor compartment (Wieland, Gleason, Serafini, & Rothman, 1987). The packaging of soluble or membrane proteins by bulk flow is a 'default' transport, since every single vesicle can transport membrane or soluble proteins as part of the fluid phase. This is for instance the case for the ER export of the secreted proteins amylase or chymotrypsinogen, which are not concentrated in COPII-coated vesicles (Martínez-Menárguez, Geuze, Slot, & Klumperman, 1999).

However, this transport strategy is particularly inefficient and the bulk transport only contributes to a marginal fraction of the anterograde transport (Malkus, Jiang, & Schekman, 2002). Nevertheless, a study made on the C-terminal domain of the Semliki Forest virus capside protein suggests that the efficiency of the bulk flow transport may be underestimated (Thor, Gautschi, Geiger, & Helenius, 2009). Clearly, the participation of bulk flow in the transport of proteins is not well established.

D. Sorting of proteins in the early secretion pathway: the specific case of transmembrane domains

Transmembrane proteins are characterized by the fact that they span the lipid bilayer. Membrane proteins account for 20-30% of the proteins encoded by eukaryotic genomes and achieve many functions within a cell. They can for example form channels allowing molecules to freely diffuse across the lipid bilayer, transport actively specific proteins from one compartment to another, or interact with an extracellular ligand to trigger signalling pathways inside the cell.

Transmembrane proteins are characterized by the presence of at least one hydrophobic segment (transmembrane domain, TMD) that usually folds as an alpha-helix in the hydrophobic lipidic environment. The general organization of TMDs is described in section D.1. The sorting determinants found in TMDs are discussed in section D.2. Finally, mechanisms governing the sorting of TMDs are presented in section D.3.

1. Structure of TMDs

Transmembrane proteins cross the lipid bilayer thanks to one or several stretches of 26 ± 6 generally hydrophobic aminoacids (Bowie, 1997; Gimpelev, Forrest, Murray, & Honig, 2004; Martin B. Ulmschneider, Tieleman, & Sansom, 2005). TMDs are most of the time organized as α-helices. Although membrane thickness differs depending on the compartment considered (with the plasma membrane thicker than the ER membrane), a typical membrane is ≈35Å thick, and it is crossed by an alpha-helical

segment of approximately 20 residues. Aminoacids are not equally frequent in a TMD: hydrophobic aminoacids are strongly preferred in TMDs (e.g. leucine, alanine, or valine) while hydrophilic residues (e.g. methionine, histidine, glutamine or cysteine) are less frequent (M. B. Ulmschneider & Sansom, 2001). However, it is not exceptional to find some hydrophilic residues within a TMD. For instance, serine residues, despite being hydrophilic, are often found in TMDs. Glycine residues are also found, and they are involved in interactions between TMDs (Fink, Sal-Man, Gerber, & Shai, 2012; Teese & Langosch, 2015) and ensure protein oligomerization or folding of multispanning proteins.

In a few cases of bacterial, chloroplastic or mitochondrial proteins, transmembrane segments are organized as β-barrels (Naveed, Xu, Jackups, & Liang, 2012), an uncommon situation not detailed in this manuscript.

2. Sorting motifs found in TMDs

Although many studies were focused on sorting determinants found in the soluble domains of membrane proteins, it has been known for decades that the TMD itself contains information governing the protein localization (J. S. Bonifacino, Cosson, Shah, & Klausner, 1991; Machamer, 1993). It is also generally accepted that TMD-based localization determinants contribute very significantly to the intracellular localization of many transmembrane proteins (Karsten, Hegde, Sinai, Yang, & Joiner, 2004). Sorting motifs in TMDs are recognized by different mechanisms that are described in the next section.

a) Length of TMDs

Membrane protein localization is largely influenced by the length of its TMD (Sharpe, Stevens, & Munro, 2010). First, ER resident membrane proteins tend to have shorter-than-average TMDs. On the contrary, plasma membrane proteins exhibit longer TMDs, which may better fit the thicker plasma membrane. This was exemplified in 2000, when the length of yeast TMDs was compared. This revealed a constant increase in the length of TMDs along the secretory pathway (from 15 residues in the ER to 23 at the plasma membrane), reflecting a thickening of the corresponding membranes (Levine, Wiggins, & Munro, 2000). Secondly, the length of a TMD is sufficient to target it to different sub-compartment. Indeed, the 17 residues TMD of syntaxin 5 is sufficient to target chimeric proteins to the early secretion pathway whereas the longer 25 residues TMD of syntaxin 3 targets reporter proteins to the plasma membrane (Watson & Pessin, 2001). Finally, changing the length of a TMD greatly impacts its localization. Indeed, lengthening the TMD of the Golgi resident sialyltransferase increases gradually its localization at the plasma membrane (Munro, 1995). Conversely, shortening the TMD of a plasma membrane protein induces its relocalization in intra-cellular compartments.

b) Charged residues in TMDs

Potentially charged residues are also important for the sorting of TMDs. For instance, each of the eight subunits of the T-cell receptor (TCR) is a type-I membrane protein, with at least one charged residue, a very unusual feature in TMDs. Only a fully assembled TCR is exported to the cell surface, whereas unassembled subunits are

retained in the ER. Actually, in the fully assembled complex, all the charges are hidden for the hydrophobic environment of the lipid bilayer, allowing therefore the surface localization (J. S. Bonifacino, Cosson, & Klausner, 1990). The unassembled subunits are retained in the ER only because of their charge, since the mutation of this polar residue abolishes the ER retention. This process is however weakened by the length of the TMD: a long TMD decreases the retention effect of a charged residue (Lankford, Cosson, Bonifacino, & Klausner, 1993).

Moreover, the introduction of a charged residue (arginine or aspartic acid for instance) in a surface localized TMD induces its ER localization (J. S. Bonifacino et al., 1991).

c) Global composition of TMDs

Finally, the global composition of a given TMD greatly affects its localization. TMDs found in the ER and the Golgi complex exhibit shorter regions enriched in highly hydrophobic residues, whereas plasma membrane TMDs contain large stretches of hydrophobic residues (Sharpe et al., 2010).

To illustrate this observation, the transmembrane domain of the μ chain has been shown to be localized in the ER of IgM-secreting cells. This TMD of the μ chain is particularly enriched in hydrophilic residues (in particular serine and threonine residues). However, their mutation into more hydrophobic residues, such as alanine or valine residues confers plasma membrane localization to these mutated TMDs (Williams, 1990).

3. Mechanisms ensuring the sorting of TMDs

The sorting motifs found in TMDs described above are recognized by different mechanisms that allow their final localization. They can be mediated by cargo receptors (a), or alternatively do not rely on the action of a cargo receptor (b).

a) Receptor-based sorting of TMDs

TMD motifs can be specifically recognized by cargo receptors, through intra-membrane interactions. As a result, the cargo receptor/client complex is specifically concentrated into COPII-coated or COPI-coated vesicles. This process is exemplified here with the cargo receptor Erv14, implied in the anterograde transport, and Rer1, the only cargo receptor implicated in the retrograde transport of TMDs identified yet. These receptors act at the ER/Golgi complex interface, but no cargo receptor ensuring the Golgi localization of TMDs has been identified yet. Since cargo receptors do exist to ensure the transport of TMDs out of the ER, or to ensure ER localization, similar mechanisms probably localize TMDs in the Golgi complex, but have never been described so far. Therefore, a lot of efforts are still needed to decipher the precise mechanisms ensuring the localization of TMDs in the Golgi complex.

(1) Erv14, an example of cargo receptor ensuring the anterograde transport of TMDs

The *Drosophila* Cornichon proteins, and the yeast homolog Erv14, were the first cargo receptors recognizing motifs in TMD to be identified in the anterograde transport.

Erv14 (Endoplasmic reticulum-derived vesicle 14 protein) is a membrane protein containing three TMDs, which localizes to ER and Golgi membranes, as well as in COPII-coated vesicles. Erv14 contains COPII-export, as well as COPI-retrieval motifs (Pagant, Wu, Edwards, Diehl, & Miller, 2015; Jacqueline Powers & Barlowe, 2002). Indeed, it was described that the IFRTL motif, found in the cytoplasmic domain of Erv14 was responsible for its binding with Sec24 (Pagant et al., 2015).

Initial studies in yeast showed that Erv14-deficient cells accumulate the plasma membrane protein Axl2 in the ER (J. Powers & Barlowe, 1998). Later studies identified more than 30 clients, which depend on Erv14 to be efficiently transported to the surface. These clients are single- or multiple-membrane spanning proteins, and include for example AMPA receptors, or G protein-coupled receptors (GPCRs). The particularity of most of these transmembrane cargoes is that they are found in the late secretory pathway, or at the plasma membrane, and have long TMDs. Therefore, Erv14 was proposed to be a cargo receptor for long TMDs (Herzig, Sharpe, Elbaz, Munro, & Schuldiner, 2012), helping them to be concentrated into COPII-coated vesicles, through the interaction of Erv14 with Sec24. The second TMD of Erv14 was shown to be responsible for the binding of some, but not all, cargoes (Pagant et al., 2015). A simplified mode of action of Erv14 is proposed in Figure 19.

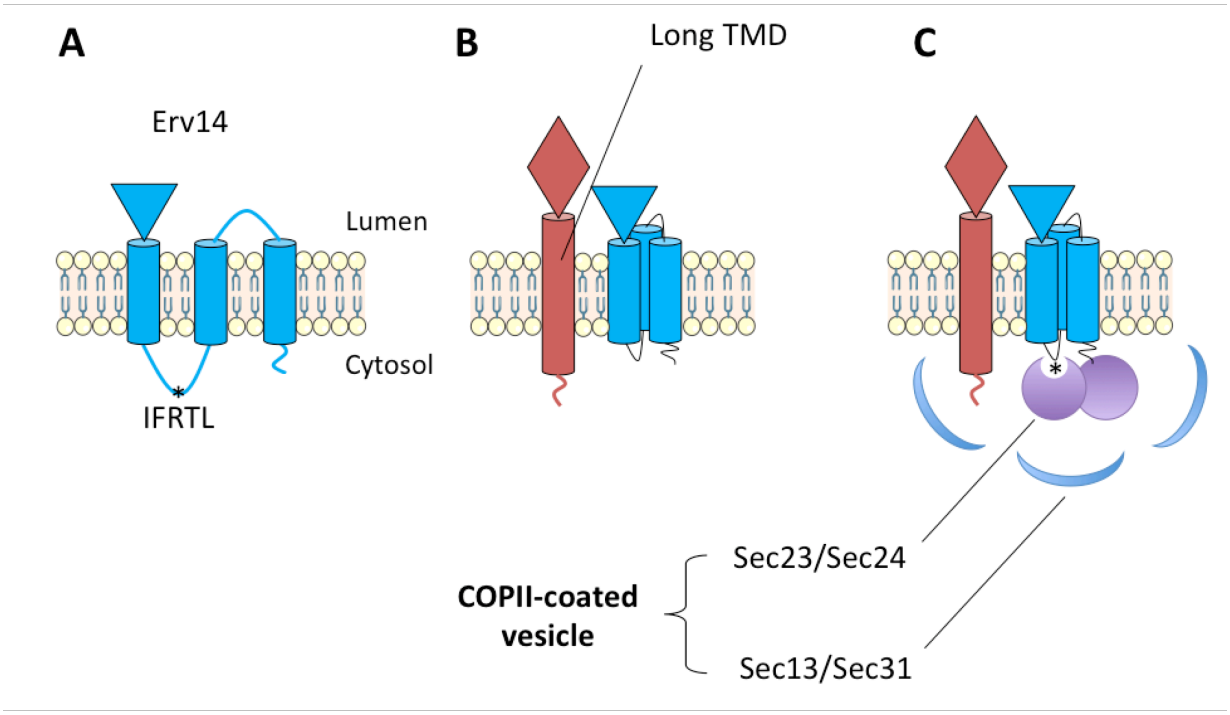

Figure 19: mode of action of Erv14. (A) Erv14 exhibits three transmembrane domains, and possesses an IRFTL motif in its cytosolic loop, thought to interact with COPII-coat components. (B) Erv14 interacts with long TMDs in the ER. (C) The interaction of Erv14 with long TMDs recruits COPII-coat components that transport the complex out of the ER.

(2) Rer1, an example of cargo receptor ensuring the retrograde transport of TMDs

Some cargo receptors are responsible for the specific retrograde transport of certain TMDs. This is notably the case for Rer1 (Retention in Endoplasmic Reticulum 1), first described in 1993 in yeast (Nishikawa & Nakano, 1993). Mutant defective for Rer1 were initially described as unable to retain the membrane protein Sec12 in the ER, but didn't show any defect in the ER localization of the KDEL-containing soluble BiP protein. Rer1 was later shown to interact specifically and directly with the transmembrane

domain of Sec12 (M. Sato, Sato, & Nakano, 1996), and more particularly with the polar residues found in the TMD of Sec12 (K. Sato, Sato, & Nakano, 2001).

It was then demonstrated that the retrieval of numerous membrane proteins (including the unassembled subunits of the γ–secretase, (Park et al., 2012)) was dependent on Rer1 (Massaad, Franzusoff, & Herscovics, 1999; K. Sato, Sato, & Nakano, 1997; M. Sato et al., 1996), all these proteins exhibiting different topologies. Moreover, the retrieval achieved by Rer1 was shown to rely on the COPI-coat machinery (Boehm et al., 1997; K. Sato et al., 1997), and the C-terminal part of Rer1 is responsible for the binding with COPI-coat components (K. Sato et al., 2001).

At steady-state, Rer1 is found in the Golgi complex but cycles between the ER and the Golgi complex. Indeed, when expressed in a thermosensitive *sec13* yeast mutant, a GFP form of Rer1 localizes in the Golgi complex at a permissive temperature but rapidly shifts in the ER at 37°C (K. Sato et al., 2001).

A model to explain the retrieval by Rer1 would be that Rer1 interacts with the TMD of its partner in the Golgi complex, as illustrated in Figure 20. This association would in turn recruit the COPI-coat components to retrieve back the receptor and its cargo. Once in the ER, the cargo is released from Rer1, which is then recycled back to the Golgi complex.

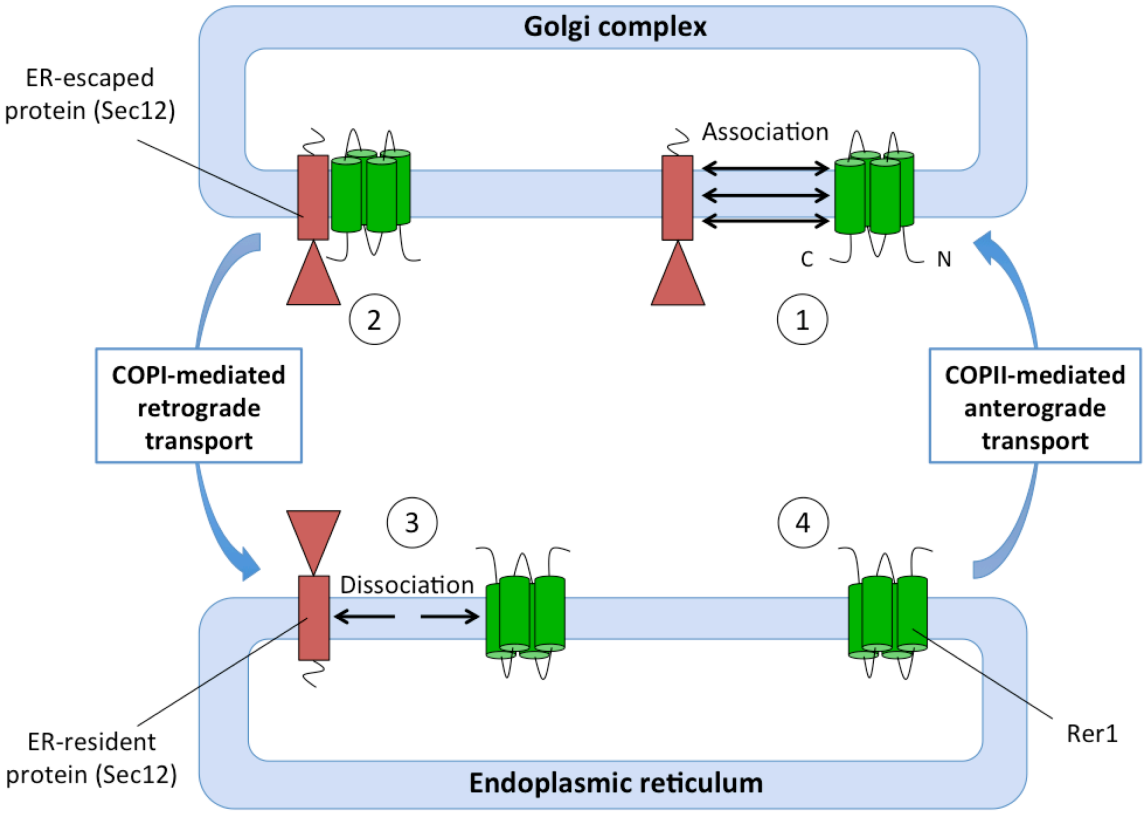

Figure 20: a simplified view of the function of Rer1. Rer1 interacts with short TMDs in the Golgi complex (1). This association recruits the COPI-coat components that package this complex into COPI-coated vesicles (2). Once in the ER, Rer1 dissociate from its substrate as indicated in (3). Free Rer1 is then recycled back in the Golgi complex by COPII-coat components and is available for another round of transport (4).

b) Non-receptor based sorting of TMDs

Sorting of TMD may also be achieved by mechanisms that do not involve any receptor. It is not clear today if receptor-independent mechanism of sorting play a minor or a major role in sorting of TMDs. In any case, they are thought to function in conjunction with specific receptors to ensure efficient targeting of TMDs in the secretory pathway. Here I depict the example of the sorting of TMDs by lipid partitioning (1) and the protein aggregation (2).

(1) Lipid partitioning

Partitioning of TMDs into lipid microdomains may participate in their sorting. This depends on the intrinsic properties of the TMDs, and their interaction with the lipid bilayer. The experimental evidence came from a study where short TMDs were shown to segregate together, thus forming microdomains, in reconstituted membranes, and in a cholesterol-dependent manner (Kaiser et al., 2011). Long TMDs were, on the contrary, excluded from these microdomains. Such microdomains were described to be abundant at the cell surface of yeasts (Spira et al., 2012).

In the ER of mammalian cells, short TMDs were differently distributed, compared with long TMDs (Ronchi, Colombo, Francolini, & Borgese, 2008). Indeed, a 17-residue TMD fluorescent reporter protein segregated differently from a 22-residues TMD within the ER. Moreover, short TMDs were shown to be excluded from COPII-coated vesicles, contrary to long TMDs, which were recruited to the ERES.

(2) Protein aggregation

It has been reported that glycosyltransferases aggregate, or at least oligomerize, in the Golgi complex. They are therefore excluded from the anterograde and retrograde transport vesicles. This process is mediated by both TMDs and soluble parts of these proteins, and is dependent on the pH and the concentration of the protein (Fenteany & Colley, 2005). For instance, TMDs of Golgi proteins are enriched in cysteine residues that form disulfide bonds, as well as in polar residues, that may favour interactions between TMDs (Machamer, 1993; Munro, 1995; Sousa et al., 2003). This aggregation-based

mechanism is called the kin-recognition model (T. Nilsson et al., 1994; T. Nilsson, Slusarewicz, Hoe, & Warren, 1993). Moreover, the aggregation of glycosyltransferases is important for their function (Giraudo, Daniotti, & Maccioni, 2001).

In this way, Golgi-resident enzymes are excluded from transport vesicles. This mechanism contributes therefore to the sorting of Golgi enzymes. This mechanism has also been described to occur in the ER, preventing the entry of aggregated proteins into COPII-coated vesicles (McCaffrey & Braakman, 2016).

In this introduction, I presented the different mechanisms that eukaryotic cells use to recognize, sort and transport proteins in the early secretion pathway. The aim of this introduction was to present the different strategies used by the cell to transport or localize correctly nascent proteins within the early secretory pathway and illustrate them with a few examples. It is important to stress the fact that the final localization of a given protein is usually not the result of one single sorting event, but results from the concerted recognition of many different sorting determinants.

II. RESULTS

II. Results

A. Publication: Immunofluorescence labeling of cell surface antigens in *Dictyostelium*

1. Introduction

Dictyostelium discoideum is a social amoeba used in laboratories to study many facets of cellular biology, such as phagocytosis, motility or signal transduction. The 34Mb genome of *Dictyostelium* has been entirely sequenced, is divided into 6 chromosomes and is particularly rich in A/T. *Dictyostelium discoideum* possess about 10'000 genes highly similar to their vertebrate's orthologs (Baldauf & Doolittle, 1997; Eichinger et al., 2005). One particularity of the *Dictyostelium* genome is that it is haploid, meaning that the disruption of one allele is not compensated by the second allele, making genetics much easier. *Dictyostelium* cells can be easily cultured in liquid medium at 21°C, and divide every 12 hours.

During my PhD, I had the opportunity to study the localization of a reporter protein in *Dictyostelium discoideum* and showed that the method of permeabilization used greatly influences the detection of cell surface antigens. Quantitative detection of cell surface antigens can be of particular importance. Therefore, we propose in this study a two-step protocol to avoid detection artefacts.

2. Publication

Immunofluorescence labeling of cell surface antigens in *Dictyostelium*

Alexandre Vernay and Pierre Cosson,

BMC Research Notes (2013), **6**:317

In this publication, I contributed to all the figures, as well as the writing of the paper.

Vernay and Cosson *BMC Research Notes* 2013, **6**:317
http://www.biomedcentral.com/1756-0500/6/317

BMC Research Notes

TECHNICAL NOTE Open Access

Immunofluorescence labeling of cell surface antigens in *Dictyostelium*

Alexandre Vernay* and Pierre Cosson

Abstract

Background: Immunolocalization of cellular antigens typically requires fixation and permeabilization of cells, prior to incubation with antibodies.

Findings: Assessing a test protein abundantly present at the cell surface of *Dictyostelium* cells, we show that in fixed cells, permeabilization extracts almost completely this cell surface antigen. The extent of this artifact is variable depending on the procedure used for labeling and permeabilization, as well as on the antigen considered.

Conclusions: An optimized protocol for labeling both surface and intracellular antigens without significant loss of labeling is proposed.

Background

In order to detect the presence of a protein in eukaryotic cells, and to determine its intracellular localization, it is common to label cells with specific fluorescent antibodies following cell fixation and permeabilization. Permeabilization must disrupt the cell membranes sufficiently to allow the passage of antibodies, while preserving the structure and protein composition of these same membranes. The problem is exacerbated at the level of the plasma membrane, which is the cellular membrane most exposed to solvents or detergents used to permeabilize cells.

Dictyostelium discoideum is a soil amoeba frequently used to study cell biology, in particular cell motility, endocytosis, cell adhesion or phagocytosis [1]. For many of these studies it is critical to determine if membrane proteins implicated in these processes are located in intracellular compartments or exposed at the cell surface. Protocols used to permeabilize and stain *Dictyostelium* cells are fundamentally similar to those used with mammalian cells, with the caveat that *Dictyostelium* membranes can be more resistant to mild permeabilizing detergents like saponin [2].

In the course of our studies, we observed that different immunofluorescence protocols detected very different levels of proteins at the cell surface. In this study we show that permeabilization procedures remove a large amount of cell surface antigens. We also propose an optimal procedure to label both the cell surface and intracellular compartments.

Methods

Cells and reagents

Dictyostelium discoideum DH1-10 cells [3] were grown at 21°C in HL5 medium (14.3 g/L Bactopeptone, 7.15 g/L Yeast Extract, 18 g/L Maltose monohydrate, 3.6 mM $Na_2HPO_4.2H_2O$ and 3.6 mM KH_2PO_4). Paraformaldehyde was purchased from by AppliChem, Saponin from Sigma and Triton X-100 was from Fluka.

The plasmid allowing expression of a fusion protein composed of the csA extracellular domain fused to the transmembrane domain of SibA and a short cytoplasmic domain (RRRSMAAA) was transfected in DH1-10 cells by electroporation. Transfected cells were then selected and grown in HL5 medium supplemented by G418 (10 μg/mL). For simplicity this fusion protein is referred to here as csA-SA. To detect csA-SA we used a mouse monoclonal antibody (41-71-21) directed to the csA extracellular domain [4]. When indicated, p23, p25 and p80 membrane proteins were detected using H194, H72, and H161 mouse monoclonal antibodies [5]. The unidentified H36 surface antigen recognized by the H36 monoclonal antibody was also described previously [6].

* Correspondence: Alexandre.Vernay@unige.ch
Department of Cell Physiology and Metabolism, Geneva Faculty of Medicine, Centre Médical Universitaire, 1 rue Michel Servet, Geneva 4 CH1211, Switzerland

Immunofluorescence

For all immunofluorescence procedures, 10^6 *Dictyostelium* cells expressing csA-SA were allowed to attach to a 22×22 mm glass coverslip for 10 minutes at room temperature in 2 mM Na_2HPO_4, 14.7 mM KH_2PO_4, pH6.0 supplemented with 0.5% HL5, 100 mM sorbitol, and 100 μM $CaCl_2$. This buffer allows optimal attachment of *Dictyostelium* cells to their substrate, while preserving optimally their general organization [7]. Cells were then fixed for 10 minutes at room temperature in PBS containing 4% paraformaldehyde, then washed in PBS containing 20 mM NH_4Cl, and in PBS containing 0.2% BSA (PBS-BSA).

In the immunofluorescence procedure referred to as "Classical", cells were then washed twice in PBS, permeabilized in methanol at −20°C for 2 seconds, washed twice in PBS and once in PBS-BSA. When indicated, methanol was replaced with Triton X-100 (0.07% in PBS for 2 minutes at room temperature) or with saponin (0.2% in PBS for 10 minutes). Permeabilized cells were incubated with a mouse anti-csA antibody in PBS-BSA for 1 hour, washed twice in PBS-BSA, incubated for 1 hour with an Alexa-488-coupled anti-mouse immunoglobulin antibody in PBS-BSA, washed twice in PBS-BSA, once in PBS and mounted in Möwiol. Cells were visualized using a LSM700 confocal microscope (Zeiss). In each experiment, pictures from different samples were taken consecutively using identical settings.

In the procedure referred to as "Surface labeling", non-permeabilized fixed cells were incubated with an anti-csA antibody in PBS-BSA for 1 hour, washed twice in PBS-BSA, incubated 1 hour with an Alexa-488-coupled anti-mouse antibody diluted in PBS-BSA. Finally, cells were washed twice in PBS-BSA, once in PBS and mounted in Möwiol.

In the procedure referred to as "Two-step" the surface of fixed cells was labeled as described above in the "Surface labeling" procedure. After surface labeling, cells were fixed again in paraformaldehyde, washed in PBS-NH_4Cl, twice in PBS-BSA, twice in PBS before permeabilization in methanol at −20°C. Permeabilized cells were rinsed twice in PBS and once in PBS-BSA. Intracellular csA was then labeled for 1 hour with a mouse anti-csA antibody diluted in PBS-BSA, washed twice in PBS-BSA and revealed using an Alexa-488-coupled anti-mouse antibody. Finally, cells were washed twice in PBS-BSA, once in PBS and mounted in Möwiol.

Findings and discussion

The csA-SA fusion protein used in this study is a single-pass type I transmembrane protein composed of the extracellular domain of the contact site A protein, fused to a single transmembrane domain and a short cytoplasmic domain. Cell surface labeling revealed that this protein was abundantly present at the surface of *Dictyostelium* cells (Figure 1A). In order to detect csA-SA both at the cell surface and in intracellular compartments, we followed a classical procedure, variations of which are most often used in many laboratories: cells were fixed with paraformaldehyde, permeabilized with methanol, then incubated sequentially with a mouse antibody against the csA moiety, and with a fluorescent secondary anti-mouse antibody. Surprisingly this procedure detected only a very small amount of protein at the cell surface (Figure 1B). This suggested that a significant amount of csA protein was lost during the procedure, particularly at the cell surface. These two procedures differ mostly by the fact that the cells are permeabilized in the latter, and previous studies have shown that a sandwich of primary and secondary antibodies can prevent loss of a cell surface protein during permeabilization [8]. Accordingly, we tested a two-step labeling procedure: cells were fixed and incubated with antibodies prior to permeabilization, then fixed again, permeabilized and intracellular antigens were labeled. This two-step procedure resulted in a prominent staining of the cell surface (Figure 1C). Together, these results indicate that the csA antigen was lost from the cell surface during cell permeabilization, unless it was stabilized by the binding of two layers of antibodies.

Since methanol solubilizes and extracts cellular lipids, it may be particularly disruptive to the integrity of biological membranes. This consideration led us to test the effect of alternative permeabilization procedures. Triton X-100 is a non-ionic detergent capable of solubilizing membrane lipids. Saponin is a mild detergent extracting cholesterol from membranes and has been reported to be less prone than methanol or triton X-100 to extracting membrane

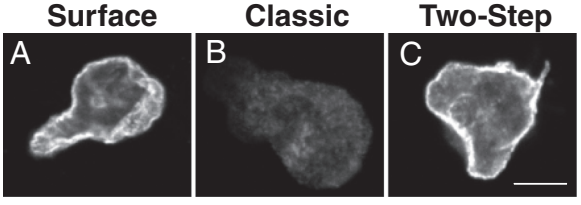

Figure 1 Cell permeabilization affects detection of a surface antigen. (A) Surface csA-SA was detected by incubating unpermeabilized fixed cells with an anti-csA antibody and a fluorescent secondary antibody. **(B)** Fixed cells were permeabilized with methanol at −20°C and then incubated sequentially with the anti-csA antibody and secondary antibodies. Surface labeling was almost entirely lost when following this classical staining procedure. **(C)** The surface of fixed cells was labeled prior to permeabilization, then cells were permeabilized and intracellular antigen stained. This two-step procedure allowed the simultaneous labeling of both surface and intracellular antigens. All the pictures presented in this figure were taken sequentially with the same microscope and with identical settings. Scale bar: 5 μm.

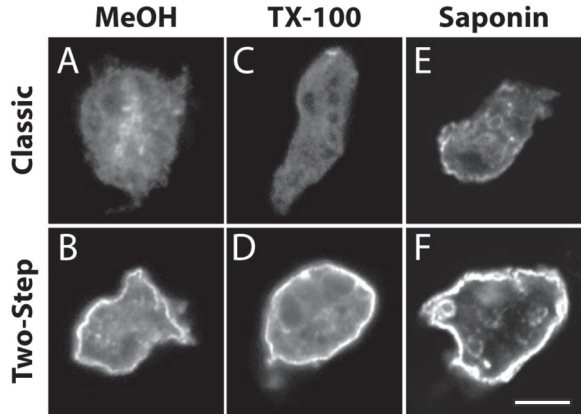

Figure 2 Three distinct permeabilization procedures affect strongly detection of surface csA-SA. Cells were treated following a classical immunofluorescence procedure (upper panel) or a two-step procedure (lower panel). Cells were permeabilized using methanol at −20°C **(A, B)** triton X-100 **(C, D)** or saponin **(E, F)**. All the pictures presented in this figure were taken sequentially with the same microscope and with identical settings. Scale bar: 5 μm.

proteins [9,10]. We observed that permeabilization with triton X-100 and saponin also resulted in a marked loss of surface csA-SA labeling (Figure 2C and E), although with the saponin a weak surface staining was still detectable (Figure 2E). In all cases, a two-step labeling procedure resulted in a prominent labeling of the cell surface (Figure 2B, D, F).

Since the csA fusion protein analyzed in this study exhibits a single transmembrane domain and a very short cytosolic domain, it may be particularly prone to be extracted from cellular membranes. In order to study this point, we stained other membrane proteins following either a classical immunofluorescence protocol, or a two-step procedure. We used for this a collection of monoclonal antibodies recognizing antigens present at the surface of *Dictyostelium* cells [5,6]. Similar to the csA-SA protein, we observed that a classical immunofluorescence procedure resulted in a strong decrease in the cell surface labeling of the p23, p25 and H36 antigens compared to a two-step procedure (Figure 3A-H). However for these three proteins, some surface protein was still detectable even after a classical immunofluorescence staining, suggesting that they were less readily extracted from the cell surface than the csA-SA protein. The p80 protein has been shown to be a polytopic protein present at a low level at the cell surface and at a higher concentration in endosomal and lysosomal compartments [5]. The surface staining of p80 was not visibly increased by prelabeling the cell surface (Figure 3I-J), suggesting that it was not extracted from cellular membranes upon permeabilization, maybe due to the fact that this protein exhibits three transmembrane domains.

In summary, we tested here three distinct procedures to permeabilize fixed cells prior to immunofluorescence staining: methanol, triton X-100 and saponin. All three methods resulted in a marked loss of cell surface labeling of the csA-SA protein. The csA-SA protein likely represents an extreme case since it is anchored to the cell membrane only by one transmembrane domain followed by a short cytoplasmic domain. When other surface proteins were tested, some (p23, p25, H36) were also largely extracted from the cell surface although they remained detectable. On the contrary, p80, maybe due to its three transmembrane domains, was not detectably extracted from the cell surface upon permeabilization.

These results suggest that when assessing the surface localization of a protein by immunofluorescence, it is best to compare results obtained using several alternative protocols in order to ascertain that no loss of labeling is caused by the permeabilization procedure. Ideally, a surface

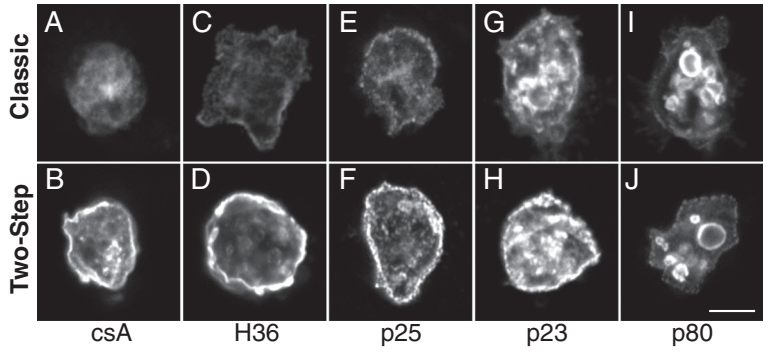

Figure 3 Loss of different surface antigens during permeabilization. Cells were labeled following a classical immunofluorescence procedure (upper panel) or a two-step procedure (lower panel). In both cases, cells were permeabilized with methanol at −20°C. The antibodies used detected csA-SA **(A, B)**, H36 **(C, D)**, p25 **(E, F)**, p23 **(G, H)** or p80 **(I, J)**. The effect of cell permeabilization on surface labeling differed for the various antigens considered. Scale bar: 5 μm.

immunofluorescence of non-permeabilized cells should be performed. In some situations, it will be difficult to detect reliably a protein of interest at the cell surface, for example if no antibodies directed to the extracellular domain of the protein are available. It may then be necessary to define the most adequate compromise to perform immunofluorescence detection: one option is to use mild detergents like saponin to reduce the amount of protein lost from the cell surface upon permeabilization. Using very low concentrations of detergents may be an alternative approach, and sufficient permeabilization may even be achieved simply by paraformaldehyde fixation with no further permeabilization [11]. It should however be kept in mind that very mild permeabilization procedures may result in incomplete permeabilization of some cellular membranes, as shown previously for saponin permeabilization in *Dictyostelium* [2]. Use of alternative methodological approaches (e.g. cell surface biotinylation followed by biochemical analysis or expression of GFP-tagged proteins in live cells) not sensitive to the same type of artifacts may be necessary to detect and quantify unambiguously the presence of a protein at the cell surface.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
AV and PC contributed to the conception and the design of the study. AV performed the experiments, which were analyzed by AV and PC. AV and PC contributed to writing and final approval of the manuscript.

Acknowledgments
This research was supported by a grant from the Swiss National Science Foundation (number 31003A-135789 to P.C.).

Received: 8 May 2013 Accepted: 9 August 2013
Published: 12 August 2013

References
1. Cosson P, Soldati T: **Eat, kill or die: when amoeba meets bacteria.** *Curr Opin Microbiol* 2008, **11:**271–276.
2. Mercanti V, Cosson P: **Resistance of Dictyostelium discoideum membranes to saponin permeabilization.** *BMC Res Notes* 2010, **3:**120.
3. Cornillon S, *et al*: **Phg1p is a nine-transmembrane protein superfamily member involved in dictyostelium adhesion and phagocytosis.** *J Biol Chem* 2000, **275:**34287–34292.
4. Bertholdt G, *et al*: **Carbohydrate and other epitopes of the contact site A glycoprotein of Dictyostelium discoideum as characterized by monoclonal antibodies.** *Cell Differ* 1985, **16:**187–202.
5. Ravanel K, *et al*: **Membrane sorting in the endocytic and phagocytic pathway of Dictyostelium discoideum.** *Eur J Cell Biol* 2001, **80:**754–764.
6. Mercanti V, *et al*: **Selective membrane exclusion in phagocytic and macropinocytic cups.** *J Cell Sci* 2006, **119:**4079–4087.
7. Smith EW, *et al*: **Effect of starvation on the endocytic pathway in Dictyostelium cells.** *Eukaryotic cell* 2010, **9:**387–392.
8. Lamvik J, *et al*: **Nonlabeled secondary antibodies augment/maintain the binding of primary, specific antibodies to cell membrane antigens.** *Cytometry* 2001, **45:**187–193.
9. Goldenthal KL, *et al*: **Postfixation detergent treatment for immunofluorescence suppresses localization of some integral membrane proteins.** *J Histochem Cytochem: Official J Histochem Soc* 1985, **33:**813–820.
10. Lecoeur H, *et al*: **A cytofluorometric method for the simultaneous detection of both intracellular and surface antigens of apoptotic peripheral lymphocytes.** *J Immunol Methods* 1998, **217:**11–26.
11. Hinners I, *et al*: **The orientation of membrane proteins determined in situ by immunofluorescence staining.** *Anal Biochem* 1999, **276:**1–7.

doi:10.1186/1756-0500-6-317
Cite this article as: Vernay and Cosson: Immunofluorescence labeling of cell surface antigens in *Dictyostelium*. *BMC Research Notes* 2013 **6**:317.

B. Publication: TM9 family proteins control surface targeting of glycine-rich transmembrane domains

1. Introduction

Due to its haploid genome, *Dictyostelium* is an outstanding model for the study of genes function. Our former studies revealed the role of the Phg1a protein in phagocytosis. Indeed, genetic disruption of Phg1a resulted in a decrease in the phagocytosis of latex beads and bacteria. This defect was actually due to a decreased adhesion capacity, which is the first step of the phagocytosis, raising the hypothesis that Phg1a was an adhesion molecule (Cornillon et al., 2000). Phg1a belong to the highly conserved TM9 SuperFamily (TM9SF), characterized by a large N-terminal domain, 9 transmembrane domains and a C-terminal domain. Three TM9 proteins exist in *Dictyostelium* (named Phg1a, b and c), in *Saccharomyces cerevisiae* (named TMN1, 2 and 3), in *Drosophila melanogaster* (TM9SF2, 3 and 4) and four members in human (TM9SF1, 2, 3 and 4). In *Dictyostelium* and human, based on sequence analysis, Phg1a, TM9SF2 and TM9SF4 belong to the group A of TM9 proteins, Phg1b, TM9SF1 and TM9SF3 belong to the group B, and Phg1c is defined as non-A and non-B (See Figure 21).

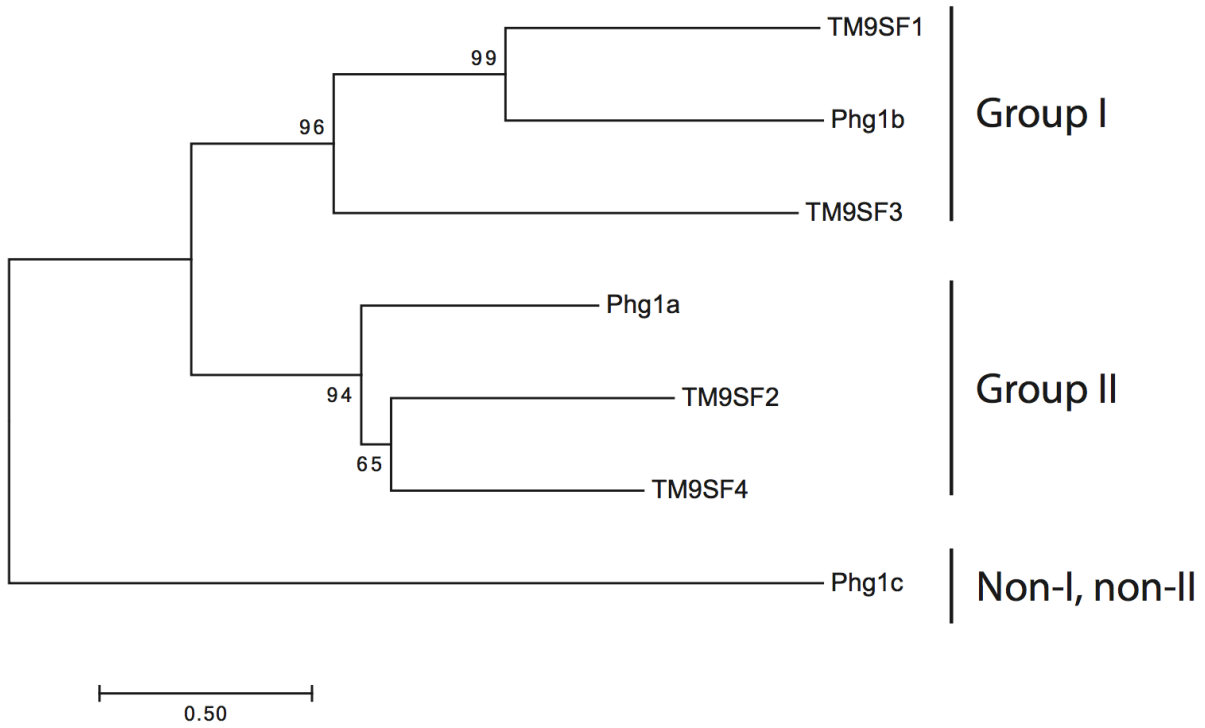

Figure 21: Phylogenetic tree of the TM9 proteins family. This tree has been obtained using the Mega7 software, by comparing the aminoacid sequences of human (TM9SF1, TM9SF2, TM9SF3 and TM9SF4) and *Dictyostelium* (Phg1a, Phg1b and Phg1c) TM9 proteins. As indicated TM9 proteins can be divided into 3 groups. TM9SF1, TM9SF3 and Phg1b belong to the group I. TM9SF2, TM9SF4 and Phg1a belong to the group II, whereas Phg1c cannot be integrated in one of these groups, and is therefore considered as non-I and non-II.

Later studies revealed that the cell surface composition was changed in the absence of Phg1a (Benghezal et al., 2003). This suggests that Phg1a is not an adhesion molecule, but rather a regulator of cell adhesion. Later studies indeed showed that the cell surface expression of SibA, an adhesion molecule presenting characteristics of β integrins, was largely decreased in cells deficient for Phg1a. This raised the possibility that Phg1a was controlling the sorting of SibA in the endocytic pathway (Froquet et al., 2012).

Besides its defect in phagocytosis, cells deficient for Phg1a showed a decreased killing rate of *Klebsiella* (Benghezal et al., 2006). This is actually due to the decreased

amount and stability of the Golgi sulfotransferase Kil1, which is thought to regulate the activity of yet unidentified effectors that directly kill *Klebsiella* (Le Coadic et al., 2013).

Finally, the *Drosophila* orthologs of Phg1a, TM9SF2 and TM9SF4, have been described to be required for the plasma membrane targeting of the PGRP-LC phagocytic receptor (Perrin et al., 2015).

Taken together, these data show that Phg1a regulates the adhesion, the phagocytosis and the killing in *Dictyostelium* cells by controlling the cell surface expression of the adhesion molecule SibA, and the stability of the Golgi sulfotransferase Kil1. This may indicate that Phg1a regulates directly the sorting of these proteins.

During my PhD, I had the opportunity to be co-first author in a study that focused on the role of Phg1a in the sorting of SibA. More precisely, we showed that Phg1a was responsible for the transport of SibA to the cell surface, and that SibA was interacting with Phg1a through its transmembrane domain (TMD). One notable feature of the TMD of SibA is the presence of five glycine residues. These glycine residues were sufficient to confer selectivity towards Phg1a, as the mutation of these residues into leucine abolished the association and the transport of SibA by Phg1a. Moreover, the presence of five glycine residues in the TMD of a chimeric protein was sufficient to induce its specific recognition by Phg1a and Phg1a-dependent transport. These features were conserved through evolution, since TM9SF4, one ortholog of Phg1a in human exhibits similar characteristics. Taken together, these results suggest that Phg1a/TM9SF4 acts as cargo receptor, which transports glycine-rich TMDs along the secretion pathway.

2. Publication

TM9 family proteins control surface targeting of glycine-rich transmembrane domains

Jackie Perrin*, Marion Le Coadic*, Alexandre Vernay*, Marco Dias, Navin Gopaldass, Hajer Ouertatani-Sakouhi and Pierre Cosson

* These authors contributed equally to this work.

Journal of Cell Science (2015) 128, 2269-2277

In this publication, I contributed to the experiments in *Dictyostelium*, and more particularly to the Figure 1, Figure 2 and Figure 4. I almost entirely performed the part of the study concerning mammalian cells (Figure 5, Figure S2 and Figure S3). I was also implicating in the writing of the manuscript.

© 2015. Published by The Company of Biologists Ltd | Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

RESEARCH ARTICLE

TM9 family proteins control surface targeting of glycine-rich transmembrane domains

Jackie Perrin[1,*], Marion Le Coadic[1,*], Alexandre Vernay[1,*], Marco Dias[1], Navin Gopaldass[2], Hajer Ouertatani-Sakouhi[1] and Pierre Cosson[1,‡]

ABSTRACT

TM9 family proteins (also named Phg1 proteins) have been previously shown to control cell adhesion by determining the cell surface localization of adhesion proteins such as the *Dictyostelium* SibA protein. Here, we show that the glycine-rich transmembrane domain (TMD) of SibA is sufficient to confer Phg1A-dependent surface targeting to a reporter protein. Accordingly, in *Dictyostelium phg1A*-knockout (KO) cells, proteins with glycine-rich TMDs were less efficiently transported out of the endoplasmic reticulum (ER) and to the cell surface. Phg1A, as well as its human ortholog TM9SF4 specifically associated with glycine-rich TMDs. In human cells, genetic inactivation of TM9SF4 resulted in an increased retention of glycine-rich TMDs in the endoplasmic reticulum, whereas TM9SF4 overexpression enhanced their surface localization. The bulk of the TM9SF4 protein was localized in the Golgi complex and a proximity-ligation assay suggested that it might interact with glycine-rich TMDs. Taken together, these results suggest that one of the main roles of TM9 proteins is to serve as intramembrane cargo receptors controlling exocytosis and surface localization of a subset of membrane proteins.

KEY WORDS: Phg1, Secretory pathway, Sorting, TM9 protein, Transmembrane domain

INTRODUCTION

Localization of membrane receptors at the surface of eukaryotic cells is often essential for them to perform their specific functions, such as cellular adhesion (e.g. integrins), nutrients capture (e.g. transferrin receptor) or signal transduction (e.g. EGF receptor). Along the exocytic pathway, transport of newly synthesized proteins from the endoplasmic reticulum (ER) to the cell surface is tightly controlled to ensure the selective delivery of functional protein complexes (Geva and Schuldiner, 2014). Similarly, selective and regulated endocytosis of surface receptors controls surface residency of membrane proteins, allowing adaptation of cellular physiology to changes in the environment.

Our previous studies have revealed the role of a new class of proteins in the control of cell surface localization of membrane proteins. Both Phg1A (Cornillon et al., 2000), Phg1B (Benghezal et al., 2003) and SadA (Fey et al., 2002) were initially characterized in *Dictyostelium* as membrane proteins necessary for efficient cellular adhesion and phagocytosis. Later studies showed that these three proteins are essential for the efficient surface localization of SibA, a cell surface adhesion molecule with integrin β features (Cornillon et al., 2006; Froquet et al., 2012). The Phg1 family, also referred to as the TM9 family, is characterized by a high degree of sequence similarity, an N-terminal luminal domain preceded by a signal sequence and followed by nine transmembrane domains. There are three members in the Phg1/TM9 family in *Dictyostelium* (Phg1A, Phg1B and Phg1C) (Benghezal et al., 2003), three in *S. cerevisiae* [TMN1 (also known as Emp70), TMN2 and TMN3] (Froquet et al., 2008), three in *Drosophila* (TM9SF2, TM9SF3 and TM9SF4) (Bergeret et al., 2008) and four in humans (TM9SF1 to TM9SF4) (Chluba-de Tapia et al., 1997; Schimmöller et al., 1998). In *Drosophila*, TM9SF2 and TM9SF4, which are highly similar to *Dictyostelium* Phg1A, have also been shown to control phagocytosis by determining the cell surface expression of the phagocytic receptor PGRP-LC (Perrin et al., 2015). Intriguingly, SadA, which is also necessary for efficient cell surface targeting of SibA, exhibits the same general organization as Phg1/TM9 proteins (one signal sequence followed by a large extracellular domain and nine transmembrane domains), but shows no sequence homology to Phg1/TM9 proteins.

Here, we studied the mechanism by which TM9 proteins control surface localization of membrane proteins like SibA. Our results indicate that the transmembrane domain (TMD) of SibA is sufficient to confer Phg1A-dependent surface localization to a reporter protein. This property is due to the presence of glycine residues in the TMD of SibA, to which Phg1A specifically associates. Human TM9SF4 shows the same propensity to associate with glycine-rich TMDs and to ensure their localization at the cell surface. This study suggests that TM9 proteins function as cargo receptors ensuring surface localization of proteins harboring glycine-rich transmembrane domains.

RESULTS

Surface localization of glycine-rich TMDs is dependent on Phg1A

Previous experiments have demonstrated that in *Dictyostelium*, a chimera composed of the extracellular domain of contact site A (csA) fused to the transmembrane and cytosolic domain of SibA is localized at the cell surface less efficiently in *phg1A*-knockout (KO) cells than in wild-type (WT) cells (Froquet et al., 2012). To identify the feature of SibA that confers differential sorting in WT and *phg1A* KO cells, we expressed in these two cell lines a chimeric protein composed of the csA extracellular domain fused to the TMD of SibA and to a very short cytosolic domain (denoted csA-A5G) (Fig. 1A, see also Table 1). The surface localization of the csA

[1]Department of Cell Physiology and Metabolism, Faculty of Medicine, University of Geneva, 1 rue Michel Servet, Geneva 4 CH-1211, Switzerland. [2]Department of Biochemistry, Sciences II, University of Geneva, 30 quai Ernest-Ansermet, Geneva 4 CH-1211, Switzerland.
*These authors contributed equally to this work

‡Author for correspondence (Pierre.Cosson@unige.ch)

Received 23 October 2014; Accepted 30 April 2015

Journal of Cell Science

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

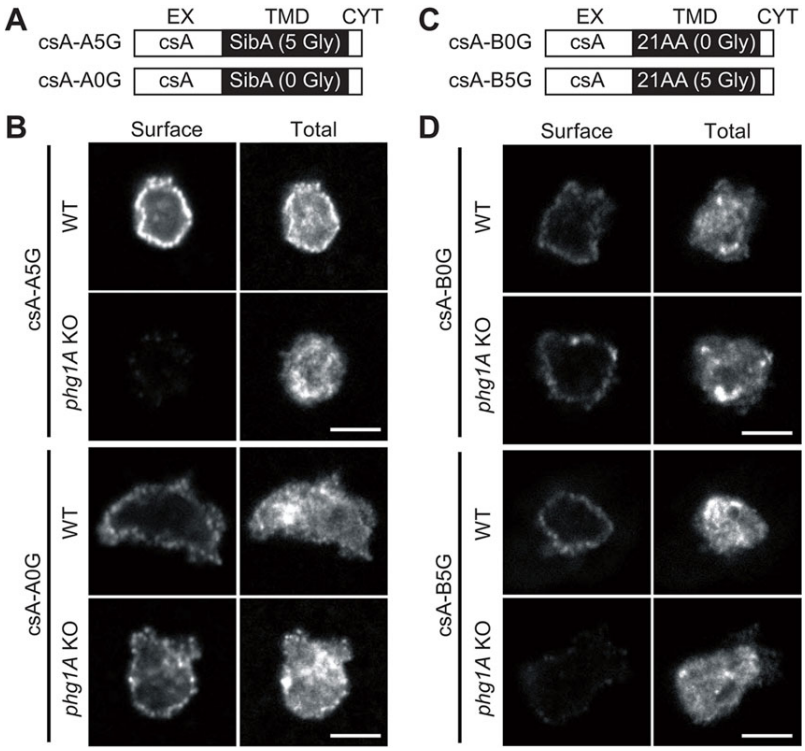

Fig. 1. Phg1A ensures efficient cell surface localization of proteins harboring the SibA glycine-rich TMD. All pictures were taken with the same confocal microscope (Zeiss LSM700) and the same setting allowing direct comparison. Scale bar: 5 µm. (A) The csA-A fusion proteins are composed of the extracellular domain of csA, the glycine-rich TMD of SibA (csA-A5G) or a mutated form without glycine residues (csA-A0G), and a short cytoplasmic tail (see Table 1). (B) Fusion proteins were labeled before (Surface) and after (Total) permeabilization by immunofluorescence in WT or *phg1A* KO cells, using an antibody specific for the csA extracellular domain. (C) CsA-B fusion proteins are composed of the extracellular domain of csA, a hydrophobic TMD without glycine residues (csA-B0G) or a mutated form with five glycine residues added (csA-B5G), followed by a short cytoplasmic tail (see also Table 1). (D) Fusion proteins were expressed in *Dictyostelium* WT or *phg1A* KO cells and labeled before (Surface) and after (Total) permeabilization by immunofluorescence.

fusion proteins was assessed by immunofluorescence. For this, we labeled, with different fluorescent antibodies in non-permeabilized cells, the csA fusion protein exposed at the cell surface and, after permeabilization, the total cellular csA (surface+intracellular) (Fig. 1B). When cells with similar total expression levels of csA were compared, the cell surface localization of csA-A5G was readily detectable in WT cells, but was much lower in *phg1A* KO cells (Fig. 1B). This result indicated that the TMD of SibA is sufficient to render the surface targeting of a reporter membrane protein dependent on Phg1A.

The most remarkable feature of the SibA TMD is the presence of five glycine residues, conserved in SibB, SibC, SibD and SibE (Cornillon et al., 2006). When these five residues were mutated to leucine (Fig. 1A; Table 1), the resulting fusion protein (csA-A0G) was targeted to the cell surface as efficiently in WT and *phg1A* KO cells (Fig. 1B). This observation suggests that the multiple glycine residues in the SibA TMD are necessary for Phg1A-dependent surface localization of the protein.

To test this hypothesis further, we assessed the surface localization of csA-B0G, a fusion protein with a 21-residue hydrophobic TMD containing no glycine residues derived from the human CD1b molecule (Mercanti et al., 2010) (Fig. 1C; Table 1). As described previously (Froquet et al., 2012), we observed that this protein is present at the surface of both WT and *phg1A* KO cells at similar levels (Fig. 1D). We then introduced five glycine residues in the TMD of csA-B0G (Fig. 1C; Table 1), and assessed the surface localization of the resulting fusion protein (csA-B5G) in WT and *phg1A* KO cells. CsA-B5G was present at the surface of WT cells, but it was detected at very low levels at the surface of *phg1A* KO cells (Fig. 1D), suggesting that the presence of glycine residues is sufficient to make surface targeting of a TMD dependent on Phg1A.

Table 1. Amino acids sequence of the transmembrane and cytosolic domains of the csA and Tac chimeric proteins

Name	Luminal domain	Transmembrane domain	Cytoplasmic domain
csA-A5G	--- GT TKENNNK	TVLTGAIAGAAAGAGLLAAGAWFLL	RRRSMAAA
csA-A0G	--- GT TKENNNK	----L---L---L-L----L-----	RRRSMAAA
csA-B0G	--- GT DILYWRNPTESD	SIVLAIIVPSLLLLLCLALLWYM	RRRSMAAA
csA-B3G	--- GT DILYWRNPTESD	-----G--G---G----------	RRRSMAAA
csA-B4G	--- GT DILYWRNPTESD	-----G--G---G--G-------	RRRSMAAA
csA-B5G	--- GT DILYWRNPTESD	-----G--G---G--G---G---	RRRSMAAA
csA-C1G	--- GT DILYWRNPTESDL	QVAVAACVFLLIAVLLLSGLTWQL	RRRSMAAA
Tac-C1G	--- EY	QVAVAACVFLLIAVLLLSGLTWQ	RRQRKSRRTI
Tac-C3G	--- EY	-----G---G--------G----	RRQRKSRRTI
Tac-C4G	--- EY	-----G---G--G-----G----	RRQRKSRRTI
Tac-C5G	--- EY	-----G---G--G---G-G----	RRQRKSRRTI
Tac-C6G	--- DL	-----G---G--G---G-G----G	RRQRKSRRTI
Tac-C1G-KKxx	--- EY	QVAVAACVFLLIAVLLLSGLTWQ	RRQRLETFKKTN

Abbreviations list: csA, contact site A; Sib, Similar to integrin β; TM9SF, transmembrane 9 superfamily; TMD, transmembrane domain.

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

In the experiments described above, cells with similar total expression levels were selected, to allow meaningful comparison between different cells. To obtain more quantitative data, we analyzed fluorescently labeled cells by flow cytometry. This technique measures surface and total labeling simultaneously in each individual cell. This analysis revealed the existence of a linear relationship between the total cellular level of csA protein, and the level found at the cell surface (Fig. 2A): higher total levels of csA protein resulted in higher levels at the cell surface. When comparing WT and *phg1A* KO cells, it is also apparent that at equivalent expression levels, a higher amount of csA-A5G is detected at the surface of WT cells than at the surface of *phg1A* KO cells (Fig. 2A). Based on these results, for each chimeric protein, we calculated the relative efficiency of surface targeting in WT and *phg1A* KO cells (surface csA in WT cells divided by surface csA in *phg1A* KO cells). Accordingly, csA-A5G, which exhibits the glycine-rich TMD of SibA, was targeted three times more efficiently to the surface of WT cells than to the surface of *phg1A* KO cells (Fig. 2B). By contrast, three different fusion proteins (csA-A0G, csA-B0G and csA-C1G) with a hydrophobic TMD exhibiting at most one glycine residue were localized as efficiently to the surface of WT and of *phg1A* KO cells (Fig. 2B). Addition of five glycine residues to the csA-B0G TMD (csA-B5G) made it dependent on Phg1A for cell surface targeting, whereas surface localization of proteins with three or four glycine residues (csA-B3G and csA-B4G) was partially dependent on Phg1A (Fig. 2B).

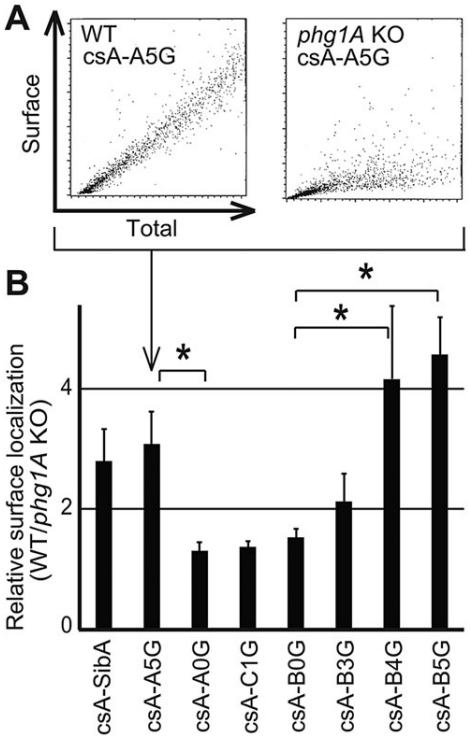

Fig. 2. Efficient surface targeting of glycine-rich TMDs by Phg1A. (A) Analysis of total expression and cell surface levels of csA-A5G chimeras in WT (left panel) and *phg1A* KO *Dictyostelium* cells (right panel) by flow cytometry. For each cell, the total and the surface levels of csA fusion protein were determined. In order to extrapolate a numeric value from the dot plot analysis, a linear regression was used. (B) We calculated the surface targeting (in arbitrary units) by dividing the slope obtained in A for WT cells and for *phg1A* KO cells. The mean±s.e.m. of at least eight experiments are indicated. **P*<0.01 (Student's *t*-test).

To verify the validity of these observations, the presence of csA-A5G and csA-A0G at the surface of WT and *phg1A* KO cells was also assessed biochemically. For this, cell surface proteins were labeled with biotin and purified on neutravidin beads. The amount of csA in the cell (total cell lysate) and at the surface (biotinylated proteins) was then assessed by western blotting, using antibodies specific for the csA extracellular domain. One representative experiment is shown (Fig. 3A), which indicated that 20.5% of csA-A5G was present at the surface of WT cells, but only 7.0% at the surface of *phg1A* KO cells (see supplementary material Fig. S1A for

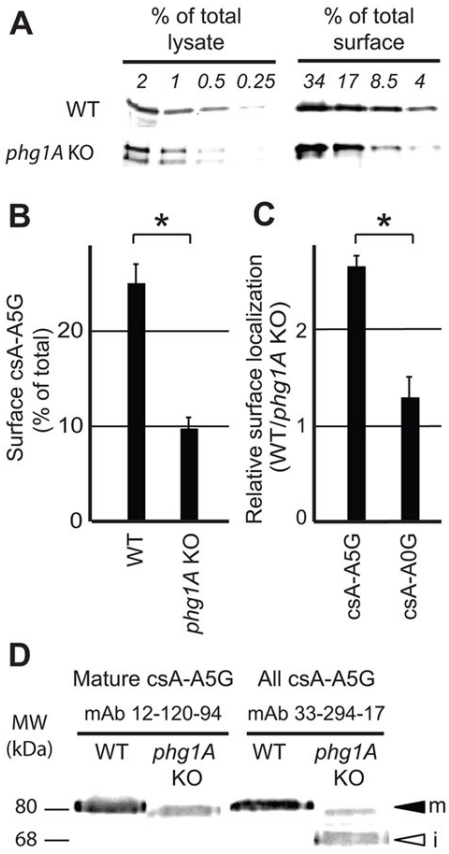

Fig. 3. Biochemical analysis confirms differential surface localization of glycine-rich TMDs in WT versus *phg1A* KO cells. (A) Following cell surface biotinylation, surface csA-A5G fusion protein was purified on neutravidin beads, and its level compared to the total cellular level by western blotting after serial dilution. The percentage of total or surface proteins loaded on each lane is indicated. Based on these results, the percentage of csA-A5G fusion protein present at the cell surface was determined in WT (20.5% of total) and *phg1A* KO cells (7% of total) (see supplementary material Fig. S1A for details of the quantification). (B) The mean±s.e.m. of three independent experiments as described in A was determined. CsA-A5G was significantly more abundant at the surface of WT cells than at the surface of *phg1A* KO cells. (C) Relative targeting of each fusion protein in WT versus *phg1A* KO cells was obtained by dividing the percentage of surface csA in WT cells by that in *phg1A* KO cells (e.g. 20.5/7.0 in A). The mean±s.e.m. of at least three independent experiments are indicated. Cell surface localization of csA-A5G, but not of csA-A0G, was decreased in *phg1A* KO cells relative to WT cells. **P*<0.01 (Student's *t*-test). (D) Export of csA-A5G out of the ER is affected in *phg1A* KO cells. Cellular lysates of WT or *phg1A* KO *Dictyostelium* cells were analyzed by western blotting using antibody 12-120-94, which detects only mature csA, and antibody 33-294-17, which detects both mature and immature csA. The mature csA-A5G (m) exhibited a molecular mass of 80 kDa, whereas the low-molecular-mass form (68 kDa) detected in *phg1A* KO cells corresponded to the immature ER form (i).

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

details of the quantification procedure). Quantification of three independent experiments confirmed that csA-A5G was significantly more abundant at the surface of WT cells than at the surface of *phg1A* KO cells (Fig. 3B). Based on these results, we calculated the relative surface targeting in WT and *phg1A* KO cells and confirmed that surface localization of csA-A5G was dependent on Phg1A (2.5 times more surface csA in WT than in *phg1A* KO cells) (Fig. 3C). By contrast, surface targeting of csA-A0G was similar in WT and in *phg1A* KO cells (equivalent surface levels of csA in WT and *phg1A* KO cells) (Fig. 3C).

Taken together, these results demonstrate that efficient surface targeting of a membrane protein exhibiting a glycine-rich TMD requires the presence of the Phg1A protein.

Efficient ER exit of proteins with glycine-rich TMDs is dependent on Phg1A

The csA protein is inserted co-translationally in the ER, where it is rapidly converted into a glycosylated 68-kDa protein. It is then further modified as it passes through the Golgi complex, to reach its final molecular mass of 80 kDa (Hohmann et al., 1985). For most csA fusion proteins, at steady state, an 80-kDa form, presumably corresponding to the mature form was predominant, and the immature 68-kDa form was hardly detectable (see for example supplementary material Fig. S1A). By contrast, as can be seen in Fig. 3A, for csA-A5G, the 68-kDa form, presumably corresponding to the immature form of csA, was abundant in *phg1A* KO cells. Quantification of several independent experiments revealed that 17±8% of csA-A5G is immature in WT cells versus 57±2% in *phg1A* KO cells (mean±s.e.m.; $P<0.001$; $n=5$). To verify that the lower-molecular-mass csA-A5G protein detected in *phg1A* KO cells was truly an immature form, we used a monoclonal antibody (12-120-94) which recognizes specifically the mature form of csA (Ochiai et al., 1982). As expected, only the 80-kDa form of csA-A5G was detected by the 12-120-94 antibody (Fig. 3D), confirming that the lower-molecular-mass form corresponds to an immature protein present in the ER.

A low level of mature csA fusion protein might indicate that the protein is retained in the ER, or that it is rapidly degraded after exiting the ER. To distinguish between these two possibilities, we measured the stability of csA fusion proteins in cells where protein synthesis was inhibited. These experiments revealed no significant difference between the stability of csA-A5G or csA-A0G in WT versus *phg1A* KO cells (supplementary material Fig. S1B,C). Taken together, these experiments indicate that csA-A5G is transported inefficiently out of the ER in *phg1A* KO cells, suggesting that Phg1A facilitates transport of glycine-rich TMDs at early stages of the secretory pathway.

Phg1A specifically interacts with glycine-rich TMDs

Previous attempts to detect an interaction between SibA and Phg1A have been unsuccessful (Froquet et al., 2012). However, a transient or low-affinity interaction would easily escape detection. Recently, the *Drosophila* TM9SF2 and TM9SF4 have been shown to control the cell surface expression of PGRP-LC, and to interact with it (Perrin et al., 2015). To establish more finely whether Phg1A associates specifically with glycine-rich TMDs, we used a system previously designed to assess the interactions between transmembrane proteins (Cosson et al., 1991): the *phg1A* gene was fused to the coding sequence of β-galactosidase (Phg1A-Gal) in a vector allowing transient expression in transfected COS7 cells. A reporter protein, the α-subunit of the interleukin-2 receptor (Tac antigen) was engineered to introduce three to six glycine residues in its TMD (T-C3G, T-C4G, T-C5G and T-C6G, see Table 1), and these various Tac fusion proteins were co-expressed with Phg1A-Gal (Fig. 4A). After immunoprecipitation of the Tac antigen, the amount of co-precipitated β-galactosidase was measured and compared to the amount of total cellular β-galactosidase. Only one glycine residue is present in the Tac TMD (T-C1G), and less than 0.5% of the β-galactosidase was co-precipitated with T-C1G (Fig. 4B). The

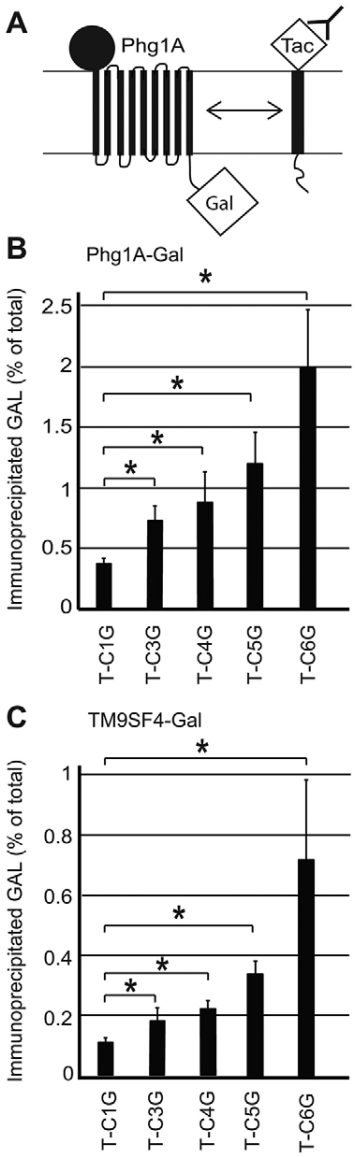

Fig. 4. *Dictyostelium* Phg1A and human TM9SF4 associate preferentially with glycine-rich TMDs. (A) To reveal a putative association of Phg1A with glycine-rich TMDs, COS7 cells were co-transfected with plasmids encoding the Phg1A protein fused to β-galactosidase (Gal) and Tac fusion proteins fused to various TMDs. Tac fusion proteins were immunoprecipitated and the amount of co-precipitated β-galactosidase activity assessed to reveal the degree of association with Phg1A. (B) Phg1A–Gal was co-expressed with T-C1G (one glycine in the TMD) or with Tac mutants with three (T-C3G), four (T-C4G), five (T-C5G) or six (T-C6G) glycine residues in their TMD. Addition of glycine residues in the TMD of Tac gradually increased its interaction with Ph1A–Gal. The mean±s.e.m. of at least six experiments are indicated. (C) Interaction between human TM9SF4 and glycine-rich TMDs was determined as described in B. A specific interaction was detected between TM9SF4 and glycine-rich TMDs. The mean±s.e.m. of at least eight experiments are indicated. *$P<0.05$ (Student's *t*-test).

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

amount of co-precipitated β-galactosidase increased gradually when three, four, five or six glycine residues were introduced (Fig. 4B). Interestingly, TM9SF4, the human ortholog of Phg1A, showed a virtually identical pattern of association with glycine-rich TMDs (Fig. 4C), suggesting that it is also capable of interacting specifically with glycine-rich TMDs, and that the role of TM9 proteins in intracellular sorting might be conserved from *Dictyostelium* to human cells.

TM9SF4 escorts glycine-rich TMDs to the cell surface in human cells

To assess the role of TM9SF4 in intracellular transport in human cells, we created three independent targeted CRISPR/Cas9-knockout cell lines by inactivating the *TM9SF4* gene in HEK293T cells (supplementary material Fig. S2A). In these three knockout cell lines, a fraction of the cells appeared multinucleated (supplementary material Fig. S2B), an observation similar to that made in *Dictyostelium*, where inactivation of two members of the TM9 family resulted in a cytokinesis defect (Benghezal et al., 2003). In these *TM9SF4* KO cells, as well as in cells overexpressing a Flag-tagged TM9SF4, the general organization of the ER, as visualized by expressing an ER-targeted fluorescent protein (ER–YFP), appeared unperturbed (supplementary material Fig. S2C). Similarly, the general organization of the Golgi complex visualized with an antibody against giantin (also known as GOLGB1) was not altered (supplementary material Fig. S2D).

We next expressed in WT or *TM9SF4* KO cells a series of chimeric proteins exhibiting an increasing number of glycine residues in their TMDs (Table 1), and we determined by immunofluorescence their intracellular localization. Fewer than five glycine residues allowed massive localization of Tac at the cell surface. By contrast, the presence of five or six glycine residues in the TMD (T-C5G or T-C6G, respectively) reduced strongly their surface localization (supplementary material Fig. S3A). The intracellular T-C5G and T-C6G were found mostly in the ER, as can be deduced from their perinuclear localization, and as assessed by colocalization with a marker of the ER (ER–YFP) (supplementary material Fig. S3B). This result indicates that glycine residues in the TMD of membrane proteins can affect their surface targeting in mammalian cells.

We focused our subsequent studies on the transport of T-C6G to the cell surface, because this protein associated most efficiently with TM9SF4 (see Fig. 4C). The intracellular localization of T-C6G was assessed in WT or *TM9SF4* KO cells expressing similar levels of fusion protein. In both cell types, the majority of T-C6G was found in the ER (Fig. 5A; supplementary material Fig. S3B). The surface level of T-C6G was weak in WT cells, and appeared even weaker in *TM9SF4* KO cells (Fig. 5A). To quantify these observations, the surface level of T-C6G was determined relative to the total expression level in individual cells, and it was indeed significantly reduced in *TM9SF4* KO cells compared to parental cells (Fig. 5B). Conversely, when TM9SF4 was overexpressed in parental and in *TM9SF4* KO, cell surface targeting of T-C6G was significantly increased (Fig. 5B). Taken together, these observations suggest that, like *Dictyostelium* Phg1A, human TM9SF4 associates specifically with glycine-rich TMDs and ensures that they are transported to the cell surface. Interestingly, neither genetic inactivation of *TM9SF4*, nor overexpression of TM9SF4 affected the ER localization of a Tac chimeric protein that was retained in the ER by virtue of a cytosolic dilysine motif (supplementary material Fig. S3C). This result indicates that no general defect in ER retention is observed in *TM9SF4* KO cells or in cells overexpressing TM9SF4. Similarly, we did not detect the induction of an ER stress response in *TM9SF4* KO cells or in cells overexpressing TM9SF4 (supplementary material Fig. S4). Thus, TM9SF4 controls the intracellular transport of a small set of proteins, without altering the general organization and function of the ER.

In order to assess further the function of the TM9SF4 protein, we expressed a tagged version of TM9SF4. Overexpression of the Flag-tagged TM9SF4 also increased cell surface expression of T-C6G (data not shown), indicating that it functions like the wild-type TM9SF4 protein. By assessing immunofluorescence, we found that

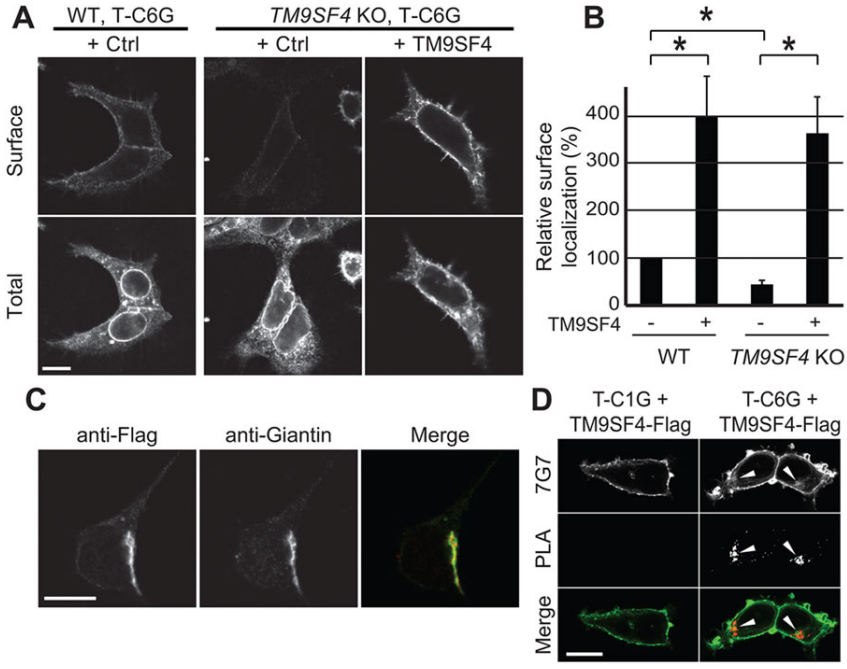

Fig. 5. TM9SF4 controls surface localization of glycine-rich TMDs in human cells. (A) A Tac protein with six glycine residues in its TMD (T-C6G) was expressed in parental (WT) or *TM9SF4* KO mammalian HEK293T cells. The protein was labeled before (Surface) and after (Total) cell permeabilization. Large amounts of T-C6G were detected in the ER in both cell types, as evidenced by the clearly visible nuclear envelope. Lower cell surface levels were observed in *TM9SF4* KO cells than in parental cells. Moreover, overexpression of TM9SF4 (+TM9SF4) increased strongly cell surface levels of T-C6G. All the pictures presented here were taken sequentially with identical settings. (B) Quantification of the surface targeting of T-C6G in parental and *TM9SF4* KO HEK293T cells, and in cells overexpressing TM9SF4. The mean±s.e.m. of four independent experiments are presented. **P*<0.01 (Student's *t*-test). (C) A Flag-tagged version of TM9SF4 was expressed in HEK293T cells, and detected by immunofluorescence. It colocalized with giantin, a marker of the Golgi complex. (D) T-C1G (upper panel) or T-C6G (lower panel) were co-expressed with a Flag-tagged version of TM9SF4 in HEK293T WT cells. A proximity ligation assay (PLA) revealed a specific signal in the Golgi complex when T-C6G and TM9SF4–Flag were co-expressed (arrowheads). Scale bars: 10 µm.

Journal of Cell Science

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

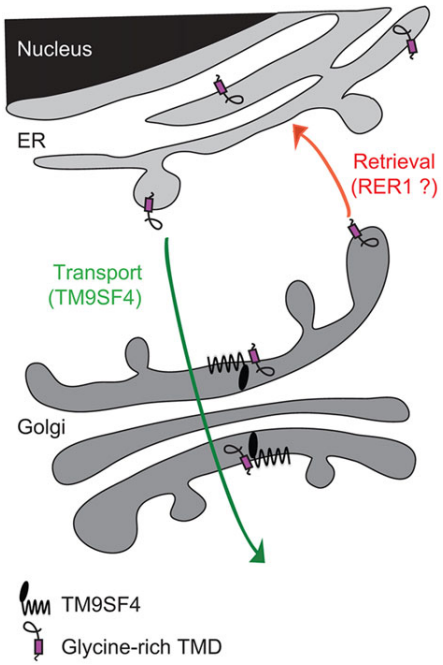

Fig. 6. A speculative model of the role of TM9SF4 in the sorting of glycine-rich TMDs. Glycine-rich TMDs can ensure ER localization of a membrane protein, presumably by interacting with the well-characterized Rer1 retrieval receptor. Human TM9SF4 is localized in the Golgi complex and associates specifically with glycine-rich TMDs to ensure their transport to the cell surface.

Flag-tagged TM9SF4 was present in a juxtanuclear compartment, where it was colocalized with giantin (also known as ???), a marker of the Golgi complex (Fig. 5C).

Given that TM9SF4 is present in the Golgi complex, whereas the majority of the T-C6G was found in the ER (in WT cells) or at the cell surface (in cells overexpressing TM9SF4), we tested directly whether a fraction of T-C6G colocalized with TM9SF4. For this, we used a proximity-ligation assay, which detects very close proximity between two proteins (below 30 nm) (Söderberg et al., 2006). A small fraction of T-C6G was detected in close proximity with TM9SF4 (Fig. 5D), suggesting that at any given time a fraction of T-C6G is colocalized with TM9SF4 in the Golgi complex, where the two proteins are in close proximity and might interact.

Overall our results suggest that TM9SF4 interacts with glycine-rich TMDs and allows them to be transported through the Golgi complex and to the cell surface (Fig. 6).

DISCUSSION

Numerous studies have established that TMDs can be essential elements in controlling intracellular transport of individual membrane proteins (reviewed in Cosson et al., 2013). Length and hydrophilicity are key features defining the influence of a given TMD on intracellular sorting, but its exact amino acid sequence can also be a determining factor (Sharpe et al., 2010). Molecular mechanisms ensuring differential sorting of TMDs have been best studied in the early steps of the secretory pathway, where short or hydrophilic TMDs are retained in the ER. Specific ER localization of short or hydrophilic TMDs might result from an inefficient exit out of the ER (Ronchi et al., 2008), coupled to a continuous retrieval of escaped proteins back to the ER after their recognition in the cis-Golgi by the Rer1 receptor (Sato et al., 2001). By contrast, proteins with long hydrophobic TMDs are recognized by Erv14 in the ER and efficiently packaged in COPII-coated secretory vesicles (Herzig et al., 2012). Recognition of TMDs by specific cargo receptors, like Erv14, or retrieval receptors, like Rer1, might control their intracellular transport at each transport step in eukaryotic cells. It is worth noting that, to date, virtually all our knowledge of the molecular mechanisms ensuring sorting of TMDs is based on results obtained in yeast cells, and potential cargo or retrieval receptors have not been characterized to the same extent in mammalian cells.

Our results suggest that Phg1A and TM9SF4 act as a cargo receptor enabling specifically transport of glycine-rich TMDs to the cell surface, both in *Dictyostelium* and in human cells. Another interpretation, not mutually exclusive with the first one, is that Phg1A and TM9SF4 might act as a chaperone facilitating folding of glycine-rich TMDs and facilitating their transport along the secretory pathway. These two interpretations would account for the observation that surface localization of proteins with glycine-rich TMDs is reduced by the genetic inactivation of Phg1A or TM9SF4, and increased by TM9SF4 overexpression in mammalian cells. Both Phg1A and TM9SF4 show a specific propensity to associate with glycine-rich TMDs. More specifically, in order to assist transport of glycine-rich TMDs, Phg1A and TM9SF4 would need to associate with them in the early secretory pathway and release them in later compartments (Golgi complex or cell surface). In agreement with this proposal, TM9SF4 is localized mostly in the Golgi complex, and a proximity-ligation assay suggests that it might interact with glycine-rich TMDs in this compartment. Our results do not exclude the possibility that Phg1A or TM9SF4, or other TM9 proteins, might also control transport of proteins at other steps of intracellular transport. Indeed, there have been scattered reports that TM9 proteins are present both in the Golgi complex and in endocytic compartments in various species (discussed in Pruvot et al., 2010), but the functional significance of these observations remains to be firmly established.

Glycine-rich TMDs are not exceptional in eukaryotic cells. Like hydrophilic residues, glycine residues allow specific interactions between TMDs, thus driving the formation of homo- or hetero-oligomeric complexes and playing a key role in the function of many membrane receptors (Fink et al., 2012). For example, GxxxG motifs, where two glycine residues are placed on the same face of the TMD helix participate in the dimerization of the α and β subunits of some integrin molecules (Kim et al., 2009). Higher numbers of glycine residues on the same face of the TMD helix also drive the homo-oligomerization of glycophorin A, or the hetero-oligomerization of class II major histocompatibility complex proteins (Cosson and Bonifacino, 1992; Lemmon and Engelman, 1994). In the human amyloid precursor protein, at least three glycine residues in the TMD have been implicated in formation of homodimers, thus controlling the intracellular fate of the protein and the generation of amyloid peptides (Kienlen-Campard et al., 2008; Munter et al., 2007). Our results suggest that TM9 proteins might participate in the intracellular transport of these proteins. In SibA itself, glycine residues are mostly placed on the same face of the TMD helix (Cornillon et al., 2006), a disposition favoring interactions with other glycine-rich TMDs. Several TMDs of TM9 proteins exhibit GxxG or GxxxG motifs (notably TMDs 4 and 9 in Phg1A and TM9SF4) and could thus directly associate with glycine-rich TMDs to assist their transport. In addition to SibA, the intracellular sorting of two proteins so far has been shown to depend on TM9 proteins: the *Dictyostelium* Kil1 protein and the *Drosophila* PGRP-LC receptor. Both Kil1 and PGRP-LC are type II

Journal of Cell Science (2015) 128, 2269-2277 doi:10.1242/jcs.164848

transmembrane proteins, and their TMDs do not exhibit multiple glycine residues. Thus, it is possible that the specificity of TM9 proteins is not limited to glycine-rich TMDs.

The finding that TM9 proteins participate in the intracellular transport of a subset of proteins with glycine-rich TMDs might account for the variety of phenotypes observed in cells with altered levels of TM9 proteins. A complete understanding of the proteins placed under the control of TM9 proteins might be essential to fully understand their role in intracellular transport. In *D. discoideum*, where the function of TM9 proteins has been most extensively studied, *phg1A* KO cells have been shown to be defective in cellular adhesion (due to depletion of surface SibA) (Froquet et al., 2012), in intracellular bacterial killing (due to depletion of Kil1, a Golgi sulfotransferase) (Benghezal et al., 2006; Le Coadic et al., 2013), and in intracellular targeting of lysosomal enzymes (Froquet et al., 2008). Similarly in *S. cerevisiae*, *D. melanogaster* and human cells, a variety of functions (adhesion, phagocytosis, intracellular transport, autophagy, nutrient sensing etc.) are known to be altered by disruption or overexpression of TM9 proteins. It is also worth noting that the genetic alterations described in this study did not generate all-or-none effects: typically, genetic inactivation of Phg1A or TM9SF4 resulted in a three-fold decrease of the surface targeting of glycine-rich TMDs. Previous studies have shown that at least three proteins facilitate surface targeting of SibA in *Dictyostelium*: Phg1A, Phg1B and SadA. Functional redundancy might thus attenuate the effect of a single gene inactivation. In the future, it will be interesting to determine whether other TM9 proteins interact specifically with different subsets of TMDs, and if they act at the same steps of intracellular transport.

MATERIALS AND METHODS

Cell lines and media

Dictyostelium discoideum DH1-10 cells (Cornillon et al., 2000) were grown in HL5 medium at 23°C and are referred to as wild-type (WT). *Phg1A* KO cells were as described previously (Benghezal et al., 2003). These two strains were modified by introducing a plasmid encoding the Rep protein, allowing replication of plasmids bearing the Ddp2 origin of replication (Shammat and Welker, 1999). Cells were diluted twice a week to maintain cultures at a maximal density of 1.5×10^6 cells/ml.

Surface targeting of csA chimeric proteins in *Dictyostelium*

The plasmid allowing the expression of a csA–SibA chimera was as described previously (Froquet et al., 2012). In this study, all csA chimeric proteins were obtained similarly by subcloning PCR fragments in the csA-0 plasmid digested with *Kpn*I and *Xba*I. The sequence of the TMD of each construct is shown in Table 1. Plasmids were transfected in WT or *phg1A* KO cells as described previously (Alibaud et al., 2003). Cells expressing csA chimeric proteins were selected and maintained in HL5 supplemented with G418 (12.5 µg/ml).

To analyze cell surface levels of csA fusion proteins in *Dictyostelium*, 0.5×10^6 cells were allowed to attach on a glass coverslip for 10 min at room temperature in phosphate buffer (2 mM Na_2HPO_4, 14.7 mM KH_2PO_4 pH 6.0 supplemented with 100 mM sorbitol, 100 µM $CaCl_2$ and 0.5% HL5) (Smith et al., 2010). Cells were washed in phosphate buffer, labeled for 2 min with a monoclonal antibody (41-71-21) against native csA (Bertholdt et al., 1985), washed in phosphate buffer and fixed for 10 min in phosphate buffer containing 4% paraformaldehyde. After washing in PBS containing 20 mM NH_4Cl, and in PBS containing 0.2% BSA (PBS-BSA), cells were incubated for 20 min with an Alexa-Fluor-647-coupled anti-mouse-IgG antibody in PBS-BSA. After a washing step in PBS-BSA, cells were permeabilized with Triton X-100 (0.07% in PBS for 2 min), and washed in PBS-BSA. This procedure is optimal for staining of cell surface proteins (Vernay and Cosson, 2013). The whole cellular csA content was labeled with the 41-71-21 antibody in PBS-BSA for 20 min, then cells were washed twice in PBS-BSA and incubated for 20 min with an Alexa-Fluor-488-coupled anti-mouse-IgG antibody in PBS-BSA. Finally, cells were washed twice in PBS-BSA, once in PBS and mounted in Mowiol for visualization in an LSM700 confocal microscope (Zeiss), or scrapped from the coverslip and resuspended in PBS for flow cytometry analysis.

To label cell surface proteins with biotin, 30×10^6 cells were harvested and washed in 10 ml SB-Sorbitol (phosphate buffer Na_2HPO_4-KH_2PO_4 17 mM pH 6.0 containing 120 mM sorbitol). The pellet was resuspended in 2 ml of SB-Sorbitol pH 8.0 containing 1 mg of NHS-SS-biotin (Pierce) and incubated on ice for 10 min. Cells were harvested and resuspended in 10 ml PBS supplemented with 100 mM Glycine for 5 min on ice. Biotinylated cells were washed four times in 10 ml SB-Sorbitol pH 6.0 and an aliquot of 500 µl was collected to analyze the total amount of csA. Cells were then lysed for 15 min in 1 ml RIPA buffer [50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS and proteases inhibitors (20 µg/ml leupeptin, 10 µg/ml aprotinin, 18 µg/ml PMSF and 18 µg/ml iodoacetic acid)] and the lysate cleared by centrifugation (5 min, 9300 ***g***, 4°C). After centrifugation, 900 µl of the supernatant was incubated with neutravidin beads overnight on a rotating wheel at 4°C. Beads were washed twice with 1 ml of RIPA buffer, incubated for 15 min at 4°C in 1 ml RIPA buffer, incubated in 1 ml of urea 6 M for 15 min at 4°C, and washed three times with 1 ml RIPA buffer. Biotinylated surface proteins were eluted in 50 µl sample buffer for 15 min at room temperature followed by 5 min at 60°C. 20 µl of sample were loaded for each dilution on a 9% SDS-PAGE gel, transferred onto nitrocellulose, and revealed with a mouse anti-csA monoclonal antibody 33-294-17 (Bertholdt et al., 1985). Serial dilutions of the precipitated material were analyzed to quantitatively assess the amounts of csA proteins.

Tac fusion proteins and association assays

We used a pCDM8-based vector containing the coding sequence of the α chain of the interleukin-2 receptor (Tac) and a *Bgl*II site in the membrane-proximal area (Cosson et al., 1991). The indicated constructs were obtained by inserting the sequence coding for the TMD of interest (see Table 1) in this vector digested with *Bgl*II and *Xba*I. Plasmids were propagated in bacteria MC1061/P3 after chemical transformation and selection on LB agar plates containing ampicillin 12.5 µg/ml and tetracyclin 7.5 µg/ml.

COS7 cells were co-transfected with Tac constructs and a human-codon-optimized version of Phg1A fused to β-galactosidase. Cells were washed in PBS, lysed for 15 min at 4°C in lysis buffer [PBS containing 0.5% Triton X-100 and protease inhibitors (leupeptin 20 µg/ml, aprotinin 10 µg/ml, PMSF 18 µg/ml and iodoacetic acid 18 µg/ml)], then centrifuged for 15 min at 4°C (10,000 ***g***). The supernatant was collected and is referred to as total lysate; an aliquot of 10 µl was kept for analysis of the total amount of β-galactosidase activity. Total lysate was incubated for 1 h at 4°C with protein-A–agarose beads previously coated with an anti-Tac mouse antibody (7G7) (Rubin et al., 1985), then the beads were washed five times with PBS with 0.1% Triton X-100. The β-galactosidase activity was assessed in the total lysate and in the immunoprecipitated sample. We used the substrate Chlorophenol Red-β-D-galactopyranoside and quantified the product of the reaction by absorbance at 600 nm. To assess association of Tac proteins with TM9SF4, a similar procedure was followed, but using HeLa cells.

Analysis of human cells

HEK293T cells expressing Tac proteins were washed in ice-cold DMEM and the Tac protein at the cell surface was labeled for 15 min with 7G7 antibody (Rubin et al., 1985) in DMEM at 4°C. Cells were then washed in DMEM at 4°C, fixed for 10 min in PBS containing 4% paraformaldehyde, washed in PBS containing 20 mM NH_4Cl and incubated for 30 min with an Alexa-Fluor-647-coupled anti-mouse-IgG antibody (Life Technologies, A21235) in PBS-BSA at room temperature. Cells were washed three times with PBS-BSA, permeabilized for 10 min in PBS containing 0.2% saponin, washed with PBS-BSA and the intracellular Tac was labeled with 7G7 antibody in PBS-BSA for 30 min. When indicated, samples were also incubated with a rabbit anti-Flag antibody (Sigma, F7425) or a mouse anti-Giantin antibody (Nizak et al., 2003). Cells were washed three times in PBS-BSA and incubated for 30 min with an Alexa- Fluor-488-coupled anti-mouse-IgG antibody (Life Technologies, A11029), an Alexa-Fluor-488-

coupled anti-rabbit-IgG antibody (Life Technologies, A11034), or an Alexa-Fluor-647-coupled anti-mouse-IgG antibody (Life Technologies, A21235). Cells were washed again three times with PBS-BSA, once with PBS and mounted in Mowiol. When indicated, the endoplasmic reticulum was identified by expressing YFP-KDEL (ER-YFP, a kind gift of Nicolas Demaurex, University of Geneva, Switzerland).

Surface labeling was quantified with ImageJ software (http://rsb.info.nih.gov/ij/). For each cell, three lines were drawn perpendicular to the cell surface and the height of the peak of fluorescence was averaged, which represented the surface labeling (in arbitrary units). Three lines of defined length were drawn inside the cell, the area under the curve was calculated and averaged, which represented the intracellular labeling (in arbitrary units). The surface:intracellular ratio was then calculated for each individual cell. In each independent experiment, at least 20 cells were quantified.

We used the previously described CRISPR/Cas9 method (Mali et al., 2013) to generate *TM9SF4*-knockout HEK293T cells. In the plasmid purchased from Sigma-Aldrich, the 20-nt guide RNA was designed towards the signal sequence, in the second exon of *TM9SF4* (CCCTGA*TGTGTGAAACA-AGCGC, where * is the cutting site of the Cas9 nuclease). 2 µg of plasmid were transfected into HEK293T cells using Lipofectamine 2000 (Life Technologies, 11668-027). After 2 days, GFP-positive cells were sorted by flow cytometry, cloned and allowed to grow, before genomic DNA extraction (QIAamp DNA Blood Mini Kit, Qiagen, 51104). 54 individual clones were screened by PCR using the following primers (primer 1, 5′-CGGTTT-TGGAGAAACTTGTAGG-3′; primer 2, 5′-CTTGTTTCACACATCAG-GGAG-3′; primer 3, 5′-CTCCCTGATGTGTGAAACAAG-3′; primer 4, 5′-CCAAGGAAAAGAGACGTTCAC-3′ – primers 2 and 3 anneal at the expected site of mutagenesis and should not anneal anymore if mutations occur, and primer 1 and 4 anneal 500 bp on either sides of the Cas9 nuclease cutting site). Pairs of primers are used as follows: 1+2, 3+4 and 1+4. At least one of the PCR amplification was defective in 13 clones, and the genomic region was amplified with primers 1 and 4 and sequenced. Only clones with mutations inducing a frame shift in both alleles were kept.

To perform a proximity ligation assay, HEK293T WT cells were transfected with a Flag-tagged version of TM9SF4 and T-C1G or T-C6G. Cells were fixed, permeabilized, and incubated with the corresponding primary antibodies as described above. The proximity ligation assay was performed according to the manufacturer's protocol (Sigma, DUO92101).

Acknowledgements
We thank Thierry Soldati and Oliver Hartley for discussions and comments on the manuscript.

Competing interests
The authors declare no competing or financial interests.

Author contributions
J.P., M.L.C., A.V., M.D., H.O.-S. and P.C. performed experiments. N.P. set up the conditions for cell surface biotinylation of proteins. P.C., J.P. and A.V. wrote the manuscript. All authors read and corrected the manuscript.

Funding
This work was supported by the Swiss National Foundation for Scientific Research [grant number 31003A-153326]; the Doerenkamp-Zbinden Foundation and the Fondation Egon Naef pour la Recherche in Vitro. Deposited in PMC for immediate release.

Supplementary material
Supplementary material available online at
http://jcs.biologists.org/lookup/suppl/doi:10.1242/jcs.164848/-/DC1

References
Alibaud, L., Cosson, P. and Benghezal, M. (2003). Dictyostelium discoideum transformation by oscillating electric field electroporation. *Biotechniques* **35**, 78-80, 82-73.

Benghezal, M., Cornillon, S., Gebbie, L., Alibaud, L., Brückert, F., Letourneur, F. and Cosson, P. (2003). Synergistic control of cellular adhesion by transmembrane 9 proteins. *Mol. Biol. Cell* **14**, 2890-2899.

Benghezal, M., Fauvarque, M.-O., Tournebize, R., Froquet, R., Marchetti, A., Bergeret, E., Lardy, B., Klein, G., Sansonetti, P., Charette, S. J. et al. (2006). Specific host genes required for the killing of Klebsiella bacteria by phagocytes. *Cell. Microbiol.* **8**, 139-148.

Bergeret, E., Perrin, J., Williams, M., Grunwald, D., Engel, E., Thevenon, D., Taillebourg, E., Bruckert, F., Cosson, P. and Fauvarque, M.-O. (2008). TM9SF4 is required for Drosophila cellular immunity via cell adhesion and phagocytosis. *J. Cell Sci.* **121**, 3325-3334.

Bertholdt, G., Stadler, J., Bozzaro, S., Fichtner, B. and Gerisch, G. (1985). Carbohydrate and other epitopes of contact site A glycoprotein of Dictyostelium discoideum as characterized by monoclonal antibodies. *Cell Differ.* **16**, 187-202.

Chluba-de Tapia, J., de Tapia, M., Jäggin, V. and Eberle, A. N. (1997). Cloning of a human multispanning membrane protein cDNA: evidence for a new protein family. *Gene* **197**, 195-204.

Cornillon, S., Pech, E., Benghezal, M., Ravanel, K., Gaynor, E., Letourneur, F., Bruckert, F. and Cosson, P. (2000). Phg1p is a nine-transmembrane protein superfamily member involved in dictyostelium adhesion and phagocytosis. *J. Biol. Chem.* **275**, 34287-34292.

Cornillon, S., Gebbie, L., Benghezal, M., Nair, P., Keller, S., Wehrle-Haller, B., Charette, S. J., Bruckert, F., Letourneur, F. and Cosson, P. (2006). An adhesion molecule in free-living Dictyostelium amoebae with integrin beta features. *EMBO Rep.* **7**, 617-621.

Cosson, P. and Bonifacino, J. S. (1992). Role of transmembrane domain interactions in the assembly of class II MHC molecules. *Science* **258**, 659-662.

Cosson, P., Lankford, S. P., Bonifacino, J. S. and Klausner, R. D. (1991). Membrane protein association by potential intramembrane charge pairs. *Nature* **351**, 414-416.

Cosson, P., Perrin, J. and Bonifacino, J. S. (2013). Anchors aweigh: protein localization and transport mediated by transmembrane domains. *Trends Cell Biol.* **23**, 511-517.

Fey, P., Stephens, S., Titus, M. A. and Chisholm, R. L. (2002). SadA, a novel adhesion receptor in Dictyostelium. *J. Cell Biol.* **159**, 1109-1119.

Fink, A., Sal-Man, N., Gerber, D. and Shai, Y. (2012). Transmembrane domains interactions within the membrane milieu: principles, advances and challenges. *Biochim. Biophys. Acta.* **1818**, 974-983.

Froquet, R., Cherix, N., Birke, R., Benghezal, M., Cameroni, E., Letourneur, F., Mosch, H.-U., De Virgilio, C. and Cosson, P. (2008). Control of cellular physiology by TM9 proteins in yeast and Dictyostelium. *J. Biol. Chem.* **283**, 6764-6772.

Froquet, R., le Coadic, M., Perrin, J., Cherix, N., Cornillon, S. and Cosson, P. (2012). TM9/Phg1 and SadA proteins control surface expression and stability of SibA adhesion molecules in Dictyostelium. *Mol. Biol. Cell* **23**, 679-686.

Geva, Y. and Schuldiner, M. (2014). The back and forth of cargo exit from the endoplasmic reticulum. *Curr. Biol.* **24**, R130-R136.

Herzig, Y., Sharpe, H. J., Elbaz, Y., Munro, S. and Schuldiner, M. (2012). A systematic approach to pair secretory cargo receptors with their cargo suggests a mechanism for cargo selection by Erv14. *PLoS Biol.* **10**, e1001329.

Hohmann, H. P., Gerisch, G., Lee, R. W. and Huttner, W. B. (1985). Cell-free sulfation of the contact site A glycoprotein of Dictyostelium discoideum and of a partially glycosylated precursor. *J. Biol. Chem.* **260**, 13869-13878.

Kienlen-Campard, P., Tasiaux, B., Van Hees, J., Li, M., Huysseune, S., Sato, T., Fei, J. Z., Aimoto, S., Courtoy, P. J., Smith, S. O. et al. (2008). Amyloidogenic processing but not amyloid precursor protein (APP) intracellular C-terminal domain production requires a precisely oriented APP dimer assembled by transmembrane GXXXG motifs. *J. Biol. Chem.* **283**, 7733-7744.

Kim, C., Lau, T.-L., Ulmer, T. S. and Ginsberg, M. H. (2009). Interactions of platelet integrin alphaIIb and beta3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood* **113**, 4747-4753.

Le Coadic, M., Froquet, R., Lima, W. C., Dias, M., Marchetti, A. and Cosson, P. (2013). Phg1/TM9 proteins control intracellular killing of bacteria by determining cellular levels of the Kil1 sulfotransferase in Dictyostelium. *PLoS ONE* **8**, e53259.

Lemmon, M. A. and Engelman, D. M. (1994). Specificity and promiscuity in membrane helix interactions. *FEBS Lett.* **346**, 17-20.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826.

Mercanti, V., Marchetti, A., Lelong, E., Perez, F., Orci, L. and Cosson, P. (2010). Transmembrane domains control exclusion of membrane proteins from clathrin-coated pits. *J. Cell Sci.* **123**, 3329-3335.

Munter, L.-M., Voigt, P., Harmeier, A., Kaden, D., Gottschalk, K. E., Weise, C., Pipkorn, R., Schaefer, M., Langosch, D. and Multhaup, G. (2007). GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Abeta42. *EMBO J.* **26**, 1702-1712.

Nizak, C., Martin-Lluesma, S., Moutel, S., Roux, A., Kreis, T. E., Goud, B. and Perez, F. (2003). Recombinant antibodies against subcellular fractions used to track endogenous Golgi protein dynamics in vivo. *Traffic* **4**, 739-753.

Ochiai, H., Stadler, J., Westphal, M., Wagle, G., Merkl, R. and Gerisch, G. (1982). Monoclonal antibodies against contact sites A of Dictyostelium discoideum: detection of modifications of the glycoprotein in tunicamycin-treated cells. *EMBO J.* **1**, 1011-1016.

Journal of Cell Science

Perrin, J., Mortier, M., Jacomin, A. C., Viargues, P., Thevenon, D. and Fauvarque, M. O. (2015). The nonaspanins TM9SF2 and TM9SF4 regulate the plasma membrane localization and signalling activity of the peptidoglycan recognition protein PGRP-LC in drosophila. *J. Innate Immun*. **7**, 37-46.

Pruvot, B., Laurens, V., Salvadori, F., Solary, E., Pichon, L. and Chluba, J. (2010). Comparative analysis of nonaspanin protein sequences and expression studies in zebrafish. *Immunogenetics* **62**, 681-699.

Ronchi, P., Colombo, S., Francolini, M. and Borgese, N. (2008). Transmembrane domain-dependent partitioning of membrane proteins within the endoplasmic reticulum. *J. Cell Biol.* **181**, 105-118.

Rubin, L. A., Kurman, C. C., Biddison, W. E., Goldman, N. D. and Nelson, D. L. (1985). A monoclonal antibody 7G7/B6, binds to an epitope on the human interleukin-2 (IL-2) receptor that is distinct from that recognized by IL-2 or anti-Tac. *Hybridoma* **4**, 91-102.

Sato, K., Sato, M. and Nakano, A. (2001). Rer1p, a retrieval receptor for endoplasmic reticulum membrane proteins, is dynamically localized to the Golgi apparatus by coatomer. *J. Cell Biol.* **152**, 935-944.

Schimmöller, F., Díaz, E., Mühlbauer, B. and Pfeffer, S. R. (1998). Characterization of a 76kDa endosomal, multispanning membrane protein that is highly conserved throughout evolution. *Gene* **216**, 311-318.

Shammat, I. M. and Welker, D. L. (1999). Mechanism of action of the Rep protein from the Dictyostelium Ddp2 plasmid family. *Plasmid* **41**, 248-259.

Sharpe, H. J., Stevens, T. J. and Munro, S. (2010). A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **142**, 158-169.

Smith, E. W., Lima, W. C., Charette, S. J. and Cosson, P. (2010). Effect of starvation on the endocytic pathway in Dictyostelium cells. *Eukaryot. Cell* **9**, 387-392.

Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.-J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.-G. et al. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* **3**, 995-1000.

Vernay, A. and Cosson, P. (2013). Immunofluorescence labeling of cell surface antigens in Dictyostelium. *BMC Res. Notes* **6**, 317.

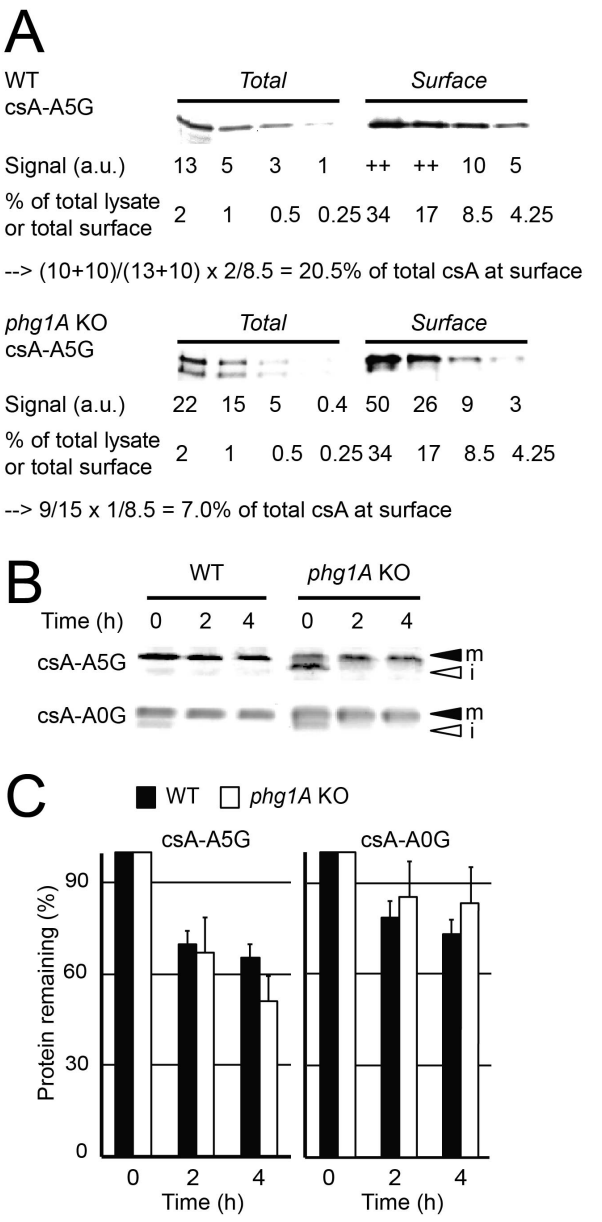

Fig. S1 (Perrin et al.)

Figure S1: Phg1A-dependent targeting of glycine-rich TMDs to the cell surface. (A) Quantification of surface csA fusion protein following surface biotinylation. To determine the percentage of csA protein present at the cell surface, serial dilutions of total or surface proteins were migrated on a SDS-PAGE gel and revealed by Western blot with an antibody recognizing the csA extracellular domain. Only the dilutions

that were in the linear range were compared. The percentages of the total cell lysate (2 to 0.25 %) and of the surface proteins (34 to 4 %) loaded on each lane are indicated. The signal corresponding to each band was quantified using the ImageJ software (http://rsb.info.nih.gov/ij/). To determine the percentage of csA-A5G present at the cell surface, we used the signal corresponding to 4.25 % and 8.5 % of the surface protein, and to 2 % and 1 % of the total protein.

% at the cell surface = $(2\text{x}Surface_{4.25\%} + Surface_{8.5\%}) / (Total_{2\%} + 2\text{x}Total_{1\%})\text{x}2/8.5$

$= (2\text{x}5+10)/(2\text{x}5+13)\text{x}2/8.5=20.5$ %

(B) csA-A5G and csA-A0G exhibit similar stability in WT and *phg1A* KO cells. Stability of csA-A5G and csA-A0G was assessed in WT and *phg1A* KO *Dictyostelium* cells. To determine the turnover of csA, 5 x 10^6 cells were incubated in HL5 containing 2 mM cycloheximide. Aliquots of 1.5 x 10^6 cells were collected after 0, 2 and 4 hours and resuspended in Sample Buffer (0.103 g/ml sucrose, 50 mM Tris, pH 6.8, 5 mM EDTA, 0.5 mg/ml bromophenol blue, 2 % SDS, 10 % β-mercaptoethanol). Samples were migrated on a 9 % acrylamide gel, and transferred to nitrocellulose using a semi-dry transfer system (Invitrogen, Carlsbad, CA). The membrane was incubated overnight in PBS containing 0.1 % Tween 20 and 7 % milk, then incubated successively with a mouse anti-csA monoclonal antibody (33-294-17) (Bertholdt et al., 1985), and a horseradish peroxidase-coupled goat anti-mouse IgG (Bio-Rad, 1706516). The signal was revealed by ECL. Mature csA has a molecular weight of 80 kDa (arrow m), while the partially glycosylated form has an molecular weight of 68 kDa (arrow i)

(C) The quantification with time of the total amount of csA (mature and immature bands) for csA-A5G and csA-A0G indicates that these two proteins exhibit the same stability in WT and *phg1A* KO cells.

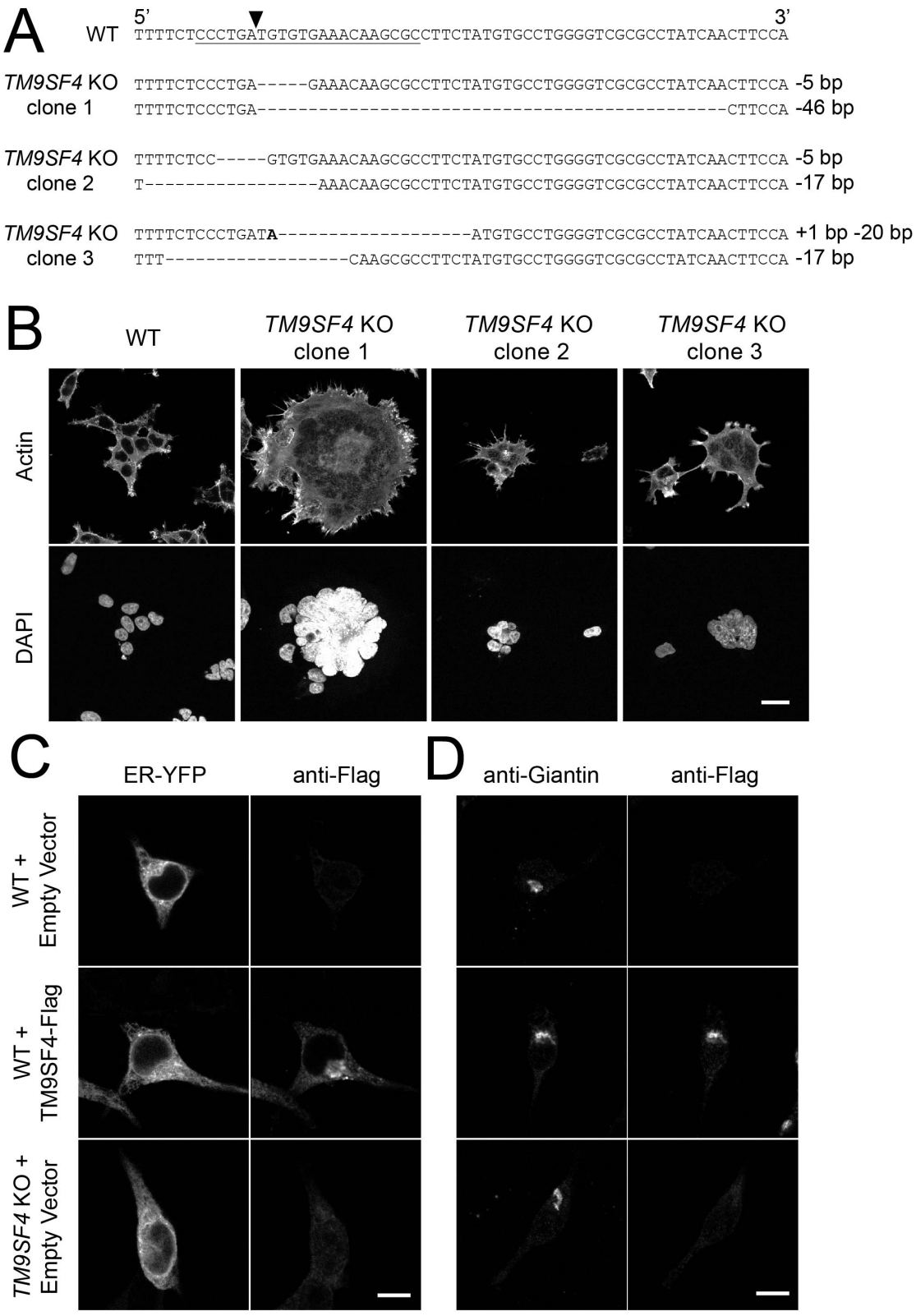

Fig. S2 (Perrin et al.)

Figure S2: Human *TM9SF4* KO cells.

(A) Sequence of *TM9SF4* in human. The sequence targeted by the guide RNA is underlined and the arrow represents the cutting site of the Cas9 nuclease. The sequence of each allele is represented for three independent mutant clones. Deletions are symbolized by (-) and insertions are in bold.

(B) The presence of plurinucleated cells was assessed by actin and nuclei staining. To stain actin and nuclei, HEK293T (WT) or *TM9SF4* KO cells were fixed in PBS containing 4 % paraformaldehyde for 30 minutes, and washed with 20 mM NH_4Cl in PBS. Cells were then permeabilized with PBS containing 0.2 % saponin for 10 minutes and washed in PBS-BSA. Actin was stained with TRITC-labeled phalloidin (Sigma, P-1951) in PBS for 1 hour. Cells were then washed three times with PBS, and nuclei were stained with DAPI (Life Technologies, D1306) in PBS for 5 minutes. Cells were finally washed three times with PBS and mounted in Mowiol. Scale bar: 30 µm.

(C-D) The endoplasmic reticulum and the Golgi apparatus are not affected by the loss or the overexpression of TM9SF4. The endoplasmic reticulum or the Golgi apparatus were revealed using (C) a soluble YFP-KDEL, or (D) an antibody against giantin, in HEK293T cells (WT) (upper panel), in cells overexpressing a Flag-tagged version of TM9SF4 (TM9SF4-Flag, intermediate panel), or in *TM9SF4 KO* cells (lower panel). Scale bar: 10 µm.

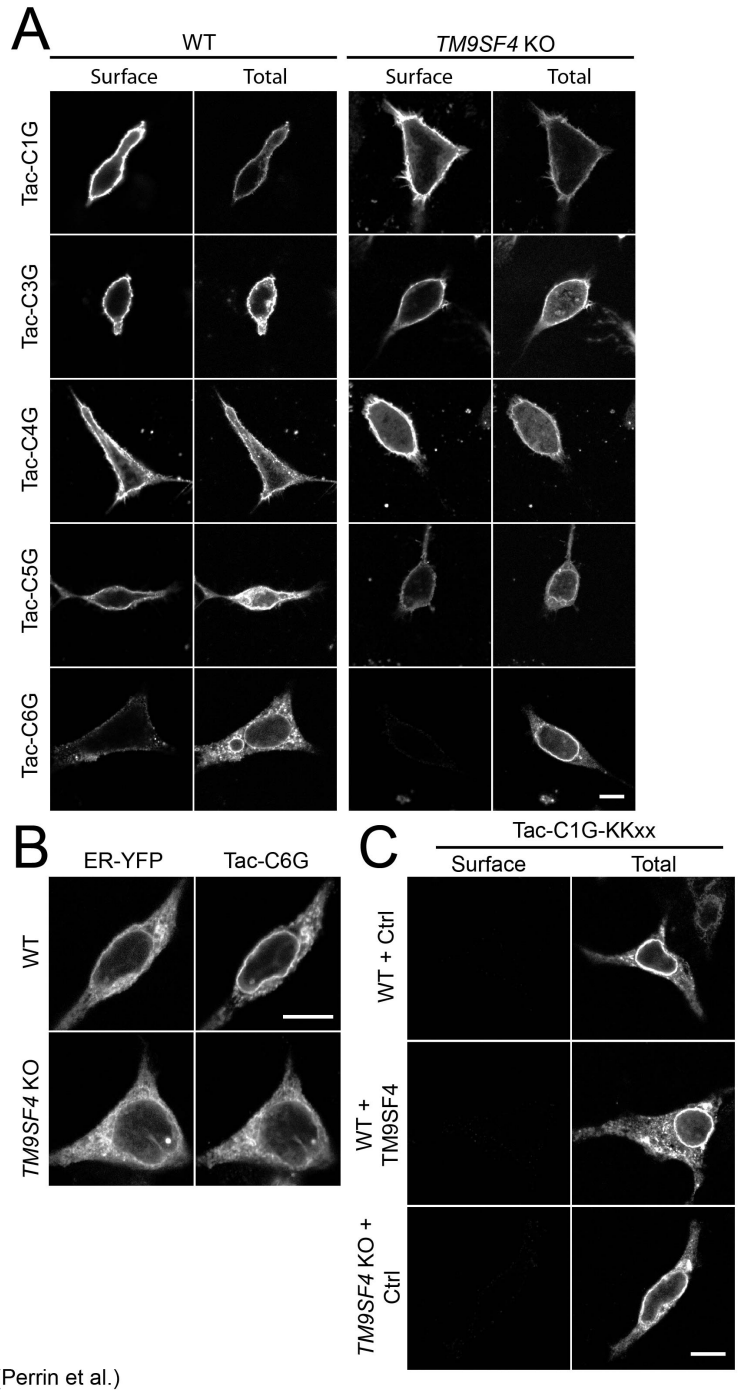

Fig. S3 (Perrin et al.)

Figure S3: TM9SF4 ensures specifically localization of glycine-rich TMDs to the cell surface.

(A) Multiple glycine residues in the TMD of Tac proteins ensure its localization in the endoplasmic reticulum in human cells. Tac chimeric proteins with 1, 3, 4, 5 or 6 glycine residues in their TMD were expressed in HEK293T cells (WT) or in *TM9SF4*

KO cells. The Tac fusion proteins were labeled before (Surface) or after (Total) cell permeabilization.

(B) The majority of Tac-C6G colocalizes with a co-expressed ER-localized soluble YFP-KDEL in WT (upper panel) or *TM9SF4* KO cells (lower panel). Scale bar: 10 µm.

(C) Loss or overexpression of TM9SF4 does not affect dilysine-mediated ER retention. Tac-C1G-KKxx was expressed in WT cells (upper panel), WT cells overexpressing TM9SF4 (intermediate panel) or in *TM9SF4* KO cells (lower panel). The Tac fusion protein was labeled before (Surface) or after (Total) cell permeabilization. Scale bar: 10 µm.

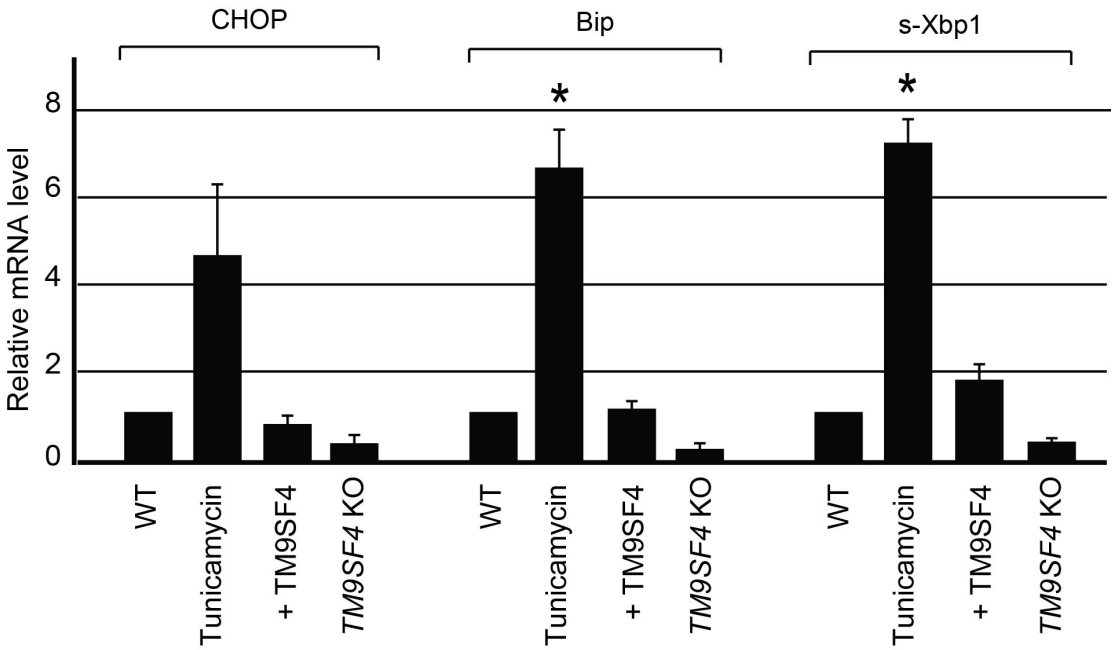

Fig. S4 (Perrin et al.)

Figure S4: ER stress response genes are not induced in human cells when TM9SF4 is lost or overexpressed.

Total RNAs from HEK293T human cells were extracted (RNeasy Mini kit, Qiagen). RT-PCR was realized with 1µg of total RNAs (Qscript cDNA synthesis, Quanta Biosciences). Analysis of mRNA levels of CHOP, Bip and spliced-Xbp1 (s-Xbp1) was realized by real-time quantitative PCR (StepOne System, Life Technologies). The primers used are described in (Oslowski and Urano, 2011). Cells treated with Tunicamycin (5 µg/ml) during 5 hours were used as positive control. The average and SEM of at least three independent experiments are indicated. * indicates a significant difference (Student *t* test $p < 0.01$).

C. Article in preparation: TM9SF4 ensures specific recognition and Golgi retention of transmembrane domains

1. Introduction

In this study entitled TM9 family proteins control surface targeting of glycine-rich transmembrane domains, we show that TM9SF4 controls the surface localization of glycine-rich TMDs. The work presented below follows this study. Here, we show that in addition to glycine-rich TMDs, TM9SF4 also associate with potentially charged TMDs, suggesting that it could interact with a large set of TMDs. We also show that TM9SF4 localizes TMDs in the Golgi complex, suggesting that TM9SF4 could be a mechanism that localizes TMDs in this organelle. Such mechanisms have never been described yet.

Most of the results presented below are preliminary and need to be confirmed. Additional experiments are also necessary to fully understand the role of TM9SF4. In this study, several Tac constructs are the same as the ones used in the second publication of this manuscript. Here, T-H0 corresponds to T-C1G, T-G4 corresponds to T-C6G and T-H0-KKxx corresponds to T-C1G-KKxx.

2. Article in preparation:

TM9SF4 ensures specific recognition and Golgi retention of transmembrane domains

Alexandre Vernay, Otmane Lamrabet, Tania Jauslin, Jackie Perrin, Pierre Cosson*

Faculty of Medicine, University of Geneva, Centre Médical Universitaire, 1 rue Michel Servet, CH1211 Geneva 4, Switzerland

*Corresponding author: Pierre.Cosson@unige.ch

I contributed to all the figures, as well as the writing of this study.

ABSTRACT

Previous studies have shown that TM9SF4 interacts with glycine-rich transmembrane domains (TMDs), escorts them out of the endoplasmic reticulum (ER) and ensures their surface localization. In this study, we tested whether TM9SF4 has a wider role in the sorting of TMDs in the secretory pathway. Our results indicate that TM9SF4 interacts with a variety of TMDs, in particular TMDs containing charged residues. A TMD carrying a positively charged residue (T-R8) was localized in the ER in WT HEK cells, but was partially relocalized to the Golgi complex upon overexpression of TM9SF4. Conversely, the HCT116 cell line has been shown to contain a high level of TM9SF4, and in these cells, T-R8 is found both in the ER and in the Golgi complex. Upon genetic inactivation of TM9SF4 in HCT116 cells, T-R8 is relocalized to the ER. These results suggest that the intracellular level of TM9SF4 determines the localization of a subset of proteins at the ER-Golgi interface. Remarkably, in HCT116 cells, overexpression of a Golgi resident enzyme displaces T-R8 from the Golgi to the ER, suggesting that TM9SF4 may also be involved in the Golgi retention of Golgi resident enzymes.

The existence of a cellular mechanism recognizing a few specific TMDs and ensuring their retention in the Golgi complex has long been hypothesized. Our results suggest that TM9SF4 may be an essential element of the machinery ensuring recognition and retention of a subset of TMDs in the Golgi complex.

INTRODUCTION

Membrane proteins destined to the cell surface are inserted co-translationally in the membrane of the endoplasmic reticulum (ER). They are then transported by vesicular intermediates to the Golgi complex, then to the cell surface. Although the Golgi complex is traversed by a continuous flow of proteins destined to the cell surface, it maintains its specific composition. This observation demonstrates the existence of specific sorting mechanisms that recognize Golgi-targeted proteins and ensure their localization in the Golgi membrane.

Many enzymes ensuring the posttranslational modification of glycoproteins in the Golgi complex are integral membrane proteins. For several of them it has been demonstrated that their localization in the Golgi complex is ensured by information contained in their cytosolic domain and their transmembrane domain (TMD) (Banfield, 2011). A short peptide motif in the cytosolic domain of several Golgi proteins is bound by the Vps74 protein, and participates in the Golgi localization of some proteins (Schmitz et al., 2008), although the exact mechanism by which Vps74 acts is still debated (Cai et al., 2014; Eckert et al., 2014). For virtually all Golgi resident proteins, the TMD has been shown to contain information essential for Golgi localization. Indeed the composition of the TMDs of Golgi-targeted enzymes differs from that of surface proteins: it is shorter (Bretscher and Munro, 1993) and exhibits a specific composition (Sharpe et al., 2010). Altered Golgi proteins with mutations in their TMD fail to be efficiently localized in the Golgi complex (Munro, 1995). To date, the cellular mechanism ensuring recognition and targeting of TMDs in the Golgi complex remains unelucidated. It is not even clear if Golgi targeting results from the specific retention of a subset of proteins in stable Golgi cisternae, or

from the continuous retrograde transfer of Golgi enzymes from one maturing cisterna to another less mature cisterna.

Previous studies have demonstrated that Phg1A is essential for efficient expression of SibA at the surface of *Dictyostelium* cells (Froquet et al., 2012). Phg1A was later shown to associate with the glycine-rich TMD of SibA and to facilitate its transport from the ER to the cell surface in *Dictyostelium* cells (Perrin et al., 2015). In human HEK cells, a glycine-rich TMD is mostly retained in the ER and surface expression is low. Genetic inactivation of TM9SF4, the human ortholog of Phg1A, lowers even more the level of surface expression, while TM9SF4 overexpression increases it significantly (Perrin et al., 2015). Overall these results suggest that TM9SF4, which is mainly localized in the Golgi complex, plays a key role in the sorting and transport of a subset of TMDs.

In this study our aim was to determine if TM9SF4 also recognized other ER-retained proteins and influenced their intracellular localization. Our results suggest that TM9SF4 recognizes a large set of TMDs, and ensures their relocalization from the ER to the Golgi complex. TM9SF4 may represent the first identified element in the machinery ensuring targeting of specific TMDs to the Golgi complex.

RESULTS

Overexpression of TM9SF4 in HEK cells relocates ER-targeted TMDs to the Golgi complex

In order to test if the levels of TM9SF4 influence the intracellular sorting of various TMDs, we expressed in HEK cells a fusion protein exhibiting a TMD with a potentially charged arginine residue (T-R8), which has been shown previously to act as an efficient ER-retention motif (Bonifacino et al., 1991) (Fig. 2C, upper panel). We then tested by immunofluorescence the intracellular localization of this protein.

For this, we overexpressed TM9SF4 in cells expressing T-R8, T-G4, or T-KKxx (Fig. 2). The Tac constructs used in this study are detailed in Fig. 1 and the aminoacid sequences are presented in Table 1. In all situations, a large fraction of the Tac protein was found in the endoplasmic reticulum, identified in particular by the fact that is delimitates the nucleus. We also determined the presence of these various proteins in the Golgi complex, identified by the presence of the giantin protein and found that in agreement with our previous observations (Perrin et al., 2015) a certain amount of T-G4 was detected in the Golgi complex (Fig. 2A). This result was quantified by comparing the relative intensity of the signal detected in the nuclear envelope, an easily identified domain of the endoplasmic reticulum, and the Golgi complex (Fig. 2B). Interestingly, overexpression of TM9SF4 increased significantly the percentage of T-G4 found in the Golgi complex (Fig. 2A and B). The effect of TM9SF4 on the Golgi localization of T-R8 was even more striking, since T-R8 was virtually absent from the Golgi in HEK cells, and

relocalized strongly to the Golgi complex in cells overexpressing TM9SF4 (Fig. 2C and D). On the contrary, a protein localized in the ER by virtue of a cytosolic KKxx ER retrieval motif was restricted to the ER, even in cells overexpressing TM9SF4 (Fig. 2E and F).

TM9SF4 interacts with charged TMDs

Our previous studies showed that TM9SF4 was capable of intramembrane interactions with glycine-rich TMDs. To test whether TM9SF4 was able to specifically interact with a broader range of TMDs, we co-expressed TM9SF4 fused to β-galactosidase (TM9SF4-βGal) and variants of Tac proteins exhibiting different TMDs. The Tac protein was then immunoprecipitated and the amount of β-galactosidase co-precipitated was determined. As previously demonstrated, a higher β-galactosidase activity was co-precipitated with T-G3, compared to T-H0 (Fig. 3A). A similar level of interaction was detected between TM9SF4 and T-R8, suggesting a specific interaction of TM9SF4 with T-R8. No interaction was detected between TM9SF4 and a Tac protein containing a negative residue in its TMD (T-D10) (Fig. 3A). Tac chimeric proteins exhibiting a cluster of hydrophilic threonine residues (T-H3) or a strongly hydrophilic glutamine (T-Q8) in their TMD also interacted efficiently with TM9SF4 (Fig. S2). Since the Tac proteins used in this study are mainly localized in the ER and TM9SF4 localizes in the Golgi complex, we reasoned that these constructs may not be present in the same sub-compartments. To ensure the same ER localization of Tac constructs and TM9SF4-βGal, we exchanged the extracellular domain of TM9SF4 with the extracellular domain of the δ-chain of the T-cell receptor, known to confer ER retention (Klausner et al., 1990). A Flag-tagged version of this construct (δ-TM9SF4-Flag) localized both in the ER and in the Golgi complex (data not

shown). We used this construct fused to the β-galactosidase (δ-TM9SF4-βGal) to detect interactions with Tac constructs in the ER. δ-TM9SF4-βGal showed intramembrane interactions with T-G3, T-R8 and T-D10 and not with T-H0 (Fig. 3B). Taken together, these results demonstrate that TM9SF4 interacts with a large set of TMDs.

Golgi localization of ER-targeted TMDs in HCT116 cells is due to TM9SF4 expression

HCT116 cells are aggressive melanoma cells, characterized by a very high expression of TM9SF4 (Lozupone et al., 2015). In order to test the role of TM9SF4 in determining intracellular localization of Tac proteins, we expressed various chimeric proteins in these cells. T-G4 was found both in the ER and in the Golgi complex (Fig. 4A, arrowheads and quantification in B). We then generated a specific TM9SF4 KO cell line using the CRISPR/Cas9 method (Fig. S1A). In these cells, the general aspect of the ER and the Golgi complex is the same as in WT cells (Fig. S1B). We observed that in TM9SF4 KO cells the localization of T-G4 in the Golgi complex was not impacted by the loss of TM9SF4 (Fig. 4A and B). However, similar to HEK cells overexpressing TM9SF4, in HCT116 cells T-R8 was largely localized to the Golgi complex. Interestingly, genetic inactivation of TM9SF4 resulted in the relocalization of T-R8 in the ER (Fig. 4C, arrowheads and quantification in D). On the contrary, the ER localization of T-H0-KKxx, a construct localized to the ER by virtue of its dilysine motif, was not impacted by the loss of TM9SF4. Together, these data confirm, in a different cellular setting that high levels of TM9SF4 allow localization of T-R8 in the Golgi complex.

TM9SF4 also localizes a TMD with a negative charge in the Golgi complex

In order to extend these observations to a wider range of TMDs, we also tested the effect of TM9SF4 on the localization of a TMD containing a negatively-charged aminoacid residue (T-D10). The organization of this Tac chimeric protein is depicted in Fig. 5A. T-D10 in mainly localized in the ER in HEK cells (Fig. 5B, upper panel). However, this constructs largely relocalized in the Golgi complex upon overexpression of TM9SF4 (Fig. 5B, lower panel, quantification in Fig. 5C). These results indicated that TM9SF4 also controls the access of T-D10 to the Golgi complex.

An endogenous Golgi protein can compete T-R8 out of the Golgi complex

In the course of our experiments, we tried to express GFP-tagged Golgi proteins as a means to detect the Golgi complex, without using anti-giantin antibodies. For this we used the beta-1,4-galactosyltransferase 1, an endogenous Golgi resident enzyme, fused to the GFP (B4GALT1-GFP). To our surprise, in HCT116 WT cells expressing both T-R8 and B4GALT1, the Golgi localization of T-R8 was abolished (Fig. 6A upper panel, arrowheads and quantification in B). These results suggest that overexpression of B4GALT1-GFP competed T-R8 out of the Golgi. Since our previous observations suggest that the Golgi localization of T-R8 is achieved by an interaction with TM9SF4, this result suggested that B4GALT1-GFP competed with T-R8 for interaction with TM9SF4.

To confirm that the Golgi complex conserves its integrity upon overexpression of B4GALT1, HCT116 WT cells were co-transfected with T-R8 and a RFP-tagged B4GALT1, and the Golgi complex was revealed with an anti-giantin antibody. As shown in Fig. 6C,

B4GALT1-RFP colocalized with the giantin marker (arrowheads). Moreover, T-R8 was not detected in the Golgi complex and appeared restricted to the ER in these cells. (Fig. 6C and D). Taken together, these results show that in HCT116 WT cells, T-R8 is excluded from the Golgi complex upon overexpression of B4GALT1.

DISCUSSION

The results presented in this study are highly reproducible. Indeed, almost every one of them was obtained in at least three independent experiments. They regard them however as preliminary in the sense that the interpretations that we are proposing have not all been verified by additional experiments.

Mechanisms ensuring Golgi localization of proteins are largely unknown. To the best of our knowledge and as mentioned above, Vps74 is the only protein described that interacts with a cytosolic motif of Golgi resident proteins and allows their correct localization. But many proteins do not rely on Vps74 to be localized in the Golgi complex. On the contrary, mechanisms ensuring Golgi localization of TMDs are largely unknown. Our observations strongly suggest that TM9SF4 could be one of the mechanisms ensuring recognition and Golgi localization of a specific set of TMDs. This interpretation is based on the observation that TM9SF4 interacts both with glycine-rich and potentially charged TMDs, and allows their Golgi localization. At this stage, two interpretations are possible. According to the first explanation, TM9SF4 would interact with TMDs in the ER and ensure their packaging into COPII-coated vesicles, and as a consequence facilitate their transfer from the ER to the Golgi. However, this model does not account for the observation that overexpression of a Golgi enzyme prevents TM9SF4 from localizing T-R8 in the Golgi complex. Thus we favour the second hypothesis: TM9SF4 captures TMDs in the Golgi complex and ensures their retention in this organelle.

One prediction of this interpretation is that TM9SF4 must interact very efficiently with B4GALT1 and ensure its localization in the Golgi complex. This prediction should be tested directly.

Another prediction of this model is that, the localization of Golgi-resident enzymes should be affected by the loss or overexpression of TM9SF4. Our first qualitative observations did not detect a change in the Golgi localization of B4GALT1 in cells depleted of TM9SF4. However this point should be carefully tested by expressing various Golgi-resident enzymes in cells expressing or not TM9SF4. It may also be necessary to express proteins targeted to the Golgi complex solely by their TMD, to avoid other Golgi-localization mechanisms (e.g. conferred by Vps74-binding cytosolic sequences) to participate in Golgi localization.

The interpretations proposed above are based on the assumption that Golgi cisternae are stable entities, and that Golgi localization results mostly from the specific retention of Golgi enzymes. Our observations are more difficult to reconcile with the possibility that Golgi proteins are dynamically localized in this organelle through various rounds of anterograde/retrograde transport. This remains however a formal possibility. If the conclusion can be reached that TM9SF4 plays a key role in determining localization of some TMDs in the Golgi complex, this may allow to design new experiments to test whether Golgi localization is obtained mainly by a dynamic transport, or by specific retention mechanisms.

MATERIALS AND METHODS

Cell culture and reagents

Human embryonic kidney (HEK) 293T cells were cultures in Dulbecco's modified Eagle's medium, in which 10% Fetal Bovine Serum and Penicillin-Streptomycin was added. HCT116 cells were grown in Roswell Park Memorial Institute medium supplemented with 10% Fetal Bovine Serum and Penicillin-Streptomycin.

HCT116 *tm9sf4* knockout cells were generated using the CRISPR/Cas9 method, as described before (Perrin et al., 2015). Two individual clones were obtained with mutations leading to a frameshift in both alleles, and were used in parallel in this study.

Plasmids were obtained as previously described (Perrin et al., 2015).

Localization of Tac proteins

HEK 293T and HCT116 cells were transfected 2 days before the experiment, using polyethylenimine (PEI) as previously described (Longo et al., 2013). Cells were first fixed for 30 min in 4% paraformaldehyde, then permeabilized for 10 min with 0,2% Saponin. Samples were then incubated with the mouse 7G7 antibody (Rubin et al., 1985) for 30 min. When indicated, cells were also incubated with a human anti-giantin antibody (Nizak et al., 2003) or a rabbit anti-Flag antibody (Sigma, F7425). Finally, cells were incubated with an Alexa-Fluor-647-coupled anti-mouse-IgG antibody (Life Technologies, A11029) and when indicated, with an Alexa-Fluor-488-coupled anti-human-IgG antibody (Jackson ImmunoResearch, 709-545-149) or an Alexa-Fluor-555-

coupled anti-rabbit-IgG antibody (Molecular Probes, A-11035), before being mounted in Möwiol. Samples were analysed using a LSM700 confocal microscope (Zeiss).

For surface localization of Tac proteins, samples were first incubated with the 7G7 antibody for 15 min at 4°C, before fixation. Surface Tac protein was revealed with an Alexa-Fluor-647-coupled anti-mouse-IgG antibody before permeabilization. If so, the intracellular content is revealed with an Alexa-Fluor-488-coupled anti-mouse IgG.

Association assays

Experiments were performed as previously described (Perrin et al., 2015). Briefly, HeLa cells were co-transfected with Tac chimeric proteins and TM9SF4 fused to the β-galactosidase. Cells were washed in PBS before lysis in PBS containing 0,5% Triton X-100 and a cocktail of protease inhibitors (leupeptin 20µg/mL, aprotinin 10µg/mL, PMSF 18µg/mL and iodoacetic acid 18µg/mL). Samples were centrifuged for 15 min at 4°C (10'000g) and the supernatants were kept. 10µL of lysate was kept to determine the total amount of β-galactosidase activity in each sample. Tac protein from the lysate was then immunoprecipitated thanks to protein A-agarose beads previously coated with the anti-Tac 7G7 antibody. The β-galactosidase activity is revealed upon addition of its substrate, Chlorophenol Red-β-D-galactopyranoside and quantified by absorbance at 600nm. The percentage of β-galactosidase immunoprecipitated was then determined.

REFERENCES

Banfield, D.K., 2011. Mechanisms of protein retention in the Golgi. Cold Spring Harb. Perspect. Biol. 3, a005264. doi:10.1101/cshperspect.a005264

Bonifacino, J.S., Cosson, P., Shah, N., Klausner, R.D., 1991. Role of potentially charged transmembrane residues in targeting proteins for retention and degradation within the endoplasmic reticulum. EMBO J. 10, 2783–2793.

Bretscher, M.S., Munro, S., 1993. Cholesterol and the Golgi apparatus. Science 261, 1280–1281.

Cai, Y., Deng, Y., Horenkamp, F., Reinisch, K.M., Burd, C.G., 2014. Sac1-Vps74 structure reveals a mechanism to terminate phosphoinositide signaling in the Golgi apparatus. J. Cell Biol. 206, 485–491. doi:10.1083/jcb.201404041

Eckert, E.S.P., Reckmann, I., Hellwig, A., Röhling, S., El-Battari, A., Wieland, F.T., Popoff, V., 2014. Golgi phosphoprotein 3 triggers signal-mediated incorporation of glycosyltransferases into coatomer-coated (COPI) vesicles. J. Biol. Chem. 289, 31319–31329. doi:10.1074/jbc.M114.608182

Froquet, R., le Coadic, M., Perrin, J., Cherix, N., Cornillon, S., Cosson, P., 2012. TM9/Phg1 and SadA proteins control surface expression and stability of SibA adhesion molecules in Dictyostelium. Mol. Biol. Cell 23, 679–686. doi:10.1091/mbc.E11-04-0338

Klausner, R.D., Lippincott-Schwartz, J., Bonifacino, J.S., 1990. The T cell antigen receptor: insights into organelle biology. Annu. Rev. Cell Biol. 6, 403–431. doi:10.1146/annurev.cb.06.110190.002155

Longo, P.A., Kavran, J.M., Kim, M.-S., Leahy, D.J., 2013. Transient mammalian cell transfection with polyethylenimine (PEI). Methods Enzymol. 529, 227–240. doi:10.1016/B978-0-12-418687-3.00018-5

Lozupone, F., Borghi, M., Marzoli, F., Azzarito, T., Matarrese, P., Iessi, E., Venturi, G., Meschini, S., Canitano, A., Bona, R., Cara, A., Fais, S., 2015. TM9SF4 is a novel V-ATPase-interacting protein that modulates tumor pH alterations associated with drug resistance and invasiveness of colon cancer cells. Oncogene 34, 5163–5174. doi:10.1038/onc.2014.437

Munro, S., 1995. An investigation of the role of transmembrane domains in Golgi protein retention. EMBO J. 14, 4695–4704.

Nizak, C., Martin-Lluesma, S., Moutel, S., Roux, A., Kreis, T.E., Goud, B., Perez, F., 2003. Recombinant antibodies against subcellular fractions used to track endogenous Golgi protein dynamics in vivo. Traffic Cph. Den. 4, 739–753.

Perrin, J., Le Coadic, M., Vernay, A., Dias, M., Gopaldass, N., Ouertatani-Sakouhi, H., Cosson, P., 2015. TM9 family proteins control surface targeting of glycine-rich transmembrane domains. J. Cell Sci. 128, 2269–2277. doi:10.1242/jcs.164848

Rubin, L.A., Kurman, C.C., Biddison, W.E., Goldman, N.D., Nelson, D.L., 1985. A monoclonal antibody 7G7/B6, binds to an epitope on the human interleukin-2 (IL-2) receptor that is distinct from that recognized by IL-2 or anti-Tac. Hybridoma 4, 91–102. doi:10.1089/hyb.1985.4.91

Schmitz, K.R., Liu, J., Li, S., Setty, T.G., Wood, C.S., Burd, C.G., Ferguson, K.M., 2008. Golgi localization of glycosyltransferases requires a Vps74p oligomer. Dev. Cell 14, 523–534. doi:10.1016/j.devcel.2008.02.016

Sharpe, H.J., Stevens, T.J., Munro, S., 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. Cell 142, 158–169. doi:10.1016/j.cell.2010.05.037

FIGURE LEGENDS

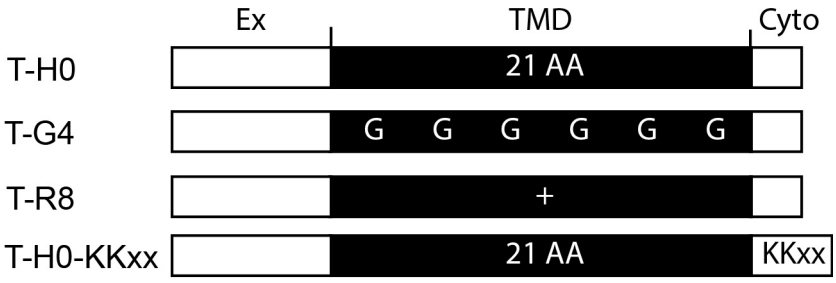

Figure 1: Organization of the Tac constructs used in this study. T-H0 is composed by the extracellular domain of the Tac protein (Ex), a 21-hydrophobic residues transmembrane domain (TMD) and a short cytosolic domain (Cyto). T-G4 contains the same TMD than T-H0 in which 6 residues have been mutated into glycine residues. T-R8 exhibits the same TMD than T-H0 with an arginine residue in position 8. Finally, T-H0-KKx has the same TMD than T-H0 but contains a dilysine motif (KKxx) in its cytosolic tail. The aminoacid sequences of these constructs are available in Table 1.

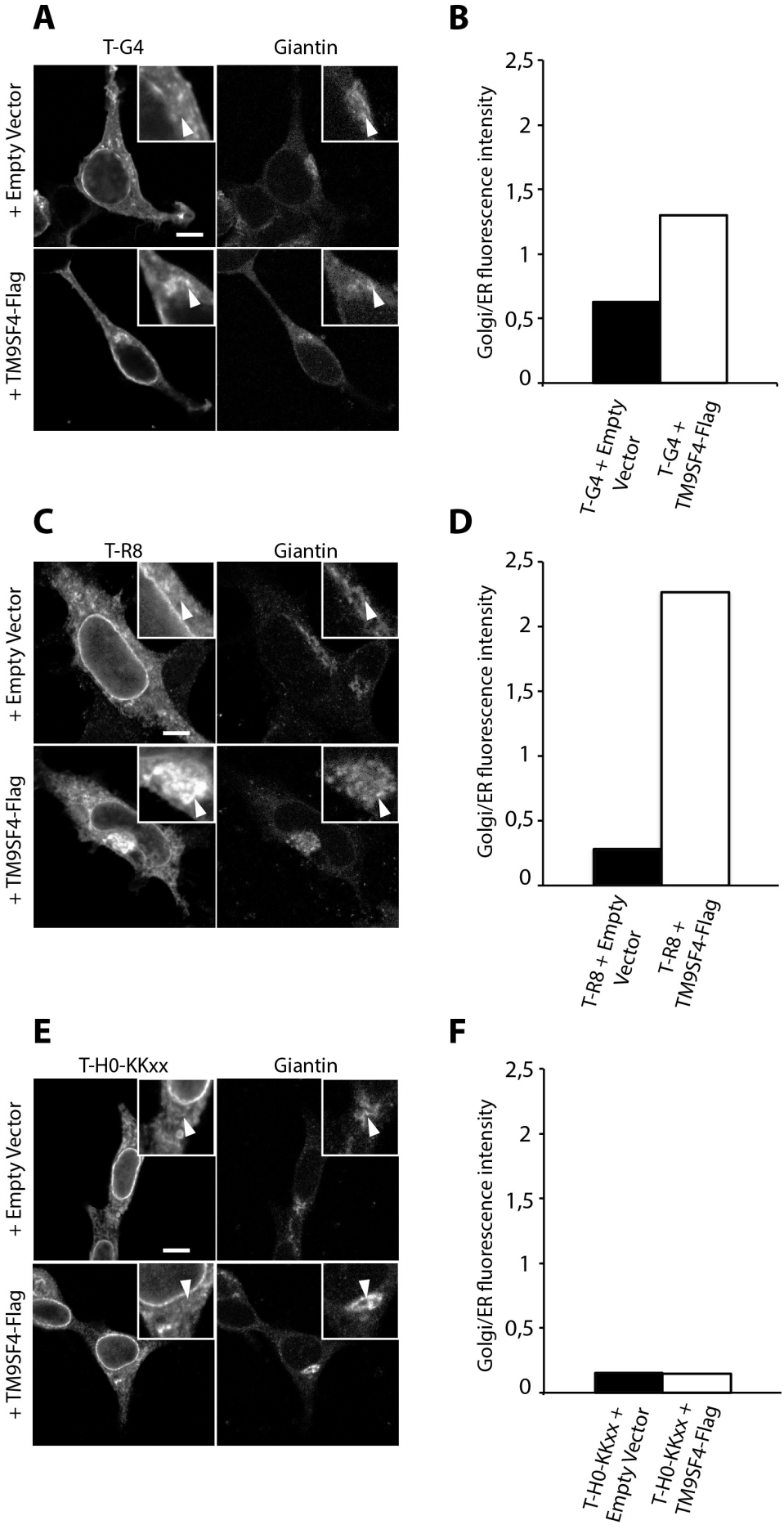
A
T-G4
Giantin
+ Empty Vector
+ TM9SF4-Flag
B
Golgi/ER fluorescence intensity
2,5
2
1,5
1
0,5
0
T-G4 + Empty Vector
T-G4 + TM9SF4-Flag
C
T-R8
Giantin
+ Empty Vector
+ TM9SF4-Flag
D
Golgi/ER fluorescence intensity
2,5
2
1,5
1
0,5
0
T-R8 + Empty Vector
T-R8 + TM9SF4-Flag
E
T-H0-KKxx
Giantin
+ Empty Vector
+ TM9SF4-Flag
F
Golgi/ER fluorescence intensity
2,5
2
1,5
1
0,5
0
T-H0-KKxx + Empty Vector
T-H0-KKxx + TM9SF4-Flag

Figure 2: Localization of Tac proteins in HEK cells. (A) T-G4 was transfected in HEK 293T cells in which an empty vector (upper panel) or TM9SF4-Flag (lower panel) was co-expressed. The Golgi complex was localized with an anti-giantin antibody. A magnification of the Golgi complex area is represented in the upper right corner. Stronger degree of colocalization of T-G4 with the Golgi complex was observed when TM9SF4 is overexpressed (arrowheads). (B) Quantification of the relative signal intensity of T-G4 in the Golgi complex compared to the relative intensity in the nuclear envelope. (C) The same experiment was performed with T-R8. A stronger colocalization of T-R8 with the Golgi complex was observed when TM9SF4-Flag is overexpressed (arrowheads). (D) Quantification of the levels of T-R8 found in the Golgi complex compared to the ER. (E) The same experiment was conducted with T-H0-KKxx. No colocalization was observed with the Golgi complex (arrowheads), as presented in the quantification in (F). Scale bars: 10μm. All the experiments presented here have been performed at least three times.

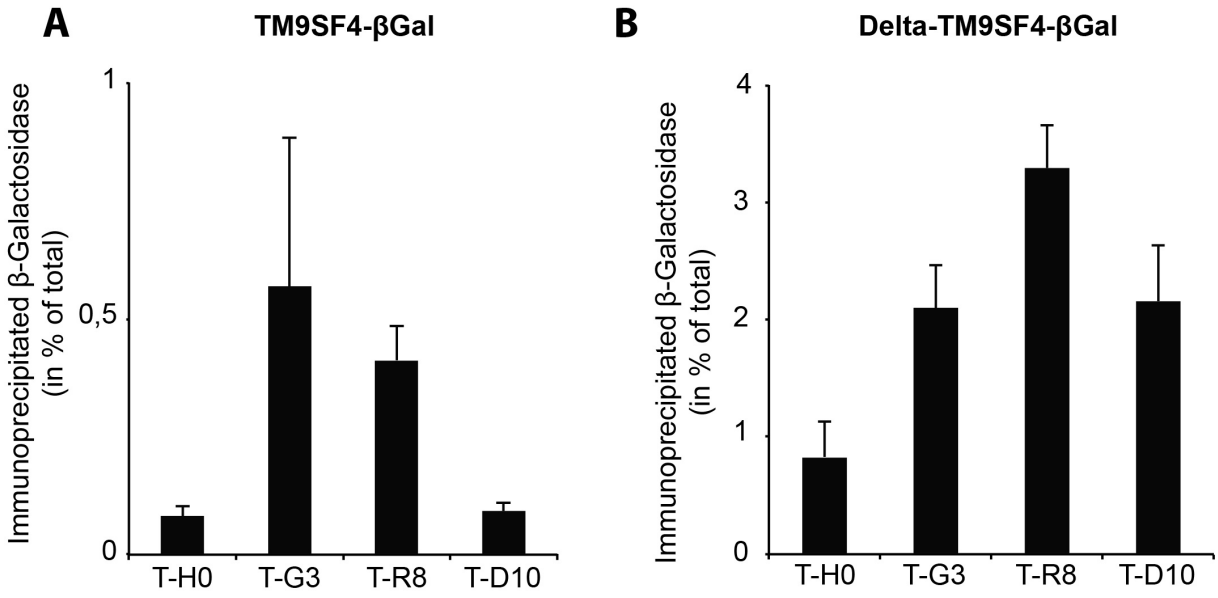

Figure 3: TM9SF4 makes intra membrane interactions with a large set of TMDs. HeLa cells were co-transfected with TM9SF4 fused to the β-galactosidase (TM9SF4-βGal) and Tac chimeric proteins. The Tac proteins were immunoprecipitated and the amount of co-precipitated β-galactosidase activity was assessed and reveals the level of interaction with TM9SF4. TM9SF4-βGal shows a higher degree of association with T-G4 and T-R8, compared to T-H0. The mean ± SEM of five independent experiments are indicated.

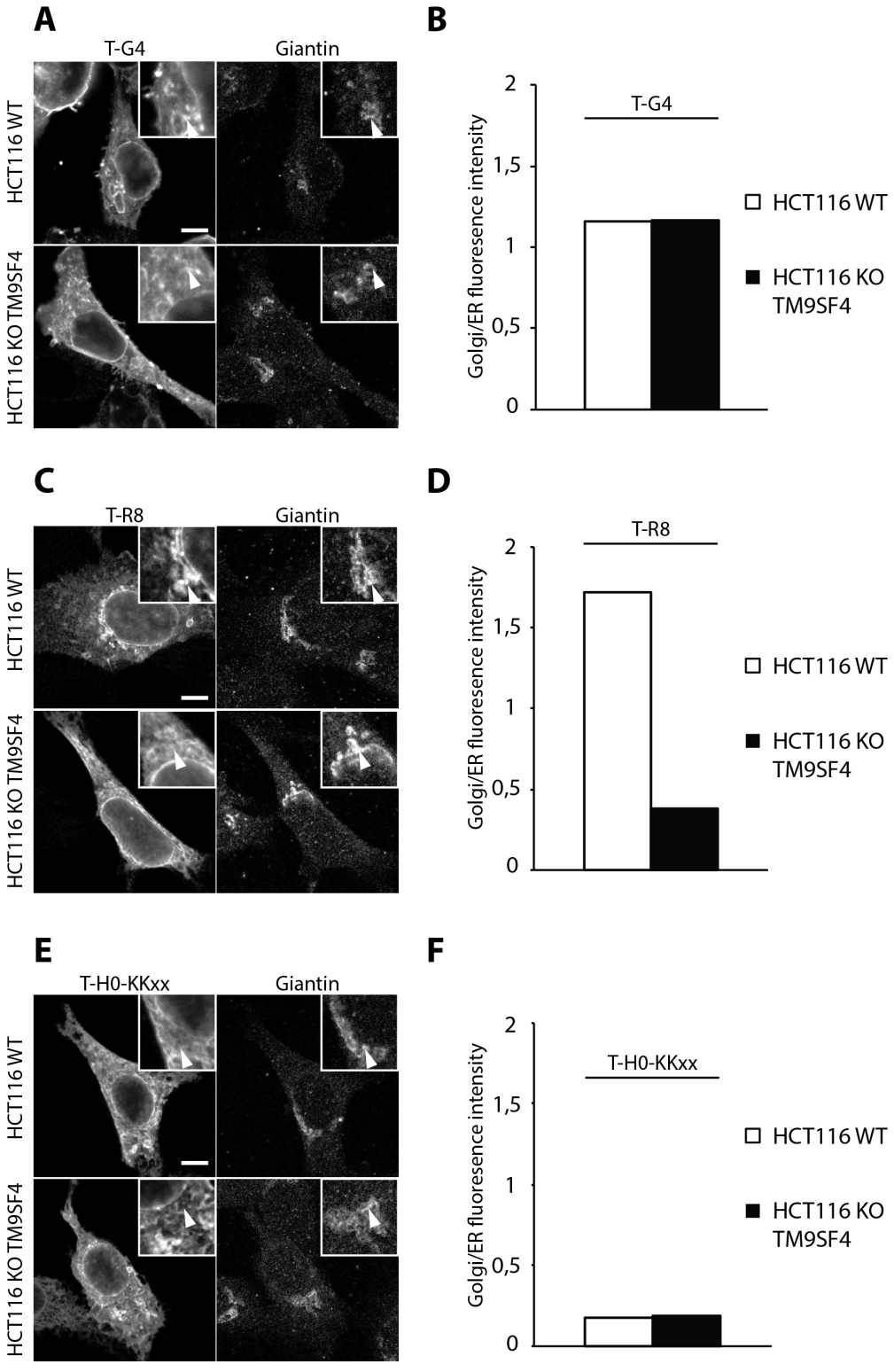

Figure 4: Localization of Tac proteins in parental HCT116 cells and deficient for TM9SF4. (A) T-G4 was transfected in parental HCT116 cells (WT, upper panel) or in cells deficient for TM9SF4 (HCT116 KO TM9SF4, lower panel). The Golgi complex is revealed thanks to

an anti-giantin antibody. A magnification of the Golgi complex area is represented in the upper right corner. T-G4 is present in the Golgi complex and the ER of both WT and cells deficient for TM9SF4 (arrowheads). A quantification of this result is represented in (B), where the relative intensity of T-G4 in the Golgi complex is compared to the relative intensity found in the ER. The same experiment was performed with T-R8 (C) and T-H0-KKxx (E), with the corresponding quantifications in (D) and (E), respectively. Scale bars: 10μm. The data presented here are the result of at least three independent experiments.

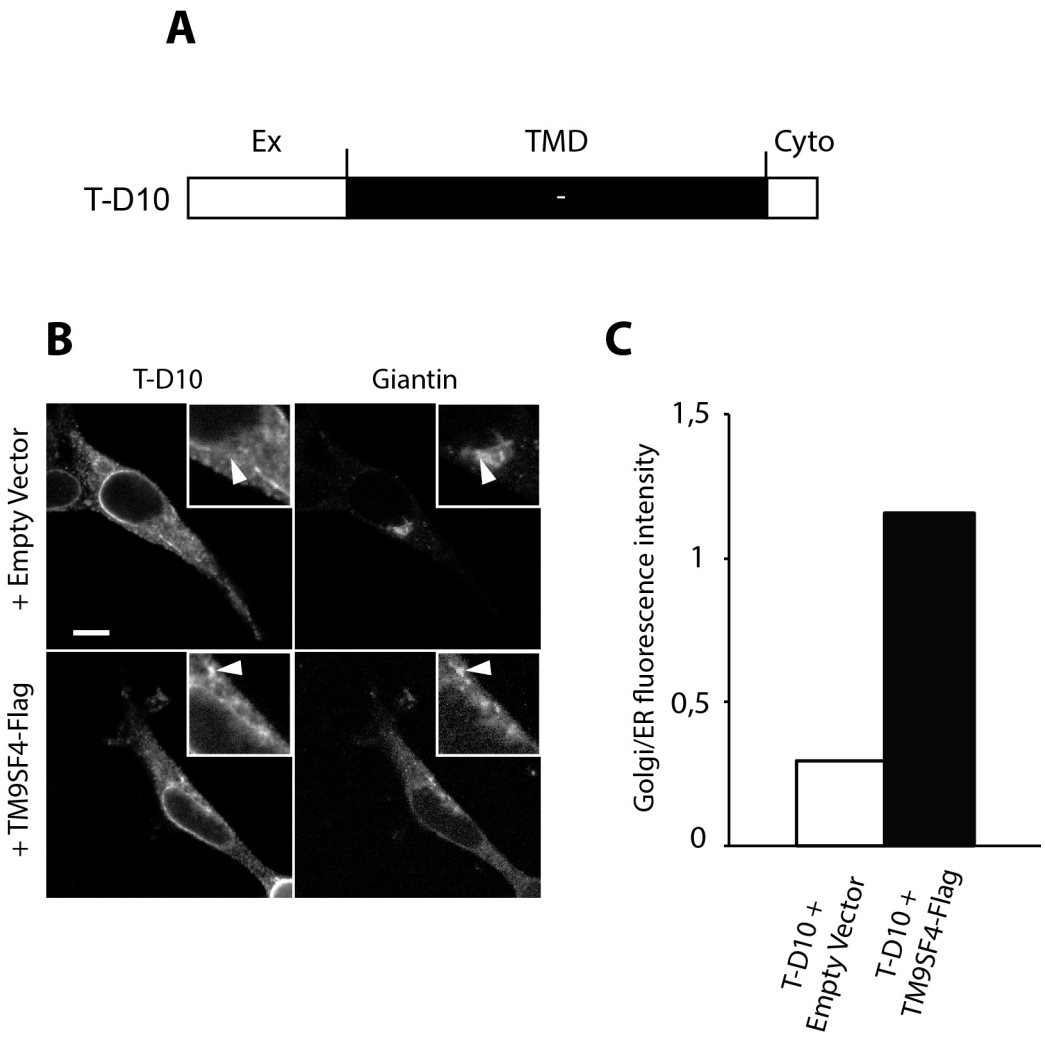

Figure 5: TM9SF4-dependent localization of negatively charged TMDs. (A) Organization of T-D10, composed by the extracellular domain of the Tac protein, the same TMD than T-H0 with a leucine mutated into aspartic acid, and a short cytosolic tail. (B) T-D10 was co-transfected with an Empty Vector (upper panel) in HEK cells. The Golgi complex is localized thanks to an antibody against the giantin. T-D10 mainly localizes in the ER. When TM9SF4-Flag is overexpressed (lower panel), T-D10 partially colocalizes with the giantin marker (arrowheads). A quantification of this result is presented in (D). Scale bar: 10µm. The experiment was performed 3 times.

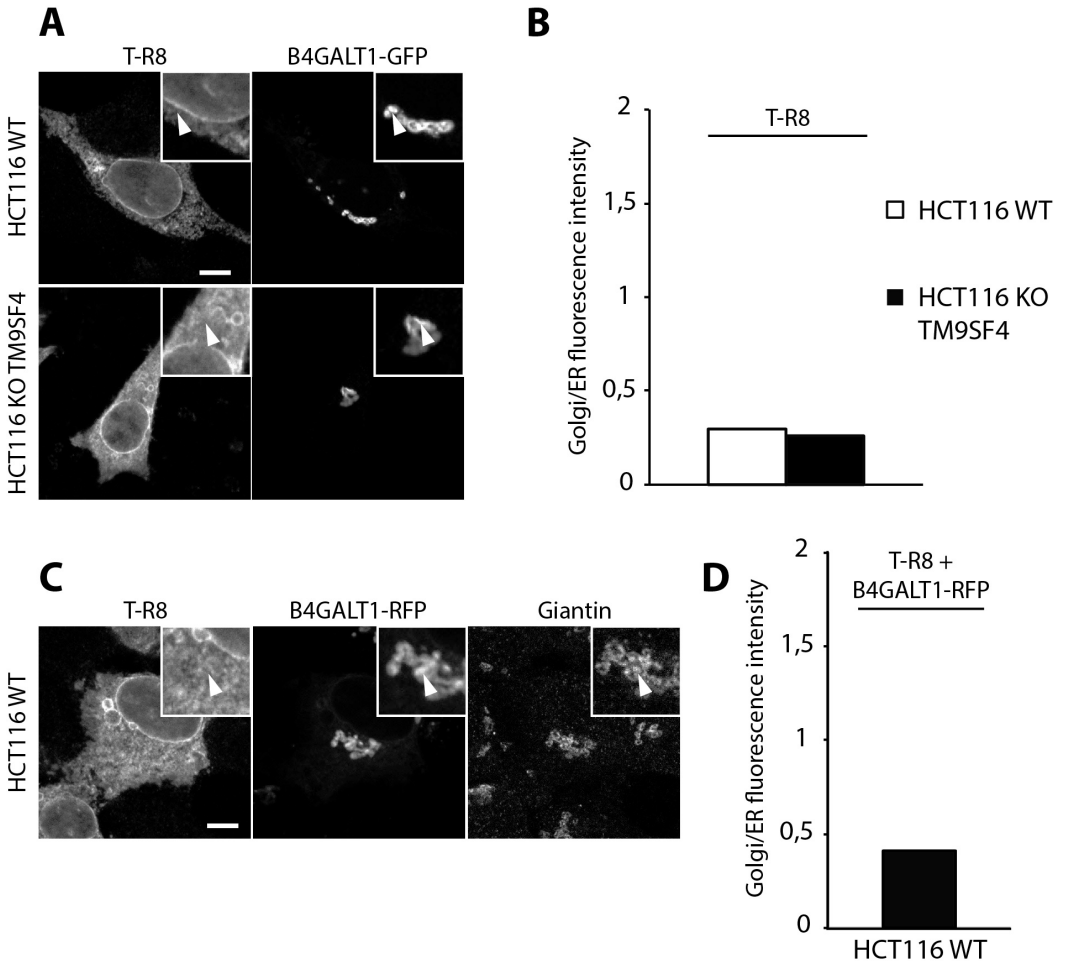

Figure 6: Competition of T-R8 and B4GALT1 for TM9SF4. (A) The localization of T-R8 was assessed by immunofluorescence in parental cells (WT, upper panel) and in cells deficient for TM9SF4 (KO TM9SF4, lower panel). Here, the Golgi complex is revealed with the B4GALT1 fused to the GFP (B4GALT1-GFP). A magnification of the Golgi complex area is represented in the upper right corner. When B4GALT1-GFP is overexpressed, T-R8 is only found in the ER and not in the Golgi complex anymore (arrowheads). This result is quantified in (B) where the relative intensity of T-R8 in the Golgi complex is compared to the relative intensity found in the ER. This data is the result of three independent experiments. (C) Parental HCT116 cells were co-transfected with T-R8 and B4GALT1 fused to the RFP. To control the Golgi complex integrity, this

organelle is also revealed thanks to an anti-giantin antibody. A magnification of the Golgi complex area is represented in the upper right corner. The giantin signal colocalizes with B4GALT1-RFP, and T-R8 is not present in this area (arrowheads). A quantification of this result is presented in (D). Scale bar: 10µm. Data presented here are the results of two independent experiments.

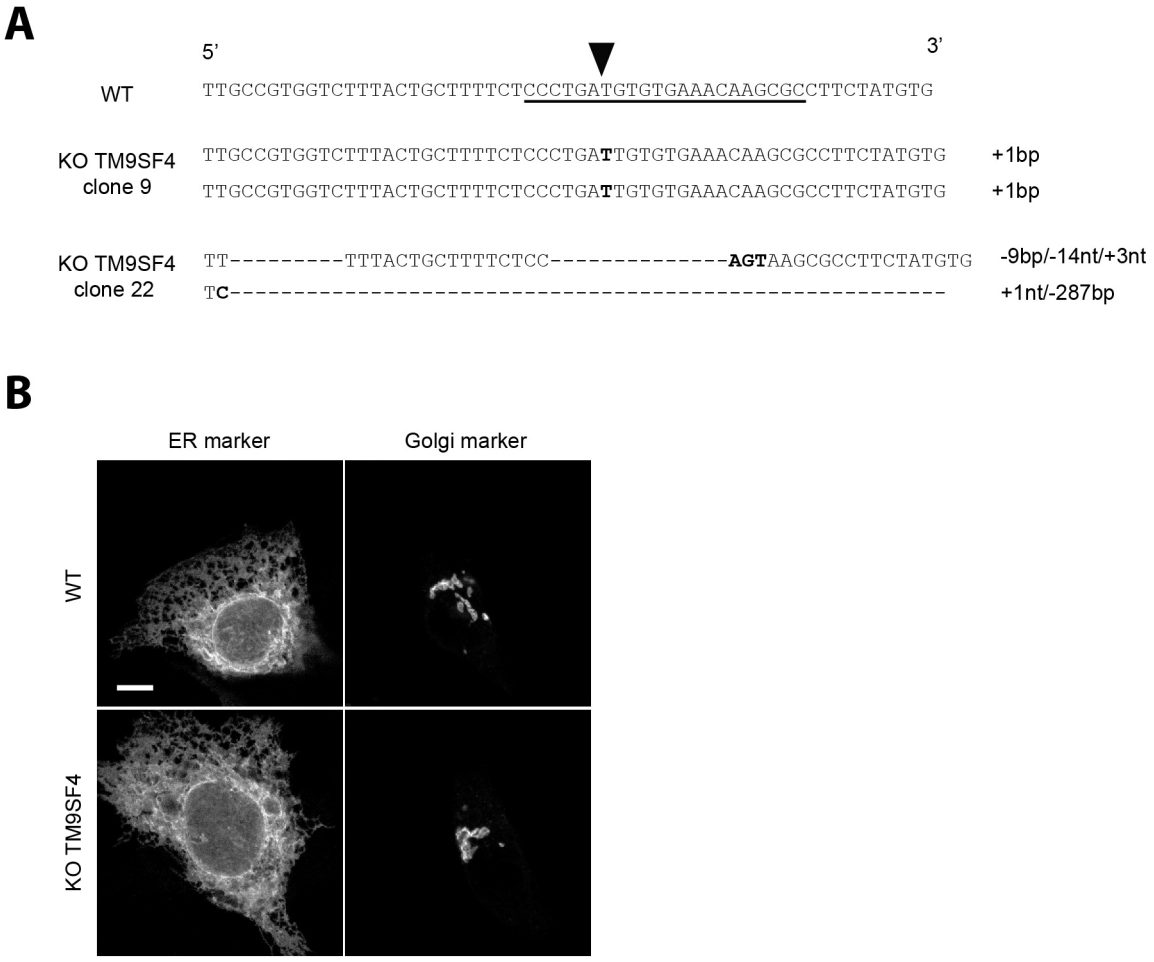

Figure S1: HCT116 cells deficient for TM9SF4. (A) The genomic DNA of HCT116 parental cells (WT) and deficient for TM9SF4 are represented. The target of the Cas9 nuclease is underlined and the arrow represents its cutting site. The sequences of the alleles of two independent clones are presented. Bold character stands for an insertion whereas (-) symbolizes a deletion. These two clones were analysed in parallel in this study with identical results. (B) The general organization of the early secretory pathway in HCT116 cells is not affected by the loss of TM9SF4. Cells were transfected with a 17-residue TMD fused to the GFP to reveal the ER and with B4GALT1-GFP to reveal the Golgi complex. Scale bar: 10µm.

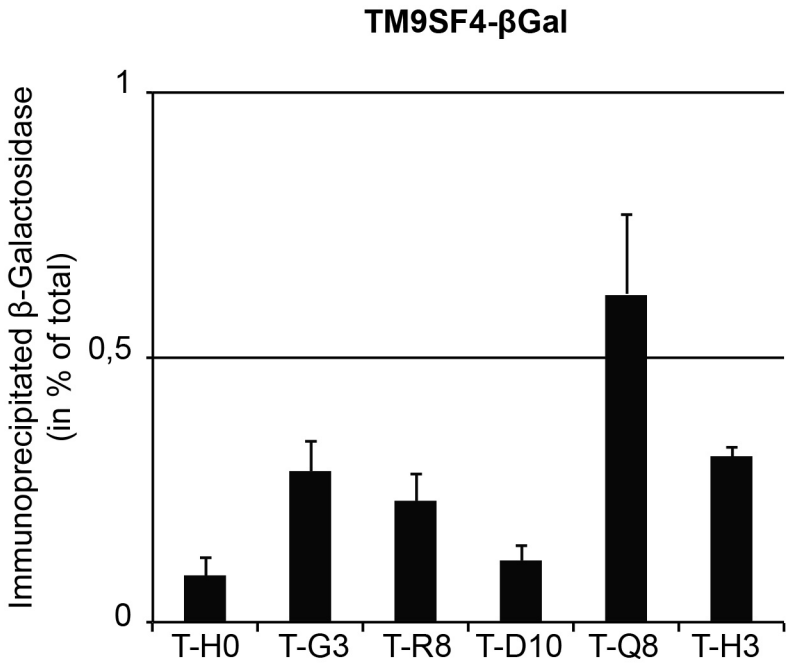

Figure S2: TM9SF4 interacts with a large set of TMDs. The same experiment than Figure 3 was performed. TM9SF4 also interacts with T-Q8 and T-H3 (the aminoacid sequences are presented in Table 1). The mean ± SEM of at least three independent experiments are indicated.

Name	Luminal	Transmembrane	Cytoplasmic
T-H0	…EY	QVAVAACVFLLIAVLLLSGLTWQ	RRQRKSRRTI
T-G3	…EY	QVAVAGCVFGLIGVLLGSGLTWQ	RRQRKSRRTI
T-G4	…DL	QVAVAGCVFGLIGVLLGSGLTWQGL	RRQRKSRRTI
T-R8	…EY	QVAVAACVRLLIAVLLLSGLTWQ	RRQRKSRRTI
T-D10	…EY	QVAVAACVFLDIAVLLLSGLTWQ	RRQRKSRRTI
T-H0KKxx	…EY	QVAVAACVFLLIAVLLLSGLTWQ	RRQRLETFKKTN
T-Q8	…EY	QVAVAGCVQLLISVLLLSGLTWQ	RRQRKSRRTI
T-H3	…EY	QVAVAGCVTTASTVLLLSGLTWQ	RRQRKSRRTI

Table 1: Aminoacid sequences of the Tac proteins used in this study.

III. DISCUSSION

III. DISCUSSION

Intracellular transport and sorting of transmembrane proteins are essential in eukaryotic cells to determine and maintain the biochemical composition and the specific functions of each cellular compartment. During my thesis, I had the opportunity to study protein sorting mechanisms both in *Dictyostelium* and in human cells. We showed that TM9 proteins are implicated in sorting of transmembrane proteins, based on determinants found in TMDs. Genetic inactivation of Phg1a in *Dictyostelium* induced a decrease in the surface localization of the adhesion molecule SibA, and in the stability of the Golgi sulfotransferase Kil1.

Detection of cell surface proteins was particularly important in these studies. As shown in my first publication (Immunofluorescence labeling of cell surface antigens in *Dictyostelium*), precautions must be taken in order to avoid the loss of cell surface antigens during the permeabilization procedures. As proposed in this publication, a two-step protocol largely reduces this type of artefact. This protocol is still not perfect and it must be used with care: it may lead to overestimate the surface amount of a given protein, and the intensity of the staining obtained cannot be directly compared to the intracellular labeling. Using this method, we showed that both Phg1a and TM9SF4 control the cell surface localization of glycine-rich TMDs. Indeed, in the absence of Phg1a/TM9SF4, glycine-rich TMDs are retained in intracellular compartments, mainly the ER.

One possible interpretation of these results was that Phg1a/TM9SF4 proteins may help glycine-rich TMDs to exit the ER. Accordingly, Phg1a/TM9SF4 proteins would interact with glycine-rich TMDs in the ER, and package them into COPII-coated vesicles.

This would lead to the efficient anterograde transport of glycine-rich TMDs, and would explain their surface localization seen when TM9SF4 is abundant. This model is described in Figure 22A. The dilysine-like motif (VKID) found in the C-terminal part of the human TM9SF4 could in principle induce its recycling, due to a direct interaction with COPI-coat components, but this remains to be formally demonstrated. In this context, the IFRTL-like sequence found in TM9SF4, as the one found in Erv14 for the anterograde transport, may be responsible for the concentration into COPII-coated vesicles. TM9SF4 exhibits a LYRTL sequence (highly similar to an IFRTL motif), in the cytosolic loop connecting TMD3 and TMD4. It would be interesting to mutate this motif to see if the anterograde transport of glycine-rich TMDs is still effective. In this hypothesis, the final localization of TM9SF4 is determined by the motifs for both anterograde and retrograde transport, that dynamically localizes TM9SF4 in the Golgi complex at steady state. It is however not excluded that other sorting motifs could be contained within the soluble or transmembrane parts of TM9SF4.

Another hypothesis would be that TM9SF4 captures TMDs in the Golgi complex. This would imply that TM9SF4 acts by retaining some TMDs in the Golgi complex and thus primarily plays a role in determining the composition of the Golgi complex. According to this model, this would increase the surface localization of a few proteins by allowing them to escape retrieval from the Golgi complex to the ER. Several lines of evidence support this hypothesis: (1) in *Dictyostelium,* the cellular levels of the Golgi sulfotransferase Kil1 are highly decreased in *phg1a* KO. This enzyme is degraded much faster in the absence of Phg1a suggesting that Phg1a localizes Kil1 in the Golgi complex. A direct interaction between TM9SF4/Phg1 and Kil1 has not however been demonstrated. (2) TMDs containing charged residues (such as T-R8, which exhibits an

arginine residue) are mainly retained in the ER in HEK cells. However, the overexpression of TM9SF4 causes their partial relocalization into the Golgi complex. In HCT116 cells, which overexpress TM9SF4, T-R8 is partially localized in the Golgi complex. This construct is largely relocalized in the ER by genetic inactivation of TM9SF4. (3) TM9-dependent Golgi localization of T-R8 is disturbed by the overexpression of the protein B4GALT1. Since B4GALT1 exhibits the TMD of an endogenous type-II Golgi enzyme, it could be a 'better' substrate for TM9SF4. However, the interaction between B4GALT1 and TM9SF4 is still to be confirmed. This second model is presented in the Figure 22B. TM9SF4 would act as a Golgi retention mechanism that stably localizes proteins in the Golgi complex, through their TMDs.

This model is based on the assumption that Golgi cisternae are stable compartments rather than transient compartments that perpetually mature. Indeed, the cisternal maturation model is not easily reconciled with our observation that TM9SF4 localizes TMDs in the Golgi complex.

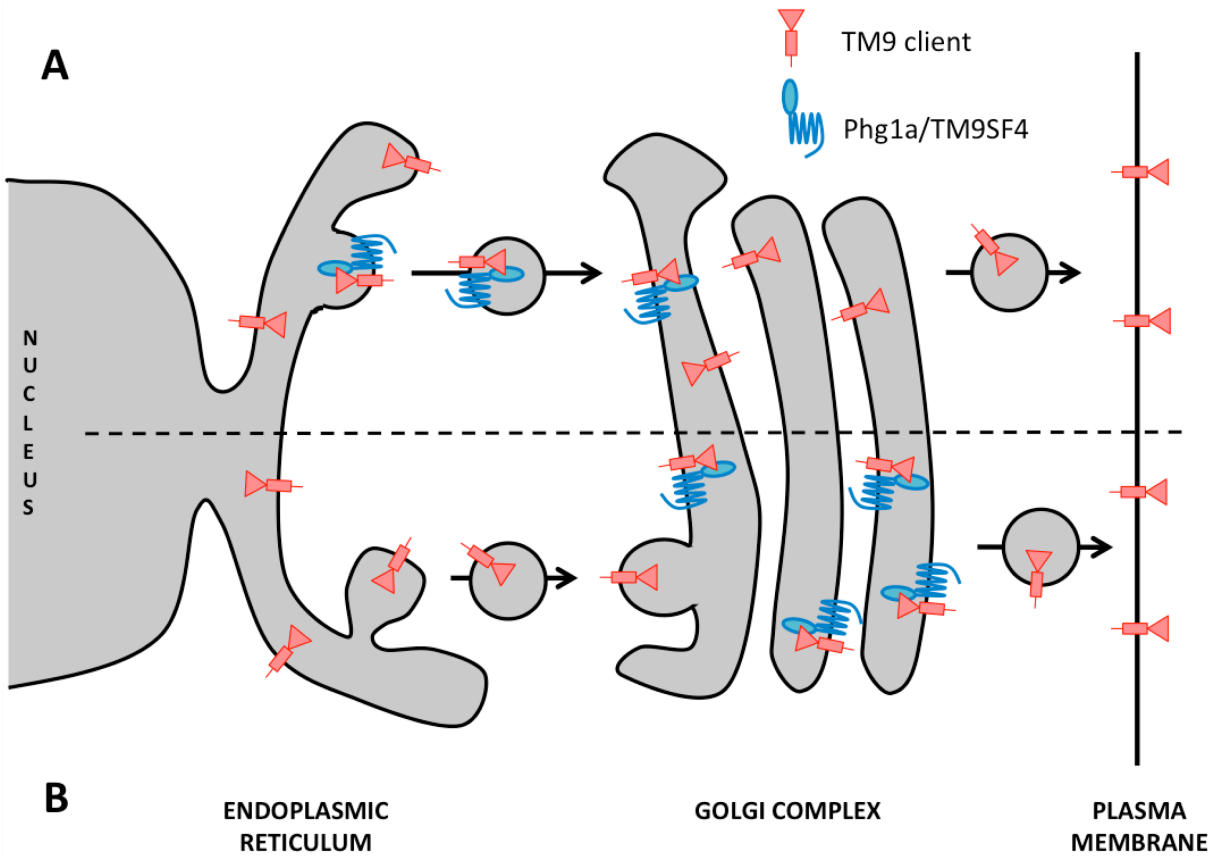

Figure 22: sorting of TMDs by Phg1a/TM9SF4. (A) TM9SF4 interacts with its TMD substrates in the ER and allows their packaging into COPII-coated vesicles. The complex TM9SF4-TMD is then transported to the Golgi complex, where these two proteins dissociate. Transported proteins can then progress through the Golgi complex and are now able to be localized in the Golgi complex. (B) An alternative model is that TMDs are localized in the ER, but a small fraction 'leaks' in the Golgi complex, where they interact with TM9SF4. As a consequence, TMDs are stabilized there and are now localized in the Golgi complex. In this model TM9SF4 acts as a Golgi complex localization mechanism for a subset of TMDs.

As mentioned in the introduction, the cisternae maturation model implies that Golgi enzymes are constantly recycled back to the previous compartment. This would imply that these enzymes are concentrated into the rims of the Golgi cisternae, where the vesicles form. This has been contradicted in a study showing by electron microscopy that mannosidase II, a Golgi resident enzyme, is excluded from peri-Golgi vesicles (Pierre Cosson, Amherdt, Rothman, & Orci, 2002). This suggests that Golgi enzymes are maintained in the Golgi complex by selective retention rather than constant recycling. In our model, TM9SF4 could be a mechanism by which Golgi enzymes are specifically retained in the Golgi complex.

The fact that glycine-rich TMDs are able to reach the plasma membrane when TM9SF4 is overexpressed must be reconciled with the fact that TM9SF4 localizes TMDs in the Golgi complex. One interpretation could be that the concentration of glycine-rich TMDs in the Golgi complex drastically increases in the presence of TM9SF4, and a little portion 'leaks' at the cell surface.

Attempts to find other cell surface proteins affected by the loss or the overexpression of TM9SF4 were unsuccessful. Indeed, biotinylation of all the cell surface antigens were performed both in HEK and HCT116 cells. Samples were then separated by SDS-PAGE and stained by silver staining (See Figure 23). No differences were observed in HEK overexpressing TM9SF4, compared to WT. Same results were observed in HCT116 KO TM9SF4, when compared to WT cells. However, the capacity to distinguish proteins is highly limited to the most abundant surface proteins with this technique. Moreover, our results suggest that the composition of the Golgi complex should be more impacted by the loss or the overexpression of TM9SF4 than the cell surface.

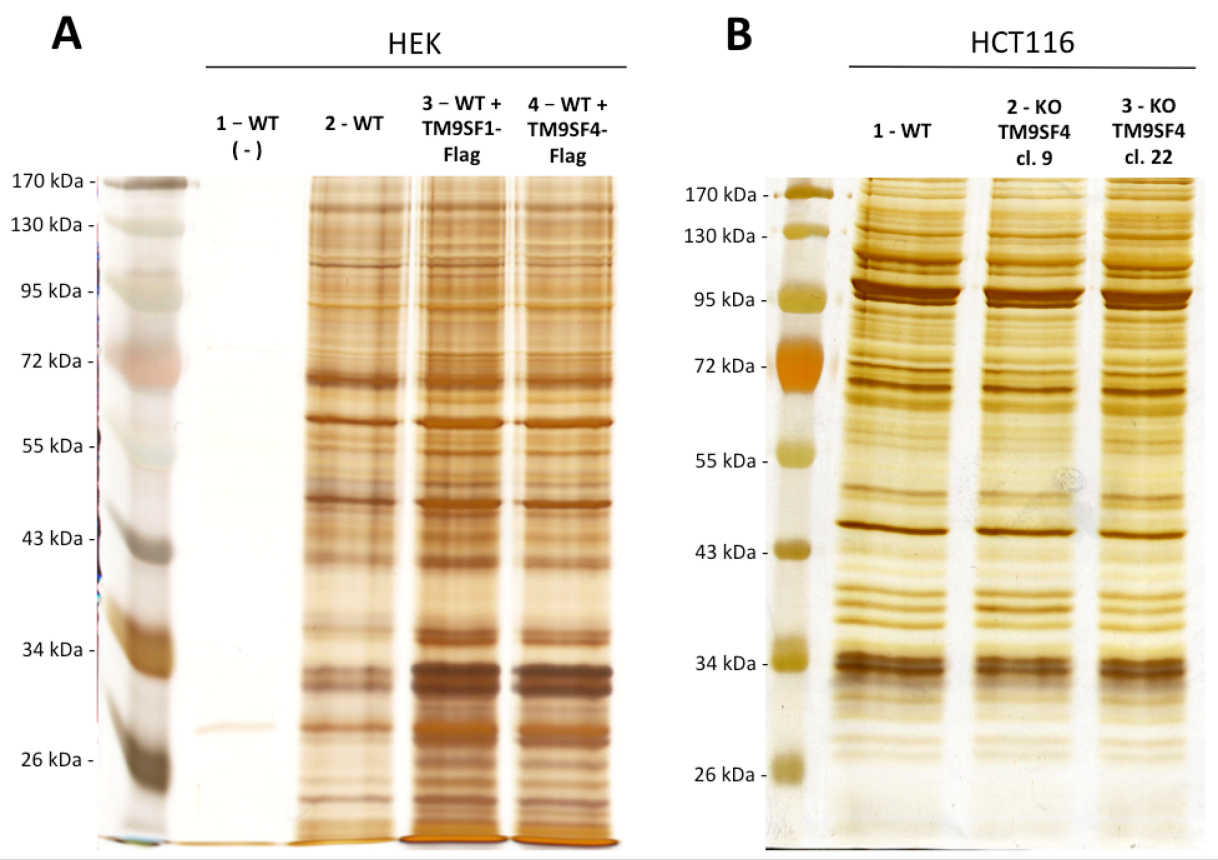

Figure 23: analysis of the cell surface composition of WT cells and cells overexpressing or depleted for TM9 proteins. (A) Surface proteins of HEK WT (line 2), overexpressing TM9SF1-Flag (line 3) or TM9SF4-Flag (line 4) were biotinylated, and precipitated with neutravidin beads. Purified membrane proteins were separated on SDS-PAGE and proteins were silver stained. A non-biotinylated control was also added (line 1). No obvious difference appears when TM9SF1-Flag or TM9SF4-Flag was overexpressed, reflecting the same plasma membrane composition between all the samples. (B) The same experiment was conducted in HCT116 WT (line 1) and deficient for TM9SF4 (two clones are presented, line 2 and line 3). Here again, the cell surface composition was the same in all the three conditions.

How does TM9SF4 interact with TMDs? It is not known if the recognition is based on one transmembrane domain of TM9SF4 or several. Moreover, the interaction could be direct or indirect. Concerning glycine-rich TMDs, TM9SF4 exhibits numerous glycine residues among its nine TMDs (1 residue in TMD1, 2 residues in TMD2, 3 residues in TMD3 and TMD4, 2 residues in TMD5, 3 residues in TMD6, 0 residue in TMD7, 2 residues in TMD8 and 1 residue in TMD9). As mentioned in the second publication in the Results part (TM9 family protein control surface targeting of glycine-rich

transmembrane domains), glycine residues in TMDs participate in specific association between TMDs, resulting in the formation of multimeric structures. This is the case for instance of several subunits of α and β integrins (C. Kim, Lau, Ulmer, & Ginsberg, 2009) or the amyloid precursor protein (APP), in which the three glycine residues present in its TMD were described to be necessary for the formation of homodimers (Kienlen-Campard et al., 2008). Therefore, the glycine residues of TM9SF4 may be of importance in the interaction with glycine-rich TMDs. Mutation of a few residues in TM9SF4 may abolish the interaction with glycine-rich TMDs. This would suggest a direct interaction between TM9SF4 and its substrate, rather than an indirect interaction. Concerning positively charged TMDs, the interaction with TM9SF4 may be mediated by the glutamic acid (E) found in the sixth TMD of TM9SF4, the only negatively charged aminoacid found in TM9SF4's TMDs. Mutation of this aminoacid may abolish the interaction with positively charged TMDs. Charged residues in TMDs are much less frequent than glycine residues. Only a handful of type I transmembrane proteins exhibit a TMD with a charged residue. Positively charged residues can for instance be found in the α chain of the T-Cell Receptor (TCR), or the envelope protein of the HIV, whereas negatively charged residues can be found in the δ and the ε chains of the TCR.

In addition to localizing TMDs in the Golgi complex, TM9SF4 may have other functions. Indeed, in cancer cells, TM9SF4 has been described to interact with the ATP6V1H subunit of the vacuolar-ATPase (V-ATPase), one of the proton pump responsible for the abnormal acidification of the extracellular medium, observed in malignant cancer cells (Lozupone et al., 2015). This unusual acidic environment is favourable for tumor progression and metastasis. Moreover, the silencing of TM9SF4 using small interfering RNA reduced the assembly of the V-ATPase complex, associated

with a decrease in the invasive ability of cancer cells. Taken together, these data suggest that TM9SF4 regulates the assembly of the V-ATPase. We could speculate that TM9SF4 interacts with the transmembrane ATP6V1H subunit through TMDs, and stabilizes it. It is however worth noting that in this study, TM9SF4 was described to be found in endosomal vesicles, whereas our studies indicate that the majority of TM9SF4 is present in the Golgi complex. Another study recently suggested that TM9SF4 was mediating the adhesion of cells to fibronectin. Moreover, this study suggests that TM9SF4 is sensitive to hypoxia, its expression decreasing under low-oxygen conditions, a condition occurring in the cancer microenvironment. These experiments were performed in acute myeloid leukaemia cells. In order to reconcile our observations with these results, we hypothesize that TM9SF4 can indirectly control the composition, and hence the function of many cellular compartments.

We speculate that TM9SF4 may be able to transduce signals. For instance, a yet unidentified ligand could bind the large N-terminal domain, induce a conformational change that is transmitted to the C-terminal part. It has been previously shown that TM9SF4 was able to bind a β-adrenergic agonist (Sugasawa et al., 2001). This could be interesting to study for instance the phosphoproteome of cells WT and deficient for TM9SF4 to see if any change is visible.

The role of the other members of the TM9 family is still unknown. During my thesis, I also studied the role TM9SF1 in the sorting of TMDs. TM9SF1 belongs to the group B of TM9 proteins whereas TM9SF4 belongs to the group A (see Figure 22, in the Results parts). TM9SF1 has been described to induce autophagy, since its overexpression led to an increase in the number of the autophagosomal marker LC3-

GFP dots (He et al., 2009; E. C. Kim, Meng, & Jun, 2013). In our results, a Flag-tagged version of TM9SF1 localizes mainly in the Golgi complex, but is also present in other compartments which could be endosomes or lysosomes (Figure 24A). TM9SF1 also interacts with glycine-rich TMDs as exemplified in Figure 24B, even stronger than TM9SF4. However, neither the loss of expression of TM9SF1, nor its overexpression have any effect on the localization of glycine-rich TMDs (See Figure 24C, and the quantification in Figure 24D). This could mean that TM9SF1 is not implicated at all in the localization of TMDs, or that it regulates their transport at another step of the intracellular transport. For instance, it could be involved in the endocytosis of glycine-rich TMDs, which would then be targeted to the endosomes. This process could have been invisible to us since the Tac protein used as reporter protein is rapidly degraded in the endo/lysosomes (Marks et al., 1995). Another reporter protein could be use to solve this problem of degradation, such as CD1b, much more resistant to proteolysis. This could answer the question of whether TM9SF1 is implicated in the endocytosis of TMDs.

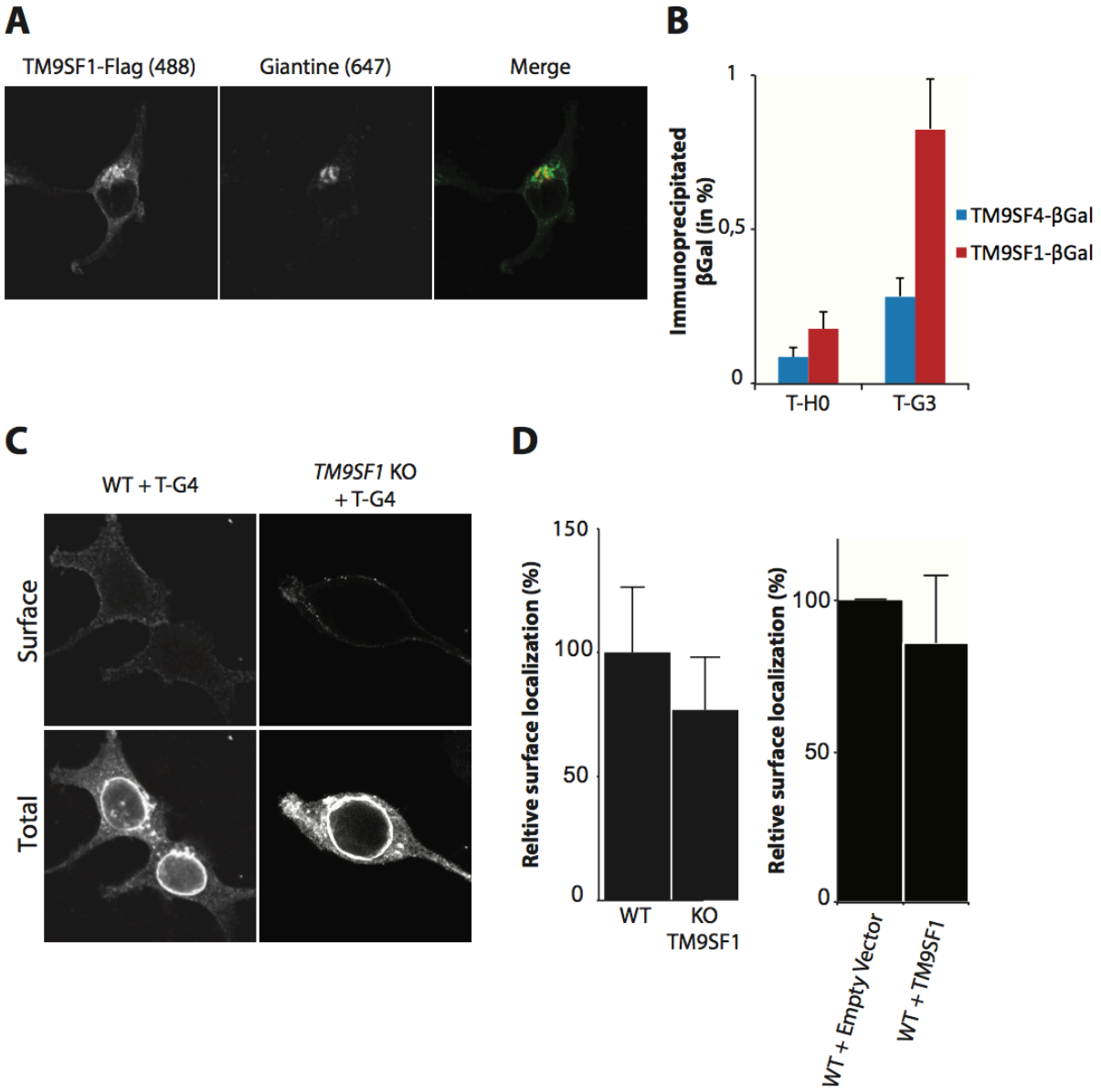

Figure 24: TM9SF1 and glycine-rich TMDs. (A) a FLAG-tagged version of TM9SF1 was transfected in HEK cells, and revealed by an anti-Mouse Alexa-647-coupled antibody. The Golgi complex is marked by an anti-Giantine antibody and revealed with an anti-human Alexa-488-coupled antibody. (B) The association between TM9SF1 fused to the β-Galactosidase and glycine-rich TMDs has been assessed, as it is described in the second publication of the Results part. Briefly, a Tac chimeric protein with 1 glycine residue (T-H0) or 5 glycine residues (T-G3) in its TMD is co-expressed with TM9SF1 fused to the β-Galactosidase, in HeLa cells. The Tac protein is immunoprecipitated and the percentage of β-Galactosidase acivity co-precipitated is determined. The association between TM9SF1-βGal and T-G3 is much stronger than with T-H0. For comparison, the association with TM9SF4-βGal is indicated in blue. (C) An immunofluorescence of T-G4 (a Tac construct containing 6 glycine residues in its TMDs) was performed. The Tac protein at the surface is first labelled before permeabilization and labelling of the intracellular content, as indicated in the second publication of the Results part. This experiment has been performed both in HEK WT cells and in cells deficient for TM9SF1. The loss of TM9SF1 did not change the surface localization of T-G4. (D) The surface localization of T-G4 is not impacted by the loss (left panel) or the overexpression (right panel) of TM9SF1.

Little is known about TM9SF2. A myc-tagged version of TM9SF2 has been described to be found in the endosomes, since it colocalizes with the transferrin receptors and some mannose 6-phosphate receptors (Schimmöller, Díaz, Mühlbauer, & Pfeffer, 1998). More recently and in *Drosophila*, TM9SF2 has been shown to be necessary for the correct activity of the peptidoglycan recognition protein (PGRP)-LC, an important transmembrane protein of the *Drosophila* immune system. TM9SF4, the closest homolog of TM9SF2, has been shown to be implicated in the surface localization of PGRP-LC (Perrin et al., 2015). These results are highly consistent with ours, and may indicate partial redundancy between these two TM9 proteins. PGRP-LC exhibits at most two glycine-residues in its TMD, which is probably too low to confer specificity towards TM9SF4, but additional determinants may be recognized by TM9SF2/4.

Concerning TM9SF3, it has been reported to be localized in the Golgi complex in male germ cells (Au et al., 2015). In another study, TM9SF3 has been shown to be a potential therapeutic target for scirrhous-type gastric cancer, since (1) the expression of TM9SF3 is correlated with poor prognosis and (2) silencing of TM9SF3 in cancer cells decreases their invasion capacity (Oo et al., 2014).

During my PhD work, I had the opportunity to use the recently described CRISPR/Cas9 system. I made use of this method to create knock-out cell lines for TM9SF1, TM9SF4 and mitoNEET in HEK, HCT116 and MEF cells. At this time, this technique was really new, and it was extremely stimulating to learn how it was working. The obtention of these KO cell lines was critical for the studies performed for this work. Indubitably, this new method is highly powerful and becomes nowadays indispensable in a growing number of laboratories around the world. This method should be however

used carefully. Indeed, the off-target effects, meaning the mutation in the genomic DNA at one or several unexpected site(s) should be taken into account. To circumvent this problem, we decided to study several independent clones to validate the phenotype of the different knock-out cell lines. Definitely, this technique will take more and more importance in the coming years.

In conclusion to this work, my work shed new light on the function of TM9 proteins. My results have been obtained first in the *Dictyostelium* model, and then in human cells. Evidence shows that TM9SF4 interacts with a subset of TMDs and allows their progression in the secretion pathway. Our results suggest that TM9SF4 may localize TMDs in the Golgi complex. To the best of our knowledge, TM9SF4 would therefore be the first mechanism described that ensures the Golgi localization of transmembrane domains. These results need to be confirmed, and this work will be continued in the laboratory.

IV. APPENDICES

IV. APPENDICES

A. Publication: MitoNEET-dependent formation of inter-mitochondrial junctions

1. Introduction

During my PhD, I had the opportunity to study the role of mitoNEET in the formation of inter-mitochodrial junctions (IMJs). We showed that mitoNEET was a tethering factor, responsible for the establishment of the mitochondrial network in the cell. This work is currently under revsion in Proceedings of the National Academy of Sciences (PNAS).

2. Publication

MitoNEET-dependent formation of inter-mitochondrial junctions

Alexandre Vernay, Anna Marchetti, Ayman Sabra, Manon Rosselin, Philipp E. Scherer, Nicolas Demaurex, Lelio Orci, Pierre Cosson

Under revision in PNAS

In this study, I contributed to the Figure 1, Figure 3, Figure 4, Figure S1 and Figure S2, as well as the writing of the manuscript.

MitoNEET-dependent formation of inter-mitochondrial junctions

Alexandre Vernay[1], Anna Marchetti[1], Ayman Sabra[1], Manon Rosselin[1], Philipp E. Scherer[2], Nicolas Demaurex[1], Lelio Orci[1], Pierre Cosson[1]*

[1] Faculty of Medicine, University of Geneva, Centre Médical Universitaire, 1 rue Michel Servet, CH1211 Geneva 4, Switzerland

[2] Touchstone Diabetes Center, Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390–8549, USA

*Corresponding author: Pierre.Cosson@unige.ch

ABSTRACT

MitoNEET (mNEET) is a dimeric mitochondrial outer membrane protein implicated in many facets of human pathophysiology, notably diabetes and cancer, but its molecular function remains poorly characterized. In this study we generated and analyzed *mNEET KO* cells, and found that in these cells the mitochondrial network was disturbed, and mitochondrial respiration decreased. Detailed ultrastructural analysis revealed that genetic inactivation of mNEET did not affect the size of mitochondria, but that the frequency of inter-mitochondrial junctions was reduced. In the same cells, the frequency of contacts between mitochondria and the endoplasmic reticulum was not decreased, and the regulation of mitochondrial fusion and fission by cycloheximide or H_2O_2 was still operating. Conversely, overexpression of mNEET increased strongly the formation of contacts between mitochondria. Our results suggest that mNEET plays a specific role in the formation of inter-mitochondrial junctions, a mechanism distinct from tethering of mitochondria to ER or from fusion and fission of mitochondria.

INTRODUCTION

Mitochondria play a key role in many facets of cellular physiology, notably metabolism and the production of ATP, and storage of calcium. Mitochondrial dysfunction has been linked to a wide variety of human pathologies including diabetes (1), neurodegeneration, and cancer (2). In many pathological situations, alterations of mitochondrial morphology have been observed, and may play an important role in the development of the observed physiological anomalies (3). The molecular mechanisms ensuring the control of mitochondrial morphology may thus represent targets for therapeutic interventions in these various pathologies. Several key elements of the fusion/fission machinery have been identified and characterized, in particular four dynamin-related proteins: mitofusin 1 and mitofusin 2 promote fusion of mitochondrial outer membranes, optic atrophy 1 promotes fusion of inner membranes, and DRP1 ensures fission of outer mitochondrial membranes (4).

MitoNEET (mNEET) was originally identified as a mitochondrial binding site of pioglitazone, an insulin sensitizer (5). Although it was later revealed that the main mitochondrial binding site of pioglitazone was the Mcp mitochondrial pyruvate carrier complex (6), this initial observation launched a series of studies to unravel the molecular function of the mNEET protein. mNEET belongs to a family of three proteins exhibiting a CDGSH domain, together with miner1 and miner2. mNEET and miner2 are integral proteins of the outer mitochondrial membrane, with their CDGSH domain located in the cytosol. Miner1 is an integral protein of the endoplasmic reticulum (7). Genetic inactivation of mNEET reduces by approximately 30% the maximal capacity of heart mitochondria to transport electrons and carry out oxidative phosphorylation (7).

The CDGSH of mNEET binds a 2Fe-2S cluster that can undergo oxidation and reduction. In addition to its putative role in diabetes, mNEET has been implicated in tumor development: mNEET and miner1 are overexpressed in breast cancer cells, and their down-regulation reduces cell proliferation and tumor growth (8). mNEET has also been overexpressed experimentally in adipocytes and in pancreatic cells in order to modulate their physiology (9, 10). The exact function of mNEET remains however elusive, as well as the effect of its overexpression on cellular physiology.

Overexpression of mNEET in pancreatic β cells induced the formation of mitochondrial clusters (10), raising the intriguing possibility that mNEET may participate in the control of the mitochondrial network morphology. The current study was designed in order to test this hypothesis. Our results suggest that mNEET is involved in the formaton of inter-mitochondrial junctions (IMJs), thus playing an important part in the establishment of the mitochondrial network in the cell.

RESULTS

Loss of mNEET decreases the frequency of inter-mitochondrial junctions

In order to assess the role of mNEET we used the CRISPR-Cas9 method to generate MEF cells in which the gene coding for mNEET was genetically inactivated (Fig. S1A)(7). The general organization of the endoplasmic reticulum (ER), the Golgi complex and the tubulin cytoskeleton was indistinguishable between WT and *mNEET KO* cells (Fig. S1B), but mitochondrial respiration was reduced by approximately 30% (7)(Fig. S1C). We then analyzed the morphology of mitochondria in parental (WT) MEFs and in *mNEET KO* cells expressing a mitochondrially-targeted fluorescent fusion protein. Compared to parental cells, the mitochondrial network appeared less connected in *mNEET KO* cells (Fig. 1A). Since the morphology of mitochondria varied significantly from cell to cell, this observation was quantified by counting the percentage of cells with various types of mitochondrial network, from fully fragmented (type 4) to fully connected (type 2) or hyperconnected (type 1)(Fig. S2). Compared with parental cells, three independent clones of *mNEET KO* cells exhibited a significantly less connected mitochondrial network (Figure 1B). For simplicity, in this study, the degree of connectivity of the mitochondrial network was defined by a single number between 0 (no connectivity) and 200 (hyperconnected) (Fig. 1C). In live unfixed cells, a similar fragmentation of the mitochondrial network was observed (data not shown).

A decrease in the connectivity of the mitochondrial network could in principle reflect a fragmentation of mitochondria, characterized by a decrease in the size of individual mitochondria. However the optimal resolution of classical fluorescence microscopy is

200nm, and does not allow to delimitate unambiguously single mitochondria. In order to obtain 3-dimensional reconstructions of cells with adequate resolution, we used a dual-beam focused ion beam/scanning electron microscope (FIB/SEM). The focused ion beam gradually trims the sample (section thickness 10nm), while the scanning electron beam visualizes the milled surface with a lateral resolution of approximately 4nm (Fig. 2). To assess the size of mitochondria, the number of sections traversed by each individual mitochondria was determined (z-size). Contrary to our expectations, the z-size of mitochondria in *mNEET KO* cells (1.33±0.08µm, n=103) and in WT cells (1.25±0.09µm, n=106) was not significantly different (p=0.48 ; Student t test) (Fig. 2A and Table S1). These observations suggest that changes in the size of individual mitochondria do not account for the modification of the mitochondrial network observed by fluorescence microscopy in *mNEET KO* cells.

As previously reported (11), mitochondria can establish close contacts, referred to as inter-mitochondrial junctions (IMJs) and this could in principle increase the apparent connectivity of the mitochondrial network. IMJs were readily observed by FIB/SEM microscopy (Fig. 2B, arrowheads) and this technique also allowed to ascertain that no fusion between mitochondria was taking place next to the regions of close contacts. In order to evaluate the significance of IMJs in connecting the mitochondrial network, we first determined the frequency of IMJs in WT cells. In WT cells, 52% of mitochondria were in contact with at least one other mitochondria, and 16% with at least two. IMJs represented 0.8 % of the total surface of mitochondria (Table S1). These contacts, which cannot by fluorescence microscopy be distinguished from fused mitochondria, would thus contribute significantly to the apparent connectivity of the mitochondrial network.

In *mNEET KO* cells, contacts between mitochondria were significantly less abundant than in WT cells: only 34% of mitochondria participated in at least one contact (p=0.012 Fisher's exact test), and 0.52% of the total mitochondrial surface was engaged in IMJs (Fig. 2C and Table S1). This observation suggested that loss of mNEET affects the connectivity of the mitochondrial network by decreasing the formation of IMJs.

We next examined conventional epon-embedded sections by transmission electron microscopy and assessed the frequency of intermitochondrial contacts, defined as regions where two apposed mitochondrial membranes were separated at most by 20nm (Fig. 2D, arrowheads). This technique provides a better resolution than FIB/SEM microscopy (lateral resolution 1nm), and allows the examination of sections collected from a larger number of cells. In WT cells, 6.3% of individual mitochondrial sections were apposed to another mitochondrial section (Fig. 2E and Table S2), and 0.8% of the mitochondrial surface was engaged into contacts with other mitochondria (Table S2). In sections of *mNEET KO* cells, the frequency of IMJs was significantly decreased compared to parental cells (2.2%; p=0.0009 Fisher's exact test)(Fig. 2E and Table S2). The surface of mitochondria involved in IMJs was also decreased to 0.4% (Table S2).
Overall, these observations indicate that genetic ablation of mNEET decreases significantly IMJs, suggesting that mNEET plays a key role in the establishement of inter-mitochondrial junctions.

Overexpression of mNEET increases the frequency of IMJs

We next assessed the effect of mNEET overexpression in cultured MEF cells. For this, we transfected *mNEET KO* cells and first assessed the mitochondrial morphology by

fluorescence microscopy in cells expressing low levels of mNEET-GFP. In these cells, a tubulated network of mitochondria was restored (Fig. 3A and B), confirming that the fragmentation of mitochondrial network in *mNEET KO* cells was caused by the loss of expression of mNEET. Cells expressing high levels of mNEET-GFP were not taken into account in that quantification. In these cells, mitochondria were frequently clustered (Fig. 3A), suggesting that overexpression of mNEET-GFP further increased contacts between mitochondria, eventually causing a collapse of the whole mitochondrial network.

In order to quantify contacts between mitochondria in cells overexpressing mNEET-GFP, WT cells were transfected with mNEET-GFP, sorted by flow cytometry and fixed one day later. Sections were then prepared for observation by conventional transmission electron microscopy. Contacts between mitochondria were much more frequent in cells overexpressing mNEET and mitochondria engaged in contacts with several other mitochondria were commonly observed (Fig. 3C). In some instances, zipper-like structures appeared to tether the adjacent mitochondrial membranes at sites of IMJs (Fig. 3C, arrowheads). Quantification revealed that more than 40% of individual mitochondrial profiles showed at least one contact (Fig. 3D and Table S3). Similarly, the percentage of the mitochondrial surface engaged in IMJs was strongly increased in cells overexpressing mNEET (13%) compared to parental cells (0.8%)(Table S3).

The fact that overexpression of mNEET leads to an increase in IMJs reinforces the suggestion that mNEET participates in the tethering of mitochondria and formation of IMJs.

mNEET-dependent formation of IMJs and mitochondrial fusion are mechanistically distinct

We next assessed whether the effect of mNEET overexpression was dependent on mitofusins. For this we used *mfn2 KO* and *mfn1/2* double KO MEF cells. Note that unlike *mNEET KO* cells described above, these cells are not derived directly from the MEF WT cells, so that phenotypes of WT, *mfn2 KO* and *mfn1/2 KO* can at best only be roughly compared. In both cell lines, overexpression of mNEET resulted in a very strong increase in IMJs (Fig. 3E-F), indicating that the ability of mNEET to promote formation of IMJs does not require the presence of mitofusins.

The mitochondrial network is rapidly fragmented in cells exposed to an oxidative stress (12, 13). On the contrary, other types of stress, and in particular inhibition of protein synthesis by cycloheximide, increase mitochondrial connectivity (14). In both cases, changes in the connectivity of the mitochondrial network are due to changes in the fusion and fission rates of mitochondria. Exposure to cycloheximide increased the connectivity of the mitochondrial network in both parental and *mNEET KO* cells (Fig. 4A). Conversely, exposure to increasing concentrations of H_2O_2 decreased the mitochondrial network connectivity in both parental cells and *mNEET KO* cells (Fig. 4B). These observations suggest that mechanisms regulating fusion and fission of mitochondria still operate in cells devoid of mNEET.

ER-mitochondria contact sites in *mNEET KO* cells

In addition to forming IMJs and participating in the connectivity of the mitochondrial network, mNEET could in principle tether mitochondria to the ER, in particular since the ER membrane contains miner1, a protein highly homologous to mNEET (7). To test this hypothesis, we visualized by conventional electron microscopy ER-mitochondria contacts (Fig. 5A) and quantified their frequency (15). The frequency of ER-mitochondria contacts was not diminished, and actually slightly elevated in *mNEET KO* cells compared to WT cells (Fig. 5B and Table S4). In addition, overexpression of mNEET-GFP did not increase ER-mitochondria contacts (Fig. 5B and Table S4). These results indicate that mNEET does not play a critical role in the establishment of ER-mitochondria contact sites.

DISCUSSION

In this study we observed and quantified intimate contacts between mitochondria. To the best of our knowledge, the intermitochondrial contacts observed in this study are identical to or apparented to inter-mitochondrial junctions (IMJs) described recently (11). As previously seen, we observed a close apposition of mitochondrial membranes and in some instances an increase in electron density at the level of inter-mitochondrial contacts. We also noted that, as previously reported, IMJs are more prominent in some cell types than in others. For example, in Hela cells, 2.3% of the mitochondrial surface was engaged in IMJs (773 mitochondrial sections analyzed), against 0.8% in MEF cells. No molecular mechanisms have been described so far that ensure close apposition of mitochondria.

Our observations strongly suggest that mNEET is involved in the formation of IMJs, since the frequency of IMJs is decreased in *mNEET KO* cells, and increased in mNEET-

overexpressing cells. The most simple mechanism by which mNEET could participate in the formation of IMJs is by establishing a direct link between two mNEET molecules on adjacent mitochondria. Indeed, structural studies have established that mNEET forms dimers, and that the linker connecting the dimerizing CDGSH domain with the mitochondrial membrane exhibits considerable flexibility (16). These features may allow mNEET molecules on neighbouring mitochondria to form dimers tethering two adjacent membranes. An indirect role of mNEET in formation of IMJs cannot however formally be ruled out at this stage. Three lines of evidence suggest that the mechanisms ensuring mNEET-dependent formation of IMJs and fusion/fission of mitochondria are largely distinct. First, the size of mitochondria is unaffected in *mNEET KO* cells compared to WT cells. Second the regulation of fusion/fission by changes in cellular physiology still operates in *mNEET KO* cells. Third, mNEET overexpression increases IMJs even in cells devoid of mitofusin 1 and 2.

From a methodological point of view, this study exemplifies the power of ultrastructural analysis to assess variations in the morphology of the mitochondrial network. First, electron microscopy provides the resolution necessary to delimitate mitochondria, and to visualize IMJs and ER-mitochondria contact sites. Second, it is amenable to reliable quantification. Third, recent technical progress in ultrastructural analysis now allows the whole cellular mitochondrial network to be visualized and quantified. Quantitative ultrastructural analysis may in some situations be an indispensable tool to understand situations where the mitochondrial network is affected.

From a functional point of view, the role of mNEET and of IMJs remains to be established. Previous reports have indicated that mNEET is overexpressed in tumor

cells (8) and depleted in cells from cystic fibrosis patients (17). Our results suggest that in these pathological conditions, changes in the levels of cellular mNEET could alter the connectivity and the metabolic function of the mitochondrial network. In addition to pioglitazone, small molecules targeting mNEET or miner1 have been characterized (18, 19) and may be used to treat cancer, diabetes or other pathological conditions. In this perspective, it is essential to understand fully the cellular function(s) of mNEET, miner1 and miner2 to determine the effect(s) of such compounds on cellular physiology. Our results suggest that compounds targeting mNEET may affect the formation of inter-mitochondrial junctions as well as the bioenergetics of mitochondria. The functional consequences of such alterations at the cellular and at the organism level remain to be established.

MATERIALS AND METHODS

Cell culture and reagents

Mouse embryonic fibroblasts (MEF) cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% Fetal Bovine Serum, Penicilin-Streptomycin, and Non-Essential Amino Acids.

We generated *mNEET*-knockout cell lines using the CRISPR/Cas9 method, as described previously (21). Briefly, a plasmid purchased from DNA2.0 was used, the guide RNA targeting the first exon of *mNEET* (AGCTCCAACTCCGCTGT*GCGAG, * representing the cutting site of the Cas9 nuclease). Eight individual clones were screened for mutations by PCR. We amplified 500bp upstream and downsteam the cutting site on the genomic DNA (purified using the QIAamp DNA Blood MiniKit, Qiagen), using the following primers: primer 1, 5'-GTGTAACTTATTACCAAAAGT-3' and primer 2, 5'-CAGTCAGTCACGCATATC-3'. Fragments obtained were then purified and sequenced. Three individual clones were obtained with indels inducing a frameshift in both alleles. These three clones were used in parallel with very similar results.

The plasmid allowing overexpression of mNEET-GFP is based on a pEGFP-N1 backbone (7) and was a kind gift of Dr. S. E. Wiley (University of California San Diego, USA).

Fluorescence microscopy

To visualize the mitochondrial network, cells were grown on 20mm glass coverslips and transfected with mitoRFP or mNEET-GFP 2 days before the experiment, using Polyethylenimine (PEI) as described (20). Cells were then directly fixed in 4%

paraformaldehyde for 30 min, permeabilized with 0.2% Saponin, washed in PBS and mounted in Möwiol. When challenged with an oxidative stress, cells were treated for 1h with hydrogen peroxide before fixation. Hyperfusion of mitochondria was induced by incubating the cells in the presence of 25µg/mL cycloheximide for 6h prior to fixation. Fluorescence was imaged by confocal microscopy (LSM700, Zeiss). As shown on Supplementary Figure 2, for each sample we examined 150 to 200 individual cells, and scored the aspect of the mitochondrial network (4: totally fragmented ; 3: partially fragmented ; 2: tubular ; 1: hyperconnected). To summarize this result into a single number, we calculated the Tubulation Index: 2x(percentage of cells with an hyperconnected network) + (percentage of cells with a tubular network) + 0.5x(percentage of cells with a partially fragmented network). Importantly, all quantifications were performed on blinded samples to avoid experimental biaises.

Electron microscopy

To obtain Epon-embedded sections for conventional electron microscopy, cells were grown in 35 mm plastic dishes, fixed with 0.1M sodium phosphate, pH 7.4 containing 2% glutaraldehyde, postfixed with osmium tetroxide, stained with uranyl acetate, dehydrated in ethanol and embedded in Epon resin. After sectioning, the samples were observed in a Morgagni electron microscope. The iTEM software was used for quantification.

In order to generate high-resolution 3-dimensional reconstructions, we analyzed samples using an Helios DualBeam NanoLab 660 scanning electron microscope. Images were further analyzed using the AMIRA software. To evaluate the size of mitochondria, we counted for each mitochondria the number of sections through which it extended to determine its size along the z axis (z-size).

Mitochondrial respiration

The oxygen consumption rate (OCR) was measured using an XF96 extracellular analyzer (Seahorse Bioscience) according to the manufacturer's protocol. Briefly, 10'000 cells were seeded in 96-well plates for 24h in culture medium. Cells were then incubated in Krebs-Ringer bicarbonate HEPES buffer (KRBH, 135mM NaCl, 3.6mM KCl, 10mM HEPES, pH 7.4, 2mM $NaHCO^{3-}$, 0.5mM NaH_2PO_4, 0.5mM $MgCl_2$, 1.5mM $CaCl_2$ supplemented with 10mM Glucose) for 45 min at 37°C in a no-CO_2 incubator. Cells were then transferred into a Seahorse analyzer at 37°C and were sequentially treated with 1µM Oligomycin, 300nM Carbonyl cyanide-p-trifluoromethoxyophenylhydrazone (FCCP) and 0.5µM of a mixture of Rotenone and Antimycin A. OCR was automatically mesured after addition of each compound, as the average of 2 readings from 9 wells. At the end of each experiment, the amount of protein in each well was quantified, and the oxygen consumption values were corrected to avoid differences caused by variations in the number of cells.

REFERENCES

(1) Montgomery MK, Turner N. Mitochondrial dysfunction and insulin resistance: an update. Endocr Connect 2015: 4: R1-R15.

(2) Giampazolias E, Tait SW. Mitochondria and the hallmarks of cancer. FEBS J 2015.

(3) Archer SL. Mitochondrial dynamics--mitochondrial fission and fusion in human diseases. N Engl J Med 2013: 369: 2236-2251.

(4) Hoppins S, Lackner L, Nunnari J. The machines that divide and fuse mitochondria. Annu Rev Biochem 2007: 76: 751-780.

(5) Colca JR, McDonald WG, Waldon DJ, et al. Identification of a novel mitochondrial protein ("mitoNEET") cross-linked specifically by a thiazolidinedione photoprobe. Am J Physiol Endocrinol Metab 2004: 286: E252-260.

(6) Colca JR, McDonald WG, Cavey GS, et al. Identification of a mitochondrial target of thiazolidinedione insulin sensitizers (mTOT)--relationship to newly identified mitochondrial pyruvate carrier proteins. PLoS One 2013: 8: e61551.

(7) Wiley SE, Murphy AN, Ross SA, van der Geer P, Dixon JE. MitoNEET is an iron-containing outer mitochondrial membrane protein that regulates oxidative capacity. Proc Natl Acad Sci U S A 2007: 104: 5318-5323.

(8) Sohn YS, Tamir S, Song L, et al. NAF-1 and mitoNEET are central to human breast cancer proliferation by maintaining mitochondrial homeostasis and promoting tumor growth. Proc Natl Acad Sci U S A 2013: 110: 14676-14681.

(9) Kusminski CM, Holland WL, Sun K, et al. MitoNEET-driven alterations in adipocyte mitochondrial activity reveal a crucial adaptive process that preserves insulin sensitivity in obesity. Nat Med 2012: 18: 1539-1549.

(10) Kusminski CM, Chen S, Ye R, et al. MitoNEET-Parkin Effects in Pancreatic alpha- and beta-Cells, Cellular Survival and Intra-Insular Crosstalk. Diabetes 2016.

(11) Picard M, McManus MJ, Csordas G, et al. Trans-mitochondrial coordination of cristae at regulated membrane junctions. Nat Commun 2015: 6: 6259.

(12) Jendrach M, Mai S, Pohl S, Voth M, Bereiter-Hahn J. Short- and long-term alterations of mitochondrial morphology, dynamics and mtDNA after transient oxidative stress. Mitochondrion 2008: 8: 293-304.

(13) Iqbal S, Hood DA. Oxidative stress-induced mitochondrial fragmentation and movement in skeletal muscle myoblasts. Am J Physiol Cell Physiol 2014: 306: C1176-1183.

(14) Tondera D, Grandemange S, Jourdain A, et al. SLP-2 is required for stress-induced mitochondrial hyperfusion. EMBO J 2009: 28: 1589-1600.

(15) Cosson P, Marchetti A, Ravazzola M, Orci L. Mitofusin-2 independent juxtaposition of endoplasmic reticulum and mitochondria: an ultrastructural study. PLoS One 2012: 7: e46293.

(16) Conlan AR, Paddock ML, Axelrod HL, et al. The novel 2Fe-2S outer mitochondrial protein mitoNEET displays conformational flexibility in its N-terminal cytoplasmic tethering domain. Acta Crystallogr Sect F Struct Biol Cryst Commun 2009: 65: 654-659.

(17) Taminelli GL, Sotomayor V, Valdivieso AG, Teiber ML, Marin MC, Santa-Coloma TA. CISD1 codifies a mitochondrial protein upregulated by the CFTR channel. Biochem Biophys Res Commun 2008: 365: 856-862.

(18) Bai F, Morcos F, Sohn YS, et al. The Fe-S cluster-containing NEET proteins mitoNEET and NAF-1 as chemotherapeutic targets in breast cancer. Proc Natl Acad Sci U S A 2015: 112: 3698-3703.

(19) Geldenhuys WJ, Leeper TC, Carroll RT. mitoNEET as a novel drug target for mitochondrial dysfunction. Drug Discov Today 2014: 19: 1601-1606.
(20) Longo PA, Kavran JM, Kim MS, Leahy DJ. Transient mammalian cell transfection with polyethylenimine (PEI). Methods Enzymol 2013: 529: 227-240.
(21) Perrin J, Le Coadic M, Vernay A, et al. TM9 family proteins control surface targeting of glycine-rich transmembrane domains. J Cell Sci 2015: 128: 2269-2277.

Acknowledgments

We thank the Pôle Facultaire de Microscopie Electronique for preparation of EM sections, and Dr I. Nikonenko for acquisition of images with FIB/SEM microscope. We also thank Pr P. Maechler and L. Oberhauser for their help concerning the Seahorse assay. This work was supported by the Swiss National Science Foundation (grant 31003A_153326 to PC).

Author contributions

A.V., A.M., A.S., M.R., L.O. and P.C. designed and performed the experiments and analyzed the results.

A.V., L.O. and P.C. wrote the manuscript. All authors discussed the results, read and corrected the manuscript.

Competing financial interests

The authors declare no competing financial interests.

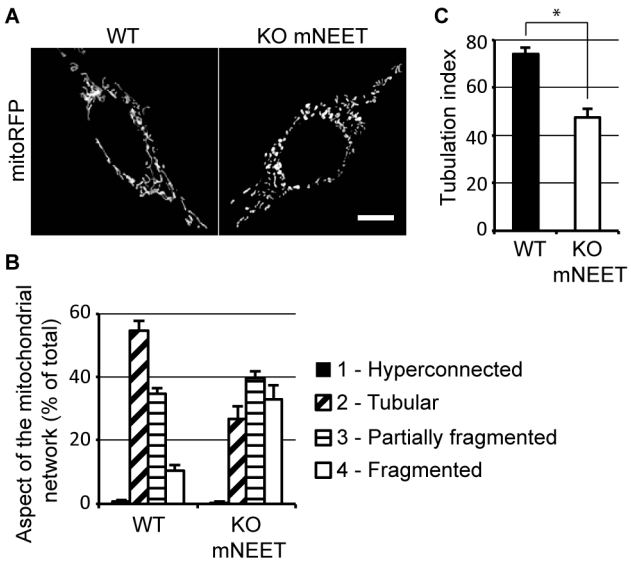

Figure 1: genetic inactivation of mNEET decreases connectivity of the mitochondrial network. (A) Mito-RFP was expressed in WT (left panel) or in *mNEET KO* MEFs cells (right panel) and cells were observed by fluorescence microscopy. Scale bar: 10µm. (B) For each cell, the connectivity of the mitochondrial network was graded 4 (totally fragmented), 3 (partially fragmented, 2 (tubular), or 1 (hyperconnected). The quantification was performed on blinded samples. The average and S.E.M. of 14 independent experiments are indicated. (C) The tubulation index was calculated as described in Materials and Methods. *: $p<0.001$ (student t-test).

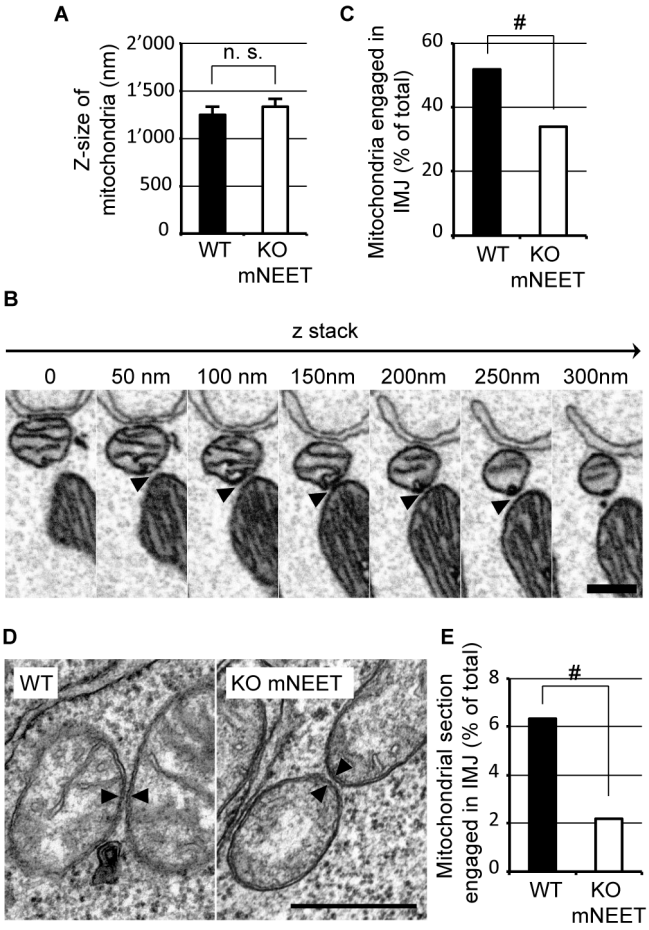

Figure 2: Inter-mitochondrial junctions are less abundant in *mNEET KO* cells than in WT cells. WT or *mNEET KO* cells were fixed, processed for electron microscopy and analyzed in a Helios Dualbeam scanning electron microscope to generate complete sets of images scanning the whole cell volume. Mitochondria from three independent experiments were analyzed for WT and for *mNEET KO* cells. (A) The size of mitochondria was evaluated by counting the number of sections through which individual mitochondria extended along the z axis (z-size) and was not significantly different in WT and *mNEET KO* cells. (B) A selection of serial pictures showing an inter-mitochondrial junction. Pictures were taken with a 10nm-interval, and one picture every 5 sections is shown. Arrowheads indicate regions of close contact between two adjacent mitochondria. The distance from each section to the first section shown is indicated. Scale bar: 50nm (C)

The percentage of mitochondria engaged in IMJs diminished significantly in *mNEET KO* cells compared to WT cells. #: p=0.012 Fisher's exact test. (D) WT or *mNEET KO* cells were fixed and sections were visualized in a transmission electron microscope. IMJs were defined as regions of close contact between two mitochondria (<20nm ; arrowheads). Scale bar: 500nm. (E) The frequency of IMJs was significantly decreased in *mNEET KO* cells compared to WT cells. #: p=0.0009 Fisher's exact test.

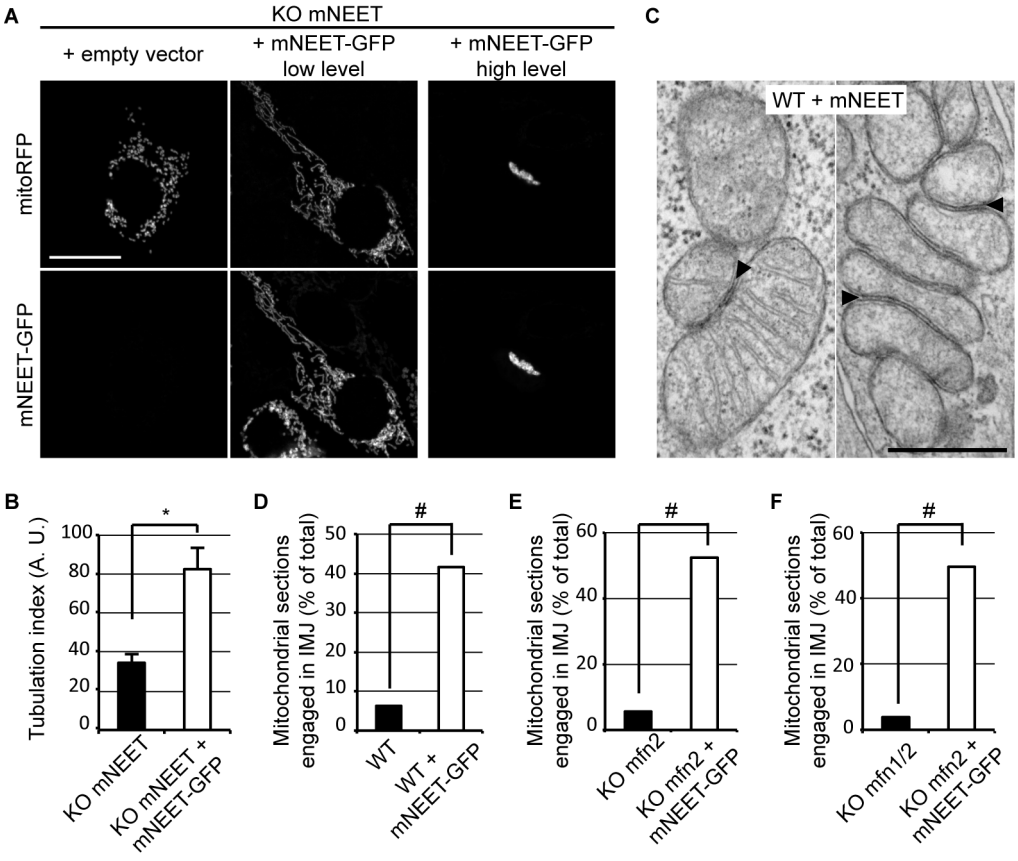

Figure 3: Overexpression of mNEET increases the connectivity of the mitochondrial network. (A) *mNEET KO* cells were co-transfected with plasmids expressing mitoRFP (upper panel) and mNEET-GFP (lower panel). As a control, an empty vector replaced the mNEET-GFP plasmid (left column). A mild expression of mNEET-GFP (middle column) restored the connectivity of the mitochondrial network whereas a high overexpression (right column) resulted in the collapse of the mitochondrial network. Scale bar: 10µm. (B) The connectivity of the mitochondrial network in cells expressing low levels of mNEET-GFP was determined as described in the Legend to Figure 1. The mean ± S.E.M of 7 independent experiments are presented. *: $p<0.001$ (student t-test). Expression of mNEET-GFP increased the connectivity of the mitochondrial network to a WT level. (C) To analyze the effect of mNEET overexpression on IMJs, WT MEF cells were transfected with a plasmid expressing mNEET-GFP. Cells expressing mNEET-GFP were sorted by

flow cytometry, then fixed and processed for electron microscopy one day later. Arrowheads indicate inter-mitochondrial contacts where electron-dense structures tethering apposed membranes are visible. Scale bar: 500nm. (D) The frequency of IMJs was quantified as described in the Legend to Figure 2. IMJs were more abundant in transfected cells than in WT cells. (E) The frequency of IMJs was determined in *mfn2 KO* cells and in *mfn2 KO* cells overexpressing mNEET-GFP as described above. (F) Similar experiments were performed using cells where both *mfn1* and *mfn2* were genetically inactivated. #: $p<0.001$ Fisher's exact test.

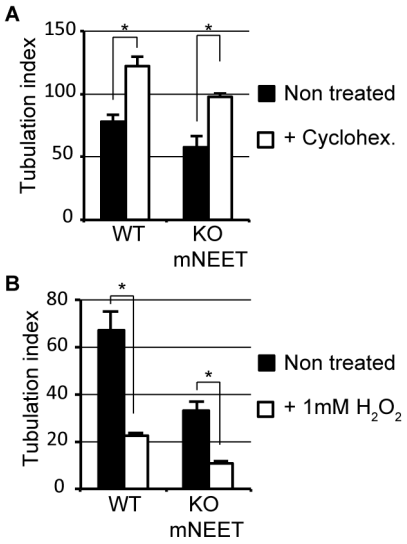

Figure 4: Regulation of mitochondrial fusion and fission is unaffected in *mNEET KO* cells. (A) WT or *mNEET KO* cells expressing mito-RFP were incubated for 6h in medium containing 25µg/mL cycloheximide or not. Cells were then fixed and the connectivity of the mitochondrial network was determined as described in the legend to Figure 1. The mean ± S.E.M. of 3 (WT) and 8 (*mNEET KO*) independent experiments is indicated. Mitochondrial fusion was stimulated by cycloheximide in both WT and *mNEET KO* cells. (B) WT and *mNEET KO* cells expressing mito-RFP were exposed to 1mM H_2O_2 for 1h. They were then fixed and examined. The mean ± S.E.M. of 6 (WT) and 3 (*mNEET KO*) independent experiments is indicated. Mitochondrial fragmentation upon exposure to H_2O_2 was observed in both WT and *mNEET KO* cells. *: $p<0.05$ (student t-test).

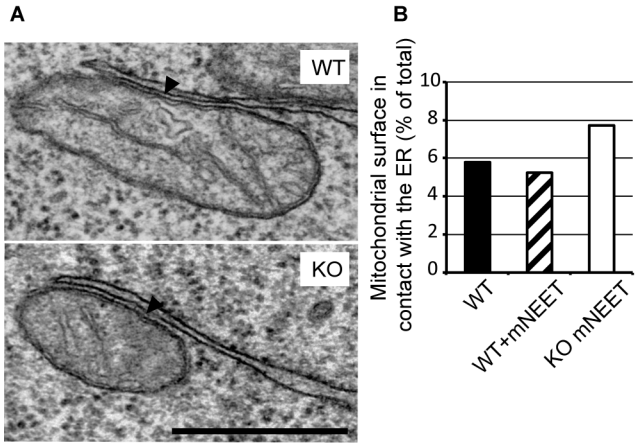

Figure 5: Establishment of ER-mitochondrial contact sites is independent on mNEET. (A) WT, *mNEET KO* or mNEET-overexpressing cells were fixed and processed for conventional electron microscopy. Sites of juxtaposition of ER and mitochondrial membranes were visualized (arrowheads), and quantified. (B) Approximately 6% of mitochondrial membrane was engaged into contacts with the ER in WT cells. This figure did not change significantly in cells overexpressing mNEET, and increased slightly in *mNEET KO* cells.

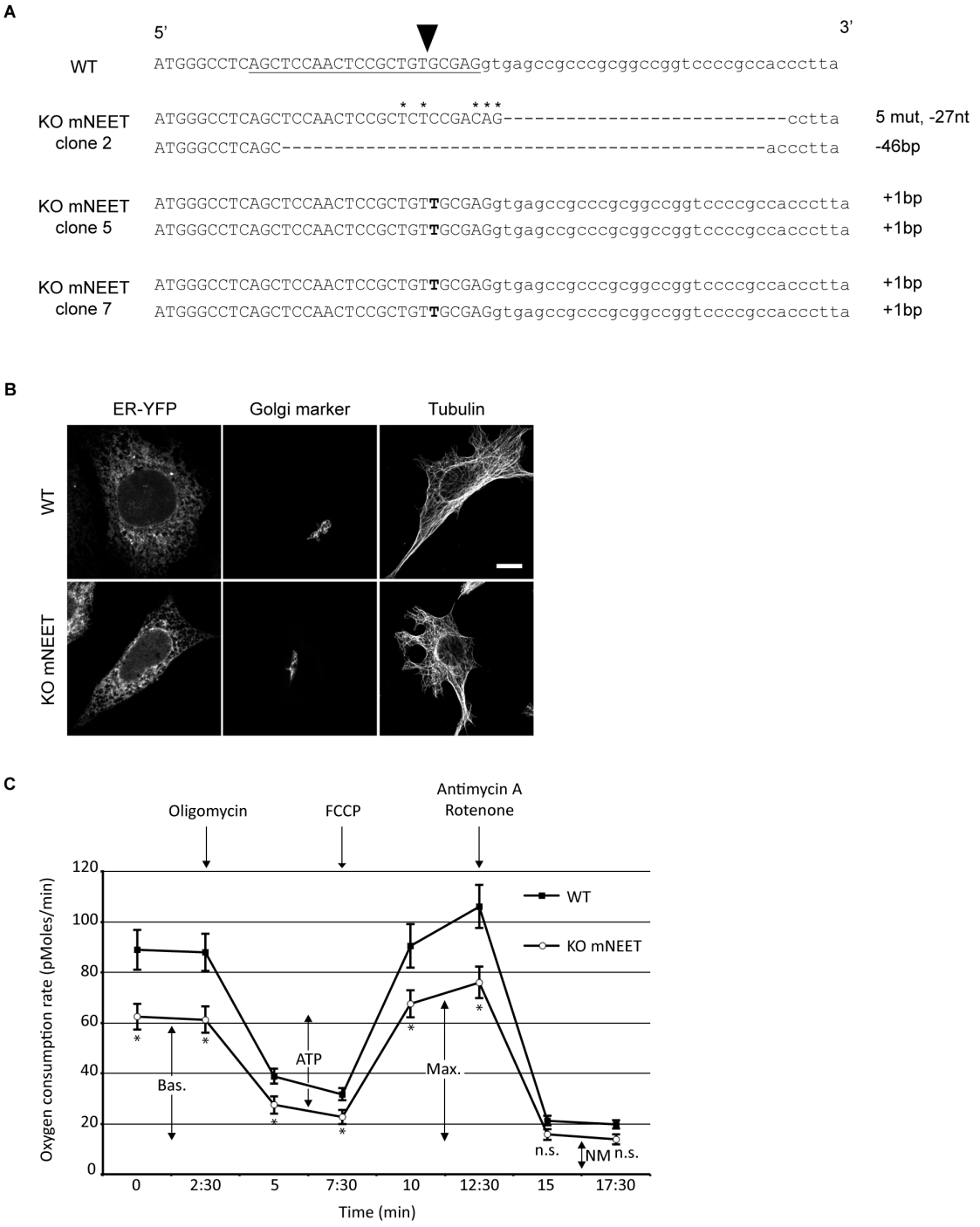

Figure S1: MEF *mNEET KO* cells. (A) Genomic sequence of the murine mNEET. The first exon is in upper case, and the following intron in lower case. The sequence targeted by the guide RNA is underlined and the cutting site of the nuclease is indicated with an arrow. The sequence of each allele is indicated for three independent mutant clones. (-) indicates a deletion, (*) a mutation and bold characters an insertion. (B) In order to

check the general organization of WT cells (upper panel) and mNEET KO cells (lower panel), both cell lines were transfected with a marker of the ER (left panel, YFP-KDEL), or of the Golgi complex (middle panel, B4GALT1-GFP) or the tubulin was stained with an anti-tubulin antibody, then revealed by an Alexa Fluor 488-coupled secondary antibody. Scale bar: 10µm. (C) Mitochondrial respiration of WT or *mNEET KO* cells was assessed. The basal respiration (Bas.) was first calculated as the oxygen consumption at steady-state. The addition of 1µM oligomycin allows the identification of the oxygen consumption associated with the production of ATP (ATP). The addition of 300nm FCCP reveals the maximal amount of oxygen consumed by the mitochondria (Max.). Finally, the addition of 0.5µM antimycin A and rotenone enables the calculation of the non-mitochondrial respiration (NM). In each well the amount of protein was determined, and the oxygen consumption corrected accordingly. *: $p<0.05$ (Mann-Whitney test, n: 25 independent samples in 3 independent experiments).

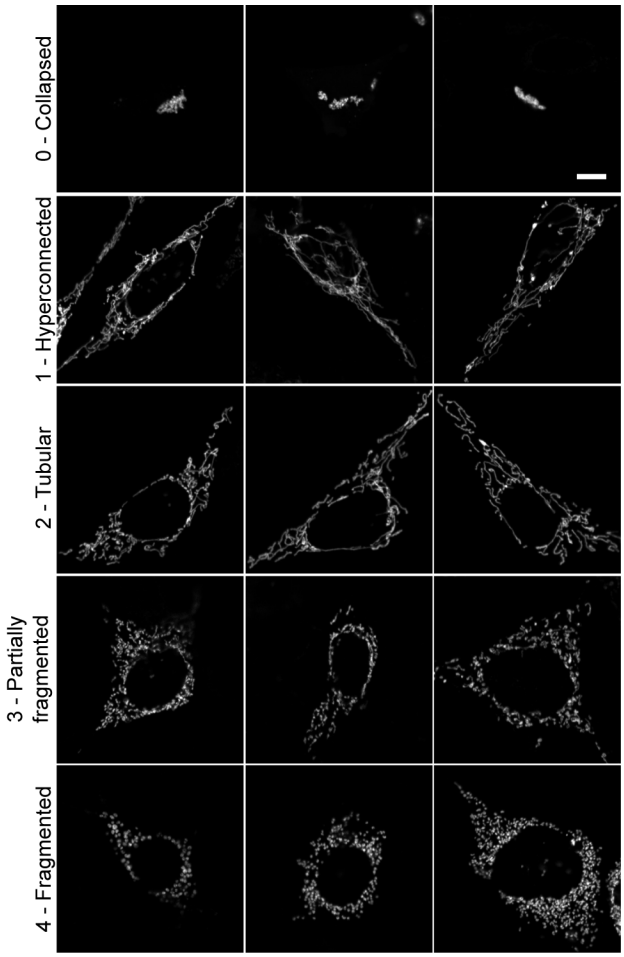

Figure S2: Evaluation of mitochondrial network connectivity. Cells expressing mito-RFP were examined individually in blinded samples. The mitochondrial network of each cells was scored as hyperconnected (score 1) when mitochondria appeared very long and highly connected; tubular (score 2) when mitochondria were mostly elongated and only a few appeared shorter ; partially fragmented (score 3) when most of the mitochondria were short and only a few of them tubular ; totally fragmented (score 4) when all mitochondria were short and round. Scale bar: 10µm. Collapsed mitochondria (score 0) were only observed in cells overexpressing mNEET-GFP. Three representative pictures corresponding to these different morphologies are shown.

Table S1 : Analysis of mitochondrial size and IMJs in electron microscopy 3-dimensional reconstructions

	Mitochondria			Inter-mito contacts			
	Nb	total	individual	Nb	total size	frequency	% surface
	(n)	length (z-µm)	length (z-µm)	(n)	(z-µm)	(n/µm)	(%)
WT1	37	50.95	1.38	25	0.300	0.49	0.59
WT2	39	57.83	1.48	35	0.502	0.61	0.87
WT3	30	23.64	0.79	19	0.254	0.80	1.07
Total WT	106	132.42	1.25	79	1056	0.60	0.80
KO1	33	42.20	1.28	11	0.183	0.26	0.43
KO2	31	53.35	1.72	19	0.233	0.36	0.44
KO3	39	41.68	1.07	22	0.297	0.53	0.71
Total KO mNEET	103	137.23	1.33	52	0.713	0.38	0.52

Table S1: Analysis of mitochondrial size and IMJs in electron microscopy 3-dimensional reconstructions

Table S2 : Frequency of IMJs analyzed by conventional electron microscopy in thin sections

	Total Mitos	Engaged in IMCs		Mito surface	IMC surface	
	(n)	(n)	%	(nm)	(nm)	(% total)
WT1	80	7	8.8	56861	287	0.50
WT2	126	4	3.2	77619	248	0.32
WT3	169	12	7.1	128911	1592	1.23
WT4	78	2	2.6	65066	314	0.48
WT5	131	12	9.2	130231	1246	0.96
Total WT	584	37	6.34%	458688	3687	0.80%
KO1	34	0	0	28254	0	0
KO2	169	5	3.0	112716	610	0.54
KO3	85	0	0	76043	0	0
KO4	78	2	2.6	63816	117	0.18
KO5	137	4	2.9	107882	796	0.74
Total KO mNEET	503	11	2.19%	388711	1523	0.39%

Table S2: Frequency of IMJs analyzed by conventional electron microscopy in thin sections

Table S3 : Effect of overexpression of mNEET-GFP on IMJs

	Total Mitos	Engaged in IMCs		Mito surface	IMC surface	
	(n)	(n)	%	(nm)	(nm)	(% total)
WT	584	37	6.34%	458688	3687	0.80%
WT+mNEET	267	111	41.6%	165448	21290	12.9%
mfn2KO	178	10	5.62%	164720	898	0.55%
mfn2KO+mNEET	189	99	52.4%	132868	12382	9.32%
mfn1/2KO	382	15	3.93%	527130	2328	0.44%
mfn1/2KO+mNEET	274	136	49.6%	312876	56605	18.1%

Table S3: Effect of overexpression of mNEET-GFP on IMJs

Table S4 : Effect of mNEET genetic inactivation or overexpression on the formation of ER-mitochondria contact sites

	Mito surface (nm)	ER contacts surface (nm)	(% mito surf)
WT1	56861	4721	8.3
WT2	77619	4809	6.2
WT3	128911	7566	5.9
WT4	65066	3169	4.9
WT5	130231	6287	4.8
WT Total	458688	26552	5.79%
+mNEET1	95618	5331	5.6
+mNEET2	69830	3345	4.8
+mNEET Total	165448	8676	5.24%
KO1	28254	4267	15.1
KO2	112716	9318	8.3
KO3	76043	3180	4.2
KO4	63816	4277	6.7
KO5	107882	8969	8.3
KO Total	388711	30011	7.72%

Table S4: Effect of mNEET genetic inactivation or overexpression on the formation of ER-mitochondria contact sites

V. REFERENCES

V. REFERENCES

Adams, C. M., Goldstein, J. L., & Brown, M. S. (2003). Cholesterol-induced conformational change in SCAP enhanced by Insig proteins and mimicked by cationic amphiphiles. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(19), 10647–10652. https://doi.org/10.1073/pnas.1534833100

Adams, C. M., Reitz, J., De Brabander, J. K., Feramisco, J. D., Li, L., Brown, M. S., & Goldstein, J. L. (2004). Cholesterol and 25-hydroxycholesterol inhibit activation of SREBPs by different mechanisms, both involving SCAP and Insigs. *The Journal of Biological Chemistry*, *279*(50), 52772–52780. https://doi.org/10.1074/jbc.M410302200

Adolf, F., Herrmann, A., Hellwig, A., Beck, R., Brügger, B., & Wieland, F. T. (2013). Scission of COPI and COPII vesicles is independent of GTP hydrolysis. *Traffic (Copenhagen, Denmark)*, *14*(8), 922–932. https://doi.org/10.1111/tra.12084

Anantharaman, V., & Aravind, L. (2002). The GOLD domain, a novel protein module involved in Golgi function and secretion. *Genome Biology*, *3*(5), research0023.

Andrei, C., Dazzi, C., Lotti, L., Torrisi, M. R., Chimini, G., & Rubartelli, A. (1999). The secretory route of the leaderless protein interleukin 1beta involves exocytosis of endolysosome-related vesicles. *Molecular Biology of the Cell*, *10*(5), 1463–1475.

Andrei, C., Margiocco, P., Poggi, A., Lotti, L. V., Torrisi, M. R., & Rubartelli, A. (2004). Phospholipases C and A2 control lysosome-mediated IL-1 beta secretion: Implications for inflammatory processes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(26), 9745–9750. https://doi.org/10.1073/pnas.0308558101

Anelli, T., Alessio, M., Bachi, A., Bergamelli, L., Bertoli, G., Camerini, S., … Sitia, R. (2003). Thiol-mediated protein retention in the endoplasmic reticulum: the role of ERp44. *The EMBO Journal*, *22*(19), 5015–5022. https://doi.org/10.1093/emboj/cdg491

Anelli, T., Ceppi, S., Bergamelli, L., Cortini, M., Masciarelli, S., Valetti, C., & Sitia, R. (2007a). Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. *The EMBO Journal*, *26*(19), 4177–4188.

https://doi.org/10.1038/sj.emboj.7601844

Anelli, T., Ceppi, S., Bergamelli, L., Cortini, M., Masciarelli, S., Valetti, C., & Sitia, R. (2007b). Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. *The EMBO Journal*, *26*(19), 4177–4188. https://doi.org/10.1038/sj.emboj.7601844

Anelli, T., & van Anken, E. (2013). Missing links in antibody assembly control. *International Journal of Cell Biology*, *2013*, 606703. https://doi.org/10.1155/2013/606703

Aoe, T., Cukierman, E., Lee, A., Cassel, D., Peters, P. J., & Hsu, V. W. (1997). The KDEL receptor, ERD2, regulates intracellular traffic by recruiting a GTPase-activating protein for ARF1. *The EMBO Journal*, *16*(24), 7305–7316. https://doi.org/10.1093/emboj/16.24.7305

Aoe, T., Lee, A. J., van Donselaar, E., Peters, P. J., & Hsu, V. W. (1998). Modulation of intracellular transport by transported proteins: insight from regulation of COPI-mediated transport. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(4), 1624–1629.

Appenzeller, C., Andersson, H., Kappeler, F., & Hauri, H. P. (1999). The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nature Cell Biology*, *1*(6), 330–334. https://doi.org/10.1038/14020

Appenzeller-Herzog, C. (2006). The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. *Journal of Cell Science*, *119*(11), 2173–2183. https://doi.org/10.1242/jcs.03019

Appenzeller-Herzog, C., Nyfeler, B., Burkhard, P., Santamaria, I., Lopez-Otin, C., & Hauri, H.-P. (2005). Carbohydrate- and conformation-dependent cargo capture for ER-exit. *Molecular Biology of the Cell*, *16*(3), 1258–1267. https://doi.org/10.1091/mbc.E04-08-0708

Appenzeller-Herzog, C., Roche, A.-C., Nufer, O., & Hauri, H.-P. (2004). pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. *The Journal of Biological Chemistry*, *279*(13), 12943–12950. https://doi.org/10.1074/jbc.M313245200

Aridor, M., Bannykh, S. I., Rowe, T., & Balch, W. E. (1999). Cargo can modulate COPII vesicle formation from the endoplasmic reticulum. *The Journal of Biological Chemistry*, *274*(7), 4389–4399.

Au, C. E., Hermo, L., Byrne, E., Smirle, J., Fazel, A., Simon, P. H. G., … Bergeron, J. J. M. (2015). Expression, sorting, and segregation of Golgi proteins during germ cell differentiation in the testis. *Molecular Biology of the Cell*, *26*(22), 4015–4032. https://doi.org/10.1091/mbc.E14-12-1632

Baldauf, S. L., & Doolittle, W. F. (1997). Origin and evolution of the slime molds (Mycetozoa). *Proceedings of the National Academy of Sciences of the United States of America*, *94*(22), 12007–12012.

Barlowe, C. (2003). Molecular recognition of cargo by the COPII complex: a most accommodating coat. *Cell*, *114*(4), 395–397.

Barlowe, C., & Helenius, A. (2016). Cargo Capture and Bulk Flow in the Early Secretory Pathway. *Annual Review of Cell and Developmental Biology*, *32*, 197–222. https://doi.org/10.1146/annurev-cellbio-111315-125016

Barlowe, C., Orci, L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., … Schekman, R. (1994). COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell*, *77*(6), 895–907.

Beck, R., Adolf, F., Weimer, C., Bruegger, B., & Wieland, F. T. (2009). ArfGAP1 activity and COPI vesicle biogenesis. *Traffic (Copenhagen, Denmark)*, *10*(3), 307–315. https://doi.org/10.1111/j.1600-0854.2008.00865.x

Behnke, J., Feige, M. J., & Hendershot, L. M. (2015). BiP and its nucleotide exchange factors Grp170 and Sil1: mechanisms of action and biological functions. *Journal of Molecular Biology*, *427*(7), 1589–1608. https://doi.org/10.1016/j.jmb.2015.02.011

Belden, W. J., & Barlowe, C. (1996). Erv25p, a component of COPII-coated vesicles, forms a complex with Emp24p that is required for efficient endoplasmic reticulum to Golgi transport. *The Journal of Biological Chemistry*, *271*(43), 26939–26946.

Belden, W. J., & Barlowe, C. (2001a). Distinct roles for the cytoplasmic tail sequences of Emp24p and Erv25p in transport between the endoplasmic reticulum and Golgi complex. *The Journal of Biological Chemistry*, *276*(46), 43040–43048. https://doi.org/10.1074/jbc.M108113200

Belden, W. J., & Barlowe, C. (2001b). Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science (New York, N.Y.)*, *294*(5546), 1528–1531. https://doi.org/10.1126/science.1065224

Benghezal, M., Cornillon, S., Gebbie, L., Alibaud, L., Brückert, F., Letourneur, F., & Cosson,

P. (2003). Synergistic control of cellular adhesion by transmembrane 9 proteins. *Molecular Biology of the Cell*, *14*(7), 2890–2899. https://doi.org/10.1091/mbc.E02-11-0724

Benghezal, M., Fauvarque, M.-O., Tournebize, R., Froquet, R., Marchetti, A., Bergeret, E., … Cosson, P. (2006). Specific host genes required for the killing of Klebsiella bacteria by phagocytes. *Cellular Microbiology*, *8*(1), 139–148. https://doi.org/10.1111/j.1462-5822.2005.00607.x

Bentivoglio, M. (1998). 1898: the Golgi apparatus emerges from nerve cells. *Trends in Neurosciences*, *21*(5), 195–200.

Bi, X., Corpina, R. A., & Goldberg, J. (2002). Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature*, *419*(6904), 271–277. https://doi.org/10.1038/nature01040

Bi, X., Mancias, J. D., & Goldberg, J. (2007). Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. *Developmental Cell*, *13*(5), 635–645. https://doi.org/10.1016/j.devcel.2007.10.006

Blum, R., Feick, P., Puype, M., Vandekerckhove, J., Klengel, R., Nastainczyk, W., & Schulz, I. (1996). Tmp21 and p24A, two type I proteins enriched in pancreatic microsomal membranes, are members of a protein family involved in vesicular trafficking. *The Journal of Biological Chemistry*, *271*(29), 17183–17189.

Boehm, J., Letourneur, F., Ballensiefen, W., Ossipov, D., Démollière, C., & Schmitt, H. D. (1997). Sec12p requires Rer1p for sorting to coatomer (COPI)-coated vesicles and retrieval to the ER. *Journal of Cell Science*, *110 (Pt 8)*, 991–1003.

Bonifacino, J. S., Cosson, P., & Klausner, R. D. (1990). Colocalized transmembrane determinants for ER degradation and subunit assembly explain the intracellular fate of TCR chains. *Cell*, *63*(3), 503–513.

Bonifacino, J. S., Cosson, P., Shah, N., & Klausner, R. D. (1991). Role of potentially charged transmembrane residues in targeting proteins for retention and degradation within the endoplasmic reticulum. *The EMBO Journal*, *10*(10), 2783–2793.

Bonifacino, J. S., & Glick, B. S. (2004). The mechanisms of vesicle budding and fusion. *Cell*, *116*(2), 153–166.

Bowie, J. U. (1997). Helix packing in membrane proteins. *Journal of Molecular Biology*, *272*(5), 780–789. https://doi.org/10.1006/jmbi.1997.1279

Brown, A. J., Sun, L., Feramisco, J. D., Brown, M. S., & Goldstein, J. L. (2002). Cholesterol addition to ER membranes alters conformation of SCAP, the SREBP escort protein that regulates cholesterol metabolism. *Molecular Cell*, *10*(2), 237–245.

Brown, M. S., & Goldstein, J. L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell*, *89*(3), 331–340.

Caramelo, J. J., & Parodi, A. J. (2008). Getting in and out from calnexin/calreticulin cycles. *The Journal of Biological Chemistry*, *283*(16), 10221–10225. https://doi.org/10.1074/jbc.R700048200

Chiu, V. K., Bivona, T., Hach, A., Sajous, J. B., Silletti, J., Wiener, H., … Philips, M. R. (2002). Ras signalling on the endoplasmic reticulum and the Golgi. *Nature Cell Biology*, *4*(5), 343–350. https://doi.org/10.1038/ncb783

Ciufo, L. F., & Boyd, A. (2000). Identification of a lumenal sequence specifying the assembly of Emp24p into p24 complexes in the yeast secretory pathway. *The Journal of Biological Chemistry*, *275*(12), 8382–8388.

Cornillon, S., Pech, E., Benghezal, M., Ravanel, K., Gaynor, E., Letourneur, F., … Cosson, P. (2000). Phg1p is a nine-transmembrane protein superfamily member involved in dictyostelium adhesion and phagocytosis. *The Journal of Biological Chemistry*, *275*(44), 34287–34292. https://doi.org/10.1074/jbc.M006725200

Cosson, P., Amherdt, M., Rothman, J. E., & Orci, L. (2002). A resident Golgi protein is excluded from peri-Golgi vesicles in NRK cells. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(20), 12831–12834. https://doi.org/10.1073/pnas.192460999

Cosson, P., Lefkir, Y., Démollière, C., & Letourneur, F. (1998). New COP1-binding motifs involved in ER retrieval. *The EMBO Journal*, *17*(23), 6863–6870. https://doi.org/10.1093/emboj/17.23.6863

Cosson, P., & Letourneur, F. (1994). Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science (New York, N.Y.)*, *263*(5153), 1629–1631.

D'Angelo, G., Vicinanza, M., Di Campli, A., & De Matteis, M. A. (2008). The multiple roles of PtdIns(4)P -- not just the precursor of PtdIns(4,5)P2. *Journal of Cell Science*, *121*(Pt 12), 1955–1963. https://doi.org/10.1242/jcs.023630

De Matteis, M. A., & Luini, A. (2008). Exiting the Golgi complex. *Nature Reviews. Molecular Cell Biology*, *9*(4), 273–284. https://doi.org/10.1038/nrm2378

Deshaies, R. J., & Schekman, R. (1987). A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum. *The Journal of Cell Biology*, *105*(2), 633–645.

Dolman, N. J., & Tepikin, A. V. (2006). Calcium gradients and the Golgi. *Cell Calcium*, *40*(5–6), 505–512. https://doi.org/10.1016/j.ceca.2006.08.012

Dominguez, M., Dejgaard, K., Füllekrug, J., Dahan, S., Fazel, A., Paccaud, J. P., … Nilsson, T. (1998). gp25L/emp24/p24 protein family members of the cis-Golgi network bind both COP I and II coatomer. *The Journal of Cell Biology*, *140*(4), 751–765.

Duden, R. (2003). ER-to-Golgi transport: COP I and COP II function (Review). *Molecular Membrane Biology*, *20*(3), 197–207. https://doi.org/10.1080/0968768031000122548

Eichinger, L., Pachebat, J. A., Glöckner, G., Rajandream, M.-A., Sucgang, R., Berriman, M., … Kuspa, A. (2005). The genome of the social amoeba Dictyostelium discoideum. *Nature*, *435*(7038), 43–57. https://doi.org/10.1038/nature03481

Ellgaard, L., & Helenius, A. (2003). Quality control in the endoplasmic reticulum. *Nature Reviews. Molecular Cell Biology*, *4*(3), 181–191. https://doi.org/10.1038/nrm1052

Emery, G., Rojo, M., & Gruenberg, J. (2000). Coupled transport of p24 family members. *Journal of Cell Science*, *113 (Pt 13)*, 2507–2516.

Farhan, H., Weiss, M., Tani, K., Kaufman, R. J., & Hauri, H.-P. (2008). Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. *The EMBO Journal*, *27*(15), 2043–2054. https://doi.org/10.1038/emboj.2008.136

Fasshauer, D., Sutton, R. B., Brunger, A. T., & Jahn, R. (1998). Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(26), 15781–15786.

Fenteany, F. H., & Colley, K. J. (2005). Multiple signals are required for alpha2,6-sialyltransferase (ST6Gal I) oligomerization and Golgi localization. *The Journal of Biological Chemistry*, *280*(7), 5423–5429. https://doi.org/10.1074/jbc.M412396200

Fink, A., Sal-Man, N., Gerber, D., & Shai, Y. (2012). Transmembrane domains interactions within the membrane milieu: principles, advances and challenges. *Biochimica Et Biophysica Acta*, *1818*(4), 974–983.

https://doi.org/10.1016/j.bbamem.2011.11.029

Foley, D. A., Sharpe, H. J., & Otte, S. (2007). Membrane topology of the endoplasmic reticulum to Golgi transport factor Erv29p. *Molecular Membrane Biology*, *24*(4), 259–268. https://doi.org/10.1080/09687860601178518

Fossati, M., Colombo, S. F., & Borgese, N. (2014). A positive signal prevents secretory membrane cargo from recycling between the Golgi and the ER. *The EMBO Journal*, *33*(18), 2080–2097. https://doi.org/10.15252/embj.201488367

Froquet, R., le Coadic, M., Perrin, J., Cherix, N., Cornillon, S., & Cosson, P. (2012). TM9/Phg1 and SadA proteins control surface expression and stability of SibA adhesion molecules in Dictyostelium. *Molecular Biology of the Cell*, *23*(4), 679–686. https://doi.org/10.1091/mbc.E11-04-0338

Füllekrug, J., Suganuma, T., Tang, B. L., Hong, W., Storrie, B., & Nilsson, T. (1999). Localization and recycling of gp27 (hp24gamma3): complex formation with other p24 family members. *Molecular Biology of the Cell*, *10*(6), 1939–1955.

Gaynor, E. C., Graham, T. R., & Emr, S. D. (1998). COPI in ER/Golgi and intra-Golgi transport: do yeast COPI mutants point the way? *Biochimica Et Biophysica Acta*, *1404*(1–2), 33–51.

Gee, H. Y., Noh, S. H., Tang, B. L., Kim, K. H., & Lee, M. G. (2011). Rescue of ΔF508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell*, *146*(5), 746–760. https://doi.org/10.1016/j.cell.2011.07.021

Gimpelev, M., Forrest, L. R., Murray, D., & Honig, B. (2004). Helical packing patterns in membrane and soluble proteins. *Biophysical Journal*, *87*(6), 4075–4086. https://doi.org/10.1529/biophysj.104.049288

Giraudo, C. G., Daniotti, J. L., & Maccioni, H. J. (2001). Physical and functional association of glycolipid N-acetyl-galactosaminyl and galactosyl transferases in the Golgi apparatus. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 1625–1630. https://doi.org/10.1073/pnas.031458398

Girod, A., Storrie, B., Simpson, J. C., Johannes, L., Goud, B., Roberts, L. M., … Pepperkok, R. (1999). Evidence for a COP-I-independent transport route from the Golgi complex to the endoplasmic reticulum. *Nature Cell Biology*, *1*(7), 423–430. https://doi.org/10.1038/15658

Glick, B. S., & Luini, A. (2011). Models for Golgi traffic: a critical assessment. *Cold Spring Harbor Perspectives in Biology*, *3*(11), a005215.

https://doi.org/10.1101/cshperspect.a005215

Glick, B. S., & Nakano, A. (2009). Membrane traffic within the Golgi apparatus. *Annual Review of Cell and Developmental Biology*, *25*, 113–132. https://doi.org/10.1146/annurev.cellbio.24.110707.175421

Goder, V., Junne, T., & Spiess, M. (2004). Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Molecular Biology of the Cell*, *15*(3), 1470–1478. https://doi.org/10.1091/mbc.E03-08-0599

Graifer, D., & Karpova, G. (2015). Roles of ribosomal proteins in the functioning of translational machinery of eukaryotes. *Biochimie*, *109*, 1–17. https://doi.org/10.1016/j.biochi.2014.11.016

Hamman, B. D., Hendershot, L. M., & Johnson, A. E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. *Cell*, *92*(6), 747–758.

Hammond, A. T., & Glick, B. S. (2000). Dynamics of transitional endoplasmic reticulum sites in vertebrate cells. *Molecular Biology of the Cell*, *11*(9), 3013–3030.

Hardwick, K. G., Boothroyd, J. C., Rudner, A. D., & Pelham, H. R. (1992). Genes that allow yeast cells to grow in the absence of the HDEL receptor. *The EMBO Journal*, *11*(11), 4187–4195.

He, P., Peng, Z., Luo, Y., Wang, L., Yu, P., Deng, W., … Ma, D. (2009). High-throughput functional screening for autophagy-related genes and identification of TM9SF1 as an autophagosome-inducing gene. *Autophagy*, *5*(1), 52–60.

Herzig, Y., Sharpe, H. J., Elbaz, Y., Munro, S., & Schuldiner, M. (2012). A systematic approach to pair secretory cargo receptors with their cargo suggests a mechanism for cargo selection by Erv14. *PLoS Biology*, *10*(5), e1001329. https://doi.org/10.1371/journal.pbio.1001329

Holtzman, J. L. (1997). The roles of the thiol:protein disulfide oxidoreductases in membrane and secretory protein synthesis within the lumen of the endoplasmic reticulum. *Journal of Investigative Medicine: The Official Publication of the American Federation for Clinical Research*, *45*(2), 28–34.

Horton, J. D. (2002). Sterol regulatory element-binding proteins: transcriptional activators of lipid synthesis. *Biochemical Society Transactions*, *30*(Pt 6), 1091–1095. https://doi.org/10.1042/

Jackson, C. L. (2009). Mechanisms of transport through the Golgi complex. *Journal of Cell*

Science, *122*(Pt 4), 443–452. https://doi.org/10.1242/jcs.032581

Jackson, L. P., Lewis, M., Kent, H. M., Edeling, M. A., Evans, P. R., Duden, R., & Owen, D. J. (2012). Molecular basis for recognition of dilysine trafficking motifs by COPI. *Developmental Cell*, *23*(6), 1255–1262. https://doi.org/10.1016/j.devcel.2012.10.017

Jackson, M. R., Nilsson, T., & Peterson, P. A. (1990). Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *The EMBO Journal*, *9*(10), 3153–3162.

Jahn, R., & Scheller, R. H. (2006). SNAREs--engines for membrane fusion. *Nature Reviews. Molecular Cell Biology*, *7*(9), 631–643. https://doi.org/10.1038/nrm2002

Johnson, A. E., & van Waes, M. A. (1999). The translocon: a dynamic gateway at the ER membrane. *Annual Review of Cell and Developmental Biology*, *15*, 799–842. https://doi.org/10.1146/annurev.cellbio.15.1.799

Kaiser, H.-J., Orłowski, A., Róg, T., Nyholm, T. K. M., Chai, W., Feizi, T., … Simons, K. (2011). Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(40), 16628–16633. https://doi.org/10.1073/pnas.1103742108

Karamyshev, A. L., Kelleher, D. J., Gilmore, R., Johnson, A. E., von Heijne, G., & Nilsson, I. (2005). Mapping the interaction of the STT3 subunit of the oligosaccharyl transferase complex with nascent polypeptide chains. *The Journal of Biological Chemistry*, *280*(49), 40489–40493. https://doi.org/10.1074/jbc.M509168200

Karsten, V., Hegde, R. S., Sinai, A. P., Yang, M., & Joiner, K. A. (2004). Transmembrane domain modulates sorting of membrane proteins in Toxoplasma gondii. *The Journal of Biological Chemistry*, *279*(25), 26052–26057. https://doi.org/10.1074/jbc.M400480200

Kienlen-Campard, P., Tasiaux, B., Van Hees, J., Li, M., Huysseune, S., Sato, T., … Octave, J.-N. (2008). Amyloidogenic processing but not amyloid precursor protein (APP) intracellular C-terminal domain production requires a precisely oriented APP dimer assembled by transmembrane GXXXG motifs. *The Journal of Biological Chemistry*, *283*(12), 7733–7744. https://doi.org/10.1074/jbc.M707142200

Kim, C., Lau, T.-L., Ulmer, T. S., & Ginsberg, M. H. (2009). Interactions of platelet integrin alphaIIb and beta3 transmembrane domains in mammalian cell membranes and

their role in integrin activation. *Blood*, *113*(19), 4747–4753. https://doi.org/10.1182/blood-2008-10-186551

Kim, E. C., Meng, H., & Jun, A. S. (2013). Lithium treatment increases endothelial cell survival and autophagy in a mouse model of Fuchs endothelial corneal dystrophy. *The British Journal of Ophthalmology*, *97*(8), 1068–1073. https://doi.org/10.1136/bjophthalmol-2012-302881

Kirchhausen, T. (2000). Three ways to make a vesicle. *Nature Reviews. Molecular Cell Biology*, *1*(3), 187–198. https://doi.org/10.1038/35043117

Kleene, R., & Berger, E. G. (1993). The molecular and cell biology of glycosyltransferases. *Biochimica Et Biophysica Acta*, *1154*(3–4), 283–325.

Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., … Aebi, M. (2006). N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science (New York, N.Y.)*, *314*(5802), 1148–1150. https://doi.org/10.1126/science.1134351

Langhans, M., Marcote, M. J., Pimpl, P., Virgili-López, G., Robinson, D. G., & Aniento, F. (2008). In vivo trafficking and localization of p24 proteins in plant cells. *Traffic (Copenhagen, Denmark)*, *9*(5), 770–785. https://doi.org/10.1111/j.1600-0854.2008.00719.x

Lankford, S. P., Cosson, P., Bonifacino, J. S., & Klausner, R. D. (1993). Transmembrane domain length affects charge-mediated retention and degradation of proteins within the endoplasmic reticulum. *The Journal of Biological Chemistry*, *268*(7), 4814–4820.

Le Coadic, M., Froquet, R., Lima, W. C., Dias, M., Marchetti, A., & Cosson, P. (2013). Phg1/TM9 proteins control intracellular killing of bacteria by determining cellular levels of the Kil1 sulfotransferase in Dictyostelium. *PloS One*, *8*(1), e53259. https://doi.org/10.1371/journal.pone.0053259

Lee, M. C. S., Orci, L., Hamamoto, S., Futai, E., Ravazzola, M., & Schekman, R. (2005). Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. *Cell*, *122*(4), 605–617. https://doi.org/10.1016/j.cell.2005.07.025

Letourneur, F., Gaynor, E. C., Hennecke, S., Démollière, C., Duden, R., Emr, S. D., … Cosson, P. (1994). Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell*, *79*(7), 1199–1207.

Levine, T. P., Wiggins, C. A., & Munro, S. (2000). Inositol phosphorylceramide synthase is

located in the Golgi apparatus of Saccharomyces cerevisiae. *Molecular Biology of the Cell*, *11*(7), 2267–2281.

Lewis, M. J., & Pelham, H. R. (1990). A human homologue of the yeast HDEL receptor. *Nature*, *348*(6297), 162–163. https://doi.org/10.1038/348162a0

Lewis, M. J., & Pelham, H. R. (1992). Sequence of a second human KDEL receptor. *Journal of Molecular Biology*, *226*(4), 913–916.

Lin, R. C., & Scheller, R. H. (1997). Structural organization of the synaptic exocytosis core complex. *Neuron*, *19*(5), 1087–1094.

Losev, E., Reinke, C. A., Jellen, J., Strongin, D. E., Bevis, B. J., & Glick, B. S. (2006). Golgi maturation visualized in living yeast. *Nature*, *441*(7096), 1002–1006. https://doi.org/10.1038/nature04717

Lozupone, F., Borghi, M., Marzoli, F., Azzarito, T., Matarrese, P., Iessi, E., … Fais, S. (2015). TM9SF4 is a novel V-ATPase-interacting protein that modulates tumor pH alterations associated with drug resistance and invasiveness of colon cancer cells. *Oncogene*, *34*(40), 5163–5174. https://doi.org/10.1038/onc.2014.437

Ma, W., & Goldberg, J. (2013). Rules for the recognition of dilysine retrieval motifs by coatomer. *The EMBO Journal*, *32*(7), 926–937. https://doi.org/10.1038/emboj.2013.41

Machamer, C. E. (1993). Targeting and retention of Golgi membrane proteins. *Current Opinion in Cell Biology*, *5*(4), 606–612.

MacKenzie, A., Wilson, H. L., Kiss-Toth, E., Dower, S. K., North, R. A., & Surprenant, A. (2001). Rapid secretion of interleukin-1beta by microvesicle shedding. *Immunity*, *15*(5), 825–835.

Malkus, P., Jiang, F., & Schekman, R. (2002). Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. *The Journal of Cell Biology*, *159*(6), 915–921. https://doi.org/10.1083/jcb.200208074

Marks, M. S., Roche, P. A., van Donselaar, E., Woodruff, L., Peters, P. J., & Bonifacino, J. S. (1995). A lysosomal targeting signal in the cytoplasmic tail of the beta chain directs HLA-DM to MHC class II compartments. *The Journal of Cell Biology*, *131*(2), 351–369.

Martínez-Menárguez, J. A., Geuze, H. J., Slot, J. W., & Klumperman, J. (1999). Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell*, *98*(1), 81–90.

https://doi.org/10.1016/S0092-8674(00)80608-X

Martinez-Menárguez, J. A., Prekeris, R., Oorschot, V. M., Scheller, R., Slot, J. W., Geuze, H. J., & Klumperman, J. (2001). Peri-Golgi vesicles contain retrograde but not anterograde proteins consistent with the cisternal progression model of intra-Golgi transport. *The Journal of Cell Biology*, *155*(7), 1213–1224. https://doi.org/10.1083/jcb.200108029

Marzioch, M., Henthorn, D. C., Herrmann, J. M., Wilson, R., Thomas, D. Y., Bergeron, J. J., … Rowley, A. (1999). Erp1p and Erp2p, partners for Emp24p and Erv25p in a yeast p24 complex. *Molecular Biology of the Cell*, *10*(6), 1923–1938.

Massaad, M. J., Franzusoff, A., & Herscovics, A. (1999). The processing alpha1,2-mannosidase of Saccharomyces cerevisiae depends on Rer1p for its localization in the endoplasmic reticulum. *European Journal of Cell Biology*, *78*(7), 435–440.

Matsuoka, K., Orci, L., Amherdt, M., Bednarek, S. Y., Hamamoto, S., Schekman, R., & Yeung, T. (1998). COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell*, *93*(2), 263–275.

Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K., & Nakano, A. (2006). Live imaging of yeast Golgi cisternal maturation. *Nature*, *441*(7096), 1007–1010. https://doi.org/10.1038/nature04737

McCaffrey, K., & Braakman, I. (2016). Protein quality control at the endoplasmic reticulum. *Essays In Biochemistry*, *60*(2), 227–235. https://doi.org/10.1042/EBC20160003

McGrath, J. P., & Varshavsky, A. (1989). The yeast STE6 gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. *Nature*, *340*(6232), 400–404. https://doi.org/10.1038/340400a0

McMahon, H. T., & Mills, I. G. (2004). COP and clathrin-coated vesicle budding: different pathways, common approaches. *Current Opinion in Cell Biology*, *16*(4), 379–391. https://doi.org/10.1016/j.ceb.2004.06.009

Mellman, I., & Simons, K. (1992). The Golgi complex: in vitro veritas? *Cell*, *68*(5), 829–840.

Michaelis, S. (1993). STE6, the yeast a-factor transporter. *Seminars in Cell Biology*, *4*(1), 17–27.

Miller, E. A., Beilharz, T. H., Malkus, P. N., Lee, M. C. S., Hamamoto, S., Orci, L., & Schekman, R. (2003). Multiple cargo binding sites on the COPII subunit Sec24p

ensure capture of diverse membrane proteins into transport vesicles. *Cell*, *114*(4), 497–509.

Mironov, A. A., Mironov, A. A., Beznoussenko, G. V., Trucco, A., Lupetti, P., Smith, J. D., … Luini, A. (2003). ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. *Developmental Cell*, *5*(4), 583–594.

Missiaen, L., Dode, L., Vanoevelen, J., Raeymaekers, L., & Wuytack, F. (2007). Calcium in the Golgi apparatus. *Cell Calcium*, *41*(5), 405–416. https://doi.org/10.1016/j.ceca.2006.11.001

Mossessova, E., Bickford, L. C., & Goldberg, J. (2003). SNARE selectivity of the COPII coat. *Cell*, *114*(4), 483–495.

Munro, S. (1995). An investigation of the role of transmembrane domains in Golgi protein retention. *The EMBO Journal*, *14*(19), 4695–4704.

Munro, S., & Pelham, H. R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, *48*(5), 899–907.

Naveed, H., Xu, Y., Jackups, R., & Liang, J. (2012). Predicting three-dimensional structures of transmembrane domains of β-barrel membrane proteins. *Journal of the American Chemical Society*, *134*(3), 1775–1781. https://doi.org/10.1021/ja209895m

Nichols, W. C., Seligsohn, U., Zivelin, A., Terry, V. H., Hertel, C. E., Wheatley, M. A., … Ginsburg, D. (1998). Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell*, *93*(1), 61–70.

Nickel, W. (2011). The unconventional secretory machinery of fibroblast growth factor 2. *Traffic (Copenhagen, Denmark)*, *12*(7), 799–805. https://doi.org/10.1111/j.1600-0854.2011.01187.x

Nickel, W., & Rabouille, C. (2009). Mechanisms of regulated unconventional protein secretion. *Nature Reviews. Molecular Cell Biology*, *10*(2), 148–155. https://doi.org/10.1038/nrm2617

Nie, Z., Hirsch, D. S., & Randazzo, P. A. (2003). Arf and its many interactors. *Current Opinion in Cell Biology*, *15*(4), 396–404.

Nilsson, T., Au, C. E., & Bergeron, J. J. M. (2009). Sorting out glycosylation enzymes in the Golgi apparatus. *FEBS Letters*, *583*(23), 3764–3769.

https://doi.org/10.1016/j.febslet.2009.10.064

Nilsson, T., Hoe, M. H., Slusarewicz, P., Rabouille, C., Watson, R., Hunte, F., … Warren, G. (1994). Kin recognition between medial Golgi enzymes in HeLa cells. *The EMBO Journal*, *13*(3), 562–574.

Nilsson, T., Slusarewicz, P., Hoe, M. H., & Warren, G. (1993). Kin recognition. A model for the retention of Golgi enzymes. *FEBS Letters*, *330*(1), 1–4.

Nishikawa, S., & Nakano, A. (1993). Identification of a gene required for membrane protein retention in the early secretory pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(17), 8179–8183.

Nufer, O., Kappeler, F., Guldbrandsen, S., & Hauri, H.-P. (2003). ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. *Journal of Cell Science*, *116*(Pt 21), 4429–4440. https://doi.org/10.1242/jcs.00759

Nyfeler, B., Zhang, B., Ginsburg, D., Kaufman, R. J., & Hauri, H.-P. (2006). Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. *Traffic (Copenhagen, Denmark)*, *7*(11), 1473–1481. https://doi.org/10.1111/j.1600-0854.2006.00483.x

Okamoto, M., Yoko-o, T., Miyakawa, T., & Jigami, Y. (2008). The cytoplasmic region of alpha-1,6-mannosyltransferase Mnn9p is crucial for retrograde transport from the Golgi apparatus to the endoplasmic reticulum in Saccharomyces cerevisiae. *Eukaryotic Cell*, *7*(2), 310–318. https://doi.org/10.1128/EC.00333-07

Oo, H. Z., Sentani, K., Sakamoto, N., Anami, K., Naito, Y., Oshima, T., … Yasui, W. (2014). Identification of Novel Transmembrane Proteins in Scirrhous-Type Gastric Cancer by the Escherichia coli Ampicillin Secretion Trap (CAST) Method: TM9SF3 Participates in Tumor Invasion and Serves as a Prognostic Factor. *Pathobiology*, *81*(3), 138–148. https://doi.org/10.1159/000357821

Orci, L., Amherdt, M., Ravazzola, M., Perrelet, A., & Rothman, J. E. (2000). Exclusion of golgi residents from transport vesicles budding from Golgi cisternae in intact cells. *The Journal of Cell Biology*, *150*(6), 1263–1270.

Orci, L., Ravazzola, M., Mack, G. J., Barlowe, C., & Otte, S. (2003). Mammalian Erv46 localizes to the endoplasmic reticulum-Golgi intermediate compartment and to cis-Golgi cisternae. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(8), 4586–4591.

https://doi.org/10.1073/pnas.0730885100

Orci, L., Ravazzola, M., Meda, P., Holcomb, C., Moore, H. P., Hicke, L., & Schekman, R. (1991). Mammalian Sec23p homologue is restricted to the endoplasmic reticulum transitional cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(19), 8611–8615.

Orci, L., Stamnes, M., Ravazzola, M., Amherdt, M., Perrelet, A., Söllner, T. H., & Rothman, J. E. (1997). Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell*, *90*(2), 335–349.

Otero, J. H., Lizák, B., & Hendershot, L. M. (2010). Life and death of a BiP substrate. *Seminars in Cell & Developmental Biology*, *21*(5), 472–478. https://doi.org/10.1016/j.semcdb.2009.12.008

Otte, S., & Barlowe, C. (2004). Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nature Cell Biology*, *6*(12), 1189–1194. https://doi.org/10.1038/ncb1195

Otte, S., Belden, W. J., Heidtman, M., Liu, J., Jensen, O. N., & Barlowe, C. (2001). Erv41p and Erv46p: new components of COPII vesicles involved in transport between the ER and Golgi complex. *The Journal of Cell Biology*, *152*(3), 503–518.

Otto, H., Hanson, P. I., & Jahn, R. (1997). Assembly and disassembly of a ternary complex of synaptobrevin, syntaxin, and SNAP-25 in the membrane of synaptic vesicles. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(12), 6197–6201.

Pagano, A., Letourneur, F., Garcia-Estefania, D., Carpentier, J. L., Orci, L., & Paccaud, J. P. (1999). Sec24 proteins and sorting at the endoplasmic reticulum. *The Journal of Biological Chemistry*, *274*(12), 7833–7840.

Pagant, S., Wu, A., Edwards, S., Diehl, F., & Miller, E. A. (2015). Sec24 is a coincidence detector that simultaneously binds two signals to drive ER export. *Current Biology: CB*, *25*(4), 403–412. https://doi.org/10.1016/j.cub.2014.11.070

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. *Science (New York, N.Y.)*, *189*(4200), 347–358.

Park, H.-J., Shabashvili, D., Nekorchuk, M. D., Shyqyriu, E., Jung, J. I., Ladd, T. B., … Kim, S.-H. (2012). Retention in endoplasmic reticulum 1 (RER1) modulates amyloid-β (Aβ) production by altering trafficking of γ-secretase and amyloid precursor protein (APP). *The Journal of Biological Chemistry*, *287*(48), 40629–40640.

https://doi.org/10.1074/jbc.M112.418442

Pelham, H. R. (1988). Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *The EMBO Journal*, *7*(4), 913–918.

Pelham, H. R. (1994). About turn for the COPs? *Cell*, *79*(7), 1125–1127.

Pelham, H. R., Hardwick, K. G., & Lewis, M. J. (1988). Sorting of soluble ER proteins in yeast. *The EMBO Journal*, *7*(6), 1757–1762.

Pellett, P. A., Dietrich, F., Bewersdorf, J., Rothman, J. E., & Lavieu, G. (2013). Inter-Golgi transport mediated by COPI-containing vesicles carrying small cargoes. *eLife*, *2*, e01296. https://doi.org/10.7554/eLife.01296

Perrin, J., Mortier, M., Jacomin, A.-C., Viargues, P., Thevenon, D., & Fauvarque, M.-O. (2015). The nonaspanins TM9SF2 and TM9SF4 regulate the plasma membrane localization and signalling activity of the peptidoglycan recognition protein PGRP-LC in Drosophila. *Journal of Innate Immunity*, *7*(1), 37–46. https://doi.org/10.1159/000365112

Powers, J., & Barlowe, C. (1998). Transport of axl2p depends on erv14p, an ER-vesicle protein related to the Drosophila cornichon gene product. *The Journal of Cell Biology*, *142*(5), 1209–1222.

Powers, J., & Barlowe, C. (2002). Erv14p directs a transmembrane secretory protein into COPII-coated transport vesicles. *Molecular Biology of the Cell*, *13*(3), 880–891. https://doi.org/10.1091/mbc.01-10-0499

Qu, Y., Franchi, L., Nunez, G., & Dubyak, G. R. (2007). Nonclassical IL-1 beta secretion stimulated by P2X7 receptors is dependent on inflammasome activation and correlated with exosome release in murine macrophages. *Journal of Immunology (Baltimore, Md.: 1950)*, *179*(3), 1913–1925.

Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E. G., Warren, G., & Nilsson, T. (1995). Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *Journal of Cell Science*, *108 (Pt 4)*, 1617–1627.

Radhakrishnan, A., Sun, L.-P., Kwon, H. J., Brown, M. S., & Goldstein, J. L. (2004). Direct binding of cholesterol to the purified membrane region of SCAP: mechanism for a sterol-sensing domain. *Molecular Cell*, *15*(2), 259–268. https://doi.org/10.1016/j.molcel.2004.06.019

Raykhel, I., Alanen, H., Salo, K., Jurvansuu, J., Nguyen, V. D., Latva-Ranta, M., & Ruddock, L. (2007). A molecular specificity code for the three mammalian KDEL receptors.

The Journal of Cell Biology, *179*(6), 1193–1204. https://doi.org/10.1083/jcb.200705180

Risselada, H. J., & Grubmüller, H. (2012). How SNARE molecules mediate membrane fusion: recent insights from molecular simulations. *Current Opinion in Structural Biology*, *22*(2), 187–196. https://doi.org/10.1016/j.sbi.2012.01.007

Ronchi, P., Colombo, S., Francolini, M., & Borgese, N. (2008). Transmembrane domain-dependent partitioning of membrane proteins within the endoplasmic reticulum. *The Journal of Cell Biology*, *181*(1), 105–118. https://doi.org/10.1083/jcb.200710093

Rothblatt, J. A., Deshaies, R. J., Sanders, S. L., Daum, G., & Schekman, R. (1989). Multiple genes are required for proper insertion of secretory proteins into the endoplasmic reticulum in yeast. *The Journal of Cell Biology*, *109*(6 Pt 1), 2641–2652.

Rothman, J. E. (1994). Mechanisms of intracellular protein transport. *Nature*, *372*(6501), 55–63. https://doi.org/10.1038/372055a0

Rothman, J. E., & Wieland, F. T. (1996). Protein sorting by transport vesicles. *Science (New York, N.Y.)*, *272*(5259), 227–234.

Rubartelli, A., Cozzolino, F., Talio, M., & Sitia, R. (1990). A novel secretory pathway for interleukin-1 beta, a protein lacking a signal sequence. *The EMBO Journal*, *9*(5), 1503–1510.

Ryu, J.-K., Min, D., Rah, S.-H., Kim, S. J., Park, Y., Kim, H., … Yoon, T.-Y. (2015). Spring-loaded unraveling of a single SNARE complex by NSF in one round of ATP turnover. *Science (New York, N.Y.)*, *347*(6229), 1485–1489. https://doi.org/10.1126/science.aaa5267

Sato, K., Sato, M., & Nakano, A. (1997). Rer1p as common machinery for the endoplasmic reticulum localization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(18), 9693–9698.

Sato, K., Sato, M., & Nakano, A. (2001). Rer1p, a retrieval receptor for endoplasmic reticulum membrane proteins, is dynamically localized to the Golgi apparatus by coatomer. *The Journal of Cell Biology*, *152*(5), 935–944.

Sato, M., Sato, K., & Nakano, A. (1996). Endoplasmic reticulum localization of Sec12p is achieved by two mechanisms: Rer1p-dependent retrieval that requires the transmembrane domain and Rer1p-independent retention that involves the

cytoplasmic domain. *The Journal of Cell Biology*, *134*(2), 279–293.

Scheel, A. A., & Pelham, H. R. (1998). Identification of amino acids in the binding pocket of the human KDEL receptor. *The Journal of Biological Chemistry*, *273*(4), 2467–2472.

Schimmöller, F., Díaz, E., Mühlbauer, B., & Pfeffer, S. R. (1998). Characterization of a 76 kDa endosomal, multispanning membrane protein that is highly conserved throughout evolution. *Gene*, *216*(2), 311–318.

Schimmöller, F., Singer-Krüger, B., Schröder, S., Krüger, U., Barlowe, C., & Riezman, H. (1995). The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *The EMBO Journal*, *14*(7), 1329–1339.

Schoberer, J., Runions, J., Steinkellner, H., Strasser, R., Hawes, C., & Osterrieder, A. (2010). Sequential depletion and acquisition of proteins during Golgi stack disassembly and reformation. *Traffic (Copenhagen, Denmark)*, *11*(11), 1429–1444. https://doi.org/10.1111/j.1600-0854.2010.01106.x

Schweizer, A., Fransen, J. A., Matter, K., Kreis, T. E., Ginsel, L., & Hauri, H. P. (1990). Identification of an intermediate compartment involved in protein transport from endoplasmic reticulum to Golgi apparatus. *European Journal of Cell Biology*, *53*(2), 185–196.

Semenza, J. C., Hardwick, K. G., Dean, N., & Pelham, H. R. (1990). ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell*, *61*(7), 1349–1357.

Sengupta, P., Satpute-Krishnan, P., Seo, A. Y., Burnette, D. T., Patterson, G. H., & Lippincott-Schwartz, J. (2015). ER trapping reveals Golgi enzymes continually revisit the ER through a recycling pathway that controls Golgi organization. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), E6752-6761. https://doi.org/10.1073/pnas.1520957112

Sharpe, H. J., Stevens, T. J., & Munro, S. (2010). A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell*, *142*(1), 158–169. https://doi.org/10.1016/j.cell.2010.05.037

Shibuya, A., Margulis, N., Christiano, R., Walther, T. C., & Barlowe, C. (2015). The Erv41-Erv46 complex serves as a retrograde receptor to retrieve escaped ER proteins. *The Journal of Cell Biology*, *208*(2), 197–209.

https://doi.org/10.1083/jcb.201408024

Sitia, R., & Meldolesi, J. (1992). Endoplasmic reticulum: a dynamic patchwork of specialized subregions. *Molecular Biology of the Cell*, *3*(10), 1067–1072.

Sitia, R., Neuberger, M., Alberini, C., Bet, P., Fra, A., Valetti, C., … Milstein, C. (1990). Developmental regulation of IgM secretion: the role of the carboxy-terminal cysteine. *Cell*, *60*(5), 781–790.

Sohn, K., Orci, L., Ravazzola, M., Amherdt, M., Bremser, M., Lottspeich, F., … Wieland, F. T. (1996). A major transmembrane protein of Golgi-derived COPI-coated vesicles involved in coatomer binding. *The Journal of Cell Biology*, *135*(5), 1239–1248.

Söllner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H., & Rothman, J. E. (1993). A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell*, *75*(3), 409–418.

Sousa, V. L., Brito, C., Costa, T., Lanoix, J., Nilsson, T., & Costa, J. (2003). Importance of Cys, Gln, and Tyr from the transmembrane domain of human alpha 3/4 fucosyltransferase III for its localization and sorting in the Golgi of baby hamster kidney cells. *The Journal of Biological Chemistry*, *278*(9), 7624–7629. https://doi.org/10.1074/jbc.M209325200

Spira, F., Mueller, N. S., Beck, G., von Olshausen, P., Beig, J., & Wedlich-Söldner, R. (2012). Patchwork organization of the yeast plasma membrane into numerous coexisting domains. *Nature Cell Biology*, *14*(6), 640–648. https://doi.org/10.1038/ncb2487

Stamnes, M. A., Craighead, M. W., Hoe, M. H., Lampen, N., Geromanos, S., Tempst, P., & Rothman, J. E. (1995). An integral membrane component of coatomer-coated transport vesicles defines a family of proteins involved in budding. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(17), 8011–8015.

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

Stephens, D. J., & Pepperkok, R. (2001). Illuminating the secretory pathway: when do we need vesicles? *Journal of Cell Science*, *114*(6), 1053.

Stephens, D. J., & Pepperkok, R. (2002). Imaging of procollagen transport reveals COPI-dependent cargo sorting during ER-to-Golgi transport in mammalian cells.

Journal of Cell Science, *115*(Pt 6), 1149–1160.

Steringer, J. P., Bleicken, S., Andreas, H., Zacherl, S., Laussmann, M., Temmerman, K., … Nickel, W. (2012). Phosphatidylinositol 4,5-bisphosphate (PI(4,5)P2)-dependent oligomerization of fibroblast growth factor 2 (FGF2) triggers the formation of a lipidic membrane pore implicated in unconventional secretion. *The Journal of Biological Chemistry*, *287*(33), 27659–27669. https://doi.org/10.1074/jbc.M112.381939

Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R., & Schekman, R. (1992). Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Molecular Biology of the Cell*, *3*(2), 129–142.

Storrie, B., Pepperkok, R., & Nilsson, T. (2000). Breaking the COPI monopoly on Golgi recycling. *Trends in Cell Biology*, *10*(9), 385–391.

Strating, J. R. P. M., & Martens, G. J. M. (2009). The p24 family and selective transport processes at the ER-Golgi interface. *Biology of the Cell*, *101*(9), 495–509. https://doi.org/10.1042/BC20080233

Sugasawa, T., Lenzen, G., Simon, S., Hidaka, J., Cahen, A., Guillaume, J. L., … Nahmias, C. (2001). The iodocyanopindolol and SM-11044 binding protein belongs to the TM9SF multispanning membrane protein superfamily. *Gene*, *273*(2), 227–237.

Sun, L.-P., Li, L., Goldstein, J. L., & Brown, M. S. (2005). Insig required for sterol-mediated inhibition of Scap/SREBP binding to COPII proteins in vitro. *The Journal of Biological Chemistry*, *280*(28), 26483–26490. https://doi.org/10.1074/jbc.M504041200

Takida, S., Maeda, Y., & Kinoshita, T. (2008). Mammalian GPI-anchored proteins require p24 proteins for their efficient transport from the ER to the plasma membrane. *The Biochemical Journal*, *409*(2), 555–562. https://doi.org/10.1042/BJ20070234

Teese, M. G., & Langosch, D. (2015). Role of GxxxG Motifs in Transmembrane Domain Interactions. *Biochemistry*, *54*(33), 5125–5135. https://doi.org/10.1021/acs.biochem.5b00495

Thor, F., Gautschi, M., Geiger, R., & Helenius, A. (2009). Bulk flow revisited: transport of a soluble protein in the secretory pathway. *Traffic (Copenhagen, Denmark)*, *10*(12), 1819–1830. https://doi.org/10.1111/j.1600-0854.2009.00989.x

Todorow, Z., Spang, A., Carmack, E., Yates, J., & Schekman, R. (2000). Active recycling of yeast Golgi mannosyltransferase complexes through the endoplasmic reticulum.

Proceedings of the National Academy of Sciences of the United States of America, *97*(25), 13643–13648. https://doi.org/10.1073/pnas.250472397

Townsley, F. M., Wilson, D. W., & Pelham, H. R. (1993). Mutational analysis of the human KDEL receptor: distinct structural requirements for Golgi retention, ligand binding and retrograde transport. *The EMBO Journal*, *12*(7), 2821–2829.

Trahey, M., & Hay, J. C. (2010). Transport vesicle uncoating: it's later than you think. *F1000 Biology Reports*, *2*, 47.

Tu, L., Chen, L., & Banfield, D. K. (2012). A conserved N-terminal arginine-motif in GOLPH3-family proteins mediates binding to coatomer. *Traffic (Copenhagen, Denmark)*, *13*(11), 1496–1507. https://doi.org/10.1111/j.1600-0854.2012.01403.x

Tu, L., Tai, W. C. S., Chen, L., & Banfield, D. K. (2008). Signal-mediated dynamic retention of glycosyltransferases in the Golgi. *Science (New York, N.Y.)*, *321*(5887), 404–407. https://doi.org/10.1126/science.1159411

Ulmschneider, M. B., & Sansom, M. S. (2001). Amino acid distributions in integral membrane protein structures. *Biochimica Et Biophysica Acta*, *1512*(1), 1–14.

Ulmschneider, M. B., Tieleman, D. P., & Sansom, M. S. P. (2005). The role of extra-membranous inter-helical loops in helix-helix interactions. *Protein Engineering, Design & Selection: PEDS*, *18*(12), 563–570. https://doi.org/10.1093/protein/gzi059

Velloso, L. M., Svensson, K., Schneider, G., Pettersson, R. F., & Lindqvist, Y. (2002). Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. *The Journal of Biological Chemistry*, *277*(18), 15979–15984. https://doi.org/10.1074/jbc.M112098200

Venditti, R., Wilson, C., & De Matteis, M. A. (2014). Exiting the ER: what we know and what we don't. *Trends in Cell Biology*, *24*(1), 9–18. https://doi.org/10.1016/j.tcb.2013.08.005

Villeneuve, J., Duran, J., Scarpa, M., Bassaganyas, L., Van Galen, J., & Malhotra, V. (2017). Golgi enzymes do not cycle through the endoplasmic reticulum during protein secretion or mitosis. *Molecular Biology of the Cell*, *28*(1), 141–151. https://doi.org/10.1091/mbc.E16-08-0560

Vincent, M. J., Martin, A. S., & Compans, R. W. (1998). Function of the KKXX motif in

endoplasmic reticulum retrieval of a transmembrane protein depends on the length and structure of the cytoplasmic domain. *The Journal of Biological Chemistry*, *273*(2), 950–956.

Vollenweider, F., Kappeler, F., Itin, C., & Hauri, H. P. (1998). Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *The Journal of Cell Biology*, *142*(2), 377–389.

Volpicelli-Daley, L. A. (2005). Isoform-selective Effects of the Depletion of ADP-Ribosylation Factors 1-5 on Membrane Traffic. *Molecular Biology of the Cell*, *16*(10), 4495–4508. https://doi.org/10.1091/mbc.E04-12-1042

Walter, P., Gilmore, R., Müller, M., & Blobel, G. (1982). The protein translocation machinery of the endoplasmic reticulum. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *300*(1099), 225–228.

Walter, P., & Johnson, A. E. (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annual Review of Cell Biology*, *10*, 87–119. https://doi.org/10.1146/annurev.cb.10.110194.000511

Watson, R. T., & Pessin, J. E. (2001). Transmembrane domain length determines intracellular membrane compartment localization of syntaxins 3, 4, and 5. *American Journal of Physiology. Cell Physiology*, *281*(1), C215-223.

Wegmann, D., Hess, P., Baier, C., Wieland, F. T., & Reinhard, C. (2004). Novel Isotypic / Subunits Reveal Three Coatomer Complexes in Mammals. *Molecular and Cellular Biology*, *24*(3), 1070–1080. https://doi.org/10.1128/MCB.24.3.1070-1080.2004

Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K., & Martoglio, B. (2002). Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science (New York, N.Y.)*, *296*(5576), 2215–2218. https://doi.org/10.1126/science.1070925

Wieland, F. T., Gleason, M. L., Serafini, T. A., & Rothman, J. E. (1987). The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell*, *50*(2), 289–300.

Wilkinson, B., & Gilbert, H. F. (2004). Protein disulfide isomerase. *Biochimica Et Biophysica Acta*, *1699*(1–2), 35–44. https://doi.org/10.1016/j.bbapap.2004.02.017

Williams, G. T. (1990). The sequence of the mu transmembrane segment determines the tissue specificity of the transport of immunoglobulin M to the cell surface. *Journal of Experimental Medicine*, *171*(3), 947–952. https://doi.org/10.1084/jem.171.3.947

Wilson, D. W., Lewis, M. J., & Pelham, H. R. (1993). pH-dependent binding of KDEL to its receptor in vitro. *The Journal of Biological Chemistry*, *268*(10), 7465–7468.

Wood, C. S., Schmitz, K. R., Bessman, N. J., Setty, T. G., Ferguson, K. M., & Burd, C. G. (2009). PtdIns4P recognition by Vps74/GOLPH3 links PtdIns 4-kinase signaling to retrograde Golgi trafficking. *The Journal of Cell Biology*, *187*(7), 967–975. https://doi.org/10.1083/jcb.200909063

Wu, M. M., Grabe, M., Adams, S., Tsien, R. Y., Moore, H. P., & Machen, T. E. (2001). Mechanisms of pH regulation in the regulated secretory pathway. *The Journal of Biological Chemistry*, *276*(35), 33027–33035. https://doi.org/10.1074/jbc.M103917200

Wu, M. M., Llopis, J., Adams, S., McCaffery, J. M., Kulomaa, M. S., Machen, T. E., … Tsien, R. Y. (2000). Organelle pH studies using targeted avidin and fluorescein-biotin. *Chemistry & Biology*, *7*(3), 197–209.

Yang, J., Nune, M., Zong, Y., Zhou, L., & Liu, Q. (2015). Close and Allosteric Opening of the Polypeptide-Binding Site in a Human Hsp70 Chaperone BiP. *Structure (London, England: 1993)*, *23*(12), 2191–2203. https://doi.org/10.1016/j.str.2015.10.012

Ye, J., Rawson, R. B., Komuro, R., Chen, X., Davé, U. P., Prywes, R., … Goldstein, J. L. (2000). ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Molecular Cell*, *6*(6), 1355–1364.

Yoo, J.-S., Moyer, B. D., Bannykh, S., Yoo, H.-M., Riordan, J. R., & Balch, W. E. (2002). Non-conventional trafficking of the cystic fibrosis transmembrane conductance regulator through the early secretory pathway. *The Journal of Biological Chemistry*, *277*(13), 11401–11409. https://doi.org/10.1074/jbc.M110263200

Zhang, B., Cunningham, M. A., Nichols, W. C., Bernat, J. A., Seligsohn, U., Pipe, S. W., … Ginsburg, D. (2003). Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nature Genetics*, *34*(2), 220–225. https://doi.org/10.1038/ng1153

Zhang, Y. C., Zhou, Y., Yang, C. Z., & Xiong, D. S. (2009). A review of ERGIC-53: its structure, functions, regulation and relations with diseases. *Histology and Histopathology*, *24*(9), 1193–1204.

Zhang, Y., Motamed, M., Seemann, J., Brown, M. S., & Goldstein, J. L. (2013). Point mutation in luminal loop 7 of Scap protein blocks interaction with loop 1 and abolishes movement to Golgi. *The Journal of Biological Chemistry*, *288*(20),

14059–14067. https://doi.org/10.1074/jbc.M113.469528

Zhao, M., Wu, S., Zhou, Q., Vivona, S., Cipriano, D. J., Cheng, Y., & Brunger, A. T. (2015). Mechanistic insights into the recycling machine of the SNARE complex. *Nature*, *518*(7537), 61–67. https://doi.org/10.1038/nature14148

Zhuravleva, A., & Gierasch, L. M. (2015). Substrate-binding domain conformational dynamics mediate Hsp70 allostery. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(22), E2865-2873. https://doi.org/10.1073/pnas.1506692112

Zimmermann, R., Eyrisch, S., Ahmad, M., & Helms, V. (2011). Protein translocation across the ER membrane. *Biochimica Et Biophysica Acta*, *1808*(3), 912–924. https://doi.org/10.1016/j.bbamem.2010.06.015