UNIVERSITY OF OXFORD

Characterising the rhomboid-like protein TMEM115

Xin Shen

Merton College

A thesis submitted to the Division of Medical Sciences,

University of Oxford, in partial fulfilment of the requirement for the Degree of Doctor of Philosophy in Interdisciplinary Bioscience DTP

Trinity Term 2017

Sir William Dunn School of Pathology

University of Oxford

Statement of originality

I certify that the thesis and research is the product of my own work, conducted between July 2014 and August 2017 at Sir William Dunn School in the University of Oxford. Any ideas and quotations from the work of others are fully acknowledged in accordance with the standard referencing practices of the discipline. It has not been submitted, either in whole or in part, for a degree or qualification at any other institute.

Acknowledgments

Firstly, I would like to thank my supervisor, Prof Matthew Freeman, for giving me this opportunity to working on this exciting project and his kindness support and insightful advice throughout my DPhil study. It is such a great honour and pleasure to be Matthew's tenth DPhil student.

I would also like to thank all my colleagues for creating a positive and enjoyable environment in our laboratory. I would like to give special thanks to Dr Angela Moncada Pazos, Dr Adam Grieve and Dr Clemence Levet, not only for their patience and useful discussions in my project but also for their support and friendship.

My warmest gratitude goes to my parents, Prof Li Xiao and Mr Yongming Shen, and my grandparents, Mr Longhai Xiao, Mrs Jinlan Wong, Mr Jiuchen Shen and Mrs Zhenmei Zhou, for their irreplaceable lifelong love and support.

My special gratitude goes to my husband, Dr Yang Liu, for his love, understanding and support not only during this course but also for the past six years.

Finally, I would like to thank the University of Oxford Medical division, Dunn School of Pathology and Doctoral Training Centre for organising courses and their support for students. I would like to thank Chinese Scholarship Council for the full fellowship they offered to support my whole DPhil study.

Table of content

Abstract

The rhomboid-like superfamily of proteins comprises transmembrane proteins with an ancient evolutionary origin. They include both intramembrane serine proteases and inactive pseudoproteases, and they have diverse cellular and pathophysiological functions. These include regulation of growth factor signalling, protein quality control, trafficking, and mitochondrial dynamics. TMEM115, a recently recognised enzymatically inactive member of the rhomboid-like superfamily (confirmed in this thesis), is conserved from yeast to human and is ubiquitously expressed in all tissues. The absence of TMEM115 in both *Drosophila* and mice causes severe phenotypes. These compelling preliminary data indicate that TMEM115 has important cellular functions. To capitalise on these preliminary observations, the overall aim of my PhD project was to characterise the mammalian TMEM115 both structurally and functionally.

Structurally, I performed a topology study for TMEM115, and with a combination of bioinformatic and experimental analysis, proved that TMEM115 indeed has a six TMD structure. Using the HHpred and Phyre algorithms, which identify structural similarity among proteins, a high degree of homology was identified between the TMD regions of TMEM115 and other rhomboid-like proteins. The above analysis together with the topology of TMEM115 definitively positions TMEM115 in the rhomboid-like superfamily.

To elucidate the biological role of TMEM115, I started with a proteomic approach, a BioID proximity screen, to identify novel interactors for TMEM115 under different physiologically relevant conditions. A number of binding partners were identified in the BioID screen and validated with co-immunoprecipitation. These indicate intriguing possible functions of TMEM115, including regulating lipid biology, protein trafficking, protein degradation and ion channels. Given that several candidates are involved in protein trafficking, I used a secretome profile analysis method, the SPECS, for identifying TMEM115 dependent secreted proteins.

A role of TMEM115 in regulating the fundamental *de novo* lipogenesis pathway, the SCAP-SREBP pathway (Freeman lab unpublished data), had been identified during the course of my project. Pursuing this, I investigated the possible relationship between TMEM115, SCAP and the novel interactor, p62; the results suggest that TMEM115 may be involved in regulating lipid homeostasis by modulating SCAP levels through the proteasomal degradation machinery.

List of abbreviations

aa	amino acid
ADAM	a disintegrin and metalloprotease
AREG	amphiregulin
BioID	proximity-dependent biotin identification
BirA*	mutant of Bifunctional ligase/repressor (R118G)
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
co-IP	co-immunoprecipitation
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
ER	endoplasmic reticulum
ERAD	ER-associated degradation
EREG	epiregulin
ERGIC53	ER-Golgi intermediate compartment 53 kDa protein
FCS	fetal calf serum
FW	forward
GFP	green fluorescent protein
GM130	Golgin subfamily A member 2
HA	hemagglutinin
HEK cells	human embryonic kidney (cells)
HHpred	Helix-Helix profile server
HRP	horseradish peroxidase
IP	immunoprecipitation
IRHD	iRhom homology domain
KO	knockout

LIR motif	LC3 interacting region
ManNAZ	N-azidoacetylmannosamine tetraacylated
MAPK	mitogen-activated protein kinase
MCF-7	Michigan Cancer Foundation-7
MEF	mouse embryonic fibroblast
mRNA	messenger RNA
p62	sequestosome 1
PARL	Presenilins-associated rhomboid-like protein
PB domain	proteasome binding domain
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PD	Parkinson's disease
PINK1	PTEN-induced putative kinase 1
PMA	phorbol 12-myristate 13-acetate
PPI	protein-protein interaction
PVDF	polyvinylidene fluoride
RNA	ribonucleic acid
RV	reverse
SCAP	sterol regulatory element-binding protein cleavage-activating protein
SDS	sodium dodecyl sulfate
siRNA	small interfering RNA
SPECS	secretome protein enrichment with click sugars
SREBP	sterol regulatory element-binding protein
Tet	tetracycline
TGFα	Transforming growth factor alpha
TGN46	Trans-Golgi network integral membrane protein 2
TMD	transmembrane domain
TMEM115	transmembrane protein 115
TNFR	TNFα receptor
TNFα	tumour necrosis factor alpha
TX-100	Triton X-100
UAB domain	ubiquitin-associated domain

VANG1	Vang-like protein VANG1
VCP/p97	valosin-containing protein
WR motif	tryptophan-arginine motif
WT	wild-type
β-COP	Coatomer subunit beta

Chapter 1 Literature of the rhomboid-like superfamily

1.1 Overview

The rhomboid-like superfamily was discovered in 2001 (Urban et al. 2001), and it comprises polytopic membrane proteins that are found in all kingdoms of life. Phylogenetic analysis indicates that rhomboid proteins originated very early in evolution and the first ancestor was possibly a rhomboid protease. Members of the rhomboid-like superfamily are widely conserved; they have distinctive and crucial cellular functions, including regulating protein quality control, growth signalling, trafficking and mitochondrial dynamics (Wasserman et al. 2000, Urban et al. 2001, Adrain et al. 2011, Greenblatt et al. 2011, Meissner et al. 2011, Zettl et al. 2011, Adrain et al. 2012, Freeman 2014, Luo et al. 2016). The first members of the rhomboid-like superfamily to be identified were the rhomboid intramembrane serine proteases, which are polytopic membrane proteases with their catalytic site embedded in the lipid bilayer (Lemberg et al. 2005, Lemieux et al. 2007). The first identified member, *Drosophila* Rhomboid-1, was found to regulate the epidermal growth factor receptor (EGFR) signalling pathway through promoting the cleavage and release of EGFR ligands from their precursors (Wasserman et al. 2000, Urban et al. 2001). Intriguingly, many rhomboid-like proteins that have been subsequently identified have lost crucial catalytic residues during evolution; therefore they no longer have catalytic activity. However, instead of being useless relic proteins, they adopted a wide range of non-proteolytic functions, including regulating the metalloprotease TACE, the epidermal growth factor receptor (EGFR), ER-associated degradation (ERAD) and other important cellular signalling pathways. Reflecting their common history, members of the rhomboid-like superfamily share structural similarities. Specifically, by definition, they all contain a six transmembrane domain (TMD) rhomboid-like core. Some members have an additional TMD present at either the amino- (PARL) or carboxy-terminus (RHBDL1-3 and iRhoms) of the core Rhomboid-like domain (Lemberg 2013, Dusterhoft et al. 2017).

In this introduction, I will review what is known about the rhomboid-like superfamily, to which I will demonstrate that TMEM115 belongs. I will review how rhomboid-like members regulate distinctive intracellular signalling pathways through proteolysis, trafficking and protein degradation (Urban et al. 2001, Adrain et al. 2011, Greenblatt et al. 2011, Meissner et al. 2011, Zettl et al. 2011, Adrain et al. 2012, McIlwain et al. 2012), which may provide some insight into how TMEM115 regulates its interactors. Since TMEM115 is a newly identified member of the superfamily (Freeman 2014), there is very little known about this protein; only two papers have referred to it (Ivanova et al. 2008, Ong et al. 2014). I will also review what is known about TMEM115 both from the literature and unpublished preliminary data from Freeman lab members. Since our recent evidence suggests a role for TMEM115 in lipid homeostasis, I will also introduce the regulation of lipid biosynthesis.

1.2 Members of the rhomboid-like superfamily

In spite of their conservation, the rhomboid-like superfamily shows extensive sequence divergence, which hinders the comprehensive identification of its members with simple sequence homology tools like BLAST. With more sophisticated bioinformatic analysis, eight major branches of proteins have been identified in the mammalian rhomboid-like superfamily (Figure1.1) (Adrain et al. 2012, Freeman 2014). The common ancestor of rhomboid-like proteins is thought to be an active intramembrane protease, despite the fact that the majority of members now are pseudoproteases, which have lost the crucial catalytic residues (Freeman 2014). In mammals, the superfamily includes the active rhomboid proteases, RHBDL1-4, that reside within the secretory pathway and PARL, which targets mitochondria. Evolutionarily, RHBDL1-3 and the iRhoms are closely related to each other, and share a similar topology, with the rhomboid-like domain located in the first six TMDs and an additional C-terminal, seventh TMD. RHBDL4 is an active protease that is more distantly related; it has a six TMD topology. The mitochondrial rhomboid protease, PARL, has a seventh TMD N-terminal to the core six TMD rhomboid-like domain. In addition to the proteases and the iRhoms,

the superfamily includes a number of much more distant non-proteases, including derlins, UBAC2, RHBDD2 and RHBDD3, all of which have six TMDs (Lemberg 2013, Freeman 2014). In Chapter 4, I will provide evidence that TMEM115 is also a bona fide member of the rhomboid-like superfamily.

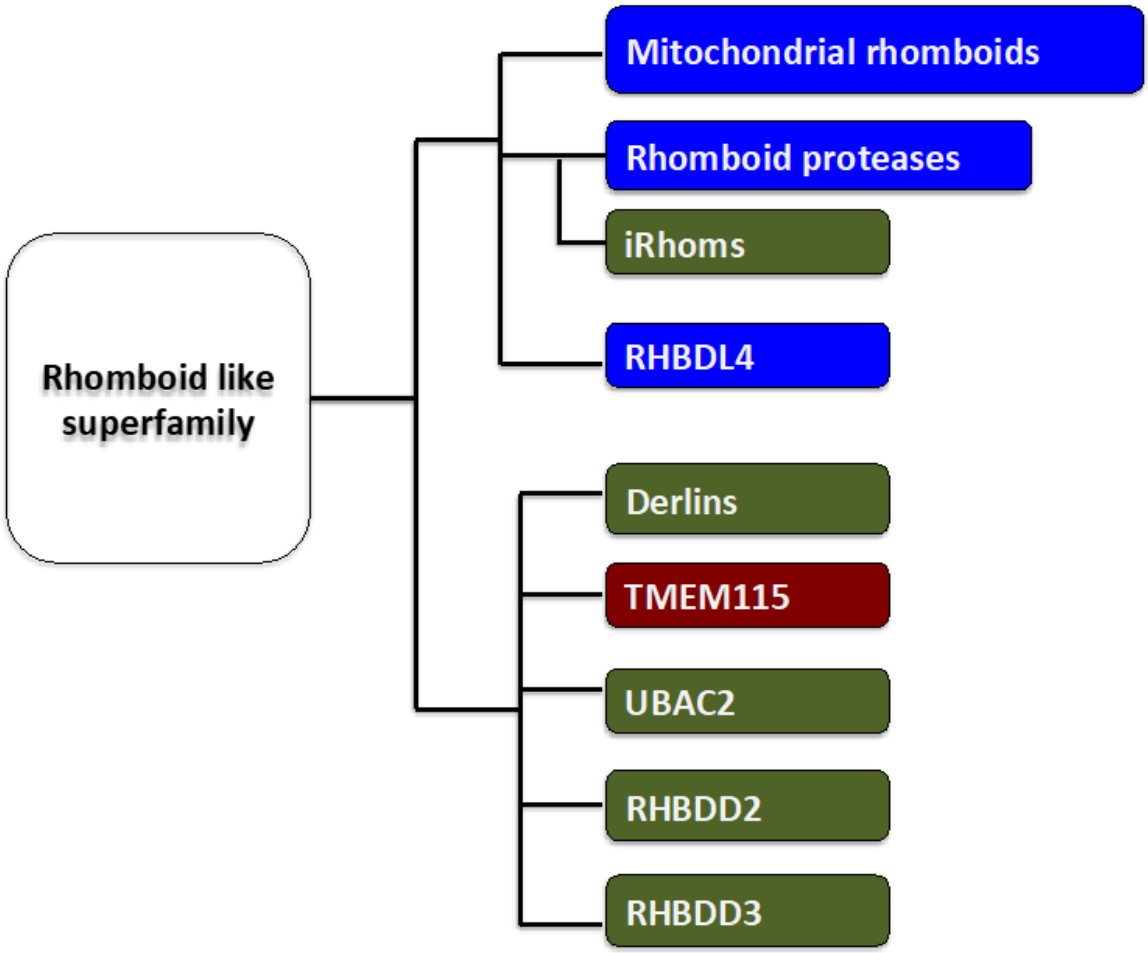


Figure 1: The eukaryotic rhomboid-like superfamily

All the catalytically active rhomboids are labelled in blue. The catalytically inactive members are labelled in green. TMEM115, my protein of interest, is labelled in red.

1.3 Structural features of rhomboid-like proteins

Despite the relatively low sequence similarity among rhomboid-like proteins, the core domain of six TMDs is conserved. The structure of the *E. coli* rhomboid protease, GlpG, which has been solved to atomic resolution, revealed the major structural features of a rhomboid-like domain (Wang et al. 2006, Wu et al. 2006). Interestingly, the amino terminus of the fourth TMD is about 10Å below the membrane surface, creating a cavity inside the protein. The catalytic dyad residues, serine and histidine, are located in TM4 and TM6 respectively, and are located in a water-filled environment surrounded by transmembrane helices; this shields them from the lipid bilayer (Wang et al. 2006, Wu et al. 2006). Other widely but not universally conserved features include a tryptophan-arginine (WR) motif, which is proposed to have a structural role within the first luminal loop at the water-membrane interface, and the helix-helix dimerization motif (GxxxG) in the sixth TMD, which may aid the packing of TMD4 and TMD6 to stabilise the alignment of the two catalytic residues (Russ et al. 2000, Wang et al. 2006, Wu et al. 2006).

1.4 Rhomboid proteases

Rhomboids are intramembrane serine proteases that were first discovered in *Drosophila* (Urban et al. 2001). The catalytic activity is important for the molecular function of rhomboid proteases; therefore, the key to unlocking their biological function is to identify their natural substrates. In *Drosophila*, Rhomboids1-3 control the EGFR pathway through cleavage and release of active EGF ligands from their membrane-tethered forms (Urban et al. 2001). In mammals, five rhomboid proteases have been identified, among which, PARL is present in mitochondria, and the other four, RHBDL1-3 and RHBDL4, are distributed through the secretory pathway (Cipolat et al. 2006, Lemberg et al. 2007, Lastun et al. 2016).

The mammalian mitochondrial rhomboid, PARL, has been studied extensively because of a proposed link with type II diabetes (Civitarese et al. 2010) and Parkinson's disease (PD) (McQuibban et al. 2006, Meissner et al. 2011). Among the proposed substrates of PARL, the focus has been on PINK1 due to the

potential medical significance of the link between PD and PINK1 cleavage (McIlwain et al. 2012). In healthy cells, PARL activity degrades PINK1 through cleavage, thereby preventing the accumulation of PINK1. This cleavage is dependent on an intact mitochondrial membrane potential and can be abolished by mutations in the PINK1 TMD region, or mitochondrial damage that disrupts the membrane potential (Meissner et al. 2011). The failure of PINK1 turnover leads to the recruitment of E3 ubiquitin ligase Parkin and ubiquitination of mitochondrial proteins, triggering mitophagy, a specific form of autophagy that contributes to the development of PD (Martinez-Vicente 2017, Rodolfo et al. 2017).

Four mammalian rhomboid proteases are present in the secretory pathway, RHBDL1-4. RHBDL1-3 are more closely related to each other than to RHBDL4 (Lemberg et al. 2007); they all have a similar seven TMD topology, with an additional TMD attached C-terminally to the core rhomboid-like domain. In contrast, RHBDL4 forms an independent clan in the rhomboid-like superfamily and contains only six TMDs. Regarding cellular localisation, RHBDL1 is predominantly located in the Golgi apparatus (Lohi et al. 2004), while RHBDL3 is located predominantly in endosomes, but can also be detected in the Golgi and on the plasma membrane (Zettl et al. 2011). RHBDL4 is ER-localised (Fleig et al. 2012), while RHBDL2 is mostly localised at the plasma membrane (Adrain et al. 2011). Among the four mammalian secretase rhomboids, no substrates have yet been identified for RHBDL1 and 3, while RHBDL2 and 4 are better characterised, both having at least some known substrates and cellular functions, although in neither case is there a full knowledge of the extent of their roles.

RHBDL2 shares some proteolytic specificity with Drosophila Rhomboid1-3, in that it can cleave known substrates of Drosophila rhomboids, Spitz and Gurken (Lohi et al. 2004, Lemberg et al. 2005). Mammalian RHBDL2 can cleave EGF but not other EGF-like ligands (Adrain et al. 2011); in mammals, the major processing enzymes for EGF and other related ligands are the metalloproteases ADAM10 and ADAM17 (Sahin et al. 2004). The cleavage of EGF by RHBDL2 only occurs when metalloprotease activity is inhibited, raising as yet unanswered questions

about the physiological significance of RHBDL2-driven EGF cleavage (Adrain et al. 2011). Interestingly, enhanced cleavage of EGF by RHBDL2 is observed in several cancer cell lines, suggesting the possible pathological relevance of RHBDL2-driven EGF cleavage (Adrain et al. 2011, Cheng et al. 2014). It has also been reported that RHBDL2 can also cleave the EGF receptor itself (Liao et al. 2012), which has the potential to add complexity to the relationship between RHBDL2 and the EGFR signalling pathway. Note, however, that again the physiological significance of EGF receptor cleavage is not clear. In mammalian cells, a range of novel substrates that are specifically cleaved by RHBDL2 have been identified, including interleukin-6 receptor (IL6R), cell surface protease inhibitor Spint-1, the collagen receptor tyrosine kinase DDR1, N-Cadherin, CLCP1/DCBLD2, KIRREL, BCAM etc. (Johnson et al. 2017), which suggests a role of RHBDL2 in epithelial homeostasis. Lastly, a potential role for RHBDL2 in endothelial cells has also been described with the identification of C-type lectin family members thrombomodulin (Lohi et al. 2004) and CLEC14a (Noy et al. 2016) as RHBDL2 substrates.

RHBDL4 is the only mammalian rhomboid protease that has six TMDs. It is implicated in a number of important pathways including ER-associated degradation (ERAD), exosomal secretion and apoptosis (Lespagnol et al. 2008, Wang et al. 2008, Fleig et al. 2012, Ren et al. 2013, Song et al. 2015, Wunderle et al. 2016). The C-terminus of RHBDL4 contains a ubiquitin interaction motif, which facilitates the binding between RHBDL4 and a number of ubiquitinated ERAD substrates, leading to their proteolysis and degradation (Fleig et al. 2012). Fleig et al. showed that RHBDL4 cleaves the α chain of the pre-T cell receptor (pTCRα) in multiple sites, both within the TMD of pTCRα and other regions that are far away from the lipid membrane (Fleig et al. 2012). A possible role in regulating exosomal secretion is proposed based on RHBDL4-driven cleavage and inactivation of TSAP6, which plays a role in exosomal secretion, and the release of full length exosomal TGFα (Lespagnol et al. 2008, Wunderle et al. 2016). Given that RHBDL4 is associated with ERAD, it cleaves ERAD substrates at multiple sites, in a similar manner to TSAP6 cleavage. This raises a question of whether TSAP6 is

also an ERAD substrate. Regarding the possible role of RHBDL-4 in anti-apoptosis, high expression levels of RHBDL4 have been identified in a number of cancer cell lines, while the loss of RHBDL4 is associated with decreased cell proliferation, cancer growth and induced apoptosis in colorectal cancer (Wei et al. 2014, Han et al. 2015, Song et al. 2015).

1.5 functions of non-proteases in the rhomboid-like superfamily

After the discovery of rhomboid proteases, non-proteases were identified within the rhomboid-like superfamily (Greenblatt et al. 2011, Zettl et al. 2011). Due to the absence of crucial catalytic residues, they cannot perform their biological functions through proteolysis. Six clans of rhomboid-like non-proteases have been identified: iRhoms (which form a sub-clan of the secretase rhomboids, due to structural similarity), derlins, UBAC2, TMEM115, RHBDD2 and RHBDD3 (Freeman 2014). The catalytically inactive rhomboid-like proteins regulate distinct signalling pathways, using a variety mechanisms, quite often through regulating trafficking or protein degradation of their binding partners. TMEM115 has also been putatively identified as a rhomboid-like non-protease, although its function has not been extensively studied. Although our current knowledge of the rhomboid-like superfamily is not sufficiently complete to be confident, it is possible to speculate that the core function of the rhomboid-like fold may be TMD recognition.

1.5.1 iRhoms

iRhoms are the best-studied non-proteases in the rhomboid-like superfamily. They regulate essential signalling, including ERAD, TNF-α, EGFR and STING pathways (Zettl et al. 2011, Adrain et al. 2012, McIlwain et al. 2012, Christova et al. 2013, Li et al. 2015, Luo et al. 2016, Grieve et al. 2017). iRhoms have seven TMDs, an extended cytosolic N-terminus, a well-conserved luminal loop between TMD 1 and 2 (named the iRhom homology domain, IRHD) and a short luminal C-terminus (Lemberg et al. 2007). Although they lack the residues necessary for rhomboid protease activity, iRhoms exhibit clear topological and sequence similarities to the active rhomboids (Lemberg et al. 2007).

Drosophila has one iRhom, which is expressed predominantly in neuronal cells (Zettl et al. 2011). Deletion of *Drosophila* iRhom causes a “sleep-like” phenotype in flies that is very similar to mutants with a gain-of-EGFR signalling in the central nervous system (Zettl et al. 2011). Consistent with the iRhom knockout phenotype, genetic interaction assays have revealed that *Drosophila* iRhom counteracts the function of active rhomboids by inhibiting the EGFR signalling pathway. Intriguingly, *Drosophila* iRhom utilises the primary ER protein quality control machinery to regulate this important intercellular signalling pathway: it binds to the membrane-tethered form of EGF family ligands and retains them in the ER, promoting their degradation by ERAD (Zettl et al. 2011).

Two iRhoms are present in mammals, iRhom1 and iRhom2, which appear to have largely redundant cellular functions. iRhom1 is widely expressed, while iRhom2 is predominantly expressed in macrophages, macrophage-like microglia and epithelia (Adrain et al. 2012, McIlwain et al. 2012, Li et al. 2015, Lichtenthaler et al. 2015, Maruthappu et al. 2017). The different tissue expression patterns of iRhom1 and iRhom2 may explain the differences in the phenotype of null mutant mice. iRhom1 knockout mice have multiple defects at birth and prematurely die (Christova et al. 2013). However, Li et al. reported a different phenotype of iRhom1 null mice, which is variable (Li et al. 2015). The inconsistency in phenotypes of iRhom1 mice published by two different groups may be due to different deletion regions when generating iRhom1 null mice, the iRhom1 null mice reported by Li et al. has a shorter deletion region (from exon4 to exon11) than the iRhom1 knockout mice reported by Christova et al. (from exon1 to exon18, the whole exon region of iRhom1) (Christova et al. 2013, Li et al. 2015). iRhom2 knockout mice appear to be healthy but have profound defects in their immune system (Adrain et al. 2012). iRhom2 knockout mice macrophages are unable to produce the primary inflammatory cytokine TNF upon immune challenge such as bacterial infection (Adrain et al. 2012, McIlwain et al. 2012). The observed defects in TNF production in iRhom2 knockout macrophages are caused by the absence of mature TNFα converting enzyme (TACE), which releases active TNFα from its membrane-tethered precursor. iRhoms are essential for TACE activation in

multiple contexts (Adrain et al. 2012, Siggs et al. 2012, Li et al. 2015, Grieve et al. 2017). They localise predominantly in the ER, where they bind to immature TACE and promote its forward trafficking from the ER to the Golgi apparatus, where pro-protein convertases remove the inhibitory prodomain of TACE (Adrain et al. 2012, McIlwain et al. 2012, Issuree et al. 2013). Several recently published suggest that iRhom2 is involved in regulating TACE activity beyond its maturation. For instance, iRhom2 controls the substrate selectivity of TACE (Maretzky et al. 2013). Grieve et al. reported regulatory roles of iRhom2 on TACE activity at the plasma membrane. Once mature TACE is at the cell surface, its activity requires iRhom2 phosphorylation. In addition, the interaction between iRhom2 and mature TACE at the plasma membrane stabilises TACE, by preventing lysosomal degradation (Grieve et al. 2017). The extended cytosolic N-terminus of iRhom2 is crucial for this regulation of TACE (Grieve et al. 2017). Another important binding partner of iRhom2, STING, is a central adaptor for innate immune responses to DNA virus infection (Luo et al. 2016). Luo et al. reported that iRhom2 regulates DNA-virus-triggered induction of type I interferons through promoting the trafficking of STING from the ER to perinuclear microsomes, thereby stabilising STING by promoting the removal of its K48-linked polyubiquitin chains (Luo et al. 2016). Notably, although iRhom2 regulates two distinctive pathways through different binding partners, the fundamental mechanisms of how iRhom2 regulates TACE and STING have striking degrees of similarity.

1.5.2 Derlins

Derlins were first discovered in yeast (Der1) to be essential components of ERAD (Knop et al. 1996) and the role is conserved in mammals (Lilley et al. 2004, Oda et al. 2006). Derlins were not recognised as rhomboid-like proteins originally, and a four TMD topology model of derlins was reported by several groups (Hitt et al. 2004, Lilley et al. 2004). Greenblatt et al. identified that derlins were rhomboid-like proteins in 2011; with a combination of sequence alignment, structural prediction and topology identification, a six TMD topology was reported for Derlin-1 (Greenblatt et al. 2011). With homology identification servers, HHpred (Soding et

al. 2005) and Phyre (Kelley et al. 2009), a high degree of homology was identified between the TMD regions of Derlin-1 and GlpG, an *E.coli* bacteria rhomboid protease with known crystal structure.

The deletion of Der1 in yeast leads to a profound deficiency in the degradation of ERAD luminal substrates, which results in increased ER stress and the activation of UPR signalling (Knop et al. 1996, Taxis et al. 2003, Hitt et al. 2004, Vashist et al. 2004). In mammals, three derlins, Derlin-1 (DERL1), Derlin-2 (DERL2), and Derlin-3 (DERL3), have been identified (Oda et al. 2006). Derlin-1 and Derlin-2 are ubiquitously expressed, while the expression of Derlin-3 is relatively restricted in small intestine, pancreas, spleen and placenta (Oda et al. 2006). All derlins localise within the ER (Lilley et al. 2004). Unlike their homologues in yeast, mammalian derlins are involved in ERAD of both luminal and membrane proteins (Lilley et al. 2004, Oda et al. 2006, Hampton et al. 2012). Mammalian Derlin-1 is the primary candidate for ERAD retrotranslocalisation (Lilley et al. 2004, Ye et al. 2004), with two additional functions, substrate interaction and VCP/p97 recruitment (Lilley et al. 2004, Ye et al. 2004, Greenblatt et al. 2011). The rhomboid-like domain of Derlin-1 is responsible for interacting with ERAD substrates (Greenblatt et al. 2011) and the cytosolic C-terminus of Derlin-1 provides a membrane-binding site in the ER for the recruitment of AAA-ATPase VCP/p97 (Ye et al. 2004), a crucial cytosolic ATPase responsible for extraction of ERAD substrates (Ye et al. 2001). Both Derlin-1 and Derlin-2 are essential components of the Hrd1 ERAD complex, and deletion of Derlin-1 in mice leads to embryonic lethality (Eura et al. 2012). Derlin-2 and Derlin-3 are responsible for the degradation of glycoproteins through ERAD, by cooperating with VCP/p97 and the ER degradation-enhancing alpha-mannosidase-like proteins (EDEM1 and 2) (Oda et al. 2006), which are receptor-like proteins for extracting misfolded glycoproteins for ERAD (Hosokawa et al. 2001).

1.5.3 Other non-catalytic members

The other four non-catalytic members of the rhomboid-like superfamily that have been identified in mammals are RHBDD2, RHBDD3, and UBAC2 (Freeman 2014);

and, confirmed in this thesis, TMEM115. They are poorly studied when compared with iRhoms or derlins. UBAC2 is another ERAD associated member; it is a central element of the gp78 complex for the recruitment of UBXD8 (Christianson et al. 2011). UBAC2 is also known to regulate lipid droplet synthesis, possibly also through the interaction with UBXD8 (Olzmann et al. 2013). RHBDD3 is reported to be a crucial regulator of dendritic cell activation by selectively inhibiting TLR-triggered NF-κB activation and the production of IL6, functioning at least in part via regulating the K27-linked ubiquitination form NEMO (NF-κB essential modulator) (Liu et al. 2013, Liu et al. 2014). RHBDD2 is found misregulated in breast and colorectal cancers, but not much is known about its molecular function (Abba et al. 2009, Lacunza et al. 2012, Lacunza et al. 2014).

1.6 TMEM115

TMEM115 (also called PL6) is a polytopic membrane protein that is well conserved across evolution (present from yeast to human). It was first identified as a potential tumour suppressor gene (Ivanova et al. 2008), on the basis that it is localised on chromosome 3p21.3, which is the critical region involved in several common cancers (Martinez et al. 2000). Bioinformatic analyses suggest that TMEM115 is a distant member of the rhomboid-like superfamily. However, the only paper about TMEM115 topology presented a four TMD model (Ong et al. 2014), which is inconsistent with it being placed in the rhomboid-like superfamily, given that all rhomboid-like proteins have at least six TMDs. The only common result between the two papers published about TMEM115 is its Golgi localisation (Ivanova et al. 2008, Ong et al. 2014), despite slight differences in the reported subcellular localisation. Ong et al. reported that TMEM115 co-localises well with medial-Golgi and trans-Golgi markers, while Ivanova et al. showed co-localisation between TMEM115 with a cis-Golgi marker, GPP-130. Preliminary data from the Freeman lab have shown that in HeLa cells, the endogenous TMEM115 signal is co-localised best with COPI vesicles and the cis-Golgi marker, GM130.

Regarding TMEM115 function, Ong et al. reported some physical interactors for TMEM115 identified by co-immunoprecipitation, including β-COP, ERGIC53

and COG3 (Ong et al. 2014). Another *in-vitro* binding assay in the same paper showed that TMEM115 can bind to all eight subunits of the COG complex, with the highest binding affinity to COG4 (Ong et al. 2014). Given that both the COG complex and β-COP are key players in mediating Golgi-to ER retrograde trafficking (Oka et al. 2004), the COG complex is also crucial for regulation of O-linked glycosylation (Smith et al. 2008). In line with these activities, Ong et al. reported a role of TMEM115 in the regulation of retrograde trafficking from the Golgi apparatus and O-linked glycosylation (Ong et al. 2014). They showed that TMEM115 is required for BFA-induced Golgi-to-ER retrograde transport. Over-expressing or knockdown of TMEM115 delayed BFA-induced retrograde transport in HeLa cells (Ong et al. 2014). They also reported that silencing of *TMEM115* alters O-linked glycosylation (Ong et al. 2014).

Preliminary data from the Freeman lab and collaborators suggest that TMEM115 may have important cellular functions linked with lipid biology. In yeast, our collaborators in the lab of Maya Schuldiner at the Weizmann Institute identified five lipid-related proteins that are interactors with TMEM115. FAT1 and AYR1 were identified in a genetic interaction screen: they were mislocalised in the absence of TMEM115. The other three proteins, NEO1, LPP1 and CPT1, were identified as physical interactors of TMEM115. Among the five lipid biology related interactors of TMEM115, FAT1 and NEO1 are responsible for lipid trafficking. FAT1 transports very long chain fatty acyl-CoA synthesis and long chain fatty acid (Faergeman et al. 1997). In yeast, NEO1 flips phosphatidilserine (PS) and phosphatidilethanolamine (PE) from the outer membrane to the inside and releases PE to the cytosol (Hua et al. 2003). ATPase II is the human homologue of NEO1; it is a group of ion pumps as well as PS flippers. The rest three lipid-related interactors of TMEM115, AYR, LPP1 and CPT1 have roles in lipid biosynthesis. AYR1 is a NADPH-dependent 1-acyl dihydroxyacetone phosphate reductase, which is involved in phosphatidic acid biosynthesis (Athenstaedt et al. 2000). LPP1 is a lipidphosphate phosphatase. It catalyzes the dephosphorylation of diacylglycerol phosphate (DGPP) to phosphatidate (PA) and the subsequent dephosphorylation of PA to diacylglycerol (DAG) (Tomsig et al. 2009). It has three

human homologues, LPP1, 2 and 3. And finally CPT1 is involved in the last step of phosphatidylcholine (PC) synthesis (Morash et al. 1994).

The biological importance of TMEM115 is further supported by phenotypes of null mutant animals. Both CG9536 (the TMEM115 orthologue in Drosophila) knockout files and *Tmem115* knockout mice have severe phenotypes. The absence of CG9536 in flies have reduced longevity, together with reduced activity. Regarding *Tmem115* knockout mice, they are much smaller than wild type mice and have pre-weaning lethality. Soon after birth, *Tmem115* knockout animals already exhibited significantly lower weight than wild type or heterozygous littermates, a condition maintained throughout their short lives. The few homozygous mutant animals that survived beyond 14 days failed to gain weight, started to lose body mass, and eventually died.

When I first started this project, no biological role was identified for TMEM115; we started from the observed severe phenotypes of TMEM115 mutant animals. During my project, a role of TMEM115 in lipid homeostasis has been identified by members in Freeman lab (unpublished data). Lipid homeostasis is an essential machinery of metabolic control, and the disruption of this machinery is linked to a wide range of diseases (Grundy 2004, Agmon et al. 2017). In vertebrates, lipid biosynthesis is regulated by the multipass protein SCAP and the SREBP transcription factors (Nohturfft et al. 1999, Horton et al. 2002). A feedback circuit tightly controls the SCAP-SREBP pathway, which senses changes in nutrient levels and responds acutely (Botolin et al. 2003). In low lipid environments, membrane-bound forms of SREBPs are transported together with SCAP from the endoplasmic reticulum (ER) to the Golgi apparatus, where the active transcription factor portion of SREBPs are released from their membrane anchors upon cleavage by site-1 and site-2 proteases (Horton et al. 2002). I will discuss the background of lipid homeostasis in more detail in Chapter 5.

Given that TMEM115 is an important but not well-studied protein, I decided to focus on it for my DPhil project. Much of the work described in this thesis was done by me alone but the experiments in Chapter 5 on identifying the biological

role of TMEM115 in regulating lipid homeostasis, were performed in collaboration with Dr Angela Moncada-Pazos, a postdoc in the Freeman lab. Dr Moncada-Pazos is investigating the physiological role of TMEM115 using animal models, with a particular focus on lipid biology. The aim of my DPhil project was to determine whether TMEM115 is indeed a bona-fide rhomboid-like protein and to use biochemical approaches to discover the biological role of this very highly conserved Golgi membrane protein. The specific questions I have addressed are as follows.

1. By identifying protein binding partners of TMEM115, can I shed light on its fundamental cellular role?

2. What is the topology of TMEM115 and is it a true rhomboid-like protein?

3. Given what I learned from the protein interaction screen, combined with results from Dr Moncada-Pazos, what is the relationship between TMEM115 and the SCAP-SREBP pathway?

I used a proteomic approach to identify novel interactors for TMEM115 under both basal and low lipid conditions. This identified multiple binding partners that indicate possible roles of TMEM115 in lipid biology, trafficking, protein degradation and the regulation of ion channels. I then validated and pursued the best hits to allow myself to focus on the most promising leads. One binding partner of TMEM115, p62, was validated and pursued for a possible role in the regulation of the SREBP-SCAP pathway. In addition, I performed a topology study for TMEM115, and with a combination of experimental methods and bioinformatic analyses, proved that in contrast to the previous topology reported by Ong et al., it indeed has a six TMD structure, consistent with it being a rhomboid-like protein. Finally, I investigated the possible relationship between TMEM115, SCAP and p62; the results suggest that TMEM115 may be involved in regulating lipid homeostasis by modulating SCAP levels through the proteasomal degradation machinery. In the final chapter of my thesis, I discuss the overall implications of my work and the future directions that will need to be followed to confirm and extend my conclusions.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 Reagents

For cell culture

Bovine serum albumin (BSA) (Sigma, A7906); tetracycline-free bovine serum albumin (Clontech, 631106); oleic acid-albumin conjugated (Sigma, O3008); mevastatin (Sigma, M2537); Dulbecco's Modified Eagle's medium (DMEM, Sigma, D5796); Penicillin Streptomycin (Life technologies, 15140122); L-glutamine (Life technologies, 25030-024); Trypsin (Life technologies, 12604-013); 25-hydroxycholesterol (Sigma H1015); puromycin (Life technologies, A1138-03); PBS (Lonza, BE17-516F); MG132 (Sigma, 474787); bafilomycin (Sigma, B1793); polyethyleneimine, linear, mw=25,000 (Pei, Polysciences 23966); Fugene HD(Promega, E2312); Lipofectamine RNAimax (Thermo Fisher, 13778); siRNA for human TMEM115 (Thermo Fisher, HSS117142); biotin (Sigma, B4501); siRNA for human p62 (dharmacon, L-010230-00-0005); doxycycline (Sigma, D9891); tetraacetyl-N-azidoacetyl mannosamine (ManNAz, Life technologies, C33366); DBCO-PEG-Biotin (Life technologies,); Cycloheximide (Sigma, C104450).

For cloning

In-Fusion HD cloning kit (Clontech, 121416); Gateway™ LR Clonase™ II Enzyme mix (Thermo Fisher, 11791100); Platinum Taq DNA Polymerase (Invitrogen 100021274); T4 ligase (NEB, M0202); Pfu polymerase (Agilent Technology, 600136); DpnI (New England Biolabs, R0176L); All the restriction enzymes were from New England Biolabs. Human cDNA libraries were provided by Ulrike Kuenzel. All the primers were ordered from Sigma. All the primers that have been used in this thesis are listed in table 2.1.

Table 2.1 list of primers

List of primers of BioID	
N-TMEM115 for Tet-one FW	GGAGATCTTTACAGCGTCGGGGGAGC
N-TMEM115 for Tet-one RV	CTCTTCCTCATCGAGTAGCTCTATGCAA CGTGCCCTGCCAG
N-BirA* for Tet-one FW	AGAGCTACTCGATGAGGAAGACTCGAG CTTCTCTGCGCTTC
N-BirA* for Tet-one RV	GGGAATTCATGGAACAAAAACTCATCTC AGAAG
BirA*-TMEM115 for pLenti-entry Flag FW	AAGCAGGCTCCACCATGGATGGAACAAA AACTCATCTCAGAAGAG
BirA*-TMEM115 for pLenti-entry Flag RV	GTGCGGCCGCGAATTCTTACAGCGTCG GGGGAGC
BirA* for pLenti-entry Flag FW	AAGCAGGCTCCACCATGGATGGAACAAA AACTCATCTCAGAAG
BirA* for pLenti-entry Flag RV	GTGCGGCCGCGAATTCCTACTCGAGCTT CTCTGCGC
C-TMEM115 for pLenti-entry Flag FW	CAAAAAAGCAGGCTCCACCATGGATGCA ACGTGCCCTGC
C-TMEM115 for pLenti-entry Flag RV	AGAGCTACTCGATGAGGAAGACAGCGT CGGGGGAGCTG
C-BirA* for pLenti-entry Flag FW	TCTTCCTCATCGAGTAGCTCTATGGAAC AAAAACTCATCTCAGAAG
C-BirA* for pLenti-entry Flag RV	GATATCTCGAGTGCGGCCGCGAATTCCT

	ACTCGAGCTTCTCTGCGC
Cloning for BioID list validation	
ERGIC-53 FW	AAGCTTGGTACCGAGCTCGGATCCATGG CGGGATCCAGGCAAAG
ERGIC-53 RV	TCATTTATCATCGTCATCCTTATAATCCT TGTCGTCATCGTCTTTGTAGTCAAAGAAT TTTTTGGCAGCTGCTTC
SEQUESTOSOME 1 FW	AAGCTTGGTACCGAGCTCGGATCCATGG CGTCGCTCACCGTGA
SEQUESTOSOME 1 RV	TCATTTATCATCGTCATCCTTATAATCCT TGTCGTCATCGTCTTTGTAGTCCAACGG GGGATGCTTT
RAB7A FW	AAGCTTGGTACCGAGCTCGGATCCATGA CCTCTAGGAAGAAAGTGTTG
RAB7A RV	TCATTTATCATCGTCATCCTTATAATCCT TGTCGTCATCGTCTTTGTAGTCGCAACT GCAGCTTTCTGCCG
list of primers for topology study	
Myc tag system	
TMEM115 position 1 FW	CTGGACCCTGGCCACCCATGAGCAGAA ACTCATCTCTGAAGAGGATCTGGGGCTG ATGGAGCAGCATG
TMEM115 position 1 RV	CATGCTGCTCCATCAGCCCCAGATCCTC TTCAGAGATGAGTTTCTGCTCATGGGTG GCCAGGGTCCAG

TMEM115 position 2 FW	CTCACCTACATGGCTTCCTTCGAGCAGAAACTCATCTCTGAAGAGGATCTGAACCTGGTCTACCTGTTCAC
TMEM115 position 2 RV:	GTGAACAGGTAGACCAGGTTCAGATCCTCTTCAGAGATGAGTTTCTGCTCGAAGGAAGCCATGTAGGTGAG
TMEM115 position 3 FW	GGACTGTGTGGTCCTGGAGCAGAAACTCATCTCTGAAGAGGATCTGCGAGTGCCCCAGGTGCGCGTCAG
TMEM115 position 3 RV	CTGACGCGCACCTGGGGCACTCGCAGATCCTCTTCAGAGATGAGTTTCTGCTCCAGGACCACACAGTCC
TMEM115 position 4 FW	CACGCTGCTCCAGAGCCCGGAGCAGAAACTCATCTCTGAAGAGGATCTGGCGCTGGCTTCCTATGGC
TMEM115 position 4 RV	GCCATAGGAAGCCAGCGCCAGATCCTCTTCAGAGATGAGTTTCTGCTCCGGGCTCTGGAGCAGCGTG
TMEM115 position 5 FW	GCCATAGCCGGGGCCGAGAGCAGAAACTCATCTCTGAAGAGGATCTGGGGGACATGGCTGACC
TMEM115 position 5 RV	GGTCAGCCATGTCCCCCAGATCCTCTTCAGAGATGAGTTTCTGCTCTCGGCCCCGGCTATGGC
Flag tag system	
TMEM115 flag position 1 FW	CTGGACCCTGGCCACCCATGACTACAAAGACGATGACGACAAGGGGCTGATGGAG

	CAGCATG
TMEM115 flag position 1 RV	CATGCTGCTCCATCAGCCCCTTGTCGTCATCGTCTTTGTAGTCATGGGTGGCCAGGGTCCAG
TMEM115 flag position 2 FW	CTCACCTACATGGCTTCCTTCGACTACAAAGACGATGACGACAAGAACCTGGTCTACCTGTTCAC
TMEM115 flag position 2 RV:	GTGAACAGGTAGACCAGGTTCTTGTCGTCATCGTCTTTGTAGTCGAAGGAAGCCATGTAGGTGAG
TMEM115 flag position 3 FW	GGACTGTGTGGTCCTGGACTACAAAGACGATGACGACAAGCGAGTGCCCCAGGTGCGCGTCAG
TMEM115 flag position 3 RV	CTGACGCGCACCTGGGGCACTCGCTTGTCGTCATCGTCTTTGTAGTCCAGGACCACACAGTCC
TMEM115 flag position 4 FW	CACGCTGCTCCAGAGCCCGGACTACAAAGACGATGACGACAAGGCGCTGGCTTCCTATGGC
TMEM115 flag position 4 RV	GCCATAGGAAGCCAGCGCCTTGTCGTCATCGTCTTTGTAGTCCGGGCTCTGGAGCAGCGTG
TMEM115 flag position 5 FW	GCCATAGCCGGGGCCGAGACTACAAAGACGATGACGACAAGGGGGACATGGCTGACC
TMEM115 flag position 5 RV	GGTCAGCCATGTCCCCCTTGTCGTCATCGTCTTTGTAGTCTCGGCCCCGGCTATGG

	C
Multi Myc tag system	
TMEM115 position 1 FW	CTGGACCCTGGCCACCCATGACTACAAAGACGATGACGACAAGGACTACAAAGACGATGAC
TMEM115 position 1 RV	GTCATCGTCTTTGTAGTCCTTGTCGTCATCGTCTTTGTAGTCATGGGTGGCCAGGGTCCAG
TMEM115 position 2 FW	CTCACCTACATGGCTTCCTTCGACTACAAAGACGATGACGACAAGGACTACAAAGACGATGAC
TMEM115 position 2 RV:	GTCATCGTCTTTGTAGTCCTTGTCGTCATCGTCTTTGTAGTCGAAGGAAGCCATGTAGGTGAG
TMEM115 position 3 FW	GGACTGTGTGGTCCTGGAGCAGAAACTCATCTCTGAAGAGGATCTGGAGCAGAAACTC
TMEM115 position 3 RV	GAGTTTCTGCTCCAGATCCTCTTCAGAGATGAGTTTCTGCTCCAGGACCACACAGTCC
TMEM115 position 4 FW	CGCTGCTCCAGAGCCCGGAGCAGAAACTCATCTCTGAAGAGGATCTGGAGCAGAAACTC
TMEM115 position 4 RV	GAGTTTCTGCTCCAGATCCTCTTCAGAGATGAGTTTCTGCTCCGGGCTCTGGAGCAGCG

TMEM115 position 5 FW	GCCATAGCCGGGGCCGAGACTACAAAGACGATGACGACAAGGACTACAAAGACGATGAC
TMEM115 position 5 RV	GTCATCGTCTTTGTAGTCCTTGTCGTCATCGTCTTTGTAGTCTCGGCCCCGGCTATGGC
TMEM115 position 6 FW	CAGATCCTCTTCAGAGATGAGTTTCTGCTCCAGATCCTCTTCAGAGATGAGTTTCTGCTCGCTGCTGCCGCTGCCGCAGCCTGTGTCCACGGC
TMEM115 position 6 RV	GAGCAGAAACTCATCTCTGAAGAGGATCTGGAGCAGAAACTCATCTCTGAAGAGGATCTGGGCAGCGGCAGCTGGCGGT
list of primers for TMEM115 truncations	
TM1-205 FW	ATGTCCCAGACTACGCAGAATTCATGCAACG
TM1-205 RV Step 1	TTAAAGATATACCCAACTGGAGAG
TM1-205 RV Step 2	AACGGGCCCTCTAGACTCGAGTTAAAGATATACCCAACTGGAGA
TM1-229 FW	ATGTCCCAGACTACGCAGAATTCATGCAACG
TM1-229 RV Step1	TTAAGGGAAGAAAGTGGCGAAA
TM1-229 RV Step2	AACGGGCCCTCTAGACTCGAGTTAAGGGAAGAAAGTGGCGAAA

For immunofluoresce

Paraformaldehyde (PFA) solution 16% (Electron Microscopy Sciences, 15710); digitonin (Sigma, D141); Triton X-100 (Sigma, X-100/1L); fish gelatin solution (Sigma, G7765) and DAPI mounting solution (eBioscience, 00-4959-52)

For western blot

Complete protease inhibitors cocktail (Roche, 11836145001); DTT (Sigma, D0632); LDS sample buffer (4X, life technologies, NP0008); NuPAGE Bis-Tris Precast Gels (life technologies, NP0336BOX, NP0321BOX, NP0322BOX); dried skimmed milk powder (Marvel, Tesco ID: 254902569); tween-20 (Sigma, P1379); 10XPBS (Sigma, D1408).

For immunoprecipitation

Pierce™ Anti-HA Magnetic Beads (Life technologies, 88837); Dynabeads® MyOne™ Streptavidin C1 (Life technologies, 65002); DynaMag™-2 Magnet rack (Life technologies, 12321D)

2.1.2 Antibodies

Rabbit anti-TMEM115 (Atlas Antibodies, HPA015497); rabbit anti-βCOP (Abcam, ab2899); mouse anti-VCP/p97 (Abcam, ab111433); rabbit anti-p62 (GeneTex, GTX111393); goat anti-SCAP (Santa Cruz, sc-9675); mouse anti-ERGIC53 (anti-LMAN1, Novus Biologicals, NBP2-03381); mouse anti-GM130 (BD Biosciences, 610823); mouse anti-β-actin (Sigma, A3854); Rabbit anti-Myc (Santa Cruz); rat anti-HA (Roche, 12013819001); sheep anti-TGN46 (Serotec, AHP500); chicken anti-GFP (Abcam, ab13970); Goat anti-Myc HRP (Abcam, ab1261); rat polyclonal anti-HA HRP (Roche, 12013819001); chicken anti-BirA (Abcam, ab14002); rabbit anti-BirA (Creative diagnostics, DPAB/PT1113RH), Neutravidin HRP (Sigma, Sigma-GERPN1231). All secondary antibodies for western blot were HRP conjugated from Santa Cruz and cell signalling. All secondary fluorescent antibodies were Alexa Fluor-conjugated from Thermo Fisher.

2.1.3 Constructs

3XHA-TMEM115 (N-terminally triple HA-tagged TMEM115) and TMEM115-3XHA (C-terminally triple HA-tagged TMEM115) were cloned by Dr Christian Luginsland. Myc-BirA* in pcDNA 3.1 construct was provided by Dr Kyojiro Ikeda. Tet-One construct was purchased from Clonetech (634301). BirA* in Tet-One vector was cloned by Dr Adam Grieve. HA-P62 (human p62 with N-terminally tagged HA) was purchased from Addgene (Plasmid 28027). Flag-SCAP (human SCAP with N-terminally tagged Flag) was a kind gift from Johan Swinnen. Flag p-lenti entry-Flag vector, plenti-dest-CMV/TO puro vector, pCMV-VSV-G and pCMV-dR8.2 dvpr were kind gifts from Dr Ian Gibbs-Seymour. I cloned all the rest constructs that have been described in this thesis.

2.2 Methods

2.2.1 Bioinformatics

The sequences of human, mouse, *Drosophila* and yeast TMEM115 were obtained from the UniProt protein database. Clustal Omega was used for aligning the protein sequences among TMEM115 orthologues. TMD regions were predicted using the algorithms Toppred, SOSUI, TMHMM 2.0 and MEMSAT. Phyre2 protein folding prediction server was used to predict the 3D structure of human TMEM115. HHpred protein structure and function prediction server was used to identify homologues of TMEM115 TMD region and for searching possible functional domains at the C-terminus of TMEM115.

2.2.2 Cloning

For the study of TMEM115 topology, using 3XHA-TMEM115 plasmid as a template cloned Pfu polymerase (Agilent Technology) was used to insert one Myc (EQKLISEEDL) or Flag (DYKDDDDK) epitope tag at each of the six insertion sites, independently. Dpn1 (NEB) was used to digest the original template construct after each mutagenesis PCR at 37 °C for one hour. For inserting multiple Myc tags in each of the five positions, a second step PCR was used to create two pieces of truncated TMEM115 with compliment sequences to the pcDNA vector at both

termini, and double Myc tag at the designed tag region. The two truncated pieces of TMEM115 were inserted into pcDNA3.1 vector between EcoRI/XhoI with InFusion cloning kit. The Dpn1 digested PCR products or InFusion reaction products were then transformed into chemically competent cells (Bioline, Golden efficiency) and plated on agar plates containing 100 µg/ml Ampicillin for overnight culture. Single colonies were picked up, and each was cultured in 2 mL of LB medium (with 100 µg/mL Ampicillin) overnight. Plasmids were isolated from the overnight culture using QIAGEN spin miniprep kit and sent for sequencing (SourceBiosciences).

For BioID, full length Myc-BirA* and TMEM115 cDNAs were amplified by PCR and linked by overlapping PCR with a flexible linker (SSSSSSS) using Platinum Taq polymerase (Life technologies). Myc-BirA*-TMEM115 was cloned into the empty Tet-One vector (Clontech) between EcoRI/BglII with T4 ligase (NEB). The ligation product was transformed into chemically competent cells and processed as described above for plasmids purification and sequencing. For the lentiviral system, BirA*-TMEM115, BirA* and TMEM115-BirA* were cloned into pEntry-Flag vector between NcoI/EcoRI. The LR recombination was performed using the Gateway™ LR Clonase™ II Enzyme mix (Thermo Fisher) with 1 µl of miniprep DNA for each of the Entry and Destination vector, 4 µl of TE pH 8.0, 2 µl of LR buffer and 2 µl of LR clonase. Reactions were incubated at room temperature from 2 h to overnight. The proteinase K digestion step was omitted from the manufacturer's protocol, and 2 µl of the reaction were used to transform 25 µl of competent *E.coli* Stbl3 (homemade).

In-Fusion cloning kit was used to clone truncated forms of TMEM115 (3XHA-TM1-205 and 3XHA-TM1-229) and the candidate proteins obtained from the BioID screen. Genes of interests were amplified with Platinum Taq DNA Polymerase from the HEK293t cDNA library, which was kindly provided by Ulrike Kuenzel. 2X Flag tags were added at the C-terminus of each candidate, and 20 bp of compliment sequences to the pcDNA vector were added at both termini of the genes of interest to facilitate the following steps to fuse to linearised pcDNA3.1

between EcoRI/Xhol sites.

2.2.3 Cell culture

HeLa cells, HEK293T cells, wild type and *Tmem115* knockout immortalised MEFs, and MCF-7 cells were cultured in DMEM with 10% FBS, 1% Penicillin-Streptavidin and 1% L-Glutamine (all Life Technologies). For BioID approach, MCF-7 cells were cultured in tetracycline-free (Tet-free) medium for the initial tests. For evaluation of the effect of lipid depletion, cells were cultured in DMEM with 0.5% FBS plus 300 µM oleic acid and 1 µg/ml 25-hydroxycholesterol (+ lipids, non-inducing) or in DMEM with 0.5% FBS plus mevastatin 0.5 µM (- lipids, inducing) for 16 hours.

2.2.4 Transfection and transformation

FuGENE transfection

For transfection of HeLa and MCF-7 cells (BioID transient transfection in Section 3.1 and topology mapping in Section 4.4-4.7), 1 µg of DNA and 4 µl of FuGENE HD (Promega) were used for transfecting each single well of a 6 well plate. 0.2 µg of DNA and 0.8 µl of FuGENE HD were used for 8 well culture chambers. DNA and FuGENE HD were diluted in serum-free medium (200 µl for 6 well dish and 20 µl for 8 well culture chamber) and incubated for 15 minutes before adding to cells.

PEi transfection

For transfection of MEFs in 6 well/ 10cm plates, 2/ 10 µg of DNA and 6/ 30 µL of a 1 mg/ml PEi solution were mixed in 250/500 µl of serum-free medium. For transfection of HeLa/ HEK293T cells in 10 cm tissue culture plates, 2.5 µg of DNA and 7.5 µl of a 1 mg/ml PEi solution (2.5 µg per construct and 15 µl Pei for con-transfection) were mixed in 500 µl of serum-free medium. For full scale BioID with transient transfection protocol, 15 µg of DNA and 45 µl of Pei were added per 15 cm dish. The transfection complex was incubated for 15 minutes and added to cells in serum-free medium. 5 h after transfection, serum-free medium was replaced with fresh complete medium.

SiRNA knockdown

Lipofectamine RNAimax was used for knocking down *Tmem115* in wild type MEFs according to manufacturer's instructions; two rounds of transfection were made in a total time of 96 hours. MEFs were kept in low lipid medium for the last 16 hours before harvest.

For knocking down *p62* in Hela cells, two rounds of transfection were made with Lipofectamine RNAimax in a total time of 48 hours. HeLa cells were kept either in DMEM with 0.5% FBS plus 300 μM oleic acid and 1 μg/ml 25-hydroxycholesterol (+ lipids, non-inducing) or in DMEM with 0.5% FBS plus mevastatin (- lipids, inducing) for the last 16 hours before harvest.

Transduction

For the production of the lentiviruses for Tet-repressor, BirA*-TMEM115, BirA* and TMEM115-BirA*, $2x10^6$ of HEK293T cells were seeded per 10cm plate 24 hours prior transfection in antibiotic-free full medium. HEK293T cells were co-transfected with the corresponding plenti-destination constructs (5 μg per 10 cm dish), together with two packaging vectors pCMV-VSV-G (1 μg per 10 cm dish) and pCMV-dR8.2 (5 μg per 10 cm dish). The plasmids described above were mix with 460 μl of serum-free medium and 20 μl FuGENE, followed by the protocol described before. 15 hours later, full medium with transfection complex was replaced with antibiotic-free full medium. The generated recombinant viruses were collected 48 hours later to infect MCF-7 cells. To generate a tetracycline inducible cell line, Tet- repressor was transformed first into MCF-7, followed by seven days of blasticidin selection (15 μg/ml) for obtaining a stable cell line for Tet-repressor. The Tet-repressor stable cell line was used as the target cell line for generating BirA*-TMEM115, BirA* and TMEM115-BirA* stable cell lines. After virus transduction, 5 μg/ml of puromycin was used for seven days for obtaining the described stable cell lines.

2.2.5 SDS-PAGE and western blot

Cells were lysed with RIPA buffer containing Complete protease inhibitors cocktail (Roche). Protein samples were heated at 65 ^oC for 10 minutes with LDS sample

buffer (4X, Life technologies). Proteins were separated by SDS-PAGE with 4%-12% Bis-Tris gradient gel and MOPS running buffer (Life technologies) at 150V. Proteins were transferred to polyvinylidene difluoride (PVDF) membranes (Millipore) or nitrocellulose membranes at 100V for 90 minutes. Membranes were blocked in PBST (PBS with 0.05% Tween 20) and 5% milk (or with 5% BSA for Neutravidin-HRP blot), before detection with the indicated primary antibodies and species-specific HRP-coupled secondary antibodies. Western-blot protocol was then applied to analyse protein samples using anti-Myc HRP, goat polyclonal (1 in 30,000 dilution, Abcam); anti-HA HRP, rat polyclonal (1 in 2000 dilution, Roche); anti-BirA, chicken polyclonal (1 in 2000 dilution, Abcam); anti-BirA, rabbit polyclonal (1 in 4000 dilution, Creative diagnostics); Neutravidin-HRP (1 in 2000 dilution in 5% BSA, Sigma-Aldrich); Rabbit anti-TMEM115 (1 in 500 dilution, Atlas Antibodies); rabbit anti-βCOP (1 in 1000, Abcam); mouse anti-VCP/p97 (1 in 2000, Abcam); rabbit anti-p62 (1 in 2000, GeneTex); goat anti-SCAP (1 in 500 for lysate and 1 in 200 for co-immunoprecipitation eluates, Santa Cruz, sc-9675); mouse anti-ERGIC53 (anti-LMAN1, 1 in 1000, Novus Biologicals); mouse anti-β-actin-HRP (1 in 30,000, Sigma); All the secondary antibodies from Santa Cruz and cell signalling were used as 1 in 5000 dilution. All the antibodies, unless specially described, were diluted with 5% milk in PBST. Band visualisation was achieved with Enhanced Chemiluminescence using X-ray films.

2.2.6 Immunofluorescence staining and confocal microscopy

For determining TMEM115 membrane topology by conditional permeabilisation, $1x10^4$ HeLa or MCF-7 cells were seeded per well in an 8 well culture chamber (Corning) 24 hours before transfection. HeLa cells were transiently transfected using FuGENE HD according to the manufacturer's instructions. 0.2 µg of DNA were used in 8-chamber glass slides. The slides were collected 24 hours post-transfection and washed with warm PBS. Cells were fixed with 4% PFA for 15 minutes and washed with PBS for 3 times. Cells were either treated with 0.2% TX-100 for 15 min to permeabilise all membranes or 25 µg/mL digitonin in PBS for 5 min to permeabilise the plasma membrane only and washed with PBS for 3

times. Cells were then incubated in blocking buffer (1% fish gelatin in PBS) for 1 h at room temperature followed by incubation with primary antibody (5 hours) and secondary antibody (45 minutes), each followed by three washes with blocking buffer. Primary antibodies were used at a dilution of 1:300 for anti-cMyc (Santa Cruz), 1:500 for anti-TGN46 (Serotec) and 1:1000 for anti-HA (Roche). Fluorescent secondary antibodies were used at 1:500 dilution. Cells were mounted onto slides using mounting medium containing DAPI. All the immunofluorescence images were obtained with an inverted Olympus FV1000 confocal system with a 60x/1.40 Oil UPlanSApo objective.

For testing the cellular location of TM1-205 and TM1-229, HeLa cells were transfected with PEi and cultured for 24 hours. Slides were prepared as the protocol described above except for the permeabilisation step. HeLa cells were treated with only 0.2% TX-100 for 15 minutes. Primary antibodies were used at a dilution of 1:1000 for anti-GM130 and 1:1000 for anti-HA. Fluorescent secondary antibodies were used at 1:500 dilutions. Images were acquired with the same confocal microscope.

For testing the cellular location of BirA*-TMEM115, MCF-7 cells were transfected with FuGene and cultured for 24 hours in tetracycline-free medium under four different conditions: no supplements, supplemented with 50 μM biotin only, supplemented with 1 μg/mL of Doxycycline only or supplemented with both 1 μg/mL of Doxycycline and 50 μM biotin. For testing the cellular location of BirA*-TMEM115 and TMEM115-BirA* in MCF-7 stable cell lines, MCF-7 cells were plated in 8 well chamber and cultured for 24 hours in full medium under two different conditions: no supplements or supplemented with both 1 μg/mL of Doxycycline and 50 μM biotin. For optimising transient BioID condition, HEK293T cells were transfected with Pei (0.2 μg DNA and 0.6 μg Pei per well of the 8 well chamber) and cultured for 24 hours in full medium supplemented with 50 μM biotin under five different conditions: no supplements, supplemented with 10, 20, 40 or 1000 ng/mL of Doxycycline. Slides were prepared as the protocol described above. Primary antibodies were used at a dilution of 1 in 1000 for both anti-BirA (0.55

µg/mL, chicken polyclonal, Abcam) and anti-GM130 (0.25 µg/mL, mouse monoclonal, BD Science). For secondary antibodies, neutravidin-568 was used at 1 in 1000 (1 µg/mL Thermo Scientific) and Alexa Fluor 488 or 647 (4 µg/ml, Life Technologies) were used at 1:500 dilution. Images were acquired with the same confocal microscope.

2.2.7 Immunoprecipitation

For conventional immunoprecipitations (IPs), HeLa/ HEK293T/ wild type/ *Tmem115* knockout MEFs cells were plated in 10 cm tissue culture plates and transiently transfected using PEi as described above. 48 hours after transfection, cells were placed on ice and washed 3x with ice-cold PBS and then lysed in 1 ml TX-100 lysis buffer (1% Triton X-100, 150 mM NaCl, 50 mM Tris-HCl, pH 7.4) supplemented with protease inhibitor cocktail (Roche). Cell lysates were cleared by centrifugation at 10,000 x g for 10 min at 4°C. The lysates were then immunoprecipitated overnight with 20 µg pre-washed HA antibody-coupled beads at 4°C on a rotor. After 5 washes with lysis buffer, the immunocomplexes were incubated at 65°C for 10 min in 1x LDS sample buffer. Typically, 75% of the immunoprecipitates and 1% of lysates were resolved on SDS-PAGE gels for subsequent western blotting.

2.2.8 BioID protocol for mass spec analysis

Cells were incubated for 24 h in complete media supplemented with 1 µg/ml doxycycline and 50 µM biotin. After three PBS washes, cells (for small-scale analysis, one 15 cm dish, $2x10^7$; for large scale analysis, 3 15 cm dish, 6×10^7) were lysed at 25°C in 2 ml lysis buffer per dish (50 mM Tris, pH 7.4, 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM DTT, and 1x Complete protease inhibitor (Roche) and sonicated (20% for 10s). Triton X-100 was added to 2% final concentration. After further sonication, three equal volume of 4°C 50 mM Tris (pH 7.4) was added before additional sonication (15% for 5s X4 times, subsequent steps at 4°C) and centrifugation at 16,000 relative centrifugal force. Supernatants were incubated with 100 µl Dynabeads per 15 cm dish of cell lysate (MyOne Streptavidin C1; Invitrogen) overnight. Beads were collected and washed twice for 8 min at 25°C

(all subsequent steps at 25°C) in 1 ml wash buffer 1 (2% SDS in dH2O). This was repeated once with wash buffer 2 (0.1% deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, and 50 mM Hepes, pH 7.5), once with wash buffer 3 (250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, and 10 mM Tris, pH 8.1) and twice with wash buffer 4 (50 mM Tris, pH 7.4, and 50 mM NaCl) (six times with wash buffer 4 for samples prepared for mass spec analysis). Bound proteins were removed from the magnetic beads with 50 µl of SDS elution buffer per dish (10X elution buffer: 1M Tris-HCl pH6.8 60ml, SDS 20g, 2-ME 4ml, 100mM 96ul, add H2O to 100ml final volume) with 50uM biotin at 98°C for 10 minutes. 5% of the beads eluate was reserved for Western blot analysis. 95% of the beads eluate was reserved for silver staining or mass spectrometry analysis.

2.2.9 Quantification and statistical analysis

All statistical analysis was done using GraphPad prism software. The specific tests performed in each section are indicated in the main text and figure legends.

2.2.10 SPECS approach and mass spectrometry

40 million of wild type or *Tmem115* KO MEFs were plated in 2 of the 15 cm dish 24 hours before treatment. For metabolic sugar labelling, MEFs were incubated for 48 hours in 15ml DMEM supplemented with 10% FCS (or Opti-MEM) and 50 µM of tetraacetyl-N-azidoacetyl mannosamine (ManNAz, Life technologies) for each dish. Conditioned media were collected and filtered through 0.22 mm PVDF filter (Millex) into a VivaSpin 20 ultrafiltration unit (cutoff, 30 kDa) at 4 °C. VivaSpin 20 columns were centrifuged at 4500 rpm at 4 °C until the retentate was 0.5ml to remove nonmetabolized ManNAz. The retentate was filled with 20 ml autoclaved ddH_2O and centrifuged at 4500 rpm at 4 °C. This procedure was repeated three times until the retentate was 0.5ml at the last time. 2µl of DBCO-PEG12-biotin (50mMClick-chemistry tools) was diluted in 0.5ml of 0.2% formic acid and added to the retentate for biotinylating azido-labelled glycoproteins. Columns were incubated overnight at 4 °C for the click chemistry reaction. To remove the unreacted DBCO-PEG12-Biotin, VivaSpin20 columns were subject to three times of centrifugation with 20 ml autoclaved ddH_2O. After the last centrifugation step,

the retentate (0.5ml) was diluted in 9.5 ml of PBS. For purification of biotinylated proteins, the sample was loaded on a 10 ml polyprep column with 300 µl high capacity neutravidin agarose (life technologies) for overnight incubation at 4°C. After binding of proteins, neutravidin beads were washed 3 times with 10 ml PBS supplemented with 2% SDS. To elute the biotinylated and azido-labeled glycoproteins, neutravidin beads were boiled with 200 µl of SDS elution buffer (10X elution buffer: 1M Tris-HCl pH6.8 60ml, SDS 20g, 2-ME 4ml, 100mM 96ul, add H2O to 100ml final volume) with 50uM biotin at 98°C for 10 minutes. 5% of the elution was reserved for Western blot analysis. The rest eluates were submitted to mass spectrometry facility for sample preparation and peptide identification.

Chapter 3 Identifying TMEM115 interacting partners to reveal novel functions of TMEM115

Introduction

To elucidate the biological role of TMEM115, I used a proteomics screen to identify proteins that interact with TMEM115 in mammalian cells. Previously, researchers in the Freeman lab had tried multiple times with conventional co-immunoprecipitation strategies followed by mass spectrometry analysis to identify physical binding partners of TMEM115, but they obtained surprisingly few candidates and the reproducibility was poor. We envisaged two reasons for these results. Either the immunoprecipitation conditions need to be carefully optimised for TMEM115, or that TMEM115 interacts with other proteins in a transient manner, or through loose physical interactions. If the latter was the case, co-immunoprecipitation would not be the ideal strategy for identifying weak interactions. Therefore, I decided to apply a proximity-dependent biotin identification (BioID) screen for identifying TMEM115 interactors. Unlike co-immunoprecipitation, BioID is a proximity tagging method that does not require strong physical interaction between the target protein and its potential interactors. Instead, it relies only on physical proximity of two proteins and is a powerful method to identify loose physical interactors and transient interactors (Roux 2013). In this assay, TMEM115 was fused with BirA* (R118G), a mutant *E. coli* biotin

ligase that biotinylates proteins within a radius of approximately 10nm (a diagram is presented in Figure 3.1) (Roux 2013). The biotinylated proteins are then purified on streptavidin-coated beads and analysed by mass spectrometry. In principle, BioID suffers from its indiscriminate labelling of any nearby protein, but previous experience within the Freeman group and careful study of published work indicated that in practice non-specific hits are not a great problem as long as the experiment is properly controlled.

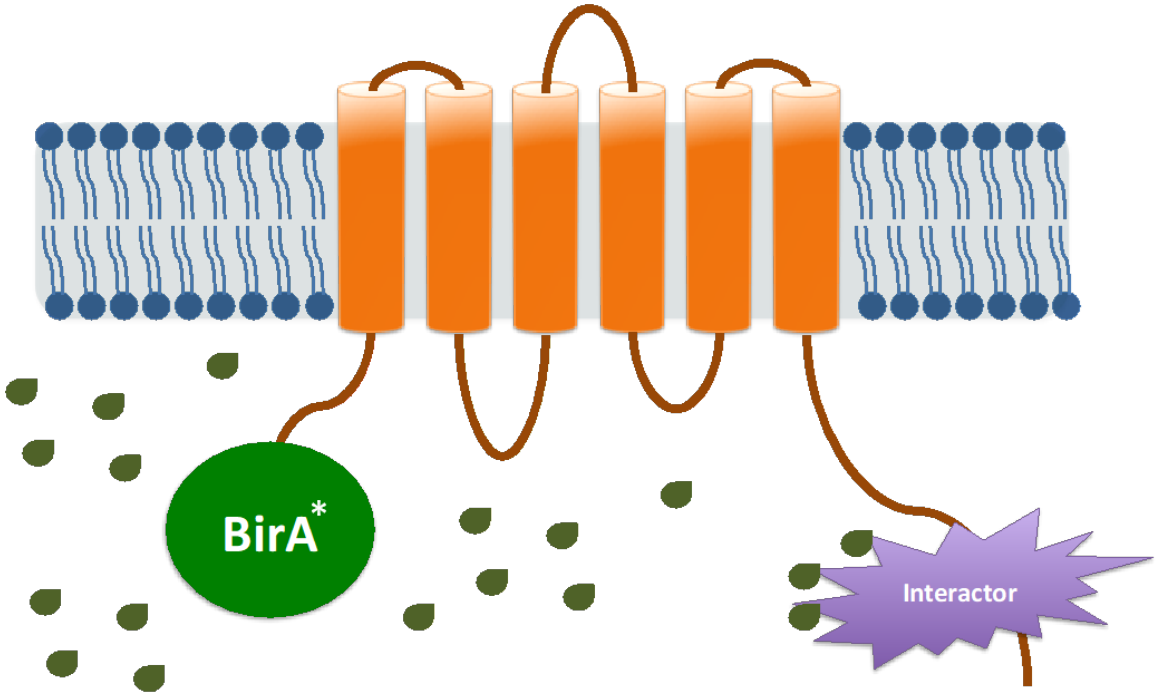


Figure 3.1: Model of BioID screen

TMEM115 was fused with BirA* at the N-terminus. BirA* releases active biotin that biotinylates nearby protein through their Lysine residues within a radios of 10 nm.

For BioID to work successfully, it is essential that the BirA* fusion protein is expressed at an appropriate level (Kim et al. 2016, Mehus et al. 2016). The excessive expression will result in unspecific labelling, which increases the difficulty of further validation. On the other hand, too low a level of the fusion protein may not provide enough copies of biotinylated interactors to detect by mass spectrometry. I, therefore, had to develop a system, in which the production level of BirA* tagged TMEM115 can be tightly controlled.

3.1 Testing the production and localisation of BirA* tagged TMEM115

As mentioned before, an appropriate level of BirA*-TMEM115 is essential for identifying physiologically relevant interactors for TMEM115 in the BioID screen. To be able to control the production level of BirA*-TMEM115, the fusion protein was cloned into a Tet-One vector, in which the Tet 3Gs promoter controls the production of the downstream protein tightly upon doxycycline stimulation in a dose-dependent manner.

I tested the production and localisation of BirA*-TMEM115 under this system. Upon doxycycline induction, the fusion protein was produced at its predicted size (around 74 kDa (37 kDa+37 kDa), Figure 3.2A) and it biotinylated protein in the presence of supplemented biotin. I was also able to observe that BirA*-TMEM115 localised predominantly within the Golgi apparatus upon doxycycline stimulation and biotinylated nearby proteins in a proximity-dependent manner. Therefore, I decided to carry on with the BirA*-TMEM115 system for my BioID screen.

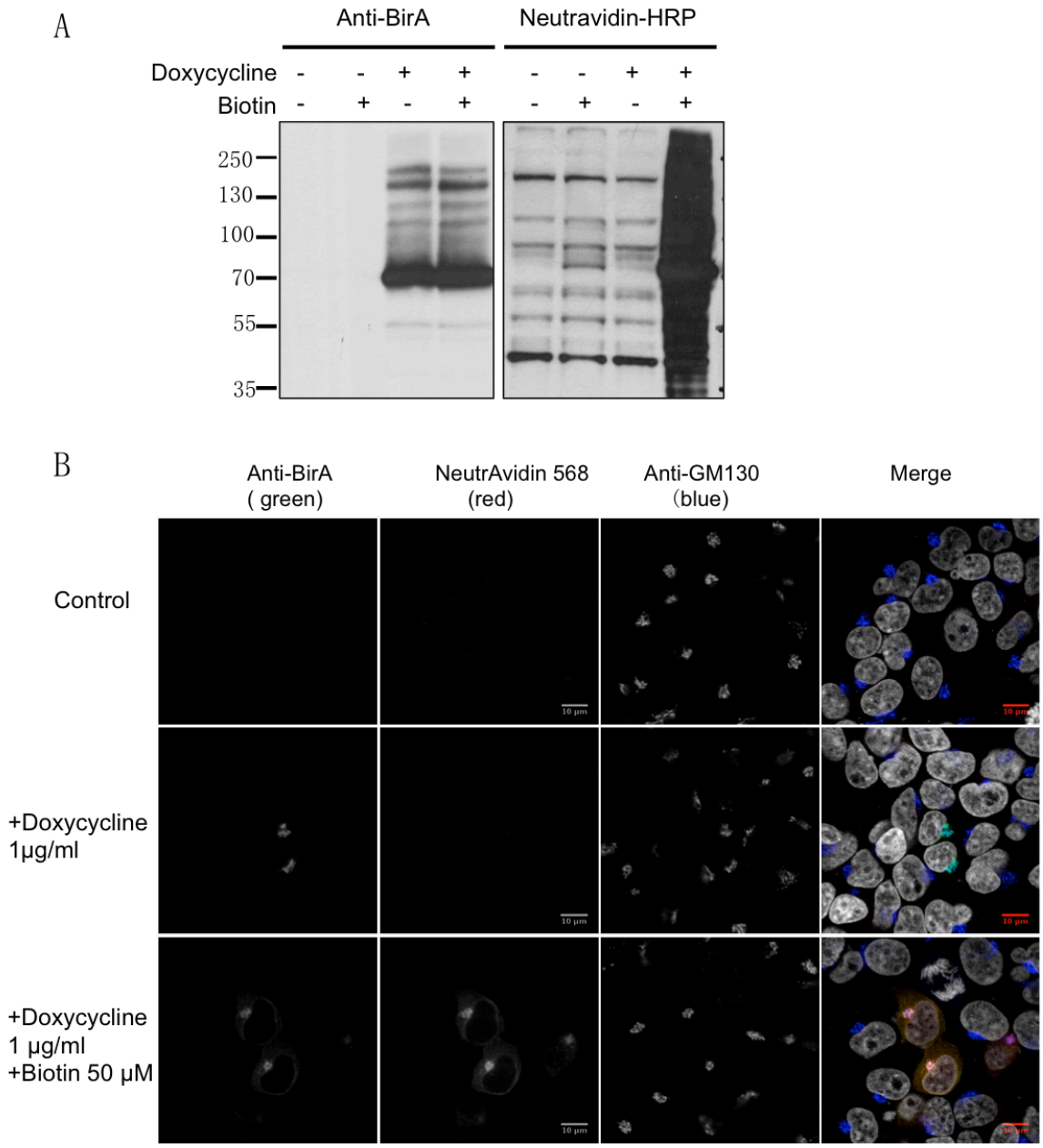


Figure 3.2: The production of BirA-TMEM115 is tightly controlled
A. BirA*-TMEM115 (74 kDa band) was only detected upon Doxycycline induction. Biotinylated proteins (blotted with Neutravidin-HRP) were detectable in the presence of fusion protein and supplemented biotin.
B. GM130 was used as a Golgi maker. Biotinylated proteins labelled with Neutravidin 568 co-localise with BirA*-TMEM115, which shows biotinylation is proximity dependent. BirA*-TMEM115 was only expressed in the presence of Doxycycline and was predominantly localised in the Golgi. Scale bar (at the bottom left of the images):10 µm

3.2 Generating a stable cell line for the BioID screen

After a series of initial assessments of the BirA*-TMEM115 construct through transient transfection, I decided to generate clonal HEK293T and MCF-7 cell lines that stably express BirA*-TMEM115 upon doxycycline induction. Compared to transient transfection approaches, clonal cell lines offer a more controlled system for expression, in which the production level of fusion protein between individual cells is homogenous. Using clonal cell lines, the reproducibility of the BioID screen is therefore expected to be high. The negative side, however, is that clonal cell lines offer less diversity in terms of expression level when compared with polyclonal cell lines or transient transfection systems. In addition, there is a potential risk of selecting a clone, which does not represent a typical condition. Therefore, one is at the risk of ignoring insightful information provided by a diversity of fusion protein production levels in different cells.

I used the Tet-One constructs that contain BirA*-TMEM115 or BirA* alone to generate stable clonal cell lines in both MCF-7 and HEK293T cells, but after several attempts, it was not successful. In MCF-7, six days after transfection, cells were clearly producing BirA*-TMEM115 upon doxycycline induction. However, after 20 days of puromycin selection and clonal cell lines generation, I had only one positive clonal cell line that expressed soluble BirA* after screening more than 20 clones for each construct (part of the screen result is shown in Figure 3.3), with no clones that stably express BirA*-TMEM115, despite plenty of clones surviving under puromycin treatment. It was the same scenario with HEK293T cells: after several attempts, I could not get any clonal cell lines that stably produced BirA*-TMEM115 using this method. This problem may have been a consequence of the specific design of the Tet-One vector I was using. To increase the induction sensitivity of the pTRE3G promoter, no antibiotic resistant gene was built within the vector to avoid any potential interference with basal expression of the gene of interest. Therefore, to generate a stable cell line, I needed to co-transfect my construct with a linear selection marker, which contained the puromycin resistance

gene. In an ideal situation, after selection, both my gene of interest and the puromycin resistance gene would be integrated into the genome of the target cells with similar efficiency. However, the false positive cells clearly contained the puromycin resistance gene only, with no trace of expression of my protein of interest under doxycycline induction. There was a huge difference in terms of integration efficiency between the linear selection marker and the constructs, despite the fact that I co-transfected the constructs with the linear selection marker at a 20 to 1 ratio, as recommended by the user manual. The linear selection marker is four times (1.7kb) smaller than my construct (6.5kb), which may contribute to its much higher integration. In addition, linearisation may increase the integration efficiency. Since I was not able to reduce the size of my constructs, I tried to linearise them. This reduced the transfection efficiency greatly, however, it did not solve the integration issue. Given that cell lines that stably express BirA* was obtained with the same system, another possibility could be that overexpression of TMEM115 is slightly toxic for cells, that under the selection of puromycin, positive cells that integrated with BirA*-TMEM115 were more sensitive to selection and could not survive eventually. For these reasons, although the Tet-One vector was an attractive system to tune the production level of the fusion protein, and due to the challenge of generating a stable cell line, I decided to adopt a different system.

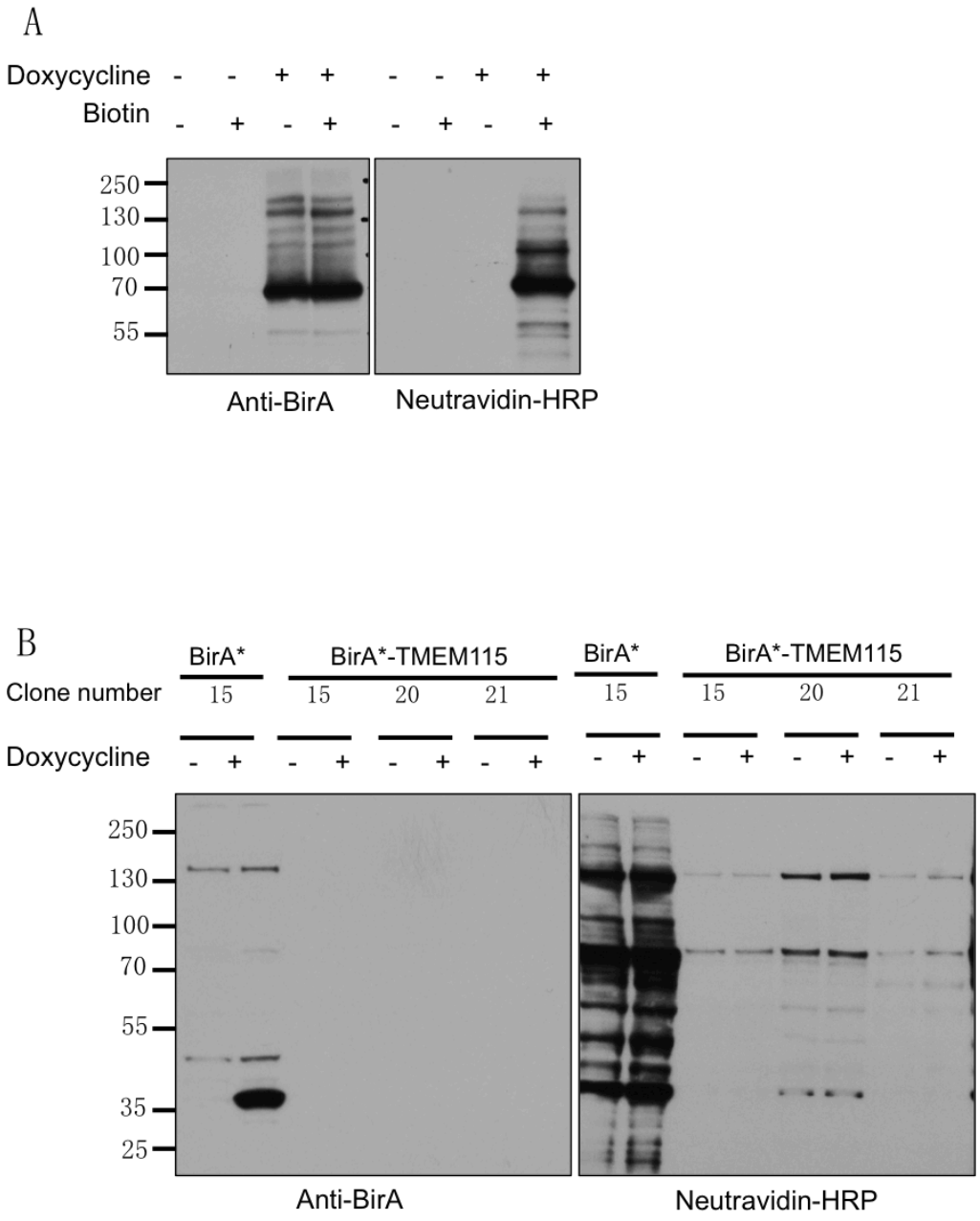


Figure 3.3 Generating Clonal cell lines with Tet-One vector

A. Western blot of mix pool of MCF-7 cells co-transfected with BirA*-TMEM115 and puromycin linear selection marker (6 days after transfection). BirA*-TMEM115 was produced upon doxycycline induction (main band at 74KDa), and additional biotinylation was detectable.

B. Western blot of clonal stable cell lines for BirA* and BirA*-TMEM115 (20 days under puromycin selection). BirA* was detected in clone BirA*15 upon doxycycline induction, none of the BirA*-TMEM115 clonal cell lines was expressing BirA*-TMEM115.

I next tried a lentiviral system for stable cell line generation. In this system, in addition to using the N-terminus BirA* tagged TMEM115 that I tested before, I also made a C-terminus tagged version of TMEM115 (TMEM115-BirA*). BirA*, BirA*-TMEM115 and TMEM115-BirA* were cloned into a CMV/TO-puromycin destination vector, which in theory was a doxycycline inducible system for tuning the production level of proteins of interest. MCF-7 cells were transduced with lentivirus and levels of target proteins were detected by western blot after seven days of puromycin selection (Figure 3.4). In the generated stable cell lines, the production of BirA* (at 37 kDa) was not doxycycline dependent, indicating a leakage in the system. BirA*-TMEM115 (at around 74 kDa) was produced only in the presence of doxycycline, which indicated a much tighter control of the CMV/TO promoter in the BirA*-TMEM115 stable cell line. Regarding the biotinylation pattern in those samples, however, there was rather a low level of additional biotinylation in comparison with what was shown previously with the Tet-One system, using the same concentration of doxycycline and biotin. With anti-TMEM115 antibody, I also compared the level of BirA*-TMEM115 with endogenous TMEM115. I found that the overexpressed protein was expressed at around five-fold the endogenous level. In summary, the lentiviral tetracycline inducible system could not be controlled tightly with doxycycline in some of the cell lines; on the contrary, there was hardly any inducible effect with supplemented doxycycline in the BirA* polyclonal cell lines. In addition, BirA*-TMEM115 was produced at a close to endogenous level, and the additional biotinylation was rather modest.

In the following step, I used immunofluorescence to detect the localisation of the fusion protein and the biotinylation pattern in the polyclonal MCF-7 cell lines. As expected, BirA*-TMEM115 co-localised with the Golgi marker, GM130. However, the pattern of biotinylated proteins was not detectable under low laser power. After increasing the laser power to a much higher level, which was enough to detect all the naturally biotinylated carboxylases, I was still not able to detect additional biotinylation patterns that could partially co-localise with BirA*-TMEM115. However, after reading the new guidelines for BioID updated by

the inventor of this method, the Roux lab, the explanation was since the fusion protein was overexpressed at only a modest level, the additional biotinylated proteins might not detectable by immunofluorescence. But even this modest level of additional biotinylation can be detected by mass spectrometry after further enrichment steps from a large number of cells.

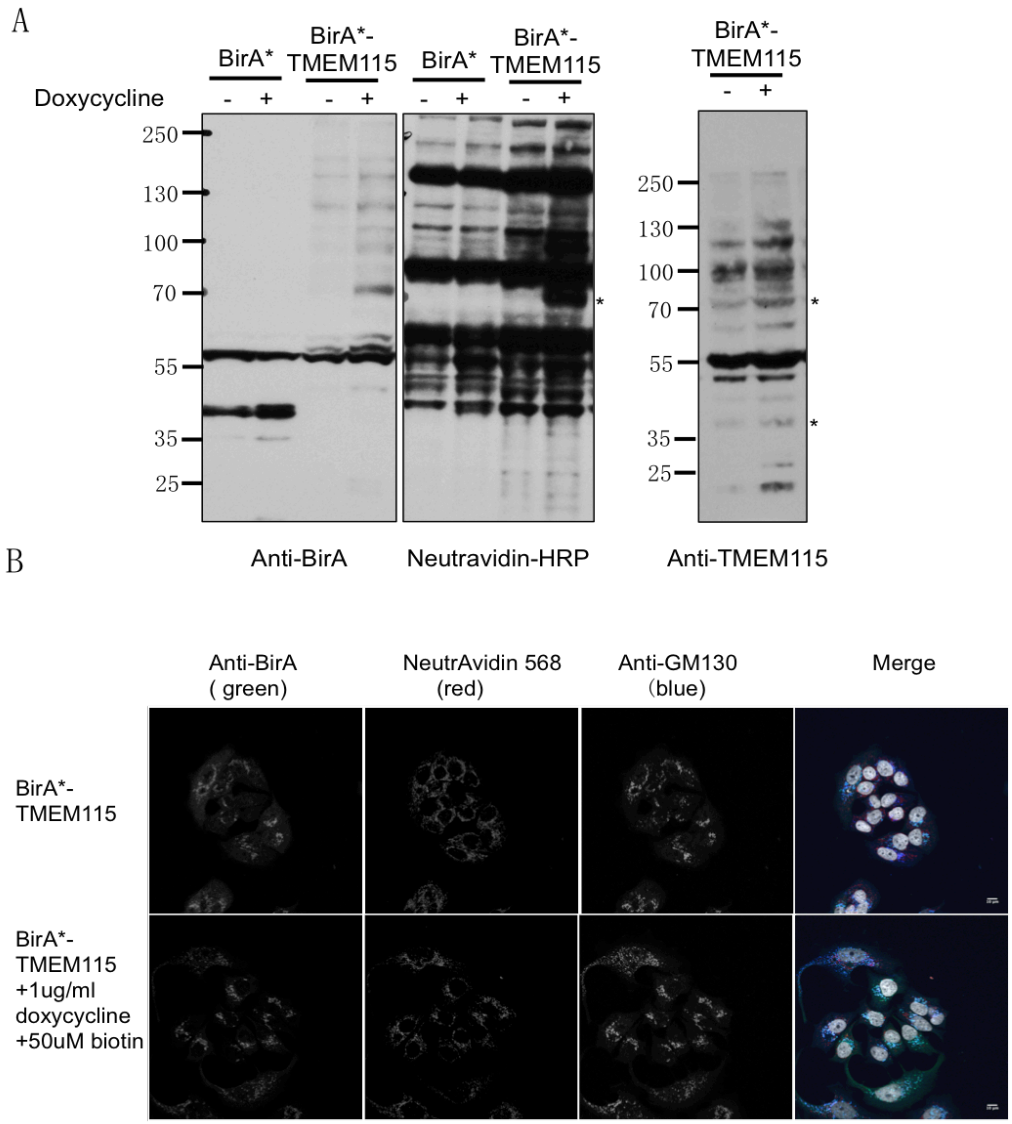


Figure 3.4 Generating stable cell lines with lentiviral system

A. Western blot of lentiviral transduced BirA* and BirA*-TMEM115 MCF-7 cell lines after seven days of puromycin selection. Supplemented Biotin (50 μM) was present in all the samples. BirA* (37 kDa) was produced with a doxycycline independent manner. The production of BirA*-TMEM115 (74 kDa, indicated with an asterisk) was controlled by doxycycline. The level of additional biotinylation was modest in BirA*-TMEM115 cell line with supplemented doxycycline (1 μg/ml) and biotin (Neutravidin-HRP blot). The level of exdogense BirA*-TMEM115 was around five folds of the endogenous TMEM115 level (Anti-TMEM115 blot, indicated with asterisks).

B. Immunofluorescences images for examining the localisation of BirA*-TMEM115 in the stable cell line. BirA*-TMEM115 co-localised with Golgi marker GM130. No obvious additional biotinylation was detected with supplemented doxycycline (1 μg/ml) and biotin (50 μM). The biotinylation pattern

in BirA*-TMEM115 MCF-7 cell line was not co-localising with BirA*-TMEM115. Scale bar (at the bottom left of the images): 10 μm

I also tested the C-terminus BirA* tagged TMEM115 in stable cell lines. However, the C terminally BirA* tagged TMEM115, TMEM115-BirA*, formed many truncated versions with multiple bands, which were much lower than the predicted full-length fusion protein (around 74 kDa). The production level of TMEM115-BirA* was also lower than BirA*-TMEM115. TMEM115-BirA* was also no longer localised within the Golgi. All the above results indicate that the BirA* tag at the C terminus of TMEM115 disrupts the folding and stability of TMEM115 (Figure 3.5).

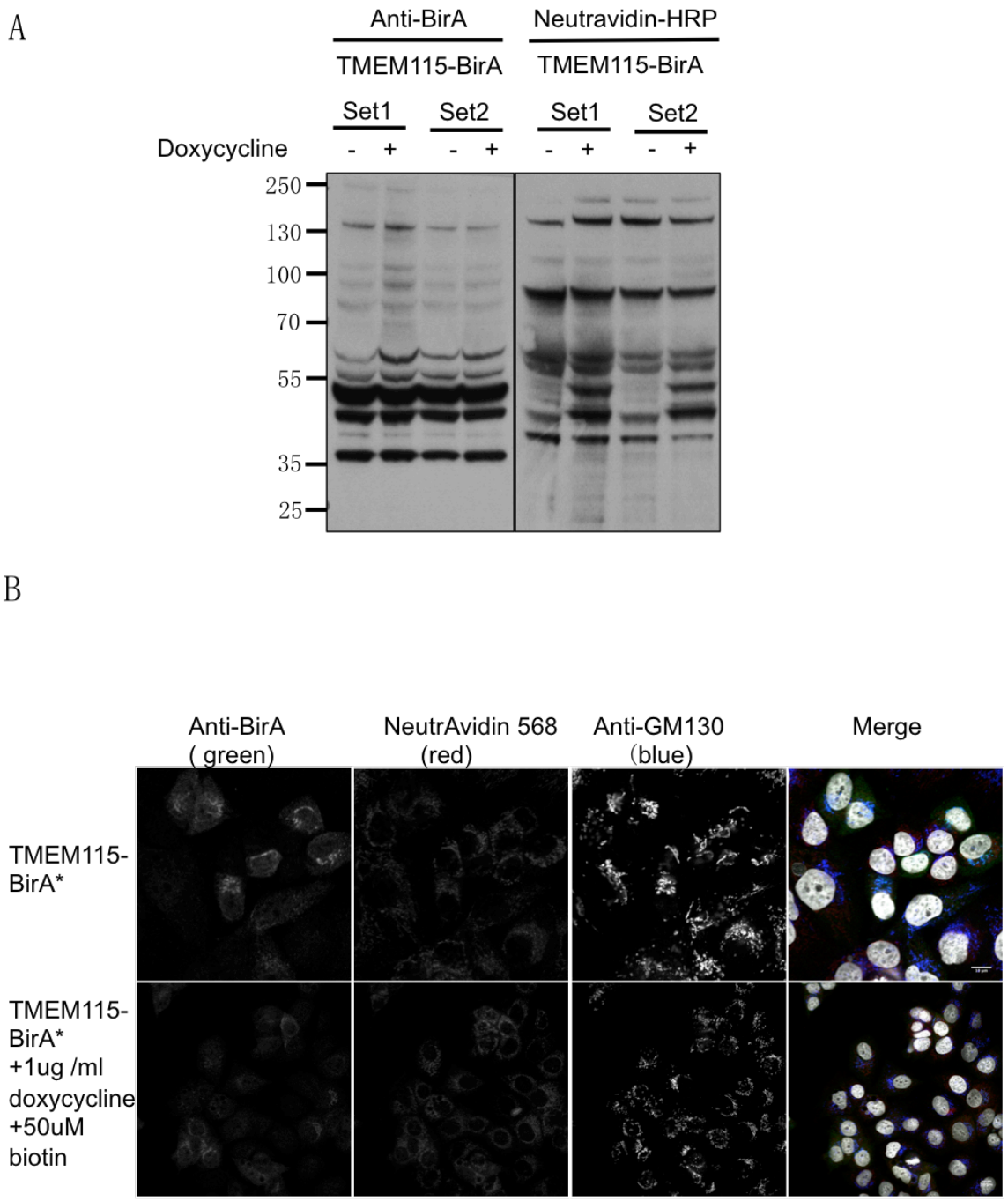


Figure 3.5 TMEM115-BirA* MCF-7 stable cell line

A. Western blot of lentiviral transduced TMEM115-BirA* MCF-7 cell lines after seven days of puromycin selection. TMEM115-BirA* was produced in truncated forms (No fusion protein was detected at 74 kDa). Almost no additional biotinylation was detected.

B. Immunofluorescences images for examining the localisation TMEM115-BirA*. TMEM115-BirA* was not co-localising with Golgi marker GM130. Scale bar (at the

bottom left of the images): 10 μm

I performed trials of BioID sample preparation on a small scale with the MCF-7 cell lines stably expressing BirA*-TMEM115 or BirA* alone. Western blot and silver staining were employed to detect biotinylated materials in the material eluted from the steptavidin beads (data not shown). Western blotting with Neutravidin-HRP showed that the biotinylated materials in the flow through were almost depleted, which suggested the high efficiency of protein purification. In the silver staining, I was not able to observe bands that were specifically present in BirA*-TMEM115 samples, which was consistent with my earlier observation of only modest biotinylation. One set of BioID samples were processed and analysed by mass spectrometry, but it did not produce any useful information about potential interacting proteins of TMEM115. In total, 51 proteins were identified in BirA*-TMEM115 sample with low peptide count, and 151 proteins were identified in BirA* alone sample. After eliminating the common hits that were shared with the BirA* alone sample, only one protein survived in BirA*-TMEM115 sample. It was highly unlikely that TMEM115 is only proximate to one protein. Furthermore, TMEM115 itself was not in the mass spec list of BirA*-TMEM115 sample, which is an indicator that the BioID screen was not efficient. BirA* biotinylates nearby proteins through their lysine residues in a proximity dependent manner, so TMEM115 fused with BirA* should be heavily biotinylated, as it has plenty of lysines in cytosolic regions of the protein. Therefore, TMEM115 should be one of the most abundant biotinylated proteins present, as has been the case for most BioID screens with rhomboid-like superfamily members that have been undertaken by researchers in the Freeman lab.

In summary, although a modest overexpression of BirA-TMEM115 at close to endogenous level is desired for identifying biological relevant interactors, the MCF-7 cell line that stably produces BirA*-TMEM115 at a comparable level to endogenous TMEM115, was not able to provide enough additional biotinylated materials at a mass spectrometry detectable level with the current sample scale. I would need either to increase the sample scale or adopt a different system that

could provide more biotinylated proteins for mass spectrometry analysis. Simply increasing sample scale would, in theory, provide more copies of biotinylated proteins in the mass spectrometry samples, but also would increase background proteins accordingly, as carboxylase proteins in the mitochondrial are naturally biotinylated. Overall, these results were not encouraging, and I decided to develop a new approach.

3.3 Optimising BioID screen conditions with Tet-One vector under transient transfection condition

After spending one year on these unsuccessful attempts to generate stable cell lines for the BioID screen, I decided to set up a protocol for the BioID screen with transiently transfected cells. In order to obtain a biologically meaningful and reproducible BioID candidates list, the transient transfection protocol needed to be carefully optimised. Firstly, the production level of overexpressed fusion protein should be relatively homogeneous among transfected cells; second the transfection rate needed to be high enough to provide enough biotinylated proteins for the mass spec analysis, without being so high as to drive non-specific background biotinylation; third, the overexpressed fusion protein had to be appropriately localised within the Golgi apparatus.

I used the Tet-One vector that was described in Section 3.1 of this chapter for this transfection protocol, as the protein production can be tightly controlled by doxycycline in a dosage-dependent manner. By applying a combination of western blotting and immunofluorescence methods, I was able to choose the optimal concentration of doxycycline for the final screen.

I titrated the dosage of doxycycline in HEK293T cells transiently transfected with BirA*-TMEM115 and detected both the level of the fusion protein and biotinylated proteins under different doses (ranging from 10 ng/ml to 1 μg/ml) of doxycycline treatment by western blot. Ideally, I wanted to choose a dose of doxycycline that provided sufficient fusion protein for production of additional biotinylated proteins that would enable identification by mass spectrometry analysis. Under these conditions, the Tet-One system responded well to

doxycycline, the level of BirA*-TMEM115 and the amount of additional biotinylated proteins showed a dose dependence at the lower concentration range from 10 ng/ml to 80 ng/ml (Figure 3.6A).

I used immunofluorescence to detect the localisation of BirA*-TMEM115 under different concentrations of doxycycline (0,10,20,40 and 1000 ng/ml), as well as to assess the transfection efficiency in HEK293T cells (Figure 3.6B). Since I did not culture the cells in tetracycline-free serum, I could already detect BirA*-TMEM115 in some cells without supplemented doxycycline. At 10 and 20 ng/ml of doxycycline induction, BirA*-TMEM115 co-localised with the Golgi marker GM130. However, starting from 40 ng/ml of doxycycline treatment, I was able to observe an ER-like distribution of BirA*-TMEM115 in some cells. This suggested that the overexpression level of the fusion protein was too high in some cells, leading to a significant portion of BirA*-TMEM115 being present in the ER. This would cause a problem for the proximity biotinylation labelling. At 10ng/ml doxycycline treatment, only modest additional biotinylation was observed. Therefore, for a trade-off of a low level of the correctly localised fusion protein and a reasonable level of additional biotinylated proteins, I chose a concentration of 20ng/ml of doxycycline for the final BioID screen. I was also able to comment about transfection efficiency with immunofluorescence: overall, more than 70% of cells expressed BirA*-TMEM115.

Figure 3.6A

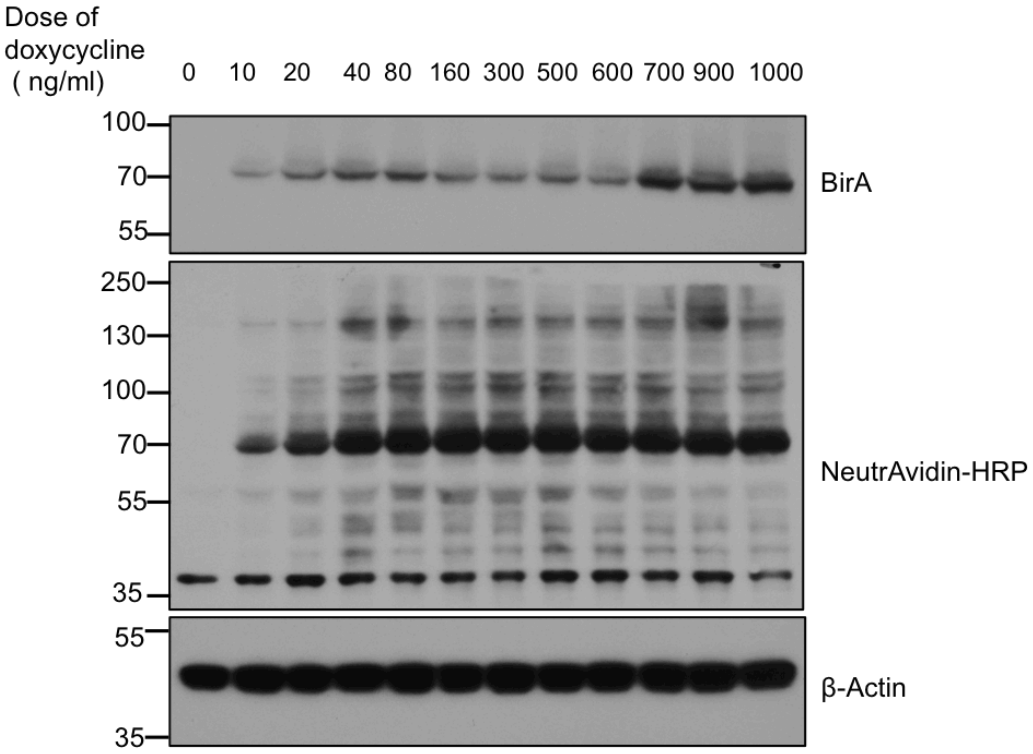


Figure 3.6 Optimising transient transfection condition for BioID in HEK293T cells

HEK293T cells were transiently transfected with BirA*-TMEM115 (Tet-One vector) and examined after 24 hours of doxycycline induction with supplemented biotin (50 μM)

A. Titrating doses of doxycycline for the optimal expression level of BirA*-TMEM115 and additional biotinylation. BirA*-TMEM115 (74 KDa main band) was produced upon doxycycline induction with a does dependent manner. Additional biotinylated proteins were observed in samples cultured with supplemented biotin and doxycycline (NeutrAvidin-HRP blot)

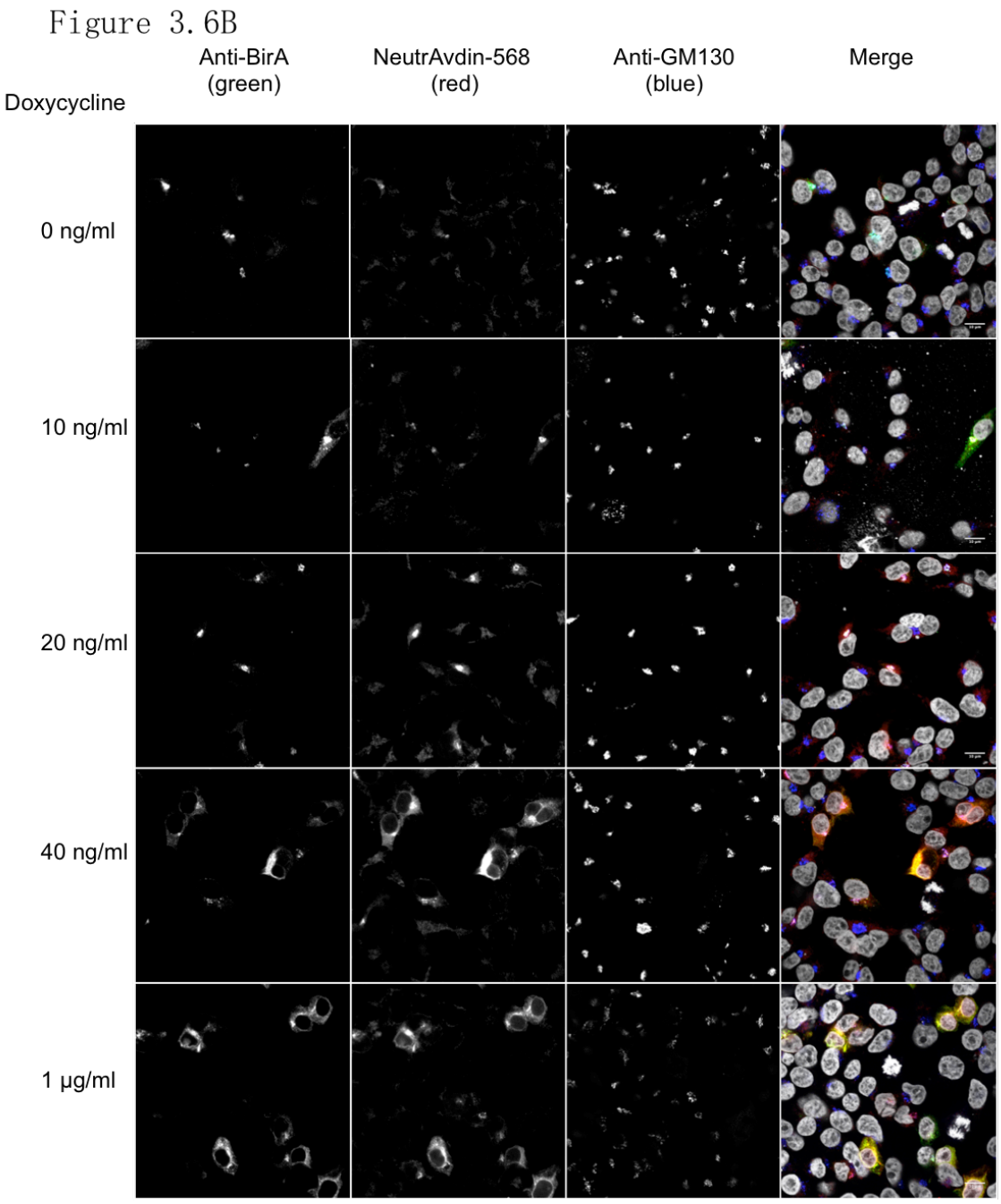


Figure 3.6 Optimising transient transfection condition for BioID in HEK293T cells

B. Immunofluorescences images for examining the cellular localisation of BirA*-TMEM115 upon 0, 10, 20, 40 and 1000 ng/ml of doxycycline induction. Starting from 40 ng/ml, an excessive amount of BirA*-TMEM115 was detected in some cells with an ER-like distribution pattern. GM130 was used as a Golgi marker. Scale bar (at the bottom left of the images): 10 µm

Before moving to the full scale BioID, I piloted BioID sample preparation at a small scale, using HEK293T cells that were transiently transfected with BirA* or BirA*-TMEM115, followed by western blot and silver staining. In Figure 3.7A, the western blotting with Neutravidin-HRP highlighted all the biotinylated proteins in the sample; there was clearly more biotinylated material in the BirA*-TMEM115 sample than the soluble BirA* alone sample. Very little biotinylated material was detected in the flow though after overnight incubation with streptavidin beads, proving that the capture efficiency was high and the amount of beads I had used was suitable. Compared with the silver staining result from the stable cell system I described in Section 3.3, where no bands could be observed that were specific for the BirA*-TMEM115 sample, it was also reassuring that I was able to observe additional bands clearly in BirA*-TMEM115 sample by silver staining (see Figure 3.7B). The strong band at around 74 kDa (arrowed in Figure 3.7B) in the BirA*-TMEM115 sample could be BirA*-TMEM115 itself. When comparing the empty vector sample, there were many additional biotinylated proteins in both BirA* and BirA*-TMEM115 samples. Overall, these data confirmed that the transient transfection system was effective at identifying potential interactors of TMEM115 over background proteins.

In summary, I used a doxycycline inducible vector for establishing a transient transfection system in HEK293T cells for the BioID screen. In this system, the transfection efficiency was high, and the level of overexpressed protein could be tightly controlled by doxycycline. This system provided enough copies of additional biotinylated proteins at a mass spectrometry detectable level, indicated by observable bands on silver stained gels.

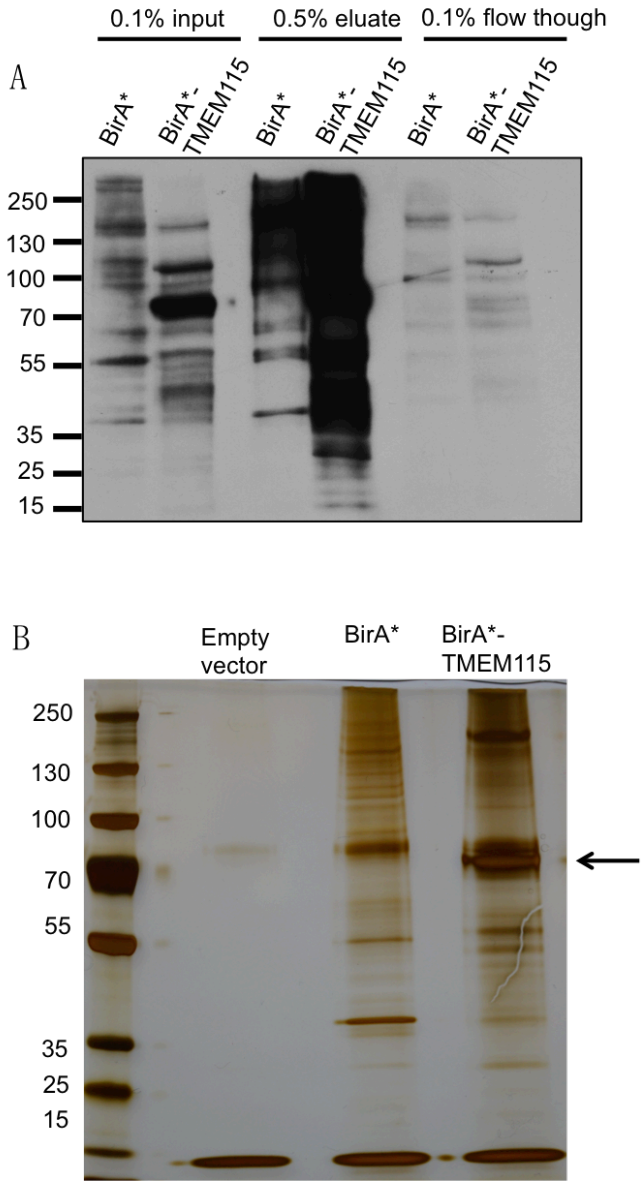


Figure 3.7 Streptavidin capture of Biotinylated proteins for BioID screen
HEK293T cells were transiently transfected with BirA* or BirA*-TMEM115 (Tet-One vector, 15 µg per 15 cm dish) and examined after 24 hours of doxycycline induction (20 ng/ml).
A. Western blot of BirA* and BirA*-TMEM115 samples after overnight streptavidin beads capture (100 µl beads per 15 cm dish of the sample). Enriched biotinylated proteins were detected in beads eluate for both samples.
B. Silver staining of empty vector, BirA* and BirA*-TMEM115 beads eluate after overnight streptavidin beads capture. The additional band at around 74 kDa in BirA*-TMEM115 sample (with arrow) could be biotinylated BirA*-TMEM115 itself.

3.4 BioID screen result under basal condition

After performing the BioID labelling and mass spectrometry, I obtained three sets of TMEM115 potential interactors from each experiment. A few elimination steps were required to produce the final list of TMEM115 potential interactors. Soluble BirA* was employed in the BioID screen to eliminate all the background hits that were either naturally biotinylated, such as mitochondrial carboxylases, or abundant cytosolic proteins that were unspecifically biotinylated by BirA*. After all these background hits were removed, to select the potential TMEM115 interacting proteins that were reproducible among triplicates of BioID, I did a cross comparison among the three lists of proteins that had been identified as being either specifically or highly enriched (at least four-fold higher in terms of spectral count in BirA*-TMEM115 sample compared with BirA* samples). The final hits are shown in Supplementary Table 1 and 2.

In summary, unsurprisingly, TMEM115 was at the top of every list. The other top hits were highly reproducible among the triplicates. In total, 116 proteins were reproducibly identified that were either only present or highly enriched in BirA*-TMEM115 samples. Some published interactors (Ong et al. 2014), including ERGIC53, beta-COP and COG4, were also present in the list. Rhomboid-like proteins are expected to interact with other membrane proteins, and TMEM115 was not an exception. More than 50% of proteins in the potential interactors list are membrane proteins, of which 35 are single-pass transmembrane proteins, and 22 are multi-pass transmembrane proteins (Figure 3.8A). In terms of protein functions, as defined by GO annotation, 27 of the potential interacting proteins are involved in trafficking and nine are involved in lipid biology. These were expected possible roles of TMEM115, as the preliminary data about TMEM115 in Freeman lab have already suggested that TMEM115 has a role in lipid biology, indicated by the lipid-related phenotypes of *Tmem115* knockout mice. TMEM115 is a Golgi protein; Ong et al. have reported it may have a role in retrograde trafficking (Ong et al. 2014). The other two functional clusters of potential interactors were completely unexpected: ten proteins are either components of ion channels or

their regulators, and there are six proteins involved in protein degradation. Does this suggest that TMEM115 may have roles in regulating ion channels or protein degradation? More validation steps and follow up experiments are required to provide a firm answer for that, but it provides interesting possible avenues to pursue.

The potential interacting proteins list from the BioID screen provides interesting hints of TMEM115 functions, but it needed further validation steps to highlight the top candidates to pursue. Given TMEM115 is a Golgi protein, BirA*-TMEM115 may label many abundant Golgi resident proteins only because they are localised in close proximity within the Golgi. Therefore, I adopted another Golgi protein as an organelle control to eliminate all the Golgi proteins that had been unspecifically biotinylated. RHBDD2 was used as a control, as it is also a rhomboid-like protein and predominantly localises within the Golgi. Dr Nina Jajcanin Jozic, a postdoc in the Freeman lab, has obtained a high quality candidate list for RHBDD2 with BioID. Comparison of the BioID list of TMEM115 with RHBDD2, I was able to eliminate all the shared candidates between these two proteins, leaving seventy-five proteins in the list that are unique candidates for TMEM115. Notably, among the top ten candidates ranked by spectral counts in the original list, nine were shared hits with RHBDD2. Many of the shared top candidates are abundant Golgi resident proteins, often with high molecular weight, such as the golgins. When classified by protein function, many of the potential interactors that involved in trafficking or lipid biology are present in both the TMEM115 and RHBDD2 BioID lists. On the contrary, almost all the protein degradation or ion channel component/regulators were unique hits for TMEM115. I decided to focus my efforts on following up those proteins that are unique hits for TMEM115, at the risk of dismissing some of the shared interactors between TMEM115 and RHBDD2 that could carry biological meaning.

In summary, I obtained a high quality list of unique potential interactors for TMEM115 from the BioID screen, after eliminating both background caused by BirA* and Golgi specific background (top 15 TMEM115 unique candidates listed in

Table 3.1). Some of the BioID hits have suggested novel and exciting potential roles of TMEM115, including protein quality control and regulating ion channels. It was slightly disappointing, however, that not many potential interactors are involved in lipid biology, given preliminary data suggest that TMEM115 has a role in lipid biology. One possible explanation is that TMEM115 only interacts with those lipid-related interactors upon conditions where the lipid-related pathways were induced. Based on this hypothesis, I performed a BioID screen under low lipid conditions, which I will discuss in the next section.

Figure 3.8

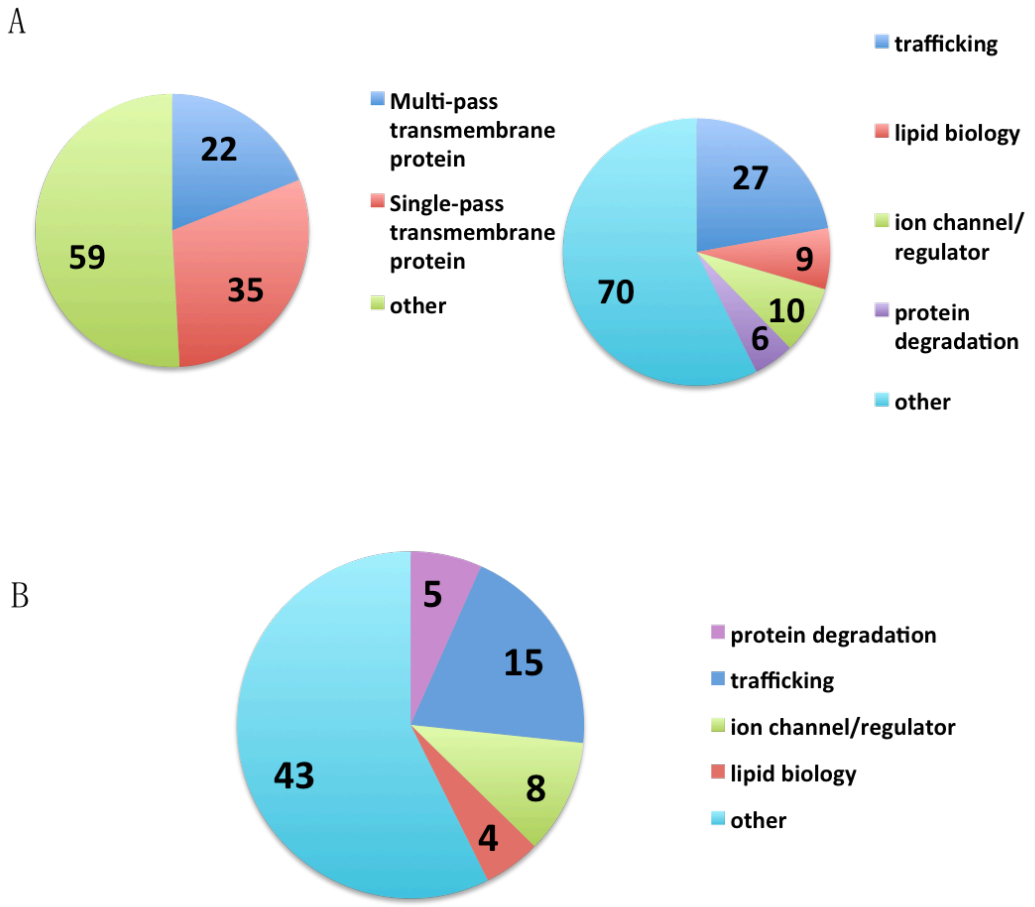


Figure 3.8 Summary of BioID under basal condition
A. 116 proteins were reproducibly identified that were either unique or highly enriched in BirA*-TMEM115 sample. Potential interactors of TMEM115 classified by protein topology (left panel) or protein functions (right panel, based on Go annotation).

B. After removing common hits with RHBDD2, 76 proteins were unique candidates for TMEM115. Potential interactors of TMEM115 classified by protein functions (based on Go annotation).

Table 3.1 List of top 15 TMEM115 unique BioID candidates (ranked by average peptide count in triplicated BioID experiments)

Protein	Description	Average Spectral count
sp|P49257|LMAN1_HUMAN	Protein ERGIC-53 OS=Homo sapiens GN=LMAN1 PE=1 SV=2	34.57
sp|P42566|EPS15_HUMAN	Epidermal growth factor receptor substrate 15 OS=Homo sapiens GN=EPS15 PE=1 SV=2	22.98
sp|Q13501|SQSTM_HUMAN	Sequestosome-1 OS=Homo sapiens GN=SQSTM1 PE=1 SV=1	20.97
sp|P05023-4|AT1A1_HUMAN	Isoform 4 of Sodium/potassium-transporting ATPase subunit alpha-1 OS=Homo sapiens GN=ATP1A1	17.84
sp|P50851-2|LRBA_HUMAN	Isoform 2 of Lipopolysaccharide-responsive and beige-like anchor protein OS=Homo sapiens GN=LRBA	17.69
sp|P08240|SRPR_HUMAN	Signal recognition particle receptor subunit alpha OS=Homo sapiens GN=SRPR PE=1 SV=2	14
sp|Q9H2J7|S6A15_HUMAN	Sodium-dependent neutral amino acid transporter B(0)AT2 OS=Homo sapiens GN=SLC6A15 PE=1 SV=1	13.98
sp|P05067-10|A4_HUMAN	Isoform APP639 of Amyloid beta A4 protein OS=Homo sapiens GN=APP	13.98
sp|O43399-5|TPD54_HUMAN	Isoform 5 of Tumor protein D54 OS=Homo sapiens GN=TPD52L2	12.00
sp|Q5VV42|CDKAL_HUMAN	Threonylcarbamoyladenosine tRNA methylthiotransferase OS=Homo sapiens GN=CDKAL1 PE=1 SV=1	12.00
sp|Q99460-2|PSMD1_HUMAN	Isoform 2 of 26S proteasome non-ATPase regulatory subunit 1 OS=Homo sapiens GN=PSMD1	12.00
sp|P51149|RAB7A_HUMAN	Ras-related protein Rab-7a OS=Homo sapiens GN=RAB7A PE=1 SV=1	11.93
sp|P53365-3|ARFP2_HUMAN	Isoform 3 of Arfaptin-2 OS=Homo sapiens GN=ARFIP2	11.00
sp|Q01968-2|OCRL	Isoform B of Inositol polyphosphate 5-phosphatase	10.89

_HUMAN	OCRL-1 OS=Homo sapiens GN=OCRL	
sp|P17987|TCPA_HUMAN	T-complex protein 1 subunit alpha OS=Homo sapiens GN=TCP1 PE=1 SV=1	9.95

3.5 BioID screen under low lipid conditions

As described in later chapters, the phenotype of TMEM115 knockout mice and flies suggests a possible role of TMEM115 in lipid biology. In brief, CG9536 (TMEM115 homologue in *Drosophila)* knockout larval fat body cells are much smaller than wild type, which indicates problems in lipid homeostasis, *Tmem115* knockout mice have problems gaining body mass and start losing weight significantly before weaning, hardly any white adipose tissue was found in *Tmem115* knockout mice died right before or at weaning stage. As mentioned in the introduction, a known physical interactor of TMEM115, SCAP, is one of the key regulators for sterol homeostasis. Preliminary data from the Freeman lab has shown that the interaction between SCAP and TMEM115 was significantly increased when assessed in low lipid conditions. This allowed me to hypothesise that TMEM115 may have a role in regulating lipid homeostasis, and the interactions between TMEM115 and its potential lipid-related interactors may dependent on specific physiological conditions. I, therefore, performed the BioID screen again, this time under low lipid cell culture conditions, hoping to identify lipid-environment dependent interactors of TMEM115 by comparing the interactome profiles under basal and low lipid conditions.

To perform the whole BioID screen under low lipid starvation conditions, I used a cell line that responds well to low lipid treatment, and which is easily transfected. Ideally, I wanted to use human pre-adipocyte cell lines, but it is very difficult to achieve high transfection efficiency with those cells. Dr Angela Moncada Pazos in the Freeman group has established a successful protocol to treat HeLa cells with low lipid medium and mevastatin (a compound that lowers hepatic production of cholesterol, for inhibiting alternative lipid synthesis in HeLa cells) to induce the sterol synthesis pathway, to study TMEM115's role in lipid biology. I adapted this protocol for low lipid induction and tested HeLa cells. SCAP was used

as a positive control: if cells react to the low lipid conditions as expected, I should be able to observe increased SCAP interaction with TMEM115. In Figure 3.9, the interaction between TMEM115 and SCAP was, as expected, substantially increased upon lipid depletion.

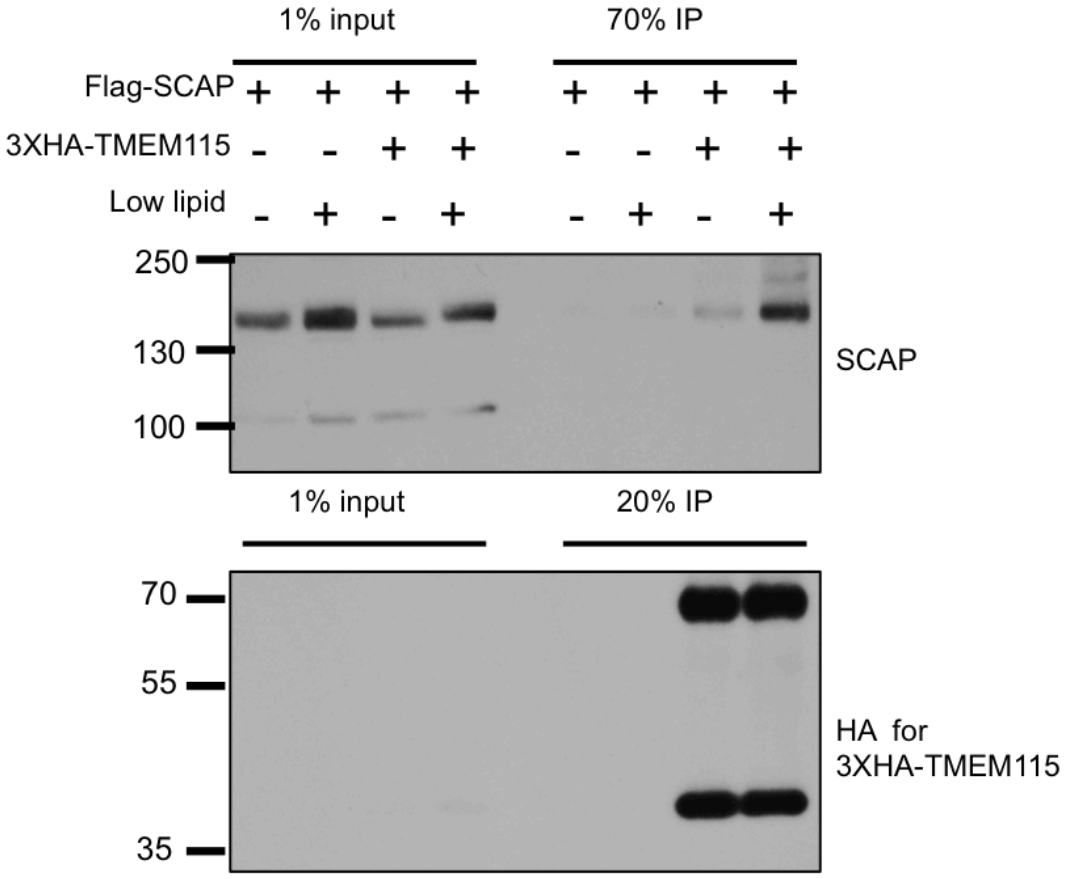


Figure 3.9 Setting up low lipid BioID screen in HeLa cells

HeLa cells were co-transfected with empty vector and Flag-SCAP, or with Flag-SCAP and 3XHA-TMEM115. HeLa cells were cultured in full lipid medium (DMEM with 0.5% FBS + 300 μM of oleic acid + 1 μg/ml of 25-hydroxycholesterol) or low lipid medium (0.5% FBS + 0.5 μM of mevastatin) for 16 hours before HA immunoprecipitation. Cell lysates and HA beads eluate were immunoblotted for SCAP and HA (3XHA-TMEM115). An elevated level of SCAP was co-immunoprecipitated with 3XHA-TMEM115 upon low lipid treatment.

I performed the low lipid BioID in HeLa cells with a similar protocol to that used for basal conditions. BirA* in low lipid culture medium and BirA*-TMEM115 under the previous conditions (referred to as basal conditions) were both used as controls in this experiment. After the first attempt, the result was quite disappointing. Firstly, SCAP, the positive control that I employed to set up the low lipid treatment, was not in the low-lipid BioID candidate list. In addition, I was not able to identify any lipid-related proteins that were either unique or enriched in low lipid BioID. Instead, the BioID candidates' list I obtained from the low lipid BioID screen was almost identical to the previous BioID list from HEK293T cells under basal condition. Although this reproducibility was reassuring about the overall BioID approach, it suggested that the approach may not be an efficient method for identifying subtle changes in interactome profiles. I, therefore, decided not to continue with the low lipid BioID.

In summary, I performed a BioID screen under low lipid conditions, hoping for identification of lipid-related interactors of TMEM115. However, it was not very useful, as no lipid-related interactors were identified that were only present in the BirA*-TMEM15 sample under low lipid conditions. In the future alternative methods, which can capture subtle changes in interactome profiles, would be more appropriate for this type of study.

3.6 Validating potential interactors obtained from BioID screen with co-immunoprecipitated

A number of appealing candidates have been identified through the BioID screen, which could provide insight into the cellular functions of TMEM115. However, the real challenge was to select the best candidate for follow up. Unlike conventional co-immunoprecipitations, which provide physical binding information between the protein of interest and its interactors, a BioID screen is a proximity biotinylation labeling approach, so it only reveals that candidate proteins that are localised

within close proximity of BirA*-TMEM115 during the 24 hours biotinylation time-course; whether those candidates bind stably to TMEM115 is not addressed. I, therefore, decided to apply a co-immunoprecipitation strategy to examine if the top candidates that I obtained from BioID are also physical binding partners of TMEM115.

The list of unique potential interactors of TMEM115 from the BioID screen contains 75 proteins. After ranking them by spectral count, I picked some of the top hits for the first round validation, including ERGIC-53, SRPR, TPD54 and RAB7A. ERGIC-53 (also named as LMAN1) is a cargo receptor that localised in the ER-Golgi intermediate compartment (ERGIC) compartment. ERGIC-53 is responsible for the ER-to-Golgi forward trafficking of glycoproteins (Hauri et al. 2000). SRPR is a component of SRP (signal recognition particle) receptor that localised predominantly within the ER, which ensures the recognition and correct delivery of newly synthesised membrane and secretory proteins to the target cellular membrane (Koch H. G et al. 2003). TPD54 (tumor protein D54) has been reported to be involved in the prognosis of several cancers, while its cellular function has not been clearly revealed (Ligang Ren et al. 2017). RAB7 belongs to the RAB family, which is a branch of the conserved Ras-like GTPase superfamily (Flora Guerra et al. 2017). Rab7 is possibly one of the best-studied RAB proteins, and in mammals, there are two RAB 7 proteins, RAB7A and RAB7B (Pereira-Leal JB et al. 2001). The two forms of RAB7 have different sub-cellular localisations and controls different steps of transport (Chua CE 2011, Agola JO, 2011, Flora Guerra et al. 2017). RAB7A, the potential interactor of TMEM115, is localised predominately in late endosomes and regulates transport to late endocytic compartments such as transport from early endosomes to late endosomes and from late endosomes to lysosomes (Bucci C 2000). In contrast, RAB7B, which controls mainly endosomes to Golgi transport, localised both in the trans-Golgi-network and late endosomes (Progida C, 2010).

By co-overexpression of 3xHA-TMEM115 and 2xFlag-tagged candidates, followed by immunoprecipitation with anti-HA beads, I was able to confirm whether

the BioID candidates bind to TMEM115 through detection of the presence of Flag-tagged proteins in bead eluates. In Figure 3.10A, using Flag antibody, all the above candidates were detected at the predicted size in cell lysates. TMEM115 only co-immunoprecipitated with ERGIC-53, suggesting that, at least when co-expressed, ERGIC-53 is a physical binding partner of TMEM115. The interaction between p62 and TMEM115 was also validated by co-immunoprecipitation (Figure 3.10B). HA-p62 and Flag-TMEM115 were co-overexpressed in HEK293T cells, followed by enrichment of HA-p62 with HA beads; Flag-TMEM115 was detected in the beads eluate of HA-p62. In addition, I also validated the interaction between TMEM115 and beta-COP (Figure 3.11A), which is another published interactor of TMEM115 (Ong et al. 2014), albeit one nears the bottom of the BioID list.

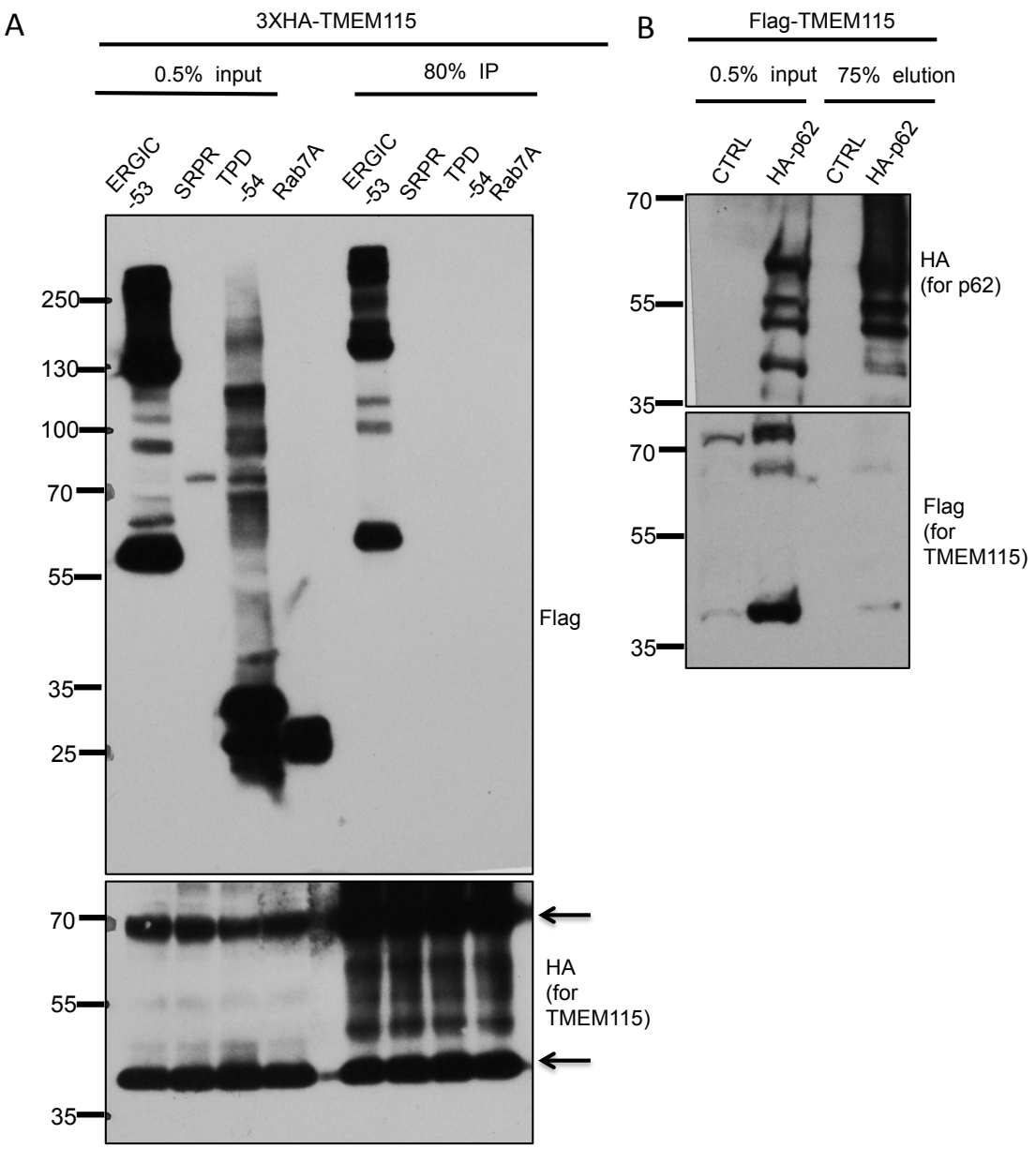


Figure 3.10 BioID candidates validation

A. Co-immunoprecipitation between overexpressed TMEM115 and ERGIC-53. 24 hours before HA immunoprecipitation, HEK293T cells were co-transfected with 3XHA-TMEM115 and ERGIC-53-2XFlag, SRPR-2XFlag, TPD-54-2XFlag and Rab7A-2XFlag, respectively. Cell lysates and HA beads eluate were immunoblotted for Flag (BioID candidates) and HA (3XHA-TMEM115). Only ERGIC-53-2XFlag was detected in HA beads eluate. The band at 38KDa is the monomer of 3XHA-TMEM115 (pointed out with an arrow). The band at around 70KDa is likely to be dimer of HA tagged TMEM115 (pointed out with an arrow).

B. Co-immunoprecipitation between overexpressed TMEM115 and p62. 24 hours before HA immunoprecipitation, HEK293T cells were either co-transfected with empty vector and Flag-TMEM115, or with Flag-TMEM115 and HA-p62. Cell

lysates and HA beads eluate were immunoblotted for Flag (Flag-tmem115) and HA (HA-p62). Flag-TMEM115 was detected only in the HA beads eluate of HA-p62 sample.

The interactions between TMEM115 and both ERGIC-53 and beta-COP were interesting findings but not completely novel, as Ong et al. have already described the interaction. However, it supports the suggestion that TMEM115 may have a role in trafficking since ERGIC-53 is a cargo receptor for the ER-to-Golgi forward trafficking of glycoproteins (Hauri et al. 2000); and beta-COP is a component of the COP-I vesicle that is essential for Golgi-to-ER retrograde trafficking (Beck et al. 2009). In further support of this, there are quite a few candidates in the TMEM115 BioID list that have known roles in regulating protein trafficking, including SCAMP3, COG3 and RAB7A. SCAMP3 (Secretory carrier-associated membrane protein 3) acts as a recycling carrier in the post-Golgi pathways, such as recycles and prevents the lysosomal degradation of EGFR (Quyen L.Aoh et al. 2009). COG3 is a subunit of the highly conserved COG complex, which is a Golgi localised peripheral membrane protein complex that controls processes such as protein sorting, glycosylation, and Golgi integrity (Richard D. Smith et al. 2009). RAB7A is crucial for vesicle transportation across the late endocytic compartments (Bucci C 2000).

The interaction between p62 and TMEM115 has never been identified before. To avoid the possible artefacts caused by co-overexpressing the two proteins, I performed additional IP experiments, in which only one protein was overexpressed. The results were still positive, as endogenous TMEM115 was able to bind to overexpressed p62 and vice versa (Figure 3.11). p62 is a well-studied protein with multiple functions; it is involved in both the ubiquitin-proteasome system (UPS) and macroautophagy/autophagy (Sanchez et al. 1998, Liu et al. 2016, Demishtein et al. 2017). In addition, p62 is also responsible for endosomal organisation necessary to maintain efficient cargo transport (Jongsma et al. 2016). Does the interaction suggest that TMEM115 is involved in one or several of the above functions? Or is a less interesting reason that the interaction is simply

because p62 is acting as a cargo receptor that sends TMEM115 for degradation? To test the latter hypothesis, I performed another IP in which I overexpressed both proteins with supplemented proteasomal inhibitor (MG132) or lysosomal inhibitor (Bafilomycin). No change in the interaction was observed (Figure 3.12). Notably, upon co-overexpression with p62, the level of TMEM115 increases, which is inconsistent with the idea that p62 is responsible for sending TMEM115 for degradation. The biological meaning of the interaction between TMEM115 and p62 is addressed further in Chapter 5.

An intriguing motif (P273-A290) was identified at the cytosolic C-terminus of TMEM115 with HHpred (Soding et al. 2005), which shares high similarity with the VCP/P97 interacting motif at the cytosolic C-terminal tail of RHBDL4 (Lim et al. 2016). Therefore, I tested the interaction between TMEM115 and VCP/p97 with IP; despite that VCP/p97 was not present in the TMEM115 BioID list. The IP result showed that overexpressed TMEM115 was able to bind to endogenous VCP/p97 Figure 3.11A). VCP/p97 is a cytosolic AAA ATPase that governs critical steps in ubiquitin-dependent protein quality control and intracellular signalling pathways (Kobayashi et al. 2007, Ballar et al. 2008, Meyer et al. 2012). VCP/p97 binds to a number of ER localised rhomboid-like proteins that are essential for ERAD, including derlins and RHBDL4 (Ye et al. 2005, Lim et al. 2016). Given VCP/p97 has multiple functions; it is very difficult to predict at this stage the exact function of VCP/p97 in the context of TMEM115. It is unlikely, however, that the loss of TMEM115 leads to a general defect in VCP/p97 related functions. VCP/p97 may act as a cofactor of TMEM115 to regulate, for instance, the degradation of TMEM115 binding partners.

In summary, I identified three physical binding partners for TMEM115, two were from the BioID list, using conventional IP. Of these, ERGIC-53 is a known partner, but p62 and VCP/p97 are completely novel interactors of TMEM115. All of them have suggested possible functions of TMEM115 that require further investigation.

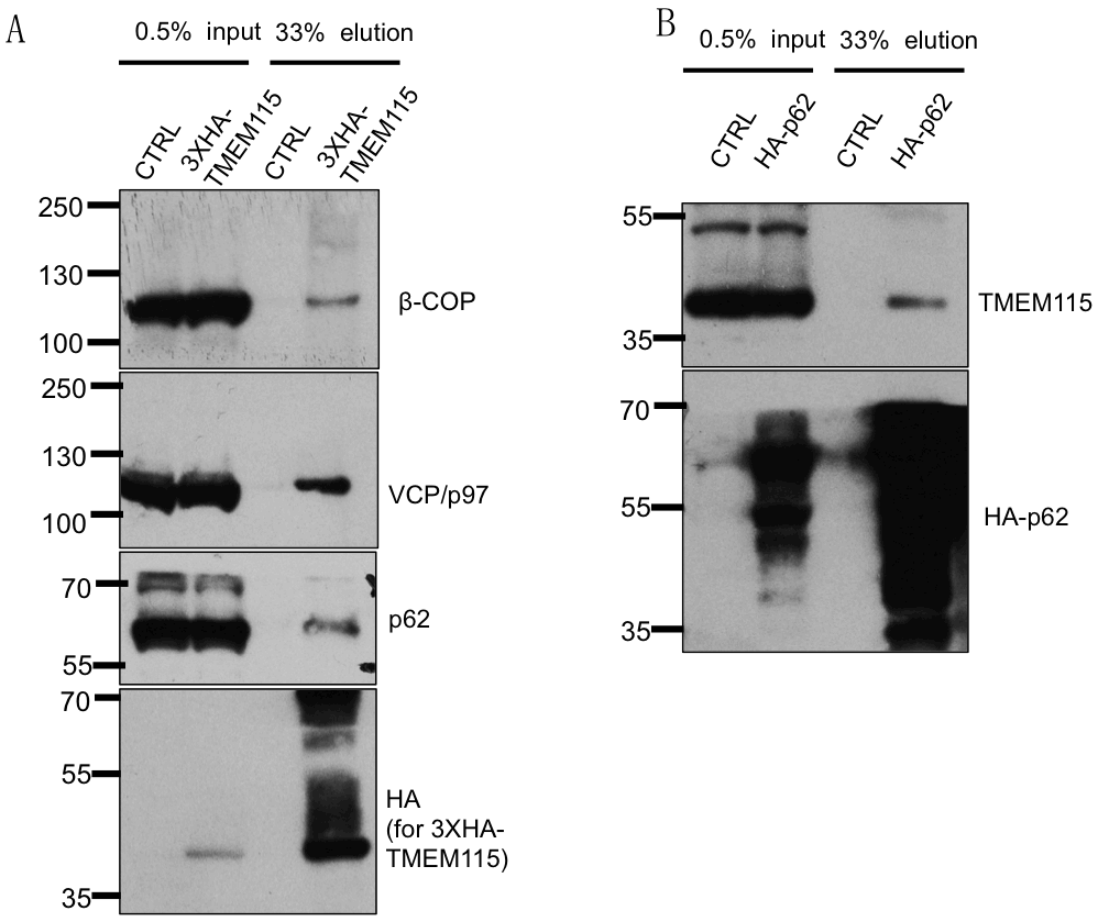


Figure 3.11 BioID candidates validation

A. Co-immunoprecipitation between overexpressed TMEM115 and endogenous β-COP, VCP/p97 and p62. 24 hours before HA immunoprecipitation, HEK293T cells were transfected either with empty vector or with 3XHA-TMEM115. Cell lysates and HA beads eluate were immunoblotted for β-COP, VCP/p97, p62 and HA (3XHA-TMEM115). Endogenous β-COP, VCP/p97 and p62 were detected in HA beads eluate of 3XHA-TMEM115 sample.

B. 24 hours before HA immunoprecipitation, HEK293T cells were either transfected with empty vector or with HA-p62. Cell lysates and HA beads eluate were immunoblotted for TMEM115 and HA (HA-p62). Endogenous TMEM115 were detected in HA beads eluate of HA-p62 sample.

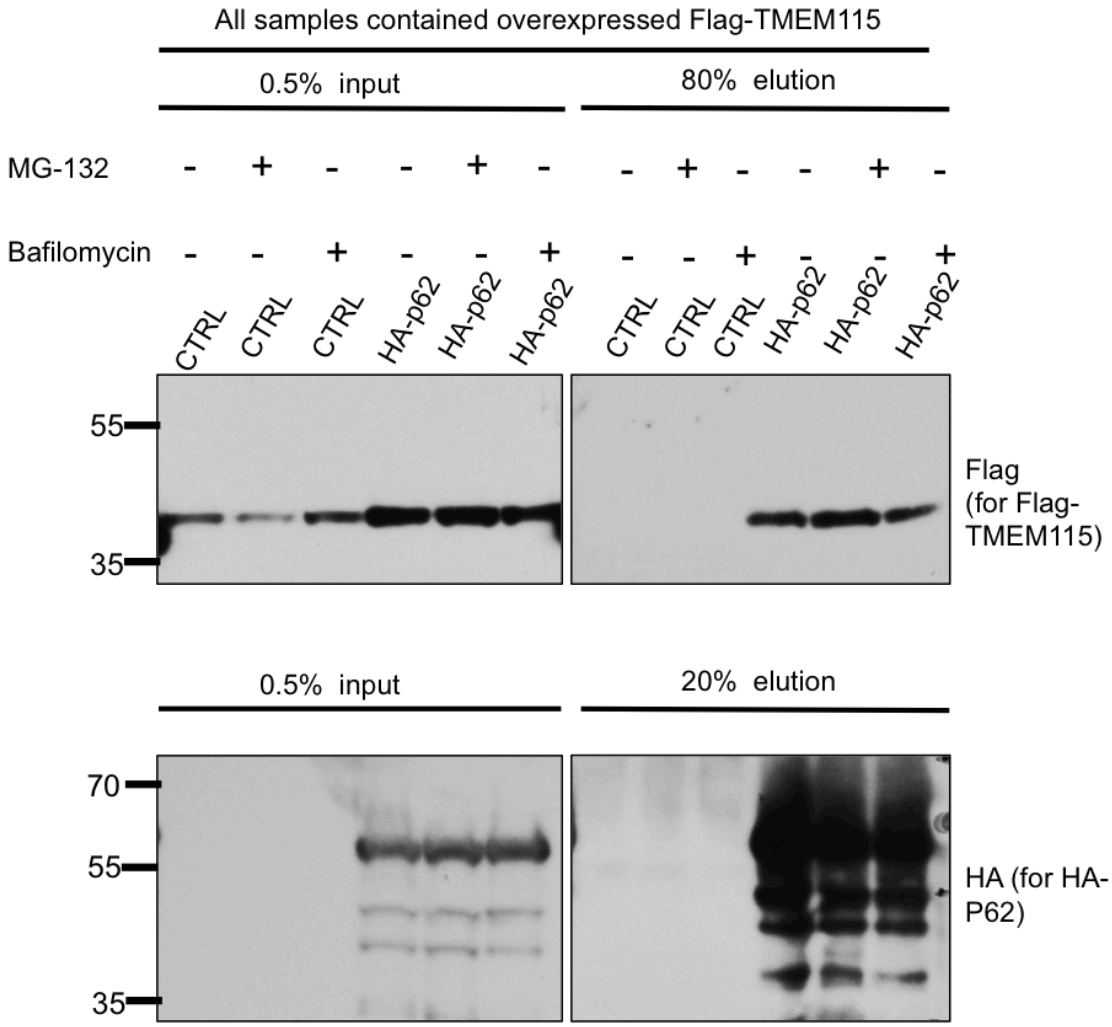


Figure 3.12 The interaction between TMEM115 and p62 does not change upon inhibition of proteasomal or lysosomal degradation
24 hours before HA immunoprecipitation, HEK293T cells were either co-transfected with empty vector and Flag-TMEM115, or with Flag-TMEM115 and HA-p62. Cells were treated with DMSO, MG-132 (10 μM) or Bafilomycin (0.1 μM) 16 hours before HA immunoprecipitation. Cell lysates and HA beads eluate were immunoblotted for Flag (for Flag-TMEM115) and HA (HA-p62). No differences were observed in levels of Flag-TMEM115 in HA beads eluate of MG-132 or bafilomycin samples.

3.7 Investigating the role of TMEM115 in trafficking with SPECS

The Golgi apparatus is the central organelle for protein glycosylation and secretion. TMEM115 is localised predominately within the Golgi, and Ong et al. (Ong et al. 2014) have reported that it may interact with the COG complex, which is the core

component for regulating protein glycosylation within the Golgi (Smith et al. 2008). In addition, fifteen trafficking related proteins were identified in the BioID screen as specific potential interactors for TMEM115. Among these, the physical interactions between TMEM115 with ERGIC-53 and β-COP were validated by co-IP. ERGIC-53 is a mannose-specific membrane lectin that localises specifically in the ER-Golgi intermediate compartment. It is a cargo receptor for transporting glycoproteins from the ER to the ERGIC compartment. The absence of functional ERGIC-53 leads to a selective defect of glycoprotein secretion in cultured cells (Hauri et al. 2000). β-COP is the best known component of the COPI vesicle, which has a vital role in multiple steps of membrane trafficking, especially Golgi-to-ER retrograde traffic (Beck et al. 2009). All the above preliminary data suggested that TMEM115 may have a role in the secretory pathway, such as regulating the protein secretion or glycosylation pathways. However, preliminary data in the Freeman lab has shown that no general glycosylation defect was observed in the absence of Tmem115 in cultured MEFs. The above observations have ruled out the possibility of any widespread or general defect in the secretory pathway linked with loss of TMEM115. However, TMEM115 could still be involved in the trafficking or glycosylation of specific proteins, which could be informative about the possible role of TMEM115 in specific pathways. Therefore, a detailed secretome analysis would be useful to reveal any subtle differences of secretome profiles caused by the absence of TMEM115.

To this end, I employed a recently described secretome analysis, SPECS, to identify potential TMEM115-dependent secreted proteins (Kuhn et al. 2012, Kuhn et al. 2015). A diagram of the standard SPECS screen is shown in Figure 3.13. Approximately 70% of secreted proteins are glycosylated. The traditional secretome analysis method has a fundamental difficulty: high levels of serum proteins and abundant cytosolic proteins released from broken or apoptotic cells often mask the low concentration of secreted proteins. SPECS combines metabolic glycan labelling and click chemistry-mediated biotinylation to enrich for glycosylated secreted proteins. Wild type and *Tmem115* knockout MEFs cultured in the presence of azido sugars produce specifically labelled glycoproteins; these

were further isolated and modified with DBCO-PEG-biotin using a click-chemistry reaction. The biotinylated secreted proteins were then enriched and analysed by mass spectrometry. By comparing the secretome profiles between wild type and *Tmem115* knockout MEFs, I was able to acquire a list of TMEM115-dependent secreted proteins.

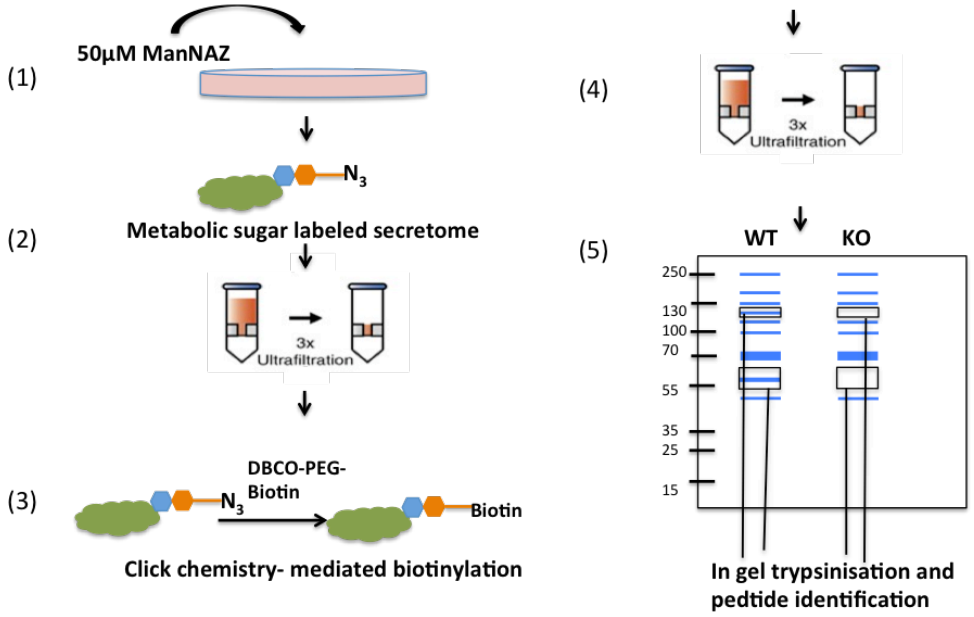


Figure 3.13 Diagram of SPECS approach
ManNAZ is an azido sugar that cells can update and utilised at the terminal position of both N- and O- linked glycosylation. Wild type and *Tmem115* knockout MEFs were cultured in full medium with supplemented ManNAZ (50μM) for 48 hours. Cell medium that contained azido sugar labelled secreted proteins was collected and furthered enriched. The enriched specifically labelled glycoproteins were modified with DBCO-PEG-biotin using a click-chemistry reaction， followed by further enrichment, in gel trypsinisation and Mass spectrometry for peptide identification.

A western blot for biotinylated SPECS samples is shown in Figure 3.12A. Firstly, the biotinylated material in the azido sugar treated samples is higher than untreated samples, which showed the efficiency of the click-chemistry reaction. Notably, there was no obvious global difference between wild type and *Tmem115* knockout MEFs in terms of general glycosylation of secreted proteins: the amount of biotinylated proteins in azido sugar treated wild type and *Tmem115* knockout MEFs samples were similar. The results are consistent among six replicates. The

preliminary results suggest that there are no general glycosylation or secretion defects in *Tmem115* knockout MEFs, which is consistent with the data from our group but contradicts the model suggested by Ong et al. (Ong et al. 2014) that general glycosylation, especially O-linked glycosylation, is affected in TMEM115 knockdown cells. Note, however, that I used MEFs in my assay, while Ong et al. used an actively secreting cell type, primary neurons.

I performed a full-scale SPECS assay, followed by in-gel digestion and mass spectrometry analysis. This attempt was unsuccessful, as the samples were too diluted to produce any meaningful readout. A few optimisation steps were therefore required before I could obtain secretome profiles. According to the suggestions from experts in the mass spectrometry facility, changing sample preparation method, such as processing eluate from beads instead of in-gel digestion, may be effective in sample enrichment. However, after consulting the group who established SPECS, it was impossible to avoid in-gel digestion step. In theory, the click chemistry reaction has specificity such that the DBCO-PEG-biotin should only biotinylate glycoproteins with the azido sugar modification. Unfortunately, it suffers from an off target reaction with a free cysteine residue in albumin, which is abundant in the culture medium. The in-gel digestion step was, therefore, necessary to eliminate the abundant biotinylated albumin at around 60 kDa (arrow in Figure 3.12).

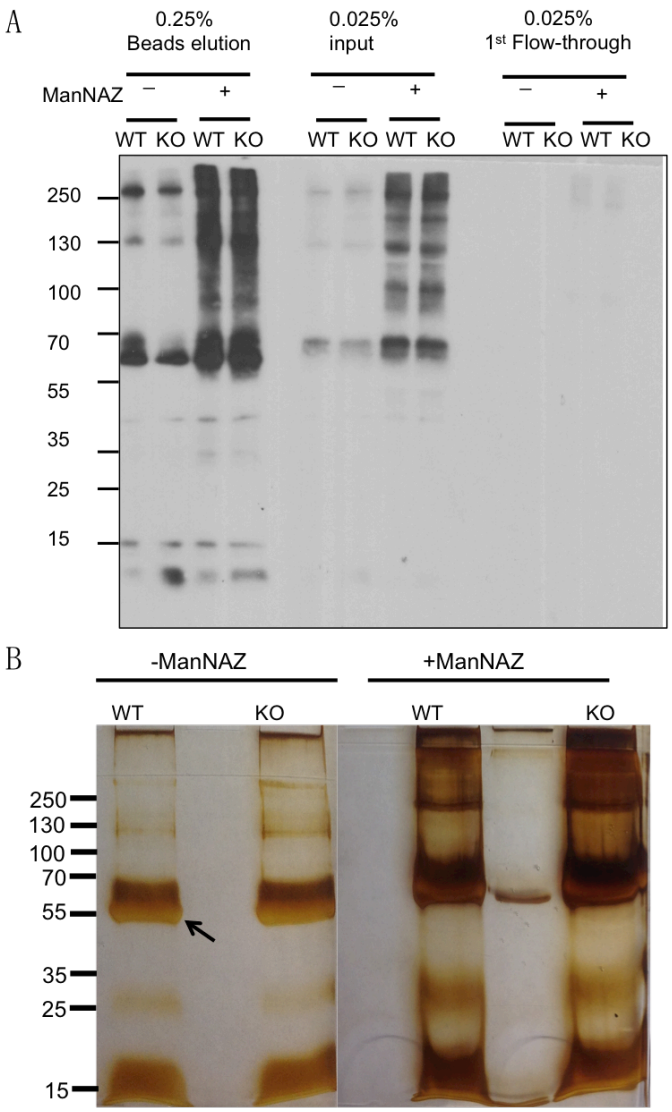


Figure 3.14 SPECS samples from in-gel digestion protocol
Wild type and *Tmem115* knockout MEFs were cultured in full medium with supplemented ManNAZ for 48 hours.
A. Western blot of biotinylated SPECS samples.
B. Silver staining of beads eluate of SPECS samples. The arrow indicts heavily biotinylated albumin due to the off-target effect of click-chemistry in the samples.

I optimised the metabolic sugar labelling condition to eliminate the off-target effect of albumin. Instead of performing the metabolic sugar labelling in full medium (DMEM with 10% FBS), I kept MEFs in Opti-MEM medium, which did not contain albumin. Under this optimised condition, not only was the albumin background eliminated, but also the click-chemistry efficiency was massively increased, probably because no albumin was present to compete for DBCO-PEG-biotin (Figure 3.13). I performed triplicates of SPECS using the optimised protocol, using two different samples of wild type and *Tmem115* knockout cells. Although the results were not reproducible between the two sets of MEFs, after cross comparison, no TMEM115 dependent secreted proteins were identified. Although it does not rule out the possibility that TMEM115 may regulate the trafficking of some proteins if they are not glycosylated, or at very low levels, or not secreted to the extracellular compartment, we can conclude that the absence of TMEM115 is unlikely to have a substantial defect in glycosylation or protein secretion.

In summary, I was not able to identify any reproducible, measurable differences in the level of proteins secreted from wild type and *Tmem115* knockout MEFs cells. Therefore, my data imply that this Golgi-localised protein does not have a role in controlling general protein secretion or glycosylation.

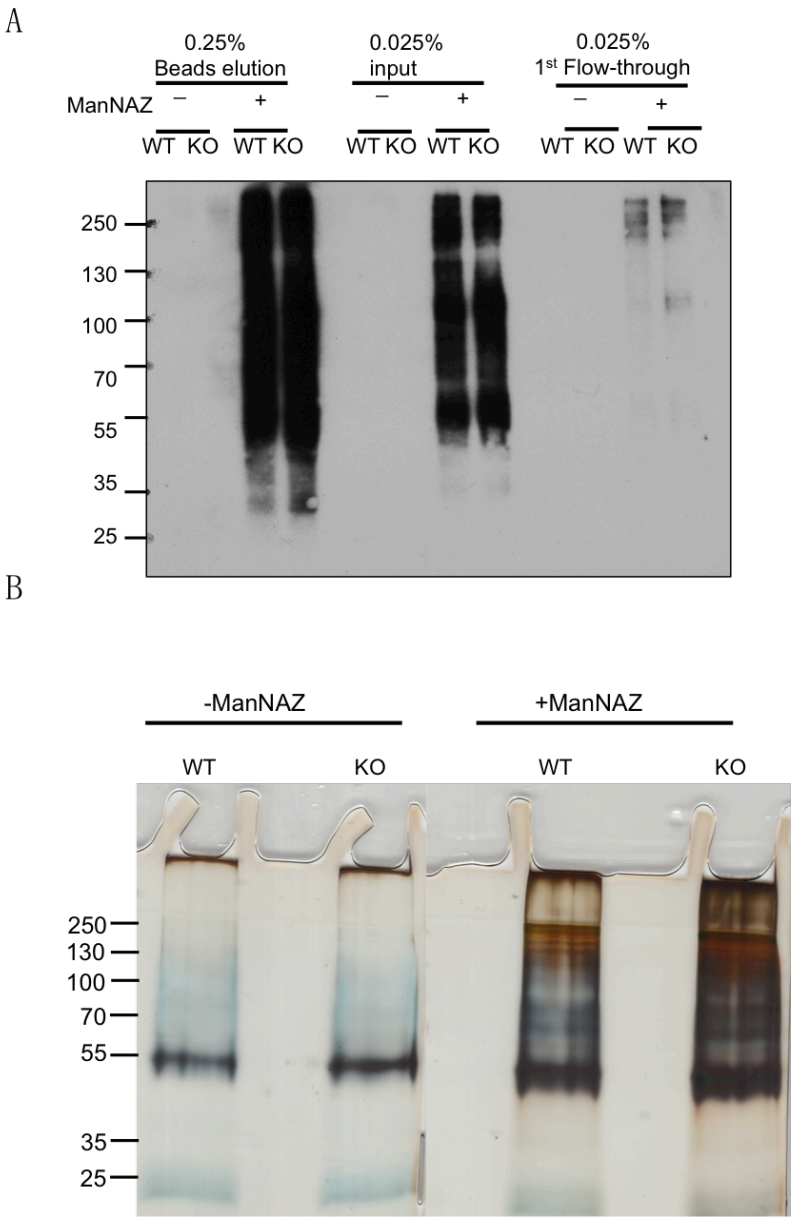


Figure 3.15 SPECS samples for optimised protocol
Wild type and *Tmem115* knockout MEFs were cultured in FBS free medium with supplemented ManNAZ for 48 hours.
A. Western blot of biotinylated SPECS samples.
B. Silver staining of beads eluate of SPECS samples.

3.8 Discussion

3.8.1 BioID is a powerful tool for identifying interactors to reveal novel functions of proteins

Proteins usually perform biological functions through forming complexes, or by protein-protein interactions (PPIs). PPIs can be either static or dynamic (Nooren et al. 2003). To reveal the biological functions of a protein, many approaches, including yeast two-hybrid assay, phage display and co-immunoprecipitation, have been established. These reveal information about PPIs and disclose the composition of protein complexes, mostly from the static prospective. Despite the value of these approaches, none of these methods captures real time dynamic information of PPIs. Several Proximity-dependent labelling approaches have been proposed to address this issue, among which, BioID is the currently most extensively used method.

As a proximity-dependent labelling strategy, compared to co-immunoprecipitation, BioID is particularly useful for identifying novel interactors for insoluble membrane protein complexes, as well as identifying transient and weak interactors (Roux et al. 2012, Kim et al. 2014, Kim et al. 2016, Kim et al. 2016). Co-immunoprecipitation is a widely used antibody-based pull down method for studying PPIs. It relies on the solubilisation of stable intact protein complexes for selective affinity purification by antibodies. The lysis buffers for co-immunoprecipitation assays usually contain a low percentage of mild detergent to retain the PPIs, while membrane proteins are hardly resolved in mild lysis buffers, which reduces the efficiency of interacting protein capture significantly. Another disadvantage of co-immunoprecipitation is the capture is performed in cell lysates, and cannot distinguish direct interactions and unspecific binding caused by the nature of membrane proteins, which increases numbers of background proteins for further validation steps (Berggard et al. 2007). The labelling process of BirA*, however, is performed *in vivo* and is not affected by the solubility of the protein complexes. In additional, BirA* biotinylates nearby proteins with a proximity dependent manner, thereby capturing strong, weak and transient PPIs.

The sequential biotinylated protein capture with streptavidin is performed in stringent buffers that contains a much higher concentration of detergents to reduce unspecific binding during the capture step, as the binding affinity between streptavidin and biotinylated proteins is extremely high(Weber et al. 1989, Roux et al. 2012).

BioID is a robust tool for studying PPIs, however, it has some disadvantages. The large BirA* (321 amino acids) tag at terminal positions may affect the folding and function of target proteins. In the case of TMEM115, the C-terminus BirA* tagged TMEM115 was not expressed at the predicted size (74 kDa) and failed to localise within the Golgi, indicating that the correct folding of TMEM115 was disrupted by BirA*. To overcome this issue, the Roux lab improved BioID and created BioID2, which was based on a smaller biotin ligase (233 amino acids) *from A. aeolicus* (Kim et al. 2016). BioID2 is comparable to BioID in terms of biotinylation activity, and the subcellular localisation of some proteins of interest improved with this smaller biotin ligase in their study (Kim et al. 2016). In addition, the proximity labelling has to rely on the presence of lysine residues in candidate proteins at an accessible region, which may lead to false negatives. For instance, in our TMEM115 BioID screen, BirA was tagged at the cytosolic N-terminus of TMEM115; it would not be able to biotinylated candidate proteins without any lysine residues, or which have lysine residues only in luminal domains. Another disadvantage is the background biotinylation caused by the localisation of the target protein. As mentioned in Section 3.5, TMEM115 and another Golgi rhomboid-like protein, RHBDD2, share 90% of the top twenty candidates in their BioID lists ranked by spectral counts, most of the shared candidates are abundant Golgi resident proteins. Apart from the possibility that TMEM115 and RHBDD2 share many mutual interactors, given that both of them are rhomboid-like proteins, another possible explanation for the high percentage of overlap in interactome profiles is that BirA*-TMEM115 labels abundant Golgi proteins unspecifically. Therefore, in addition to having soluble BirA* as a control, a more useful control is a BirA* tagged protein from another family that localises in the same subcellular compartment, for the elimination of organelle-specific background.

3.8.2 Possible biological roles of TMEM115 according to BioID result

The BioID screen aimed to identify interacting proteins that may help to reveal the cellular functions of TMEM115. Four possible roles emerged after classifying BioID candidates by functions defined by GO annotation: trafficking, lipid regulation, protein degradation and ion channel regulation.

The possible role of TMEM115 in trafficking has been investigated by Ong et al., and they reported that proper expression level of TMEM115 is required for Brefeldin A-induced Golgi-to-ER retrograde transport, together with a role in correct O-linked glycosylation (Ong et al. 2014). However, Dr Moncada Pazos was not able to observe general defects in O-linked glycosylation or Golgi-to-ER retrograde transport in *Tmem115* knockout MEFs. Consistent with these observations, the SPECS screen (section 3.7) was also not able to identify any reproducible measurable differences regarding secretome profiles between wild type and *Tmem115* knockout MEFs. My data imply that the loss of TMEM115 does not cause a defect in general glycosylation or secretion. While it certainly does not rule out the possibility that TMEM115 regulates the trafficking of some of its interactors, given that it is a rhomboid-like protein, several members of the same superfamily regulate crucial pathways through regulating the transportation of their interactors. For instance, mammalian iRhoms promote the forward trafficking of TACE and activation of the downstream TNF pathway (Adrain et al. 2012).

A role of TMEM115 in lipid regulation is supported by the phenotypes of both *Tmem115* knockout mice and files, together with other preliminary data from Freeman lab. This suggested that SCAP-SREBP pathway, which is crucial for sterol homeostasis in vertebrates (Nohturfft et al. 1999), is dysregulated in the absence of TMEM115. However, SCAP, as a known binding partner, was not present in any of the BioID lists. It could be due to the low abundance of biotinylated endogenous SCAP, or it may be explained by the unrevealed mechanism of how TMEM115 regulates SCAP. For instance, if TMEM115 interacts with SCAP to coordinate the degradation of SCAP, it would be difficult to

identify biotinylated SCAP without inhibitors for blocking protein degradation. The possible role of TMEM115 in lipid biology and its regulation of SCAP is further addressed in Chapter 5.

Protein degradation and ion channel regulation are two possible novel roles of TMEM115. Five proteins that are involved in protein degradation (GO annotation) have been identified in BioID. In addition, HHpred, the Helix-Helix profile server for structural homologue searches (Soding et al. 2005), identified a region in TMEM115 C-terminus (Pro273 to Ala290) with high structural similarity with the well characterised VCP-interacting motif (VIM) (Hanzelmann et al. 2011). The interaction between TMEM115 and VCP/p97 was validated with co-immunoprecipitation. Many rhomboid-like family members are involved in protein quality control machinery through both proteasomal and lysosomal degradation (Zettl et al. 2011, Fleig et al. 2012, Freeman 2014, Mehnert et al. 2014). Derlins, iRhom and UBAC2 are either components of ERAD, or utilize ERAD for regulating downstream signalling pathways (Greenblatt et al. 2011, Zettl et al. 2011, Christianson et al. 2012), while Grieve et al. also reported a novel role of mammalian iRhom2 in protecting mature TACE at the plasma membrane from lysosomal degradation (Grieve et al. 2017). Therefore, it is possible that TMEM115 may also be involved in some form of degradation machinery, which is explored further in Chapter 5. The last possible function is that TMEM115 has a role in regulating ion channels, but we do not yet have any supporting data for this idea. Interestingly, among the eight TMEM115 unique candidates that linked with ion channel functions, five are involved in regulating calcium transport (according to Go annotations). It is an intriguing possibility that TMEM115 may be related to calcium signalling. More experiments are required to investigate the possible relationship between calcium signalling and TMEM115 functions.

Chapter 4 Topology study of TMEM115 and analysis of the contribution of specific domains to its known interactions

Introduction:

4.1 The importance of studying TMEM115 topology

All members of the rhomboid-like superfamily have at least have six TMDs (some members have a seventh extended TMD), which constitutes the core conserved rhomboid-like domain across the whole superfamily (Freeman 2014). To unambiguously assign TMEM115 as a member of the rhomboid-like superfamily, it is necessary to prove the presence of this rhomboid-like domain. By sequence similarity, TMEM115 does indeed appear to be a member of the superfamily (Lemberg 2013, Freeman 2014), but, in a recent paper about TMEM115, Ong et al. reported a four TMD structure and did not report the possibility of it being a rhomboid-like protein (Ong et al. 2014). Whether TMEM115 has six TMDs, or is the first rhomboid-like protein that has been identified with only four TMDs, or was in fact not a true member of the superfamily was a crucial question to address. Therefore, the topology of TMEM115 needed to be mapped experimentally. Moreover, the detailed topology information of TMEM115 would also facilitate the further functional study of this protein, for instance, to map the domains of TMEM115 essential for binding to its interactors.

In this topology study section, I aimed to answer several questions.

1. How many transmembrane domains does TMEM115 have and where are they located?

2. What are the lengths of both termini of TMEM115? What are the orientations of both termini?

3. What are the locations and sizes of the loops? What are their orientations?

4.2 The importance of interaction domain study

Members of the rhomboid-like superfamily are known to interact with other membrane domain containing proteins, often through their TMDs. Indeed, it is possible that the characteristic functional feature of the rhomboid-like domain is TMD recognition (Freeman 2004). In the case of rhomboid proteases, they are intramembrane serine proteases that cleave their substrates within the TMD. Most of their known substrates are single-pass transmembrane proteins. What determines the substrate specificity is the recognition motif in their TMD region, which is recognised by the core TMD region of rhomboids (Strisovsky et al. 2009). Replacing the TMD of a single pass transmembrane protein, which cannot be cleaved by rhomboids, with a known rhomboid substrate TMD, is sufficient to create a rhomboid substrate (Urban et al. 2003). Therefore, TMD-based interactions are crucial for the functions of rhomboid proteases and, following an evolutionary logic, may also be crucial for many inactive rhomboid-like proteins.

Although the TMD region is the key domain for rhomboid-like proteins, the less conserved regions, including extended termini and large cytosolic or luminal loops of some rhomboid-like proteins are also known to have important regulatory functions. In the case of iRhom2, one of its most important interactors, TACE, is regulated and trafficked by iRhom2 from ER to the plasma membrane, where TACE cuts and releases its substrates from their membrane-tethered precursors. The long cytosolic N-terminus of iRhom2 is shown to have a crucial role in interacting and regulating the function of mature TACE, while the first loop, the iRhom IRHD, is required for binding to the immature form of TACE (Grieve et al. 2017). Regarding rhomboid proteases, the conserved six-TMD domain is sufficient for their proteolytic activity, while their cytoplasmic domains are crucial for modulating enzyme activity. The amino-terminal cytoplasmic domain of *Drosophila* rhomboid-4 contains EF hands, which co-ordinate binding of calcium ions, which limit its proteolytic activity substantially, possibly through modulating gating to active sites (Baker et al. 2015). Accordingly, removal of the EF-hands causes

upregulated proteolysis of the normally calcium-stimulated activity of rhomboid-4 against its substrate Spitz.

Given the known importance of the transmembrane domains, the N- and C-termini, and the large inter-TMD loops in regulating the functions of other rhomboid-like proteins, I made different truncations of TMEM115 to assess their contribution to known interactions. A number of TMEM115 interacting proteins, including both transmembrane proteins and soluble proteins, were identified by BioID and were validated by the co-immunoprecipitation methods, as described in Chapter 3. With the help of different truncations, I was able to investigate the interacting domain of TMEM115 with its soluble or transmembrane interactors. The approach has brought useful information of the interaction regions between TMEM115 and its interactors.

Results

4.2 Bioinformatic analysis for predicting TMEM115 topology

Ong et al. based their proposed topology of TMEM115 on the use of a single prediction tool, TMpred (Soding et al. 2005). Predicting the TMDs of a polytopic membrane protein is difficult because the TMD-TMD packing tends to allow more polar and charged residues than would normally present in a single TMD, where the whole helix is in contact with the lipid bilayer of the membrane. I, therefore, used several bioinformatics tools in an attempt to predict the topology of TMEM115. A protein sequence alignment was performed between TMEM115 and GlpG, an *E. coli* rhomboid of known structure. However, the alignment did not provide meaningful information about transmembrane domain identification, due to these two proteins being too evolutionarily distant from one other. Since transmembrane domains are the most conserved regions amongst TMEM115 orthologues, I did an alignment of the sequences of TMEM115 among four different species, human, mouse, *Drosophila* and yeast (Figure 4.1A). According to the alignment, TMEM115 is a highly conserved protein across species. Therefore, it is challenging to identify the most conserved region across the whole protein, which should be the location of transmembrane domains. More

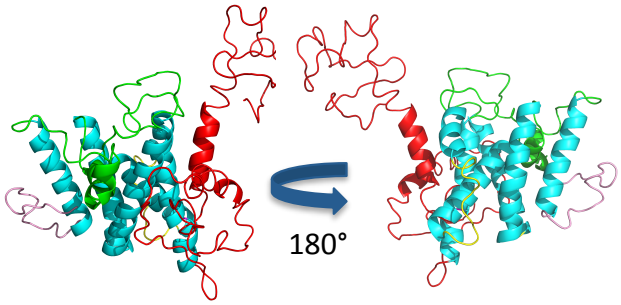
comprehensive domain and motif identification tools were needed to predict the locations and numbers of transmembrane domains for TMEM115.

In the next step, to gain more insight about the topology of TMEM115 and to look for structural similarity between TMEM115 and the rhomboid-like proteins, I submitted the protein sequence of TMEM115 to the Phyre2 protein folding structure prediction server to predict its 3D structure. This tool found that the *E.coli* rhomboid with a known crystal structure, GlpG, was the protein (among the whole protein crystal structure database) that shares the highest structure similarities with TMEM115. It also provided a TMEM115 3D structure prediction with six TMDs, using the GlpG structure as a base and with over 80% confidence for the first 51% of the full-length protein. Regarding the details of the 3D structure model, it predicts that three relatively large luminal loops (labelled in green in Figure 4.1 B), in which, loop 1 is the largest, containing 30 amino acids. The two cytosolic loops (labelled in yellow in Figure 4.1B), on the contrary, are very small. The cytosolic loop 2 only consists of one amino acid. What is consistent with the four topology prediction models is the length of cytosolic C-terminus. The high structural similarity between large regions of the two proteins, TMEM115 and GlpG, which are quite distant to each other evolutionary, suggests that TMEM115 has the rhomboid-like domain that is conserved across the rhomboid-like superfamily.

A

```
sp|Q12239|TM115_YEAST       MQYSSRFLELNIPDSFLNINKIPDATKFITVTYICLTATLFCIRRSLYNKLVLEDPNLDY
tr|Q9VMD2|Q9VMD2_DROME      ---MSAQLARNWPYIWQQLNALLHNTS-PVITLICVVTTFGYL-------LSFSETAI-L
sp|Q12893|TM115_HUMAN       -------MQRALPGARQHLGAILASAS-VVVKALCAAVLFLYL-------LSFAVDT--G
sp|Q9WUH1|TM115_MOUSE       -------MQRALPGARQHLGAILASAS-VVVKALCAVVLFLYL-------LSFAVDT--G

sp|Q12239|TM115_YEAST       NLITSPLLQMVPSQIWRYPTSLVLSNFIDTKAWKVVVNLLNLIIGGSFIERNWNSSKEMF
tr|Q9VMD2|Q9VMD2_DROME      LLSVTPGYILPNGKFWIWT--AFTFCFIELHWWEVAVDVVTVGLCGKMLEPLWGQL-EMF
sp|Q12893|TM115_HUMAN       CLAVTPGYLFPP-NFWIWT--LATHGLMEQHVWDVAISLTTVVVAGRLLEPLWGAL-ELL
sp|Q9WUH1|TM115_MOUSE       CLAVTPGYLFPP-NFWIWT--LATHGLMEQHVWDVAISLATVVVAGRLLEPLWGAL-ELL

sp|Q12239|TM115_YEAST       KFIIVLGSLTNVLII---MLTLLVSFFSNKVRLDIPLDGNYTILIGFPIIYRQLLPETTI
tr|Q9VMD2|Q9VMD2_DROME      KFFALSNFGVSLLTTVYYLFYYMV-TKNPTILFEVHIHGLAGYVAGICVAVRQIMPDHLI
sp|Q12893|TM115_HUMAN       IFFSVVNVSVGLLGAFAYLLTYMA-SFNLVYLFTVRIHGALGFLGGVLVALKQTMGDCVV
sp|Q9WUH1|TM115_MOUSE       IFFSVVNVSVGLLGALAYLLTYMA-SFNLVYLFTIRIHGALGFLGGVLVALKQTMGDCVV

sp|Q12239|TM115_YEAST       IHLKTPQFLAKNFRFKLLPIFVMFTMTVTQIIWFHHFA---QLFSIWVTFFASWSYLRFF
tr|Q9VMD2|Q9VMD2_DROME      FKTRYGRLTNRNVPLTVLIM--------AIILWAIGLLDGTYPAMFASGSLVSWIYLRFY
sp|Q12893|TM115_HUMAN       LRVPQVRVS--VMPMLLLAL--------LLLLRLATLLQSPALASYGFGLLSSWVYLRFY
sp|Q9WUH1|TM115_MOUSE       LRVPQVRVS--VVPMLLLAL--------LLLLRLATLLQSPALASYGFGLLSSWVYLRFY

sp|Q12239|TM115_YEAST       QKLAPLNCPSLPTTNSQGGQEILVGDASDTFQLIYFFPDLIKPILRPIFNFIYNVVVVKF
tr|Q9VMD2|Q9VMD2_DROME      QHHPN----------------GRGDSSESFTFVSFFPNVSQPFISVLVNPIYNCCLRAG
sp|Q12893|TM115_HUMAN       QRHSR----------------GRGDMADHFAFATFFPEILQPVVGLLANLVHSLLVKVK
sp|Q9WUH1|TM115_MOUSE       QRHSR----------------GRGDMADHFAFATFFPEILQPVVGLLANLVHGLLVKVK

sp|Q12239|TM115_YEAST       KVI--KPFHDIDIDIGNTIAESRGAKKIMTVEERRQLALQVLEERMVNP----------
tr|Q9VMD2|Q9VMD2_DROME      VVKTPTPLRTISTASLTSVSVQMPG-VDPHDIERRRQIALKALSERLKATDSSRHAQLPK
sp|Q12893|TM115_HUMAN       ICQ--KTVKRYDVGAPSSITISLPG-TDPQDAERRRQLALKALNERLKRVEDQSI-----
sp|Q9WUH1|TM115_MOUSE       ICQ--KTVKRYDVGAPSSITISLPG-TDPQDAERRRQLALKALNERLKRVEDQSA-----

sp|Q12239|TM115_YEAST       ------------------------------------------------------------
tr|Q9VMD2|Q9VMD2_DROME      SFPQQQQLQHHHHQHAPHQQQKHHSHGTGHSHSHGAGHSHSHGAGHSHGPGPAQMPQPDF
sp|Q12893|TM115_HUMAN       -WPSMDDDE--------------EESGAKVDS-------PLPSDKAPTPPGKGAAPESSL
sp|Q9WUH1|TM115_MOUSE       -WPSMDDDE--------------EEAGAKTDS-------PLPLEEASTPPGKVTVPESSL

sp|Q12239|TM115_YEAST       ------------------------------------------------------------
tr|Q9VMD2|Q9VMD2_DROME      LKSTGSASQLPITTSRAEPRMISTMSPIAIPMPAPPPKEGNVPGQGAESSSSGGVEATLI
sp|Q12893|TM115_HUMAN       ITFEAA-----------PPTL------------------------------------
sp|Q9WUH1|TM115_MOUSE       ITLETA-----------PLL------------------------------------

sp|Q12239|TM115_YEAST       ------------
tr|Q9VMD2|Q9VMD2_DROME      NLDDVPTTSSMA
sp|Q12893|TM115_HUMAN       ------------
sp|Q9WUH1|TM115_MOUSE       ------------
```

B



MQRALPGARQHLGAILASASVVVKALCAAV
LFLYLLSFAVDTGCLAVTPGYLFPPNFWIW
TLATHGLMEQHVWDVAISLTTVVVAGRLLE
PLWGALELLIFFSVVNVSVGLLGAFAYLLT
YMASFNLVYLFTVRIHGALGFLGGVLVALK
QTMGDCVVLRVPQVRVSVMPMLLLALLLLL
RLATLLQSPALASYGFGLLSSWVYLRFYQR
HSRGRGDMADHFAFATFFPEILQPVVGLLA
NLVHSLLVKVKICQKTVKRYDVGAPSSITI
SLPGTDPQDAERRRQLALKALNERLKRVED
QSIWPSMDDDEEESGAKVDSPLPSDKAPTP
PGKGAAPESSLITFEAAPPTL

Figure 4.1 The topology of TMEM115 by bioinformatics predictions
A. The alignment of TMEM115 among human, mouse, Drosophila and yeast.
B. Phyre 3D structural prediction of TMEM115. Both termini are labelled in pink. All the predicted TMDs are in blue. Luminal loops are in green and cytosolic loops are in yellow.

Finally, I applied four different topology prediction algorithms for TMEM115 and obtained four different possible models. After aligning the predicted results, despite the differences in the locations and numbers of TMDs, it was consistent that in all of the models, both termini face the cytosol. Another consensus among the four models was that the TMDs were predicted to be predominantly localised in the N-terminal portion of TMEM115. Of the four programs I employed, Toppred, TMHMM and MEMSAT, predicted models with six TMDs and relatively short small loop regions. However, each TMD is predicted to be located slightly differently across the protein. The last software, SOSUI, predicted a four TMD model with extended loops, not dissimilar to that stated in Ong et al. (Figure 4.2). When compared with the predicted 3D structural model, in spite of some overlapping TMD regions, none of the four predicted topology models has exactly the same TMD and loop regions as the 3D model.

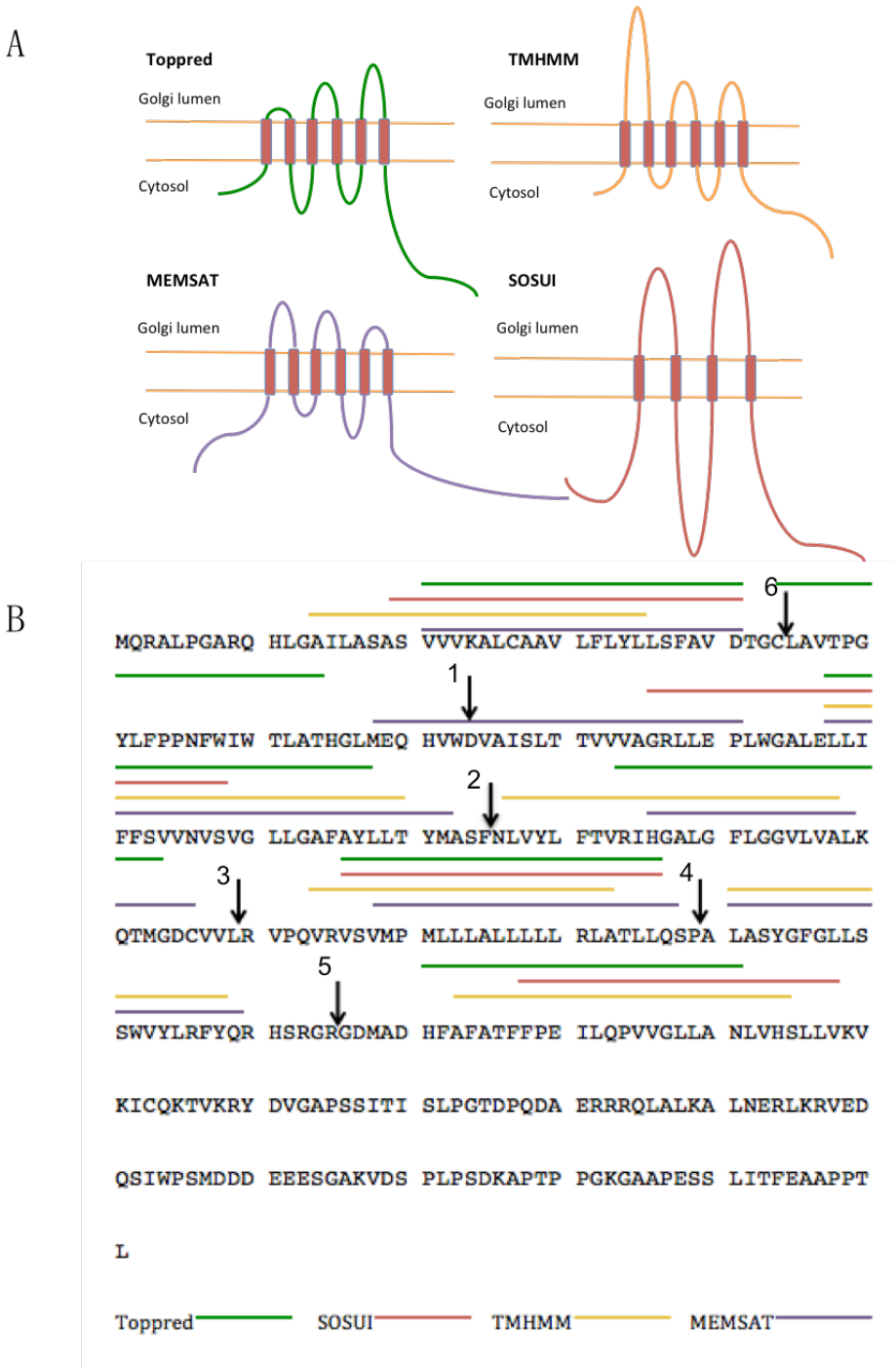


Figure 4.2 The topology model of TMEM115 by four different prediction algorithms

A. Predicted TMEM115 topology by Toppred, TMHMM, MEMSAT and SOSUI

B. The alignment of four predicted topology models from four different algorithms. The coloured lines are the predicted TMD regions of each model. The arrows indicate the selected sites for inserting tags.

Altogether, TMD prediction algorithms and the 3D structure prediction show that although TMEM115 is evolutionarily distant from rhomboid proteases, and is not highlighted by simple sequence similarity searches (like BLAST), TMEM115 probably is a rhomboid-like protein with six TMDs. The preferred model of Ong et al. is a four TMDs model (Ong et al. 2014), although notably, they do not provide any experimental support for this proposal. Despite this set of new predictions, the inconsistencies between different models, the disagreement with what had already been published, and the overall need to assign the protein topology to allow rigorous identification of TMEM115 as a member of the rhomboid-like family, made me decide to examine the topology of TMEM115 experimentally.

4.4 Establishing selective permeabilisation conditions to localise TMEM115 domains

Based on the different predictions, I initially designed five tagged constructs in which epitope tags were inserted in different positions within the loops predicted by different models. By assessing whether each of these tags was cytoplasmic or luminal, I was able to distinguish among different predicted models. To determine whether an inserted epitope tag is localised at the cytosolic or luminal part of a Golgi membrane protein, I established a selective permeabilisation strategy, in which two different detergents were used. The lipid composition of the cellular membrane system is not uniform, but rather distinctive. For instance, cholesterol is enriched in the plasma membrane but not in the Golgi membrane. Some detergents, such as digitonin, have a preference to extract particular types of lipid from membranes; digitonin selectively extracts membrane cholesterol, while other detergents, such as Triton X-100 (TX-100), are not very selective and extract all membrane lipids. By using different detergents, I either partially (only permeabilising the plasma membrane) or fully permeabilised the cells. TGN46, a single-pass type I Golgi membrane protein, was used as a Golgi luminal control to set up the conditions. TGN46 has a large luminal domain, a transmembrane domain and a small cytosolic tail. The TGN46 antibody I used in this assay can only recognise the luminal domain of the protein. Therefore, TGN46 can only be

detected by immunofluorescence if both the plasma membrane and the Golgi membrane are permeabilised. With 0.2% TX-100, cells are fully permeabilised so that both luminal and cytosolic tags are accessible to antibodies. Under this treatment, TGN46 can be detected by immunofluorescence. Conversely, at low concentrations of digitonin, the plasma membrane is permeablised but the Golgi membrane is left intact, a condition whereby luminal protein regions are inaccessible to antibodies. Under this condition, TGN46 cannot be detected since the Golgi membrane is impermeable for anti-TGN46 (a diagram for this conditional permeabilisation is presented in Figure 4.3 A).

I optimised the concentration of digitonin suitable for conditional permeabilisation, using TGN46 labelling as readout. Two different digitonin concentrations were used in similar studies in published papers: 25 and 40 μg/ml. I found that TGN46 is accessible to anti-TGN46 under 0.2% TX-100 and 40 μg/ml of digitonin but not with 25 μg/ml of digitonin treatment, which shows the partial permeabilisation condition is set up correctly (Figure 4.3B). 25 μg/ml of digitonin was then chosen for conditional permeabilisation experiments that are described next.

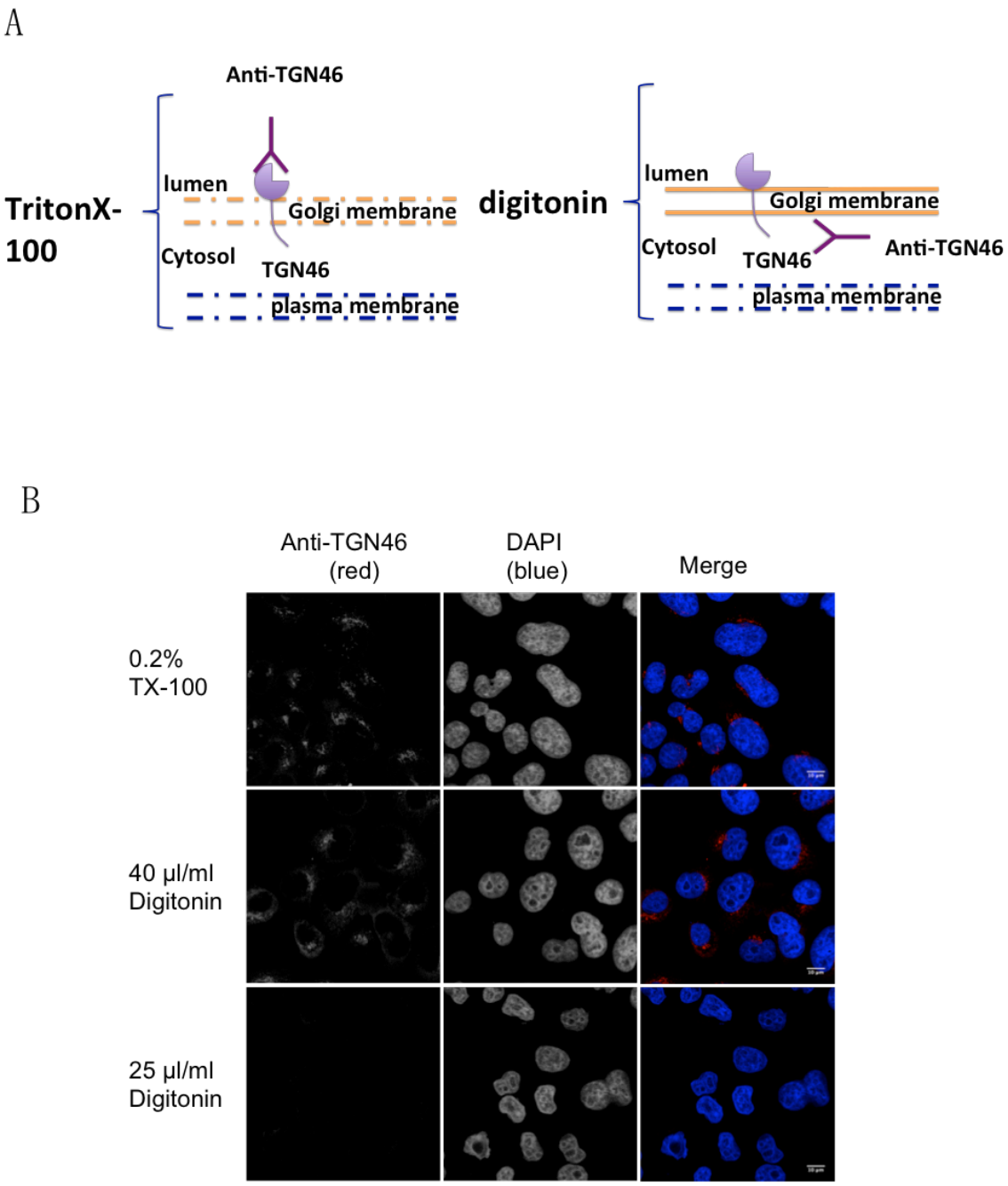


Figure 4.3 Optimising conditions for TMEM115 topology study
A. Model of conditional permeabilisation with TX-100 or digitonin. The endogenous Golgi protein, TGN46, is selected as a luminal control.
B. HeLa cells were permeabilisated with 0.2% TX-100, 40 or 25 µl/ml of digitonin. TGN46 was used to optimise the concentration of digitonin for conditional permeabilisation. Scale bar (at the bottom left of the images): 10 µm

4.5 The localisation of both TMEM115 termini

According to all the predicted TMEM115 topology models, both termini of TMEM115 are cytosolic. I was able to examine the orientation of both termini of TMEM115 experimentally with the established selective permeabilisation conditions. In this study, TMEM115 was either N-terminally (3xHA-TMEM115) or C-terminally (TMEM115-3xHA) tagged with triple HA tags. By western blot, I tested whether the triple HA tags added at either terminus of TMEM115 would affect the production level of the protein. In Figure 4.4, both 3xHA-TMEM115 and TMEM115-3XHA were expressed at the expected size (around 40 kDa). Notably, the expression level of 3XHA-TMEM115 and TMEM115-3XHA are equivalent when detected by the anti-TMEM115 antibody. Interestingly, if the same protein samples were blotted with the anti-HA antibody (Figure 4.4A), TMEM115-3XHA gave much more intensive signal than 3XHA-TMEM115. This may be because the C-terminus of TMEM115 is much longer than its N-terminus (based on all the predicted models), which would provide easier accessibility for antibody to detect its epitopes. It indicates that the position where tags are inserted across a protein can affect signal intensity significantly in some cases.

I applied the conditional permeabilisation method to examine the location of both termini of TMEM115. TGN46 was used both as a Golgi marker and as a luminal control in every experiment to confirm that the partial permeabilisation was successful. Under 0.2% TX-100 permeabilisation (Figure 4.4B), HA signal of both termini of TMEM115, together with TGN46, can be detected. In addition, the different tagged versions of TMEM115 co-localised perfectly with both the Golgi marker TGN46 and endogenous TMEM115. This indicated that tags added at either terminus of TMEM115 did not affect the folding and the cellular localisation of the protein. Under 25 μg/ml digitonin treatment, TGN46 signal was undetectable as expected, while HA signal for both termini of TMEM115 was observed, indicating that both termini are cytosolic. This is consistent with all the predicted models and the observation of Ong et al. (Ong et al. 2014) that the C-terminus is cytosolic.

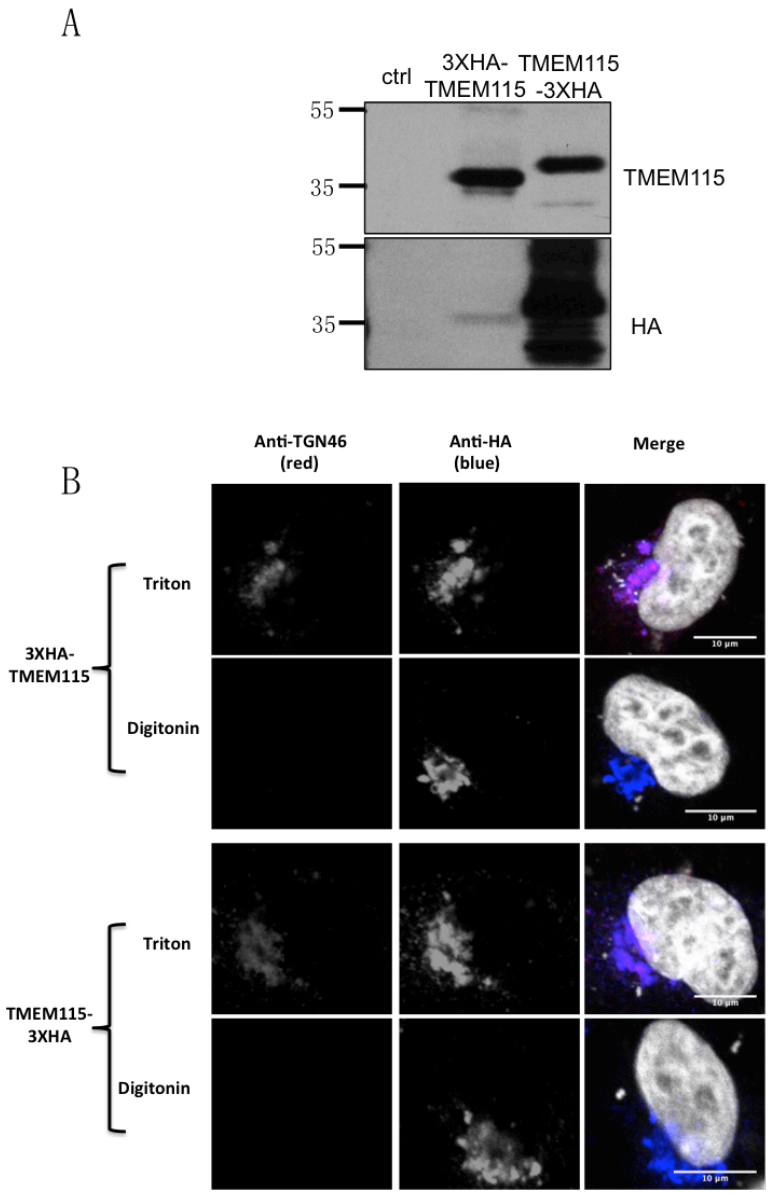


Figure 4.4 Both termini of TMEM115 are cytosolic

A. Western blot for lysates of HeLa cells that were transfected with 3XHA-TMEM115 and TMEM115-3XHA. Cell lysates were immunoblotted for TMEM115 and HA (for 3XHA-TMEM115 and TMEM115-3XHA). Both tagged forms of TMEM115 were produced at the predicted size.

B. Immunofluorescence images for 3XHA-TMEM115 and TMEM115-3XHA under different permeabilisation treatments. Both termini have cytosolic orientation. Scale bar (at the bottom left of the images): 10 µm

4.6 Construct design for mapping TMEM115 topology

After aligning the predicted models, I designed five tagged constructs, in which epitope tags were inserted in different positions within the predicted loops (arrows in Figure 4.2B). By assessing whether each of these tags was luminal or cytoplasmic, I planned to distinguish among the several models.

Initially, I used a single tag strategy to minimise the possible disruption of protein folding by the additional sequences. I started with inserting a single Myc tag in each of the selected positions. Considering the potential difficulty of detecting a single epitope tag, the constructs were also tagged with triple HA at the N-terminus for an alternative way of detecting the over-expressed tagged proteins. Unexpectedly, different versions of overexpressed tagged proteins were neither detectable with anti-Myc antibody in western blot, nor in immunofluorescence (data not shown). However, with an anti-HA antibody, different versions of over-expressed tagged TMEM115 were visible via their triple HA tags at the N-terminus. There were two possibilities to explain this result: either the anti-Myc antibody was not sensitive enough to detect the single Myc tag inserted in the middle of the loop region, or the tag was masked by protein structure or membrane proximity.

To test the first possibility, I made another set of constructs with a single Flag tag inserted in the same designed positions and triple HA tags added at the N-terminus. These five different constructs were again detectable with HA antibody, but not with anti-Flag antibody. This indicated that the single Myc or Flag inserted in loops of TMEM115 were likely to be masked by protein structure or membrane proximity, which is consistent with the small size of loops predicted by most models.

I decided to progress with a triple tag strategy to solve the tag detection problem. Triple Myc tags were inserted in the same designed positions. However, this carried the potential risk was that the long insertion within the loop region might affect the natural structure of TMEM115. Inserting tags in the middle of a multi-span membrane protein is quite challenging, especially if the protein is

predicted to have small loops. There is a risk that TMD folding and architecture may be disrupted by changes in protein sequence. Misfolded protein is unlikely to pass the protein quality control system in the ER; it will either accumulate within the ER or be degraded by the proteasome by ERAD. Therefore, it was necessary to check if different versions of 3xHA-TMEM115 with triple Myc tags were still able to fold properly. The conventional methods to examine the correct folding of a protein are either through some functional output or expression level and cellular localisation patterns. Since no functional assay for TMEM115 has yet been established, I examined the expression level and localisation of the different tagged versions of TMEM115. I performed a transient transfection and checked the expression level of multi-Myc tagged TMEM115 in different positions by western blot. 3xHA-TMEM115 was used as a control of protein expression level since it was the original template construct for all the mutagenesis. In Figure 4.5A, different versions of tagged over-expressed TMEM115 can be detected with both anti-Myc and anti-HA antibodies. As expected, all the multi-Myc tagged proteins in position 1-6 were slightly larger than the 3xHA-TMEM115 (39 kDa); 42 kDa for position 1-5 and 41kDa for position 6 (Figure 4.5A).

I also used immunofluorescence to test if these versions of triple-Myc tagged TMEM115 (Figure 4.5B) were localised within the Golgi. TGN46 was used as a trans-Golgi marker. All versions of Myc-tagged 3xHA TMEM115 were shown to co-localise with TGN46, which proves that they predominantly localised within the Golgi. The fact that these proteins passed ER protein quality control and were trafficked to the Golgi indicates that the inserted triple Myc tags at different loop regions do not affect the correct folding of TMEM115.

In summary, I discovered that multiple copies of the Myc tag worked well in this strategy. The triple Myc tags inserted in the loops give detectable signals in both western blot and immunofluorescence, indicating that these long insertions do not substantially affect the folding and stability of TMEM115.

4.7 Examining the orientation of inserted tags

I performed the final experiments to explore the orientation of inserted tags with HeLa cells under transient transfection conditions. As mentioned above, all constructs for this TMEM115 topology study have a dual tagging system, triple HA at the N-terminus and multiple Myc tags inserted in one of the predicted loops. The triple HA tag at the cytosolic N-terminus of TMEM115 served as a marker for overexpressed protein and as a cytosolic control, while TGN46 was used as a Golgi marker and a luminal control.

In the case of triple-Myc tags inserted in position 1, 3 and 5 (Figure 4.5B), TGN46 was detectable under 0.2% TX-100 treatment but not with 25 μg/ml digitonin, which showed that the Golgi membrane was intact under digitonin but fully permeabilised with TX-100. In the same cell line, Myc tags at position 1, 3 and five can be detected using the same laser power under both treatments; it indicates that these positions are cytosolic.

In the case of Myc tags inserted at position 2 and 4 (Figure 4.5B), they can be observed with TX-100 permeabilisation, but not with digitonin treatment. However, the overexpressed protein can be detected with anti-HA antibody in the samples treated with digitonin, indicating that these positions are luminal.

Once I worked out the orientation of each single position, I started to map the final topology of TMEM115. Given position 1 (Asp74) is cytosolic and position 2 (Phe125) is luminal, there must be an odd number of TMDs between the two positions. Since the length of a single TMD in the Golgi is shorter than that in the ER or plasma membrane, at around 18aa (18.2 +/- 2.9; mean +/- s.d.) (Quiroga et al. 2013), we can rule out the possibility that there are three TMDs between position 1 and 2. I applied the same analysis to the rest of the positions, which are position 3 (Leu159, cytosolic), position 4 (Ser189, luminal) and position 5 (Arg215, cytosolic). In total, I was able to map the location of four TMDs precisely: those between position 1 and 2, 2 and 3, 3 and 4, plus a final TMD between 4 and 5 (Figure 4.6).

To determine whether there were TMDs between the N-terminus of TMEM115 and position 1, or between position 5 and C-terminus of TMEM115, was more complicated. Based on the topology prediction results, all the algorithms I have employed have highlighted the N-terminal two-thirds of the protein to be where the TMDs are located. Therefore, I chose a sixth position, Cys44, and inserted a double Myc tag at it. After applying the same methodology, I was able to determine that position 6 is luminal (Figure 4.5B). Given the N-terminus of TMEM115 is cytosolic, there is a TMD between the N-terminus and position 6; while the cytosolic location of position 1 has shown that the second TMD is between position 6 and position 1.

Figure 4.5A

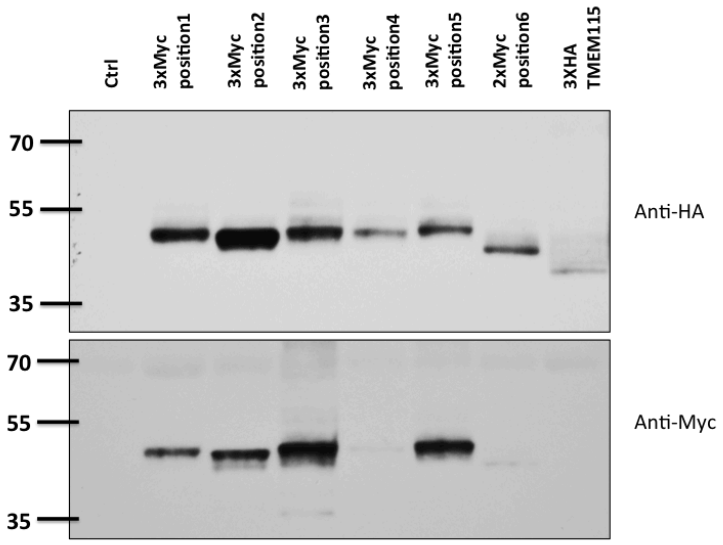


Figure 4.5B

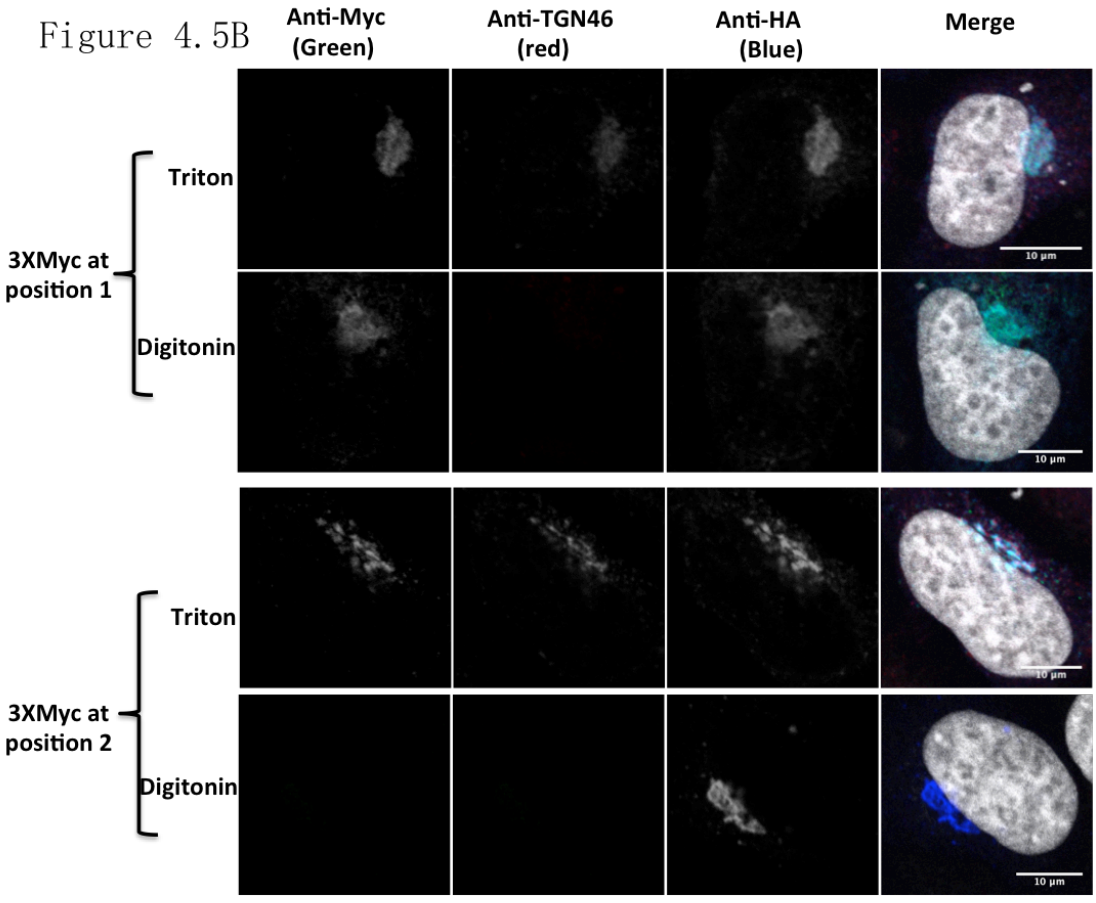


Figure 4.5B

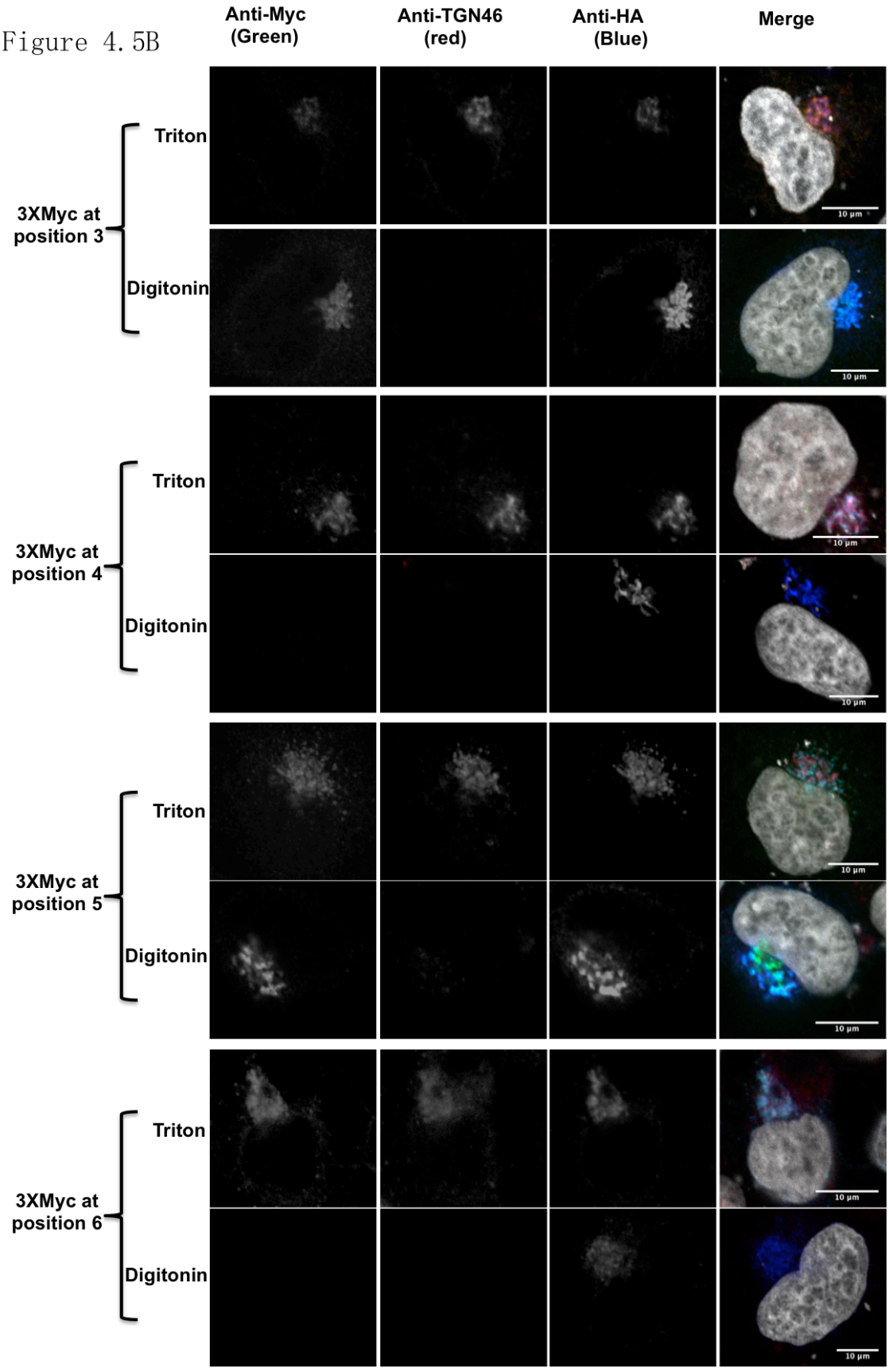

Anti-Myc (Green)
Anti-TGN46 (red)
Anti-HA (Blue)
Merge
3XMyc at position 3
Triton
Digitonin
3XMyc at position 4
Triton
Digitonin
3XMyc at position 5
Triton
Digitonin
3XMyc at position 6
Triton
Digitonin
10 µm

Figure 4.5 Examine the orientation of each inserted tags in six selected positions
A. Western blot for lysates of HeLa cells that were transfected with six multi-Myc tagged forms of 3XHA-TMEM115. Cell lysates were immunoblotted for Myc (for detecting multi-Myc tags in the loop region) and HA (3XHA tag at the N-terminus). All tagged forms of TMEM115 were produced at the predicted size.
B. Immunofluorescence images for six multi-Myc tagged forms of 3XHA-TMEM115 under different permeabilisation treatments. TGN46, the Golgi luminal control, was only detectable under 0.2% TX-100 permeabilisation. The cytosolic HA tag at the N-terminus of TMEM115 was detectable under both permeabilisation treatments. All the multi-Myc tags inserted in the predicted loop regions were detectable under 0.2% TX-100 permeabilisation. Only Myc tags at position 1,3 and 5 were detectable under digitonin treatment, indicating the cytosolic location of these three positions. Scale bar (at the bottom left of the images): 10 μm

A model of TMEM115 topology is presented in Figure 4.6B. In summary, I have shown that TMEM115 has six TMDs and five relatively small loops (90 amino acids in total), especially the three loops that face the lumen. This explains why it was a challenge to tag them and detect the tagged protein in the first place. The six TMDs are within the N-terminal two-thirds of TMEM115, which is the more conserved region of the protein. Most of the C-terminus of the protein, more than a third (~150 amino acids), is cytosolic. The C-terminus of TMEM115 is considerably longer than the N-terminus (which is ~20 amino acids). The fact that TMEM115 has six TMDs, and that Phyre 3D structure prediction has found that GlpG, the only Rhomboid with a known crystal structure, has the highest structural similarity with TMEM115, together provide strong evidence that TMEM115 is indeed a member of the rhomboid-like superfamily. Therefore, the six TMDs that have been identified in this study are likely to be the rhomboid-like domain, which has been described in the introduction to be the structural feature that unifies the whole rhomboid-like superfamily.

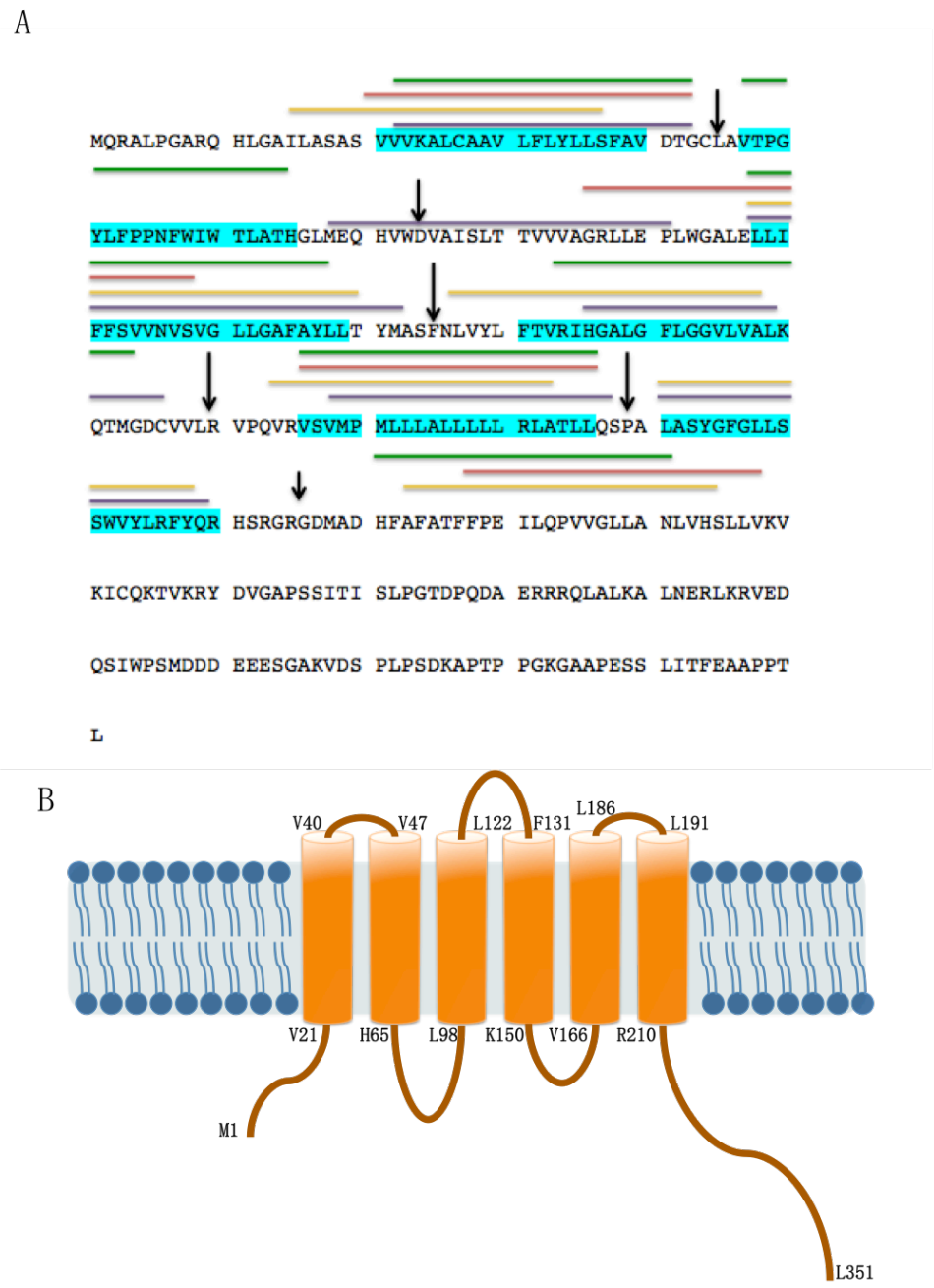


Figure 4.6 Final TMEM115 topology model

A. Alignment of the four predicted topology models from the different algorithms. The coloured lines indicate the predicted TMD regions of each model; the arrows indicate the selected sites for inserting tags. Areas labelled with blue background are where the TMD regions are likely to locate after topology mapping.

B. Topology model for TMEM115. First and last residues of the protein are indicated, as well as the predicted first and last amino acid of each of the six TMDs.

4.9 TMEM115 C-terminal truncations and their cellular localisations

After mapping the topology of TMEM115, I made two C-terminally truncated versions of TMEM115. In the first truncation, TMEM115 (TM1-205), I deleted the whole C-terminus and possibly part of the sixth TMD, starting from R206. For the second truncation, TMEM115 (TM1-229), I retained all six TMDs and the reported Golgi targeting region (Ong et al. 2014), R206-P229, distal to the final TMD. This deletion started at A230. I transiently transfected these constructs to check the production levels of the two truncated forms of TMEM115. Both truncated TMEM115 proteins were expressed at a normal level and at their predicted sizes (around 20KDa and 23 kDa) (Figure 4.7B).

To assess whether these truncations affected cellular localisation, I checked them with immunofluorescence. GM130, a cis-Golgi protein, was used as a Golgi marker in Figure 4.7C. 3xHA-TMEM115 was transiently transfected into HeLa cells to show the cellular location of full-length TMEM115, which co-localised with GM130 within the Golgi. TM1-205 did not co-localise with GM130. Instead, the cellular distribution was reminiscent of an ER-like pattern (Figure 4.7C) no ER marker was used in this set of imaging). In the case of TM1-229, which contained the published Golgi targeting sequence region, it was localised within the Golgi, as expected (Figure 4.7C).

What was slightly concerning at this stage was the ER-like distribution pattern of TM1-205. Whether it was due to the absence of the Golgi targeting region (206-229), or because the truncated protein was not folded properly, so it was not able to pass the ER protein quality control, was unknown. However, the production level of TM1-205 was comparable with the full-length protein. Since no functional assay was established to test TMEM115 functions, I decided to carry on with the interaction study with these two truncation mutants. Based on the principle that if TM1-205 could still interact with some of TMEM115's known interactors (presumably not all of them, due to the different interaction regions), it should indicate that it is folded correctly.

In summary, TM1-205, which lacks the reported Golgi-localisation signal, has an ER-like distribution pattern, while TM1-229 is localised within the Golgi apparatus. Both of the truncations were produced at a comparable level with the full-length TMEM115.

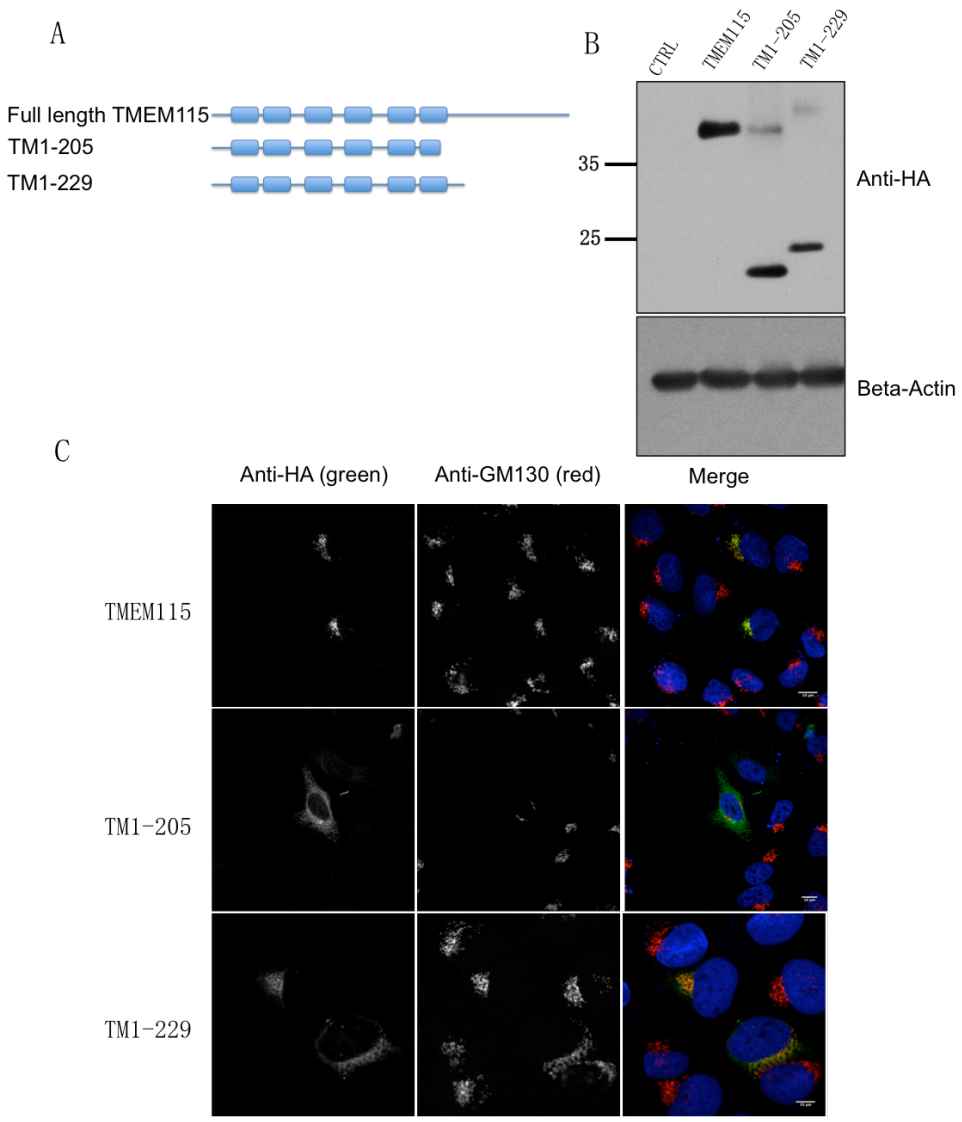


Figure 4.7: The expression pattern of Truncated form of TMEM115

A. A diagram of different truncated TMEM115

B. Western blot of full-length TMEM115, TM1-205 and TM1-229.

C. Immunofluorescences images of full-length TMEM115, TM1-205 and TM1-229. GM130 is a cis-Golgi marker. Scale bar (at the bottom left of the images): 10 μm

4.10 The interaction pattern between TMEM115 truncations and known TMEM115 interactors

As reported in Chapter 3, I identified a number of convincing TMEM115 interacting proteins, including both TMD containing proteins and soluble proteins. By removing the cytosolic C-terminus tail, which is the major cytosolic region in TMEM115, and therefore likely to be the main site for cytosolic interactions, I would be able to distinguish whether the interaction between TMEM115 with a known interactor depends on the TMD region or the cytosolic region.

I started by validating the interaction between TMD containing interactors and different versions of TMEM115. 3xHA-TMEM115, 3xHA-TM1-205 and 3xHA-TM1-229 were co-transiently transfected with Flag-ERGIC-53 into HEK293T cells. After enriching different versions of TMEM115 with HA beads, I probed for overexpressed ERGIC-53 in the material eluted from the HA beads (Figure 4.8A). Interestingly, the interaction between ERGIC-53 and TM1-205 was very strongly increased compared with full-length TMEM115 and noticeably increased with TMEM1-229. The fact that TM1-205 was still able to interact with known interactors of TMEM115 suggests that it is probably correctly folded despite being at least partially mislocalised in the ER. I repeated the same co-immunoprecipitation experiment with overexpressed TM1-205, TM1-229 and full-length TMEM115, and detected endogenous SCAP in the material eluted from the HA beads. Again interactions were detected at a higher level with the truncation mutants than with the full-length TMEM115 sample. This suggested that TMEM115 interacts with both SCAP and ERGIC-53 at least primarily through its transmembrane domain: the long C-terminal tail is unnecessary for the interactions. These changes of interaction pattern between both SCAP and ERGIC-53 with TMEM1-205 may be explained by the mislocalisation of TM1-205 in the ER. SCAP is predominately localised within the ER under normal conditions where lipids are not in short supply, and only a tiny proportion of SCAP travels to the Golgi and recycled back to ER. In the case of ERGIC-53, it is a lectin type membrane protein that is continuously recycled between the ER, the ER-Golgi

intermediate compartment (ERGIC), and the cis-Golgi. Having the truncated form of TMEM115, TM1-205, localised within the ER could facilitate the binding between TMEM115 with SCAP or ERGIC-53. The hypothesis could explain the increased binding between TM1-205 with the two tested interactors, but it was not sufficient to account for the fact that the correctly located truncation, TM1-229, which is predominately localised within the Golgi, also showed increased binding to SCAP and ERGIC-53. Another, not mutually incompatible, the possibility is that the C-terminus of TMEM115 could act as a regulatory domain for the interactions, possibly inhibiting the interaction.

A negative control, VANG1 (Vang-like protein VANG1), which contains 4 TMDs and involved in cell polarity regulation (Sawa 2012), was employed as a negative control for the co-immunoprecipitation. No interaction with TMEM115 was expected, and this was used to confirm that the increased interaction between truncated TMEM115 with SCAP and ERGIC-53 was not caused by the unspecific stickiness of transmembrane domain interactions. As shown in Figure 4.8A, VANG1 did not interact with any versions of TMEM115.

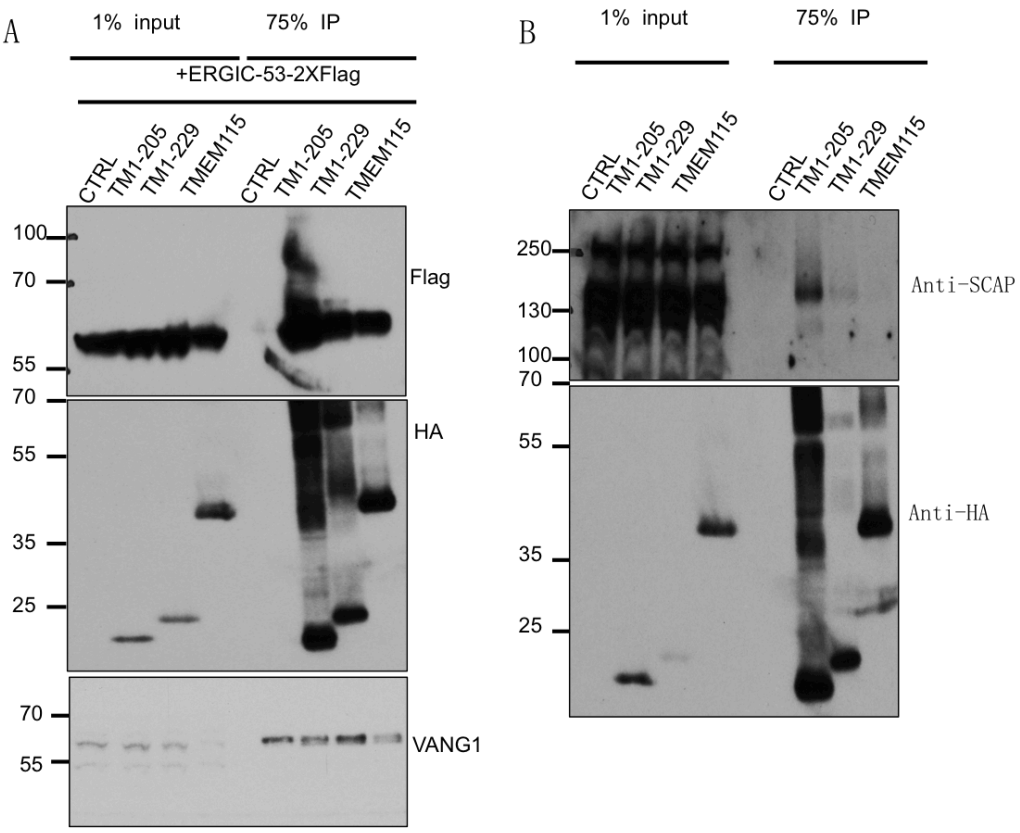


Figure 4.8: The interaction between truncated forms of TMEM115 and membrane protein interactions

A . Co-immunoprecipitation between overexpressed different truncations of TMEM115 and ERGIC-53. 24 hours before HA immunoprecipitation, HEK293T cells were co-transfected with ERGIC-53-2XFlag and empty vector, 3XHA-TM1-205,3XHA-TM1-229 or 3XHA-TMEM115, respectively. Cell lysates and HA beads eluate were immunoblotted for Flag (for ERGIC-53-2XFlag), HA (for different forms of 3XHA tagged TMEM115) and VANG1. Elevated levels of interaction between ERGIC-53 and truncated TMEM115 were observed.

B. Co-immunoprecipitation between overexpressed different truncations of TMEM115 and endogenous SCAP. 24 hours before HA inmunoprecipitation, HEK293T cells were transfected with empty vector, 3XHA-TM1-205, 3XHA-TM1-229 or 3XHA-TMEM115. Cell lysates and HA beads eluate were immunoblotted for SCAP and HA (for different forms of 3XHA tagged TMEM115). Elevated levels of interaction between endogenous SCAP and truncated TMEM115 were observed (hardly any endogenous SCAP was detected in the beads eluate of full length TMEM115 sample in this blot, I will show the interaction in figure 5.2).

I tested the interactions between several cytosolic interactors of TMEM115 with the truncations described above. Different versions of HA-tagged TMEM115 were overexpressed in HEK293t cells. Endogenous beta-COP, VCP/p97 and p62 were detected in the eluted materials from HA beads. Surprisingly, again the interactions were strongly enhanced in C-terminally truncated TMEM115 for all the soluble interactors (Figure 4.9). This phenomenon is difficult to explain, as the cytosolic C-terminus is more than one-third of the protein, and is predicted to be the primary region for interacting with cytosolic interactors. However, my data consistently and reproducibly contradicted this expectation. It was hard to explain why the truncated versions of TMEM115, which contain mostly the TMD core and a tiny N-terminal cytosolic region, could bind more strongly than full length to cytosolic proteins. There are several possible explanations. Perhaps the interactions are indirect, depending on other membrane proteins to make a link with TMEM115. Upon the removal of the C-terminus of TMEM115, the interactions between truncated TMEM115 and its transmembrane interactors, or possibly through endogenous TMEM115, could be increased, thereby potentially enriching any cytosolic interactors. To investigate this possibility, I examined TMEM115 self-interaction and whether TMEM115 truncations were still able to bind to endogenous TMEM115 in Section 4.10.

Alternatively, perhaps some of the small remaining cytosolic regions in TM1-205 and TM1-229, which I dismissed originally, including the small cytoplasmic loops or N-terminus, could provide binding sites for those soluble proteins. New versions of truncated TMEM115 are required to test this hypothesis. For instance, to make a new truncation with both termini deleted and test if the binding is affected with soluble interactors.

In summary, the truncations of TMEM115 provided interesting tools for understanding the interactions between TMEM115 with its known interaction partners. The results have suggested that upon the deletion of the cytosolic C-terminus, interactions are increased between TMEM115 and all tested partners,

irrespective of their topological localisation. However, more experiments are required to provide a reasonable explanation of the altered interaction pattern.

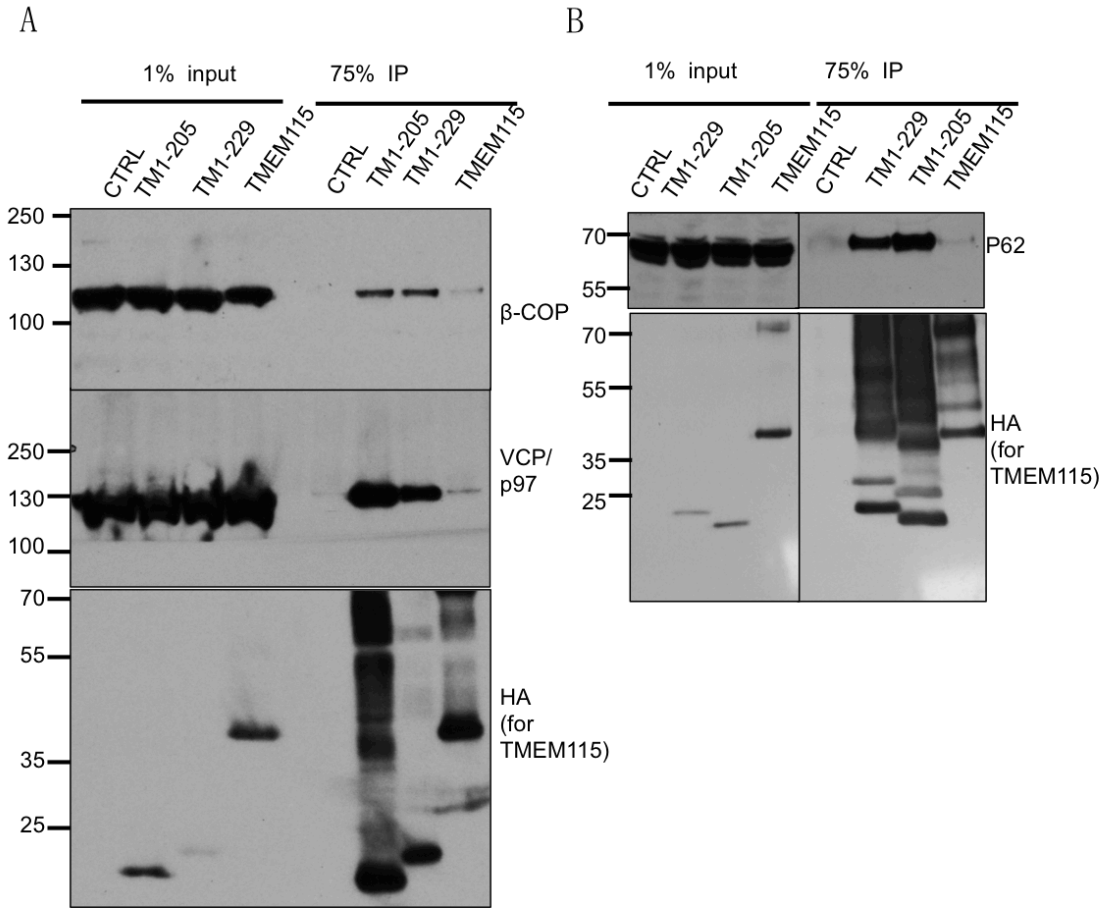


Figure 4.9: The interaction between truncated forms of TMEM115 and cytosolic interactors

24 hours before HA immunoprecipitation, HEK293T cells were transfected with empty vector, 3XHA-TM1-205, 3XHA-TM1-229 or 3XHA-TMEM115.

A. Co-immunoprecipitation between overexpressed different truncations of TMEM115 and endogenous β-COP and VCP/p97. Cell lysates and HA beads eluate were immunoblotted for β-COP, VCP/p97 and HA (for different versions of 3XHA tagged TMEM115). Elevated levels of interaction between endogenousβ-COP, VCP/p97 and truncated TMEM115 were observed.

B. Co-immunoprecipitation between overexpressed different truncations of TMEM115 and endogenous p62. Cell lysates and HA beads eluate were immunoblotted for p62 and HA (for different versions of 3XHA tagged TMEM115). Elevated levels of interaction between endogenous p62 and truncated TMEM115 was observed

4.11 TMEM115 interacts with itself through the TMD region

One possibility is that when expressed in cells, truncated TMEM115 might interact with endogenous full length TMEM115, and this might in some way underlie the unexpected increase in binding of interacting proteins when overexpressing truncated versions of TMEM115. I, therefore, addressed the question of whether TMEM115 interacts with itself and, if so, which region of the protein is responsible for such homodimerisation. 3xHA-TMEM115 and Flag-TMEM115 were co-transfected into HEK293t cells. After immunoprecipitating 3xHA-TMEM115 with HA beads, significant levels of Flag-TMEM115 were detected in the eluate indicating that TMEM115 does indeed interact with itself (Figure 4.10A).

In the next step, I investigated which domains of TMEM115 were responsible for this self-interaction. The anti-TMEM115 antibody I employed in this study only recognises part of the C-terminus of the protein, which was deleted in both of the truncations. Endogenous TMEM115 was detected in the eluted materials from HA beads of both truncated forms of TMEM115, which suggested that TMEM115 is self-interacting through its transmembrane domain region, not the long cytoplasmic C-terminus (Figure 4.10B).

To address the question whether endogenous TMEM115 contributes to the binding between truncated TMEM115 and its cytosolic interactors, I performed immunoprecipitation experiments in both wild type and *Tmem115* knockout MEFs. Full-length TMEM115 and different truncations with triple HA tags were transfected into MEFs. Unfortunately, however, the extremely low transfection efficiency prevented effective immunoprecipitation of overexpressed TMEM115. None of the tested interactors could be detected in the HA beads elution at an endogenous level (data not shown). To address this question in the future, alternative transfection methods, or co-transfecting both truncated forms of TMEM115 and its cytosolic interactors, will be required.

In summary, my data indicate that TMEM115 self-associates, probably through its TMD region. Whether this self-interaction contributes to the increased binding between truncated TMEM115 and its cytosolic interaction partners remains unclear.

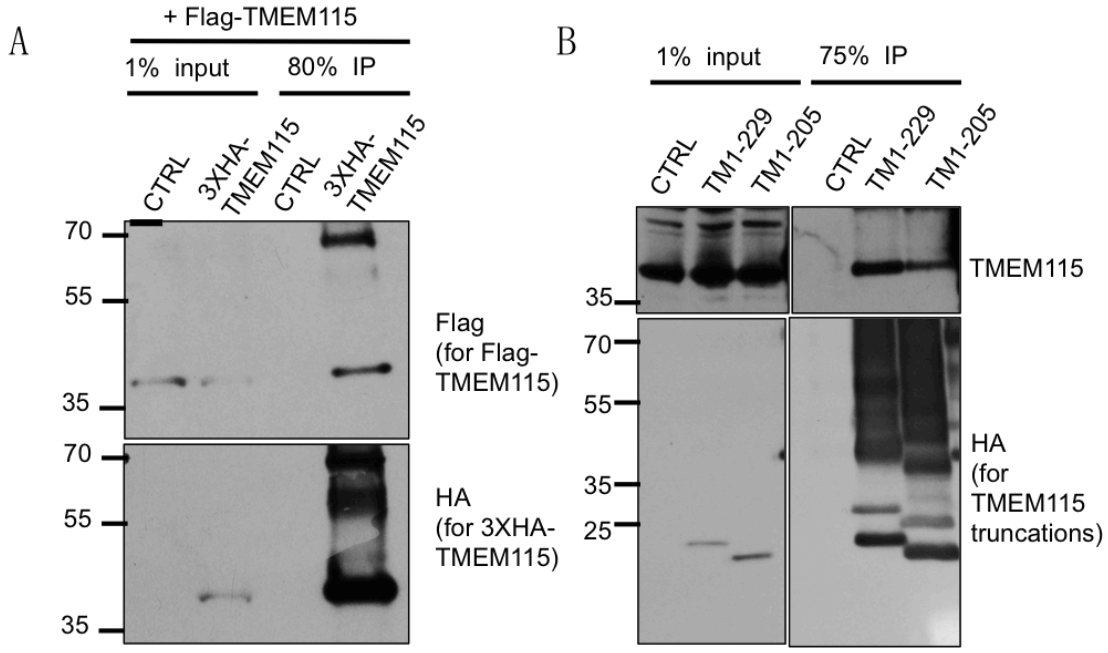


Figure 4.10: TMEM115 interacts with itself possibly through TMD region
A. TMEM115 interacts with itself. 24 hours before HA immunoprecipitation, HEK293T cells were co-transfected with Flag-TMEM115 either with empty vector or with 3XHA-TMEM115. Cell lysates and HA beads eluate were immunoblotted for Flag (Flag-TMEM115) and HA (3XHA-TMEM115). Flag-TMEM115 is only present in HA beads eluate of 3XHA-TMEM115 sample.
B. Co-immunoprecipitation between overexpressed different truncations of TMEM115 and endogenous TMEM115. 24 hours before HA immunoprecipitation, HEK293T cells were transfected with empty vector, TM1-229 or TMEM1-205. Cell lysates and HA beads eluate were immunoblotted for TMEM115 and HA (for TMEM115 truncations). Endogenous TMEM115 were detected in HA beads eluate of both TM1-229 and TMEM1-205 samples.

4.12 Discussion

4.12.1 TMEM115 is a distant member of the rhomboid-like superfamily with six TMDs

TMEM115 is one of the least studied rhomboid-like proteins. The only published paper about TMEM115 topology supports a four-pass TMD model, which is incompatible with the inclusion of TMEM115 within the rhomboid-like superfamily, given that all the known members have the rhomboid-like domain, which contains six TMDs. In this chapter, I have shown that TMEM115 does indeed have six TMDs, using a combination of bioinformatic and experimental examination. More evidence for the structural similarity between the TMDs of TMEM115 and other rhomboid-like proteins was identified with HHPred, a server for searching structural homologs (Soding et al. 2005). Using the sequence of predicted TMEM115 TMD regions to search for proteins that contain similar structural features, rhomboid-like proteins were identified as top candidates. The fact that TMEM115 does have six TMDs and the high structural similarity of the N-terminal half with the transmembrane domains of the *E. coli* rhomboid protease, GlpG, identifies TMEM115 as a *bona fide* member of the rhomboid-like superfamily.

Despite its clear membership of the superfamily, TMEM115 is evolutionarily very distant from the rhomboid protease/iRhom branch (Figure 1.1). PFAM, the protein family database, has not included TMEM115 as a member of the rhomboid-like superfamily (Finn et al. 2016). Some structural features that are described to be shared among most rhomboid-like proteins are not conserved in TMEM115. For instance, most of the active rhomboid proteases and some of the pseudoproteases, including Derlins and UBAC2, have a conserved WR motif (Q/ExWRxxS/T) in their first luminal loop (Lemberg 2013). Another commonly shared feature within the superfamily is the TM helix dimerisation motif (GxxxG) localised in the sixth TMD of the rhomboid-like domain, which enables the tight packing of the fourth and sixth TMD for aligning the active site Ser-His dyad linked (Wang et al. 2006, Lemberg 2013). All the active rhomboid proteases and many of the nonprotease branches, such as iRhoms, Derlins and UBAC2, have this

conserved TM helix dimerisation motif (Bergbold et al. 2013). TMEM115 has no identifiable WR motif, while a GxxxG motif is located at the fourth TMD (GFLGG, residue 140-143), according to the final topology model. This further highlights the distant relationship of TMEM115 with the better-studied members of the superfamily. Similar to TMEM115, the other two distant members of the rhomboid-like superfamily, RHBDD2 does not have WR motif, while RHBDD3 has neither the TM helix dimerisation motif (GxxxG) nor WR motif.

4.12.2 Methods for mapping the topology of transmembrane proteins and their limitations

Mapping the topology of membrane proteins provides preliminary information for further structure-function studies. Despite the conceptual simplicity of a mapping strategy, such a topology study of a multi-span transmembrane protein is never an easy journey. There are not many papers in which the methodologies for mapping membrane protein topology have been discussed in detail, and most of them are about single-pass membrane proteins. Several methods have been described for this type of study: bioinformatic analysis, inserting tags or glycosylation acceptor sequences in the predicted loop regions, and protein 3D structural analysis (Wang et al. 2006, Wu et al. 2006, Greenblatt et al. 2011, Carrara et al. 2012). Depending on the protein of interest, combinations of several methods have been applied to examine the topology.

The topology of Derlin-1, a rhomboid-like protein involved in ERAD, was addressed elegantly with a combination of bioinformatics tools and glycosylation acceptor sequences inserted within the middle of predicted loops (Greenblatt et al. 2011) Compared with TMEM115, which is very distant to GlpG and other rhomboid proteases, Derlin-1 is more closely related by sequence and secondary structure, and the fact that it has the WR motif and the TM helix dimerisation motif facilitates the prediction of all the loop regions. Since N-glycosylation acceptor sequences can only be recognised in the luminal side of a protein (Aebi 2013), by inserting N-glycosylation acceptor sequences in different locations across the whole protein and examining whether the tagged protein is glycosylated with

endoglycosidase H (an enzyme that can remove N-link glycosylation), all the luminal sites were therefore distinguishable from the cytosolic locations (Greenblatt et al. 2011). I could also have used this method for mapping TMEM115 topology. Given that the N-glycosylation acceptor sequence insertion (Asn-X-Ser/Thr) is quite short, it minimises the possible disruption to the folding of the protein. However, based on our current knowledge about TMEM115 topology, all the five loops are relatively small, especially the three luminal loops. Therefore, there are potential false negative risks, as the luminal glycosylation insertion could be masked by secondary structure in proximity to the membrane.

Solving the 3D structure of a membrane protein may be the most challenging way of determining its topology. Still, it provides much more structural information than other methods. For the structural study of rhomboid-like proteins, although scientists worldwide have made some limited in-roads, only the structural of GlpG in both *E. coli* and *Haemophilus* influenza have been revealed by several individual groups so far (Wang et al. 2006, Wu et al. 2006, Ben-Shem et al. 2007, Lemieux et al. 2007). This is because it is challenging to crystallise membrane proteins, especially multi-pass transmembrane proteins. The emerging structural technique of cryo-EM has sped up the solution of membrane protein structures substantially. A significant advantage of protein structural studies with cryo-EM is that crystallising a protein is no longer a necessary step. In addition, it is possible to reveal the structure of a protein in a more biologically relevant environment. Might it be possible to study TMEM115 structure using cryo-EM? The current limitation is that cryo-EM can provide structural information for proteins or complexes above a certain molecular weight (about 145 kDa) (Merk et al. 2016). Full-length TMEM115 is only 37 kDa, which is far below the limit. One possible solution would be to study TMEM115 structure in a complex with its interactors. I will discuss this possibility in the final perspective chapter of this thesis.

4.12.3 Binding regions between TMEM115 and its interactors

The C-terminus of TMEM115, comprising more than 140 amino acids, represents the majority of the TMEM115 cytosolic region. By prediction, it is likely to serve as a binding platform for cytosolic interacting partners of TMEM115. Strikingly, however, instead of abolishing the interaction between TMEM115 with some of its known binding partners, deletion of the C-terminus enhanced interactions with all the binding partners I tested, irrespective of their topologies.

Regarding the two membrane domain containing interactors, ERGIC-53 (a type-II single-pass transmembrane protein) and SCAP (a polytopic transmembrane protein with eight TMDs), deletion of TMEM115 C-terminus enhanced the two interactions. This may be due to enhanced TMD-TMD interactions, or the removal of an inhibitory region in the C-terminus. Several conserved motifs within the TMD regions are shown to drive TMD-TMD interaction; the most common one is TM helix dimerization motif (GxxxG/small amino acid xxx small amino acid) (Russ et al. 2000, Teese et al. 2015). Structural studies suggest that small amino acid xxx small amino acid motifs (typically GxxxG) maximise the van der Waals interactions and/or hydrogen bond formation by allowing the two interacting helices to be in close proximity to each other (MacKenzie et al. 1997, Smith et al. 2001). In addition, motifs with polar residues that contribute to the formation of hydrogen bonds (e.g. polar-xx-polar) (Sal-Man et al. 2005), are also frequently identified in TMD-TMD interactions. A GxxxG motif is present in the fourth TMD of TMEM115, as well as the predicted eighth TMD of SCAP, which further supports the possibility that TMD-TMD interactions promote the association of TMEM115 with SCAP. Neither TM helix dimerisation motifs nor polar residue motifs are identified in the predicted TMD region in ERGIC-53. However, the topology of ERGIC-53 suggests the interaction between ERGIC-53 and TMEM115 is likely through TMD regions, as ERGIC-53 has a large luminal domain (more than 400 amino acid), a transmembrane domain and a short cytosolic tail (12 amino acids) (Hauri et al. 2000). While the mapped TMEM115 topology shows

that it has three relatively small luminal loops. Therefore, there are few possibilities that the interaction occurs via cytosolic or luminal regions.

Nevertheless, more evidence is required for the existence of TMD-TMD based interactions. The presence of a GxxxG motif in both TMEM115 and SCAP may be an indicator of TMD-TMD interaction. Firstly, to test if the TMD region of TMEM115 is sufficient for the binding between SCAP and ERGIC-53, a new truncated form of TMEM115, which contains only the cytosolic C-terminus, is required. If the binding with ERGIC-53 and SCAP is abolished upon the deletion of TMEM115 TMD region, it will further support the view at TMEM115 interacts with the tested membrane domain containing proteins through its TMD region. In addition, to test whether the TMD regions of these two interactors are sufficient for the binding, I can replace the TMD region of, for instance, ERGIC-53, with TMD from another type-I single-pass transmembrane protein that does not bind to TMEM115. If the new fusion protein loses its ability to bind to TMEM115, it further supports the existence of a specific TMD-TMD interaction between TMEM115 and tested interactors. The cytosolic C-terminus may be a regulatory domain to control the level of binding between TMEM115 and the tested membrane protein interactors.

The increased binding between truncated TMEM115 and the tested soluble interactors is much more difficult to explain. In the results section, I have proposed two possibilities: first, the increased binding between truncated TMEM115 could cause by indirect binding, either through endogenous TMEM115 or through other membrane protein interacting partners. Second, other cytosolic regions of TMEM115 may be responsible for the binding of its soluble interactors.

Regarding the possibility of indirect binding, I have shown that TMEM115 indeed interacts with itself and the TMD region is responsible for the interaction. It is possible that endogenous TMEM115 contributes to the elevated binding between truncated forms of TMEM115 and all the tested soluble interactors. Co-immunoprecipitation experiments in *TMEM115* knockout cell lines will be useful to test whether the increased binding pattern is caused by the presence of

endogenous TMEM115. In addition, co-immunoprecipitation with a new truncated form of TMEM115 cytosolic C-terminus would be useful to examine in the absence of TMD region of TMEM115; the cytosolic C-terminus tail can bind to the tested soluble interactors. If the C-terminus of TMEM115 is not able to interact with the tested soluble interactor, while the deletion of endogenous TMEM115 does not affect the elevated binding between the two TMEM115 truncations and soluble interactors, the increased binding may reflect indirect binding caused by other membrane domain containing proteins.

Given the evidence I have now, it is still too early to draw conclusions about the contributions made by the TMD core and C-terminus of TMEM115 to its known interactions. Further investigation of interaction domains would be useful to reveal the exact mechanism of how TMEM115 regulates its interacting proteins.

Chapter 5 Exploring the relationship between TMEM115 and SCAP-SREBP pathway

5.1 Introduction

TMEM115 is widely conserved in eukaryotes and ubiquitously expressed, which suggests that it may have important roles. Data generated by a postdoc in Freeman lab, Dr Angela Moncada Pazos, has shown that the deletion of TMEM115 in both mice and *Drosophila* leads to severe phenotypes. As has been described in detail in the introduction, both phenotypes are linked with defective lipid homeostasis.

To summarise the results of these experiments, which were not carried out by me, *Tmem115* knockout MEFs were found to have a profound defect in their ability to generate lipid droplets when the cells were induced to differentiate into adipocyte-like cells. Given that the SREBP family of transcription factors are the primary regulators in vertebrates for controlling lipid homeostasis (Horton et al. 2002), the SCAP (SREBP cleavage-activating protein)-SREBP pathway was therefore closely examined in *Tmem115* knockout MEFs and mouse tissues. In *Tmem115* mutant adipocyte-like cells, the expression levels of four SREBP target genes (ACACA, FASN, HMGCS and LDLR) were strikingly low upon lipid depletion, compared with wild type cells. In livers of wild type and *Tmem115* knockout mice, the expression of FASN, HMGCS and LDLR were also significantly lower in *Tmem115* knockout mice compared to wild type mice, confirming the physiological relevance of the results in cells.

The interaction between TMEM115 and endogenous SCAP was validated with co-immunoprecipitation. In addition, the localisation of SCAP in wild type and *Tmem115* knockout MEFs were closely examined. Whereas SCAP was primarily localised in the ER of wild type cells grown in lipid-rich medium, it was much more prominently located in the Golgi apparatus in *Tmem115* knockout cells. This suggests that TMEM115 may have a function in regulating the localisation of SCAP and thereby its control of SREBP activation.

The SCAP-SREBP pathway is crucial for sterol homeostasis in vertebrates (diagram in Figure 5.1). There are two SREBPs in vertebrates; between them, they control the activation of different downstream lipid synthesis genes. Under normal conditions, where cytoplasmic sterols are relatively high, the SCAP-SREBP complex is anchored in the ER by the protein INSIG, and lipid synthetic genes are inactive. SCAP is a multi-pass transmembrane protein and has eight TMDs, a sterol-sensing domain (SSD) and seven WD domains, which are required for the binding of SREBPs through their regulatory domain. Upon lipid depletion, the SSD of SCAP undergoes a conformational change, separates from INSIG, and is loaded, in complex with bound SREBP, into COPII vesicles that traffic to the Golgi apparatus. Once in the Golgi, SREBP transcription factors are released from their membrane-tethered precursors by site-1 and site-2 protease cleavages. The transcription factors enter the nucleus and activate the expression of genes for lipid synthesis and uptake. This pathway, ultimately controlled by cellular lipid levels, is the principal mechanism of homeostatically regulating *de novo* lipid production in all vertebrate cells (Wang et al. 1994, Nohturfft et al. 1999, Horton et al. 2002).

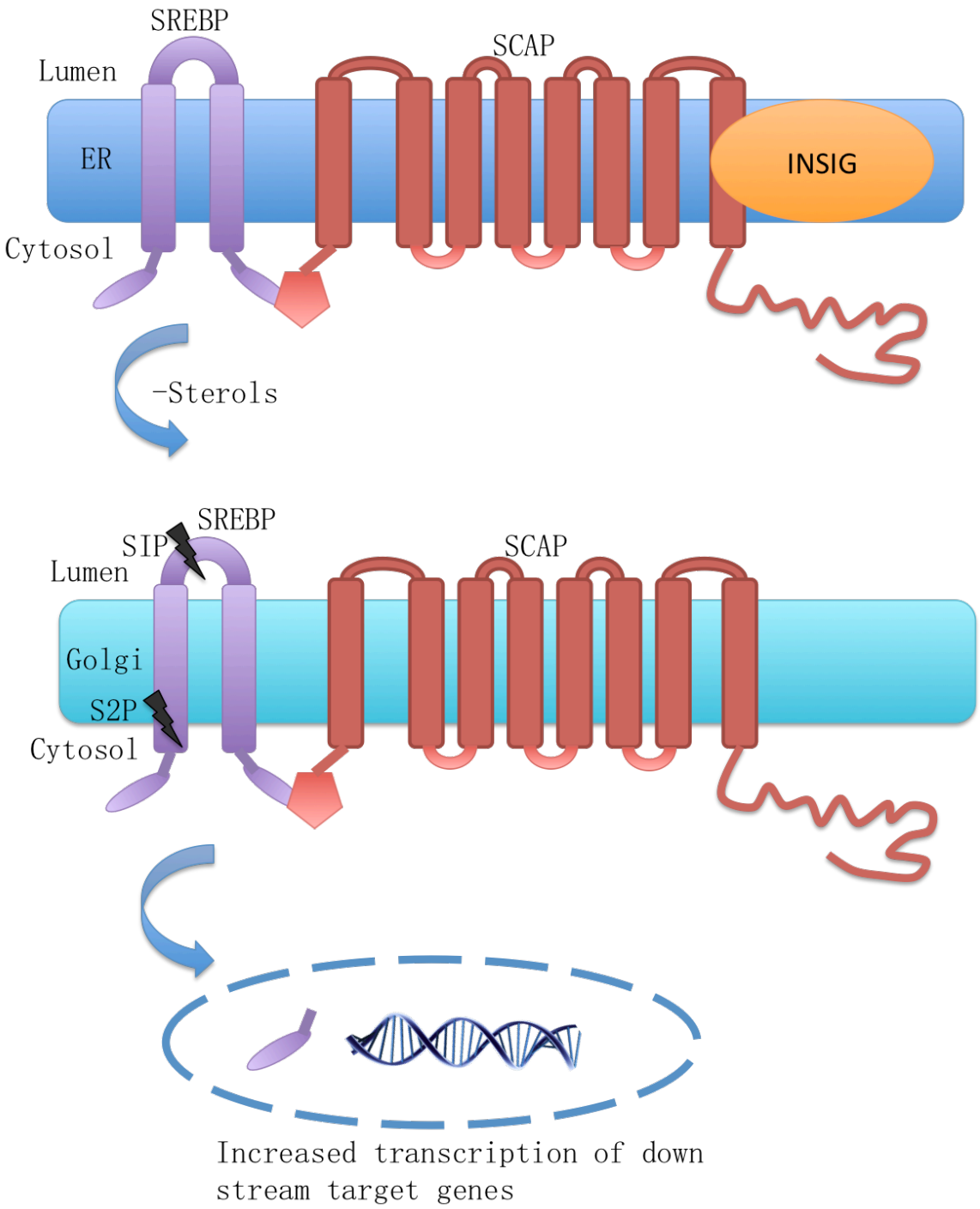


Figure 5.1A A diagram of SCAP-SREBP pathway
SCAP-SREBP complex is anchored in the ER by the protein INSIG. Only upon low sterol stimulation, SCAP goes through a conformation change, and the whole complex is forward trafficked from the ER to the Golgi. SREBPs are cleaved in Golgi by site-1 and site-2 proteases, which release the active transcriptional factors at the N-terminus of SREBPs.

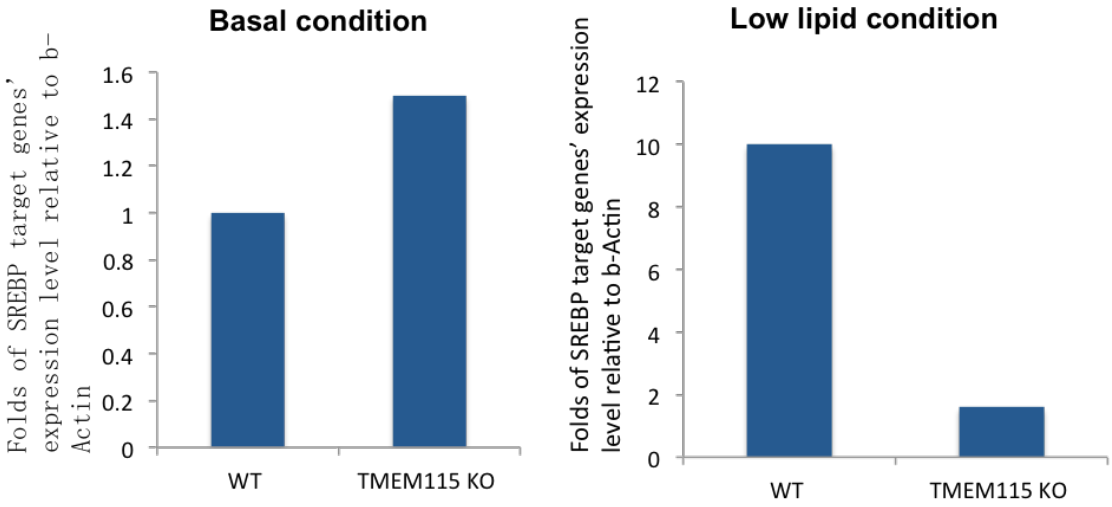


Figure 5.1B Illustrative chat of SREBP target genes' expression level in WT and *Tmem115* KO MEFs under different conditions
The expression levels of SREBP downstream target genes are dysregulated in Tmem115 KO MEFs. In *Tmem115* KO MEFs, the constitutive expression levels of SREBP target genes are slightly higher than in WT MEF. Under low lipid condition, the expression of SREBP targets in *Tmem115* KO MEFs was significantly lower than in wild-type controls.

Given Dr Moncada-Pazos's data that the SCAP-SREBP pathway is dysregulated in the absence of TMEM115 and the potential physiological significance indicated by the animal phenotypes (A illustrative chat of SREBP target genes expression level in WT and Tmem115 KO MEFs is presented in Figure 5.1B), I entered a collaboration to explore the relationship between TMEM115 and the SCAP-SREBP pathway. Since this was a joint project, in the results section, I will refer to some of Dr Moncada-Pazos's data as necessary.

Results

5.3 Changes in TMEM115 protein level affect the level of endogenous SCAP, which may affect levels of downstream SREBP targets

Preliminary data from Dr Moncada-Pazos has shown that TMEM115 physically interacts with SCAP under co-overexpressed condition. I confirmed this interaction by detecting endogenous SCAP in the HA beads eluate of cells expressing

3XHA-TMEM115 (Figure 5.2A). This interaction may underlie the phenotypic consequences of loss of TMEM115 on lipid homeostasis but, if so, the mechanism is unclear. One of the first lines I tested was whether changes of TMEM115 protein level could affect the expression level of full-length SCAP protein. TMEM115 is localised predominantly within the Golgi, while the SCAP-SREBP complex is mostly present in the ER, anchored by the protein Insig1/2 under lipid-rich conditions. Only upon low lipid conditions does SCAP undergo a conformational change, allowing the whole complex to be transported from the ER to the Golgi, where SCAP can encounter TMEM115. I decided to perform the experiments under conditions of low lipid when the SREBP pathway should be induced. Upon 96 hours of knockdown with *TMEM115* siRNA in wild type MEFs, together with low lipid culture conditions in the final 16 hours, the level of endogenous SCAP protein was significantly increased in all the nine replicated experiments (Figure 5.2B, mean=2.215, p-value=0.0078). I also examined the effect on SCAP expression of the overexpression of TMEM115 in HeLa cells, again in low lipid conditions. Overexpression of TMEM115 led to a decrease in endogenous SCAP levels in HeLa cells (Figure 5.2C, mean=0.5577, p-value=0.0865).

Combining these two experiments, my data provide clear evidence that TMEM115 can influence levels of endogenous SCAP protein under low lipid conditions. Loss of TMEM115 leads to increased SCAP levels; TMEM115 overexpression reduces SCAP levels. A possible explanation would be that TMEM115 is involved in destabilisation of SCAP. Consistent with these findings, Dr Moncada Pazos observed the constitutive activation of targets of the SCAP-SREBP pathway in human HeLa cells upon knockdown with TMEM115 siRNA (data not shown).

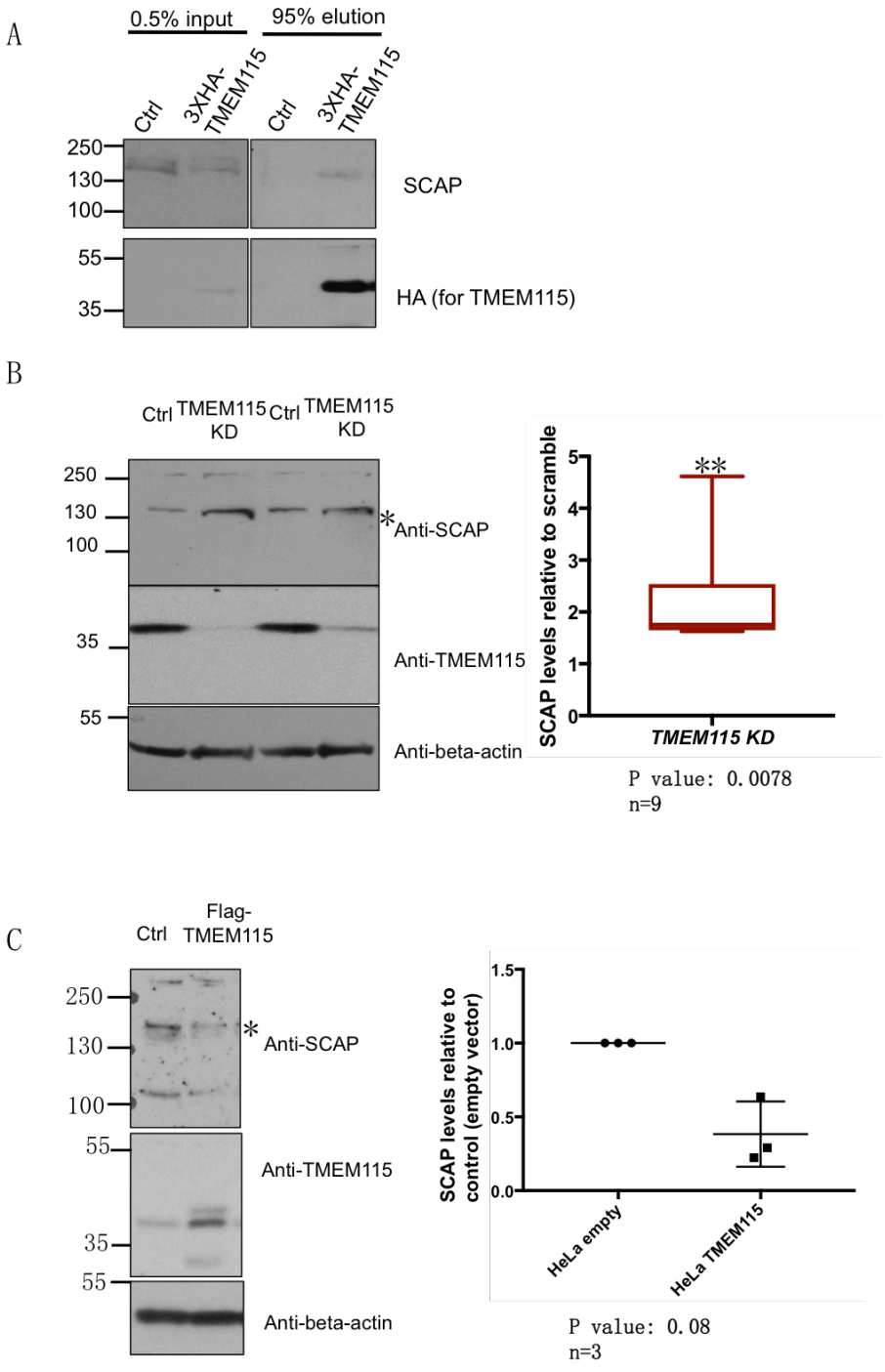


Figure 5.2 TMEM115 interacts with endogenous SCAP and regulates its levels

A. Co-immunoprecipitation between overexpressed TMEM115 and endogenous SCAP. 24 hours before HA immunoprecipitation, HEK293T cells were transfected with either empty vector or 3XHA-TMEM115. Cell lysates and HA beads eluate were immunoblotted for SCAP and HA (3XHA-TMEM115). Endogenous SCAP is only present in HA beads eluate of 3XHA-TMEM115 sample.

B. Knockdown of *Tmem115* increases endogenous SCAP level under low lipid condition. Wild type MEFs were transfected with either scramble siRNA or *Tmem115* siRNA for 96 hours, with low lipid treatment in the final 16 hours. Left panel: representative western blot of the levels of SCAP, TMEM115 and β-actin as endogenous control. Right: quantification of the levels of SCAP under knockdown (KD) of TMEM115 and relative to control cells (n=8). The p-value is the result of testing the experimental values under a Wilcoxon Signed Rank Test for theoretical median=1.
C. Overexpression of TMEM115 decreases endogenous SCAP level under low lipid condition. HeLa cells were transfected with empty vector or Flag-TMEM115 for 24 hours, with low lipid treatment in the final 16 hours. Left panel: representative western blot of the levels of SCAP, TMEM115 and β-actin as endogenous control. Right: quantification of the levels of SCAP under overexpression of TMEM115 and relative to control cells (n=3). The p-value is the result of testing the experimental values under a one sample T-Test.

5.4 TMEM115 and p62 are both physical interactors of SCAP

Changes in TMEM115 levels under low lipid affects levels of SCAP, suggesting that TMEM115 may have a role in regulating the turnover of SCAP. Along the same lines, with live imaging, Dr Moncada-Pazos observed that the deletion of TMEM115 in MEFs delays the SCAP exit from the Golgi. Importantly, when the same approach was performed in the presence of the proteasome inhibitor MG-132, the Golgi pool of SCAP in wild type cells was identical to those in the *Tmem115* knockout cells. The fact that the absence of TMEM115 had a similar effect on SCAP levels to that observed upon inhibition of proteasomal degradation led me to hypothesise that TMEM115 may regulate post-Golgi proteasomal degradation. This could be part of the response needed to reduce lipid synthesis when cells are moved from low to high lipid conditions. Although this is an attractive idea, it raises further questions. There is no well described programme of proteasomal degradation from the Golgi apparatus that would be analogous to the much-studied ER-associated degradation (ERAD). One possible model is that upon switching off the SCAP-SREBP pathway (when cells move from a low lipid cultural condition to an abundant lipid cultural condition), TMEM115 coordinates the downregulation of SCAP through proteasomal degradation. Another possibility, although our data are not fully consistent with this idea, is that TMEM115 promotes the retrograde trafficking of SCAP back to the ER, from where ERAD might occur. If the sole role of TMEM115 is to promote the retrograde trafficking of SCAP, in wild type MEFs upon inhibition of proteasomal degradation after re-addition of lipid, SCAP would be redistributed from the Golgi and readily detected in the ER (which is not observed).

In my BioID list of TMEM115 interacting proteins, five candidates have known roles in protein degradation, among which, p62 was one of the top candidates in the list, and the interaction was validated through co-immunoprecipitation. p62 is a central coordinator of both proteasomal and lysosomal degradation. p62 is a cytosolic protein with three important domains, a proteasome binding domain (PB domain), a ubiquitin-associated domain (UAB domain) and an LIR motif for

interacting with lysosomes (Liu et al. 2016). Although a soluble cytosolic protein, it is also detected associated with membranes, and has been detected at the ER, trans-Golgi region (TGN) and late endosomes (Sanchez et al. 1998).

The interaction between p62 and TMEM115 was validated with co-immunoprecipitation in Chapter 3. To examine whether p62 also interacts with SCAP, I performed a co-immunoprecipitation experiment in which I overexpressed HA-p62 and pulled down p62 with HA beads. I was able to detect both endogenous SCAP and TMEM115 in the materials eluted from the HA beads, indicating an interaction between the two proteins (Figure 5.3A). This is an exciting result, as it also suggested a link between p62 and SCAP. But whether the three proteins form a complex, or whether the interaction between p62 and SCAP is an independent event from the interaction between p62 and TMEM115, is still an open question. To test if TMEM115 is essential for the interaction between p62 and SCAP, I used wild type and *Tmem115* knockout MEFs for my co-immunoprecipitation experiment to ask whether the absence of TMEM115 affected the binding of SCAP and p62. By co-overexpressing both SCAP and HA-p62, followed by enrichment with HA beads, I was able to detect SCAP in both wild type and *Tmem115* knockout samples, with no difference in levels. This suggests that TMEM115 is not a necessary component for the interaction between p62 and SCAP, although it certainly does not rule out the possibility that TMEM115 and p62 may collaborate in regulating the SCAP-SREBP pathway.

To extend our understanding of the relationship between TMEM115 and p62, Dr Moncada-Pazos performed an immunofluorescent analysis, in which the distribution of tagged p62 was examined in wild type and *Tmem115* knockout MEFs. A clear difference in the distribution of p62 puncta was observed: the puncta were distributed significantly further away from the Golgi in the absence of TMEM115, and this could be rescued by adding back TMEM115 to knockout cells. Conversely, overexpression of TMEM115 significantly reduced the distance of p62 puncta from the Golgi. This suggests that the presence of TMEM115 in some way affects the subcellular localisation of p62.

In summary, the interaction between p62 and SCAP was validated with co-immunoprecipitation and was found to be TMEM115-independent. Dr Moncada-Pazos showed the absence of TMEM115 affects the subcellular localisation of p62. Given our results that the levels of TMEM115 affect SCAP levels and that there is involvement of proteasomal degradation, this p62 connection is very interesting and will need further investigation in the future. Specifically, it will be important to determine whether p62 acts as an adaptor between TMEM115 and the proteasome for regulating the degradation of SCAP; and whether this interaction between p62 and SCAP uses one of the known degradation pathways, or instead hints at the presence of an independent Golgi-associated degradation pathway.

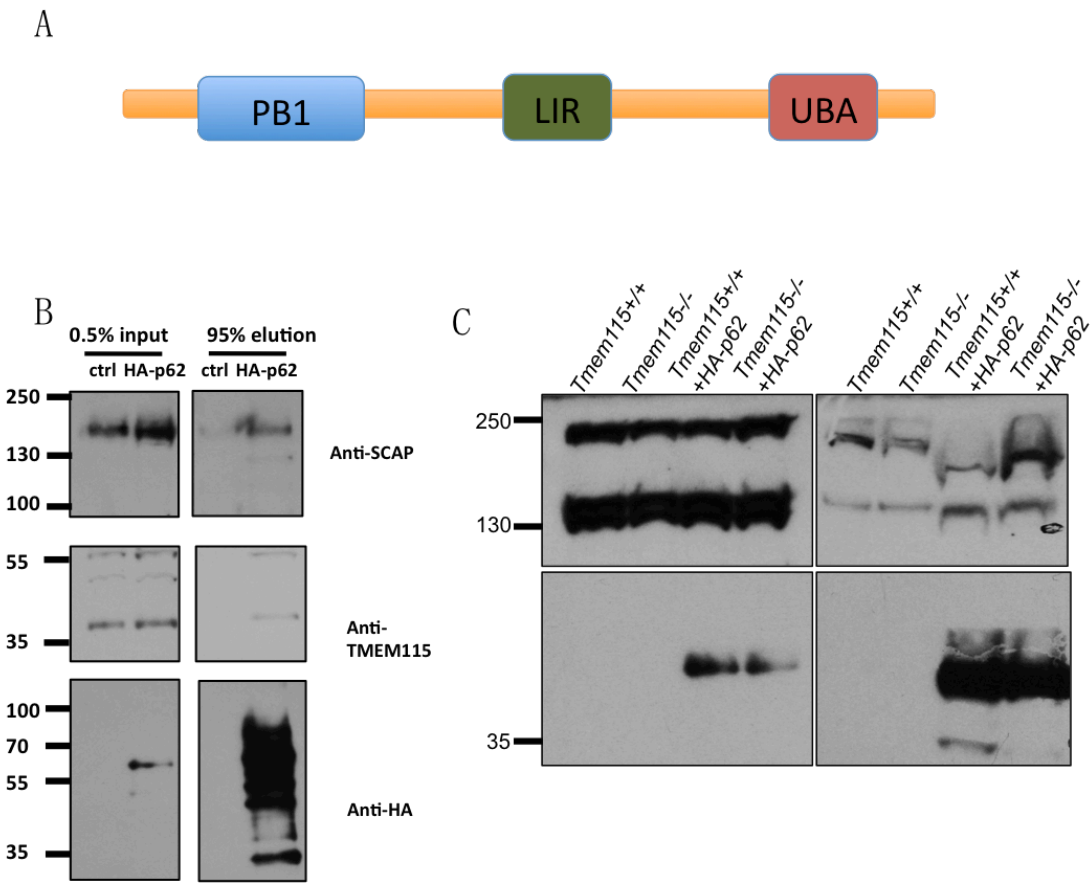


Figure 5.3 p62 is a physical interaction for both TMEM115 and SCAP
A. Model of p62 topology.
B. Overexpressed HA-p62 interacts with endogenous TMEM115 and SCAP. 24 hours before HA immunoprecipitation, HEK293T cells were transfected with either empty vector or HA-p62. Cell lysates and HA beads eluate were immunoblotted for SCAP, TMEM115 and HA (HA-p62). Endogenous SCAP and TMEM115 were detected in HA beads eluate of HA-p62 sample.
C. The interaction between SCAP and p62 is TMEM115 independent. Wild type and Tmem115 knockout MEFs were co-transfected with Flag-SCAP and empty vector or HA-p62. HA immunoprecipitation was performed 48 hours post transfection. No difference were detected in terms of SCAP level in HA beads eluate of Wild type and Tmem115 knockout MEFs.

5.5 Changes in p62 protein level affect the level of endogenous SCAP

From previous observations, TMEM115 seems to destabilise SCAP, possibly through proteasomal degradation. In addition, TMEM115 interacts with and regulates the distribution of p62, an adaptor for both proteasomal and lysosomal degradation. If p62 has a role in assisting TMEM115 to transport SCAP for proteasomal degradation, I should be able to observe that changes in p62 levels should have an effect on SCAP levels.

I performed an experiment in wild type MEFs in which cells were transfected with either TMEM115 or p62 alone, or in combination, and assayed endogenous SCAP in cell lysates and quantified it relative to beta-actin. Overexpression of p62 decreased the level of SCAP (Figure 5.4, mean=59.21% when compared with empty vector control, p-value =0.0264), as did overexpression of TMEM115 (Figure 5.4, mean=39.76% when compared with empty vector control, p-value =0.0022). When both TMEM115 and p62 were overexpressed, a sharp decrease in SCAP levels was observed (Figure 5.4, mean=28.79%, p-value= 0.0002,). The effect of co-overexpression of p62 and TMEM115 on SCAP levels could be a result of them working together to regulate SCAP stability; importantly, however, without further analysis, it is not possible to rule out the possibility that this is simply an additive effect of two distinct regulatory pathways.

In contrast to the quite clear consequences of p62 overexpression, loss of function of p62 provided surprising results that add complexity to the result and its interpretation. Upon knockdown of p62 in HeLa cells, endogenous SCAP was decreased under both normal and low lipid conditions, with a greater reduction under normal lipid conditions (Figure 5.5). The endogenous TMEM115 level is also affected upon p62 knockdown: an increase in TMEM115 protein level was observed under basal conditions. Consistent with the changes in SCAP level upon p62 knockdown under basal condition, Dr Moncada-Pazos observed that two of the SREBP downstream target genes, HMGCS and LDLR, had significantly reduced transcript levels under the basal condition in p62 knockdown samples.

In summary, both overexpression and knockdown of p62 reduce endogenous SCAP protein level under high lipid conditions. These observations add some complexity to the initial hypothesis. The reduction in endogenous SCAP level upon p62 knockdown is inconsistent with the hypothesis that p62 may be a component of the regulatory function of TMEM115 on SCAP stability. The whole regulatory system is more complicated than our initial assumption. However, given that TMEM115 level also changed upon p62 knockdown, p62 could, therefore, regulate SCAP level through TMEM115. More experiments are needed to test this hypothesis.

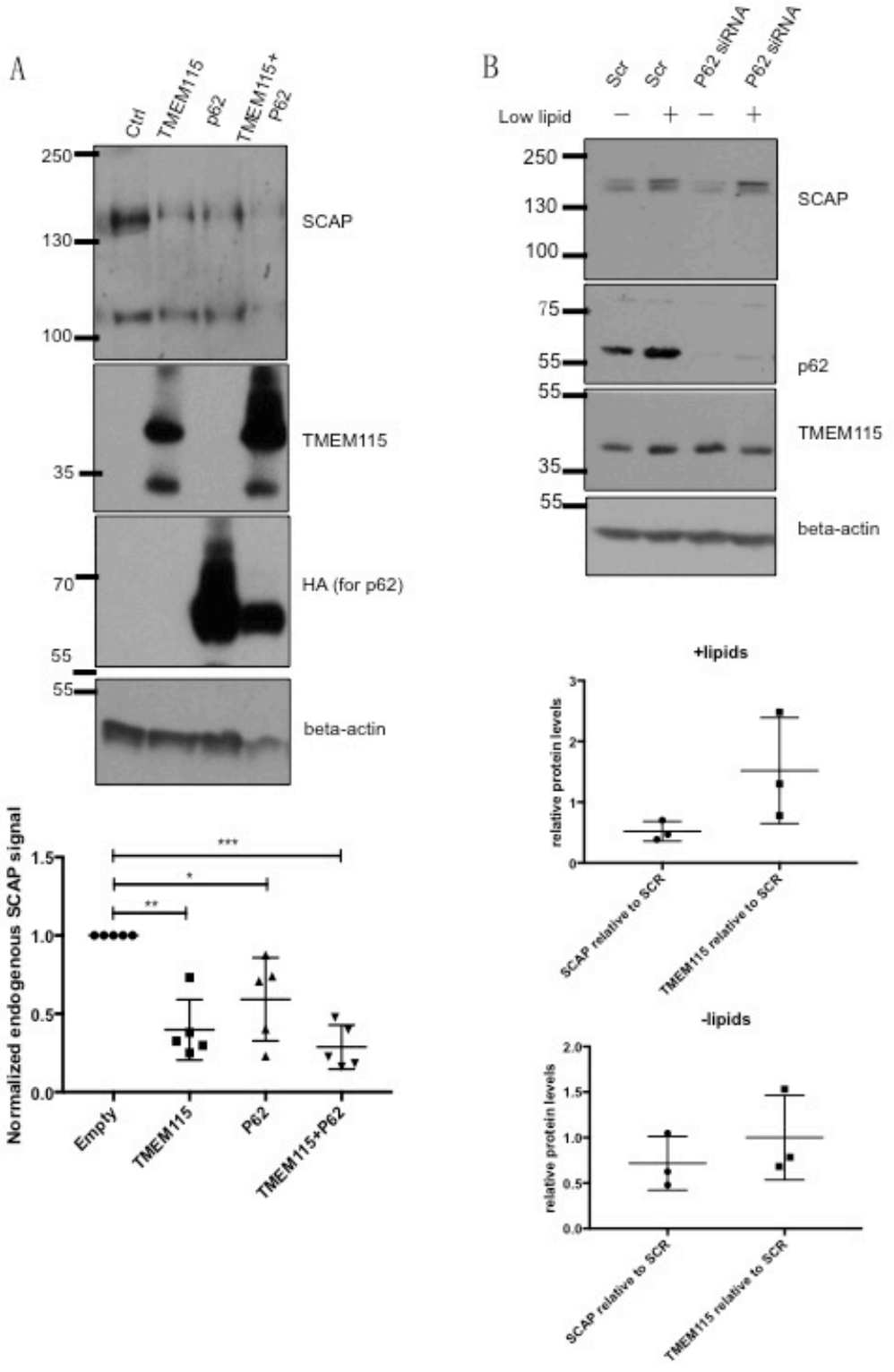


Figure 5.4 p62 regulates the level of SCAP

A. Overexpression of p62 decreases endogenous SCAP level. Wild type MEFs were transfected with empty vector, Flag-TMEM115, HA-p62 or both

Flag-TMEM115 and HA-p62 for 48 hours. Top panel: representative western blot of the levels of SCAP, HA (for p62), TMEM115 and β-actin as endogenous control. Bottom panel: quantification of the levels of SCAP under overexpression of TMEM115 and relative to control cells (n=5). The p-value is the result of testing the experimental values under a Wilcoxon Signed Rank Test for theoretical median=1. **B.** Knockdown of *p62* decreases endogenous SCAP level under both basal and low lipid conditions. HeLa cells were transfected with either scramble siRNA or *p62* siRNA for 48 hours, with low lipid treatment in the final 16 hours. Top panel: representative western blot of the levels of SCAP, p62, TMEM115 and β-actin as endogenous control. Bottom panel: quantification of the levels of SCAP and TMEM115 under knockdown (KD) of *p62* and relative to control cells (n=3). The p-value is the result of testing the experimental values under one sample T-Test.

5.6 Discussion

5.6.1 The physiological significance of TMEM115's role in maintaining lipid homeostasis

Although the molecular mechanism of how TMEM115 regulates the SCAP-SREBP pathway is not fully understood, the overall message of the results (combining mine and those of Dr Moncada-Pazos) is consistent: loss of TMEM115 causes dysregulation of lipid homeostasis at both cellular and tissue levels, which implies the physiological significance of TMEM115's role in this tightly regulated pathway. *Tmem115* knockout mice usually die at weaning or soon after, when they start losing body mass dramatically. The change of diet at weaning, from high fat milk to solid food that contains mostly protein and carbohydrates, triggers activation of the SCAP-SREBP pathway (Botolin et al. 2003) and this is defective in *Tmem115* knockout mice, which could lead to their death. Consistent with this idea, Dr Moncada-Pazos had shown that in livers of *Tmem115* knockout mice that died just before or at weaning, the expression of SREBP targets was significantly lower than in wild-type controls. Given that the SCAP-SREBP pathway is crucial in vertebrates for *de novo* lipogenesis (Horton et al. 2002), failure of activating this pathway led to inadequate lipid synthesis is consistent with the phenotype of *Tmem115* knockout mice.

5.6.2 TMEM115 regulates the SCAP-SREBP pathway in the Golgi

The SCAP-SREBP pathway is the core component for regulating sterol homeostasis, and SCAP is a crucial player in this pathway (Horton et al. 2002). The mechanisms of SCAP anchoring, release, and trafficking have been well studied (Horton et al. 2002, Sun et al. 2005, Xu et al. 2016), but much less is known about the regulation of the pathway in the Golgi apparatus after SREBP cleavage.

Several papers have described that SCAP travels back to the ER through retrograde trafficking (Shao et al. 2014, Takashima et al. 2015). Shao et al.

reported that this retrograde traffic occurs post SREBP cleavage and that upon blocking SREBP cleavage by inhibition of site-1 protease activity, instead of going back to the ER, the whole SCAP-SREBP complex is degraded in the lysosome. In a slightly different conclusion, Takashima et al. have reported that both SCAP alone and SCAP-SREBP complexes can be recycled through retrograde trafficking back to the ER. Despite the inconsistency of whether SCAP-SREBP complex can be recycled, there is a consensus that retrograde trafficking from the Golgi to the ER is an important element of SCAP regulation. Degradation of SCAP is also part of the lipid regulatory machinery. Shao et al. have reported that SCAP can be degraded by the lysosome, and a more recent paper reported that SCAP is sensitive to ubiquitin-mediated proteasomal degradation (Asano et al. 2017),

After investigation of the relationship between TMEM115 the known elements of the SCAP-SREBP pathway, what is consistent is that TMEM115 is important for the proper regulation of the SCAP-SREBP pathway under different conditions, possibly through regulating SCAP levels in the Golgi. Nevertheless, the molecular mechanism still not fully clear. *Tmem115* knockout mice do not have embryonic development defects (the frequency of knockout animal birth is Mendelian), while the phenotype of null mutations of SREBPs, SCAP and INSIG1-2, are much more severe (Shimano et al. 1997, Engelking et al. 2006, Suzuki et al. 2013, Vergnes et al. 2016, Engelking et al. 2017). The differences in phenotypes suggest that rather than being a necessary component for SCAP activity; TMEM115 is more likely to be a modulator.

What has been published about SCAP post-Golgi regulation includes retrograde transport back to the ER (Shao et al. 2014, Takashima et al. 2015), and degradation by both proteasomal (Asano et al. 2017) and/or lysosomal degradation (Shao et al. 2014). Taking together these published papers and our observations, several models can be raised for explaining the possible role of TMEM115 in the SREBP-SCAP pathway. TMEM115 may be involved in retrograde trafficking of SCAP and/or regulation of SCAP degradation. I will

discuss these models and how they support or are inconsistent with our observations.

The simplest model is that TMEM115 is only responsible for the retrograde trafficking of SCAP. It is consistent with what has been published about a role for TMEM115 in general retrograde trafficking by Ong et al., and with our observations. Under high lipid conditions during tissue culture, in the absence of TMEM115, a significant portion of SCAP was abnormally present in the Golgi. This was associated with slightly elevated constitutive activation of the SCAP-SREBP pathway in normal conditions. It also potentially explains the defect in SREBP induction under low-lipid conditions: because the Golgi pool of SCAP was not able to recycle back to the ER through retrograde trafficking, all the anchoring proteins in the Golgi were occupied, resulting in difficulties in accepting more SCAP-SREBP complexes from the ER. This model, however, is inconsistent with the observation that the proteasomal inhibitor, MG132, delays the loss of SCAP signal from the Golgi in wild type cells. If TMEM115 is only involved in retrograde trafficking of SCAP, instead of detecting SCAP in the Golgi after re-adding lipid for 2 hours, an increased signal of SCAP in the ER should be observed. Overall, the relationship between TMEM115 and SCAP is likely more complex than the simple assumption that TMEM115 is only responsible for the retrograde trafficking of SCAP.

Alternatively (although not mutually incompatible with the idea that TMEM115 promotes retrograde trafficking of SCAP), TMEM115 may be involved in targeting degradation of SCAP from the Golgi through a proteasome-dependent mechanism. Indeed, degradation of SCAP post-Golgi through the proteasome has been reported (Asano et al. 2017). This proteasomal degradation model is consistent with the live imaging result that inhibiting proteasomal degradation has a similar effect as the absence of TMEM115 on delaying the loss of Golgi SCAP signal. It would also explain the result that a reduction of TMEM115 protein level by siRNA knockdown led to an increase in endogenous SCAP, while overexpression of TMEM115 reduces endogenous SCAP levels. Nevertheless, how a Golgi protein

might regulate protein degradation through the proteasome is complicated. It is possible that TMEM115 serves as a cargo protein for trafficking SCAP for degradation. Perhaps significantly, PSMD-1, a component of the 26S proteasome (Kanayama et al. 1992), is a candidate for a TMEM115 interacting proteins in the BioID list, suggesting that BirA-TMEM115 was in close proximity with PSMD-1, or possibly the 26S proteasome, during the 24 hours biotinylation labelling period.

Another speculative possibility is that TMEM115 collaborates with one or more adaptor proteins during this process, and that the role of TMEM115 is to hand over SCAP from the Golgi exit site, and that the adaptor is in charge of transporting SCAP to the proteasome for degradation. The interactions between TMEM115, p62 and SCAP suggest that p62 may also regulate SCAP, but whether p62 is directly involved in TMEM115's regulation of the SCAP-SREBP pathway and the contribution of p62 to this pathway, is still unknown. One possible hypothesis is that TMEM115 is responsible for recruiting p62 to the Golgi to transport the Golgi pool of SCAP for degradation (a simplified diagram is presented in Figure 5.5). In the absence of TMEM115, p62 would not be able to target the Golgi correctly, and that could delay SCAP exit from the Golgi. p62 coordinates protein degradation through both the proteasome and the lysosome; together with the fact that p62 interacts with both TMEM115 and SCAP, it is an appealing hypothesis that TMEM115 and p62 regulate the degradation of SCAP through the proteasome. Another result that in the absence of TMEM115, p62 was distributed abnormally and significantly further away from the Golgi is consistent with the earlier model that TMEM115 is responsible for recruiting p62 to Golgi for transporting SCAP for proteasomal degradation. A model for all these data could, therefore, be that TMEM115 is in charge of both exporting SCAP from the Golgi and recruiting p62 to the Golgi, while p62 transports SCAP to the proteasome for degradation. Nevertheless, this idea is inconsistent with the effect of knockdown of p62 on SCAP level. The model would predict that an insufficient amount of p62 would lead to accumulation SCAP due to reduced turnover, while the result was actually the opposite. Further indirect evidence that is inconsistent with the current model is the phenotype of p62 knockout mice. Unlike *Tmem115* knockout mice,

which lose weight, p62 null mice exhibit mature-onset obesity and insulin resistance and eventually develop type-II diabetes (Okada et al. 2009). All these inconsistencies suggest that p62 may not be in the same regulatory pathway with TMEM115 in modulating SCAP functions. Still, it does not rule out the possibility that TMEM115 collaborates with other adaptors for regulating the degradation of SCAP.

Taken together, TMEM115 may regulate the degradation of SCAP through the proteasome. This is supported by my data that provide clear evidence that TMEM115 can influence levels of endogenous SCAP under low lipid conditions, as well as Dr Moncada-Pazos's live imaging result that the deletion of TMEM115 has a similar effect to proteasomal inhibition on delaying SCAP dynamics at the Golgi. However, the molecular mechanism needs further investigation. TMEM115 may also regulate the retrograde trafficking of SCAP. However, it cannot be the sole molecular function of TMEM115 in SCAP regulation.

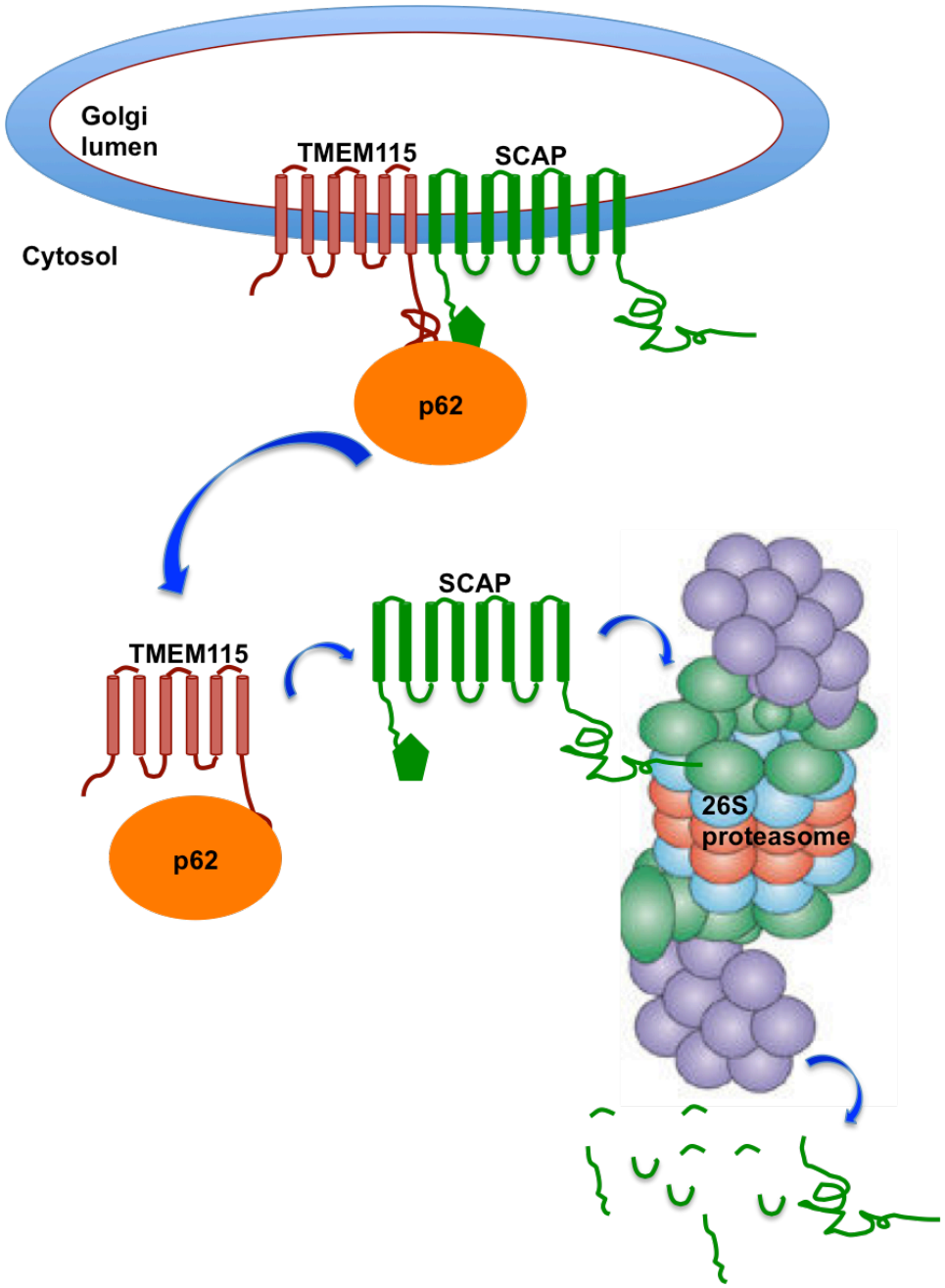


Figure 5.5 A diagram of how TMEM115 may collaborate with p62 for regulating the level of SCAP

TMEM115 may regulate the degradation of SCAP in the Golgi through interacting with p62. To switch off the SCAP-SREBP pathway promptly, TMEM115 interacts with SCAP in the Golgi and recruits p62 to the Golgi. p62 shufflers TMEM115-SCAP to the 26S proteasome, where SCAP gets degraded. However, this model is inconsistent with the result that endogenous SCAP level increases upon both overexpress and knockdown of p62.

Chapter 6 Final perspective

6.1 Overview of the project

The work in my thesis aims to add insight to both the function and structural aspects of TMEM115, a highly conserved, yet poorly studied clan of the rhomboid-like superfamily. When I started this project three years ago, there was no known function assigned for TMEM115. Using biochemical approaches, I have identified a list of candidate binding partners, which suggests intriguing possible cellular functions of TMEM115, including regulating trafficking, protein degradation, lipid biology and ion channels. I have also discovered a number of validated binding partners of TMEM115, including ERGIC-53, VCP/p97, p62, β-COP and SCAP. My topological study, together with bioinformatic analysis data obtained from Phyre and HHpred (Soding et al. 2005, Kelley et al. 2009), unambiguously positions TMEM115 in the rhomboid-like superfamily. I have also revealed possible interacting sites of TMEM115 for the tested membrane proteins and soluble binding partners. The investigation of a possible relationship between TMEM115 and SCAP suggests that TMEM115 may affect the stability of the Golgi pool of SCAP. With the efforts of both myself and Dr Moncada-Pazos, we now show that TMEM115 has a role in regulating lipid homeostasis through the fundamental lipid synthesis SCAP-SREBP pathway (Nohturfft et al. 1999).

Overall, I have made progress on characterising the novel rhomboid-like protein, TMEM115, both structurally and functionally. Nevertheless, the functional and structural aspects of TMEM115 are not yet fully understood. Structurally, the topology study of TMEM115 in combination with bioinformatic analysis provided some preliminary information about the secondary structure features of TMEM115. Functionally, the role of TMEM115 in lipid homeostasis is confirmed, but the molecular mechanism is not clear. In addition, two novel binding partners of TMEM115, p62 and VCP/P97, are both key components in protein degradation machinery (Seibenhener et al. 2004, Komatsu et al. 2007, Pankiv et al. 2007, Ballar et al. 2008). That p62 and VCP/p97 are interactors suggests that TMEM115 may have a role in regulating degradation of some of its binding partners. I will

discuss what is known about TMEM115 at this stage, together with the direction of further structural and functional study, which may help to extend our knowledge of this *bona fide* rhomboid-like protein.

6.2 Structural conclusions about TMEM115

My topology study of TMEM115 clearly showed that TMEM115 has six TMDs, which are located in the N-terminus portion of the protein. These data are inconsistent with what has been reported by Ong et al., in which a four TMDs topology model with extended loops was proposed (Ong et al. 2014). Ong et al. used TMpred for topology prediction, which provides two different topology models for TMEM115. The preferred model contains five TMDs, with a luminal C-terminal tail. The second favoured model is what had been reported in their paper, with 4 TMDs and cytosolic orientation of both termini. Ong et al. applied a conditional permeabilisation approach for mapping the localisation of C-terminus and confirmed its cytosolic orientation but did not report any further experimental confirmation of their proposal. I applied four different topology prediction programmes: Toppred, TMHMM, MEMSAT and SOSUI. Most of the predicted models have six TMDs. In addition, I used experimental methods and mapped the orientation of each predicted loop region. Consistent with the Ong et al. report, both termini of TMEM115 are cytosolic, and the TMD region is localised at the N-terminal part of the protein.

The topology study of TMEM115 has revealed a few additional structural features. According to Phyre and HHpred, the TMD region of TMEM115 shares structural similarity with other rhomboid-like proteins. While based on my data of the possible locations of each TMD, the luminal loop 1 is predicted to be surprisingly short (5 amino acids in length). It is unusual for rhomboid-like proteins, as most of they have a large luminal loop between the first and the second TMD of the core rhomboid-like domain, which contains a conserved WR motif (in rhomboid proteases, derlins and UBAC2) or an IRHD domain (iRhoms) (Wang et al. 2006, Adrain et al. 2012, Lemberg 2013). In addition, the proposed helix-helix dimerisation motif (GxxxG) (Russ et al. 2000) is localised within the fourth TMD of

TMEM115, unlike most of the rhomboid-like members, which have a similar GxxxG motif in the sixth TMD of their rhomboid-like domain (Wang et al. 2006, Lemieux et al. 2007, Lemberg 2013). Another intriguing motif is at the cytosolic C-terminus of TMEM115, which shares high similarity with the VCP/P97 interacting motif at the cytosolic C-terminal tail of RHBDL4 (Lim et al. 2016). Having the generic rhomboid-like domain together with some possible unique structural features at the core domain makes TMEM115 an appealing protein for 3D structural study.

As mentioned in Chapter 4, the revolutionary technique of cryo-EM has been used extensively in studying membrane protein structures, with a limitation in detection threshold of protein/complex size: proteins must greater than about 145 kDa (Merk et al. 2016). The size of TMEM115 is far below the threshold, but cryo-EM might still be used to resolve the structure of TMEM115 in a complex with its binding partners, for instance, SCAP (around 150 kDa) or VCP/p97 (around 100kDa). The obvious advantage of studying TMEM115 structure in a complex is the structure will be obtained in a more biological relevant condition, which is helpful to acquire detailed structural information of the interaction interface. This might not be apparent from the structures of individual components, since proteins may need to go through conformational change while interacting with different binding partners (Bhardwaj et al. 2011). The challenge, however, is to obtain high purity proteins that form a stable complex, and this may be difficult given that the interaction between TMEM115 and both SCAP and VCP/p97 are relatively modest in co-immunoprecipitation.

6.3 TMEM115 and its interactors

A number of TMEM115 interacting proteins have been identified with BioID and/or co-immunoprecipitation. The relationship between TMEM115 and SCAP has been investigated and a role of TMEM115 in lipid homeostasis is confirmed, while the molecular mechanism of how TMEM115 regulates SCAP is not yet fully understood. The biological significance linked with other binding partners is understudied, which may reveal additional roles of TMEM115.

Two proteins, ERGIC-53 and β-COP, which are crucial for membrane trafficking between the ER and the Golgi, are physical binding partners of TMEM115 (Appenzeller et al. 1999, Hauri et al. 2000, Beck et al. 2009). A role of TMEM115 in regulating general retrograde trafficking and O-linked glycosylation was also reported by Ong et al. (Ong et al. 2014). However, it was inconsistent with my SPECS data, in which I showed there were no measurable differences in secretome profiles between wild type and *Tmem115* knockout MEFs. The SPECS assay relies on metabolic sugar labelling at the terminal positions of both N- and O-linked glycan chains of secreted proteins (Kuhn et al. 2012, Kuhn et al. 2015). If a general O-linked glycosylation defect existed, I would observe significantly fewer secreted proteins with O-linked glycosylation modification in the eluate of *Tmem115* knockout MEFs, when compared with wild type cells. However, it certainly does not rule out the possibility that TMEM115 regulates forward/retrograde trafficking of some of its binding partners through interaction with ERGIC-53 or β-COP. It is also possible that the observed interactions reflect the biosynthetic trafficking of TMEM115 itself rather than a role in regulating trafficking.

The interactions between TMEM115 and both p62 and VCP/p97, which are essential components of the protein degradation machinery (Seibenhener et al. 2004, Komatsu et al. 2007, Ballar et al. 2008, Moscat et al. 2009, Moscat et al. 2009, Hanzelmann et al. 2011), are completely novel and unexpected. p62 is an adaptor protein for both proteasomal and lysosomal degradation, while it also controls the positioning and dynamics of endosomal vesicle transport (Seibenhener et al. 2004, Pankiv et al. 2007, Jongsma et al. 2016, Liu et al. 2016). Intriguingly, the absence of TMEM115 leads to a clear difference in the distribution of p62 puncta: they were significantly further away from the Golgi (unpublished data in Freeman lab), while the position of early and late endosomes relative to Golgi were unaffected. It indicates that TMEM115 does not affect the function of p62 in regulating endosomal architecture. Instead, TMEM115 may be involved in recruiting p62 to the Golgi for sequential functions, possibly regulating the degradation of some of its binding partners.

VCP/p97 is a cytosolic AAA ATPase that has multiple functions. Functional diversity of VCP/p97 is determined by its binding to different cofactors (Ye 2006, Schuberth et al. 2008, Yamanaka et al. 2012). It governs critical steps in ubiquitin-dependent protein quality control and intracellular signalling pathways (Kobayashi et al. 2007, Ballar et al. 2008, Meyer et al. 2012). In addition, VCP/p97 has been shown to regulate autophagy, endosomal sorting and protein degradation at the outer membrane of mitochondria (Ju et al. 2009, Yamanaka et al. 2012, Kirchner et al. 2013). VCP/p97 binds to a number of ER localised rhomboid-like proteins that are essential for ERAD, including derlins and RHBDL4 (Ye et al. 2005, Lim et al. 2016). What could be the biological significance of the interaction between the Golgi localised rhomboid-like protein and VCP/p97? Given VCP/p97 has multiple functions, it is very difficult to predict at this stage the exact function of VCP/p97 in the context of TMEM115. It is unlikely, however, that the loss of TMEM115 leads to a general defect in VCP/p97 related functions. VCP/p97 may act as a cofactor of TMEM115 to regulate, for instance, the degradation of TMEM115 binding partners.

6.4 TMEM115 and SCAP

A role of TMEM115 in lipid homeostasis has been shown at both cellular level and in animal models, and SCAP is the crucial binding partner for TMEM115's role in the regulation of lipid homeostasis. However, the molecular mechanism of how TMEM115 regulates SCAP is largely unclear. Nevertheless, our preliminary data together with a possible consensus of how rhomboid-like proteins regulate their binding partners support several possible models of how TMEM115 may regulate SCAP.

TMEM115 is a distant member of the rhomboid-like superfamily. Conserved structural features suggest implying that all members are evolved from the same ancestor (Koonin et al. 2003, Freeman 2014), suggesting that they might also share functional characteristics. Members of the rhomboid-like superfamily regulate distinctive and crucial intracellular signalling pathways, including cell proliferation, immune responses, protein quality control and parasites invasion

(Greenblatt et al. 2011, Zettl et al. 2011, Adrain et al. 2012, Shen et al. 2014, Luo et al. 2016). Despite their diversified roles, there is a consensus in the fundamental machinery that rhomboid-like proteins utilise in the regulation of these pathways. Most rhomboid-like proteins participate in one or a combination of the following types of machinery: proteolysis, trafficking and protein degradation. Regarding active rhomboids, they bind and cleave their substrates, releasing them from their membrane tethers or promoting their degradation. RHBDL4, for instance, binds ubiquitinated single-spanning and polytopic membrane proteins with unstable membrane helices and cleaves them before sending them for ERAD (Fleig et al. 2012). Pseudoproteases within the superfamily, which have lost their crucial catalytic residues, can promote the trafficking and regulate the degradation of their binding partners through different protein quality control systems (Greenblatt et al. 2011, Zettl et al. 2011). Derlins and UBAC2 are directly involved in the primary ER protein quality control mechanism, ERAD (Christianson et al. 2011, Zettl et al. 2011). In the case of mammalian iRhom2, it regulates two different immune response pathways, TNF signalling and innate immunity to DNA viruses, through different binding partners with strikingly similar mechanisms (Adrain et al. 2012, Luo et al. 2016, Grieve et al. 2017). IRhom2 promotes the ER to Golgi forward trafficking of both TACE and STING, and protects them from lysosomal and proteasomal degradation, respectively (Luo et al. 2016, Grieve et al. 2017).

A link between the SREBP pathway and rhomboid-like proteins can be traced back to the fission yeast, *Schizosaccharomyces pombe* (Kim et al. 2015, Hwang et al. 2016, Burr et al. 2017). In fission yeast, the SREBP pathway is much simpler when compared to the mammalian system. Instead of being cleaved by site1 and site 2 proteases, SREBPs are cleaved in the Golgi with a combination of effects of the Golgi rhomboid protease Rbd2, the essential ATPase Cdc48 (fission yeast homolog of mammalian VCP) and the Dsc E3 ligase complex (Burr et al. 2017). This is an intriguing coincidence, especially when considering our data with TMEM115, which is also a Golgi resident rhomboid-like protein that regulates the SCAP/SREBP pathway in flies and mammals, and also interacts with VCP in

mammalian cell lines. Although it may be a coincidence, it might also suggest an evolutionarily conserved relationship between rhomboid-like proteins and lipid homeostasis.

Combining our observations that TMEM115 affects the level of SCAP possibly via a proteasomal dependent mechanism, together with the identification of TMEM115 as a rhomboid-like protein, the current hypothesis is TMEM115 may regulate its binding partner, SCAP, by modulating SCAP degradation and/or trafficking. If TMEM115 promotes the degradation of SCAP, it would explain the absence of SCAP in the BioID screen. The presence of protein degradation inhibitors would be required in the proximity-labelling step to prevent biotinylated SCAP from being degraded.

The SCAP-SREBP pathway is highly dynamic and tightly modulated to ensure an appropriate response to nutrient changes in the environment. It is possible that there might be a choice/decision point that controls which pathway SCAP takes after delivery to the Golgi: recycling or degradation. Under normal conditions, when plenty of lipid is present, only a small proportion of SCAP-SREBP goes to the Golgi and it is likely to be quickly transported back to the ER (Shao et al. 2014). When the lipids run in short supply, the majority of SCAP-SREBP complexes are transported to the Golgi for the activation of the pathway, possibly with some retrograde trafficking of SCAP back to the ER to reload with more SREBPs. When cells need to switch back from low lipid conditions to normal conditions, the pathway needs to be switched off. This is when I believe it is likely that the Golgi pool of SCAP is transported out of Golgi for degradation. I hypothesise that the molecular roles for TMEM115 are both the retrograde trafficking and degradation of SCAP, but there are different priorities under different lipid conditions. When the SCAP-SREBP activation needs to be tuned down, TMEM115 sends SCAP for proteasomal degradation. In other conditions, TMEM115 regulates the retrograde trafficking of SCAP.

6.5 Conclusion mark

In conclusion, exciting progress has been made on elucidating the functional and structural aspects of the novel rhomboid-like protein, TMEM115, during my DPhil project. Nevertheless, some intriguing questions of TMEM115 are still unsolved, which would improve our knowledge of the biological roles of TMEM115 in depth.

TMEM115 is an essential regulator of lipid homeostasis at both cellular and organism levels. The fact that TMEM115 interact and regulate SCAP, the essential component of the fundamental *de novo* lipogenesis SCAP-SREBP pathway, underlying the physiological significance of TMEM115. However, whether the severe defects of *Tmem115* null animals are caused solely by dysregulated lipid homeostasis is still debatable. A list of binding partners of TMEM115 have been identified via biochemical approaches, which suggests multiple intriguing possible cellular functions of TMEM115 apart from lipid biology, including regulating trafficking, protein degradation and ion channels. Among which, VCP/p97 and p62 are two novel interactors of TMEM115 that have been validated by co-immunoprecipitation. Both novel binding partners have multiple roles in regulating protein degradation through lysosomal and proteasome. Does it suggest that TMEM115, a Golgi-localised rhomboid-like protein, is involved in regulating degradation of some of its binding partners, such as SCAP? Preliminary investigation of how TMEM115 regulates SCAP in Chapter 5 supports the hypothesis that TMEM115 regulates the degradation of SCAP post Golgi through proteasome, while the fundamental mechanism is still unrevealed. Whether TMEM115 regulates SCAP level through one of several of its binding partners, such as p62 or VCP/p97, is an interesting question to solve. Moreover, identifying new functions of TMEM115 based on the BioID candidates' list is also an intriguing avenue to follow up.

The topology study has confirmed a six TMD topology of TMEM115. It also highlighted the N-terminus part of the protein to be where the rhomboid-like domain located, which is crucial for the interactions of all the tested binding partners, irrespective of their topology. The fact that the TMD region of TMEM115

can still interact with its soluble cytosolic binding partners and with even higher affinity is a complex but intriguing question to address. Therefore, more detailed interaction region mapping with different truncated versions of TMEM115 will provide insights of how TMEM115 interact with different binding partners.

References:

Abba, M. C., E. Lacunza, M. I. Nunez, A. Colussi, M. Isla-Larrain, A. Segal-Eiras, M. V. Croce and C. M. Aldaz (2009). "Rhomboid domain containing 2 (RHBDD2): a novel cancer-related gene over-expressed in breast cancer." Biochim Biophys Acta **1792**(10): 988-997.

Adrain, C. and M. Freeman (2012). "New lives for old: evolution of pseudoenzyme function illustrated by iRhoms." Nat Rev Mol Cell Biol **13**(8): 489-498.

Adrain, C., K. Strisovsky, M. Zettl, L. Hu, M. K. Lemberg and M. Freeman (2011). "Mammalian EGF receptor activation by the rhomboid protease RHBDL2." EMBO Rep **12**(5): 421-427.

Adrain, C., M. Zettl, Y. Christova, N. Taylor and M. Freeman (2012). "Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE." Science **335**(6065): 225-228.

Aebi, M. (2013). "N-linked protein glycosylation in the ER." Biochim Biophys Acta **1833**(11): 2430-2437.

Agmon, E. and B. R. Stockwell (2017). "Lipid homeostasis and regulated cell death." Curr Opin Chem Biol **39**: 83-89.

Appenzeller, C., H. Andersson, F. Kappeler and H. P. Hauri (1999). "The lectin ERGIC-53 is a cargo transport receptor for glycoproteins." Nat Cell Biol **1**(6): 330-334.

Asano, L., M. Watanabe, Y. Ryoden, K. Usuda, T. Yamaguchi, B. Khambu, M. Takashima, S. I. Sato, J. Sakai, K. Nagasawa and M. Uesugi (2017). "Vitamin D Metabolite, 25-Hydroxyvitamin D, Regulates Lipid Metabolism by Inducing Degradation of SREBP/SCAP." Cell Chem Biol **24**(2): 207-217.

Baker, R. P. and S. Urban (2015). "Cytosolic extensions directly regulate a rhomboid protease by modulating substrate gating." Nature **523**(7558): 101-105.

Ballar, P. and S. Fang (2008). "Regulation of ER-associated degradation via p97/VCP-interacting motif." Biochem Soc Trans **36**(Pt 5): 818-822.

Beck, R., M. Rawet, F. T. Wieland and D. Cassel (2009). "The COPI system: molecular mechanisms and function." FEBS Lett **583**(17): 2701-2709.

Ben-Shem, A., D. Fass and E. Bibi (2007). "Structural basis for intramembrane

proteolysis by rhomboid serine proteases." Proc Natl Acad Sci U S A **104**(2): 462-466.

Bergbold, N. and M. K. Lemberg (2013). "Emerging role of rhomboid family proteins in mammalian biology and disease." Biochim Biophys Acta **1828**(12): 2840-2848.

Berggard, T., S. Linse and P. James (2007). "Methods for the detection and analysis of protein-protein interactions." Proteomics **7**(16): 2833-2842.

Bhardwaj, N., A. Abyzov, D. Clarke, C. Shou and M. B. Gerstein (2011). "Integration of protein motions with molecular networks reveals different mechanisms for permanent and transient interactions." Protein Sci **20**(10): 1745-1754.

Botolin, D. and D. B. Jump (2003). "Selective proteolytic processing of rat hepatic sterol regulatory element binding protein-1 (SREBP-1) and SREBP-2 during postnatal development." J Biol Chem **278**(9): 6959-6962.

Burr, R., D. Ribbens, S. Raychaudhuri, E. V. Stewart, J. Ho and P. J. Espenshade (2017). "Dsc E3 ligase localization to the Golgi requires the ATPase Cdc48 and cofactor Ufd1 for activation of Sterol Regulatory Element-binding Protein in fission yeast." J Biol Chem.

Carrara, G., N. Saraiva, C. Gubser, B. F. Johnson and G. L. Smith (2012). "Six-transmembrane topology for Golgi anti-apoptotic protein (GAAP) and Bax inhibitor 1 (BI-1) provides model for the transmembrane Bax inhibitor-containing motif (TMBIM) family." J Biol Chem **287**(19): 15896-15905.

Cheng, T. L., C. H. Lai, S. J. Jiang, J. H. Hung, S. K. Liu, B. I. Chang, G. Y. Shi and H. L. Wu (2014). "RHBDL2 is a critical membrane protease for anoikis resistance in human malignant epithelial cells." ScientificWorldJournal **2014**: 902987.

Christianson, J. C., J. A. Olzmann, T. A. Shaler, M. E. Sowa, E. J. Bennett, C. M. Richter, R. E. Tyler, E. J. Greenblatt, J. W. Harper and R. R. Kopito (2011). "Defining human ERAD networks through an integrative mapping strategy." Nat Cell Biol **14**(1): 93-105.

Christianson, J. C., J. A. Olzmann, T. A. Shaler, M. E. Sowa, E. J. Bennett, C. M. Richter, R. E. Tyler, E. J. Greenblatt, J. W. Harper and R. R. Kopito (2012). "Defining human ERAD networks through an integrative mapping strategy." Nat Cell Biol **14**(1): 93-105.

Christova, Y., C. Adrain, P. Bambrough, A. Ibrahim and M. Freeman (2013). "Mammalian iRhoms have distinct physiological functions including an essential role in TACE regulation." EMBO Rep **14**(10): 884-890.

Cipolat, S., T. Rudka, D. Hartmann, V. Costa, L. Serneels, K. Craessaerts, K. Metzger, C. Frezza, W. Annaert, L. D'Adamio, C. Derks, T. Dejaegere, L. Pellegrini, R. D'Hooge, L. Scorrano and B. De Strooper (2006). "Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via

OPA1-dependent cristae remodeling." Cell **126**(1): 163-175.

Civitarese, A. E., P. S. MacLean, S. Carling, L. Kerr-Bayles, R. P. McMillan, A. Pierce, T. C. Becker, C. Moro, J. Finlayson, N. Lefort, C. B. Newgard, L. Mandarino, W. Cefalu, K. Walder, G. R. Collier, M. W. Hulver, S. R. Smith and E. Ravussin (2010). "Regulation of skeletal muscle oxidative capacity and insulin signaling by the mitochondrial rhomboid protease PARL." Cell Metab **11**(5): 412-426.

Demishtein, A., M. Fraiberg, D. Berko, B. Tirosh, Z. Elazar and A. Navon (2017). "SQSTM1/p62-mediated autophagy compensates for loss of proteasome polyubiquitin recruiting capacity." Autophagy: 1-12.

Dusterhoft, S., U. Kunzel and M. Freeman (2017). "Rhomboid proteases in human disease: Mechanisms and future prospects." Biochim Biophys Acta.

Engelking, L. J., M. J. Cantoria, Y. Xu and G. Liang (2017). "Developmental and extrahepatic physiological functions of SREBP pathway genes in mice." Semin Cell Dev Biol.

Engelking, L. J., B. M. Evers, J. A. Richardson, J. L. Goldstein, M. S. Brown and G. Liang (2006). "Severe facial clefting in Insig-deficient mouse embryos caused by sterol accumulation and reversed by lovastatin." J Clin Invest **116**(9): 2356-2365.

Eura, Y., H. Yanamoto, Y. Arai, T. Okuda, T. Miyata and K. Kokame (2012). "Derlin-1 deficiency is embryonic lethal, Derlin-3 deficiency appears normal, and Herp deficiency is intolerant to glucose load and ischemia in mice." PLoS One **7**(3): e34298.

Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate and A. Bateman (2016). "The Pfam protein families database: towards a more sustainable future." Nucleic Acids Res **44**(D1): D279-285.

Fleig, L., N. Bergbold, P. Sahasrabudhe, B. Geiger, L. Kaltak and M. K. Lemberg (2012). "Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins." Mol Cell **47**(4): 558-569.

Freeman, M. (2004). "Proteolysis within the membrane: rhomboids revealed." Nat Rev Mol Cell Biol **5**(3): 188-197.

Freeman, M. (2014). "The rhomboid-like superfamily: molecular mechanisms and biological roles." Annu Rev Cell Dev Biol **30**: 235-254.

Greenblatt, E. J., J. A. Olzmann and R. R. Kopito (2011). "Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum." Nat Struct Mol Biol **18**(10): 1147-1152.

Grieve, A. G., H. Xu, U. Kunzel, P. Bambrough, B. Sieber and M. Freeman (2017). "Phosphorylation of iRhom2 at the plasma membrane controls mammalian TACE-dependent inflammatory and growth factor signalling." Elife **6**.

Grundy, S. M. (2004). "Obesity, metabolic syndrome, and cardiovascular disease."

J Clin Endocrinol Metab **89**(6): 2595-2600.

Hampton, R. Y. and T. Sommer (2012). "Finding the will and the way of ERAD substrate retrotranslocation." Curr Opin Cell Biol **24**(4): 460-466.

Han, J., J. Bai, Y. Yang, H. Yin, W. Gao, A. Lu, F. Liu, H. Ge, Z. Liu, J. Wang and L. Zhong (2015). "Lentivirus-mediated knockdown of rhomboid domain containing 1 inhibits colorectal cancer cell growth." Mol Med Rep **12**(1): 377-381.

Hanzelmann, P. and H. Schindelin (2011). "The structural and functional basis of the p97/valosin-containing protein (VCP)-interacting motif (VIM): mutually exclusive binding of cofactors to the N-terminal domain of p97." J Biol Chem **286**(44): 38679-38690.

Hauri, H. P., F. Kappeler, H. Andersson and C. Appenzeller (2000). "ERGIC-53 and traffic in the secretory pathway." J Cell Sci **113 (Pt 4)**: 587-596.

Hitt, R. and D. H. Wolf (2004). "Der1p, a protein required for degradation of malfolded soluble proteins of the endoplasmic reticulum: topology and Der1-like proteins." FEMS Yeast Res **4**(7): 721-729.

Horton, J. D., J. L. Goldstein and M. S. Brown (2002). "SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver." J Clin Invest **109**(9): 1125-1131.

Hosokawa, N., I. Wada, K. Hasegawa, T. Yorihuzi, L. O. Tremblay, A. Herscovics and K. Nagata (2001). "A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation." EMBO Rep **2**(5): 415-422.

Hwang, J., D. Ribbens, S. Raychaudhuri, L. Cairns, H. Gu, A. Frost, S. Urban and P. J. Espenshade (2016). "A Golgi rhomboid protease Rbd2 recruits Cdc48 to cleave yeast SREBP." EMBO J **35**(21): 2332-2349.

Issuree, P. D., T. Maretzky, D. R. McIlwain, S. Monette, X. Qing, P. A. Lang, S. L. Swendeman, K. H. Park-Min, N. Binder, G. D. Kalliolias, A. Yarilina, K. Horiuchi, L. B. Ivashkiv, T. W. Mak, J. E. Salmon and C. P. Blobel (2013). "iRHOM2 is a critical pathogenic mediator of inflammatory arthritis." J Clin Invest **123**(2): 928-932.

Ivanova, A. V., A. Vortmeyer, S. V. Ivanov, M. L. Nickerson, E. R. Maher and M. I. Lerman (2008). "Loss of PL6 protein expression in renal clear cell carcinomas and other VHL-deficient tumours." J Pathol **214**(1): 46-57.

Johnson, N., J. Brezinova, E. Stephens, E. Burbridge, M. Freeman, C. Adrain and K. Strisovsky (2017). "Quantitative proteomics screen identifies a substrate repertoire of rhomboid protease RHBDL2 in human cells and implicates it in epithelial homeostasis." Sci Rep **7**(1): 7283.

Jongsma, M. L., I. Berlin, R. H. Wijdeven, L. Janssen, G. M. Janssen, M. A. Garstka, H. Janssen, M. Mensink, P. A. van Veelen, R. M. Spaapen and J. Neefjes (2016). "An ER-Associated Pathway Defines Endosomal Architecture for Controlled Cargo Transport." Cell **166**(1): 152-166.

Ju, J. S., R. A. Fuentealba, S. E. Miller, E. Jackson, D. Piwnica-Worms, R. H.

Baloh and C. C. Weihl (2009). "Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease." J Cell Biol **187**(6): 875-888.

Kanayama, H. O., T. Tamura, S. Ugai, S. Kagawa, N. Tanahashi, T. Yoshimura, K. Tanaka and A. Ichihara (1992). "Demonstration that a human 26S proteolytic complex consists of a proteasome and multiple associated protein components and hydrolyzes ATP and ubiquitin-ligated proteins by closely linked mechanisms." Eur J Biochem **206**(2): 567-578.

Kelley, L. A. and M. J. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." Nat Protoc **4**(3): 363-371.

Kim, D. I., K. C. Birendra, W. Zhu, K. Motamedchaboki, V. Doye and K. J. Roux (2014). "Probing nuclear pore complex architecture with proximity-dependent biotinylation." Proceedings of the National Academy of Sciences of the United States of America **111**: E2453-2461.

Kim, D. I., S. C. Jensen, K. A. Noble, B. Kc, K. H. Roux, K. Motamedchaboki and K. J. Roux (2016). "An improved smaller biotin ligase for BioID proximity labeling." Mol Biol Cell **27**(8): 1188-1196.

Kim, D. I., S. C. Jensen and K. J. Roux (2016). "Identifying Protein-Protein Associations at the Nuclear Envelope with BioID." Methods Mol Biol **1411**: 133-146.

Kim, J., H. J. Ha, S. Kim, A. R. Choi, S. J. Lee, K. L. Hoe and D. U. Kim (2015). "Identification of Rbd2 as a candidate protease for sterol regulatory element binding protein (SREBP) cleavage in fission yeast." Biochem Biophys Res Commun **468**(4): 606-610.

Kirchner, P., M. Bug and H. Meyer (2013). "Ubiquitination of the N-terminal region of caveolin-1 regulates endosomal sorting by the VCP/p97 AAA-ATPase." J Biol Chem **288**(10): 7363-7372.

Knop, M., A. Finger, T. Braun, K. Hellmuth and D. H. Wolf (1996). "Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast." EMBO J **15**(4): 753-763.

Kobayashi, T., A. Manno and A. Kakizuka (2007). "Involvement of valosin-containing protein (VCP)/p97 in the formation and clearance of abnormal protein aggregates." Genes Cells **12**(7): 889-901.

Komatsu, M., S. Waguri, M. Koike, Y. S. Sou, T. Ueno, T. Hara, N. Mizushima, J. Iwata, J. Ezaki, S. Murata, J. Hamazaki, Y. Nishito, S. Iemura, T. Natsume, T. Yanagawa, J. Uwayama, E. Warabi, H. Yoshida, T. Ishii, A. Kobayashi, M. Yamamoto, Z. Yue, Y. Uchiyama, E. Kominami and K. Tanaka (2007). "Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice." Cell **131**(6): 1149-1163.

Koonin, E. V., K. S. Makarova, I. B. Rogozin, L. Davidovic, M. C. Letellier and L. Pellegrini (2003). "The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene

transfers." Genome Biol **4**(3): R19.

Kuhn, P. H., K. Koroniak, S. Hogl, A. Colombo, U. Zeitschel, M. Willem, C. Volbracht, U. Schepers, A. Imhof, A. Hoffmeister, C. Haass, S. Rossner, S. Brase and S. F. Lichtenthaler (2012). "Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons." EMBO J **31**(14): 3157-3168.

Kuhn, P. H., M. Voss, M. Haug-Kroper, B. Schroder, U. Schepers, S. Brase, C. Haass, S. F. Lichtenthaler and R. Fluhrer (2015). "Secretome analysis identifies novel signal Peptide peptidase-like 3 (Sppl3) substrates and reveals a role of Sppl3 in multiple Golgi glycosylation pathways." Mol Cell Proteomics **14**(6): 1584-1598.

Lacunza, E., R. Canzoneri, M. E. Rabassa, A. Zwenger, A. Segal-Eiras, M. V. Croce and M. C. Abba (2012). "RHBDD2: a 5-fluorouracil responsive gene overexpressed in the advanced stages of colorectal cancer." Tumour Biol **33**(6): 2393-2399.

Lacunza, E., M. E. Rabassa, R. Canzoneri, M. Pellon-Maison, M. V. Croce, C. M. Aldaz and M. C. Abba (2014). "Identification of signaling pathways modulated by RHBDD2 in breast cancer cells: a link to the unfolded protein response." Cell Stress Chaperones **19**(3): 379-388.

Lastun, V. L., A. G. Grieve and M. Freeman (2016). "Substrates and physiological functions of secretase rhomboid proteases." Semin Cell Dev Biol **60**: 10-18.

Lemberg, M. K. (2013). "Sampling the membrane: function of rhomboid-family proteins." Trends Cell Biol **23**(5): 210-217.

Lemberg, M. K. and M. Freeman (2007). "Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases." Genome Res **17**(11): 1634-1646.

Lemberg, M. K., J. Menendez, A. Misik, M. Garcia, C. M. Koth and M. Freeman (2005). "Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases." EMBO J **24**(3): 464-472.

Lemieux, M. J., S. J. Fischer, M. M. Cherney, K. S. Bateman and M. N. James (2007). "The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis." Proc Natl Acad Sci U S A **104**(3): 750-754.

Lespagnol, A., D. Duflaut, C. Beekman, L. Blanc, G. Fiucci, J. C. Marine, M. Vidal, R. Amson and A. Telerman (2008). "Exosome secretion, including the DNA damage-induced p53-dependent secretory pathway, is severely compromised in TSAP6/Steap3-null mice." Cell Death Differ **15**(11): 1723-1733.

Li, X., T. Maretzky, G. Weskamp, S. Monette, X. Qing, P. D. Issuree, H. C. Crawford, D. R. McIlwain, T. W. Mak, J. E. Salmon and C. P. Blobel (2015). "iRhoms 1 and 2 are essential upstream regulators of ADAM17-dependent EGFR signaling." Proc Natl Acad Sci U S A **112**(19): 6080-6085.

Liao, H. J. and G. Carpenter (2012). "Regulated intramembrane cleavage of the EGF receptor." Traffic **13**(8): 1106-1112.

Lichtenthaler, S. F., B. F. O'Hara and C. P. Blobel (2015). "iRhoms in the brain - a new frontier?" Cell Cycle **14**(19): 3003-3004.

Lilley, B. N. and H. L. Ploegh (2004). "A membrane protein required for dislocation of misfolded proteins from the ER." Nature **429**(6994): 834-840.

Lim, J. J., Y. Lee, T. T. Ly, J. Y. Kang, J. G. Lee, J. Y. An, H. S. Youn, K. R. Park, T. G. Kim, J. K. Yang, Y. Jun and S. H. Eom (2016). "Structural insights into the interaction of p97 N-terminus domain and VBM in rhomboid protease, RHBDL4." Biochem J **473**(18): 2863-2880.

Liu, J., C. Han, B. Xie, Y. Wu, S. Liu, K. Chen, M. Xia, Y. Zhang, L. Song, Z. Li, T. Zhang, F. Ma, Q. Wang, J. Wang, K. Deng, Y. Zhuang, X. Wu, Y. Yu, T. Xu and X. Cao (2014). "Rhbdd3 controls autoimmunity by suppressing the production of IL-6 by dendritic cells via K27-linked ubiquitination of the regulator NEMO." Nat Immunol **15**(7): 612-622.

Liu, J., S. Liu, M. Xia, S. Xu, C. Wang, Y. Bao, M. Jiang, Y. Wu, T. Xu and X. Cao (2013). "Rhomboid domain-containing protein 3 is a negative regulator of TLR3-triggered natural killer cell activation." Proc Natl Acad Sci U S A **110**(19): 7814-7819.

Liu, W. J., L. Ye, W. F. Huang, L. J. Guo, Z. G. Xu, H. L. Wu, C. Yang and H. F. Liu (2016). "p62 links the autophagy pathway and the ubiqutin-proteasome system upon ubiquitinated protein degradation." Cell Mol Biol Lett **21**: 29.

Lohi, O., S. Urban and M. Freeman (2004). "Diverse substrate recognition mechanisms for rhomboids; thrombomodulin is cleaved by Mammalian rhomboids." Curr Biol **14**(3): 236-241.

Luo, W. W., S. Li, C. Li, H. Lian, Q. Yang, B. Zhong and H. B. Shu (2016). "iRhom2 is essential for innate immunity to DNA viruses by mediating trafficking and stability of the adaptor STING." Nat Immunol **17**(9): 1057-1066.

MacKenzie, K. R., J. H. Prestegard and D. M. Engelman (1997). "A transmembrane helix dimer: structure and implications." Science **276**(5309): 131-133.

Maretzky, T., D. R. McIlwain, P. D. Issuree, X. Li, J. Malapeira, S. Amin, P. A. Lang, T. W. Mak and C. P. Blobel (2013). "iRhom2 controls the substrate selectivity of stimulated ADAM17-dependent ectodomain shedding." Proc Natl Acad Sci U S A **110**(28): 11433-11438.

Martinez, A., P. Fullwood, K. Kondo, T. Kishida, M. Yao, E. R. Maher and F. Latif (2000). "Role of chromosome 3p12-p21 tumour suppressor genes in clear cell renal cell carcinoma: analysis of VHL dependent and VHL independent pathways of tumorigenesis." Mol Pathol **53**(3): 137-144.

Martinez-Vicente, M. (2017). "Neuronal Mitophagy in Neurodegenerative Diseases." Front Mol Neurosci **10**: 64.

Maruthappu, T., A. Chikh, B. Fell, P. J. Delaney, M. A. Brooke, C. Levet, A. Moncada-Pazos, A. Ishida-Yamamoto, D. Blaydon, A. Waseem, I. M. Leigh, M. Freeman and D. P. Kelsell (2017). "Rhomboid family member 2 regulates cytoskeletal stress-associated Keratin 16." Nat Commun **8**: 14174.

McIlwain, D. R., P. A. Lang, T. Maretzky, K. Hamada, K. Ohishi, S. K. Maney, T. Berger, A. Murthy, G. Duncan, H. C. Xu, K. S. Lang, D. Haussinger, A. Wakeham, A. Itie-Youten, R. Khokha, P. S. Ohashi, C. P. Blobel and T. W. Mak (2012). "iRhom2 regulation of TACE controls TNF-mediated protection against Listeria and responses to LPS." Science **335**(6065): 229-232.

McQuibban, G. A., J. R. Lee, L. Zheng, M. Juusola and M. Freeman (2006). "Normal mitochondrial dynamics requires rhomboid-7 and affects Drosophila lifespan and neuronal function." Curr Biol **16**(10): 982-989.

Mehnert, M., T. Sommer and E. Jarosch (2014). "Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane." Nat Cell Biol **16**(1): 77-86.

Mehus, A. A., R. H. Anderson and K. J. Roux (2016). "BioID Identification of Lamin-Associated Proteins." Methods Enzymol **569**: 3-22.

Meissner, C., H. Lorenz, A. Weihofen, D. J. Selkoe and M. K. Lemberg (2011). "The mitochondrial intramembrane protease PARL cleaves human Pink1 to regulate Pink1 trafficking." J Neurochem **117**(5): 856-867.

Merk, A., A. Bartesaghi, S. Banerjee, V. Falconieri, P. Rao, M. I. Davis, R. Pragani, M. B. Boxer, L. A. Earl, J. L. Milne and S. Subramaniam (2016). "Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery." Cell **165**(7): 1698-1707.

Meyer, H., M. Bug and S. Bremer (2012). "Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system." Nat Cell Biol **14**(2): 117-123.

Moscat, J. and M. T. Diaz-Meco (2009). "p62 at the crossroads of autophagy, apoptosis, and cancer." Cell **137**(6): 1001-1004.

Moscat, J. and M. T. Diaz-Meco (2009). "To aggregate or not to aggregate? A new role for p62." EMBO Rep **10**(8): 804.

Nohturfft, A., R. A. DeBose-Boyd, S. Scheek, J. L. Goldstein and M. S. Brown (1999). "Sterols regulate cycling of SREBP cleavage-activating protein (SCAP) between endoplasmic reticulum and Golgi." Proc Natl Acad Sci U S A **96**(20): 11235-11240.

Nooren, I. M. and J. M. Thornton (2003). "Diversity of protein-protein interactions." EMBO J **22**(14): 3486-3492.

Noy, P. J., R. K. Swain, K. Khan, P. Lodhia and R. Bicknell (2016). "Sprouting angiogenesis is regulated by shedding of the C-type lectin family 14, member A (CLEC14A) ectodomain, catalyzed by rhomboid-like 2 protein (RHBDL2)." FASEB J **30**(6): 2311-2323.

Oda, Y., T. Okada, H. Yoshida, R. J. Kaufman, K. Nagata and K. Mori (2006). "Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation." J Cell Biol **172**(3): 383-393.

Oka, T., D. Ungar, F. M. Hughson and M. Krieger (2004). "The COG and COPI complexes interact to control the abundance of GEARs, a subset of Golgi integral membrane proteins." Mol Biol Cell **15**(5): 2423-2435.

Okada, K., T. Yanagawa, E. Warabi, K. Yamastu, J. Uwayama, K. Takeda, H. Utsunomiya, H. Yoshida, J. Shoda and T. Ishii (2009). "The alpha-glucosidase inhibitor acarbose prevents obesity and simple steatosis in sequestosome 1/A170/p62 deficient mice." Hepatol Res **39**(5): 490-500.

Olzmann, J. A., C. M. Richter and R. R. Kopito (2013). "Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover." Proc Natl Acad Sci U S A **110**(4): 1345-1350.

Ong, Y. S., T. H. Tran, N. V. Gounko and W. Hong (2014). "TMEM115 is an integral membrane protein of the Golgi complex involved in retrograde transport." J Cell Sci **127**(Pt 13): 2825-2839.

Pankiv, S., T. H. Clausen, T. Lamark, A. Brech, J. A. Bruun, H. Outzen, A. Overvatn, G. Bjorkoy and T. Johansen (2007). "p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy." J Biol Chem **282**(33): 24131-24145.

Quiroga, R., A. Trenchi, A. Gonzalez Montoro, J. Valdez Taubas and H. J. Maccioni (2013). "Short transmembrane domains with high-volume exoplasmic halves determine retention of Type II membrane proteins in the Golgi complex." J Cell Sci **126**(Pt 23): 5344-5349.

Ren, X., W. Song, W. Liu, X. Guan, F. Miao, S. Miao and L. Wang (2013). "Rhomboid domain containing 1 inhibits cell apoptosis by upregulating AP-1 activity and its downstream target Bcl-3." FEBS Lett **587**(12): 1793-1798.

Rodolfo, C., S. Campello and F. Cecconi (2017). "Mitophagy in neurodegenerative diseases." Neurochem Int.

Roux, K. J. (2013). "Marked by association: techniques for proximity-dependent labeling of proteins in eukaryotic cells." Cellular and Molecular Life Sciences **70**: 3657-3664.

Roux, K. J., D. I. Kim, M. Raida and B. Burke (2012). "A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells." The Journal of cell biology **196**: 801-810.

Russ, W. P. and D. M. Engelman (2000). "The GxxxG motif: a framework for transmembrane helix-helix association." J Mol Biol **296**(3): 911-919.

Sahin, U., G. Weskamp, K. Kelly, H. M. Zhou, S. Higashiyama, J. Peschon, D. Hartmann, P. Saftig and C. P. Blobel (2004). "Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands." J Cell Biol **164**(5): 769-779.

Sanchez, P., G. De Carcer, I. V. Sandoval, J. Moscat and M. T. Diaz-Meco (1998). "Localization of atypical protein kinase C isoforms into lysosome-targeted endosomes through interaction with p62." Mol Cell Biol **18**(5): 3069-3080.

Schuberth, C. and A. Buchberger (2008). "UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97." Cell Mol Life Sci **65**(15): 2360-2371.

Seibenhener, M. L., J. R. Babu, T. Geetha, H. C. Wong, N. R. Krishna and M. W. Wooten (2004). "Sequestosome 1/p62 is a polyubiquitin chain binding protein involved in ubiquitin proteasome degradation." Mol Cell Biol **24**(18): 8055-8068.

Shao, W. and P. J. Espenshade (2014). "Sterol regulatory element-binding protein (SREBP) cleavage regulates Golgi-to-endoplasmic reticulum recycling of SREBP cleavage-activating protein (SCAP)." J Biol Chem **289**(11): 7547-7557.

Shen, B., J. S. Buguliskis, T. D. Lee and L. D. Sibley (2014). "Functional analysis of rhomboid proteases during Toxoplasma invasion." MBio **5**(5): e01795-01714.

Shimano, H., I. Shimomura, R. E. Hammer, J. Herz, J. L. Goldstein, M. S. Brown and J. D. Horton (1997). "Elevated levels of SREBP-2 and cholesterol synthesis in livers of mice homozygous for a targeted disruption of the SREBP-1 gene." J Clin Invest **100**(8): 2115-2124.

Siggs, O. M., N. Xiao, Y. Wang, H. Shi, W. Tomisato, X. Li, Y. Xia and B. Beutler (2012). "iRhom2 is required for the secretion of mouse TNFalpha." Blood **119**(24): 5769-5771.

Smith, R. D. and V. V. Lupashin (2008). "Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation." Carbohydr Res **343**(12): 2024-2031.

Smith, S. O., D. Song, S. Shekar, M. Groesbeek, M. Ziliox and S. Aimoto (2001). "Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers." Biochemistry **40**(22): 6553-6558.

Soding, J., A. Biegert and A. N. Lupas (2005). "The HHpred interactive server for protein homology detection and structure prediction." Nucleic Acids Res **33**(Web Server issue): W244-248.

Song, W., W. Liu, H. Zhao, S. Li, X. Guan, J. Ying, Y. Zhang, F. Miao, M. Zhang, X. Ren, X. Li, F. Wu, Y. Zhao, Y. Tian, W. Wu, J. Fu, J. Liang, W. Wu, C. Liu, J. Yu, S. Zong, S. Miao, X. Zhang and L. Wang (2015). "Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway." Nat Commun **6**: 8022.

Strisovsky, K., H. J. Sharpe and M. Freeman (2009). "Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates." Mol Cell **36**(6): 1048-1059.

Sun, L. P., L. Li, J. L. Goldstein and M. S. Brown (2005). "Insig required for sterol-mediated inhibition of Scap/SREBP binding to COPII proteins in vitro." J Biol Chem **280**(28): 26483-26490.

Suzuki, R., H. A. Ferris, M. J. Chee, E. Maratos-Flier and C. R. Kahn (2013).

"Reduction of the cholesterol sensor SCAP in the brains of mice causes impaired synaptic transmission and altered cognitive function." PLoS Biol **11**(4): e1001532.

Takashima, K., A. Saitoh, T. Funabashi, S. Hirose, C. Yagi, S. Nozaki, R. Sato, H. W. Shin and K. Nakayama (2015). "COPI-mediated retrieval of SCAP is crucial for regulating lipogenesis under basal and sterol-deficient conditions." J Cell Sci **128**(15): 2805-2815.

Taxis, C., R. Hitt, S. H. Park, P. M. Deak, Z. Kostova and D. H. Wolf (2003). "Use of modular substrates demonstrates mechanistic diversity and reveals differences in chaperone requirement of ERAD." J Biol Chem **278**(38): 35903-35913.

Teese, M. G. and D. Langosch (2015). "Role of GxxxG Motifs in Transmembrane Domain Interactions." Biochemistry **54**(33): 5125-5135.

Urban, S. and M. Freeman (2003). "Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain." Mol Cell **11**(6): 1425-1434.

Urban, S., J. R. Lee and M. Freeman (2001). "Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases." Cell **107**(2): 173-182.

Vashist, S. and D. T. Ng (2004). "Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control." J Cell Biol **165**(1): 41-52.

Vergnes, L., R. G. Chin, T. de Aguiar Vallim, L. G. Fong, T. F. Osborne, S. G. Young and K. Reue (2016). "SREBP-2-deficient and hypomorphic mice reveal roles for SREBP-2 in embryonic development and SREBP-1c expression." J Lipid Res **57**(3): 410-421.

Wang, X., R. Sato, M. S. Brown, X. Hua and J. L. Goldstein (1994). "SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis." Cell **77**(1): 53-62.

Wang, Y., X. Guan, K. L. Fok, S. Li, X. Zhang, S. Miao, S. Zong, S. S. Koide, H. C. Chan and L. Wang (2008). "A novel member of the Rhomboid family, RHBDD1, regulates BIK-mediated apoptosis." Cell Mol Life Sci **65**(23): 3822-3829.

Wang, Y., Y. Zhang and Y. Ha (2006). "Crystal structure of a rhomboid family intramembrane protease." Nature **444**(7116): 179-180.

Wasserman, J. D., S. Urban and M. Freeman (2000). "A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling." Genes Dev **14**(13): 1651-1663.

Weber, P. C., D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme (1989). "Structural origins of high-affinity biotin binding to streptavidin." Science **243**(4887): 85-88.

Wei, X., T. Lv, D. Chen and J. Guan (2014). "Lentiviral vector mediated delivery of RHBDD1 shRNA down regulated the proliferation of human glioblastoma cells." Technol Cancer Res Treat **13**(1): 87-93.

Wu, Z., N. Yan, L. Feng, A. Oberstein, H. Yan, R. P. Baker, L. Gu, P. D. Jeffrey, S.

Urban and Y. Shi (2006). "Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry." Nat Struct Mol Biol **13**(12): 1084-1091.

Wunderle, L., J. D. Knopf, N. Kuhnle, A. Morle, B. Hehn, C. Adrain, K. Strisovsky, M. Freeman and M. K. Lemberg (2016). "Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGFalpha." Sci Rep **6**: 27342.

Xu, H. F., J. Luo, H. P. Wang, H. Wang, T. Y. Zhang, H. B. Tian, D. W. Yao and J. J. Loor (2016). "Sterol regulatory element binding protein-1 (SREBP-1)c promoter: Characterization and transcriptional regulation by mature SREBP-1 and liver X receptor alpha in goat mammary epithelial cells." J Dairy Sci **99**(2): 1595-1604.

Yamanaka, K., Y. Sasagawa and T. Ogura (2012). "Recent advances in p97/VCP/Cdc48 cellular functions." Biochim Biophys Acta **1823**(1): 130-137.

Ye, Y. (2006). "Diverse functions with a common regulator: ubiquitin takes command of an AAA ATPase." J Struct Biol **156**(1): 29-40.

Ye, Y., H. H. Meyer and T. A. Rapoport (2001). "The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol." Nature **414**(6864): 652-656.

Ye, Y., Y. Shibata, M. Kikkert, S. van Voorden, E. Wiertz and T. A. Rapoport (2005). "Recruitment of the p97 ATPase and ubiquitin ligases to the site of retrotranslocation at the endoplasmic reticulum membrane." Proc Natl Acad Sci U S A **102**(40): 14132-14138.

Ye, Y., Y. Shibata, C. Yun, D. Ron and T. A. Rapoport (2004). "A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol." Nature **429**(6994): 841-847.

Zettl, M., C. Adrain, K. Strisovsky, V. Lastun and M. Freeman (2011). "Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signaling." Cell **145**(1): 79-91.

DiRusso, C.C., et al., (2005) Comparative biochemical studies of the murine fatty acid transport proteins (FATP) expressed in yeast. J Biol Chem,.280 (17):p.16829-37.

Hua, Z. and T.R. Graham, (2003) Requirement for neo1p in retrograde transport from the Golgi complex to the endoplasmic reticulum. Mol Biol Cell,. 14(12): p.4971-83.

Tehlivets, O., et al., (2007) Fatty acid synthesis and elongation in yeast. Biochim Biophys Acta, 1771(3): p. 255-70.

Tomsig, J.L., et al., (2009) Lipid phosphate phosphohydrolase type 1 (LPP1) degrades extracellular lysophosphatidic acid in vivo. Biochem J,. 419(3): p.611-8.

Morash, S.C., et al., (1994) Studies employing Saccharomyces cerevisiae cpt1 and ept1 null mutants implicate the CPT1 gene in coordinate regulation of

phospholipid biosynthesis. J Biol Chem,. 269(46): p.28769-76.

Koch H. G., Moser M., and Müller M. (2003) Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. Rev. Physiol. Biochem. Pharmacol. 146, 55–94

Ligang Ren et al. (2017) Increased expression of tumor protein D54 is associated with clinical progression and poor prognosis in patients with prostate cancer. Oncol Lett. 14(6): 7739–7744.

Flora Guerra et al. (2016) Multiple Roles of the Small GTPase Rab7. Cells. 5(3): 34.

Chua CE et al. (2011) Involvement of members of the Rab family and related small GTPases in autophagosome formation and maturation. Cell Mol Life Sci. 68(20):3349-58.

Agola JO et al. (2011) Rab GTPases as regulators of endocytosis, targets of disease and therapeutic opportunities. Clin Genet. 80(4):305-18.

Pereira-Leal JB, et al. (2001) Evolution of the Rab family of small GTP-binding proteins. J Mol Biol. 313(4):889-901.

Bucci C, et al. (2000) Rab7: a key to lysosome biogenesis. Mol Biol Cell. 11(2):467-80.

Progida C, et al. (2010) Rab7b controls trafficking from endosomes to the TGN. J. Cell Sci. 123:1480–1491

Quyen L.Aoh et al. (2009) SCAMP3 negatively regulates epidermal growth factor receptor degradation and promotes receptor recycling. Mol Biol Cell. 20(6):1816-32.

Richard D. Smith et al. (2009) Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. Carbohydr Res. 343(12): 2024–2031

Appendix

A1. Full list of TMEM115 BioID candidates ranked by average spectral count in triplicated BirA-TMEM115 samples

Protein	Description	Spectral counts in BirA-TMEM115 sample
sp|Q12893|TM115_HUMAN	Transmembrane protein 115 OS=Homo sapiens GN=TMEM115 PE=1 SV=1	169.00
sp|Q8TBA6|GOGA5_HUMAN	Golgin subfamily A member 5 OS=Homo sapiens GN=GOLGA5 PE=1 SV=3	69.89
sp|Q9H3P7|GCP60_HUMAN	Golgi resident protein GCP60 OS=Homo sapiens GN=ACBD3 PE=1 SV=4	66.00
sp|P82094-2|TMF1_HUMAN	Isoform 2 of TATA element modulatory factor OS=Homo sapiens GN=TMF1	61.00
sp|Q13439-4|GOGA4_HUMAN	Isoform 4 of Golgin subfamily A member 4 OS=Homo sapiens GN=GOLGA4	55.41
sp|Q9Y6Y8|S23IP_HUMAN	SEC23-interacting protein OS=Homo sapiens GN=SEC23IP PE=1 SV=1	52.89
sp|P49257|LMAN1_HUMAN	Protein ERGIC-53 OS=Homo sapiens GN=LMAN1 PE=1 SV=2	34.57
sp|Q15643|TRIPB_HUMAN	Thyroid receptor-interacting protein 11 OS=Homo sapiens GN=TRIP11 PE=1 SV=3	34.00
sp|O15173-2|PGRC2_HUMAN	Isoform 2 of Membrane-associated progesterone receptor component 2 OS=Homo sapiens GN=PGRMC2	29.99
sp|O75976|CBPD_HUMAN	Carboxypeptidase D OS=Homo sapiens GN=CPD PE=1 SV=2	28.95
sp|Q96S66|CLCC1_HUMAN	Chloride channel CLIC-like protein 1 OS=Homo sapiens GN=CLCC1 PE=1 SV=1	27.97
sp|Q8NBN3|TM87A_HUMAN	Transmembrane protein 87A OS=Homo sapiens GN=TMEM87A PE=1 SV=3	27.67
sp|Q96A33|CCD47_HUM	Coiled-coil domain-containing protein 47	26.88

AN	OS=Homo sapiens GN=CCDC47 PE=1 SV=1	
sp|Q08378-2|GOGA3_HUMAN	Isoform 2 of Golgin subfamily A member 3 OS=Homo sapiens GN=GOLGA3	25.96
sp|Q6P996-4|PDXD1_HUMAN	Isoform 4 of Pyridoxal-dependent decarboxylase domain-containing protein 1 OS=Homo sapiens GN=PDXDC1	25.00
sp|Q14789-2|GOGB1_HUMAN	Isoform 2 of Golgin subfamily B member 1 OS=Homo sapiens GN=GOLGB1	24.91
sp|Q14789|GOGB1_HUMAN	Golgin subfamily B member 1 OS=Homo sapiens GN=GOLGB1 PE=1 SV=2	24.00
sp|Q9H8Y8|GORS2_HUMAN	Golgi reassembly-stacking protein 2 OS=Homo sapiens GN=GORASP2 PE=1 SV=3	24.00
sp|P42566|EPS15_HUMAN	Epidermal growth factor receptor substrate 15 OS=Homo sapiens GN=EPS15 PE=1 SV=2	22.98
sp|Q07065|CKAP4_HUMAN	Cytoskeleton-associated protein 4 OS=Homo sapiens GN=CKAP4 PE=1 SV=2	21.89
sp|Q13501|SQSTM_HUMAN	Sequestosome-1 OS=Homo sapiens GN=SQSTM1 PE=1 SV=1	20.97
sp|P05023-4|AT1A1_HUMAN	Isoform 4 of Sodium/potassium-transporting ATPase subunit alpha-1 OS=Homo sapiens GN=ATP1A1	17.84
sp|P50851-2|LRBA_HUMAN	Isoform 2 of Lipopolysaccharide-responsive and beige-like anchor protein OS=Homo sapiens GN=LRBA	17.69
sp|Q9BSJ8-2|ESYT1_HUMAN	Isoform 2 of Extended synaptotagmin-1 OS=Homo sapiens GN=ESYT1	15.93
sp|O15498|YKT6_HUMAN	Synaptobrevin homolog YKT6 OS=Homo sapiens GN=YKT6 PE=1 SV=1	15.00
sp|O15126|SCAM1_HUMAN	Secretory carrier-associated membrane protein 1 OS=Homo sapiens GN=SCAMP1 PE=1 SV=2	14.00
sp|P08240|SRPR_HUMAN	Signal recognition particle receptor subunit alpha OS=Homo sapiens GN=SRPR PE=1 SV=2	14.00
sp|Q9HCU5|PREB_HUMAN	Prolactin regulatory element-binding protein OS=Homo sapiens GN=PREB PE=1 SV=2	14.00

sp|Q9H2J7|S6A15_HUMAN	Sodium-dependent neutral amino acid transporter B(0)AT2 OS=Homo sapiens GN=SLC6A15 PE=1 SV=1	13.98
sp|P05067-10|A4_HUMAN	Isoform APP639 of Amyloid beta A4 protein OS=Homo sapiens GN=APP	13.98
sp|P50402|EMD_HUMAN	Emerin OS=Homo sapiens GN=EMD PE=1 SV=1	13.65
sp|O43399-5|TPD54_HUMAN	Isoform 5 of Tumor protein D54 OS=Homo sapiens GN=TPD52L2	12.00
sp|Q5VV42|CDKAL_HUMAN	Threonylcarbamoyladenosine tRNA methylthiotransferase OS=Homo sapiens GN=CDKAL1 PE=1 SV=1	12.00
sp|Q99460-2|PSMD1_HUMAN	Isoform 2 of 26S proteasome non-ATPase regulatory subunit 1 OS=Homo sapiens GN=PSMD1	12.00
sp|Q07960|RHG01_HUMAN	Rho GTPase-activating protein 1 OS=Homo sapiens GN=ARHGAP1 PE=1 SV=1	11.97
sp|P51149|RAB7A_HUMAN	Ras-related protein Rab-7a OS=Homo sapiens GN=RAB7A PE=1 SV=1	11.93
sp|Q13190|STX5_HUMAN	Syntaxin-5 OS=Homo sapiens GN=STX5 PE=1 SV=2	11.83
sp|P18031|PTN1_HUMAN	Tyrosine-protein phosphatase non-receptor type 1 OS=Homo sapiens GN=PTPN1 PE=1 SV=1	11.81
sp|P53365-3|ARFP2_HUMAN	Isoform 3 of Arfaptin-2 OS=Homo sapiens GN=ARFIP2	11.00
sp|Q01968-2|OCRL_HUMAN	Isoform B of Inositol polyphosphate 5-phosphatase OCRL-1 OS=Homo sapiens GN=OCRL	10.89
sp|Q04695|K1C17_HUMAN	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2	10.00
sp|P17987|TCPA_HUMAN	T-complex protein 1 subunit alpha OS=Homo sapiens GN=TCP1 PE=1 SV=1	9.95
sp|O75694-2|NU155_HUMAN	Isoform 2 of Nuclear pore complex protein Nup155 OS=Homo sapiens GN=NUP155	9.94
sp|P35580-2|MYH10_HU	Isoform 2 of Myosin-10 OS=Homo sapiens	9.00

MAN	GN=MYH10	
sp|Q9Y2H6-2|FND3A_HUMAN	Isoform 2 of Fibronectin type-III domain-containing protein 3A OS=Homo sapiens GN=FNDC3A	9.00
sp|Q92575|UBXN4_HUMAN	UBX domain-containing protein 4 OS=Homo sapiens GN=UBXN4 PE=1 SV=2	9.00
sp|P08195-2|4F2_HUMAN	Isoform 2 of 4F2 cell-surface antigen heavy chain OS=Homo sapiens GN=SLC3A2	9.00
sp|O96005-4|CLPT1_HUMAN	Isoform 3 of Cleft lip and palate transmembrane protein 1 OS=Homo sapiens GN=CLPTM1	8.99
sp|O14828|SCAM3_HUMAN	Secretory carrier-associated membrane protein 3 OS=Homo sapiens GN=SCAMP3 PE=1 SV=3	8.91
sp|Q04721|NOTC2_HUMAN	Neurogenic locus notch homolog protein 2 OS=Homo sapiens GN=NOTCH2 PE=1 SV=3	8.00
sp|Q8NEN9|PDZD8_HUMAN	PDZ domain-containing protein 8 OS=Homo sapiens GN=PDZD8 PE=1 SV=1	8.00
sp|Q9H3N1|TMX1_HUMAN	Thioredoxin-related transmembrane protein 1 OS=Homo sapiens GN=TMX1 PE=1 SV=1	8.00
sp|Q9HD26-2|GOPC_HUMAN	Isoform 2 of Golgi-associated PDZ and coiled-coil motif-containing protein OS=Homo sapiens GN=GOPC	8.00
sp|Q9Y5Z9-2|UBIA1_HUMAN	Isoform 2 of UbiA prenyltransferase domain-containing protein 1 OS=Homo sapiens GN=UBIAD1	8.00
sp|O43396|TXNL1_HUMAN	Thioredoxin-like protein 1 OS=Homo sapiens GN=TXNL1 PE=1 SV=3	8.00
sp|Q9NYM9|BET1L_HUMAN	BET1-like protein OS=Homo sapiens GN=BET1L PE=1 SV=1	7.94
sp|P16615-2|AT2A2_HUMAN	Isoform 2 of Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 OS=Homo sapiens GN=ATP2A2	7.91
sp|P53367|ARFP1_HUMAN	Arfaptin-1 OS=Homo sapiens GN=ARFIP1 PE=1 SV=2	7.90
sp|Q587I9|SFT2C_HUMA	Vesicle transport protein SFT2C OS=Homo	7.50

N	sapiens GN=SFT2D3 PE=2 SV=1	
sp|Q5HYI8|RABL3_HUMAN	Rab-like protein 3 OS=Homo sapiens GN=RABL3 PE=1 SV=1	7.00
sp|Q9HC62-2|SENP2_HUMAN	Isoform 2 of Sentrin-specific protease 2 OS=Homo sapiens GN=SENP2	7.00
sp|Q9Y679-2|AUP1_HUMAN	Isoform Short of Ancient ubiquitous protein 1 OS=Homo sapiens GN=AUP1	7.00
sp|Q8TEY7-2|UBP33_HUMAN	Isoform 2 of Ubiquitin carboxyl-terminal hydrolase 33 OS=Homo sapiens GN=USP33	7.00
sp|O00763-2|ACACB_HUMAN	Isoform Short of Acetyl-CoA carboxylase 2 OS=Homo sapiens GN=ACACB	7.00
sp|P53618|COPB_HUMAN	Coatomer subunit beta OS=Homo sapiens GN=COPB1 PE=1 SV=3	7.00
sp|Q14739|LBR_HUMAN	Lamin-B receptor OS=Homo sapiens GN=LBR PE=1 SV=2	6.99
sp|P05556|ITB1_HUMAN	Integrin beta-1 OS=Homo sapiens GN=ITGB1 PE=1 SV=2	6.99
sp|P98194-2|AT2C1_HUMAN	Isoform 2 of Calcium-transporting ATPase type 2C member 1 OS=Homo sapiens GN=ATP2C1	6.98
sp|Q86Y07|VRK2_HUMAN	Serine/threonine-protein kinase VRK2 OS=Homo sapiens GN=VRK2 PE=1 SV=3	6.98
sp|P27824-2|CALX_HUMAN	Isoform 2 of Calnexin OS=Homo sapiens GN=CANX	6.93
sp|O75864|PPR37_HUMAN	Protein phosphatase 1 regulatory subunit 37 OS=Homo sapiens GN=PPP1R37 PE=1 SV=4	6.00
sp|O15269|SPTC1_HUMAN	Serine palmitoyltransferase 1 OS=Homo sapiens GN=SPTLC1 PE=1 SV=1	6.00
sp|P31944|CASPE_HUMAN	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	6.00
sp|Q6PL24|TMED8_HUMAN	Protein TMED8 OS=Homo sapiens GN=TMED8 PE=1 SV=1	6.00
sp|Q9Y5P4-2|C43BP_HUMAN	Isoform 2 of Collagen type IV alpha-3-binding protein OS=Homo sapiens GN=COL4A3BP	6.00
sp|O14917|PCD17_HUMAN	Protocadherin-17 OS=Homo sapiens GN=PCDH17 PE=2 SV=2	6.00
sp|P51648-2|AL3A2_HUM	Isoform 2 of Fatty aldehyde dehydrogenase	5.99

AN	OS=Homo sapiens GN=ALDH3A2	
sp|P30519-2|HMOX2_HUMAN	Isoform 2 of Heme oxygenase 2 OS=Homo sapiens GN=HMOX2	5.99
sp|Q15323|K1H1_HUMAN	Keratin, type I cuticular Ha1 OS=Homo sapiens GN=KRT31 PE=2 SV=3	5.98
sp|O94830|DDHD2_HUMAN	Phospholipase DDHD2 OS=Homo sapiens GN=DDHD2 PE=1 SV=2	5.97
sp|O15258|RER1_HUMAN	Protein RER1 OS=Homo sapiens GN=RER1 PE=1 SV=1	5.97
sp|O95249|GOSR1_HUMAN	Golgi SNAP receptor complex member 1 OS=Homo sapiens GN=GOSR1 PE=1 SV=1	5.96
sp|P78381-2|S35A2_HUMAN	Isoform 2 of UDP-galactose translocator OS=Homo sapiens GN=SLC35A2	5.96
sp|P23396-2|RS3_HUMAN	Isoform 2 of 40S ribosomal protein S3 OS=Homo sapiens GN=RPS3	5.94
sp|P02751-10|FINC_HUMAN	Isoform 10 of Fibronectin OS=Homo sapiens GN=FN1	5.93
sp|Q9H9E3|COG4_HUMAN	Conserved oligomeric Golgi complex subunit 4 OS=Homo sapiens GN=COG4 PE=1 SV=3	5.93
sp|Q9NRW1|RAB6B_HUMAN	Ras-related protein Rab-6B OS=Homo sapiens GN=RAB6B PE=1 SV=1	5.93
sp|Q16658|FSCN1_HUMAN	Fascin OS=Homo sapiens GN=FSCN1 PE=1 SV=3	5.90
sp|Q9NP61-2|ARFG3_HUMAN	Isoform 2 of ADP-ribosylation factor GTPase-activating protein 3 OS=Homo sapiens GN=ARFGAP3	5.00
sp|P20700|LMNB1_HUMAN	Lamin-B1 OS=Homo sapiens GN=LMNB1 PE=1 SV=2	5.00
sp|O00264|PGRC1_HUMAN	Membrane-associated progesterone receptor component 1 OS=Homo sapiens GN=PGRMC1 PE=1 SV=3	5.00
sp|P62491-2|RB11A_HUMAN	Isoform 2 of Ras-related protein Rab-11A OS=Homo sapiens GN=RAB11A	5.00
sp|Q8TCT9-5|HM13_HUMAN	Isoform 5 of Minor histocompatibility antigen H13 OS=Homo sapiens GN=HM13	5.00
sp|P46977|STT3A_HUMA	Dolichyl-diphosphooligosaccharide--protein	5.00

N	glycosyltransferase subunit STT3A OS=Homo sapiens GN=STT3A PE=1 SV=2	
sp|P78386|KRT85_HUMAN	Keratin, type II cuticular Hb5 OS=Homo sapiens GN=KRT85 PE=1 SV=1	5.00
sp|P78310-7|CXAR_HUMAN	Isoform 7 of Coxsackievirus and adenovirus receptor OS=Homo sapiens GN=CXADR	5.00
sp|Q8NFA0|UBP32_HUMAN	Ubiquitin carboxyl-terminal hydrolase 32 OS=Homo sapiens GN=USP32 PE=1 SV=1	4.99
sp|Q13586|STIM1_HUMAN	Stromal interaction molecule 1 OS=Homo sapiens GN=STIM1 PE=1 SV=3	4.98
sp|Q9C0E8-2|LNP_HUMAN	Isoform 2 of Protein lunapark OS=Homo sapiens GN=LNP	4.98
sp|P10809|CH60_HUMAN	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2	4.97
sp|Q9HDC5|JPH1_HUMAN	Junctophilin-1 OS=Homo sapiens GN=JPH1 PE=1 SV=2	4.96
sp|P19634|SL9A1_HUMAN	Sodium/hydrogen exchanger 1 OS=Homo sapiens GN=SLC9A1 PE=1 SV=2	4.95
sp|Q9Y4P3|TBL2_HUMAN	Transducin beta-like protein 2 OS=Homo sapiens GN=TBL2 PE=1 SV=1	4.95
sp|Q8WUX9|CHMP7_HUMAN	Charged multivesicular body protein 7 OS=Homo sapiens GN=CHMP7 PE=1 SV=1	4.95
sp|P23526|SAHH_HUMAN	Adenosylhomocysteinase OS=Homo sapiens GN=AHCY PE=1 SV=4	4.00
sp|Q96AG4|LRC59_HUMAN	Leucine-rich repeat-containing protein 59 OS=Homo sapiens GN=LRRC59 PE=1 SV=1	4.00
sp|P27708|PYR1_HUMAN	CAD protein OS=Homo sapiens GN=CAD PE=1 SV=3	4.00
sp|P31151|S10A7_HUMAN	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4	4.00
sp|Q96JB2|COG3_HUMAN	Conserved oligomeric Golgi complex subunit 3 OS=Homo sapiens GN=COG3 PE=1 SV=3	4.00
sp|O43493-2|TGON2_HUMAN	Isoform TGN46 of Trans-Golgi network integral membrane protein 2 OS=Homo sapiens GN=TGOLN2	4.00
sp|O75915|PRAF3_HUM	PRA1 family protein 3 OS=Homo sapiens	4.00

AN	GN=ARL6IP5 PE=1 SV=1	
sp|P02786|TFR1_HUMAN	Transferrin receptor protein 1 OS=Homo sapiens GN=TFRC PE=1 SV=2	4.00
sp|P47756-2|CAPZB_HUMAN	Isoform 2 of F-actin-capping protein subunit beta OS=Homo sapiens GN=CAPZB	4.00
sp|Q9NRY5|F1142_HUMAN	Protein FAM114A2 OS=Homo sapiens GN=FAM114A2 PE=1 SV=4	4.00
sp|Q9NZJ5|E2AK3_HUMAN	Eukaryotic translation initiation factor 2-alpha kinase 3 OS=Homo sapiens GN=EIF2AK3 PE=1 SV=3	4.00
sp|P51572-2|BAP31_HUMAN	Isoform 2 of B-cell receptor-associated protein 31 OS=Homo sapiens GN=BCAP31	4.00
sp|P53367-2|ARFP1_HUMAN	Isoform A of Arfaptin-1 OS=Homo sapiens GN=ARFIP1	4.00

A2. List of TMEM115 unique interactors

Protein	Description	Spectral Count in BirA*-TMEM115 sample
sp|P49257|LMAN1_HUMAN	Protein ERGIC-53 OS=Homo sapiens GN=LMAN1 PE=1 SV=2	34.57
sp|P42566|EPS15_HUMAN	Epidermal growth factor receptor substrate 15 OS=Homo sapiens GN=EPS15 PE=1 SV=2	22.98
sp|Q13501|SQSTM_HUMAN	Sequestosome-1 OS=Homo sapiens GN=SQSTM1 PE=1 SV=1	20.97
sp|P05023-4|AT1A1_HUMAN	Isoform 4 of Sodium/potassium-transporting ATPase subunit alpha-1 OS=Homo sapiens GN=ATP1A1	17.84
sp|P50851-2|LRBA_HUMAN	Isoform 2 of Lipopolysaccharide-responsive and beige-like anchor protein OS=Homo sapiens GN=LRBA	17.69
sp|P08240|SRPR_HUMAN	Signal recognition particle receptor subunit alpha OS=Homo sapiens GN=SRPR PE=1 SV=2	14

sp|Q9H2J7|S6A15_HUMAN	Sodium-dependent neutral amino acid transporter B(0)AT2 OS=Homo sapiens GN=SLC6A15 PE=1 SV=1	13.98
sp|P05067-10|A4_HUMAN	Isoform APP639 of Amyloid beta A4 protein OS=Homo sapiens GN=APP	13.98
sp|O43399-5|TPD54_HUMAN	Isoform 5 of Tumor protein D54 OS=Homo sapiens GN=TPD52L2	12
sp|Q5VV42|CDKAL_HUMAN	Threonylcarbamoyladenosine tRNA methylthiotransferase OS=Homo sapiens GN=CDKAL1 PE=1 SV=1	12
sp|Q99460-2|PSMD1_HUMAN	Isoform 2 of 26S proteasome non-ATPase regulatory subunit 1 OS=Homo sapiens GN=PSMD1	12
sp|P51149|RAB7A_HUMAN	Ras-related protein Rab-7a OS=Homo sapiens GN=RAB7A PE=1 SV=1	11.93
sp|P53365-3|ARFP2_HUMAN	Isoform 3 of Arfaptin-2 OS=Homo sapiens GN=ARFIP2	11
sp|Q01968-2|OCRL_HUMAN	Isoform B of Inositol polyphosphate 5-phosphatase OCRL-1 OS=Homo sapiens GN=OCRL	10.89
sp|P17987|TCPA_HUMAN	T-complex protein 1 subunit alpha OS=Homo sapiens GN=TCP1 PE=1 SV=1	9.95
sp|O75694-2|NU155_HUMAN	Isoform 2 of Nuclear pore complex protein Nup155 OS=Homo sapiens GN=NUP155	9.94
sp|Q9Y2H6-2|FND3A_HUMAN	Isoform 2 of Fibronectin type-III domain-containing protein 3A OS=Homo sapiens GN=FNDC3A	9.00
sp|P08195-2|4F2_HUMAN	Isoform 2 of 4F2 cell-surface antigen heavy chain OS=Homo sapiens GN=SLC3A2	9.00
sp|O96005-4|CLPT1_HUMAN	Isoform 3 of Cleft lip and palate transmembrane protein 1 OS=Homo sapiens GN=CLPTM1	8.99
sp|O14828|SCAM3_HUMAN	Secretory carrier-associated membrane protein 3 OS=Homo sapiens GN=SCAMP3 PE=1 SV=3	8.91
sp|Q04721|NOTC2_HUM	Neurogenic locus notch homolog protein 2	8.00

AN	OS=Homo sapiens GN=NOTCH2 PE=1 SV=3	
sp|Q8NEN9|PDZD8_HUMAN	PDZ domain-containing protein 8 OS=Homo sapiens GN=PDZD8 PE=1 SV=1	8.00
sp|Q9H3N1|TMX1_HUMAN	Thioredoxin-related transmembrane protein 1 OS=Homo sapiens GN=TMX1 PE=1 SV=1	8.00
sp|Q9Y5Z9-2|UBIA1_HUMAN	Isoform 2 of UbiA prenyltransferase domain-containing protein 1 OS=Homo sapiens GN=UBIAD1	8.00
sp|O43396|TXNL1_HUMAN	Thioredoxin-like protein 1 OS=Homo sapiens GN=TXNL1 PE=1 SV=3	8.00
sp|Q9NYM9|BET1L_HUMAN	BET1-like protein OS=Homo sapiens GN=BET1L PE=1 SV=1	7.94
sp|P16615-2|AT2A2_HUMAN	Isoform 2 of Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 OS=Homo sapiens GN=ATP2A2	7.91
sp|P53367|ARFP1_HUMAN	Arfaptin-1 OS=Homo sapiens GN=ARFIP1 PE=1 SV=2	7.90
sp|Q5HYI8|RABL3_HUMAN	Rab-like protein 3 OS=Homo sapiens GN=RABL3 PE=1 SV=1	7.00
sp|Q9HC62-2|SENP2_HUMAN	Isoform 2 of Sentrin-specific protease 2 OS=Homo sapiens GN=SENP2	7.00
sp|Q9Y679-2|AUP1_HUMAN	Isoform Short of Ancient ubiquitous protein 1 OS=Homo sapiens GN=AUP1	7.00
sp|Q8TEY7-2|UBP33_HUMAN	Isoform 2 of Ubiquitin carboxyl-terminal hydrolase 33 OS=Homo sapiens GN=USP33	7.00
sp|O00763-2|ACACB_HUMAN	Isoform Short of Acetyl-CoA carboxylase 2 OS=Homo sapiens GN=ACACB	7.00
sp|Q14739|LBR_HUMAN	Lamin-B receptor OS=Homo sapiens GN=LBR PE=1 SV=2	6.99
sp|P05556|ITB1_HUMAN	Integrin beta-1 OS=Homo sapiens GN=ITGB1 PE=1 SV=2	6.99
sp|P98194-2|AT2C1_HUMAN	Isoform 2 of Calcium-transporting ATPase type 2C member 1 OS=Homo sapiens GN=ATP2C1	6.98
sp|O75864|PPR37_HUMAN	Protein phosphatase 1 regulatory subunit 37 OS=Homo sapiens GN=PPP1R37 PE=1 SV=4	6.00
sp|O15269|SPTC1_HUM	Serine palmitoyltransferase 1 OS=Homo	6.00

AN	sapiens GN=SPTLC1 PE=1 SV=1	
sp|P31944|CASPE_HUMAN	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2	6.00
sp|Q6PL24|TMED8_HUMAN	Protein TMED8 OS=Homo sapiens GN=TMED8 PE=1 SV=1	6.00
sp|Q9Y5P4-2|C43BP_HUMAN	Isoform 2 of Collagen type IV alpha-3-binding protein OS=Homo sapiens GN=COL4A3BP	6.00
sp|O14917|PCD17_HUMAN	Protocadherin-17 OS=Homo sapiens GN=PCDH17 PE=2 SV=2	6.00
sp|P30519-2|HMOX2_HUMAN	Isoform 2 of Heme oxygenase 2 OS=Homo sapiens GN=HMOX2	5.99
sp|Q15323|K1H1_HUMAN	Keratin, type I cuticular Ha1 OS=Homo sapiens GN=KRT31 PE=2 SV=3	5.98
sp|O94830|DDHD2_HUMAN	Phospholipase DDHD2 OS=Homo sapiens GN=DDHD2 PE=1 SV=2	5.97
sp|O15258|RER1_HUMAN	Protein RER1 OS=Homo sapiens GN=RER1 PE=1 SV=1	5.97
sp|P23396-2|RS3_HUMAN	Isoform 2 of 40S ribosomal protein S3 OS=Homo sapiens GN=RPS3	5.94
sp|P02751-10|FINC_HUMAN	Isoform 10 of Fibronectin OS=Homo sapiens GN=FN1	5.93
sp|Q9NRW1|RAB6B_HUMAN	Ras-related protein Rab-6B OS=Homo sapiens GN=RAB6B PE=1 SV=1	5.93
sp|Q16658|FSCN1_HUMAN	Fascin OS=Homo sapiens GN=FSCN1 PE=1 SV=3	5.93
sp|P20700|LMNB1_HUMAN	Lamin-B1 OS=Homo sapiens GN=LMNB1 PE=1 SV=2	5.00
sp|O00264|PGRC1_HUMAN	Membrane-associated progesterone receptor component 1 OS=Homo sapiens GN=PGRMC1 PE=1 SV=3	5.00
sp|P62491-2|RB11A_HUMAN	Isoform 2 of Ras-related protein Rab-11A OS=Homo sapiens GN=RAB11A	5.00
sp|Q8TCT9-5|HM13_HUMAN	Isoform 5 of Minor histocompatibility antigen H13 OS=Homo sapiens GN=HM13	5.00
sp|P46977|STT3A_HUMAN	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A OS=Homo	5.00

	sapiens GN=STT3A PE=1 SV=2	
sp|P78310-7|CXAR_HUMAN	Isoform 7 of Coxsackievirus and adenovirus receptor OS=Homo sapiens GN=CXADR	5.00
sp|Q8NFA0|UBP32_HUMAN	Ubiquitin carboxyl-terminal hydrolase 32 OS=Homo sapiens GN=USP32 PE=1 SV=1	4.99
sp|Q13586|STIM1_HUMAN	Stromal interaction molecule 1 OS=Homo sapiens GN=STIM1 PE=1 SV=3	4.98
sp|Q9C0E8-2|LNP_HUMAN	Isoform 2 of Protein lunapark OS=Homo sapiens GN=LNP	4.98
sp|P10809|CH60_HUMAN	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2	4.97
sp|Q9HDC5|JPH1_HUMAN	Junctophilin-1 OS=Homo sapiens GN=JPH1 PE=1 SV=2	4.96
sp|P19634|SL9A1_HUMAN	Sodium/hydrogen exchanger 1 OS=Homo sapiens GN=SLC9A1 PE=1 SV=2	4.95
sp|Q9Y4P3|TBL2_HUMAN	Transducin beta-like protein 2 OS=Homo sapiens GN=TBL2 PE=1 SV=1	4.95
sp|Q8WUX9|CHMP7_HUMAN	Charged multivesicular body protein 7 OS=Homo sapiens GN=CHMP7 PE=1 SV=1	4.95
sp|P23526|SAHH_HUMAN	Adenosylhomocysteinase OS=Homo sapiens GN=AHCY PE=1 SV=4	4.00
sp|P27708|PYR1_HUMAN	CAD protein OS=Homo sapiens GN=CAD PE=1 SV=3	4.00
sp|P31151|S10A7_HUMAN	Protein S100-A7 OS=Homo sapiens GN=S100A7 PE=1 SV=4	4.00
sp|Q96JB2|COG3_HUMAN	Conserved oligomeric Golgi complex subunit 3 OS=Homo sapiens GN=COG3 PE=1 SV=3	4.00
sp|O43493-2|TGON2_HUMAN	Isoform TGN46 of Trans-Golgi network integral membrane protein 2 OS=Homo sapiens GN=TGOLN2	4.00
sp|O75915|PRAF3_HUMAN	PRA1 family protein 3 OS=Homo sapiens GN=ARL6IP5 PE=1 SV=1	4.00
sp|P47756-2|CAPZB_HUMAN	Isoform 2 of F-actin-capping protein subunit beta OS=Homo sapiens GN=CAPZB	4.00
sp|Q9NRY5|F1142_HUMAN	Protein FAM114A2 OS=Homo sapiens GN=FAM114A2 PE=1 SV=4	4.00

sp|Q9NZJ5|E2AK3_HUMAN	Eukaryotic translation initiation factor 2-alpha kinase 3 OS=Homo sapiens GN=EIF2AK3 PE=1 SV=3	4.00
sp|P51572-2|BAP31_HUMAN	Isoform 2 of B-cell receptor-associated protein 31 OS=Homo sapiens GN=BCAP31	4.00
sp|P53367-2|ARFP1_HUMAN	Isoform A of Arfaptin-1 OS=Homo sapiens GN=ARFIP1	4.00