Charles University

Faculty of Science

Study programme: Biology

Branch of study: Immunology

Bc. Šárka Boháčová

Proximity proteome of intramembrane serine protease RHBDL4

Proximitní proteom intramembránové serinové proteázy RHBDL4

Diploma thesis

Supervisor: Ing. Kvido Stříšovský, Ph.D.

Prague, 2019

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne …….

……

Šárka Boháčová

Acknowledgements / Poděkování

Ráda bych poděkovala mému báječnému školiteli Kvidovi Stříšovskému, který mi byl během práce na mém diplomovém projektu vždy oporou. V nesnázích mi poskytl cenné rady, při úspěchu se mnou sdílel nadšení, naslouchal mým nápadům a neváhal mi je pomoci realizovat. Umožnil mi vskutku naplno se věnovat kvalitní vědě. Rovněž během sepisování této práce mi neváhal pomoci s korekturami, za což mu náleží veliký dík.

Dále bych chtěla poděkovat všem svým kolegům a kolegyním z naší laboratoře za podporu, inspiraci a skvělou přátelskou atmosféru, která u nás vždy panuje. Obzvláště bych pak chtěla poděkovat Petře Rampírové za výpomoc s klonováním a Blance Collis za pasážování buněk v mé nepřítomnosti a za jazykové korektury.

Many thanks belong to our alumnus postdoc Aymeric Morlé. Although he had left our group before my diploma project started, he actually taught me how to work in a lab during the preceding two years. Thanks to his tolerant leadership I developed from a naïve student into an aspiring researcher and mastered many techniques. I am truly sorry for all the ruined experiments from the early days!

Poděkování náleží skupině hmotnostní spektrometrie Josefa Cvačky na ÚOCHB, která se významně podílela na měření a vyhodnocení hmotnostních dat. Obzvláště děkuji Michalovi Koreckému, Martinu Hubálkovi a Aleně Křenkové. Velmi cenné konzultace ohledně hmotnostní spektrometrie a metody triple-SILAC poskytl Marek Vrbacký z Fyziologického ústavu AV ČR.

V neposlední řadě bych ráda poděkovala své rodině, blízkým a přátelům, kteří mne po celou dobu mého studia podporovali a poskytli mi potřebné zázemí, ač jak často říkávají, nemají vlastně moc tušení, co že to dělám.

Abstract

Regulated intramembrane proteolysis is an interesting process involved in a multitude of cellular pathways. Enzymes which catalyse this are termed intramembrane proteases (IMPRs), cleaving proteins passing through the membrane within their transmembrane domain. Rhomboid proteases are serine IMPRs. They are widely distributed among organisms and evolutionarily conserved, but despite many efforts, their physiological roles are largely unexplored. RHBDL4 is a mammalian rhomboid protease localised to the endoplasmic reticulum. It is involved in the development of colorectal cancer, which makes it an important focus of research, but its physiological function is not well understood. In order to explore it, I established and employed a proximity proteomics approach, termed APEX2. It is based on biotinylation of proteins in the spatial proximity of the target in the physiological environment of intact living cells. Labelled proteins are subsequently purified, identified and quantified by mass spectrometry. Exploring the physiological vicinity of RHBDL4, its interaction partners and substrates can be revealed and the detailed subcellular compartment, where RHBDL4 resides, can thus be inferred. During three independent experiments in HCT116 cell line, three proteins emerged repeatedly in the RHBDL4 proximity proteome: ANXA7, ITGB1 and PLIN3, indicating that RHBDL4 may be involved in intracellular trafficking and lipid metabolism.

Key words: intramembrane proteolysis, rhomboid protease, RHBDL4, endoplasmic reticulum, proximity proteomics, APEX2

Abstrakt

Intramembránová proteolýza je součástí celé řady buněčných drah a procesů. Je katalyzována specializovanými enzymy – intramembránovými proteázami (IMPR), které svůj substrát štěpí v rámci jeho transmembránové domény pod povrchem membrány. Rhomboidní proteázy jsou serinové IMPR. Vyskytují se napříč celým spektrem živých organismů a co do transmembránové topologie a sekvenčních motivů jsou evolučně konzervovány. Avšak navzdory dlouhodobé snaze, fyziologickou úlohu většiny rhomboidních proteáz doposud neznáme. RHBDL4 je savčí rhomboidní proteáza nacházející se v endoplasmatickém retikulu. Podílí se na vzniku kolorektálního karcinomu, což podtrhuje význam jejího studia. Fyziologická úloha RHBDL4 je však převážně neobjasněna. Za účelem jejího hlubšího prozkoumání jsem zavedla a následně aplikovala metodu proximitní proteomiky označovanou APEX2. Je založena na značení proteinů v blízkosti vybraného cíle biotinem, a to ve fyziologickém prostředí živých neporušených buněk. Značené proteiny jsou následně izolovány a detekovány pomocí hmotnostní spektrometrie. Analýzou proximitního proteomu proteázy RHBDL4 je možné odhalit její interakční partnery, substráty či lépe definovat konkrétní podoblast ER, kde se vyskytuje. V rámci tří nezávislých experimentů v buněčné linii HCT116 se tři proteiny opakovaně nacházely v blízkosti RHBDL4: ANXA7, ITGB1 a PLIN3, což naznačuje potenciální úlohu RHBDL4 v buněčném transportu a metabolismu lipidů.

Klíčová slova: intramembránová proteolýza, rhomboidní proteáza, RHBDL4, endoplasmatické retikulum, proximitní proteomika, APEX2

Table of contents

Abbreviations

ADAM a disintegrin and metalloproteinase
APP amyloid precursor protein
ATGL adipose triglyceride lipase
ANXA7 annexin A7
ßA beta amyloid
BFA Brefeldin-A
cDC conventional dendritic cell
COG conserved oligomeric Golgi tethering complex
CRC colorectal cancer
CREB3 cAMP-response element binding 3
DC dendritic cell
DC-STAMP dendritic cell specific transmembrane protein
ECM extra cellular matrix
EGF epidermal growth factor
EGFR epidermal growth factor receptor
EL eluate
EMT epithelial-to-mesenchymal transition
ER endoplasmic reticulum
ERAD ER-associated degradation
ERES ER-exit sites
FBS fetal bovine serum
FT flow-through
eGFP enhanced green fluorescent protein
Golgi Golgi apparatus
H heavy (labelled by heavy amino acids)
HA hemagglutinin tag
HCV hepatitis C virus
HRP horseradish peroxidase
IL interleukin
IFN interferon
IFN-I type I interferon
IMPR intramembrane protease
IRS1 insulin receptor substrate 1
ITGB1 integrin beta 1
KO knock-out
L light (labelled by light amino acids)
M medium (labelled by medium amino acids)
mDC myeloid dendritic cell
MHC major histocompatibility complex
MOI multiplicity of infection
MS mass spectrometry
NES nuclear export sequence
NK natural killer cell
NTF N-terminal fragment
PAA polyacrylamide
pDC plasmacytoid dendritic cell
PES polyethersulfone
PGE_2 prostaglandin E2
PLIN3 perilipin 3
PSM peptide spectrum matches
S1P site-1 protease
S2P site-2 protease
SA S144A, serine 144 mutated into alanine
SDS sodium dodecyl sulfate
SILAC stable isotope labelling with amino acids in cell culture
SPP signal peptide peptidase
SREBP sterol regulatory element binding protein
TACE TNFα-converting enzyme
TGCT testicular germ cell tumour
TGFα tumour growth factor α
TMH transmembrane helix
TNF tumour necrosis factor
UIM ubiquitin interacting motif
UPR unfolded protein response
VBM VCP (p97) binding motif
WB western blot/blotting
WT wild-type

1 Introduction and aims

RHBDL4 is an intramembrane serine protease. Intramembrane proteases catalyse a remarkable reaction in which they cleave within the transmembrane domains of their substrates underneath the surface of lipid bilayer. As a result of the cleavage, the transmembrane domain can be liberated from the membrane and fulfil various cellular roles, or it can be degraded.

RHBDL4 belongs to so-called 'rhomboid' superfamily (see Chapter 4 for the the history of the name), which comprises not only proteases but also a number of proteolytically inactive members. Rhomboids are evolutionarily widespread and their typical rhomboid fold is conserved, indicating their importance. However, for most of the rhomboids, their physiological functions are unclear. Interestingly, the catalytically inactive rhomboids are much more explored than the rhomboid proteases. It is supposed that inactive rhomboids developed from the active ones by mutation of the catalytic residues, but maybe the rhomboid fold itself is the most important attribute of the superfamily as indicated by a recent study showing that multiple rhomboid-like proteins including proteases and pseudoproteases can overcome the viscosity of membrane and diffuse incredibly fast (Kreutzberger et al. 2019).

RHBDL4 is the most studied mammalian rhomboid protease and it is localised in the endoplasmic reticulum (ER). It is involved in the development of colorectal carcinoma and it is implicated also in several other cancer types. However, the underlying molecular mechanisms are controversial as well as the physiological role of RHBDL4. Although there are several indications including involvement in ER-associated degradation (ERAD), regulation of intracellular trafficking and secretion, or regulation of apoptosis and proliferation, a consensus has not been reached yet. Therefore, we decided to employ a proximity proteomics to explore RHBDL4 in a completely new way. The proximity proteomics are currently getting popular for study of membrane proteins and membrane compartments, as they combine the physiological environment of living intact cell with sensitive detection and quantification by mass spectrometry.

In this thesis I aim to explore the proximity proteome of RHBDL4. In particular, I will:

- Test BioID2 and APEX2 proximity proteomics approaches
- Implement the proximity proteomics for study of RHBDL4
- Analyse the RHBDL4 proximity proteome in HCT116 cell line

In the first part of the thesis I will present an overview of the current state of knowledge. I will introduce the concept of intramembrane proteolysis and focus on the immune context of this process. Then I will describe the rhomboid superfamily proteins focusing on RHBDL4 protease and the related catalytically inactive rhomboids. Finally, I will briefly introduce the proximity proteomics method and sum up the rationale of this study.

2 Intramembrane proteases

Intramembrane proteases (IMPRs) are an important and interesting group of enzymes which were first described at the turn of the 21st century (Wolfe & Kopan 2004; Brown et al. 2000). As enzymes, they face many challenges due to the specific environment in which they operate. These hydrophobic polytopic transmembrane proteins have an active site embedded under the surface of the lipid bilayer. Therefore, unlike soluble proteases immersed in an aqueous environment, IMPRs need to be supplied with water molecules, which are necessary for peptide bond hydrolysis, because water molecules are normally rare within the lipid bilayer. This issue was clarified by crystallographic structures of representative IMPRs (Li et al. 2012; Wang et al. 2006; Feng et al. 2007; Manolaridis et al. 2013) showing that the catalytic site of IMPRs is located in a cavity buried in the membrane but open to the aqueous environment.

Another challenge resides in the nature of IMPR substrates (Wolfe 2009; Ha 2009). The substrates are also transmembrane proteins and their cleavage site is usually localised within their alpha helix, which anchors them inside lipid membranes. The alpha helix is a stable secondary structure supported by hydrogen bonds, which is normally inaccessible to proteolytic cleavage. Therefore, in order

to cleave them, IMPRs have to destabilise and/or unwind the alpha helix of their substrates.

The biochemical mechanism of the hydrolysis itself is similar to that of soluble proteases. Despite the lack of structural or sequence similarity between soluble and intramembrane proteases (Wolfe & Kopan 2004), the catalytic residues are analogous. Amongst IMPRs, we find metalloproteases with a zinc ion coordinated by the typical sequence motif HExxH, serine proteases with a catalytic S and H dyad, aspartyl proteases and glutamyl proteases (the soluble form of which is found for example in fungi (Fujinaga et al. 2004)). This is a typical example of convergent evolution.

IMPRs have numerous important biological functions. Firstly, IMPRs are integral to degradation of transmembrane proteins. Secondly, intramembrane cleavage liberates protein domains or peptides which act as transcription factors, growth factors or ligands and influence a multitude of cellular processes. They are involved in, for example, the unfolded protein response (UPR), lipid metabolism, cell differentiation, proliferation, immunity and many other processes. Finally, IMPRs are evolutionarily extremely widespread and can be found in organisms ranging from bacteria to humans, highlighting their biological significance.

2.1 Metallo-IMPRs

The site-2 protease (S2P) was the first IMPR to be identified, thus introducing the concept of intramembrane proteolysis (Brown et al. 2000). This metalloprotease is essential for lipid homeostasis (Rawson et al. 1997) and the UPR (Ye et al. 2000) in human cells. Its catalytic site is defined by the conserved HExxH sequence and a more distant aspartate which together coordinate a zinc ion (Fig.1) (Zelenski et al. 1999; Rawson et al. 1997; Feng et al. 2007). This zinc ion activates water molecules for nucleophilic attack of peptide bonds.

In bacteria, the S2P family of proteins mediates the response to extra-cytoplasmic stress. For example SpoIVFB protease from *B. subtilis* releases special sigma factor kappa σ^{K}, which regulates sporulation in stress conditions (Rudner et al. 1999). *E. coli* possesses a related YaeL protease (also known as RasP), which cleaves RseA factor, thus liberating σ^{E}, which is again important during stress

response (Akiyama et al. 2004). In addition, bacterial S2Ps mediate cleavage and degradation of signal peptides that had been previously removed from pre-proteins by signal peptidases (Saito et al. 2011).

In animals, S2P is a component of the SREBP (sterol regulatory element-binding protein) pathway – a feedback mechanism indispensable for maintaining cholesterol homeostasis (Brown & Goldstein 1997). SREBP is retained in the ER by sterol sensing factors if cholesterol levels in the cell are high. At lower cholesterol concentrations, SREBP is released from the ER and moves to the Golgi apparatus (Golgi). In the Golgi, site-1 protease (S1P) cleaves SREBP first, preparing it as a substrate for S2P. The second intramembrane cleavage by S2P then liberates the N-terminal transcription factor domain from SREBP. This domain regulates expression of genes involved in sterol metabolism.

The same principle operates with another membrane-anchored transcription factor, ATF6, a factor from the UPR. ATF6 is processed analogously to SREBP modification by S1P and S2P in mammals (Ye et al. 2000; Haze et al. 1999). The main difference is that ATF6 does not interact with sterol sensing proteins and its availability for proteolysis is regulated by the unmasking of a Golgi targeting sequence. This occurs due to the dissociation of the BiP chaperone from the targeting sequence upon ER stress (Shen et al. 2002).

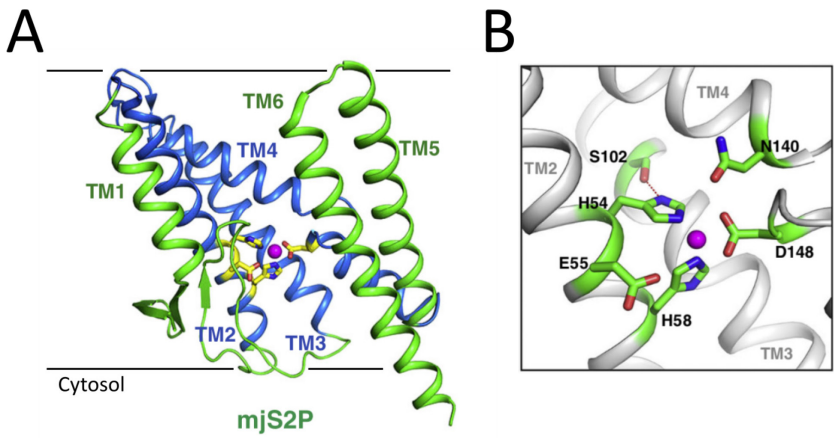

Fig.1: Structure of site-2 protease (adopted and adjusted from Sun et al. 2016).
A. Cartoon representation of site-2 protease (S2P) from *Methanocaldococcus jannaschii*. The blue helices form a stable core whereas the green helices are flexible and involved in gating of the substrate. The active site is linked to the cytosol via an aqueous channel. **B.** The active site of S2P in detail. The zinc ion, represented by a purple sphere, is coordinated by H54, E55, H55 (of the HExxH conserved motif) and by D148.

2.2 Aspartyl IMPRs

Aspartyl IMPRs were the second class to be discovered. Their active site contains two catalytic aspartates, which activate a water molecule for subsequent nucleophilic attack on the peptide bond. A subgroup of aspartyl IMPRs are the presenilins (Selkoe & Wolfe 2007; Hass et al. 2009). Presenilin, unlike other IMPRs, is active in a protein complex composed of several accessory factors (Fig.2). The entire complex is known as gamma-secretase. Gamma-secretase was studied intensively because it products beta amyloid (ßA) peptides from the amyloid precursor protein (APP), that are aggregation-prone and contribute to the formation of amyloid plaques in the brain that are associated with Alzheimer disease (Masters et al. 1985).

The second subgroup consists of signal peptide peptidase (SPP) and SPP-like proteins. SPP hydrolyses signal peptides, which had been cleaved off by signal peptidase from proteins directed into the secretory pathway through the ER (Weihofen et al. 2002). The role of SPP is not limited to removal of signal peptides. The resulting products of SPP-mediated cleavage are peptides important for the function of the immune system, which will be discussed in Chapter 3. Besides SPP, there are four SPP-like proteases encoded in the human genome: SPPL2a, SPPL2b, SPPL2c and SPPL3 (Voss et al. 2013).

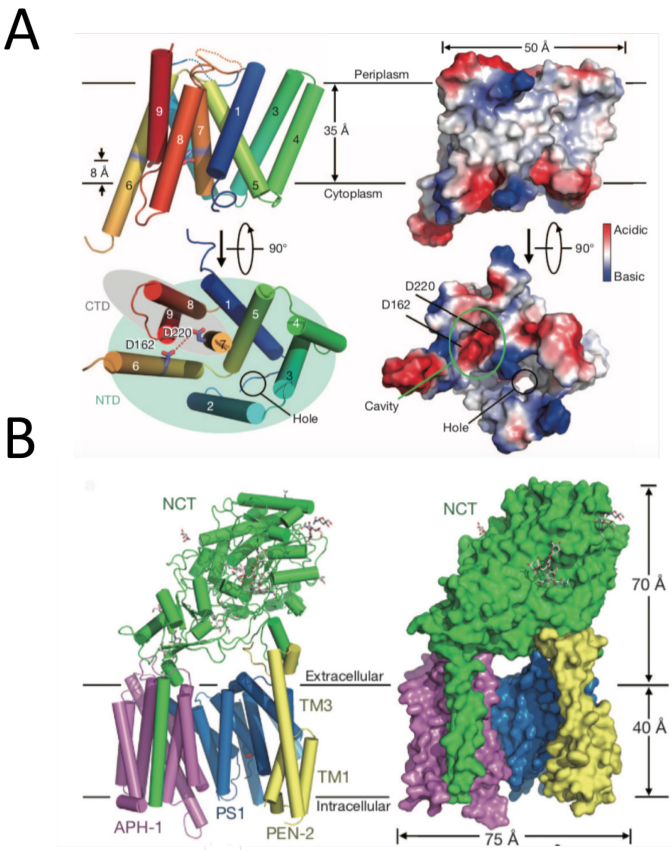

Fig.2: Structure of presenilin and gamma-secretase (adopted from Li et al. 2012 and Bai et al. 2015)
A. X-ray structure of presenilin from *Methanoculleus marisnigri* represented as a cartoon or a molecular surface. The membrane is represented by horizontal lines. The catalytic cavity is open to the cytosol. Catalytic residues D162 and D220 are at the base of the cavity approximately 8 Å beneath the surface of the membrane. In addition, there is a hydrophobic hole within the enzyme. **B.** CryoEM structure of the whole gamma secretase composed of four proteins: presenilin (PS1), PEN-2, APH-1 and nicastrin (NCT). Presenilin is the catalytic subunit of the gamma-secretase complex.

2.3 Serine IMPRs

Rhomboids are serine IMPRs (Urban et al. 2001), which catalyse protein hydrolysis via a catalytic dyad formed by serine and histidine (Lemberg et al. 2005). Rhomboid proteases do not need any accessory factors for activity and do not require substrate pre-cleavage, unlike most other IMPRs. Their substrates are full-length proteins (Lemberg et al. 2005). Rhomboid proteases were the first IMPRs whose three-dimensional structure was solved. The structure of bacterial rhomboid protease GlpG (Fig.3) (Wang et al. 2006; Wu et al. 2006) explains some mechanistic features of rhomboid intramembrane cleavage. In particular, entry of water to the catalytic site was shown to be possible, because the catalytic residues are embedded in an aqueous cavity beneath the membrane surface. However, the biological functions of rhomboids are largely unclear, except for the *Drosophila melanogaster* Rhomboid-1, which is the first discovered rhomboid, and whose developmental function is relatively well established (Lee et al. 2001). This rhomboid is the protease which cleaves Spitz (a homolog of epidermal growth factor (EGF)). Rhomboid cleavage enables secretion of Spitz and therefore growth-promoting signalling in adjacent cells (Fig.3). This function of rhomboid is apparently not conserved in mammals, where cleavage of EGF receptor (EGFR) ligands is performed predominantly by a disintegrin and metalloproteinase 17 /

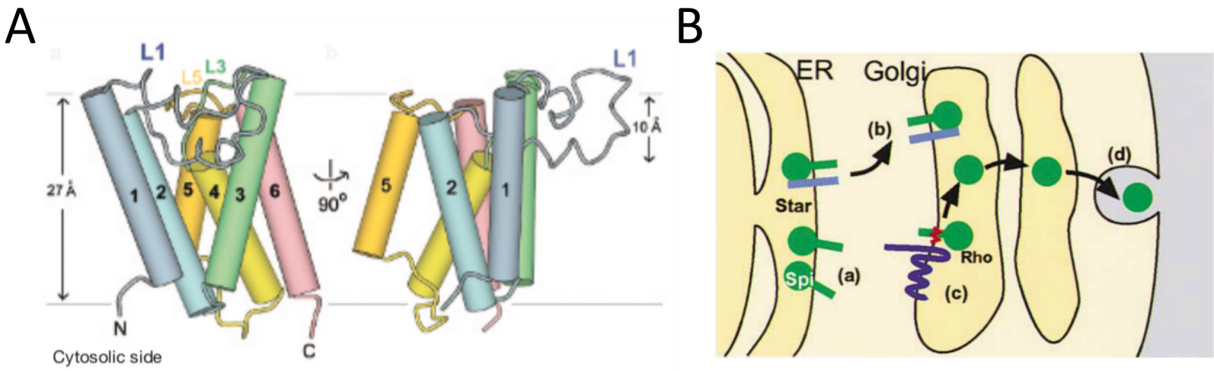

Fig.3: Rhomboid proteases structure and function
A. Structure of the 6 TMH core of rhomboid protease GlpG, front view and rotated by 90° (adopted from Wang et al. 2006). The width of the cellular membrane is indicated by the span between the two horizontal grey lines. The transmembrane helices are numbered in ascending order from the N-terminus. In the front view the catalytic cavity is visible, helix 4 forms its base. The cavity is open to the extracellular side and it is approximately 10 Å deep. At the top of helix 4 a catalytic serine is present, the catalytic histidine is within helix 6. L = loop **B.** Illustration of Rhomboid-1 function in *Drosophila* (adopted from Lee et al. 2001). Rho – Rhomboid1; Spi – Spitz, EGFR ligand; Star – factor inducing translocation of Spitz into Golgi.

TNFα-converting enzyme (ADAM17/TACE), referred to as shedding (i.e. a cleavage just above the membrane) (Lichtenthaler et al. 2018).

In eukaryotic cells, rhomboid proteases are distributed in various cell membranes. In mammalian cells there are five rhomboid proteases present – RHBDL1, RHBDL2, RHBDL3 and RHBDL4 in the secretory pathway (Lastun et al. 2016) and PARL in mitochondria (Spinazzi & De Strooper 2016). The rhomboid proteases in the secretory pathway (also termed secretases) are divided into A and B subgroups based on the number of transmembrane helices (TMHs) determined by topology prediction (Lemberg & Freeman 2007). Secretases B (RHBDL4) have only a 6 TMH structure, which is the conserved core among rhomboids. Secretases A (RHBDL1, RHBDL2, RHBDL3) have an additional C-terminal TMH, therefore seven TMHs in total. PARL-type rhomboids have an additional N-terminal TMH and inverted topology. RHBDL2, RHBDL4 and PARL are the most extensively studied mammalian rhomboid proteases, but their cellular functions are still not entirely clear. In Chapter 4.3, I will focus in depth on RHBDL4, as it is the principal topic of this thesis.

Interestingly, rhomboid proteases belong to a wider superfamily of rhomboid-like proteins, which includes several types of proteins similar to rhomboid proteases in architecture but devoid of proteolytic activity, such as iRhoms or Derlins (Dulloo et al. 2019; Tichá et al. 2018). Although proteolytically inactive, they play a key role in signalling, and membrane protein proteostasis. The rhomboid superfamily is described in detail in Chapter 4.

2.4 Glutamyl IMPRs

This group of IMPRs was discovered most recently and is exemplified by Rce1 protein and its homologs in all domains of life. Rce1 was first described in yeast as a converting enzyme for Ras and a-factor (Boyartchuk et al. 1997). Ras and a-factor are prenylated and as a final modification, the three C-terminal amino acids beyond the lipid anchor have to be cleaved off. However, it was not until 2013 that Rce1 was shown to be a glutamyl IMPR (Manolaridis et al. 2013).

The crystallographic structure of Rce1 from *Methanococcus maripaludis* revealed that this protease is an IMPR, because it has a catalytic cavity, formed by transmembrane alpha helices, embedded into the lipid membrane (Fig.4) (Manolaridis et al. 2013). The catalytic residues, glutamate and histidine, are inside the cavity, near its base. The cavity allows solvent access from the cytosol and it has a 'back door' accessible from within the membrane, which is presumably used by its substrate's prenyl anchor. The peptide chain of the substrate is inserted into the cavity in the shape of a hairpin. The catalytic mechanism of cleavage is an acid-base reaction.

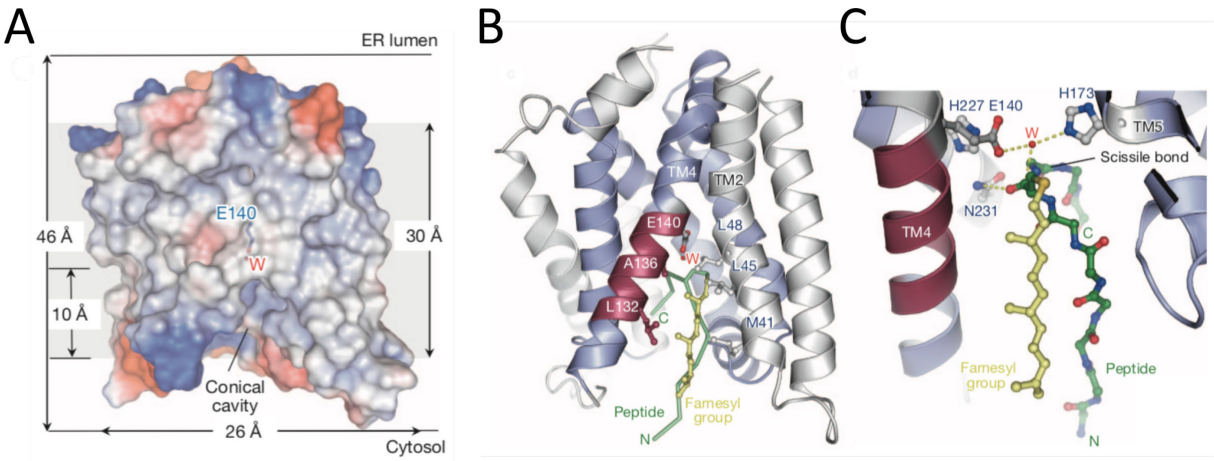

Fig.4: Structure of Rce1 (adopted and adjusted from Manolaridis et al. 2013).
A. Molecular surface model of Rce1. The red and blue regions mark negatively and positively charged amino acids, the white regions are uncharged. The catalytic cavity is open to the cytosol, the catalytic water (W) and glutamate (E140) inside the cavity are indicated. The depth of the cavity is approximately 10 Å. The membrane is represented by a grey rectangle.
B. and C. Cartoon representation of the catalytic cavity of Rce1 with farnesylated peptide modelled in it. The catalytic dyad E140 and H173 with the catalytic water molecule are indicated.

In human cells, Rce1 hydrolyses all farnesylated and geranylgeranylated proteins with a CAAX motif (where cysteine is prenylated and cleavage occurs right after it) (Otto et al. 1999). Although Rce1 has many natural substrates in the cell, only some are affected by its deficiency, suggesting that there may be some functional redundancy with other enzymes. Ras signalling is disrupted upon Rce1 inactivation (Kim et al. 1999; Burrows et al. 2009). This results in slower cell growth and limited Ras-induced transformation (Bergo et al. 2002). In contrast, one study reports accelerated K-Ras induced myeloproliferative disease as a result of Rce1 removal (Wahlstrom et al. 2007). Apart from effects on Ras, cardiomyopathy (Bergo et al. 2004), dysfunction of photoreceptor cells

(Christiansen et al. 2011) and disrupted integrity and deformation of nuclear lamina (Maske et al. 2003) were reported in Rce1 deficiency studies.

3 Intramembrane proteolysis in immunity

IMPRs are linked to the function of the immune system in many ways. The most prominent and well understood examples will be discussed, including SPP, SPP-like proteins and S2P. Moreover, a group of inactive IMPRs –rhomboid-like pseudoproteases, are particularly relevant for the immune system and will be discussed too.

3.1 SPPL2a

SPPL2a is an aspartyl IMPR which resides in a late endosomal/lysosomal cellular compartment (Friedmann et al. 2006). It has been implicated in processing of N-terminal fragments (NTF) of several membrane proteins after shedding of their soluble ectodomains by other proteases: tumour necrosis factor (TNF) (Friedmann et al. 2006; Fluhrer et al. 2006), FasL (Kirkin et al. 2007), Bri2 (Martin et al. 2008) and CD74 (Schneppenheim et al. 2013; Beisner et al. 2013; Schneppenheim et al. 2014). In some studies, the activity of SPPL2a was found to be redundant in relation to SPPL2b (Friedmann et al. 2006; Fluhrer et al. 2006; Martin et al. 2008). Intramembrane cleavage of NTFs can result in the release of an intracellular signalling peptide or in the degradation of the membrane-bound NTF. To date, only the NTF of CD74 cleavage was confirmed as a substrate of SPPL2a *in vivo* and is supported by a mouse phenotype study (Schneppenheim et al. 2013; Beisner et al. 2013).

CD74, also known as the invariant chain, protects major histocompatibility complex (MHC) class II molecules from premature peptide binding in the ER and guides MHC-II into a late endosomal compartment, also known as the MHC-II-compartment, where peptide binding to the groove of MHC-II occurs. As CD74 occupies the peptide-binding groove of MHC-II, it must be removed and exchanged for the peptide. Its removal begins with cleavage by lysosomal proteases, such as Cathepsin S. The membrane embedded NTF of CD74 cannot be degraded by cathepsins and needs to be removed by the specialized intramembrane protease SPPL2a (Schneppenheim et al. 2013; Beisner et al. 2013).

If SPPL2a is inactive or absent, the NTF of CD74 accumulates in the endosomal compartment together with MHC-II and causes several defects in mice. Transitional T2 B cells and mature B cells are significantly affected as their numbers are reduced. As a result, immunoglobulin levels are strongly decreased at steady state as well as during specific antigen response (Schneppenheim et al. 2013; Beisner et al. 2013). The numbers of myeloid dendritic cells (mDCs) are also significantly decreased, whereas plasmacytoid DCs (pDCs) are affected much less (Beisner et al. 2013). A significant rescue of this complex phenotype is possible by additional knock-out (KO) of CD74, showing that the phenotype is not caused by signalling through a CD74 fragment, but it rather results from toxic accumulation of CD74 NTF (Schneppenheim et al. 2013; Beisner et al. 2013).

The question arises as to why only mDCs, T2 and mature B cells are affected. Expression of SPPL2a is comparable among developmental stages of B cells and between pDCs and mDCs. But CD74 and MHC-II expression is high specifically in mDCs and their increase in B cells occurs during the T1 to T2 developmental transition. Therefore, it is likely that SPPL2a is specifically important in cells with high MHC-II turnover and it enables efficient antigen presentation in these cells (Beisner et al. 2013). In addition to this, in B cells, CD74 NTF accumulation leads to the abrogation of B-cell activating factor receptor (BAFFR) and B-cell receptor (BCR) signalling, both of which are important signalling pathways for B cell maturation (Schneppenheim et al. 2013). The role of SPPL2a in the degradation of NTF of CD74 was also confirmed in humans, although it is not yet clear if the B cell and DC phenotype is the same as in the mouse (Schneppenheim et al. 2014).

3.2 SPPL3

SPPL3 is another aspartyl IMPR, localised in the ER and Golgi (Friedmann et al. 2006). SPPL3 deficient mice have various phenotypes, depending on the genetic background of the mouse strain used. In the C57Bl/6J background, deficiency of SPPL3 is lethal very early after birth (Hamblet et al. 2016). This strong phenotype is reduced in a mixed C57Bl/6J:129S5 mouse, where viable

offspring suffer from hematological abnormalities, growth retardation and male sterility (Voss et al. 2014).

In the context of immunity, SPPL3 has been found to be important for natural killer (NK) cell maturation (Hamblet et al. 2016). In the mouse model with SPPL3 deficiency specific for the hematopoietic system, the levels of CD11b+ NK cells were significantly reduced, as was the ability of NK cells to clear MHC-I-negative tumour cells. However, despite an intensive effort, the causative mechanism was not revealed. It is clear that the defect is intrinsic for NK cells and dependent on SPPL3 proteolytic activity. The decreased proliferation of CD11b- NK cell precursors, coupled with the improved survival of maturing CD11b+ NK cells results in a shift in the CD11b+ to CD11b- ratio, also causing the defect in cytotoxic function on population level. However, the responsible substrate(s) of SPPL3 whose cleavage would give rise to this phenotype have not been identified.

As a protease, SPPL3 has been implicated in the cleavage of several glycosyl transferases and other Golgi factors involved in N- and O- glycosylation (Voss et al. 2014; Kuhn et al. 2015). SPPL3 deficiency results in a hyperglycosylation phenotype whereas SPPL3 overexpression causes hypoglycosylation, dependently on the catalytic activity of SPPL3. The glycosyl transferase MGAT5, could be a direct substrate of SPPL3, whose cleavage causes this phenotype (Voss et al. 2014). Surprisingly, in NK cells deficient in SPPL3 no hyperglycosylation of surface proteins was observed even though the MGAT5 protein level was high (Hamblet et al. 2016). Therefore, the function of SPPL3 in NK cell maturation is probably independent of regulation of glycosylation.

A further link between SPPL3 and the immune system exists, but is independent of SPPL3 protease activity (Makowski et al. 2015). SPPL3 directly interacts with the ER protein STIM1, which is a sensor for Ca^{2+} level in ER. When Ca^{2+} is low, STIM1 binds to the Ora1 channel on the plasma membrane (PM) and allows Ca^{2+} replenishment from the extracellular space. SPPL3 supports effective interaction of STIM1 with Ora1 and therefore maintains potent Ca^{2+} signalling, which is necessary during activation of immune cells. Depletion of SPPL3 significantly decreases NFAT activation during T cells activation through its T cell receptor and CD28 (Makowski et al. 2015). This can be rescued by overexpression

of wild-type or catalytically inactive SPPL3, representing a clear protease-independent function of SPPL3.

3.3 SPP

SPP is an aspartyl IMPR which cleaves signal peptides remaining in the membrane of the endoplasmic reticulum (Weihofen et al. 2002). Interestingly, cleavage products of the signal peptides of MHC-I molecules are important for a proper function of immune system. The N-terminal part of the MHC-I signal peptide is released by SPP into the cytosol, returned by transporter associated with antigen processing (TAP) into the ER, trimmed and subsequently loaded onto HLA-E, a non-classical and non-polymorphic MHC-I molecule. This loading is essential for HLA-E export to the plasma membrane (Lemberg et al. 2001). HLA-E molecules are recognized by CD94/NKG2A inhibitory receptors of NK cells (Braud et al. 1998). A conserved motif in the signal peptides of MHC-I molecules is a marker of cells expressing MHC-I. Therefore, SPP is vital for the protection of healthy cells against NK cells and immune surveillance.

SPP is also known to be essential for maturation of core protein from hepatitis C virus (HCV) (McLauchlan et al. 2002). HCV has a single-stranded sense RNA genome, which is translated into a long polyprotein. To divide it into individual proteins, the polyprotein is cleaved by proteases of cellular and viral origin. The core protein and envelope glycoprotein 1 (E1) are separated by a signal peptide, which directs E1 into the lumen of the ER. Signal peptidase cuts off E1, but the core protein remains attached to the ER membrane by the signal sequence. For it to be released, SPP has to cleave the signal peptide (Pène et al. 2009; McLauchlan et al. 2002). The core protein then associates with lipid droplets and assembles into new virions (Targett-Adams et al. 2008; McLauchlan et al. 2002).

This is one example of an interaction between IMPRs and viruses. However, cellular IMPRs are not only exploited by viruses, but also involved in virus defence mechanisms. This topic, reviewed in (Ye 2013), is related to the immune system, but not described in detail here, as it is not closely relevant to the work covered in this thesis.

3.4 S2P and CREB3 transcription factors

CREB3 proteins (cAMP-response element binding 3) are assumed to be substrates of regulated intramembrane proteolysis catalysed by S1P and S2P, due to their similarity to ATF6 (Bailey & O'Hare 2007). They possess a transcription factor domain at their N-terminus which faces the cytosol, next to a transmembrane domain and a C-terminal domain in the ER lumen. Their cleavage depends on the Golgi associated S1P, meaning that they have to be transported from the ER to the Golgi to be activated. Direct evidence for the role of S2P in the liberation of the N-terminal domain of CREB3 is missing (Raggo et al. 2002), but the involvement of S2P in CREB3 family activation is broadly accepted (Chan et al. 2011). Individual members of the CREB3 protein family can be transported from the ER to the Golgi in response to different types of stimuli and the cellular responses which follow transcription factor domain binding to DNA may also vary. Their functions are probably cell type specific (Chan et al. 2011).

3.4.1 CREB3 (LUMAN) and DC maturation

CREB3 is an interaction partner of DC-STAMP (dendritic cell specific transmembrane protein) (Eleveld-Trancikova et al. 2010). Both of these proteins are upregulated in immature dendritic cells. Intramembrane cleavage of CREB3 resulting in its nuclear accumulation was observed upon Brefeldin-A (BFA) treatment and during DC maturation (BFA induces dissociation of the Golgi and merging of the Golgi membranes with the ER, thus giving S1P access to ER proteins). DC-STAMP negatively regulates intramembrane cleavage of CREB3, probably by keeping CREB3 in the ER. Together, this suggests an important role for CREB3 and DC-STAMP during DC maturation, although the stimuli that force DC-STAMP to release CREB3 to be processed in the Golgi are unknown.

3.4.2 CREB3L2 and B cell activation

CREB3L2 is strongly upregulated in differentiating primary human B cells (Al-Maskari et al. 2018). Concurrently, CREB3L2 is also cleaved as B cells approach the antibody secreting plasma cellular phase. Surprisingly, when cleavage of CREB3L2 was blocked by an S1P inhibitor, there was a B cell differentiation and proliferation defect early in development, specifically between

the activated B cell and the plasmablast stages, when B cell division is rapid. Treatment with the S1P inhibitor resulted in reduced plasmablast numbers and in defective antibody secretion by CD38 low plasmablasts. S1P inhibitor treated plasmablasts also displayed markers of autophagy and an altered gene expression profile in comparison to non-treated cells.

But importantly, inhibition of S1P is not specific for CREB3L2 and influences all other substrates of S1P (e.g.: ATF6, SREBP, CREB3…) (Al-Maskari et al. 2018). Therefore, the observed pleiotropic abnormalities in B cell development could be the result of concerted action of S1P-regulated factors and not only of CREB3L2. This would agree with gene expression changes observed in plasmablasts, for example downregulation of the sterol synthesis pathway upon S1P inhibitor treatment. Nevertheless, strong upregulation and cleavage of CREB3L2 during B cell differentiation indicate the importance of this factor for B cell development.

3.5 iRhoms

The iRhoms are immune-relevant rhomboid-like proteins, but unlike rhomboid proteases, which are serine IMPRs, iRhoms are catalytically inactive.

The first immunity related client of iRhom2 to be identified in mammalian cells is **ADAM17 / TACE**. By binding to ADAM17, iRhom2 regulates its trafficking and maturation and guides it from the ER to the Golgi (McIlwain et al. 2012; Adrain et al. 2012). This role of iRhom2 is functionally redundant with iRhom1, but the two homologs are expressed differentially in different cell types (Li et al. 2015). ADAM17 is a metalloprotease which cleaves a number of substrates at the plasma membrane. To be active, ADAM17 has to be modified during its transfer through the secretory pathway (Lambrecht et al. 2018). Therefore, when iRhom2 is not present, ADAM17 is retained in the ER, where it remains non-functional and its substrates are not cleaved.

An important substrate of ADAM17 is the membrane bound pro-form of TNF. Soluble TNF, the product of ADAM17 cleavage, is an important pro-inflammatory cytokine. In the absence of iRhom2, soluble TNF is not produced, mice are insensitive to lipopolysaccharide induced septic shock and at the same

time they are defenceless against some bacterial pathogens (McIlwain et al. 2012; Siggs et al. 2012).

Many immunologically relevant molecules are cleaved by ADAM17 in cooperation with its relative ADAM10. ADAM10 activity is constitutive and is not regulated by iRhoms. This is exemplified by FasL processing in T cells (Ebsen et al. 2015), interleukin (IL)-23 receptor shedding (Franke et al. 2016), degradation of soluble interferon (IFN)γ (Kanzaki et al. 2016), or shedding of MICA and MICB as an immune escape strategy in tumors (Chitadze et al. 2013). The role of ADAMs in immunity has been reviewed recently (Lambrecht et al. 2018).

Another immunity related client of iRhom2 is **STING** (Luo et al. 2016), an ER-resident transmembrane protein. The STING pathway (Chen et al. 2016) is important for the innate immune response, as it enables recognition of pathogenic and/or our own DNA in the cytosol of cells and activates the immune system. More specifically, STING is essential for production of type I interferons (IFN-I) after stimulation with non-CpG DNA of various origin (Ishikawa et al. 2009). iRhom2 was recently added to the STING pathway as an additional STING regulator, because iRhom2 deficiency was found to have a detrimental effect on IFN-I production in response to cytosolic DNA, very similar to STING deficiency (Luo et al. 2016). iRhom2 interacts with STING and mediates its association with TRAPβ, an important regulator of STING trafficking. Also, iRhom2 supports STING stability by inhibiting its K48 polyubiquitination through the recruitment of a deubiquitinase.

Another client of iRhom2, recently identified, is **MAVS** (also known as VISA, IPS-1 or Cardif) (Luo et al. 2017). It was independently discovered by four groups at the same time (Kawai et al. 2005; Seth et al. 2005; Xu et al. 2005; Meylan et al. 2005), and each of these invented their own name for the protein. IPS-1 stands for *IFNβ promoter stimulator 1* and was identified in a screen for proteins activating the IFN-I promoter (Kawai et al. 2005). MAVS means *mitochondrial antiviral signalling* and was identified as a mitochondrial protein containing a CARD-like domain similar to the CARD domain of RNA-sensing proteins RIG-I and MDA-5 (Seth et al. 2005). VISA is an abbreviation for *virus induced signalling adaptor* (Xu et al. 2005), and was identified as a protein strongly activating the NF-κB

promoter. Cardif means *CARD adaptor inducing IFNβ* and it was examined as a potential adaptor of RIG-I with a CARD-like domain (Meylan et al. 2005).

Together, these articles revealed that MAVS is a mitochondrial protein attached to the outer mitochondrial membrane via its C-terminal transmembrane domain. It has a CARD-like domain on its N-terminus, which enables interaction with other CARD-containing proteins, specifically with RIG-I and MDA-5. RIG-I and MDA-5 are cytosolic sensors for non-self RNA molecules and MAVS transfers a signal from these sensors to the kinases activating IRF3 and NF-κB (Kell et al. 2015). Therefore, MAVS is an essential positive regulator of the cellular interferon response, resulting from infection by RNA virus or transfection by artificial dsRNA, and a principal factor of an efficient anti-viral defence.

iRhom2 interacts with MAVS in the ER-mitochondrial contact sites, a compartment where mitochondrial and ER membrane meet. According to a recent report, iRhom2 supports the stability of MAVS, therefore potentiating an interferon response following infection by RNA viruses (Luo et al. 2017). The stabilisation is mediated by iRhom2-induced degradation of two ubiquitin ligases, RNF5 and MARCH5, which target MAVS for degradation by polyubiquitination.

3.6 RHBDD3

RHBDD3 is another pseudoprotease of the rhomboid-like superfamily. It is an important regulator of NF-κB signalling specifically in conventional dendritic cells (cDCs) (Liu et al. 2014). Study of RHBDD3 deficient mice revealed a spontaneous autoimmune phenotype, very severe induced colitis, increased levels of IL-6 and IL-17, decreased number of regulatory T cells and an inflammatory phenotype in cDC (Liu et al. 2014). When characterising cDCs in detail, RHBDD3 was shown to suppress excessive NF-κB signalling. It does so by suppressing K63-linked polyubiquitination of the protein NEMO by recruitment of deubiquitinase A20. NEMO is an important regulator of IKK kinase and is substantially polyubiquitinated on various lysine residues. The K63-linked polyubiquitination which is suppressed by RHBDD3 is necessary for the activation of IKK, which subsequently activates NF-κB. RHBDD3 itself associates with NEMO via another polyubiquitin chain. Moreover, even RHBDD3 itself is ubiquitinated and via this

polyubiquitin associates with A20. In the absence of RHBDD3, NF-κB is hyperactivated in cDCs in response to TLR stimuli. As a result, cDCs then express a large amount of IL-6, TNF, MHC-II and co-stimulatory molecules, therefore supporting the Th17 pro-inflammatory immune response.

Prior to this study, another link between RHBDD3 and immunity was proposed, specifically that RHBDD3 deficient NK cells are more cytotoxic and pro-inflammatory (J. Liu et al. 2013). This increased cytotoxic ability was suggested to be intrinsic to NK cells and independent of dendritic cells. There was some indication that RHBDD3 may induce proteasomal degradation of DAP12 protein, a signalling adaptor of stimulatory NK cell receptors, and so negatively regulate NK cell cytotoxicity, but the evidence for this is not very convincing.

4 Rhomboid superfamily proteins

A rhomboid protein was first described in *Drosophila* during genetic and developmental studies (Jurgens et al. 1984). Ethyl methane sulfonate-treated *Drosophila* eggs, unable to hatch, were examined for larval cuticule abnormalities and their corresponding mutated alleles were found. Since then, the *rhomboid* gene bears its name as the mutant larva had a 'rhomboid-like' (or pointed) head skeleton. Subsequently, the *rhomboid* sequence was compared to genomic databases and several highly similar genes were found. Six more rhomboid-like genes were discovered in *Drosophila*, so the first *rhomboid* was renamed as *rhomboid-1* and the others were numbered in ascending order (Wasserman et al. 2000). In 2001, rhomboid-1 became the first intramembrane serine protease to be described (Urban et al. 2001).

Gradually, rhomboid-like genes were also identified in bacteria, archaea, yeast, nematodes, plants and mammals (Wasserman et al. 2000; Pascall & Brown 1998), but their overall sequence similarity is quite low. Therefore, to specify the definition of a true rhomboid protein, the position of the active site residues within the protein and the overall predicted transmembrane topology were together assessed (Lemberg & Freeman 2007). A true rhomboid protease was defined to consist of a six TMH core (Fig.5). At the N-terminus of TMH4 is the GxSx motif harbouring the catalytic serine, and at the N-terminus of TMH6 is the conserved

catalytic histidine, forming the catalytic dyad of rhomboid protease (Lemberg & Freeman 2007). Based on this rhomboid protein definition and on sequence homology, rhomboids from important eukaryotic model organisms were identified and subjected to phylogenetic analysis (Lemberg & Freeman 2007). Apart from the true rhomboid proteases a tight cluster of proteolytically inactive rhomboid-like proteins, where these catalytic residues were absent, was identified. These proteins are now known as iRhoms (for inactive rhomboid protease-like homologs).

The only available structure of a rhomboid protein is that of rhomboid protease GlpG from *E. coli* (Wang et al. 2006; Wu et al. 2006). The topology of all other rhomboid proteins is derived from homology modelling based on the GlpG structure. The capability to form the GlpG fold and the resulting position of the characteristic residues are assessed when a protein is classified as belonging to the rhomboid family today. Apart from the catalytic residues in TMH4 and TMH6, rhomboids usually have a conserved motif in the first loop (L1) connecting TMHs 1 and 2 and a conserved GxxxG packing motif in TMH6 (Lemberg & Freeman 2007) (Fig.5).

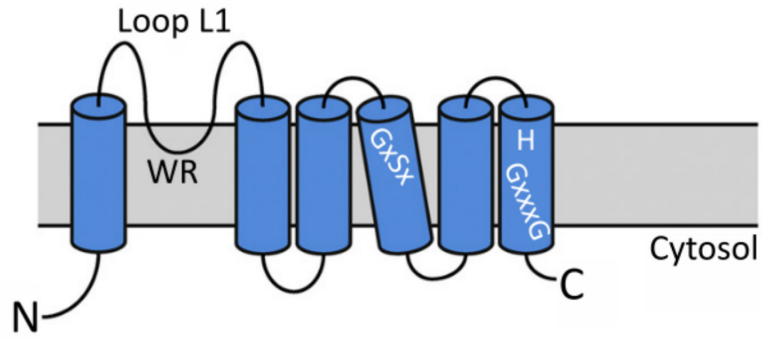

Fig.5: The conserved core of rhomboid protease (adopted and adjusted from Tichá et al. 2018).

4.1 Non-protease members (Rhomboid pseudoproteases)

iRhoms were the first identified group of rhomboid pseudoproteases. They are larger proteins than typical rhomboid proteases, because of large soluble domains at the N-terminus and the highly conserved iRhom homology domain (IRHD) between the first and second TMH. Compared to the rhomboid core, they also have an additional C-terminal TMH (Fig.6). The immunological role of iRhoms is extensively described in the Chapter 3. Other slightly more distant inactive rhomboid homologs include Derlins, RHBDD2, RHBDD3, UBAC2 and TMEM115

(Lemberg & Adrain 2016) (Fig.6). Their cellular functions will be briefly described, as their topology and suggested biological functions are predicted to be similar to rhomboid protease RHBDL4, which is the subject of this thesis. The RHBDD3 protein and its functions have been already mentioned in Chapter 3, as it is relevant for the immune system.

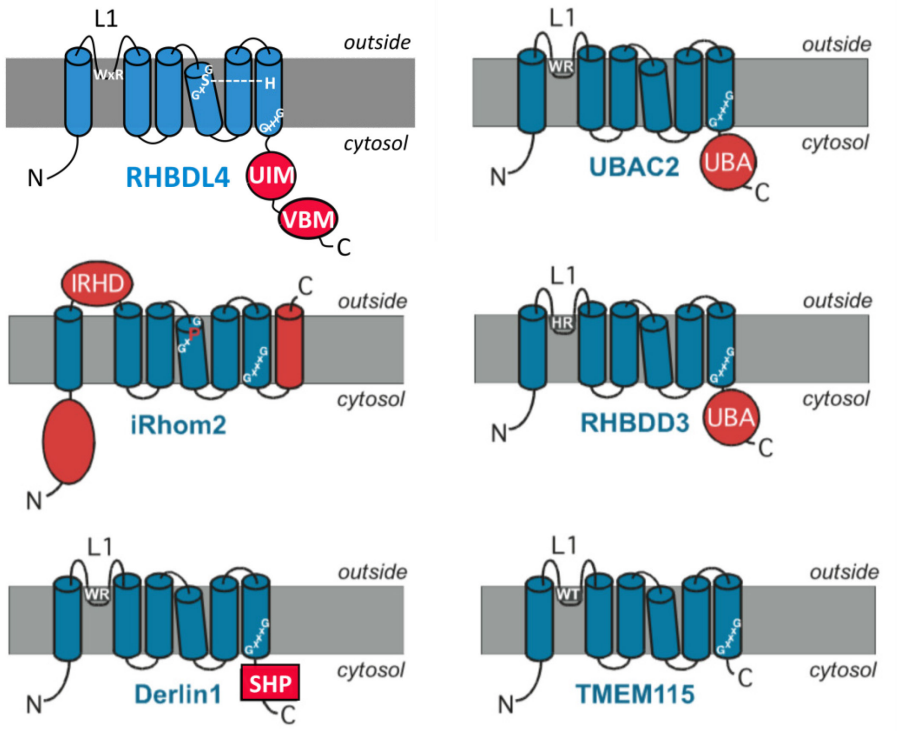

Fig.6: Mammalian rhomboid pseudoproteases and RHBDL4 (adopted and adjusted from Lemberg & Adrain 2016).
Schematic representation of mammalian rhomboids with conserved motifs depicted. The core consisting of six TMHs is blue, additional domains are coloured in red. UIM – ubiquitin interaction motif, VBM – VCP binding motif, IRHD – iRhom homology domain, SHP – another VCP binding motif, UBA – ubiquitin associated domain, L1 – loop 1

4.1.1 Derlins

Derlins are ER localised rhomboid pseudoproteases conserved among eukaryotes (Greenblatt et al. 2011). Only four TMHs were initially predicted in their structure by hydrophobicity analysis (Knop et al. 1996; Ye et al. 2004), but modeling of Derlin-1 sequence on the structure of rhomboid protease revealed a similarity to the rhomboid superfamily (Fig.6) (Greenblatt et al. 2011). Homology searches confirmed the conservation of rhomboid motifs in Derlin-1, and they were found to be necessary for activity of Derlin-1. The new 6 TMH model of Derlin-1

was also confirmed by Ramachandran analysis, the positive inside rule and by analysis of glycosylation sites (Greenblatt et al. 2011).

The role of Derlins is connected with ERAD and UPR. Der1 (degradation in the ER 1) was first identified in yeast during a screen for proteins whose mutation abrogates ERAD (Knop et al. 1996). By intracellular crosslinking, Der1 was shown to interact with yeast translocation machinery involved in ERAD, specifically with Hrd1 ubiquitin ligase and Hrd3, which is a receptor for misfolded proteins, and with soluble ERAD substrates (luminal misfolded proteins) (Mehnert et al. 2014). Der1 enables transfer of misfolded proteins from Hrd3 to Hrd1 and their subsequent translocation to the cytosol. Potentially, Der1 might form some kind of channel together with Hrd1 and catalyse the translocation of misfolded proteins from the ER lumen towards the cytosol, followed by their ubiquitination by Hrd1.

Another yeast Derlin, Dfm1, is involved in ERAD of membrane proteins, especially in their retro-translocation (Neal et al. 2018). Unlike Der1, Dfm1 possesses an additional sequence motif (the SHP motif) for interaction with the Cdc48 ATPase (also known as VCP or p97 in mammals), an ATPase which fuels the retro-translocation of misfolded proteins during ERAD. Interestingly, in Dfm1-deficient cells subjected to ER stress, the requirement for Dfm1 is bypassed by upregulation of Hrd1 (via chromosomal duplication), which can compensate for the absence of Dfm1.

In mammals, there are three Derlin homologs. The most well-studied is Derlin-1. All three are ER localised, connected with ERAD and upregulated during the UPR (Oda et al. 2006). Derlin-2 and Derlin-3 associate with EDEM proteins, which are manosidases involved in the initiation of ERAD of glycoproteins (Oda et al. 2006). Derlin-1 was discovered as a part of ER membrane complex, which serves as a receptor for p97 ATPase (Ye et al. 2004). It participates in ERAD of major histocompatibility complex one (MHC-I) heavy chain, which is strongly induced by US11 protein from human cytomegalovirus as an immune escape strategy (Ye et al. 2004; Lilley & Ploegh 2004). The association of Derlin-1 with p97 and its activity in ERAD are dependent on the cytosolic SHP motif as is the case with Dfm1 (Greenblatt et al. 2011). But unlike Dfm1, whose substrates are only membrane proteins, Derlin-1 is also involved in ERAD of luminal misfolded proteins.

How exactly Derlins participate in the retro-translocation is not yet known. Based on the topology in the GlpG structure, the TMHs do not form a real channel, which could conduct proteins from the ER lumen to the cytosol (Greenblatt et al. 2011). However, Derlins may be able to oligomerize (Mehnert et al. 2014) and they closely interact with other factors (such as Hrd1), therefore the possibility that they form a channel together with other components still exists. As the conserved rhomboid motifs are important for Derlin function, it is likely that some of these common features are required for translocation (Greenblatt et al. 2011). This may be the ability to partially unwind and destabilise their substrates or some interactions with lipids enabling insertion of the substrate into the ER membrane.

4.1.2 UBAC2

UBAC2 rhomboid pseudoprotease is ER localised (Christianson et al. 2011; Olzmann et al. 2013), with a predicted 6 TMH fold and a C-terminal polyubiquitin binding UBA domain (Fig.6) (Christianson et al. 2011). In a systematic approach to reveal the interaction network among mammalian ERAD proteins, UBAC2 was found to be a central component of the gp78 E3 ubiquitin ligase complex (Christianson et al. 2011). The whole complex consists of gp78, UBAC2, UBXD8 and Derlin-2, where UBAC2 mediates interaction with UBXD8. Because UBXD8 recruits the p97 ATPase, the absence of UBAC2 may abrogate the recruitment of p97 to the E3 complex. Indeed, knock-down of UBAC2 resulted in reduced ERAD of several model substrates (Christianson et al. 2011).

Interestingly, UBXD8 is also involved in the negative regulation of lipid turnover (Olzmann et al. 2013). By default, UBXD8 is associated with the membrane of lipid droplets and interacts with a rate-limiting enzyme of lipid hydrolysis, adipose triglyceride lipase (ATGL). UBXD8 recruits p97 to the lipid droplet and together they mediate removal of accessory cofactors from ATGL, which leads to ATGL inactivity. As a result, degradation of lipids is reduced, and lipid droplets can increase in their size and number. The role of UBAC2 is to sequester UBXD8 from lipid droplets, as UBAC2 is localised in the ER and the affinity between these two proteins is high. Therefore, the protein ratio between UBAC2 and UBXD8 determines how much of the UBXD8 pool is localised to lipid droplets and thus the relative levels of lipid degradation or accumulation.

A link between UBAC2, lipid metabolism and E3 ubiquitin ligases also exists in yeast (Lloyd et al. 2013), in which the UBAC2 homolog Dsc2 is a central component of the E3 ubiquitin ligase complex termed Dsc. Dsc is localised to the Golgi and is in many aspects similar to the mammalian gp78. Dsc is responsible for cleavage of SREBP factors in yeast and the liberation of the corresponding transcription factors regulating lipid synthesis. Mammalian-like activation of the SREBP pathway, described above, does not occur in yeast, because there are no homologs of S1P and S2P, which sequentially cleave SREBP factors in mammals.

4.1.3 RHBDD2

RHBDD2 is a rhomboid pseudoprotease assumed to be localised in the Golgi (Ahmedli et al. 2013), although one study suggests ER localisation (Abba et al. 2009). Little is known about RHBDD2 function. High expression of this protein is associated with proliferating tissues in rats (Ferretti et al. 2015) and with colorectal and breast cancers in humans (Abba et al. 2009; Lacunza et al. 2012; Lacunza et al. 2014; Canzoneri et al. 2018). In one case, knock-down of RHBDD2 in breast cancer cell lines resulted in decreased cell proliferation, decreased cell migration and increased expression of the UPR associated proteins compared to control cell lines (Abba et al. 2009; Lacunza et al. 2014). RHBDD2 may thus be involved in UPR resolution, therefore supporting growth of stressed cancer cells, but this hypothesis remains speculative.

RHBDD2 was identified in a screen for proteins relevant for proper function of the cone photoreceptors, so its expression in mouse retina was assessed (Ahmedli et al. 2013). RHBDD2 expression levels in retina vary during mouse development, suggesting a role in the development of cones. Interestingly, a missense mutation in RHBDD2 was found in a patient with familial autosomal recessive retinitis pigmentosa.

4.1.4 TMEM115

TMEM115 (also known as PL6) is considered to be another rhomboid pseudoprotease with the 6 TMH fold based on structural homology to GlpG rhomboid (Fig.6) (Lemberg & Adrain 2016). However, according to a recent study, it is predicted to consist of 4 TMHs and a long C-terminal cytosolic domain (Ong

et al. 2014). An earlier study of TMEM115 suggests a rhomboid like motif (Ivanova et al. 2008), supporting its membership in the rhomboid superfamily. TMEM115 is Golgi resident in the majority of cells and tissues (Ivanova et al. 2008; Ong et al. 2014). Upon TMEM115 knock-down, the Golgi changes morphologically to form more tightly packed cisternae, and the glycosylation pattern on the cell surface is altered (Ong et al. 2014). Upon treatment with BFA, Golgi dissociation is reduced and significantly delayed. This and the glycosylation defect are changes typically observed following knock-down of proteins of the conserved oligomeric Golgi (COG) tethering complex. The COG complex is important for retrograde transport from the endosomes to the Golgi and between different compartments of the Golgi (Fisher & Ungar 2016). It interacts with many proteins involved in intracellular trafficking, including Rab, SNARE, golgins or vesicle coat proteins, and probably links their functions. Therefore, it is presumably involved in the selection and tethering of vesicles, their docking and fusion with the correct cisternae. The defect in glycosylation, when the COG complex is not functional, is caused by mislocalisation of glycosyl transferases and other enzymes involved in glycosylation. TMEM115 can be co-precipitated with all subunits of the COG complex, when co-expressed, and the strongest interaction is with COG4 (Ong et al. 2014). But these interactions need to be confirmed at the endogenous level.

4.2 Rhomboid proteases – proteolytic mechanism and substrate recognition

The mechanism of substrate cleavage was studied largely using bacterial rhomboids. However, as rhomboids are assumed to be mechanistically related, many features can presumably be applied to rhomboids in general (Strisovsky et al. 2009; Lemberg & Freeman 2007). Rhomboids are serine proteases. Their catalytic center is inside an aqueous cavity, embedded in the membrane and open to the extracellular/luminal area (Fig.7) (Wang et al. 2006; Wu et al. 2006). They catalyse the cleavage of a peptide bond using a catalytic dyad, formed by serine and histidine (Lemberg et al. 2005; Wang et al. 2006). Due to the absence of aspartate, which typically contributes to a catalytic triad in many soluble

proteases, the catalytic efficiency of rhomboids is lower (e.g. compared to trypsin) (Strisovsky 2016).

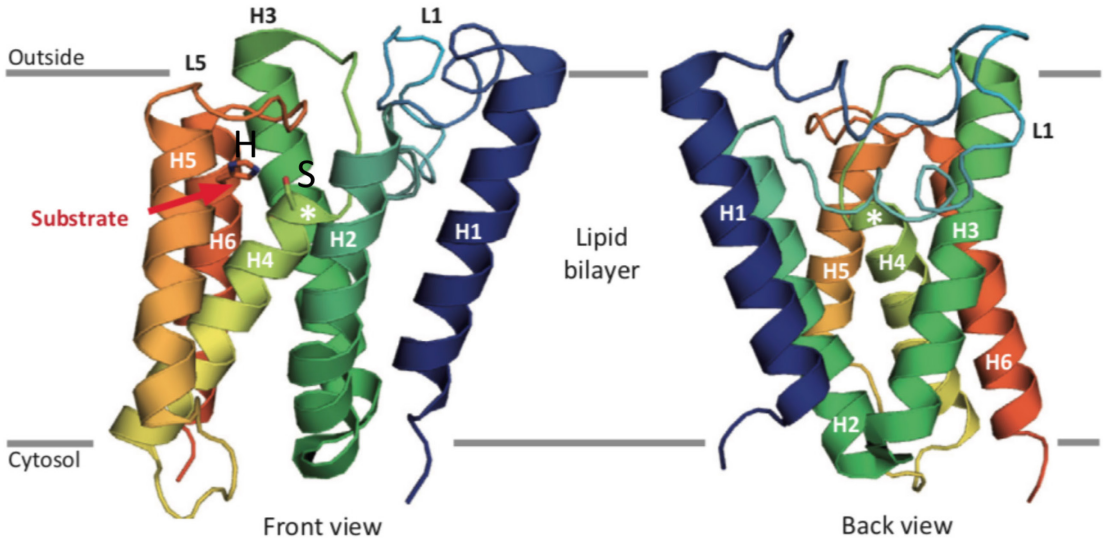

Fig.7: Structure of rhomboid protease GlpG from *E. coli* (adopted from Lemberg 2013). The core of GlpG is shown, consisting of six TMHs. This core is conserved among rhomboids. The white star marks the base of the aqueous cavity, which is open to the outside of the cell. In the front view, the catalytic residues S and H in the cavity are shown. The direction of substrate entry into the active site is indicated by the red arrow.

In the acid-base mechanism proposed for rhomboids, histidine acts as a general base and serine as a general acid (Strisovsky 2016). Based on computational analyses it was proposed that when the enzyme is substrate-free, the alkalinity of histidine is low and the catalytic serine is protonated (Uritsky et al. 2012). Substrate docking into the catalytic cavity displaces water molecules, which increases the alkalinity of histidine, hence deprotonating the serine, which can then perform a nucleophilic attack on the peptide bond. An acyl-enzyme covalent intermediate is formed, which is subsequently hydrolyzed by water. The water molecule required for the hydrolysis can access the catalytic cavity, because it is open to the aqueous environment outside the membrane. Although the substrate displaces water molecules from the catalytic cavity, there is a water retention site near the catalytic serine (Zhou et al. 2012).

Rhomboids display substrate specificity, which depends on several features of the substrate with varying levels of understanding. Rhomboid proteases recognize substrates by loose sequence preferences and poorly understood conformational preferences (Strisovsky et al. 2009; Tichá et al. 2018), precluding identification of rhomboid substrates by simple sequence searching in protein

databases. It is generally accepted that rhomboid recognizes its substrate in two steps (Fig.8) (Strisovsky et al. 2009; Strisovsky 2016). Firstly, the TMH of the substrate interacts with rhomboid in the so-called intramembrane exosite. The structural details of this interaction are unknown. Some studies indicate that the TMH of the substrate must be helically unstable, for example, by containing prolines, glycines, or polar amino acids (Urban & Freeman 2003). Other studies suggest that a lateral gate opening for the substrate might occur (Wu et al. 2006) while others dispute this (Xue & Ha 2013). Secondly, the protein chain surrounding the cleavage site, the so-called recognition motif, docks into the catalytic cavity of the rhomboid and defines the site of cleavage (Strisovsky et al. 2009) and cleavage kinetics (Strisovsky et al. 2009; Zoll et al. 2014; Ticha et al. 2017; Tichá et al. 2017). The cleavage site is located either in the TMH of the substrate near the extracellular/luminal side of the membrane, or in the extracellular/luminal domain of the substrate. If it is in the substrate TMH, the alpha helix needs to be partially unwound, so that the cleaved bond can be placed into the catalytic cavity. For such unwinding, helix-destabilising residues must be present. If the cleavage site is in an extramembrane domain, it can enter the catalytic cavity presumably without helix unwinding (Strisovsky et al. 2009).

Not surprisingly, the specificity and catalytic activity of the rhomboid can be altered by changing the membrane environment (Bondar et al. 2009; Moin & Urban 2012). Rhomboids are thought to make the membrane thinner in their immediate vicinity (Bondar et al. 2009), which was thought to participate in substrate unwinding (Strisovsky et al. 2009) and may be the reason for the unusually fast diffusion rate of rhomboid proteases in membranes (Kreutzberger et al. 2019). This recent study showed that thanks to their irregular shape, their cytosolic domains and the membrane thinning, rhomboids can overcome the viscosity of the membrane and diffuse as fast as a small protein with one TMH would. This was shown to be important for the proteolytic activity of rhomboids and could be a unifying feature of rhomboids, responsible for the strict conservation of the rhomboid fold, even among rhomboids with no catalytic activity.

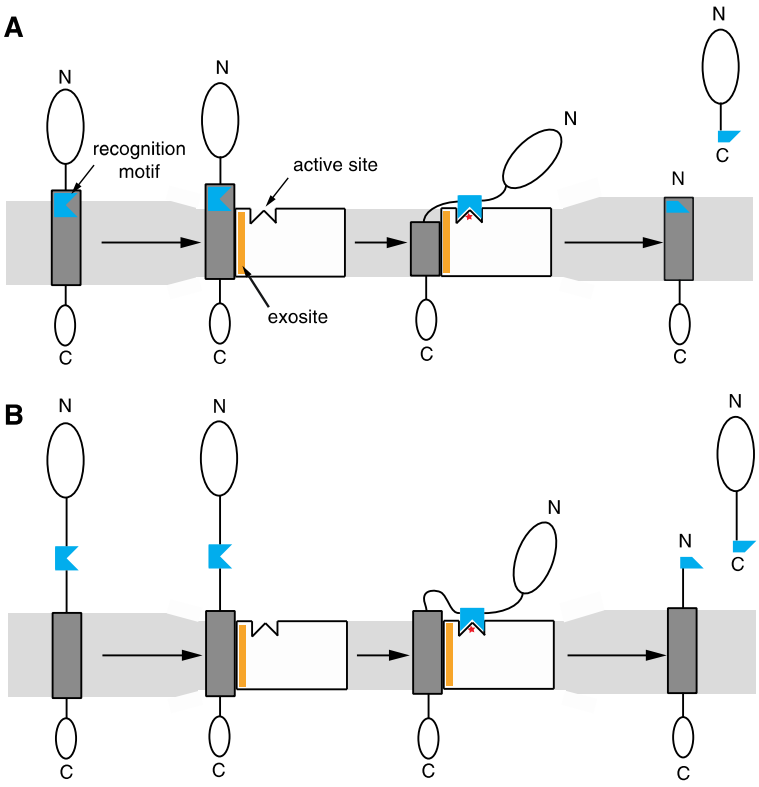

Fig.8: Substrate recognition by rhomboid protease (adopted from Strisovsky et al. 2009). The substrate TMH is recognized by the exosite (yellow) of the rhomboid. The recognition motif (blue) is recognized by the active site of the rhomboid, where catalytic residues (red dot) cleave the substrate. If the cleavage site is in the TMH, helix unwinding occurs (**A**). If the cleavage site is in the extracellular domain, it docks into the active site without helix unwinding (**B**). Typical thinning of the membrane in the vicinity of rhomboid is shown.

4.3 RHBDL4 rhomboid protease

RHBDL4 is a rhomboid protease (also known as RHBDD1) localised in the ER of mammalian cells (Fleig et al. 2012; Wunderle et al. 2016). It is conserved among eukaryotes (Lemberg & Freeman 2007) and it is one of the best studied mammalian rhomboid protease. Based on topology prediction, both RHBDL4 termini face the cytosol and the active site faces the ER lumen. The C-terminus is longer and contains two conserved regions, ubiquitin interacting motif (UIM) and the VCP (also known as p97 ATPase) binding motif (VBM) (Fleig et al. 2012; Lim et al. 2016). Of human rhomboids, RHBDL4 is the only protease with only 6 TMHs. Other rhomboid proteases possess additional TMHs and soluble domains either at the N- or at the C-terminus. Interestingly, RHBDL4 seems to be more similar to inactive rhomboid-like proteins than to rhomboid proteases in humans, considering the protein fold and sequence similarity (Fig.6,9) (Bergbold & Lemberg 2013). In addition, the VBM and UIM motifs (or their functional analogies) are

found among rhomboid pseudoproteases and the proposed cellular function of RHBDL4 combines many functions suggested for rhomboid pseudoproteases. Possibly, RHBDL4 might be also an evolutionary ancestor of more complex rhomboids due to the basic 6 TMH fold (Lemberg & Freeman 2007).

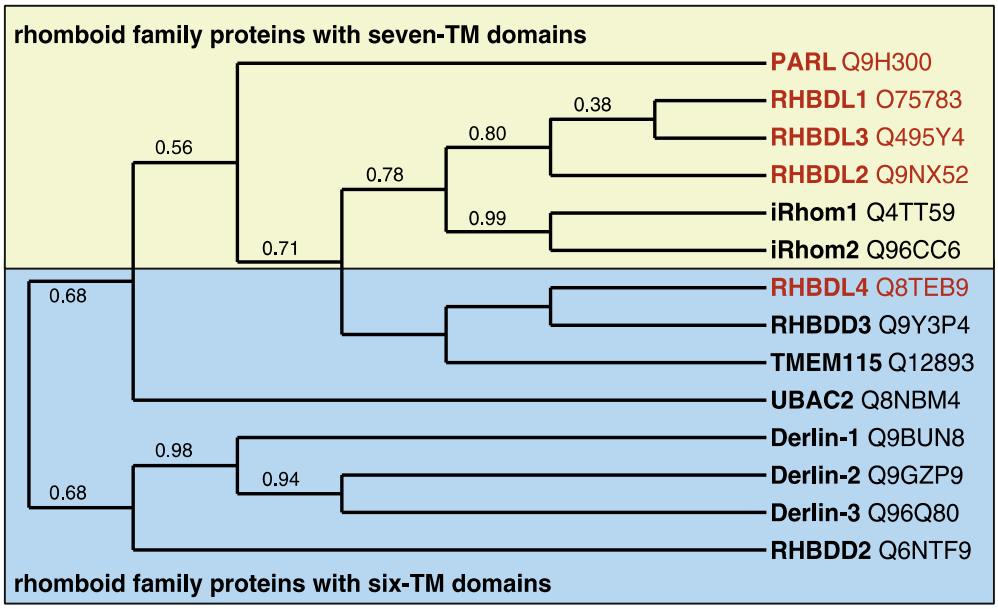

Fig.9: Phylogenetic relationship between rhomboid superfamily proteins found in the human cell (adopted from Bergbold & Lemberg 2013). Rhomboid proteases are shown in red, inactive rhomboid-like proteins are shown in black. The relationship between rhomboids is assessed according to similarity between conserved rhomboid regions.

The intracellular localisation of RHBDL4 was initially ambiguous. It was shown to be ER-localised using two independent methods (Fleig et al. 2012; Wunderle et al. 2016). Firstly, green fluorescent protein (GFP)-tagged RHBDL4 co-expressed with the ER-retained red fluorescent protein (RFP)-KDEL colocalise using confocal microscopy (Fleig et al. 2012; Wunderle et al. 2016). Secondly, cellular fractionation indicates an RHBDL4 signal in the ER fraction (Fleig et al. 2012). Previously, RHBDL4 was mistakenly thought to be localised to mitochondria (Y. Wang et al. 2008), or plasma membrane (Song et al. 2015). Wunderle *et al.* admit that they cannot completely exclude the possibility that a small fraction of RHBDL4 is outside the ER, but the majority of RHBDL4 is ER localised. RHBDL4 mislocalisation in Song *et al.* may have been caused by formaldehyde fixation of cells (Wunderle et al. 2016).

The cellular function of RHBDL4 remains ambiguous. Various roles for RHBDL4 have been proposed. RHBDL4 may be an ERAD factor (Fleig et al. 2012),

may be involved in secretion (Wunderle et al. 2016; Wan et al. 2012), in cancer and apoptosis (Song et al. 2015; M. Zhang et al. 2018; X. Zhang et al. 2018; Barrett et al. 2019) or it may act as an alternative protease for amyloid precursor protein (APP) hydrolysis (Paschkowsky et al. 2016; Recinto et al. 2018; Paschkowsky et al. 2018).

4.3.1 RHBDL4 in ERAD

RHBDL4 expression is increased upon induction of the UPR (Fleig et al. 2012). The UPR pathway helps to resolve stress in the ER caused by accumulation of unfolded proteins (Hetz 2012; Hetz & Papa 2018). Therefore, when the UPR is triggered, many proteins are typically upregulated which: 1) decrease the amount of proteins reaching the ER, 2) increase the protein folding capacity of the ER, 3) remove terminally unfolded proteins from the ER, 4) suppress apoptosis until ER stress is resolved. If ER stress is not resolved in a reasonable time, cells undergo apoptosis. Therefore, upregulation of RHBDL4 during the UPR suggests that RHBDL4 function is needed for one or more of the above-mentioned processes.

Additionally, RHBDL4 recognizes and cleaves membrane proteins, polytopic or with a single TMH, which bear two basic amino acids in their transmembrane domain (Fleig et al. 2012). This di-basic motif was originally identified as a degron recognized in orphan receptor subunits, inducing their degradation via ERAD. RHBDL4-catalysed hydrolysis also induces ERAD (Fleig et al. 2012). Upon inhibition of the proteasome, the substrates and their cleavage products were stabilised. Moreover, two ERAD relevant sequence motifs were found in the RHBDL4 C-terminal domain, the UIM and VBM. The UIM allows RHBDL4 to interact with ubiquitinated substrate proteins and cleave them efficiently. The VBM motif mediates RHBDL4 interaction with p97 ATPase, whose activity was important for transport of cleavage fragments to the proteasome.

The interaction between RHBDL4 and p97 was supported by a crystal structure of the RHBDL4 C-terminus with p97 N-terminal domain (p97N) (Lim et al. 2016). *In vitro* cross-linking and pull-down experiments show that RHBDL4 and p97N can both bind to polyubiquitin and together form a ternary complex (Lim et al. 2016). However, if p97N is already bound to the VBM motif, it cannot

simultaneously bind ubiquitin. This data led to a proposed mechanism for the interaction between RHBDL4 and p97 (Fig.10).

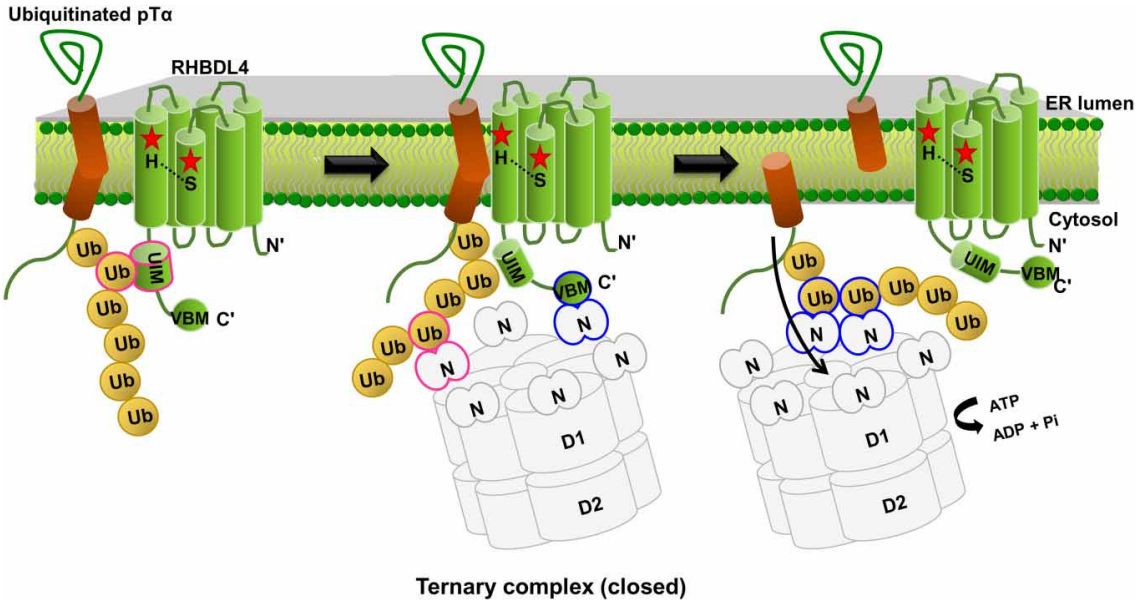

Fig.10: RHBDL4 and p97 cooperate to promote ERAD of ubiquitinated membrane substrates e.g. pTα (adopted from Lim et al. 2016).
RHBDL4 recognizes the helix destabilising residues in the TMH of the substrate and the polyubiquitin chain. Ubiquitin binds to the UIM and the p97 ATPase is subsequently recruited via the VBM. p97 N domains bind not only to VBM but also to the ubiquitin forming a ternary complex. When the substrate is cleaved by RHBDL4, it has higher affinity to p97 due to cooperative binding to multiple p97 N subunits. Therefore, it dissociates from RHBDL4 and enters the cytosol together with p97, finally reaching the proteasome machinery. The interactions highlighted in blue are three orders of magnitude stronger than the interactions highlighted in pink. pTα was investigated as a substrate of RHBDL4 destined for ERAD after cleavage (Fleig et al. 2012).

ER stress occurs naturally during the lifetime of many cells in our body, with the most extreme examples being B cells producing large amounts of antibodies or pancreatic beta cells producing insulin (Hetz 2012; Hetz & Papa 2018). As the unresolved ER stress results in apoptosis, defects in ERAD and UPR are linked to human diseases, which is explored for Derlin-1 rhomboid pseudoprotease in particular. For example, among patients with familial amyotrophic lateral sclerosis a mutation of superoxide dismutase 1 (SOD1) causes association of SOD1 with Derlin-1, abrogating Derlin-1 function and resulting in ER-stress and subsequent death of motor neurons (Nishitoh et al. 2008). On the other hand, cancer cells benefit from a functional UPR as they are protected from cell death resulting from ER stress. In many types of cancer Derlin-1 is upregulated in a way that correlates with increased tumour growth and malignance (J. Wang et al. 2008;

Tan et al. 2015; Pi et al. 2017; Dong et al. 2017). RHBDL4 is also upregulated in cancer cells in this way, as described in section 4.3.3.

4.3.2 RHBDL4 and protein secretion

The role of RHBDL4 in secretion was studied using tumour growth factor α (TGFα), a ligand of EGFR. TGFα is a well-known substrate of ADAM metalloproteases. It is expressed as a single transmembrane protein with C and N terminal pro-domains and ADAMs cleave it to release mature TGFα from plasma membrane. Upon ADAM inhibition, an unusual TGFα cleavage pattern was observed and a noncanonical secretion of TGFα occurred in a manner dependent on RHBDL4 activity (Song et al. 2015; Wunderle et al. 2016). Song *et al.* studied a 17kDa fragment of TGFα, which was (according to their experiments) a cleavage product of RHBDL4. Misplacing RHBDL4 on the plasma membrane, they inferred that RHBDL4 acts in parallel to ADAMs by cleaving TGFα on the plasma membrane, releasing it from the cell and therefore stimulating the EGFR pathway and cell growth (Song et al. 2015).

Wunderle *et al.* were unable to repeat these experiments. They observed the 17kDa fragment, but its production was independent of RHBDL4 activity and unaffected by disruptive mutations of the presumed TGFα cleavage site. Therefore, they focused on larger secreted TGFα forms, whose production was dependent on RHBDL4 activity. These large fragments were also observed but were neglected by Song *et al.* The larger form of TGFα is 37kDa, and as such cannot be a cleavage fragment. This was confirmed by mutation of potential cleavage sites in TGFα (no effect), by immunoprecipitation of a catalytically inactive RHBDL4 mutant which traps cleavage substrates (no TGFα isolated) and by replacement of TGFα TMH with the calnexin TMH. The 37kDa form of TGFα is the full-length version glycosylated by sugars typical for the late Golgi compartment and it is likely secreted in microvesicles (exosomes), which is dependent on RHBDL4 activity (Wunderle et al. 2016). One possibility to explain this observation is that RHBDL4 might function at ER-exit sites (ERES), where it could stimulate transport of TGFα from the ER to the Golgi, therefore enabling glycosylation and subsequent secretion of TGFα. However, it is difficult to conclude if RHBDL4 is directly linked to microvesicles formation, because secretion of the 37kDa TGFα is also observed

upon inhibition of ADAM metalloproteases irrespective of RHBDL4. When ADAMs are inhibited, TGFα is not cleaved and accumulates at the plasma membrane. Accumulation at the plasma membrane will also happen if anterograde transport of TGFα is stimulated via RHBDL4. Thus, microvesicle formation could simply be the result of an excess of TGFα at the plasma membrane.

Alternatively, a direct link between RHBDL4 and secretion of microvesicles has been proposed as RHBDL4 can cleave the TSAP6 protein, which is involved in the secretion of exosomes (Wan et al. 2012). In RHBDL4 deficient cells, TSAP6 is more stable and exosomes contain more proteins. Thus, RHBDL4 may regulate protein secretion via exosomes, potentially through TSAP6. However, according to these results, the effect of RHBDL4 expression on exosome secretion would be negative, what disagrees with the study concerned with TGFα (Wunderle et al. 2016). To assess the physiological significance of RHBDL4-mediated noncanonical TGFα secretion, one can look at TGFα's bioactivity. This was measured by both Song *et al.* and Wunderle *et al.*, but with conflicting results so this question remains to be clarified.

4.3.3 RHBDL4 in cancer and apoptosis

Many studies relating RHBDL4 to cancer development and progression were done. Unfortunately, in most cases, these focused only on the effects of depletion of RHBDL4 by RNA interference in cancer cell lines. RHBDL4 depletion led to reduced proliferation and increased apoptosis in cell lines derived from colorectal cancer (Song et al. 2015; Han et al. 2015), hepatocellular carcinoma (X.-N. Liu et al. 2013), glioblastoma (Wei et al. 2014) and breast cancer (X. Zhang et al. 2018). A reduction in cell mobility and invasiveness was shown for colorectal cancer (M. Zhang et al. 2018) and breast cancer (Huang et al. 2018; X. Zhang et al. 2018).

Some efforts were made to elucidate the molecular mechanism of these effects. In the case of **breast cancer**, the importance of RHBDL4 was confirmed on breast cancer tissue, where RHBDL4 was upregulated compared to adjacent healthy tissue (Huang et al. 2018; X. Zhang et al. 2018). Transcriptomic analysis of a breast cancer cell line with RHBDL4 knock-down revealed downregulation of anti-apoptotic and cell cycle related genes (X. Zhang et al. 2018). Moreover, the

cells were deficient in cell cycle progression (X. Zhang et al. 2018) and several genes involved in the epithelial-to-mesenchymal transition (EMT) were downregulated (Huang et al. 2018). Proteins that may mediate the changes following RHBDL4 depletion in cells were identified based on their expression level during cell cycle, and include CDK2, p-Akt, p27, and cyclin E1 (X. Zhang et al. 2018). The expression of RHBDL4, CDK2 and Akt correlates in breast cancer tissues.

For **colorectal cancer**, the role of RHBDL4 is best established. In human colorectal tumour, in mouse colitis associated colorectal tumour and in cell lines derived from colorectal cancer (CRC) RHBDL4 protein expression is elevated compared to adjacent healthy tissue (Song et al. 2015). Upon RHBDL4 deficiency in tumour cells, the *in vivo* growth of the tumour is slower, an effect which can be at least partially rescued by re-expression of recombinant RHBDL4. In a study of 539 human CRC patients, high levels of RHBDL4 expression correlated with worse prognosis, worse TNM stage and less differentiated tumour tissue (Song et al. 2015). In contrast, low levels of RHBDL4 correlated with better prognosis. There may be some bias, as a majority of the patients had TNM stage II or III CRC, but the statistics are robust.

Concerning the mechanism of increased tumour growth in RHBDL4 positive cells, high levels of RHBDL4 and EGFR signalling correlate well in a mouse model of CRC (Song et al. 2015). Also, TGFα depletion in a CRC cell line has the same effect on tumour growth as the depletion of RHBDL4. Therefore, **EGFR pathway** stimulation might be a mechanism for how RHBDL4 promote cancer growth. There are two possible connections between RHBDL4 and EGFR signalling.

As mentioned above, RHBDL4 induces noncanonical secretion of TGFα, but it is unclear whether this atypical TGFα can stimulate EGFR (Song et al. 2015; Wunderle et al. 2016). If the noncanonical TGFα was bioactive, this would allow RHBDL4 to directly promote cell proliferation. If the TGFα was not bioactive, other cofactors, for example ADAM metalloprotease, would be required to promote cancer. In an analogy, RHBDL4 would deliver a large number of missiles (immature TGFα) to a launch pad (the plasma membrane/exosome membrane), but RHBDL4 would not be able to fire (cleave and mature) them alone. However, it

would still be a potent mechanism to increase stimulation of EGFR as the active ADAM would fire all of the missiles available.

The second possible link between RHBDL4 and the EGFR pathway is through the **AP-1 transcription factor subunit c-Jun**. c-Jun is a pleiotropic tumour-associated transcription factor, which is known to support tumour growth, regulate apoptosis, increase angiogenesis and the metastatic abilities of cells (Eferl & Wagner 2003). The expression and phosphorylation of c-Jun is positively regulated by expression of RHBDL4 (Song et al. 2013). Therefore, decreased activation of c-Jun could explain all the effects observed in cancer cells where RHBDL4 is knocked down. Furthermore, the expression of EGFR is positively regulated by c-Jun (Miao et al. 2017).

Another cancer-related signalling pathway influenced by the depletion of RHBDL4 is the **Wnt pathway** (M. Zhang et al. 2018). Wnt signalling leads to the stabilisation and nuclear translocation of ß-catenin, which itself regulates the transcription of many target genes (Tortelote et al. 2017). The transcriptional activity of ß-catenin is downregulated in RHBDL4-depleted cells, even though its nuclear translocation is unaffected (M. Zhang et al. 2018). Also, based on transcriptome analysis of RHBDL4-depleted cells, genes associated with EMT and stemness were differentially expressed, so the gene expression pattern was more epithelial-like state and had fewer stem cell-like properties when RHBDL4 was depleted. It was suggested that RHBDL4 stimulates the Wnt pathway by increasing active ß -catenin levels, ß-catenin stimulates the expression of ZEB1 protein (an important EMT regulator), which in turn supports the metastatic abilities of the cells (M. Zhang et al. 2018).

Interestingly, in a detailed study of human **glioblastoma**, based on cultured cells isolated from human surgical biopsies, RHBDL4 was identified among 20 genes potentially useful as treatment targets (Stangeland et al. 2015). Although it was not shortlisted due to a low correlation between expression levels and patient survival, it was strongly upregulated in glioblastoma stem cells compared to neural stem cells, which is another supportive argument for RHBDL4 involvement in cancer.

Another interesting finding emerged from a study of **refractory testicular germ cell tumours** (TGCT), a small fraction of the TGCT, which is malignant (Barrett et al. 2019). This type of tumour is characterised by a low mutation rate, aneuploidy, gain of a 12p chromosome arm and mixed histology. In the refractory TGCT, additional changes in the genome were found, which might be responsible for its malignity. Among these was an intensive focal amplification of the RHBDL4 gene and of the adjacent insulin receptor substrate 1 (IRS1) gene. Although the malignity might be the result of IRS1 amplification and RHBDL4 could be present due to its proximity, it is also possible that RHBDL4 supports malignity of these tumours. Of note, RHBDL4 is expressed extensively in human and mouse testis (Y. Wang et al. 2008; Wang et al. 2009), although a link between RHBDL4 and testis function is missing.

To conclude, it is clear that RHBDL4 is an anti-apoptotic and pro-proliferative protein, and as such is relevant to cancer. Among many studies, several molecular mechanisms explaining these qualities of RHBDL4 were proposed, including the increased activity of AP-1, ß-catenin and EGFR signalling. RHBDL4 might support cancer development also as an ERAD factor, but none of these mechanisms has been well confirmed and understood in detail.

4.3.4 RHBDL4 as an alternative protease of APP

APP (amyloid precursor protein) is a single-pass transmembrane protein. It is trafficked to the plasma membrane, where it is cleaved by α-, ß- and γ-secretases. A specific cleavage fragment of APP, Aß, is known to aggregate into amyloid plaques which are associated with Alzheimer's disease (O'Brien & Wong 2011).

When APP was co-transfected into cells with catalytically active RHBDL4, atypical APP cleavage was observed, yielding N-terminal fragments of about 70 kDa and C-terminal fragments in the 10-25 kDa range (Paschkowsky et al. 2016). APP cleavage occurred independently of other types of proteases, and the catalytically inactive RHBDL4 mutant bound APP and led to its stabilisation, which further supported the role of APP as a RHBDL4 substrate. Importantly, when processed by RHBDL4, APP cleavage by canonical proteases was reduced, and so the amount of Aß peptides produced was reduced too (Paschkowsky et al.

2016). However, when APP was expressed endogenously and only RHBDL4 was transfected into the cells, the cleavage was hardly detectable, except for one cleavage product, which is almost undistinguishable from the product of cleavage by α-secretase (Paschkowsky et al. 2016). The physiological relevance of APP as a substrate of RHBDL4 is thus questionable.

Nonetheless, APP is a good model substrate for further study of the RHBDL4 cleavage mechanism. Interestingly, APP cleaved by RHBDL4 is not degraded via ERAD, even when the TMH of APP is mutated to include the aforementioned di-basic degron motif or if it is completely replaced by the TMH of a RHBDL4 substrate, which is degraded by ERAD (Recinto et al. 2018). Therefore, there has to be another feature in the luminal domain of the substrate, which determines what will happen after the cleavage.

Moreover, the pattern of APP cleavage by RHBDL4 is also interesting, because the size of cleavage fragments implies that RHBDL4 cleaves inside the luminal domain of APP (Paschkowsky et al. 2016). This is a bit unexpected for an intramembrane protease, but not impossible – the active site of RHBDL4 is open to the ER lumen. In addition, cleavage inside the luminal domain occurs also when RHBDL4 cleaves the α-chain of pre-T cell receptor, an ERAD substrate (Fleig et al. 2012). This implies that there are several motifs in proteins which are recognized by RHBDL4 and which influence the position of cleavage and the fate of the resulting cleavage products.

RHBDL4 cleavage of APP is dependent on the cholesterol level in cells (Paschkowsky et al. 2018). This could be due to direct cholesterol to RHBDL4 binding, as a potential cholesterol binding site was found in RHBDL4, which could influence the topology of the enzyme. Regulation of rhomboid activity by membrane characteristics could be envisioned but confirmation of the cholesterol effect is needed.

5 Proximity proteomics and rationale for the study

As documented in Chapter 3, IMPRs are very relevant to the immune system. Interestingly, initial assessment of the phenotype of RHBDL4 deficient mice suggests significant alterations in abundancies of T and B cell subtypes

(Bohacova & Strisovsky, unpublished data). How RHBDL4 mediates this immune phenotype is so far speculative. It might be connected with the highly secretory nature of some immune cell types and the possible role of RHBDL4 in secretion, membrane protein quality control and the UPR. Or, it might be connected with the pro-proliferative and anti-apoptotic effects observed in cancer studies. Possibly, a yet unresolved substrate of RHBDL4 might be involved. The desire to clarify this issue prompted us to focus on the physiological role of RHBDL4.

The biological function of RHBDL4, although extensively studied since 2008, is still unclear and without consensus. Therefore, it is particularly important to try to link known facts and verify them independently. One possible approach to get insight into the molecular processes in which RHBDL4 is involved, is to directly analyze the proteome of the compartment(s) where it normally resides. To this end, we decided to use proximity proteomics approaches, in particular, the BioID2 and APEX2 methods. These powerful strategies are based on biotin labelling in intact living cells and subsequent mass-spectrometry (MS) detection (Kim et al. 2016; Hung et al. 2016).

These methods have been used extensively to map the proteome of various cellular membranes and compartments as well as the interactomes of membrane proteins (Hwang et al. 2016; Varnaitė & MacNeill 2016; Hung et al. 2017; Lobingier et al. 2017; Paek et al. 2017; Zuzow et al. 2018). They have many advantages compared to traditional methods for seeking interaction partners of a membrane protein. Specifically, there is no need for purification of the membrane protein, which is in general a problematic process. Also, the interactions are captured in the natural environment of the cell and therefore unaffected by detergents or protein unfolding. Moreover, even very transient and weak interactions can be captured by these approaches. By applying proximity proteomics to rhomboid protease, we can visualize the specific sub-compartment of the cell, where it is localised, its interaction partners and even potential substrates of the rhomboid. However, as for all screening methods, the results have to be further confirmed.

The BioID2 is an engineered promiscuous biotin ligase (Kim et al. 2016). It catalyses enzymatic ligation of biotin to lysines in proteins in the proximity of

target protein, to which the BioID2 is genetically fused. By the length of a linker between BioID2 and the target, the biotinylation radius can be adjusted. The biotinylation requires supplementation of cell medium with free biotin and it lasts for many hours (16-24h) yielding a history of protein-protein interactions over this time period (Roux et al. 2018). Its predecessor BioID is a biotin ligase derived from *E. coli* (Roux et al. 2012). The upgraded BioID2 is derived from smaller biotin ligase from *A. aeolicus* and as a result it has higher activity and it influences the intracellular localisation of the target protein less (Kim et al. 2016).

The so-called APEX2 is an engineered ascorbate peroxidase catalysing the oxidation of phenol derivatives to phenoxyl radicals (Rhee et al. 2013). Its predecessor APEX was engineered from soybean ascorbate peroxidase (Rhee et al. 2013) and APEX2 is significantly improved in its enzymatic activity due to additional A134P mutation (Lam et al. 2014). Because of that, APEX2 activity is detectable even when its expression is low in cells, enabling examination of proteins without the need for overexpression. The biotin labelling in proximity of the target occurs through biotin-phenol radicals, produced by APEX2, which are short lived, with a half-time shorter than 1 ms, and interact with electron-rich amino acids, which in practice means mainly tyrosine (Hung et al. 2016). The lipid membrane cannot be crossed by the biotin radical. Compared to BioID2, the kinetics of labelling by APEX2 are much faster. The radical biotinylation can last only tens of seconds (Rhee et al. 2013) yielding a kind of 'snap shot' of the cell. This is particularly advantageous when seeking after dynamic interactions. The APEX2 method is very sensitive, but this has the downside of lower specificity. The higher background signal is filtered out based on quantitative comparison to control experiments (Hung et al. 2016).

6 Materials

6.1 Chemicals and Reagents

- PCR master mix: NEB #MO4925
- Restriction enzymes (EcoRI, NcoI, NotI, NheI): NEB (#R3101S, #R3193S, #R3189G, #R3131S)
- Cut Smart: NEB #B7204S
- Orange G: Sigma #03756
- Agarose: Serva #120274
- GelRed: Biotium #41003
- HyperLadder 1kb or 50bp: Bioline
- Gel DNA extraction kit: QIAGEN #28706
- $MgCl_2$: Affymetrix #78641
- NAD: Sigma #10127981001
- Polyethylene glycol 8000: Affymetrix #19966
- dNTP set: Sigma #11969064001
- T5 exonuclease: NEB #M0363S
- Taq DNA ligase: NEB #NEM0208L
- High-Fidelity DNA polymerase: NEB #NEM0530S
- LB Broth, powder: Sigma #L3022
- LB Broth with agar, powder: Sigma #L2897
- Ampicillin: Biotika #1808005
- Miniprep kit: QIAGEN #27106
- Midiprep kit: #740410.1
- DMEM: Gibco #2026729
- FBS: Gibco #10270106
- PBS: Gibco #70011036
- Lipofectamin 3000®: Invitrogen #1882752
- Cumate: Biosciences #PBQM100A-1
- Puromycin: Gibco #A1113803
- FuGENE® HD: Promega #E2692
- 0.45 µm and 0.22 µm PES membrane: ROTH #P667.1 and #P668.1
- Stericup 500 ml 0.22 PES: Sigma #SCGPU05RE
- Polybrene: Sigma #H9268
- Tris: Sigma #T1503
- LDS: Sigma #L9781
- Glycerol: Penta #56-81-5
- ß-mercaptoethanol: Sigma #M3148
- Bromophenol blue: Sigma #B5525
- cOmplete™ protease inhibitors: Roche #33576900
- Pierce universal nuclease: Thermo #88701
- $MgCl_2$: NEB #B0510A
- 40% Acrylamide: BioRad #1610140
- 2% Bis-acrylamide:BioRad #1610142
- SDS: BioRad #1610301
- TEMED: Roth #2367.1
- APS: Sigma #248614
- Tris: Sigma #T1503
- Glycine: Sigma #G8898
- SDS: Sigma #L3771
- NaCl: Penta #7647-14-5
- Color protein standard: NEB #P7712
- Methanol: Sigma #67-56-1
- Immobilon®-FL: Sigma #IPVH00010
- Immobilon®-P PVDF: Sigma #IPFL00010
- Blocker™ Casein in TBS: Thermo #37532
- REVERT™ Total Protein Stain: LI-COR #926-11010
- SuperSignal™ West Femto substrate: Thermo #34096
- DMEM for SILAC deficient in Arg and Lys: Thermo #88364
- Dialysed FBS: Thermo #26400044
- Biotin-phenol: Iris-Biotech #41994-02-9
- Hydrogen peroxide solution: Sigma #H1009
- Sodium ascorbate: Sigma #11140
- Trolox: Sigma #238813
- Sodium azide: Sigma #71290
- NaCl (APEX2): Sigma #S7853
- SDS (APEX2): Sigma #L6026
- Sodium deoxycholate: Sigma #D6750
- Triton X-100: Sigma #T8787
- Urea: Sigma #51456
- Streptavidin magnetic beads: Thermo #88817
- Biotin: Sigma #B4639
- DL-Dithiothreitol: Sigma #D9779
- Coomassie brilliant blue G-250: Sigma #B0770
- Aluminium sulfate-(14-18)-hydrate: Sigma #450308
- Ethanol: Penta #70390
- Orthophosphoric acid: Penta #19200
- Pierce™ 660nm Protein Assay: Thermo #22660
- Detergent compatibility: Thermo #22663
- KCl: Penta #16210
- Na_2CO_3: Penta #28010
- Precast gradient PAA gel: BioRad #4561094
- **Heavy AAs**
- $^{13}C_6$ $^{15}N_2$ L-Lysine HCl: Silantes #211604102
- $^{13}C_6$ $^{15}N_4$ L-Arginine HCl: Silantes #201604102
- **Medium AAs**
- 4,4,5,5-D_4 L-Lysine 2HCl: Silantes #211104113
- $^{13}C_6$ L- Arginine HCl: Silantes #201204102
- **Light AAs**
- L-Lysine: Sigma #L9037
- L-Arginine: Sigma #A6969
- **Antibodies**
- HSC70: Santa Cruz #K1814
- RHBDD1: Sigma #A96509
- Myc tag: CST #2278S
- GFP: CST #2956S
- HA tag: CST #3724
- FLAG tag: CST #2368S
- EGFR: Millipore #06-847
- anti-Rabbit-HRP: Santa Cruz sc-2004
- m-IgG BP-HRP: Santa Cruz sc-516102
- anti-Rabbit-IRDye 800CW: LI-COR #925-32211/3
- Streptavidin-IRDye: LI-COR #926-32230 (1/5000 dilution)
- Streptavidin-HRP: Invitrogen #S911

* Mass spectrometry equipment and reagents are not included as it was performed by specialized facility at IOCB

6.2 Solutions and buffers

Agarose gel
1% (w/v) agarose in TAE buffer; 0.006% (v/v) GelRed
Orange G loading buffer 5x
15% (v/v) Glycerol; 0.2% (w/v) Orange G
TAE buffer
40 mM Tris; 20 mM acetic acid; 1 mM EDTA, pH 8.4
Gibson assembly master mix
200 mM Tris-HCl pH 7.4; 200 mM $MgCl_2$; 20 mM dGTP; 20 mM dCTP; 20 mM dTTP; 20 mM dATP; 200 mM Dithiothreitol; 25% (w/v) Polyethylene Glycol 8000; 50 mM NAD; 1 unit T5 Exonuclease; 167 units Taq DNA Ligase; 13 units High Fidelity DNA Polymerase
LB medium
LB Broth dissolved in sterile water according to manufacturer instructions and autoclaved
Ampicillin agar plate for bacteria culture
LB Broth with agar dissolved in sterile water according to manufacturer instructions and autoclaved; Ampicillin 100 µg/ml
LDS sample buffer 4x:
277.8 mM Tris-HCl pH 6.8; 4.4% (w/v) lithium dodecyl sulphate (LDS); 44.4% (v/v) Glycerol; 10% (v/v) ß-mercaptoethanol; 0.02% (w/v) Bromophenol blue
LDS lysis buffer:
1x LDS sample buffer supplemented with cOmplete™ protease inhibitors (Roche), 5 mM MgCl2, universal nuclease (1/1000 dilution)
10% Polyacrylamide (PAA) gel:
Per one resolving gel (5 ml):
1.21 ml 40% Acrylamide; 0.83 ml 2% Bis-acrylamide; 1.25 ml 1.5 M Tris-HCl pH 8.8; 50 µl 10% (w/v) sodium dodecyl sulphate (SDS); 1.61 ml water; 2 µl TEMED; 50 µl 10% (w/v) ammonium persulfate (APS)
Per one stacking gel (2 ml):
0.19 ml 40% Acrylamide; 0.1 ml 2% Bis-acrylamide; 0.5 ml 1 M Tris-HCl pH 6.8; 20 µl 10% (w/v) SDS; 1.17 ml water; 1.6 µl TEMED; 20 µl 10% (w/v) APS
SDS-PAGE running buffer:
25 mM Tris; 192 mM Glycine; 0.1% (w/v) SDS
Western blot (WB) running buffer:
5% (v/v) Methanol; 12.5 mM Tris; 96 mM Glycine; 0.005% (w/v) SDS
TBS:
50 mM Tris; 137 mM NaCl; pH 7.6
TBS-Tween:
TBS solution with 0.1% (v/v) Tween® 20
SILAC medium:
0.798 mM Lysine, 0.398 mM Arginine, DMEM for SILAC, dialysed fetal bovine serum (FBS); assembled and filtered through polyethersulfone (PES) 0.22 µm membrane in Stericup
EDTA
0.02% (w/v) EDTA; 0.4% (w/v) Phenol Red
Trypsin solution
0.01% (w/v) EDTA; 0.2% (w/v) Phenol Red; 0.25% (w/v) Trypsin
PBS with quenchers 2x:
PBS supplemented with 20 mM sodium ascorbate, 10 mM Trolox and 20 mM sodium azide
Harvesting solution:
EDTA and PBS with quenchers 2x are mixed 1:1
RIPA buffer 2x:
100 mM Tris-HCl; 300 mM NaCl; 1% (w/v) SDS; 1% (w/v) sodium deoxycholate; 2% (v/v) Triton X-100; pH 7.5
RIPA buffer for cell lysis:
1x RIPA buffer supplemented with cOmplete™ protease inhibitors (Roche), universal nuclease (1/1000 dilution), 10 mM sodium ascorbate, 5 mM Trolox and 10 mM sodium azide
RIPA-U:
1x RIPA buffer supplemented with 4 M Urea
SDS sample buffer 6x:
350 mM Tris-HCl pH 6.8; 10% (w/v) SDS; 50% (v/v) Glycerol; 10% (v/v) ß-mercaptoethanol; 0.05% (w/v) Bromophenol blue
Elution buffer:
3x SDS sample buffer, 4 mM Biotin, 20 mM Dithiothreitol
Coomassie brilliant blue G solution: 0.02% (w/v) Coomassie brilliant blue G-250, 5% (w/v) aluminium sulfate-(14-18)-hydrate, 10% (v/v) ethanol, 2% (v/v) orthophosphoric acid

7 Methods

7.1 Cloning

To test protein stability, expression and correct cellular localization, RHBDL4 was cloned into pcDNA3.1 vector together with BioID2 or APEX2 enzyme (labelling enzyme) and an epitope tag (Table1). The whole fusion protein was cloned in frame, to encode this, in order from the N-terminus: tag, labelling enzyme, RHBDL4. In some cases, a peptide linker (GGGGSGGGGSGGGGS) was inserted between labelling enzyme and RHBDL4. Sequences of eGFP-APEX2 (#49385), APEX2-NES (#49386) and BioID2 (#74223) were obtained from Addgene. Sequence of human RHBDL4 (and RHBDL4 S144A) was derived from constructs characterized previously (Johnson et al. 2017).

To achieve a stable expression of the fusion protein in cells, sequence of HA-APEX2-linker-RHBDL4 was cloned into pM6P retroviral vector with a puromycin resistance gene for selection (Oikonomidi et al. 2018). Alternatively, into PiggyBac transposon vector with inducible protein expression regulated by cumate (Mullick et al. 2006), also with the puromycin resistance gene. For a control experiment, APEX2-NES (nuclear export sequence) was cloned into pM6P retroviral vector. PiggyBac expression vector (PBQM812A-1) and Super PiggyBac transposase expression vector (PB210PA-1) were purchased from System Biosciences. Vector pM6P was a kind gift of Dr. Felix Randow (MRC Laboratory of Molecular Biology, Cambridge, UK).

Inserts were amplified by PCR using PCR master mix according to manufacturer instructions. Primers were designed for cloning by isothermal (Gibson) assembly (Gibson et al. 2009) and purchased from GATC-Eurofins Genomics. Annealing temperatures were calculated via NEB T_m calculator. Amplified inserts were analysed via agarose gel electrophoresis. The whole PCR reaction mix was supplemented with Orange G loading buffer and separated on agarose gel in TAE buffer at constant voltage – usually 100V for 45 min. Clear DNA bands (visualized by GelRed) of correct size were cut out, weighed in Eppendorf tube and the DNA was purified from the gel via Gel Extraction Kit according to manufacturer instructions. Expression vectors (into which the inserts were to be cloned) were cleaved by appropriate restriction enzyme according to manufacturer instructions (NEB). Opened vectors were analysed via agarose gel electrophoresis and purified from the gel in the same way as amplified inserts. The ligation of appropriate inserts and expression vector was performed via Gibson assembly (Gibson et al. 2009).

Assembled vectors were amplified in *E. coli*. The vectors were transformed into competent *E. coli* strain either by electroporation or by a heat shock (Green & Sambrook 2012). Bacteria were then shaken in LB medium in 37°C for 30 min to allow expression of the antibiotic resistance gene. Subsequently, they were plated on agar plates supplemented with antibiotic and incubated overnight in 37°C. The *E. coli* strains used were competent DH5α (pcDNA3.1 vectors) and Stbl3 (pM6P and PiggyBac vectors). Several well-separated colonies were picked from the agar plate after overnight incubation by sterile tip and resuspended in 3 ml of sterile LB medium supplemented with antibiotic. The medium with bacteria was shaken overnight in 37°C. Grown bacterial culture was then spun down and expression vectors were isolated via Miniprep kit according to manufacturer instructions. The identity of assembled constructs was verified by Sanger sequencing of the insert region at GATC-Eurofins Genomics, and sequencing primers were designed according to their

instructions. The correct constructs were then further amplified in *E. coli* in bigger media volume to get sufficient amount of DNA for transfection in mammalian cell culture. The vectors were purified from bacteria via NucleoBond Xtra Midi according to manufacturer's instructions.

Expression vectors containing BioID2 and APEX2-NES were cloned by lab technician Petra Rampírová, the others were cloned by myself (Table1). APP (amyloid precursor protein, a model RHBDL4 substrate) cloned with N-terminal Myc tag in pcDNA3.1, was a kind gift of Lisa Műnter (McGill University, Canada). Expression vector encoding the endoplasmic reticulum-marker Sec61b-mCherry was obtained from Addgene (#49155).

Table 1: Expression vectors

Insert	Vector	Purpose of the vector
Myc-BioID2-hRHBDL4 S144A	pcDNA3.1	Stability, expression test
Myc-BioID2-linker-hRHBDL4 S144A	pcDNA3.1	Stability, expression test
Myc-BioID2-hRHBDL4	pcDNA3.1	Stability, expression test
Myc-BioID2-linker-hRHBDL4	pcDNA3.1	Stability, expression test
eGFP-APEX2-hRHBDL4	pcDNA3.1	Stability, intracellular localisation test
eGFP-APEX2-linker-hRHBDL4	pcDNA3.1	Stability, intracellular localisation test
eGFP-APEX2-hRHBDL4 S144A	pcDNA3.1	Stability, intracellular localisation test
eGFP-APEX2-linker-hRHBDL4 S144A	pcDNA3.1	Stability, intracellular localisation test
HA-APEX2-linker-hRHBDL4	pM6P.puro	Stable expression, constitutive
HA-APEX2-linker-hRHBDL4 S144A	pM6P.puro	Stable expression, constitutive
HA-APEX2-linker-hRHBDL4	PiggyBac.puro	Stable expression, tunable
HA-APEX2-linker-hRHBDL4 S144A	PiggyBac.puro	Stable expression, tunable
FLAG-APEX2-NES	pM6P.puro	Stable expression, control experiment

7.2 Cell culture and transfection

HCT116 cells (adherent human colorectal cancer cell line, acquired from ATCC) were cultured in DMEM 10% FBS (vol/vol). They were regularly passaged by trypsinization, checked for mycoplasma infection and maintained at low passage number (passaged for 2 months at maximum). Wild type HCT116 cells (WT) and RHBDL4 deficient HCT116 cells (KO) were used in this study. The KO cell line was prepared by alumnus postdoctoral researcher Aymeric Morlé by CRISPR/Cas9, targeting of the first coding exon of RHBDL4, and verified by DNA sequencing of the chromosomal locus and western blot.

Transfection was performed by Lipofectamin® 3000 according to manufacturer protocol. Cells were examined 36 – 48 h after transfection. The ratio between total amount of transfected DNA and volume of the transfection reagent was kept constant.

7.2.1 Generation of cell lines with stable expression of RHBDL4 in fusion protein

PiggyBac transposon system (Mullick et al. 2006) enables generation of cell lines with stable and tunable expression of desired protein. RHBDL4 deficient HCT116 cells were co-transfected with PiggyBac expression vector (0.7 µg per well of a 6-well plate) and Super PiggyBac Transposase (0.3 µg per well of a 6-well plate) and incubated for 72 h. Cells with chromosomally integrated constructs were then selected by puromycin (0.5 µg/ml), confirmed by WB and cryo-preserved in liquid nitrogen. Expression of the introduced

protein is controlled by cumate-inducible promoter/operator in cumate concentration dependent manner. For induction, the cumate was dissolved in dimethyl sulfoxide (DMSO) at a concentration of 300 mg/ml and added to DMEM to reach the concentration of 10 µg/ml. Cells were incubated in 0-10 µg/ml cumate for 48 h and then examined.

Alternatively, the cell line with stable expression of RHBDL4 fusion protein was generated by retroviral delivery of expression cassette (Randow & Sale 2006). The resulting expression of the desired protein is constitutive. Virus was prepared in HEK cells, which were co-transfected with retroviral expression vector pM6P (1.5 µg per well of a 6-well plate) and virus packaging vector pCL.10A1 (1.5 µg per well of a 6-well plate) by FuGENE® HD according to manufacturer's protocol. Viral particles were produced by HEK cells into DMEM 10% FBS medium, which was collected 2 days later. Viral medium was filtered through 0.45µm PES membrane, diluted by additional DMEM 10% FBS medium and supplemented with polybrene to a final concentration of 10 µg/ml. RHBDL4 deficient HCT116 cells plated a day before, were infected by sequentially diluted viral media to decrease the multiplicity of infection (MOI) and thus reduce the copy number of integrated retroviral construct and resulting expression of the introduced protein. After 24 h, cells were washed with PBS and clones with chromosomally integrated constructs were selected by puromycin (0.5 µg/ml). Selected cells were grown, verified by WB and cryo-preserved in liquid nitrogen.

7.3 Confocal light microscopy

Live cells transiently expressing proteins with fluorescent tag were examined by confocal microscope Zeiss LSM 780 or a custom-built one (kindly provided by the group of Dr. Josef Lazar, IOCB). The obtained images were processed in ImageJ Fiji.

7.4 Western blot analysis

7.4.1 Cell lysis

Grown cells were washed twice by cold PBS in a culture dish. Having decanted the PBS, cold LDS lysis buffer was added directly into the dish (150 µl per well of a 6-well plate). The lysis buffer was distributed on the bottom of the dish by tilting, incubated for 1 min at a room temperature and then collected into an Eppendorf tube, scraping and flushing the bottom thoroughly by a tip. The sample was then either kept on ice and directly used for analysis, or frozen and kept at -20°C until needed.

7.4.2 SDS-PAGE (Laemmli 1970; Brunelle & Green 2014)

The cell lysate sample was heated to 50°C for 10 min, with agitation in an Eppendorf Thermomixer. A 10% PAA gel was prepared in advance and kept in wet paper towel in refrigerator until needed. Wells of the gel were washed by SDS-PAGE running buffer to remove non-polymerized PAA. The samples were loaded in the gel and a coloured protein standard was loaded next to them. The gel was run in the equipment from BioRad at constant voltage. At first, the gel was run at 80 V until the proteins reached the resolving portion of the gel. Then the voltage was increased up to 180 V and kept constant until bromophenol blue, marking the leading edge of migrating proteins, migrated out of the bottom of the gel. The gel was then removed from the glass plates and further processed.

7.4.3 Western blotting (LeGendre 1990; Towbin et al. 1979)

The gel from SDS-PAGE was washed in WB running buffer. A transfer membrane (Immobilon®-FL PVDF or Immobilon®-P PVDF) was activated in methanol and equilibrated in WB running buffer for several minutes. A WB sandwich was assembled, carefully removing any bubbles and tightly packed. The transfer of proteins from the gel to the membrane was performed in BioRad wet blotting unit at constant voltage of 80 V for 1 h 30 min, with a cooling insert to decrease temperature during electrophoresis.

7.4.4 Membrane labelling and protein visualisation

After the transfer, the membrane was removed from blotting sandwich and washed in distilled water. Then, either the total protein content on the membrane was revealed, or the membrane was directly blocked in Casein solution. The total protein content was determined by REVERT™ Total Protein Stain according to manufacturer instructions (LI-COR) and scanned by Odyssey® CLx imaging system. After the scanning, the membrane was reactivated in methanol and blocked in Casein solution. The blocking was performed at room temperature for at least 30 min. Primary antibodies were added to the blocked membrane at an appropriate dilution and incubated at 4°C overnight. The primary antibodies were usually diluted 1/1000 in the Casein blocker solution with 0.1% Tween® 20, using 3 ml total volume per membrane.

The next day, the membrane was washed in TBS-Tween (5 times, 8 min each wash) to remove excessive antibody. Then, the membrane was washed in TBS to remove remaining detergent and incubated with a secondary antibody for 2 h at room temperature. The secondary antibody was usually diluted 1/10000 in Casein blocker solution with 0.1% Tween® 20, using 5 ml total volume per membrane. If the secondary antibody was conjugated to a fluorescent dye, the membrane was incubated in the dark. After incubation, the membrane was washed in TBS-Tween in the same way as after the incubation with primary antibody.

Finally, a signal on the membrane was revealed and recorded. Fluorescent signal was scanned from dried membrane on Odyssey® CLx imaging system. Chemiluminescent signal from HRP (horseradish peroxidase) was revealed by addition of Pierce™ ECL Western Blotting Substrate and recorded on UVP ChemiDoc-It™ 600 Imaging system.

7.4.5 Quantification of proteins by Western blot

The measurement of signal intensity of individual bands was done in Image Studio™ Lite software (LI-COR). For comparison between individual samples, the intensity of the given protein band was normalized to the intensity of total protein in the corresponding gel lane. Graphs were plotted in Excel (Microsoft).

7.5 APEX2 proximity proteomics

APEX2 experiment was performed based on a published protocol (Hung et al. 2016). After several rounds of testing, we included some adjustments. I will describe my modified protocol, which was used for the experiments followed by mass spectrometry detection, shown in the Result section.

7.5.1 SILAC-based experiment

For the purpose of this experiment, the SILAC (stable isotope labelling with amino acids in cell culture) approach was employed (Ong et al. 2002). Cells were grown in SILAC medium for 11 days (Table 2) to achieve isotopic labelling. All cells well cultured and passaged simultaneously. Successful labelling of cellular proteins was confirmed by mass spectrometry (assessed as 97-98%).

7.5.1.1 Biotinylation in intact cells

SILAC-labelled cells were expanded into T75 culture flask per condition. When confluence was reached, 13 ml of appropriate SILAC medium were supplemented with 500µM biotin-phenol and incubated on cells for 1h at 37°C. Then, freshly diluted H_20_2 was added to the medium to reach 1 mM concentration. The medium was briefly agitated and cells were incubated in it for 1 min at room temperature. The medium was then removed by suction and cells were thoroughly washed by 13 ml of PBS supplemented with quenchers of radicals. Washing was repeated 2 times with PBS only and 2 times with PBS and quenchers, 1 min each wash, at room temperature. For the incubation and washing of cells, an inversion technique described in the original protocol was used. Briefly, to keep the timing same for all culture flasks, they were all at once inverted when the incubation/washing was over or when it begun. Cells were then covered by 5 ml of harvesting solution and incubated for 5 min at 37°C. Cells were then easily detached by pipetting and collected into a Falcon tube. The cells were pelleted at 3000×g, 4°C, 10 min and lysed in 400 µl of RIPA buffer for cell lysis. The lysed sample was kept on ice for 5 min and centrifuged at 15000×g, 4°C, 10 min. A clear supernatant was transferred into a new tube and flash frozen by liquid nitrogen. A small part of the clarified lysate was then supplemented with SDS sample buffer and directly used for WB analysis of biotinylation reaction. The rest was applied to streptavidin-coated magnetic beads for specific isolation of biotinylated proteins.

Table 2: SILAC labelling in APEX2 proteomics

Cell line	SILAC label	Expected biotinylation pattern
HCT116 KO: APEX2-RHBDL4	Heavy (H)	Biotinylation mainly in proximity of RHBDL4
HCT116 KO: APEX2-RHBDL4	Heavy (H)	Biotinylation mainly in proximity of RHBDL4 - duplicate
HCT116 KO: APEX2-NES	Medium (M)	Non-specific biotinylation of cytosolic proteins
HCT116 KO: APEX2-NES	Medium (M)	Non-specific biotinylation of cytosolic proteins - duplicate
HCT116 KO	Light (L)	No artificial biotinylation due to the absence of APEX2
HCT116 KO: APEX2-RHBDL4	Light (L)	No artificial biotinylation due to the absence of H_2O_2

7.5.1.2 Isolation of biotinylated proteins

Before the purification of biotinylated proteins on the streptavidin-coated magnetic beads, protein concentration of individual samples was determined by Pierce™ 660nm Protein Assay according to manufacturer's protocol. Cell lysates with heavy (H), medium (M) and light (L) SILAC labels were mixed together in 1:1:1 ratio of total protein (specifically, 1.8 mg of protein from each condition were used). Samples were prepared in

duplicates (Table 2) – two H, two M and two L, leading to two mixed H+M+L samples. The only difference between duplicates is the nature of the L sample, because two different ways of generating a negative control were used.

At first, streptavidin-coated magnetic beads were washed twice by RIPA-U. The washing was performed in an Eppendorf tube, where the beads were resuspended in the washing solution, rotated for 1 min, then pelleted on a magnetic rack and the supernatant was removed. After the initial washing step, the mixed (H+M+L) cell lysate was applied to the beads and supplemented with 3.2 M Urea (83 µl of beads were used per 1 mg of protein in mixed cell lysate). Incubation was performed overnight at 4°C on a rotator.

Unbound proteins were collected after incubation and termed 'flow-through'. To wash out nonspecific binders from beads, several washes were performed: 2× RIPA-U, 2× 1 M KCl, 2× 0.1 M Na_2CO_3, 2× 4 M Urea in 10 mM Tris-HCl pH8, 2× RIPA-U. For each wash, 1ml of solution was used, beads were dispersed, rotated for 5 min and pelleted for 3 min on a magnetic rack. To elute the captured proteins from beads, they were boiled at 95°C in 50 µl of elution buffer for 15-20 min. The sample was then briefly vortexed, cooled on ice, briefly centrifuged to bring down condensation and pelleted on a magnetic rack. The supernatant was transferred into a clean tube and termed as 'eluate'. The eluate was then loaded into one well of pre-casted gradient PAA gel and subjected to SDS-PAGE. Resulting gel was agitated 3 h with Coomassie Brilliant blue G solution to reveal protein bands and used for subsequent mass spectrometry analysis. A small portion of eluate was kept for WB analysis of the purification.

7.5.2 Label free experiment

The biotinylation in unlabelled HCT116 cells and the isolation of biotinylated proteins were performed as in the SILAC-based experiment. However, importantly, the cell lysates were not mixed, but processed individually. They were also individually analysed by mass-spectrometry and compared at the level of identified proteins.

7.6 Mass spectrometry (LC-MS/MS)

Measurement performed by service MS laboratory of Josef Cvačka at IOCB.

7.6.1 Sample preparation

Each gel lane to be analysed was cut into 12 (13) slices. All of them separately were destained with 25 mM ammonium bicarbonate in 50 % acetonitrile at 30 °C for 30 min. Each gel slice was treated with 50 µl of 5 mM DTT in 50 mM ammonium bicarbonate (65 °C, 30 min). Then 50 µl of 15 mM iodoacetamide in 50 mM ammonium bicarbonate was added and incubated at room temperature for 30 min in the dark. Excess liquid was removed, gel slices were washed with 50 mM ammonium bicarbonate, washed twice by acetonitrile and dried afterwards in SpeedVac. Proteins trapped in the gel slices were digested with 1 µl (0.1 ug) of trypsin solution in 50 mM ammonium bicarbonate at 37 °C for 16 hours. Each gel slice was then transferred into a new microtube, and peptides were extracted with 50 µl of 2 % TFA for 15 min. Afterwards, 50 µl of 60 % acetonitrile was added and incubated for 15 min and sonicated. The solution was removed, the extracted tryptic peptides were dried at the SpeedVac and dissolved in 20 µl of 0.1 % formic acid.

7.6.2 Label-free experiment

Peptides dissolved in 0.1 % formic acid were analyzed on UltiMate 3000 RSLCnano system (Dionex) coupled to a TripleTOF 5600 mass spectrometer with a NanoSpray III source (AB Sciex). The instrument was operated with Analyst TF 1.6 software (AB Sciex). After the injection, the samples were trapped and desalted with 2 % acetonitrile in 0.1 % formic acid at flow rate of 5 µL/min on Acclaim PepMap100 column (5 µm, 2 cm×100 µm ID, Thermo Scientific). Eluted peptides were separated using Acclaim PepMap100 analytical column (3 µm, 15 cm×75 µm ID, Thermo Scientific). The 70 min elution gradient at constant flow of 300 nl/min employed two phases in variable proportions, phase B (0.1 % formic acid in 99.9 % acetonitrile) and phase A (0.1 % formic acid). The elution gradient started at 5 % B for the first 5 min, then a gradient elution proceeded from 5 % to 50 % B over 55 min, from 50 % to 99 % B over 5 min, stayed at 99 % B for 10 min and descended to 5 % B and remained there for 15 min. TOF MS mass range was set to 350–1600 m/z. In MS/MS mode the instrument acquired fragmentation spectra with m/z ranging from 100 to 2000.

The resulting data were used for searching against *Homo Sapiens* database by combination of three open source searching algorithms - X!tandem, OMSSA and MS-GF+ in Galaxy interface and the proteins were identified. The search was set by choosing carbamidomethyl as static modification, oxidation as dynamic modification and trypsin as the digestion agent. Working with a peptide list generated by the search software, all peptides and corresponding spectral counts (peptide spectrum matches, PSM) were assigned to each identified protein represented by a unique gene name in Excel. Whenever it was not possible to unambiguously determine to which protein the peptide belonged, proteins were grouped. By integration of the PSM of all peptides belonging to one protein, a spectral count (abundance) of the protein was obtained. By dividing the protein abundance from the real experiment by the corresponding abundance from the control experiment, quantitative ratios were obtained. Proteins detected only in the real experiment and not in the control, were assigned an arbitrary cytosolic abundance of 0.01, which was randomly chosen as a low non-zero number. Based on the calculated quantitative ratios and the protein abundance in real experiment, a cut-off was set to distinguish RHBDL4 proximity proteome from the background. The cut-off was set to protein abundance (total number of peptides identified) equal to 4 or higher, and the quantitative ratio equal to 2 or higher.

7.6.3 SILAC-based experiment

Peptides dissolved in 0.1 % formic acid were analyzed on UltiMate 3000 RSLCnano system (Dionex) coupled to an Orbitrap Fusion Lumos Tribrid mass spectrometer with an EASY-Spray source (ThermoFisher Scientific). The instrument was operated with Xcalibur software (ThermoFisher scientific). After the injection, the samples were trapped and desalted with 2 % acetonitrile in 0.1 % formic acid at flow rate of 30 µL/min on Acclaim PepMap100 column (5 µm, 5 mm×300 µm ID, ThermoFisher Scientific). Eluted peptides were separated using Acclaim PepMap100 analytical column (3 µm, 15 cm×75 µm ID, ThermoFisher Scientific). The 65 min elution gradient at constant flow of 400 nl/min employed two phases in variable proportions, phase B (0.1 % formic acid in 80 % acetonitrile) and phase A (0.1 % formic acid). The elution gradient started at 5 % B for the first 1 min, then proceeded from 5 % to 35 % B over 52 min, from 35 % to 99 % B over 5 min, stayed at 90 % B for 10 min and descended to 5 % B and remained there for 7 min. Orbitrap was set as the mass analyser for data acquisition with resolution of 120000, selecting

peptides in range 350-2000 m/z with charge state from 2+ to 6+. The masses being for 10s over threshold level 1.5×10^4 were excluded. MS2 was recorded on Orbitrap in the same range with resolution 30000 using HCD activation type with isolation window of 1.5 Da and set collision energy of 30%.

MaxQuant was used for protein identification and quantification from raw spectra using *Homo sapiens* database. The search was set by choosing carbamidomethyl as static modification, oxidation of methionine, biotin-phenoxyl of histidine/tryptophan/tyrosine, label 13C(6)15N(2) of lysine, label 13C(6)15N(4) of arginine, label 2H(4) of lysine and label 13C(6) of arginine as dynamic modifications. Trypsin was set as the digestion agent, with 2 allowed miscleavages. Both SILAC-based experiments (duplicates) were evaluated by Andromeda search algorithm in MaxQuant, allowing 20 ppm peptide tolerance in the first search and 4.5 ppm peptide tolerance in the main search. Unique and razor peptides were selected for quantification, minimal label ratio count was set as 2 and re-quantification was enabled. Missing peptide intensities were artificially imputed into the data as the normal distribution of background signal, using Perseus. This step was included to enable quantification of all proteins detected. The quantitative SILAC ratios were used to distinguish the RHBDL4 proximity proteome from background. The cut-off was set based on quantitative SILAC ratios of protein VCP, which is an established interaction partner of RHBDL4.

8 Results

8.1 Fusion protein BioID2-RHBDL4 is unstable in cells.

The first proximity proteomics approach that I tested was BioID2 (Kim et al. 2016), which uses an engineered biotin ligase from *Aquifex aeolicus* with very broad substrate specificity, able to biotinylate indiscriminately proteins in the vicinity of RHBDL4 on Lys residues. First it was necessary to test whether fusion of BioID2 (26.6 kDa protein) to RHBDL4 (35.6 kDa protein) yields a stable protein expressed in cells, because any BioID2 mislocalised from RHBDL4 would yield a large background signal. Several diversely arranged fusion protein constructs were generated, where BioID2 was connected to the N-terminus of catalytically active (WT) or inactive (S144A) RHBDL4 either directly or through a peptide linker [$(G_4S)_3$] (Fig.11a). However, none of these fusion proteins was stable in cells and species shorter than the full-length form appeared on western blots (Fig.11b). When proteasomal degradation was blocked by MG132 (orange circle in Fig.11), both the shorter forms and the full-length form accumulated. This indicates that BioID2-RHBDL4 is unstable and is normally degraded by proteasome in cells. We also tested fusion protein arranged inversely, where BioID2 is fused to the C-terminus of RHBDL4 (data not shown), but this fusion protein could not be observed in cells at all. The effect of proteasome inhibition was not analysed in this case, but it is likely that RHBDL4-BioID2 was expressed but very rapidly degraded. To conclude, the BioID2 method was found inappropriate for the analysis of the proximity proteome of RHBDL4 because of the instability of BioID2-RHBDL4 fusions in cells. Therefore, we decided to test another proximity proteomics approach using a different enzyme termed APEX2.

8.2 Fusion protein APEX2-RHBDL4 is stable and it retains proteolytic activity of RHBDL4.

APEX2 is an engineered ascorbate peroxidase recently used for the exploration of proximity proteome of various proteins (Lobingier et al. 2017; Paek et al. 2017; Cao et al. 2018; Zuzow et al. 2018; Hung et al. 2017). The enzyme catalyses formation of short-lived biotin-phenol radical, which reacts with electron-rich amino

acids and covalently labels proteins in the vicinity of APEX2 by biotin. The biotinylation reaction occurs over a period of tens of seconds in intact living cells, yielding a 'snap shot' of the natural proximity proteome of the protein / cellular compartment to which APEX2 is anchored (Fig.12).

First, I tested if APEX2 (27 kDa protein) fused to RHBDL4 can be expressed in cells in full-length form. APEX2 was fused to the N-terminus of RHBDL4 directly or through a flexible peptide linker $(G_4S)_3$ (Fig.13a). Additionally, enhanced GFP (eGFP) was connected to the N-terminus of APEX2 enabling me to visualize the fusion protein by fluorescence microscopy. When transiently transfected into RHBDL4 deficient HCT116 cells, the whole fusion protein was expressed and only partially degraded (Fig.13b). This was encouraging considering the behaviour of BioID2-RHBDL4 constructs (Fig.11b). Even more encouraging results were obtained using fusion proteins with the peptide linker/spacer between APEX2 and RHBDL4

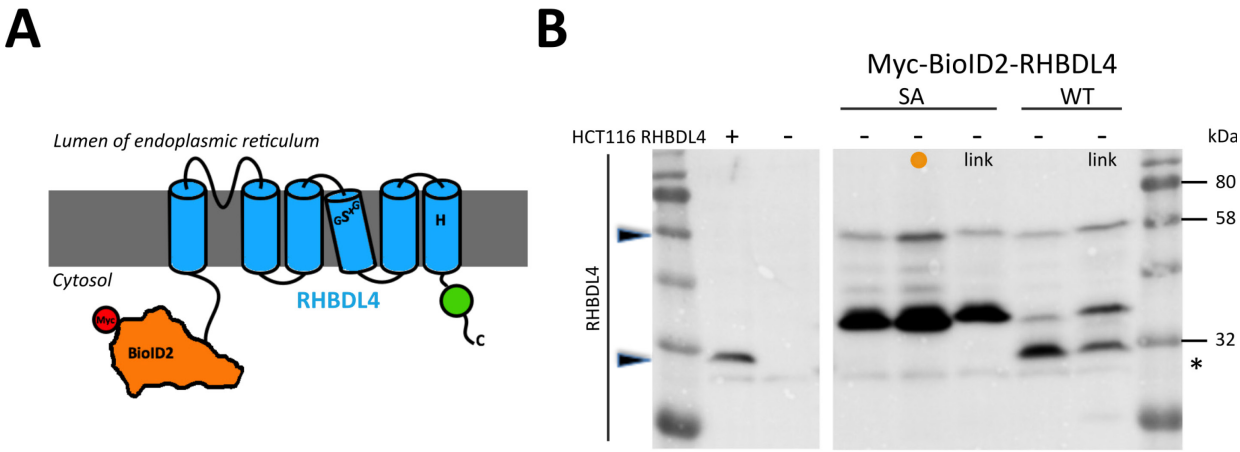

Fig.11: Fusion protein BioID2-RHBDL4 is unstable in cells.
A. An illustration of Myc-BioID2-RHBDL4 fusion protein. Catalytic residues (Ser and His) of RHBDL4 are denoted. **B.** RHBDL4 deficient HCT116 cells were transfected with BioID2 fused to the active (WT) or catalytically inactive (SA) RHBDL4. In one condition, cells were treated with proteasome inhibitor MG132 (10 μM) for 6h before cell lysis (orange circle). Cell lysates were analysed on WB by antibody against RHBDL4. HCT116 wild-type cells in the first line of WB are included as a positive control and endogenous RHBDL4 is marked by an arrow on the side (below 32 kDa marker). Right next are non-transfected RHBDL4 deficient HCT116 cells as a negative control. The full-length BioID2-RHBDL4 fusion protein is marked by the upper arrow on the side of the blot. The full-length form is stained quite weekly and large amounts of shorter (faster migrating) species are observed, independently of the presence of the linker between BioID2 and RHBDL4 (link). Shorter forms (presumably products of degradation) as well as the full-length form accumulated when proteasome was inhibited. Fusions to the active RHBDL4 were degraded further down to yield RHBDL4 alone. RHBDL4 migrates on the gel lower than it was initially expected. Its real molecular weight is 35.6 kDa. This apparent 4-5 kDa shift is observed consistently and is typical for highly hydrophobic proteins such as RHBDL4. RHBDL4 SA – catalytically inactive RHBDL4 due to S144A substitution; RHBDL4 WT – catalytically active RHBDL4; Asterisk – unspecific band; Two parts of the same membrane are shown.

- these were always expressed and mostly stable. The most problematic was the construct without peptide linker and with active RHBDL4, whose expression level was markedly low and degradation products were observed. Therefore, I decided to work further with the fusion proteins with the peptide linker (spacer). Interestingly, expression of these constructs yielded three closely co-migrating bands around 80 kDa. To test whether they differ in the glycosylation status, the cell lysate was treated with the endoglycosidase PNGase, which is able to hydrolyze N-glycans from glycoproteins, prior to WB analysis. As a result, only two bands with slightly different mobilities appeared (Fig.13c), indicating that the fusion protein, and most likely the luminal part of RHBDL4 was N-glycosylated. The identity of the two remaining bands is unknown, involving possibly some other post-translational modification(s).

The replacement of eGFP by HA tag in the fusion protein resulted in disappearance of the multiple bands (see below in Fig.15), implying that the eGFP possibly induces this effect. I did not explore this phenomenon further as it was not very relevant for the proximity proteomics study in which eGFP was not used, but I noticed that RHBDL4 actually does not possess any asparagine suitable for N-

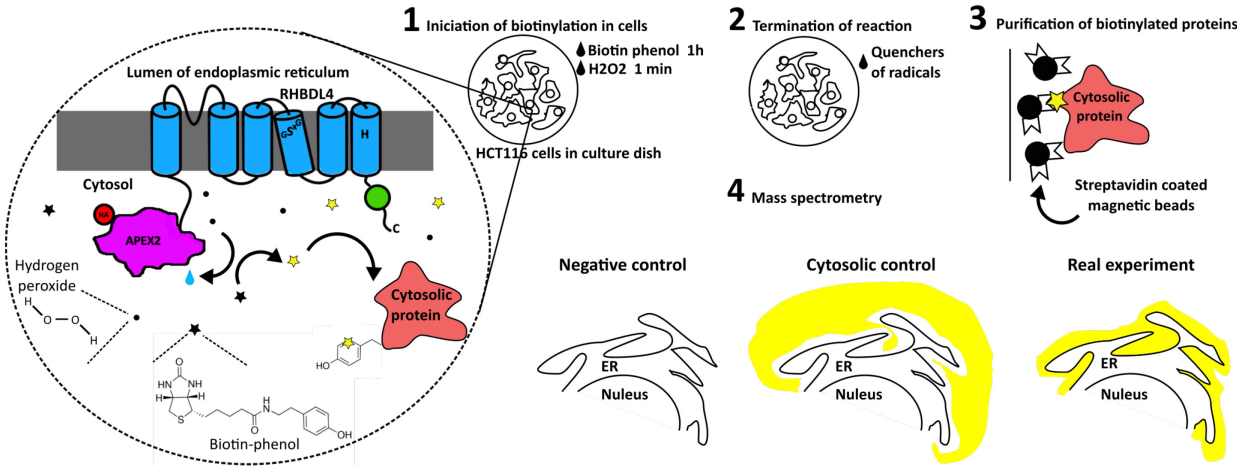

Fig.12: APEX2 proximity proteomics
APEX2 (engineered ascorbate peroxidase) catalyse the conversion of biotin-phenol (black star) into a biotin-phenol radical (yellow star), while hydrogen peroxide (dot) is turned into water. The short-lived biotin-phenol radical reacts with electron rich amino acids in proximity, mainly with tyrosine side chains. As a result, proteins spatially close to APEX2 are covalently labelled by biotin in intact living cells. The biotinylation reaction is stopped after 1 min, by washing the cells with quenchers of radicals. When the cells are lysed, biotinylated proteins are captured on streptavidin by affinity purification and identified by mass spectrometry. Two control experiments are performed in parallel with the real one. In negative control, no artificial biotinylation occurs and only endogenously biotinylated proteins (several mitochondrial enzymes) and non-specific binders to streptavidin resin will be purified from this sample. In cytosolic control, all cytosolic proteins are non-specifically biotinylated and comparison between the cytosolic control and the real experiment should increase the spatial resolution of the method.

glycosylation in its ER luminal loops. Therefore, it is exciting to speculate what actually happens, if a change in the topology of RHBDL4 occurs, or if the

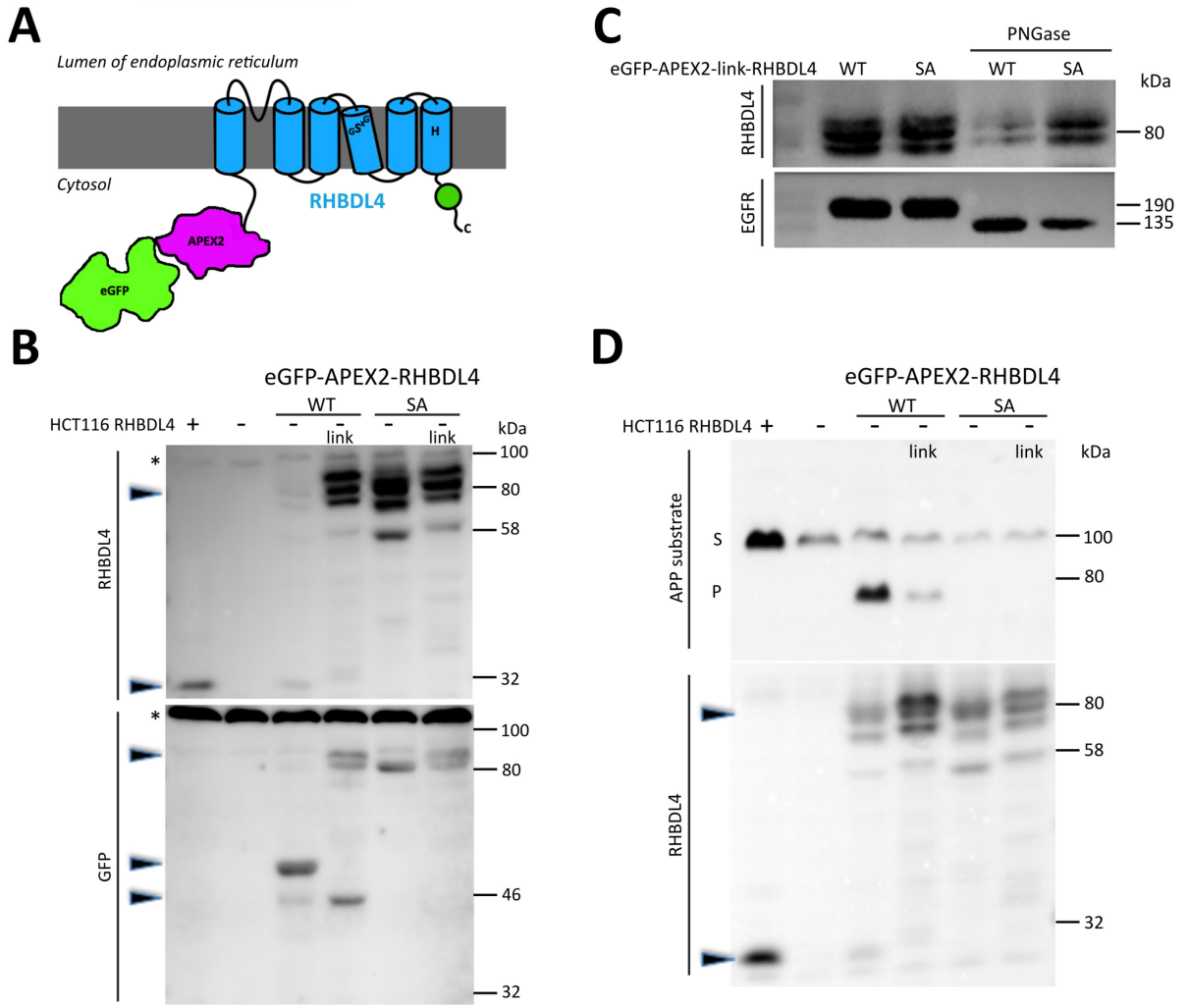

Fig.13: Fusion protein APEX2-RHBDL4 is expressed and retains proteolytic activity of RHBDL4. **A.** Scheme of eGFP-APEX2-RHBDL4 fusion protein. Catalytic residues (Ser and His) of RHBDL4 are denoted. **B.** RHBDL4 deficient HCT116 cells were transiently transfected with eGFP-APEX2-RHBDL4 WT or SA mutant. APEX2 was connected to RHBDL4 either directly or via peptide linker (link). Cell lysates were examined by WB together with wild-type or RHBDL4 deficient HCT116 cells serving as positive and negative controls, respectively. Fusion protein was visualized by antibodies against RHBDL4 and GFP. In the upper panel, endogenous RHBDL4 and RHBDL4 fusion protein are marked by arrows on the side. In the lower panel, besides the full-length fusion protein, its two degradation products are apparent, also marked by arrows on the side. **C.** Cell lysates from RHBDL4 deficient HCT116 cells transiently transfected with eGFP-APEX2-RHBDL4 were treated with PNGase to reveal whether the fusion protein was N-glycosylated. Treated lysates were analysed on WB together with untreated lysates for comparison. The fusion protein was visualized by antibody against RHBDL4. To verify that sugars were removed effectively, EGFR was visualized too. EGFR is known to be glycosylated and PNGase treatment induced apparent shift to lower molecular weight. APEX2-RHBDL4 banding pattern clearly changed too when treated with PNGase. **D.** RHBDL4 deficient HCT116 cells were transiently co-transfected with eGFP-APEX2-RHBDL4 (as in panel B) and APP, a model substrate for RHBDL4, at the ratio of 1:5 (250 ng APEX2-RHBDL4 + 1250 ng APP). As control samples, wild type or RHBDL4 deficient HCT116 cells were transfected with APP alone (total amount of DNA transfected was kept the same). Cell lysates were analysed by WB, APP was visualized by antibody against Myc tag. Full-length APP is marked as a "substrate" (S) and RHBDL4 specific cleavage product as a "product" (P). The cleavage product is present when transfected RHBDL4 is catalytically active. Asterisks denote non-specific bands.

glycosylation occurs inside the catalytic cavity of RHBDL4, or if the RHBDL4 topology prediction based on the crystal structure of bacterial GlpG (Fleig et al. 2012) is not completely accurate, which is possible.

The proteolytic activity of RHBDL4 in the fusion protein was tested against its model substrate APP. The 80 kDa product of APP cleavage was observed only when active RHBDL4 was co-transfected whereas mutation of catalytic serine in RHBDL4 (S144A) prevented the cleavage (Fig.13d). Therefore, I concluded that RHBDL4 retains its proteolytic ability when fused to APEX2 and eGFP. Surprisingly, I was not able to observe APP cleavage by endogenous RHBDL4 in HCT116 cells. I do not have a full explanation for this, but I also observed this phenomenon in HCT116 cells stably expressing RHBDL4 in fusion protein at near-endogenous level (data not shown). Hence, the absence of cleavage is most likely simply due to expression levels and consequently intracellular localisation, which would argue that APP is not a natural substrate of endogenous RHBDL4 despite claims and suggestions otherwise (Paschkowsky et al. 2016; Recinto et al. 2018). Nevertheless, it is a useful model substrate and I cannot exclude that in different cell types it may be naturally cleaved by endogenous RHBDL4.

8.3 Fusion protein APEX2-RHBDL4 is localised in the endoplasmic reticulum.

RHBDL4 has been demonstrated to be localised in the ER previously (Fleig et al. 2012; Wunderle et al. 2016). Therefore, to further verify whether RHBDL4 behaves naturally in fusion with APEX2, I made use of the eGFP-APEX2-RHBDL4 construct and observed localisation of this fusion protein in live HCT116 cells. Co-expression with Sec61-mCherry, a marker of rough ER, revealed very similar pattern of staining (Fig.14). When signals from eGFP and mCherry were overlapped, colocalisation was observed, albeit imperfect. I concluded that ER localisation of RHBDL4 is retained in fusion with APEX2, but it is not localised exactly at identical ER region as Sec61 translocon. In addition, the apparent lack of any cytosolic background GFP signal further supports the stability of APEX2-RHBDL4 fusion.

8.4 Stable expression of APEX2-RHBDL4 at near-endogenous level

Having confirmed that APEX2 can be steadily fused to RHBDL4 without affecting its activity or intracellular localisation, I proceeded to introducing APEX2-RHBDL4 fusion protein stably into the genome of RHBDL4 deficient HCT116 cells, to reach as physiological expression level as possible. I also replaced eGFP by the much smaller hemagglutinin (HA) tag for detection purposes (because anti-APEX2 antibodies were not available). By reducing the titre (i.e. the multiplicity of infection, MOI) of the retroviral particles during cell transduction, I was able to reach quite low stable expression level of HA-APEX2-RHBDL4: approximately only 11-fold (RHBDL4 WT) or 6-fold (RHBDL4 SA) higher than the endogenous expression of RHBDL4 (Fig.15a).

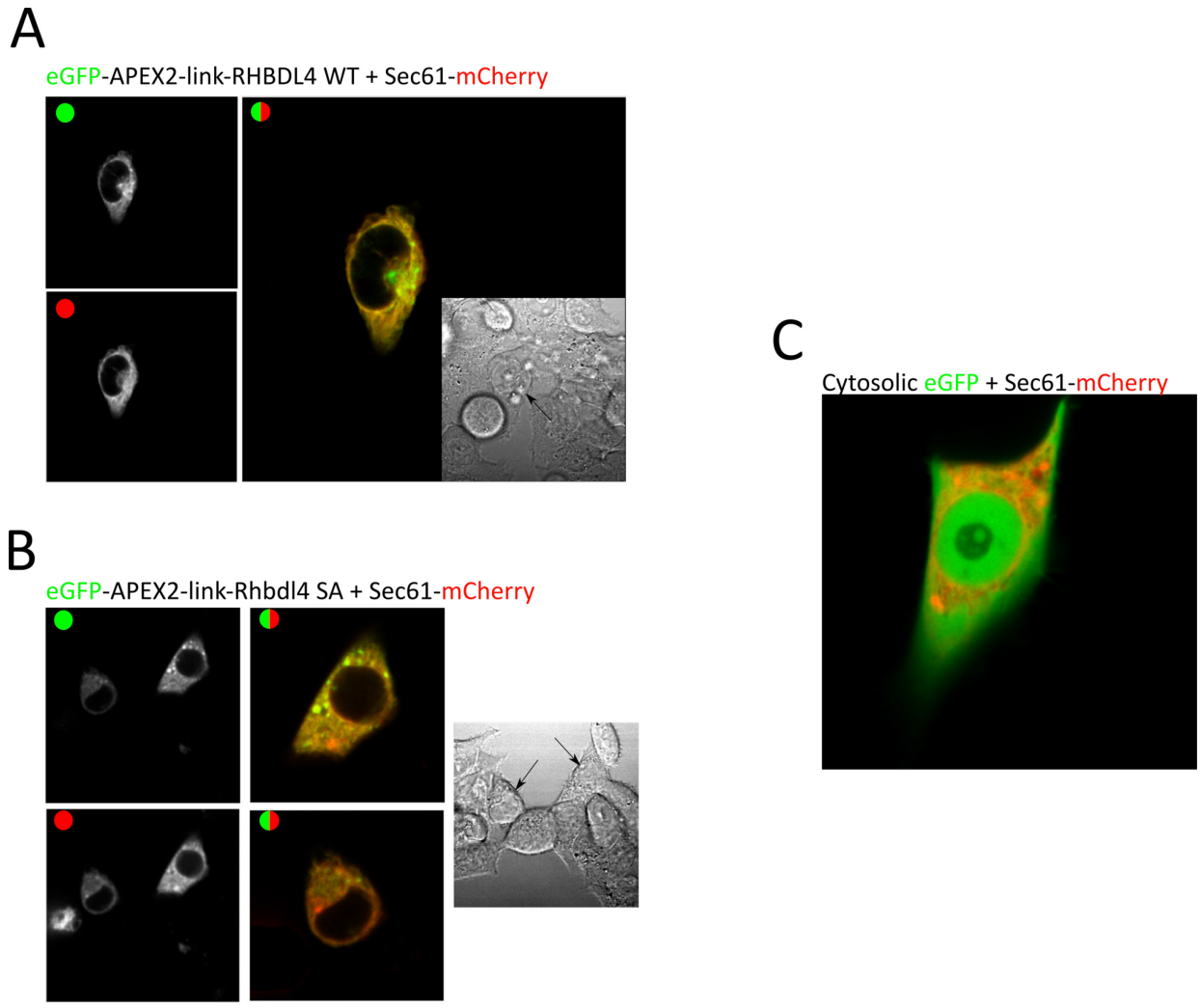

Fig.14: Fusion protein APEX2-RHBDL4 is localised to the endoplasmic reticulum.
A. and B. RHBDL4 deficient HCT116 cells were transiently co-transfected with eGFP-APEX2-link-RHBDL4 (green) and Sec61-mCherry (red). Fluorescence in live cells was examined under confocal microscope and images were processed in Fiji ImageJ. Fluorescent cells are marked by arrows in transmitted light images. **C.** As a control, HCT116 cells were co-transfected with cytosolic GFP and Sec61-mCherry.

I was particularly interested in reaching near-endogenous expression level of APEX2-RHBDL4, because the overexpression of proteins is toxic for cells in general. It can activate specific cellular pathways connected with such a stress condition completely changing the physiological setup of cell (Moriya 2015). Proteins can be mislocalised or can participate in unnatural protein interactions, and upon the overexpression of ER-targeted APEX2 even aggregation of endoplasmic reticulum was observed (Lam et al. 2014). Therefore, to explore the physiological proximity proteome of RHBDL4, its expression level has to be taken into account. To reduce the expression level of the fusion protein further to near-endogenous level of RHBDL4, I used an inducible expression system, where HA-APEX2-RHBDL4 is inserted into the genome under cumate-sensitive promoter/operator (Mullick et al. 2006) encoded in a PiggyBac transposon vector (System Biosciences). Contrary to the manufacturer's information, I did observe some protein expression even in the absence of inducer. Conveniently, this 'leaky' expression of RHBDL4 in the fusion was at a level very close to the endogenous level of RHBDL4, even about 2-fold lower, and these results were well reproducible (Fig.15b).

Encouragingly, expression of HA-APEX2-RHBDL4 from stably, chromosomally integrated constructs yielded single bands on western blots probed with RHBDL4 antibody (Fig.15). Therefore, the three bands previously observed with the eGFP-APEX2-RHBDL4 fusions (previous section) were the result of overexpression and/or fusion to eGFP. Probing the HA-APEX2-RHBDL4 fusions with anti-HA antibody, targeting the N-terminus of the whole fusion protein, revealed several weak bands beneath the prominent one and this smear was disappearing as the expression of RHBDL4 was decreasing. Absence of these additional bands on WB labelled by RHBDL4 antibody indicates that in these shorter products, the C-terminus of RHBDL4 may be slightly shortened or modified, therefore disrupting the epitope(s) recognized by the RHBDL4 antibody. These C-terminal changes might be a natural feature of RHBDL4, but indeed I am not able to observe it endogenously as the only available antibodies against RHBDL4 are against the very same C-terminal extramembrane part.

To conclude, I was able to express HA-APEX2-RHBDL4 as a stable and full-length fusion protein at near-endogenous level in HCT116 cells, which is

particularly important for the following proximity proteomics. For the subsequent small scale biotinylation experiments and preliminary proteomics experiments, I chose the retrovirus-generated cell lines marked as WT 1/12 and SA 1/12 (Fig.15a),

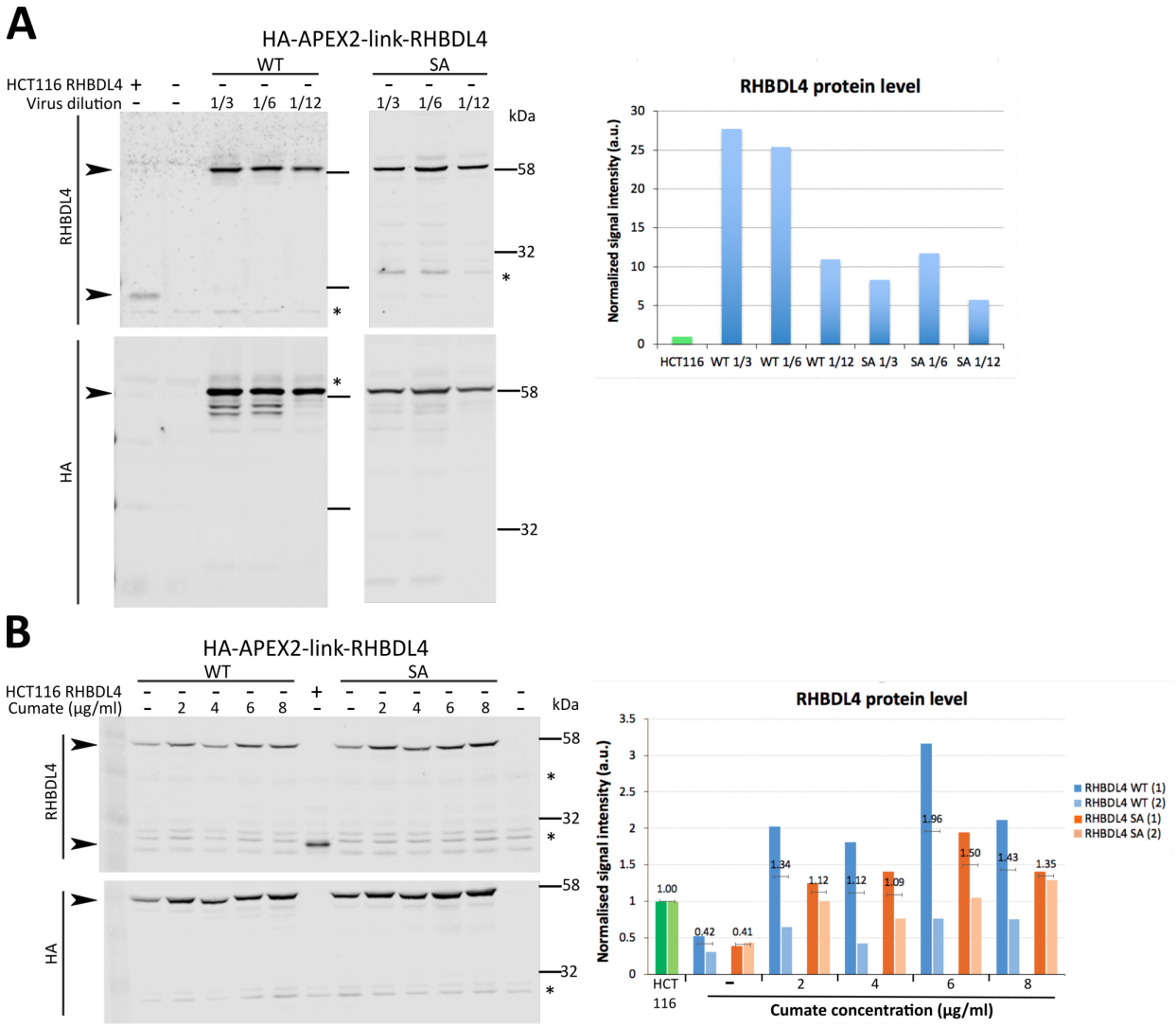

Fig.15: Stable expression of APEX2-RHBDL4 at near-endogenous level
A. RHBDL4 deficient HCT116 cells were infected by retrovirus particles bearing an HA-APEX2-link-RHBDL4 expression cassette (in vector pM6P). Virus particles were sequentially diluted to decrease the MOI and thus likely the resulting copy number of chromosomal insertion events and expression of the introduced protein. After antibiotic selection (0.5µg/ml puromycin), the expression of APEX2-RHBDL4 was analysed by WB in selected cell populations by near-infrared fluorescence detection (LI-COR). Wild type and RHBDL4 deficient HCT116 cells were used as controls. Arrows are pointing to the APEX2-RHBDL4 fusion and endogenous RHBDL4 protein. Two different WBs are shown side by side, therefore the position of individual bands is slightly shifted. The fluorescence signal of RHBDL4 was quantified, normalized to total protein in each lane, and plotted on a graph as relative values to the intensity of endogenous RHBDL4 (which was set to one). **B.** RHBDL4 deficient HCT116 cells were transfected with transposase and HA-APEX2-link-RHBDL4 expression cassette under cumate-sensitive promotor in a PiggyBac transposon vector. After antibiotic selection (0.5µg/ml puromycin), the resulting cell populations were grown at different concentrations of cumate and analysed by WB for APEX2-RHBDL4 expression. Wild type and RHBDL4 deficient HCT116 cells were used as controls. The fluorescence signal of RHBDL4 was plotted as in panel A. Two independent experiments are shown in the graph with the number of experiment indicated in brackets in the legend. Experiment 1 corresponds to the WB shown. The mean values are displayed in the graph. Asterisks in WB indicate unspecific bands.

in which the 1/12 ratio indicates the dilution factor of retroviral particle stock used for cell transduction and whose fusion protein expression was 6 to 11-fold above the endogenous level of RHBDL4. Cumate-inducible WT and SA cell lines (Fig.15b) were subsequently used in SILAC-based proteomics upon no induction, in which the fusion protein expression was about 2-fold bellow the endogenous level of RHBDL4.

8.5 Biotinylation in intact cells and enrichment of biotinylated proteins

To test the enzymatic activity of APEX2 in HA-APEX2-RHBDL4 fusion protein, I performed biotinylation reaction (Fig.12) at a small scale first using the retrovirus-generated cell lines described in the previous section. In the presence of biotin-phenol and hydrogen peroxide in the media of cells expressing APEX2, this engineered ascorbate peroxidase takes one electron away from biotin-phenol, leaving a biotin-phenol radical while hydrogen peroxide is turned into water. This short-lived biotin-phenol radical then reacts with electron-rich amino acids in proximity and as a result, proteins spatially close to APEX2 are covalently biotinylated. When cells are lysed, biotinylated proteins can be easily captured on streptavidin by affinity purification.

As expected, depending on the presence of hydrogen peroxide and APEX2 itself, a large spectrum of cellular proteins was biotinylated by APEX2 (Fig.16a). In the negative control lanes, several prominent bands represent endogenously biotinylated proteins (Hung et al. 2016). The two bands around 80 kDa probably represent alpha subunits of propionyl-CoA carboxylase and 3-methyl-crotonyl-CoA carboxylase, and a band around 130 kDa probably represents pyruvate carboxylase. Interestingly, the activity of APEX2 fused to RHBDL4 SA was about 2-fold weaker than when fused to RHBDL4 WT (Fig.16b).

Purification of biotinylated proteins on streptavidin-coated beads was successful (Fig.16c). The amount of cell lysate corresponding to 160 µg of total protein per condition was applied onto the resin. Although many biotinylated proteins remained unbound and appeared in the flow-through (FT) fraction (Fig.16d), possibly due to overloading of the column, the purified eluate (EL) was robustly depleted of non-biotinylated contaminant proteins, as indicated by the very weak total protein staining in this fraction (Fig.16e). The staining is actually very

close to background, represented by the eluate from alone streptavidin beads, on which no cell lysate was applied.

The apparent discrepancy between quantitative information gleaned from the WB in Fig.16c and chart in Fig.16d is due to very different total volumes of FT and EL fractions. Although the sample volumes loaded into each lane were equal (20 µl), this represents only 4% of the FT fraction but 67% of the EL fraction. The chart in Fig.6d takes these differences in protein concentration into account as it was produced by integrating the total HRP (horseradish peroxidase) signal from detected biotin in each lane and recalculating it to the total volume of the corresponding fraction (i.e., multiplying the signal from FT by 25 and from EL by 1.49, to correct them to 100% of the fraction volume). The quantification of HRP signal might be inaccurate at the high signal intensities due to a narrower dynamic range (as opposed to near-infrared fluorescence signal that is linear over several orders of magnitude), but this uncertainty should be of a similar magnitude for both FT and EL lanes of the blot as they have been intentionally loaded such that the HRP signal has approximately similar intensity.

To conclude, a significant proportion of biotinylated proteins can be cleanly isolated by the enrichment on streptavidin-coated beads. For the purposes of APEX2 proximity proteomics, the purity of isolated biotinylated proteins is much more important than their total yield. Total protein staining together with the detection of biotin (Fig.16c,16e) shows that very low total protein remains in EL, but these proteins are intensely biotinylated.

8.6 Proximity proteome of RHBDL4

Having thoroughly tested that APEX2 approach is well applicable on the study of RHBDL4 in its physiological cellular environment, I proceeded to large scale experiments involving MS for identification of isolated biotinylated proteins.

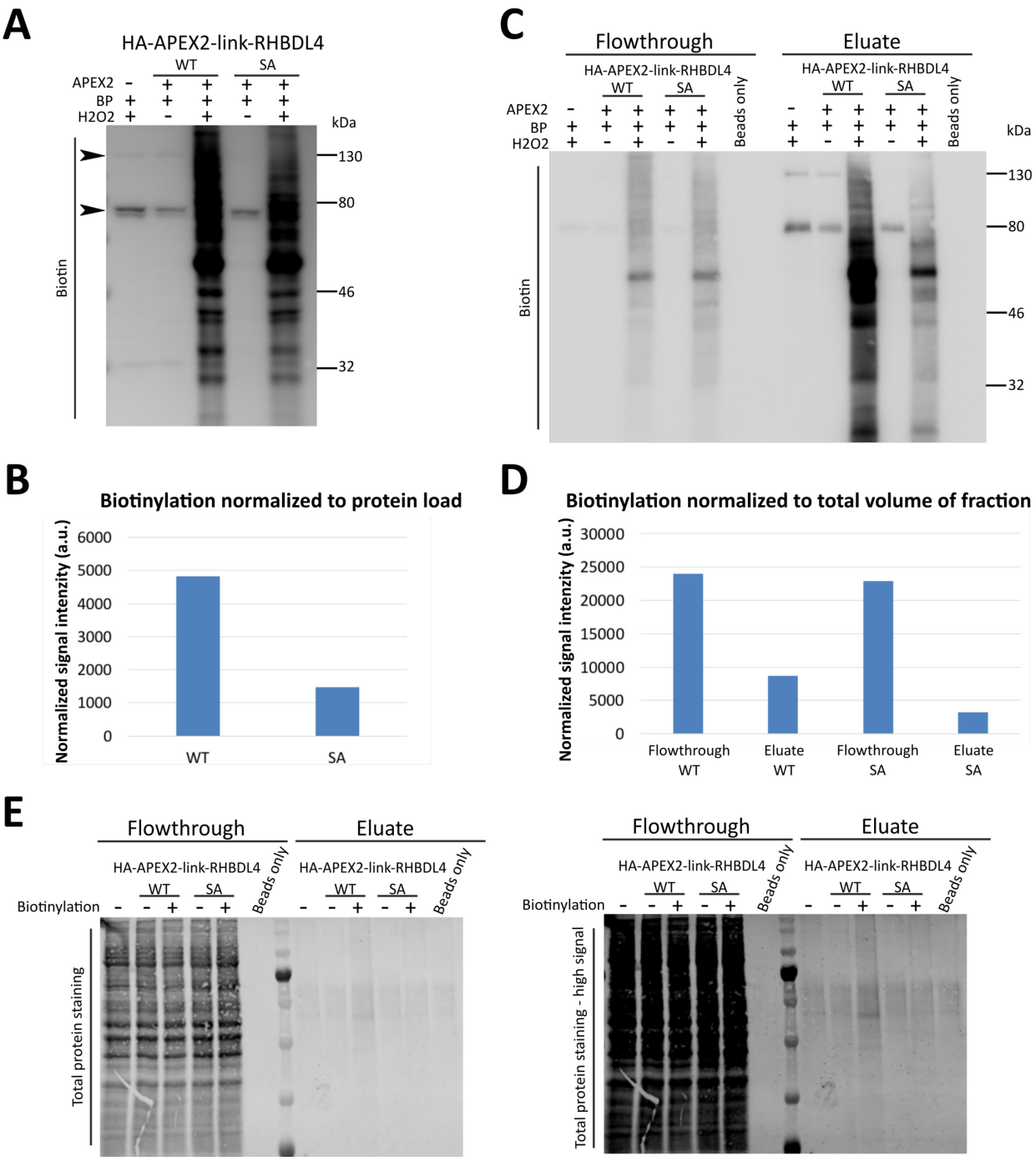

Fig.16: Biotinylation reaction catalysed by APEX2
A. Biotinylation reaction was performed in RHBDL4 deficient HCT116 cells stably expressing HA-APEX2-RHBDL4 WT or SA mutant. The specific cell lines used correspond to the WT 1/12 and SA 1/12 analysed in Fig.15. Cell lysates were directly loaded on the gel and analysed by WB via Streptavidin-HRP conjugate. In the negative controls, which were lacking H_2O_2, or APEX2 in cells (the RHBDL4 deficient HCT116 cells), only endogenously biotinylated proteins are visible around 80 kDa and 130 kDa. In contrast, when biotinylation reaction occurred via HA-APEX2-RHBDL4, broad range of cellular proteins was biotinylated. **B.** Biotin signal from WB in panel A was normalized to the amount of protein loaded in each lane. **C.** Cell lysates from panel B were loaded on streptavidin resin. Unbound proteins (flow-through, FT), and proteins captured by the resin and subsequently eluted (eluate, EL) were collected. Samples were analysed by WB via Streptavidin-HRP. **D.** Biotin signal from WB in panel C was normalized to the total volume of the FT and EL fraction. **E.** The same membrane as shown in panel C was stained to reveal total protein loaded based on near-infrared fluorescence (LI-COR), before biotin detection. The membrane is shown also in high signal intensity, to demonstrate that there is only weak signal on the side of the eluate.

Due to the nature of the radical reaction through which the biotinylation occurs it is important to include appropriate controls while performing APEX2 experiment (Fig.12). Although the probability and extent of biotinylation decrease with each nanometre distance from the APEX2 enzyme, because biotin radicals are very short-lived (<1ms) and the whole reaction lasts only tens of seconds, some distant proteins might be accidentally biotinylated too. Because there is no barrier which would surround the cytosolic RHBDL4 proximity where the radicals are produced, the radicals can potentially freely diffuse into the cytosol. Therefore, a comparison with the control experiment, in which the APEX2 is not anchored but floats within the cytosol (cytosolic control), increases the spatial resolution of the method. Randomly biotinylated distant proteins will be biotinylated also in the cytosolic control experiment to the same or larger extent, whereas the proximity proteome of RHBDL4 should be significantly enriched in the real experiment (ER-localised APEX2-RHBDL4 fusion) compared to the cytosolic control. In addition, a negative control experiment is performed in parallel, in which no artificial biotinylation occurs and it should reveal the background of this method, including endogenously biotinylated proteins (several mitochondrial enzymes) and non-specific binders to streptavidin resin. Since I wanted to compare the results from real experiment and the two controls quantitatively, SILAC approach was very appropriate (Ong et al. 2002). It enables me to pool the three samples together, so any subsequent procedures affect all of them equally, which makes the quantitative comparison more reliable.

However, before employing SILAC which uses expensive reagents, I carried out a preliminary experiment (Fig.17) without SILAC labelling, to ascertain that the procedure and my workflow work reliably also at a large scale, and to confirm that mass spectrometry is compatible with our samples. I used the retrovirus-generated cell line expressing HA-APEX2-RHBDL4 WT, described in previous sections (Fig.15a – WT 1/12; Fig.16), and cytosolic control cell line, in which APEX2 tagged by a nuclear export sequence (APEX2-NES) was expressed freely in the cytosol (cytosolic control, Fig.17a). The biotinylation reaction was performed in both cell lines in parallel yielding a broad spectrum of biotinylated proteins in cell lysates

(Fig.17a), which were individually purified on streptavidin resin and subjected to mass spectrometry. A list of identified proteins (protein hits) was obtained for each sample individually.

To perform a quantitative comparison between the real experimental sample and the cytosolic control, I performed a calculation similar to relative quantification in label-free MS by spectral counting (Zhu et al. 2010), which yielded approximate quantitative information already from this preliminary experiment. To each identified protein represented by a unique gene name I assigned all its peptides detected by mass spectrometry. If it was not possible to unambiguously determine which of the proteins the peptide belongs to, due to similarities in protein sequences, the proteins were grouped. Each peptide was characterised by its spectral count (peptide spectrum matches, PSM), which says how many times the peptide was detected by mass spectrometry. By integration of the PSM of all peptides belonging to one protein a spectral count for the protein was obtained. This value strongly correlates with the original abundance of the protein in sample (Zhu et al. 2010). Clearly, the more molecules of a given protein are present in the sample, the more peptides can be generated and detected by mass spectrometry. Therefore, I will refer to this quantitative value as protein abundance (although it is not exactly the same).

For every protein detected in the real experimental sample I calculated a quantitative ratio by dividing its abundance by the abundance found in cytosolic control. I displayed these ratios together with the abundance from real experiment in a scatter plot (Fig.17b). Theoretically, the higher the quantitative ratio, the closer might the protein be to RHBDL4 in cells, because the extent and probability of biotinylation are decreasing with the distance from APEX2 (fused to RHBDL4). Therefore, by using a ratio cut-off we can distinguish the RHBDL4 proximity proteome from randomly labelled cytosolic proteins. However, since we do not know in advance, which proteins will be present above the cut-off (i.e., we do not know a sufficiently big set of RHBDL4 interaction partners), we cannot calculate the cut-off as it was described in original protocol (i.e., set the cut-off in such a way that the positive control proteins are mostly included) (Hung et al. 2016). Therefore, I used an arbitrary ratio cut-off further described below (red lines in Fig.17b).

Defining spatial relations between the proteins above the cut-off is more complicated. Theoretically, proteins which are very close to RHBDL4 should be biotinylated in all the cells where APEX2 is connected to RHBDL4, whereas more

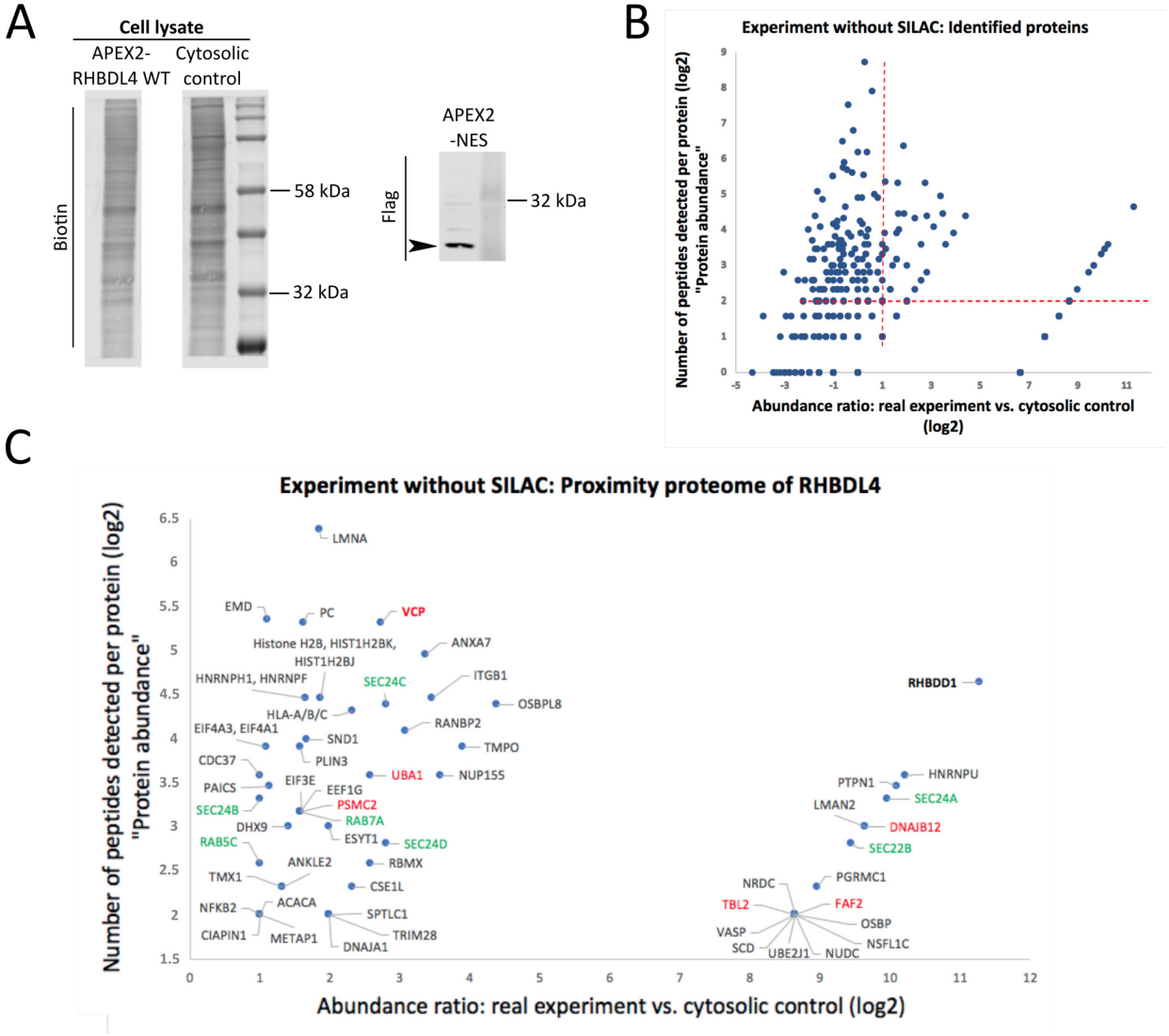

Fig.17: Proximity proteome of Rhbdl4 – preliminary experiment without SILAC
A. Biotinylation reaction analysed in Fig.16 was performed at large scale to create samples for protein identification by mass-spectrometry. As a real experimental sample, retroviral-generated cell line expressing HA-APEX2-RHBDL4 WT was used. As a cytosolic control, RHBDL4 deficient HCT116 cells stably expressing Flag-APEX2-NES were used. A small portion of cell lysates after biotinylation reaction was analysed on WB, showing successful biotinylation of a broad protein spectrum in both conditions. **B.** Biotinylated proteins were purified from the cell lysates shown in panel A and subjected to mass spectrometry analysis. All proteins identified in the real experimental sample are shown based on their quantitative ratios against cytosolic control and on their abundance in the real experimental sample. The proteins which cluster on the right side of X axis were not detected in the cytosolic sample and their cytosolic abundance was artificially set to 0.01. Red lines mark the cut-off, defining the likely proximity proteome of RHBDL4 (upper, right corner), i.e. proteins at least 2-fold enriched in the real experiment and with at least 4 peptides detected. **C.** Proteins selected to be in the proximity of RHBDL4 by the cut-off in panel B are shown in detail. The proteins are represented by dots and labelled by individual gene names. Proteins displayed in red represent ER-associated degradation factors, proteins displayed in green represent factors associated with intracellular trafficking.

distant ones should be biotinylated only in a fraction of the cellular population, based on probability. However, we can imagine that proteins which form polymers will have a higher quantitative ratio than proteins acting as monomers irrespective to the distance from RHBDL4. Therefore, the quantitative ratio mirrors also the proportion of the protein pool that localises to the proximity of RHBDL4. Also, it is possible that the close proximity of APEX2-RHBDL4 is actually saturated by biotin radicals, which limits the 'resolution' of the method.

The abundance itself (displayed on the Y axis in Fig.17b) is also largely dependent on the size (molecular weight) of the protein, on the ability of its peptides to ionize during MS and on its overall abundance in the cell. These influences, however, do not affect the ratio of abundances for a given protein, because they affect its abundance in the cytosolic control to the same extent as in the real sample. However, some proteins were detected only in the real experimental sample and the ratio could not be computed. Therefore, to get a number when calculating the quantitative ratio for them and to visually distinguish these proteins from the rest, their cytosolic abundance was arbitrarily set to 0.01, which was chosen as a practical, low non-zero number. These proteins thus become separated on the right side of X axis in Fig.17b, but because their true cytosolic abundance was not available, their ratios are influenced by all the factors mentioned above, related to their molecular weight and ionizability of their peptides. Therefore, these values should be considered with special caution. The reason why these proteins were not found in the cytosolic control can be inaccessibility by cytosolic APEX2. For example, RHBDL4 itself is found among these proteins (RHBDD1), possibly hidden between ER sheets and tubules. Another possible explanation is that cytosolic APEX2 labelled these proteins only very rarely resulting simply in undetectable MS signal. Hence, these proteins might actually belong to the best candidates for RHBDL4 proximity proteome, therefore I displayed them together with other protein hits.

The arbitrary cut-off (red lines in Fig.17b) was chosen quite strictly, considering the proteins in RHBDL4 proximity to be biotinylated at least twice more than in the cytosolic control, and to possess at least 4 detected peptides, to obtain reliable quantitative ratios. The proteins above the cut-off are displayed in detail with gene name labels (Fig.17c). The protein hits are colour-coded to highlight

proteins involved in trafficking and ER-associated degradation in which RHBDL4 might function according to literature (Fleig et al. 2012; Wunderle et al. 2016). Importantly, two positive controls were identified by this procedure among the RHBDL4 proximity proteome – RHBDL4 itself (RHBDD1) and VCP ATPase (also known as p97 or Cdc48) which is a validated interaction partner (Fleig et al. 2012; Lim et al. 2016), confirming that our approach worked well.

The following experiments were triple-SILAC-based using a growth media with isotopically labelled amino acids arginine and lysine. They were performed with the RHBDL4 deficient HTC116 cells expressing cumate-inducible HA-APEX2-RHBDL4 construct at the expression level below the endogenous level of RHBDL4, described in previous section (Fig.15b,18a), and with cytosolic control and negative control cell lines derived also from RHBDL4 deficient HTC116 cells. The real experimental cell line was labelled by SILAC medium containing heavy (see Methods) amino acids (therefore denoted as the H sample). The cytosolic control cell line was the same as in the preliminary experiment without SILAC (APEX2-NES) and it was labelled by SILAC medium with amino acids of medium (see Methods) molecular weight (the M sample). As a negative control cell line, I used either RHBDL4 deficient HCT116 cells, or the cell line used in the real experiment but without H_2O_2 in the reaction, and these cells were grown in SILAC medium with normal amino acids of light weight (the L samples). The biotinylation reaction was performed in parallel with all the samples yielding a broad spectrum of biotinylated proteins in cell lysates except for the negative controls (Fig.18b). Exploiting the advantages of SILAC labelling, cell lysates of each H-M-L set were then pooled yielding two biological replicates – Experiment 1 and Experiment 2. For each of these experiments separately, biotinylated proteins were then purified on streptavidin resin and subjected to mass spectrometry.

The results were analysed using MaxQuant, which calculated quantitative ratios between the real experimental sample and the two controls (H/L and H/M ratio; see Methods for more details). All proteins detected (3000 proteins in total) are displayed in scatter plots based on the H/L and H/M ratios (Fig.18c). Theoretically, the proteins which are in the proximity proteome of RHBDL4 should have both ratios high (at least higher than 1 would make sense, generally). However,

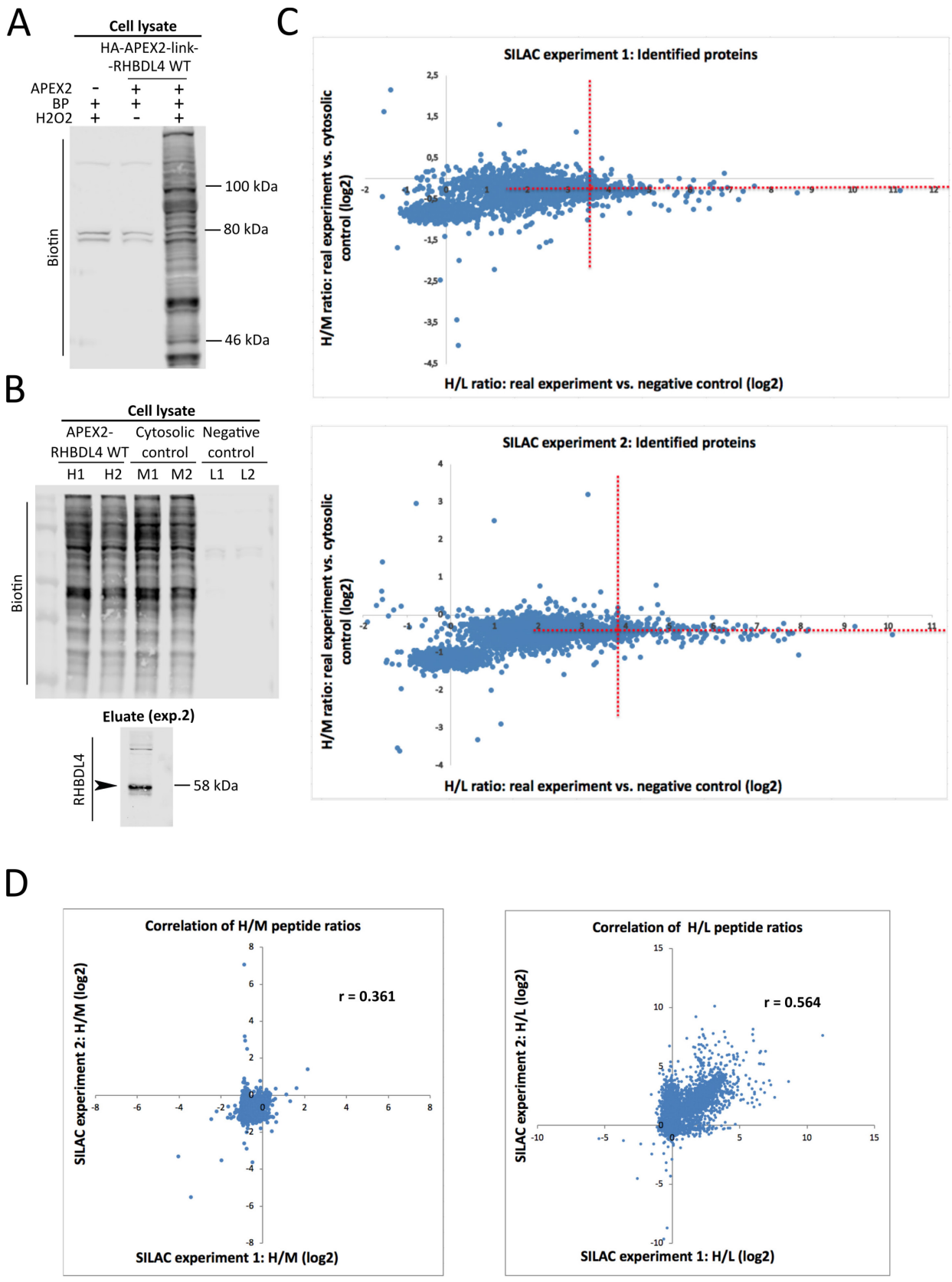

Fig.18: Proximity proteome of RHBDL4 – SILAC-based experiments
A. Biotinylation reaction was performed in cell line expressing cumate-inducible HA-APEX2-RHBDL4 under no induction. When all the necessary factors were present, a broad spectrum of proteins was biotinylated in cell lysate. **B.** Biotinylation reaction was performed on a large scale to create samples for protein identification by mass-spectrometry. As a real experimental sample, the cell line shown in panel A was used. As a cytosolic control, RHBDL4 deficient HCT116 cells expressing Flag-APEX2-NES were used. As a negative control either RHBDL4 deficient HCT116 cells or the cell line shown in panel A without H_2O_2 in the reaction were used. Small portions of cell lysates after biotinylation reaction were analysed, showing successful biotinylation of a broad protein spectrum in all conditions, except the negative controls. H, M and L means heavy, medium and light SILAC labels, respectively.

probably happened due to the very low expression level of HA-APEX2-RHBDL4 (Fig.15b), which was not sufficient for biotinylation of proteins in its vicinity above the cytosolic control (see Disscusion). Also, the correlation between H/M ratios for individual proteins between the duplicates was much worse compared to the H/L ratios (Fig.18d). Therefore, instead of setting the ratio cut-off at value, which implies sufficient enrichment over controls (as for the preliminary experiment), I set the cut-off based on the only validated cytosolic interaction partner of RHBDL4 – the VCP ATPase, and considered all proteins having better quantitative ratios as the candidate proximity proteome of RHBDL4 (Fig.18c – red lines). Surprisingly, RHBDL4 itself was not detected by mass spectrometry, although it was present in the eluate from the streptavidin column (Fig.18b).

To conclude, three independent experiments were performed in total to map the close proteomic environment of RHBDL4 in HCT116 cells. Finally, I compared the resulting RHBDL4 proximity proteomes and found several proteins that were found in two of them, and four proteins that were found in all three experiments (Table 1), yielding the most reproducible and likely proximity proteome candidates. The known binding partner of RHBDL4, VCP ATPase, is indeed among them, which makes the other three proteins – integrin beta 1 (ITGB1), annexin A7 (ANXA7) and perilipin 3 (PLIN3) – worth analysing further. They will be discussed in the next section in more depth. According to a publicly available database of Hela cell proteome (Itzhak et al. 2016), these candidate proteins are not the most abundant ones in mammalian cells, further supporting their significance and demonstrating that their cellular abundance is not the cause of their repeated identification in the proximity proteome of RHBDL4. Interestingly, only very few membrane proteins were detected repeatedly among the candidate proximity proteomes according to prediction software Phobius (Käll et al. 2004).

C. Cell lysates analysed in panel B were mixed 1:1:1 (based on protein concentration) yielding two replicate experiments (Exp.1: H1+M1+L1; Exp.2: H2+M2+L2). Biotinylated proteins were then purified from the pooled samples, subjected to mass spectrometry and evaluated using MaxQuant. All identified proteins are displayed based on their quantitative peptide ratios. Red lines mark the cut of, above which the proteins are considered to be in very proximity of RHBDL4. The cut of was set based on a known interaction partner of RHBDL4 – the VCP ATPase. **D.** Correlation for the H/L and H/M quantitative ratios between the duplicates is shown for individual proteins. The Pierson correlation coefficient is displayed (r).

	RHBDL4/neg. control peptide ratio (log_2)		RHBDL4/cytosolic control peptide ratio (log_2)			Copy number in Hela cells (percentile)	TMD prediction
Gene name	**Exp.1**	**Exp.2**	**Exp.1**	**Exp.2**	**Exp. without SILAC**	***mapofthecell.org***	***Phobius***
ITGB1	4.56	5.77	-0.11	-0.24	3.46	88.9	1x TMD
ANXA7	3.91	4.10	-0.05	-0.32	3.37	88.9	none
PLIN3	3.76	3.90	-0.08	-0.30	1.58	94.8	none
VCP	3.57	3.80	-0.22	-0.38	2.74	97.0	none
APRT	3.62	5.32	-0.20	-0.34		96.3	none
EIF5A	3.87	3.83	-0.07	-0.23		99.6	none
EZR	3.82	4.15	-0.16	-0.31		96.3	none
GPD2	4.55	5.66	-0.08	-0.33		84.2	none
GSPT1	3.69	5.80	-0.22	-0.31		94.1	none
GSR	3.95	4.07	-0.02	-0.23		91.8	none
IMPDH2	3.80	4.08	-0.13	-0.31		93.7	none
ITGA3	5.73	3.96	0.00	-0.20		43.3	1x TMD
MAGEB2	3.82	4.08	-0.21	-0.29		15.9	none
NFKB2	3.57		-0.13		1.00	75.4	none
NUDC	4.06		-0.17		8.64	96.7	none
PPA1	4.77	5.62	-0.13	-0.32		97.6	none
SFN	4.57	4.35	-0.08	-0.25		97.9	none
SYNCRIP	3.86	4.17	-0.20	-0.32		95.1	none
TAGLN2	6.58	6.29	-0.08	-0.19		99.4	none
TK1	4.08	5.99	-0.22	-0.07		80.9	none
UBA6	4.54	4.66	-0.05	-0.28		76.1	none
YWHAQ	3.92	5.07	-0.17	-0.22		97.2	none

Table 3: The candidate proximity proteome of RHBDL4 – repeatedly detected proteins.
Proteins detected in all three experiments are shown at the top in green, below are proteins detected in two of the experiments in alphabetical order. Proteins displayed in purple were detected in SILAC and preliminary non-SILAC experiment. The quantitative ratios, percentile abundance in Hela cell and TMD predictions are shown. The percentile signifies the relative abundance of protein in cell (percentile 90 means that 10% of cellular proteins are more abundant and 90% of cellular proteins are less abundant than the particular protein).

9 Discussion

9.1 The BioID2

The BioID2 appeared to be very unstable when fused to RHBDL4, therefore I did not proceed in this method. It is hard to speculate what was the cause of the instability. The APEX2 enzyme (which was stable in fusion with RHBDL4) is of the same size and it was connected to RHBDL4 in an identical way as the BioID2. Therefore, the instability of the fusion was probably due to some inherent qualities of BioID2, which were not compatible with RHBDL4.

It is interesting that the pattern of degradation products was dependent on the catalytic activity of RHBDL4, because it is not possible that RHBDL4 would cleave the fusion protein by itself (its active site is opened to the ER lumen). It is possible to speculate that RHBDL4 active site may be important for the recruitment of factors, which are involved in the degradation of BioID2-RHBDL4.

9.2 The APEX2 assay establishment

In this study, the APEX2 proximity proteomics approach was successfully implemented in our laboratory for the investigation of rhomboid superfamily protein members. APEX2 and its predecessor APEX were shown to be powerful tools for analysis of membrane proteins and cellular membrane compartments. In particular, dynamics of G-protein-coupled receptor (GPCR) signalling (Paek et al. 2017), the protein network involved in it (Lobingier et al. 2017) and the proteomes of ER and mitochondrial membranes were revealed thanks to these methods (Hung et al. 2017; Rhee et al. 2013). Compared to the original protocol for spatially resolved proteomic analysis by APEX2 (Hung et al. 2016), I adjusted the procedure for isolation of biotinylated proteins on streptavidin resin to increase the purity and reduce non-specific binding. The improved protocol is described in methods in detail. Briefly, the main improvements were:

- The binding of proteins on the resin was performed overnight in the presence of 3.2 M urea and 0.5% SDS (instead of the originally used no urea and 0.1% SDS). These more stringent conditions should reduce non-specific binding of proteins to the resin and protein aggregation.

- The subsequent washing of the resin with bound proteins was more intensive and used harsher conditions. In total, the resin was washed 10 times using 0.5% SDS and 4 M urea (instead of only 5 times with 0.1% SDS and none or 2 M urea, as originally described). These conditions should increase the purity of isolated biotinylated proteins as the unspecific binders are efficiently washed out.
- The elution of biotinylated proteins from the resin was achieved by boiling for 15-20 min in the presence of 4 mM biotin (instead of originally used 2 mM biotin and only 10 min boiling). This should assure that all proteins are eluted from the resin.

Together, these adjustments significantly reduced the amount of non-biotinylated proteins eluted from the resin.

9.3 Analysis of the proximity proteome of RHBDL4

APEX2 approach was applied to mammalian rhomboid protease RHBDL4, which is localised in the ER membrane. APEX2 was genetically fused to the N-terminus of RHBDL4, therefore facing the cytosol. As the labelling of proteins in the proximity occurs through biotin radicals for which the membrane is impermeable, only proteins with cytosolic domains could be labelled. Therefore, one would suggest to insert the APEX2 into RHBDL4 in such a way that it would face the ER lumen, into which the active site of RHBDL4 opens. However, the insertion of APEX2 into the luminal loops of RHBDL4 is challenging and risky when addressing the physiological function of RHBDL4 in particular, because it is not known which loops are functionally important. All of them are near the active site of rhomboid and the insertions could thus cause unwanted side effects. Therefore, we decided to start with APEX2 connected to the cytosolic end of RHBDL4. The N-terminus was selected in particular, as it is far from the known functionally important parts of the protein, and the N-terminal tagging does not negatively influence RHBDL4 protease activity on heterologous model substrates (Fleig et al. 2012). Whereas the C-terminus of RHBDL4 bears two conserved motifs important for its proposed function in ERAD (Fleig et al. 2012) and our previous attempts to genetically fuse BioID2 enzyme to the C-terminus of RHBDL4 completely failed.

9.3.1 Problems encountered – spatial cytosolic control

We were particularly interested in reaching very low expression level of APEX2-RHBDL4 in cells, in order to explore the function of RHBDL4 in physiological conditions avoiding overexpression artefacts. The retroviral delivery of the APEX2-RHBDL4 expression cassette yielded expression level approximately 6 to 11-fold higher than the endogenous level of RHBDL4. No further reduction of expression level was achieved by reducing MOI during cell infection/transduction, suggesting that we reached the bottom limit of this method. When the expression cassette was stably introduced into the cells using PiggyBac transposon vector, the baseline expression level from an inducible promoter (i.e. promoter repressed) was conveniently very low, approximately 2-fold lower than endogenous level of RHBDL4 expression. In both cases, the APEX2 enzyme activity was sufficient to biotinylate many cellular proteins, therefore, we decided to focus on the latter case in which APEX2-RHBDL4 was actually "underexpressed", to completely exclude the possibility of overexpression artefacts.

To increase the spatial specificity of the proximity proteomics experiment, the proteome generated by APEX2-RHBDL4 is compared to the cytosolic control proteome generated by APEX2 which is localised freely in the cytosol (henceforth termed "cytosolic APEX2"). The logic behind this is that biotin radicals produced by APEX2-RHBDL4 can label also cytosolic proteins, which are just randomly passing near RHBDL4 or are very abundant in cells. Moreover, the radicals might exceptionally escape from RHBDL4 proximity and penetrate further into the cytosol as there is no barrier, which would stop them (although they are generally very short-lived). The resulting non-specifically labelled proteins should be labelled to the same or bigger extent by cytosolic APEX2 and therefore excluded based on quantitative comparison, whereas functionally relevant proteins for RHBDL4 should be enriched above the cytosolic control. To enable robust quantification, which is absolutely key for this approach, we used triple SILAC labelling.

However, the enrichment above the spatial control does not seem to have been successful when APEX2-RHBDL4 was "underexpressed", because even the VCP ATPase which is known to interact with RHBDL4 was more abundant in the control cytosolic proteome. Therefore, it was not possible to evaluate the data based on the

spatial control, and the cut-off for distinguishing the RHBDL4 proximity proteome from non-specifically labelled proteins was set based on the VCP ATPase as a positive control.

This problem was unexpected, because the preliminary APEX2 experiment not using SILAC labelling showed much more enrichment of VCP and many other proteins. The failure of the spatial control could have been due to the very low expression of APEX2-RHBDL4, as the retroviral generated cell lines slightly overexpressing the fusion protein were successfully used in the preliminary experiment. Perhaps, differences in experimental conditions of the preliminary versus SILAC experiments might have contributed too, including significantly longer handling of the cells during the labelling reaction (because there were twice as many samples in the SILAC versus preliminary experiment), SILAC approach itself and differences in mass-spectrometry analysis (length and steepness of the HPLC gradient during LC-MS/MS may influence sample complexity and thus detection of low-abundance species).

The longer handling time with cells should not be problematic, as we used an inversion technique, which should assure the same timing of biotinylation reaction in all samples (Hung et al. 2016). The cells were cultured in sealable flasks, which can be inverted upside down. Hence, the labelling medium was first added to the flasks when turned upside down (so that the medium was not in contact with the cells), and then the flasks were all inverted back at once, so that the labelling medium came into contact with the cells evenly and at the same moment. The labelling was then stopped after exactly 1 min similarly, and all the washing steps were carried out in a similar manner.

The SILAC labelling is a frequently used tool when cellular proteomes are quantitatively compared (Ong et al. 2002). It is recommended in the protocol for APEX2-based proteomics (Hung et al. 2016) and it was used for identification of ER-membrane proteome via APEX2 (Hung et al. 2017). Therefore, it does not seem to be likely that SILAC itself would be the primary cause of the observed problems in my experiment. Rather, the simultaneous purification of biotinylated proteins from the pooled experimental and control samples might have contributed to the problems. If the abundance of biotinylated proteins was significantly higher in the

control, these proteins would be captured on the resin with higher probability, possibly competing with low-abundance proteins labelled in the proximity of RHBDL4. However, comparison of the activity of APEX2 between the experimental and control sample did not reveal a significant difference in the total abundance of biotinylated proteins (data not shown), making this possibility unlikely. Moreover, in the study of ER-membrane proteome (Hung et al. 2017) the activity of cytosolic APEX2 was apparently higher compared to the ER anchored APEX2, but it was still possible to enrich ER-associated proteins above the control using SILAC. Therefore, SILAC labelling itself and different abundances of biotinylated proteins in the spatial control and RHBDL4 proximity samples should not be the main problems.

To conclude, based on all available data, the most likely cause of the failure of enrichment above the spatial cytosolic control is the low level of expression of APEX2-RHBDL4 fusion protein. Basically, there is not enough RHBDL4 to fish out its interacting partners from cytosol and so they are predominantly labelled by cytosolic APEX2.

9.3.2 Problems encountered – RHBDL4 missing in its proximity proteome

We expected to find two control proteins in the RHBDL4 proximity proteome – VCP ATPase which is known to interact with RHBDL4, and RHBDL4 itself. Indeed, APEX2 enzyme fused to RHBDL4 should induce biotinylation of RHBDL4 part of the fusion. Interestingly, RHBDL4 was not detected by mass spectrometry in any of the SILAC-based experiments, whereas it was detected and significantly enriched over cytosolic control in the preliminary experiment (without SILAC labelling). However, we could detect RHBDL4 by western blot in the eluates from streptavidin column (Fig.17b), indicating that the absence of RHBDL4 in the SILAC experiment was due to problems on the side of mass spectrometry. As an integral membrane protein, RHBDL4 will produce only a limited number of hydrophilic peptides with good ionization properties easily detectable in the mass spectrometer. It is well known that the hydrophobic peptides originating from transmembrane regions are much harder to detect in the presence of abundant hydrophilic peptides. Therefore, the shorter HPLC gradient used for SILAC-based experiments together with the simultaneous analysis of the experimental and cytosolic control samples

could cause this problem. Moreover, this might have been aggravated by the low expression level of APEX2-RHBDL4.

Importantly, this issue highlights a weak point of mass spectrometry – we might be losing not only RHBDL4, but also other membrane proteins. This might be the reason why membrane proteins form only a small portion of the identified RHBDL4 proximity proteome. It might be possible to solve this issue by employing cell fractionation and analysing ER membranes separately from the cytosolic fraction.

9.4 The proximity proteome of RHBDL4

Despite the shortcomings in the proteomics experiments described above, which will require further work and probably modification of the procedure, we can nevertheless extract some useful information from it. The VCP ATPase was present in all three proximity proteomics data sets, and it was considered as a positive control to set the cut-off for distinguishing the RHBDL4 proximity proteome members from non-specifically labelled proteins in the SILAC-based experiments. This revealed three other proteins, which were present in the proximity proteome of RHBDL4 repeatedly in the three independent experiments – ITGB1, PLIN3, ANXA7. The first experiment was significantly different in the procedure from the other two, which further supports the likely significance of these protein hits. Several proteins were identified in two independent experiments, but their relevance should be further verified. Because of the problems with spatial control in SILAC-based experiments, there is a possibility that some randomly labelled proteins are incorrectly included in the RHBDL4 proximity proteome.

Integrin beta 1 is the most common beta subunit among integrin alpha/beta heterodimers (Moreno-Layseca et al. 2019). Integrins are transmembrane receptors for extracellular matrix (ECM) ligands such as fibronectin, collagen or laminin and interact with some cell adhesion molecules such as VCAM1, ICAM or E-cadherin. They are activated via intracellular signalling, changing their conformation to efficiently bind extracellular ligands, which in turn transmits a signal back to the cell. Therefore, integrins enable cells to respond to their environment, to adhere and move.

In addition to the regulation of their conformation, integrins are extensively trafficked within cells, what determines their availability on the cell surface (Moreno-Layseca et al. 2019). The spatiotemporal recycling is vital for cell migration and even internalized integrins can promote specific signalling pathways e.g. suppressing anoikis (Alanko et al. 2015). It is tempting to speculate that RHBDL4 might be involved in the trafficking of integrins or their assembly into complexes. Interestingly, integrins can localize to perinuclear vesicle compartments (Dozynkiewicz et al. 2012) which are overlapping with the ER, where RHBDL4 is localized (Wunderle et al. 2016; Fleig et al. 2012). Moreover, the degradation vs. recycling of integrins is regulated by ubiquitination of their cytosolic domain (Lobert et al. 2010) and RHBDL4 possess the ubiquitin binding motif facing cytosol (Fleig et al. 2012). ITGB1 may be also direct substrate of RHBDL4 protease.

Integrins are also secreted in exosomes, whose production/protein composition is influenced by RHBDL4 (Wunderle et al. 2016; Wan et al. 2012). The integrins in exosomes can be secreted together with bound ECM proteins such as fibronectin and create a "footpath" for tumour cells, assuring long-lasting oriented migration (Sung et al. 2015). Tumour derived exosomes are also able to prepare a niche for metastasis in an organ-specific fashion, provided by integrins (Hoshino et al. 2015). Exosomes with integrins are produced by immune cells too, can interact with other cells through adhesion molecules and induce a calcium response (Clayton et al. 2004). Taken together, regulation of trafficking and secretion of integrins could well connect the molecular function of RHBDL4 and its cancer promoting effects described in the literature. Moreover, in the physiological setup, integrins are vital for the proper function of immune system (Ley et al. 2007; Evans et al. 2009; Zhang & Wang 2012) which is affected in RHBDL4 deficient mouse (Bohacova & Strisovsky, unpublished data).

Perilipin 3 belongs to the perilipin protein family, which are the most prominent lipid droplet associated proteins (Kimmel & Sztalryd 2016). They surround lipid droplets, which are comprised of triacylglycerols, cholesteryl-esters, phospholipid monolayer and surface proteins, and regulate the access of other cellular proteins (e.g., lipases) to the lipids. Perilipins are also important for the stability of lipid droplets. PLIN3 is widely distributed among mammalian cell types

(Kimmel & Sztalryd 2016) and it can participate in regulation of lipolysis during starvation (Kaushik & Cuervo 2015), although the major player regulating lipid metabolism in adipocytes is perilipin 1 (Kimmel & Sztalryd 2016). PLIN3 also regulates trafficking of mannose-6-phosphate receptor which transports lysosomal hydrolases from Golgi to lysosomes, by binding to its cytoplasmic tail from the cytosolic side (Díaz & Pfeffer 1998).

Lipid droplets are formed from the ER membrane and PLIN3 is one of the first perilipins that binds to the emerging droplet at the ER (Skinner et al. 2009), therefore PLIN3 and RHBDL4 might be in direct contact. It seems unlikely that PLIN3 is a substrate of RHBDL4, because PLIN3 is believed to be only peripherally associated with the lipid droplet membrane from the cytosolic side (Kimmel & Sztalryd 2016) while the active site of RHBDL4 is opened to the opposite side of the ER membrane, into the ER lumen. Lipid droplets eventually detach from the ER, they can fuse and even be secreted (McManaman 2014), therefore there is definitely a potential for some regulatory mechanisms analogous to the vesicle trafficking from ER, in which RHBDL4 might participate (Wunderle et al. 2016).

Most importantly, ERAD in which RHBDL4 was implicated (Fleig et al. 2012), can occur in lipid-droplet-associated ER membranes, regulating the sterol synthesis via ubiquitination and degradation of HMG-CoA reductase and SREBPs (Jo et al. 2013). In addition, RHBDL4 was recently suggested to bind and be regulated by cholesterol (Paschkowsky et al. 2018). Therefore, it is possible to speculate that RHBDL4 might respond to the cellular sterol levels and support ERAD of protein machinery synthesising sterols and lipids. RHBDL4 would not be the first rhomboid family protein connected with lipid droplets and lipid metabolism – the UBAC2 rhomboid pseudoprotease interacts with UBXD8 which regulates the activity of ATGL on lipid droplets (described in detail in the section 4.1.2) (Olzmann et al. 2013).

Lipid droplets or lipid bodies are found in leukocytes, they are induced by inflammation and enable synthesis of eicosanoids (Melo et al. 2011). In LPS activated neutrophils, PLIN3 is vital for the formation of lipid droplets and prostaglandin E2 (PGE_2) production (Nose et al. 2013). In macrophages, PLIN3 and lipid droplets are stimulated by high glucose and insulin, potentially leading to the

emergence of foam cells involved in atherosclerosis (Fan et al. 2013). Therefore, if RHBDL4 was involved in lipid droplet biogenesis or its intracellular trafficking, there would be a potential connection to the immune system function, because eicosanoids are very potent inflammatory and immune regulatory molecules (Dennis & Norris 2015).

Annexin A7 belongs to the annexin protein family which share the ability to peripherally associate with lipid membranes upon calcium influx (Gerke et al. 2005). Thus, annexins are involved in various membrane-associated events accompanied by calcium signalling, such as exocytosis, endocytosis, membrane domain formation or membrane connection with cytoskeletal proteins; some annexins even have extracellular roles. The calcium-membrane binding domain is shared among annexins, but their N-terminal domains are different as well as their specific molecular functions.

The function of ANXA7 in particular is unfortunately not understood in detail. It was shown to bring two membranes close to each other, which can precede their fusion (Pollard et al. 1988), and it is localised mostly on secretory vesicles, nuclear envelope and plasma membrane (Kuijpers et al. 1992). Therefore, it might be involved in the secretory pathway, but it is not clear if it could actually co-localise with RHBDL4. Interestingly, ANXA7 and ANXA11 interact with another calcium binding protein ALG-2 (Satoh et al. 2002). Recently, ANXA11 was shown to interact with Sec-31 via ALG-2 and to support association of Sec-31 with ERES, therefore influencing the structure and function of ERES (Shibata et al. 2015). Because ANXA7 interacts with ALG-2 too, there is a possibility that it can also localise to ERES, where the interaction with RHBDL4 could occur. However, as in the case of PLIN3, ANXA7 is most likely not a substrate of RHBDL4 protease as it is supposed to be a peripheral membrane protein. Interestingly, ANXA7 negatively regulates PGE2 synthesis (Luo et al. 2015), in which PLIN3 is also involved.

To conclude, the protein candidates identified by APEX2 approach seem to have a potential to interact with RHBDL4. Moreover, they might explain the cancer promoting effects upon RHBDL4 overexpression (Song et al. 2015) and the immune aberrances observed upon RHBDL4 deficiency (Bohacova & Strisovsky, unpublished data). But a further validation of these interaction candidates is needed.

9.5 Future perspectives

Following this study, I would like to repeat the SILAC based experiments in cells expressing higher amounts of APEX2-RHBDL4 (used in the first non-SILAC experiment and created by the retroviral delivery of expression cassette), to overcome the problems in enrichment observed in the SILAC experiments presented in this thesis, further confirm the initial proximity protein candidates obtained in this thesis, and possibly widen the RHBDL4 proximity proteome.

The proximity proteome of RHBDL4 should be then verified by independent approach. Therefore, we would like to explore the co-localisation of the candidate proteins with RHBDL4 in living cells using fluorescent labels and high-resolution microscopy. In particular, super-resolution microscopy would be ideal, because the ER membranes are tightly packed and therefore it can be demanding to assess the true distance between individual fluorescent signals by conventional confocal microscopy. There is also a possibility to explore the proximity proteome of the catalytically inactive mutant of RHBDL4, which could facilitate identification of RHBDL4 substrates, because the inactive mutant can bind and trap them, while the wild-type will cleave them, thereby changing their localisation.

Finally, this study is only the first step to explain the physiological function of RHBDL4. It was performed in human colorectal cell line (HCT116), which is relevant to RHBDL4 function in cancer development but probably less so for its physiological role, considering the immune phenotype of RHBDL4 deficient mice. Therefore, we plan to proceed with this proteomic analysis in mouse cells particularly important for the phenotype of RHBDL4 deficient mouse which is concurrently studied. Ideally, both approaches will synergize, connecting the molecular function of RHBDL4 with its overall physiological relevance at the organismal level.

References

Abba, M.C. et al., 2009. Rhomboid domain containing 2 (RHBDD2): a novel cancer-related gene over-expressed in breast cancer. *Biochimica et biophysica acta*, 1792(10), pp.988–97.

Adrain, C. et al., 2012. Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE. *Science (New York, N.Y.)*, 335(6065), pp.225–8.

Ahmedli, N.B. et al., 2013. Dynamics of the rhomboid-like protein RHBDD2 expression in mouse retina and involvement of its human ortholog in retinitis pigmentosa. *The Journal of biological chemistry*, 288(14), pp.9742–54.

Akiyama, Y., Kanehara, K. & Ito, K., 2004. RseP (YaeL), an Escherichia coli RIP protease, cleaves transmembrane sequences. *The EMBO journal*, 23(22), pp.4434–42.

Al-Maskari, M. et al., 2018. Site-1 protease function is essential for the generation of antibody secreting cells and reprogramming for secretory activity. *Scientific reports*, 8(1), p.14338.

Alanko, J. et al., 2015. Integrin endosomal signalling suppresses anoikis. *Nature Cell Biology*, 17(11), pp.1412–1421.

Bai, X. et al., 2015. An atomic structure of human γ-secretase. *Nature*, 525(7568), pp.212–217.

Bailey, D. & O'Hare, P., 2007. Transmembrane bZIP Transcription Factors in ER Stress Signaling and the Unfolded Protein Response. *Antioxidants & Redox Signaling*, 9(12), pp.2305–2322.

Barrett, M.T. et al., 2019. Clonal analyses of refractory testicular germ cell tumors A. Vicha, ed. *PLOS ONE*, 14(3), p.e0213815.

Beisner, D.R. et al., 2013. The intramembrane protease Sppl2a is required for B cell and DC development and survival via cleavage of the invariant chain. *The Journal of experimental medicine*, 210(1), pp.23–30.

Bergbold, N. & Lemberg, M.K., 2013. Emerging role of rhomboid family proteins in mammalian biology and disease. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1828(12), pp.2840–2848.

Bergo, M.O. et al., 2002. Absence of the CAAX endoprotease Rce1: effects on cell growth and transformation. *Molecular and cellular biology*, 22(1), pp.171–81.

Bergo, M.O. et al., 2004. On the physiological importance of endoproteolysis of CAAX proteins: heart-specific RCE1 knockout mice develop a lethal cardiomyopathy. *The Journal of biological chemistry*, 279(6), pp.4729–36.

Bondar, A.-N., del Val, C. & White, S.H., 2009. Rhomboid Protease Dynamics and Lipid Interactions. *Structure*, 17(3), pp.395–405.

Boyartchuk, V.L., Ashby, M.N. & Rine, J., 1997. Modulation of Ras and a-factor function by carboxyl-terminal proteolysis. *Science (New York, N.Y.)*, 275(5307), pp.1796–800.

Braud, V.M. et al., 1998. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature*, 391(6669), pp.795–799.

Brown, M.S. et al., 2000. Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell*, 100(4), pp.391–8.

Brown, M.S. & Goldstein, J.L., 1997. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell*, 89(3), pp.331–40.

Brunelle, J.L. & Green, R., 2014. One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). In *Methods in enzymology*. pp. 151–159.

Burrows, J.F. et al., 2009. USP17 Regulates Ras Activation and Cell Proliferation by Blocking RCE1 Activity. *Journal of Biological Chemistry*, 284(14), pp.9587–9595.

Canzoneri, R. et al., 2018. Alternative splicing variant of RHBDD2 is associated with cell stress response and breast cancer progression. *Oncology Reports*, 40(2), pp.909–915.

Cao, Q. et al., 2018. PAQR3 Regulates Endoplasmic Reticulum-to-Golgi Trafficking of COPII Vesicle via Interaction with Sec13/Sec31 Coat Proteins. *iScience*, 9, pp.382–398.

Clayton, A. et al., 2004. Adhesion and signaling by B cell-derived exosomes: the role of integrins. *The FASEB Journal*, 18(9), pp.977–979.

Dennis, E.A. & Norris, P.C., 2015. Eicosanoid storm in infection and inflammation. *Nature reviews. Immunology*, 15(8), pp.511–23.

Díaz, E. & Pfeffer, S.R., 1998. TIP47: a cargo selection device for mannose 6-phosphate receptor trafficking. *Cell*, 93(3), pp.433–43.

Dong, Q. et al., 2017. Derlin-1 overexpression confers poor prognosis in muscle invasive bladder

cancer and contributes to chemoresistance and invasion through PI3K/AKT and ERK/MMP signaling. *Oncotarget*, 8(10), pp.17059–17069.
Dozynkiewicz, M.A. et al., 2012. Rab25 and CLIC3 collaborate to promote integrin recycling from late endosomes/lysosomes and drive cancer progression. *Developmental cell*, 22(1), pp.131–45.
Dulloo, I., Muliyil, S. & Freeman, M., 2019. The molecular, cellular and pathophysiological roles of iRhom pseudoproteases. *Open biology*, 9(3), p.190003.
Ebsen, H. et al., 2015. Subcellular localization and activation of ADAM proteases in the context of FasL shedding in T lymphocytes. *Molecular Immunology*, 65(2), pp.416–428.
Eferl, R. & Wagner, E.F., 2003. AP-1: A double-edged sword in tumorigenesis. *Nature Reviews Cancer*, 3(11), pp.859–868.
Eleveld-Trancikova, D. et al., 2010. DC-STAMP interacts with ER-resident transcription factor LUMAN which becomes activated during DC maturation. *Molecular Immunology*, 47(11–12), pp.1963–1973.
Evans, R. et al., 2009. Integrins in immunity. *Journal of Cell Science*, 122(2), pp.215–225.
Fan, B. et al., 2013. High glucose, insulin and free fatty acid concentrations synergistically enhance perilipin 3 expression and lipid accumulation in macrophages. *Metabolism*, 62(8), pp.1168–1179.
Feng, L. et al., 2007. Structure of a Site-2 Protease Family Intramembrane Metalloprotease. *Science*, 318(5856), pp.1608–1612.
Ferretti, V.A. et al., 2015. Spatiotemporal expression of Rhomboid domain containing 2 (Rhbdd2) during rat development. *Acta Histochemica*, 117(7), pp.635–641.
Fisher, P. & Ungar, D., 2016. Bridging the Gap between Glycosylation and Vesicle Traffic. *Frontiers in Cell and Developmental Biology*, 4, p.15.
Fleig, L. et al., 2012. Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. *Molecular cell*, 47(4), pp.558–569.
Fluhrer, R. et al., 2006. A γ-secretase-like intramembrane cleavage of TNFα by the GxGD aspartyl protease SPPL2b. *Nature Cell Biology*, 8(8), pp.894–896.
Franke, M. et al., 2016. Human and Murine Interleukin 23 Receptors Are Novel Substrates for A Disintegrin and Metalloproteases ADAM10 and ADAM17. *Journal of Biological Chemistry*, 291(20), pp.10551–10561.
Friedmann, E. et al., 2006. SPPL2a and SPPL2b promote intramembrane proteolysis of TNFα in activated dendritic cells to trigger IL-12 production. *Nature Cell Biology*, 8(8), pp.843–848.
Fujinaga, M. et al., 2004. The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum. *Proceedings of the National Academy of Sciences*, 101(10), pp.3364–3369.
Gerke, V., Creutz, C.E. & Moss, S.E., 2005. Annexins: linking Ca2+ signalling to membrane dynamics. *Nature Reviews Molecular Cell Biology*, 6(6), pp.449–461.
Gibson, D.G. et al., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5), pp.343–345.
Green, M.R. & Sambrook, J., 2012. *Molecular cloning: a laboratory manual*, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.
Greenblatt, E.J., Olzmann, J.A. & Kopito, R.R., 2011. Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant α-1 antitrypsin from the endoplasmic reticulum. *Nature structural & molecular biology*, 18(10), pp.1147–52.
Ha, Y., 2009. Structure and mechanism of intramembrane protease. *Seminars in cell & developmental biology*, 20(2), pp.240–50.
Hamblet, C.E. et al., 2016. NK Cell Maturation and Cytotoxicity Are Controlled by the Intramembrane Aspartyl Protease SPPL3. *Journal of immunology (Baltimore, Md. : 1950)*, 196(6), pp.2614–26.
Han, J. et al., 2015. Lentivirus-mediated knockdown of rhomboid domain containing 1 inhibits colorectal cancer cell growth. *Molecular Medicine Reports*, 12(1), pp.377–381.
Hass, M.R. et al., 2009. Presenilin: RIP and beyond. *Seminars in Cell & Developmental Biology*, 20(2), pp.201–210.
Haze, K. et al., 1999. Mammalian Transcription Factor ATF6 Is Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress P. Silver, ed. *Molecular Biology of the Cell*, 10(11), pp.3787–3799.
Hetz, C., 2012. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology*, 13(2), pp.89–102.

Hetz, C. & Papa, F.R., 2018. The Unfolded Protein Response and Cell Fate Control. *Molecular Cell*, 69(2).
Hoshino, A. et al., 2015. Tumour exosome integrins determine organotropic metastasis. *Nature*, 527(7578), pp.329–335.
Huang, C. et al., 2018. Silencing of rhomboid domain containing 1 to inhibit the metastasis of human breast cancer cells in vitro. *Iranian journal of basic medical sciences*, 21(11), pp.1161–1166.
Hung, V. et al., 2017. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *eLife*, 6.
Hung, V. et al., 2016. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nature protocols*, 11(3), pp.456–75.
Hwang, J. et al., 2016. A Golgi rhomboid protease Rbd2 recruits Cdc48 to cleave yeast SREBP. *The EMBO journal*, 35(21), pp.2332–2349.
Chan, C.-P., Kok, K.-H. & Jin, D.-Y., 2011. CREB3 subfamily transcription factors are not created equal: Recent insights from global analyses and animal models. *Cell & bioscience*, 1(1), p.6.
Chen, Q., Sun, L. & Chen, Z.J., 2016. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nature Immunology*, 17(10), pp.1142–1149.
Chitadze, G. et al., 2013. Shedding of endogenous MHC class I-related chain molecules A and B from different human tumor entities: Heterogeneous involvement of the "a disintegrin and metalloproteases" 10 and 17. *International Journal of Cancer*, 133(7), pp.1557–1566.
Christiansen, J.R. et al., 2011. RAS-converting enzyme 1-mediated endoproteolysis is required for trafficking of rod phosphodiesterase 6 to photoreceptor outer segments. *Proceedings of the National Academy of Sciences*, 108(21), pp.8862–8866.
Christianson, J.C. et al., 2011. Defining human ERAD networks through an integrative mapping strategy. *Nature cell biology*, 14(1), pp.93–105.
Ishikawa, H., Ma, Z. & Barber, G.N., 2009. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature*, 461(7265), pp.788–792.
Itzhak, D.N. et al., 2016. Global, quantitative and dynamic mapping of protein subcellular localization. *eLife*, 5.
Ivanova, A. et al., 2008. Loss of PL6 protein expression in renal clear cell carcinomas and other *VHL* -deficient tumours. *The Journal of Pathology*, 214(1), pp.46–57.
Jo, Y., Hartman, I.Z. & DeBose-Boyd, R.A., 2013. Ancient ubiquitous protein-1 mediates sterol-induced ubiquitination of 3-hydroxy-3-methylglutaryl CoA reductase in lipid droplet–associated endoplasmic reticulum membranes R. G. Parton, ed. *Molecular Biology of the Cell*, 24(3), pp.169–183.
Johnson, N. et al., 2017. Quantitative proteomics screen identifies a substrate repertoire of rhomboid protease RHBDL2 in human cells and implicates it in epithelial homeostasis. *Scientific Reports*, 7(1), p.7283.
Jurgens, G. et al., 1984. Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster. *Wilhelm Roux's Archives of Developmental Biology*, 193(5), pp.283–295.
Käll, L., Krogh, A. & Sonnhammer, E.L., 2004. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology*, 338(5), pp.1027–1036.
Kanzaki, H. et al., 2016. A-Disintegrin and Metalloproteinase (ADAM) 17 Enzymatically Degrades Interferon-gamma. *Scientific reports*, 6, p.32259.
Kaushik, S. & Cuervo, A.M., 2015. Degradation of lipid droplet-associated proteins by chaperone-mediated autophagy facilitates lipolysis. *Nature Cell Biology*, 17(6), pp.759–770.
Kawai, T. et al., 2005. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nature Immunology*, 6(10), pp.981–988.
Kell, A.M., Gale, M. & Jr, 2015. RIG-I in RNA virus recognition. *Virology*, 479–480, pp.110–21.
Kim, D.I. et al., 2016. An improved smaller biotin ligase for BioID proximity labeling Y. Zheng, ed. *Molecular Biology of the Cell*, 27(8), pp.1188–1196.
Kim, E. et al., 1999. Disruption of the mouse Rce1 gene results in defective Ras processing and mislocalization of Ras within cells. *The Journal of biological chemistry*, 274(13), pp.8383–90.
Kimmel, A.R. & Sztalryd, C., 2016. The Perilipins: Major Cytosolic Lipid Droplet–Associated Proteins and Their Roles in Cellular Lipid Storage, Mobilization, and Systemic Homeostasis. *Annual Review of Nutrition*, 36(1), pp.471–509.
Kirkin, V. et al., 2007. The Fas ligand intracellular domain is released by ADAM10 and SPPL2a cleavage in T-cells. *Cell Death & Differentiation*, 14(9), pp.1678–1687.
Knop, M. et al., 1996. Der1, a novel protein specifically required for endoplasmic reticulum

degradation in yeast. *The EMBO journal*, 15(4), pp.753–63.
Kreutzberger, A.J.B. et al., 2019. Rhomboid distorts lipids to break the viscosity-imposed speed limit of membrane diffusion. *Science (New York, N.Y.)*, 363(6426), p.eaao0076.
Kuhn, P.-H. et al., 2015. Secretome Analysis Identifies Novel Signal Peptide Peptidase-Like 3 (SPPL3) Substrates and Reveals a Role of SPPL3 in Multiple Golgi Glycosylation Pathways. *Molecular & Cellular Proteomics*, 14(6), pp.1584–1598.
Kuijpers, G.A., Lee, G. & Pollard, H.B., 1992. Immunolocalization of synexin (annexin VII) in adrenal chromaffin granules and chromaffin cells: evidence for a dynamic role in the secretory process. *Cell and tissue research*, 269(2), pp.323–30.
Lacunza, E. et al., 2014. Identification of signaling pathways modulated by RHBDD2 in breast cancer cells: a link to the unfolded protein response. *Cell Stress and Chaperones*, 19(3), pp.379–388.
Lacunza, E. et al., 2012. RHBDD2: a 5-fluorouracil responsive gene overexpressed in the advanced stages of colorectal cancer. *Tumor Biology*, 33(6), pp.2393–2399.
Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), pp.680–5.
Lam, S.S. et al., 2014. Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nature Methods*, 12(1), pp.51–54.
Lambrecht, B.N., Vanderkerken, M. & Hammad, H., 2018. The emerging role of ADAM metalloproteinases in immunity. *Nature Reviews Immunology*, 18(12), pp.745–758.
Lastun, V.L., Grieve, A.G. & Freeman, M., 2016. Substrates and physiological functions of secretase rhomboid proteases. *Seminars in cell & developmental biology*, 60, pp.10–18.
Lee, J.R. et al., 2001. Regulated Intracellular Ligand Transport and Proteolysis Control EGF Signal Activation in Drosophila. *Cell*, 107(2), pp.161–171.
LeGendre, N., 1990. Immobilon-P transfer membrane: applications and utility in protein biochemical analysis. *BioTechniques*, 9(6 Suppl), pp.788–805.
Lemberg, M.K. et al., 2005. Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *The EMBO journal*, 24(3), pp.464–72.
Lemberg, M.K. et al., 2001. The Journal of Immunology. *J. Immunol.*, 159(5), pp.2139–2146.
Lemberg, M.K. & Adrain, C., 2016. Inactive rhomboid proteins: New mechanisms with implications in health and disease. *Seminars in cell & developmental biology*, 60, pp.29–37.
Lemberg, M.K. & Freeman, M., 2007. Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases. *Genome research*, 17(11), pp.1634–46.
Ley, K. et al., 2007. Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nature Reviews Immunology*, 7(9), pp.678–689.
Li, X. et al., 2015. iRhoms 1 and 2 are essential upstream regulators of ADAM17-dependent EGFR signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 112(19), pp.6080–5.
Li, X. et al., 2012. Structure of a presenilin family intramembrane aspartate protease. *Nature*, 493(7430), pp.56–61.
Lichtenthaler, S.F., Lemberg, M.K. & Fluhrer, R., 2018. Proteolytic ectodomain shedding of membrane proteins in mammals—hardware, concepts, and recent developments. *The EMBO Journal*, 37(15), p.e99456.
Lilley, B.N. & Ploegh, H.L., 2004. A membrane protein required for dislocation of misfolded proteins from the ER. *Nature*, 429(June), pp.834–840.
Lim, J.J. et al., 2016. Structural insights into the interaction of p97 N-terminus domain and VBM in rhomboid protease, RHBDL4. *Biochemical Journal*, 473(18), pp.2863–2880.
Liu, J. et al., 2014. Rhbdd3 controls autoimmunity by suppressing the production of IL-6 by dendritic cells via K27-linked ubiquitination of the regulator NEMO. *Nature Immunology*, 15(7), pp.612–622.
Liu, J. et al., 2013. Rhomboid domain-containing protein 3 is a negative regulator of TLR3-triggered natural killer cell activation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), pp.7814–9.
Liu, X.-N. et al., 2013. Lentivirus-mediated silencing of rhomboid domain containing 1 suppresses tumor growth and induces apoptosis in hepatoma HepG2 cells. *Asian Pacific journal of cancer prevention : APJCP*, 14(1), pp.5–9.
Lloyd, S.J.-A., Raychaudhuri, S. & Espenshade, P.J., 2013. Subunit architecture of the Golgi Dsc E3 ligase required for sterol regulatory element-binding protein (SREBP) cleavage in fission yeast. *The Journal of biological chemistry*, 288(29), pp.21043–54.

Lobert, V.H. et al., 2010. Ubiquitination of alpha 5 beta 1 integrin controls fibroblast migration through lysosomal degradation of fibronectin-integrin complexes. *Developmental cell*, 19(1), pp.148–59.
Lobingier, B.T. et al., 2017. An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell*, 169(2), p.350–360.e12.
Luo, D. et al., 2015. Influence of annexin A7 on insulin sensitivity of cellular glucose uptake. *Pflügers Archiv - European Journal of Physiology*, 467(4), pp.641–649.
Luo, W.-W. et al., 2016. iRhom2 is essential for innate immunity to DNA viruses by mediating trafficking and stability of the adaptor STING. *Nature Immunology*, 17(9), pp.1057–1066.
Luo, W.-W. et al., 2017. iRhom2 is essential for innate immunity to RNA virus by antagonizing ER- and mitochondria-associated degradation of VISA. *PLoS pathogens*, 13(11), p.e1006693.
Makowski, S.L., Wang, Z. & Pomerantz, J.L., 2015. A protease-independent function for SPPL3 in NFAT activation. *Molecular and cellular biology*, 35(2), pp.451–67.
Manolaridis, I. et al., 2013. Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1. *Nature*, 504(7479), pp.301–5.
Martin, L. et al., 2008. Regulated Intramembrane Proteolysis of Bri2 (Itm2b) by ADAM10 and SPPL2a/SPPL2b. *Journal of Biological Chemistry*, 283(3), pp.1644–1652.
Maske, C.P. et al., 2003. A carboxyl-terminal interaction of lamin B1 is dependent on the CAAX endoprotease Rce1 and carboxymethylation. *The Journal of Cell Biology*, 162(7), pp.1223–1232.
Masters, C.L. et al., 1985. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 82(12), pp.4245–9.
McIlwain, D.R. et al., 2012. iRhom2 regulation of TACE controls TNF-mediated protection against Listeria and responses to LPS. *Science (New York, N.Y.)*, 335(6065), pp.229–32.
McLauchlan, J. et al., 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *The EMBO journal*, 21(15), pp.3980–8.
McManaman, J.L., 2014. Lipid transport in the lactating mammary gland. *Journal of mammary gland biology and neoplasia*, 19(1), pp.35–42.
Mehnert, M., Sommer, T. & Jarosch, E., 2014. Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane. *Nature cell biology*, 16(1), pp.77–86.
Melo, R.C.N. et al., 2011. Lipid Bodies in Inflammatory Cells. *Journal of Histochemistry & Cytochemistry*, 59(5), pp.540–556.
Meylan, E. et al., 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature*, 437(7062), pp.1167–1172.
Miao, F. et al., 2017. RHBDD1 upregulates EGFR via the AP-1 pathway in colorectal cancer. *Oncotarget*, 8(15), pp.25251–25260.
Moin, S.M. & Urban, S., 2012. Membrane immersion allows rhomboid proteases to achieve specificity by reading transmembrane segment dynamics. *eLife*, 1.
Moreno-Layseca, P. et al., 2019. Integrin trafficking in cells and tissues. *Nature Cell Biology*, 21(2), pp.122–132.
Moriya, H., 2015. Quantitative nature of overexpression experiments. *Molecular biology of the cell*, 26(22), pp.3932–9.
Mullick, A. et al., 2006. The cumate gene-switch: a system for regulated expression in mammalian cells. *BMC Biotechnology*, 6(1), p.43.
Neal, S. et al., 2018. The Dfm1 Derlin Is Required for ERAD Retrotranslocation of Integral Membrane Proteins. *Molecular Cell*, 69(2), p.306–320.e4.
Nishitoh, H. et al., 2008. ALS-linked mutant SOD1 induces ER stress- and ASK1-dependent motor neuron death by targeting Derlin-1. *Genes & Development*, 22(11), pp.1451–1464.
Nose, F. et al., 2013. Crucial Role of Perilipin-3 (TIP47) in Formation of Lipid Droplets and PGE2 Production in HL-60-Derived Neutrophils H. Yang, ed. *PLoS ONE*, 8(8), p.e71542.
O'Brien, R.J. & Wong, P.C., 2011. Amyloid precursor protein processing and Alzheimer's disease. *Annual review of neuroscience*, 34, pp.185–204.
Oda, Y. et al., 2006. Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation. *The Journal of cell biology*, 172(3), pp.383–93.
Oikonomidi, I. et al., 2018. iTAP, a novel iRhom interactor, controls TNF secretion by policing the stability of iRhom/TACE. *eLife*, 7.
Olzmann, J.A., Richter, C.M. & Kopito, R.R., 2013. Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proceedings of the National Academy of*

Sciences of the United States of America, 110(4), pp.1345–50.
Ong, S.-E. et al., 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP*, 1(5), pp.376–86.
Ong, Y.S. et al., 2014. TMEM115 is an integral membrane protein of the Golgi complex involved in retrograde transport. *Journal of Cell Science*, 127(13), pp.2825–2839.
Otto, J.C. et al., 1999. Cloning and characterization of a mammalian prenyl protein-specific protease. *The Journal of biological chemistry*, 274(13), pp.8379–82.
Paek, J. et al., 2017. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. *Cell*, 169(2), p.338–349.e11.
Pascall, J.C. & Brown, K.D., 1998. Characterization of a mammalian cDNA encoding a protein with high sequence similarity to the Drosophila regulatory protein Rhomboid. *FEBS Letters*, 429(3), pp.337–340.
Paschkowsky, S. et al., 2016. Alternative Processing of the Amyloid Precursor Protein Family by Rhomboid Protease RHBDL4. *Journal of Biological Chemistry*, 291(42), pp.21903–21912.
Paschkowsky, S. et al., 2018. Membrane cholesterol as regulator of human rhomboid protease RHBDL4. *Journal of Biological Chemistry*, 293(40), pp.15556–15568.
Pène, V. et al., 2009. Sequential processing of hepatitis C virus core protein by host cell signal peptidase and signal peptide peptidase: a reassessment. *Journal of Viral Hepatitis*, 16(10), pp.705–715.
Pi, L. et al., 2017. Elevated expression of Derlin-1 associates with unfavorable survival time of squamous cell carcinoma of the head and neck and promotes its malignance. *Journal of Cancer*, 8(12), pp.2336–2345.
Pollard, H.B., Burns, A.L. & Rojas, E., 1988. A molecular basis for synexin-driven, calcium-dependent membrane fusion. *The Journal of experimental biology*, 139, pp.267–86.
Raggo, C. et al., 2002. Luman, the cellular counterpart of herpes simplex virus VP16, is processed by regulated intramembrane proteolysis. *Molecular and cellular biology*, 22(16), pp.5639–49.
Randow, F. & Sale, J.E., 2006. Retroviral transduction of DT40. *Sub-cellular biochemistry*, 40, pp.383–6.
Rawson, R.B. et al., 1997. Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Molecular cell*, 1(1), pp.47–57.
Recinto, S.J., Paschkowsky, S. & Munter, L.M., 2018. An alternative processing pathway of APP reveals two distinct cleavage modes for rhomboid protease RHBDL4. *Biological Chemistry*, 399(12), pp.1399–1408.
Rhee, H.-W. et al., 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science (New York, N.Y.)*, 339(6125), pp.1328–1331.
Roux, K.J. et al., 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of cell biology*, 196(6), pp.801–10.
Roux, K.J. et al., 2018. BioID: A Screen for Protein-Protein Interactions. *Current protocols in protein science*, 91, p.19.23.1-19.23.15.
Rudner, D.Z., Fawcett, P. & Losick, R., 1999. A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, 96(26), pp.14765–70.
Saito, A. et al., 2011. Post-liberation cleavage of signal peptides is catalyzed by the site-2 protease (S2P) in bacteria. *Proceedings of the National Academy of Sciences*, 108(33), pp.13740–13745.
Satoh, H. et al., 2002. The penta-EF-hand domain of ALG-2 interacts with amino-terminal domains of both annexin VII and annexin XI in a Ca2+-dependent manner. *Biochimica et biophysica acta*, 1600(1–2), pp.61–7.
Selkoe, D.J. & Wolfe, M.S., 2007. Presenilin: Running with Scissors in the Membrane. *Cell*, 131(2), pp.215–221.
Seth, R.B. et al., 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell*, 122(5), pp.669–82.
Shen, J. et al., 2002. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. *Developmental cell*, 3(1), pp.99–111.
Shibata, H. et al., 2015. A New Role for Annexin A11 in the Early Secretory Pathway via Stabilizing Sec31A Protein at the Endoplasmic Reticulum Exit Sites (ERES). *Journal of Biological Chemistry*, 290(8), pp.4981–4993.

Schneppenheim, J. et al., 2014. Signal-peptide-peptidase-like 2a is required for CD74 intramembrane proteolysis in human B cells. *Biochemical and Biophysical Research Communications*, 451(1), pp.48–53.

Schneppenheim, J. et al., 2013. The intramembrane protease SPPL2a promotes B cell development and controls endosomal traffic by cleavage of the invariant chain. *The Journal of experimental medicine*, 210(1), pp.41–58.

Siggs, O.M. et al., 2012. iRhom2 is required for the secretion of mouse TNFα. *Blood*, 119(24), pp.5769–71.

Skinner, J.R. et al., 2009. Diacylglycerol enrichment of endoplasmic reticulum or lipid droplets recruits perilipin 3/TIP47 during lipid storage and mobilization. *The Journal of biological chemistry*, 284(45), pp.30941–8.

Song, W. et al., 2013. Rhomboid domain containing 1 inhibits cell apoptosis by upregulating AP-1 activity and its downstream target Bcl-3. *FEBS Letters*, 587(12), pp.1793–1798.

Song, W. et al., 2015. Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway. *Nature communications*, 6(1), p.8022.

Spinazzi, M. & De Strooper, B., 2016. PARL: The mitochondrial rhomboid protease. *Seminars in cell & developmental biology*, 60, pp.19–28.

Stangeland, B. et al., 2015. Combined expressional analysis, bioinformatics and targeted proteomics identify new potential therapeutic targets in glioblastoma stem cells. *Oncotarget*, 6(28), pp.26192–215.

Strisovsky, K., 2016. Rhomboid protease inhibitors: Emerging tools and future therapeutics. *Seminars in Cell and Developmental Biology*.

Strisovsky, K., Sharpe, H.J. & Freeman, M., 2009. Sequence-Specific Intramembrane Proteolysis: Identification of a Recognition Motif in Rhomboid Substrates. *Molecular Cell*, 36(6), pp.1048–1059.

Sun, L., Li, X. & Shi, Y., 2016. Structural biology of intramembrane proteases: mechanistic insights from rhomboid and S2P to γ-secretase. *Current Opinion in Structural Biology*, 37, pp.97–107.

Sung, B.H. et al., 2015. Directional cell movement through tissues is controlled by exosome secretion. *Nature Communications*, 6(1), p.7164.

Tan, X. et al., 2015. Derlin-1 is overexpressed in human colon cancer and promotes cancer cell proliferation. *Molecular and cellular biochemistry*, 408(1–2), pp.205–13.

Targett-Adams, P. et al., 2008. Maturation of Hepatitis C Virus Core Protein by Signal Peptide Peptidase Is Required for Virus Production. *Journal of Biological Chemistry*, 283(24), pp.16850–16859.

Ticha, A. et al., 2017. Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases. *The Journal of biological chemistry*, 292(7), pp.2703–2713.

Tichá, A. et al., 2017. General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell Chem Biol.*, 24(12), pp.1523–36.

Tichá, A., Collis, B. & Strisovsky, K., 2018. The Rhomboid Superfamily: Structural Mechanisms and Chemical Biology Opportunities. *Trends in Biochemical Sciences*, 43(9), pp.726–739.

Tortelote, G.G. et al., 2017. Complexity of the Wnt/ß-catenin pathway: Searching for an activation model. *Cellular Signalling*, 40(May), pp.30–43.

Towbin, H., Staehelin, T. & Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76(9), pp.4350–4.

Urban, S. & Freeman, M., 2003. Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Molecular cell*, 11(6), pp.1425–34.

Urban, S., Lee, J.R. & Freeman, M., 2001. Drosophila Rhomboid-1 Defines a Family of Putative Intramembrane Serine Proteases. *Bang and Kintner*, 107, pp.173–182.

Uritsky, N., Shokhen, M. & Albeck, A., 2012. The Catalytic Machinery of Rhomboid Proteases: Combined MD and QM Simulations. *J Chem Theory Comput.*, 8(11), pp.4663–71.

Varnaitė, R. & MacNeill, S.A., 2016. Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. *Proteomics*, 16(19), pp.2503–2518.

Voss, M. et al., 2014. Shedding of glycan-modifying enzymes by signal peptide peptidase-like 3 (SPPL3) regulates cellular N-glycosylation. *The EMBO journal*, 33(24), pp.2890–905.

Voss, M., Schröder, B. & Fluhrer, R., 2013. Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1828(12), pp.2828–2839.

Wahlstrom, A.M. et al., 2007. Rce1 deficiency accelerates the development of K-RAS-induced myeloproliferative disease. *Blood*, 109(2), pp.763–768.

Wan, C. et al., 2012. Exosome-related multi-pass transmembrane protein TSAP6 is a target of rhomboid protease RHBDD1-induced proteolysis. *PloS one*, 7(5), p.e37452.

Wang, J. et al., 2008. Derlin-1 is overexpressed in human breast carcinoma and protects cancer cells from endoplasmic reticulum stress-induced apoptosis. *Breast cancer research : BCR*, 10(1), p.R7.

Wang, Y. et al., 2008. A novel member of the Rhomboid family, RHBDD1, regulates BIK-mediated apoptosis. *Cellular and Molecular Life Sciences*, 65(23), pp.3822–3829.

Wang, Y. et al., 2009. GC-1 mRHBDD1 knockdown spermatogonia cells lose their spermatogenic capacity in mouse seminiferous tubules. *BMC cell biology*, 10, p.25.

Wang, Y., Zhang, Y. & Ha, Y., 2006. Crystal structure of a rhomboid family intramembrane protease. *Nature*, 444(7116), pp.179–180.

Wasserman, J.D., Urban, S. & Freeman, M., 2000. A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling. *Genes & development*, 14(13), pp.1651–63.

Wei, X. et al., 2014. Lentiviral Vector Mediated Delivery of RHBDD1 shRNA down Regulated the Proliferation of Human Glioblastoma Cells. *Technology in Cancer Research & Treatment*, 13(1), pp.87–93.

Weihofen, A. et al., 2002. Identification of Signal Peptide Peptidase, a Presenilin-Type Aspartic Protease. *Science*, 296(5576), pp.2215–2218.

Wolfe, M.S., 2009. Intramembrane-cleaving proteases. *The Journal of biological chemistry*, 284(21), pp.13969–73.

Wolfe, M.S. & Kopan, R., 2004. Intramembrane proteolysis: Theme and variations. *Science*, 305(5687), pp.1119–1123.

Wu, Z. et al., 2006. Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. *Nature Structural and Molecular Biology*, 13(12), pp.1084–1091.

Wunderle, L. et al., 2016. Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGFalpha. *Scientific reports*, 6, p.27342.

Xu, L.-G. et al., 2005. VISA is an adapter protein required for virus-triggered IFN-beta signaling. *Molecular cell*, 19(6), pp.727–40.

Xue, Y. & Ha, Y., 2013. Large Lateral Movement of Transmembrane Helix S5 Is Not Required for Substrate Access to the Active Site of Rhomboid Intramembrane Protease. *Journal of Biological Chemistry*, 288(23), pp.16645–16654.

Ye, J. et al., 2000. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Molecular Cell*, 6(6), pp.1355–1364.

Ye, J., 2013. Roles of regulated intramembrane proteolysis in virus infection and antiviral immunity. *Biochimica et biophysica acta*, 1828(12), pp.2926–32.

Ye, Y. et al., 2004. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature*, 429, pp.841–847.

Zelenski, N.G. et al., 1999. Membrane topology of S2P, a protein required for intramembranous cleavage of sterol regulatory element-binding proteins. *The Journal of biological chemistry*, 274(31), pp.21973–80.

Zhang, M. et al., 2018. RHBDD1 promotes colorectal cancer metastasis through the Wnt signaling pathway and its downstream target ZEB1. *Journal of experimental & clinical cancer research : CR*, 37(1), p.22.

Zhang, X. et al., 2018. Rhomboid domain-containing protein 1 promotes breast cancer progression by regulating the p-Akt and CDK2 levels. *Cell communication and signaling : CCS*, 16(1), p.65.

Zhang, Y. & Wang, H., 2012. Integrin signalling and function in immune cells. *Immunology*, 135(4), pp.268–75.

Zhou, Y. et al., 2012. An internal water-retention site in the rhomboid intramembrane protease GlpG ensures catalytic efficiency. *Structure*, 20(7), pp.1255–63.

Zhu, W., Smith, J.W. & Huang, C.-M., 2010. Mass Spectrometry-Based Label-Free Quantitative Proteomics. *Journal of Biomedicine and Biotechnology*, 2010, pp.1–6.

Zoll, S. et al., 2014. Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO Journal*, 33(20), pp.2408–21.
Zuzow, N. et al., 2018. Mapping the mammalian ribosome quality control complex interactome using proximity labeling approaches S. Wolin, ed. *Molecular Biology of the Cell*, 29(10), pp.1258–1269.