The Putative Protein Interactome of the rhomboid protease RHBDL4

Jacqueline Melissa Hsiao[1], Ylauna Christine Mégane Penalva[2, 3, 4#], Helen Yee-Li Wu[2, 4#], Bin Xiao[2, 4], Gregor Jansen[1], Kurt Dejgaard[1], Jason C Young[1§], Lisa Marie Munter[2, 4*]

[1]Department of Biochemistry, McGill University, Montreal, QC, Canada H3G 0B1

[2]Department of Pharmacology & Therapeutics, McGill University, Montreal, QC, Canada H3G 0B1

[3]Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada H3A 2B4

[4]Cell Information Systems group, Bellini Life Sciences Complex, McGill University, Montreal, QC, Canada H3G 0B1

CRBS

# These authors contributed equally to the manuscript.

§ Posthumous publication

Corresponding author: *Lisa Munter, Department of Pharmacology & Therapeutics, Bellini Life Sciences Complex, Room 165, McGill University, 3649 Promenade Sir William Osler, Montreal, QC, Canada H3G 0B1. lisa.munter@mcgill.ca.

Abstract

The physiological functions of the rhomboid-related protein 4 (RHBDL4) are emerging, but their molecular details remain unclear. Because increased expression of RHBDL4 has been clinically linked to poorer outcomes in cancer patients, this association urgently demands a better understanding of RHBDL4. To elucidate the molecular interactions and pathways that RHBDL4 may be involved in, we conducted proximity-dependent biotin identification (BioID) assays. Our analyses corroborated several of the expected protein interactors such as the transitional endoplasmic reticulum ATPase VCP/p97 (TERA), but they also described novel putative interactors including IRS4, PGAM5 and GORS2. Using proximity-ligation assays, we validated VCP/p97, COPB and VRK2 as proteins that are in proximity to RHBDL4. Overall, our results support the emerging functions of RHBDL4 in endoplasmic reticulum quality control, but also point towards putative RHBDL4 functions in protein membrane insertion and membrane organisation and trafficking.

Introduction

Intramembrane rhomboid proteases are conserved throughout evolution, suggesting that they carry out fundamental biological functions [1]. The rhomboid-related protein 4 (RHBDL4, Uniprot accession number Q8TEB9) is one of five catalytically active mammalian rhomboid family members and is clinically linked to cancer: patients with a higher expression of RHBDL4 in the tumour-surrounding tissues showed a lower overall survival rate in colorectal and breast cancer [2, 3]. As a potential molecular mechanism, it was proposed that RHBDL4 facilitates the proteolytic release of TGFα[2]. However, this idea was challenged by Wunderle et al.; they instead suggested that RHBDL4 promotes the transport of TGFα to the cell surface [4]. In addition, cancer-relevant signalling pathways including AKT and β-catenin were found to be altered when RHBDL4 expression was increased [3, 5, 6]. RHBDL4 has been shown to play a role in the endoplasmic reticulum (ER)-associated degradation (ERAD), where it facilitates the retro-translocation process of misfolded proteins [7]. The globular C-terminus of RHBDL4 contains a ubiquitin-interacting motif and an interaction site for the transitional endoplasmic reticulum ATPase, VCP/p97, which is known to pull misfolded proteins out of the ER to subsequently direct them to the proteasome [7, 8]. Thus, it was proposed that RHBDL4 may cleave misfolded or aggregation-prone proteins that are stalled in the retro-translocation process [9]. A number of substrates have been identified for RHBDL4, including Bcl-2-interacting killer (BIK), oligosaccharyl transferase (OST), pre T-cell receptor substrate alpha (pTα), sterol regulatory element-binding protein 1 (SREBP1), and tumour suppressor activated pathway-6 (TSAP6) [6, 7, 10-12]. A recent paper identified dozens more RHBDL4 substrates, and amongst many of them, RHBDL4 appears to control the removal of C-terminal ER-retention motifs of ER chaperones [13]. Previously, we have shown that RHBDL4 cleaves the amyloid precursor protein (APP) that bypasses APP from its amyloidogenic processing, suggesting that RHBDL4 could play a modulatory role in Alzheimer's disease [14]. To gain further insight into its potential regulation in this neurodegenerative disease, we hypothesised that a comprehensive putative protein interactome map of RHBDL4 could serve as the basis to understand its physiological roles since it is likely that interacting proteins modulate its functions. Proximity-dependent biotin identification (BioID) is a powerful proteomics tool; it screens for all protein-protein interactions, albeit

weak, transient, or indirect, within a 10-nanometer interaction range [15, 16]. Here, we analysed the putative protein interactome of RHBDL4 using BioID assays and validated several of the hits using proximity ligation assays (PLA).

Results

To account for potentially different interaction partners at the cytosolic termini, we generated two RHBDL4 constructs with a mutated biotin ligase (BirA R118G) that were either fused at the N- or C-terminus of RHBDL4 (Figure 1A). Hereon, these constructs will be referred to as N-BirA-RHBDL4 or RHBDL4-BirA-C. These fusion constructs were stably integrated into Flp-In T-REx-293 cells and expressed under a tetracycline-inducible promoter.

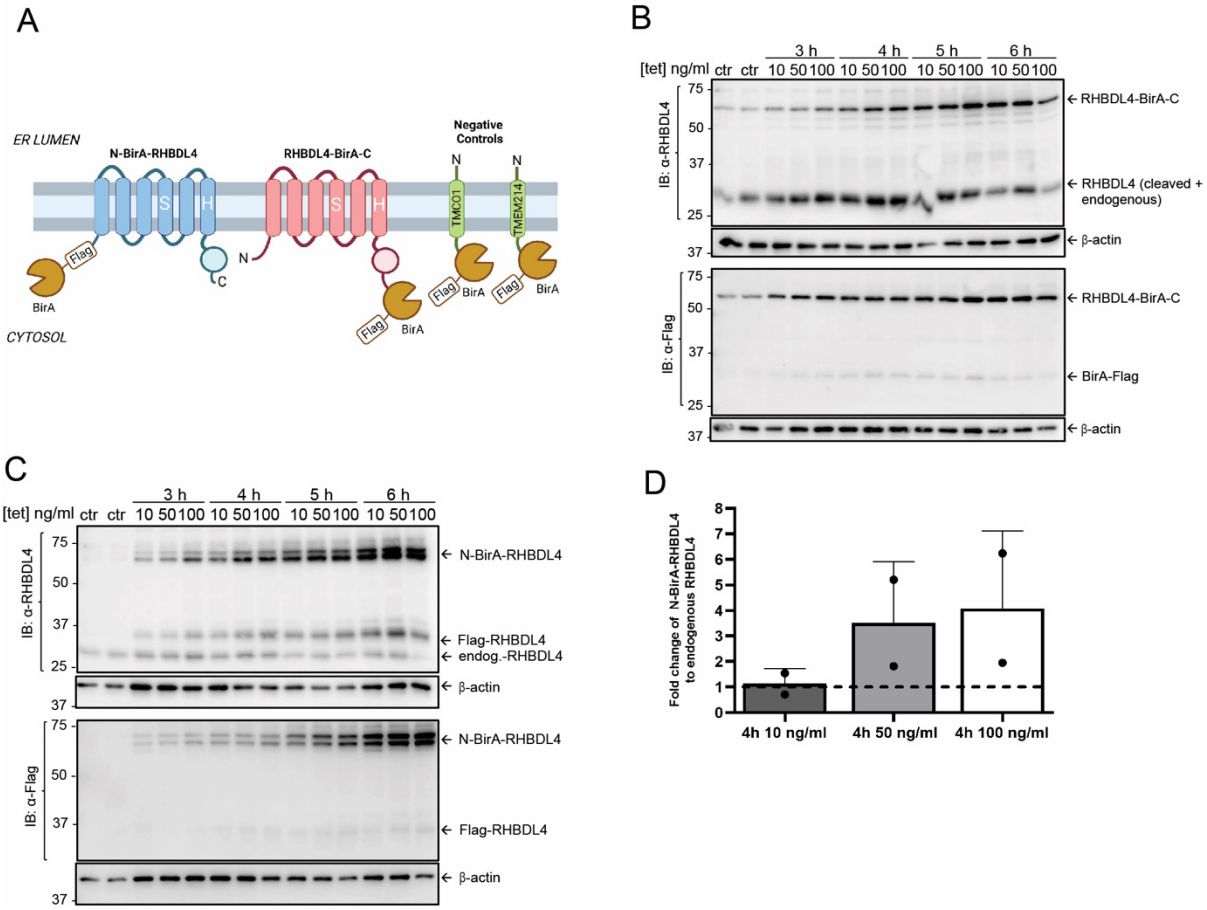


Figure 1: Optimization of RHBDL4 fusion protein expression. A: Schematic of the RHBDL4 and control fusion proteins, created with BioRender.com. **B, C:** Control (ctr) cells were treated with vehicle control (methanol). Upper band at around 60 kDa represents the intact fusion proteins, whereas signals at 37 kDa represent endogenous RHBDL4. Note that N-BirA-RHBDL4 generates an RHBDL4 signal at 42 kDa containing the Flag tag as a cleavage fragment from the fusion protein. Shown are representative Western blots of independent replicates that covered a larger range of concentrations and durations (n=3). **D:** Each independent replicate of N-BirA-RHBDL4 was quantified from the western blots and normalised to β-actin. Error bar denotes mean ±SEM (n=2).

To avoid detecting unspecific proximity partners of RHBDL4, we aimed to express RHBDL4 fusion protein levels as close to the endogenous RHBDL4 protein levels for BioID. Therefore, the expression levels of the RHBDL4 fusion proteins were first optimised to those of endogenous RHBDL4. We tested several concentrations and durations of tetracycline treatments to yield different fusion protein levels relative to endogenous RHBDL4, which is seen in the non-tetracycline-treated vehicle controls (Figure 1B-D). We observed that the fusion construct expressing N-BirA-RHBDL4 was cleaved and thereby released the exogenous RHBDL4 moiety, which migrated at a slightly higher molecular weight than endogenous RHBDL4 (Figure 1C). We thus took this exogenous higher RHBDL4 signal to optimise the expression of the fusion constructs, since the exogenous RHBDL4 cleavage fragment allowed for a direct comparison to endogenous RHBDL4 due to their similar size in the western blot. We determined that 10 ng/mL of tetracycline induction for four hours allowed the expression of the N-BirA-RHBDL4 fusion protein level to be approximately at the endogenous RHBDL4 protein level. Thus, this condition was used for all subsequent experiments.

We utilised the ER-resident transmembrane proteins, TMCO1 and TMEM214, as BioID negative controls with BirA fused to their C-termini, which are located in the cytosol and thus correspond to the localisation and orientation of the RHBDL4 fusion constructs. Assuming that unspecific protein interaction partners of RHBDL4 fusion proteins would also bind to the negative control fusion proteins when expressed at comparable levels, we matched the expression of the control fusion proteins, TMCO1-BirA and TMEM214-BirA, to the fusion protein levels of N-BirA-RHBDL4. Accordingly, the expressions of the TMCO1 and TMEM214 fusion constructs were induced with different concentrations and durations of tetracycline and subsequently compared to N-BirA-RHBDL4 (Figure 2). When both negative control fusion proteins were induced with 50 ng/mL tetracycline for a duration of four hours, their expressions were comparable to that of N-BirA-RHBDL4 fusion protein (Figure 2A, B). Thus, this condition for the negative control fusion proteins was used for all subsequent experiments.

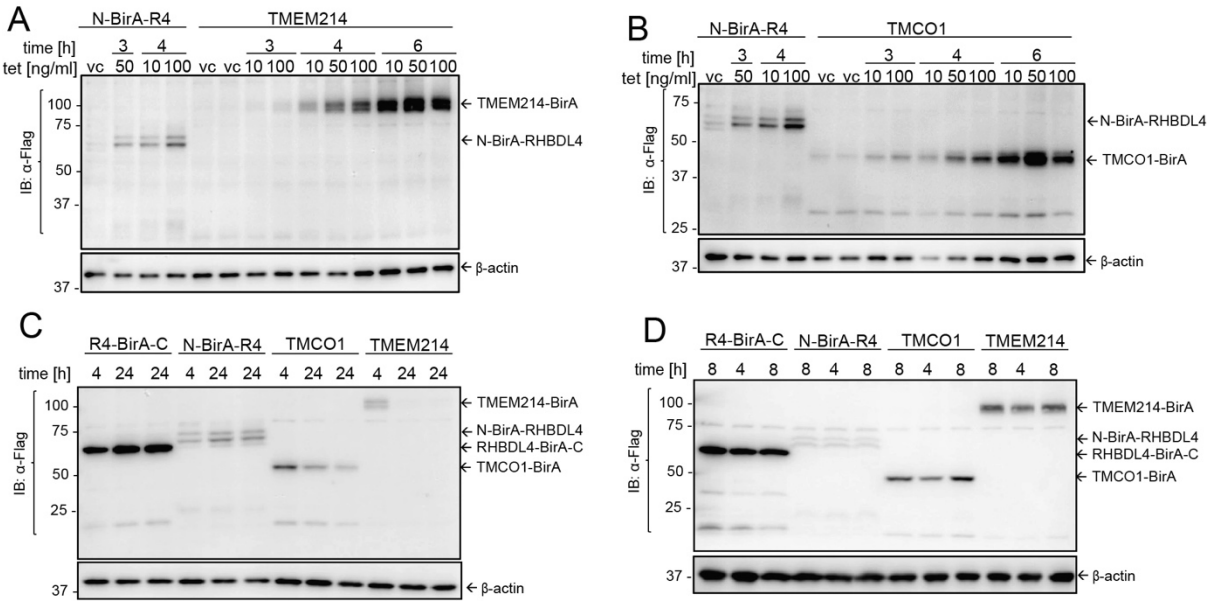


Figure 2: Optimization of control fusion proteins expression and stability assessment of fusion constructs. Lysates from Flp-In T-REx-293 cells stably expressing **A:** TMEM214-BirA or **B:** TMCO1-BirA was induced with different concentrations

and durations of tetracycline and compared to N-BirA-RHBDL4. Shown are representative Western blots of independent replicates (n=3). **C**: Fusion proteins were induced for 4 h with tetracycline. Cells were then either lysed or washed with PBS and incubated with medium for an additional 20 h, indicated as 24. Shown is a representative Western blot of independent replicates (n=3). **D**: As in A, but the additional incubation period was 4 h, indicated as 8 (n=1).

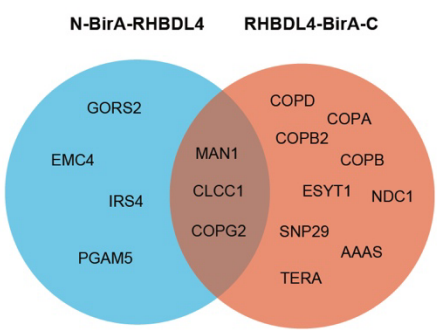
Optimization of fusion proteins for stability and biotin labelling.

The original protocol for BioID utilises a 24-hour incubation of biotin for the labelling of vicinal proteins [16, 17]. However, we observed that the TMEM214 and TMCO1 fusion proteins were unstable at 20 hours following the removal of the inducer tetracycline (Figure 2). Conversely, it was previously shown that biotinylation reactions reach saturation between six to 24 hours and yielded comparable results [17]. We therefore analysed the stability of each fusion protein at shorter time points and found that all proteins were optimally stable for an additional four hours after the removal of tetracycline (Figure 2C, D), allowing for a total biotin incubation of eight hours. Furthermore, in pilot experiments, we compared the resulting interactome of RHBDL4 and the negative controls from an eight-hour biotin incubation to a 24-hour biotin incubation and determined that there were no major differences (data not shown). Thus, we conducted all BioID assays by inducing the expression of fusion proteins and their biotinylation for four hours with tetracycline plus biotin, followed by another four-hour incubation with biotin alone.

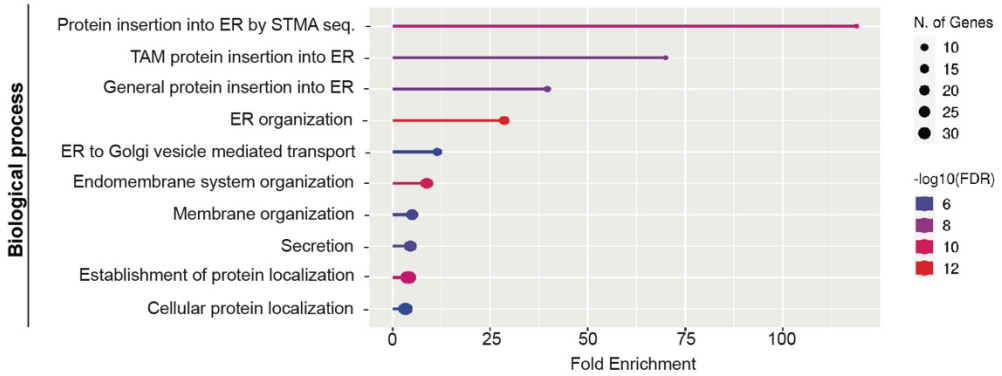
Stringent analysis of RHBDL4 BioID hits

Putative vicinal proteins of RHBDL4 were identified by mass spectrometry. Potential interactors that were shared with the negative controls as well as proteins characteristic of the BioID Contaminant Repository for Affinity Purification (CRAPome) were excluded [18]. Results were analysed in two ways: first, we performed a relatively stringent analysis where proteins were only identified as a unique potential interactor of RHBDL4 if the following criteria were met: one, the total spectral counts (TSCs) from the mass spectrometry raw data were at least five; and two, the TSCs were at least five-fold greater for the RHBDL4 fusion protein than in both controls in at least two of the three independent pull-downs. This stringent set of criteria allowed for the generation of a robust list of potential interactors that we are reporting with high confidence. A total of 16 proteins were identified as potential unique interactors of RHBDL4, of which only three proteins were found to be common interactors between the N- and C-termini of RHBDL4 (Figure 3A). Within these hits, we identified VCP/p97 (TERA) as a putative interactor of the RHBDL4 C-terminus, which we anticipated based on the VCP/p97 interacting motif located at the C-terminus of the protease [7]. Thus, this finding serves as an intrinsic positive control.

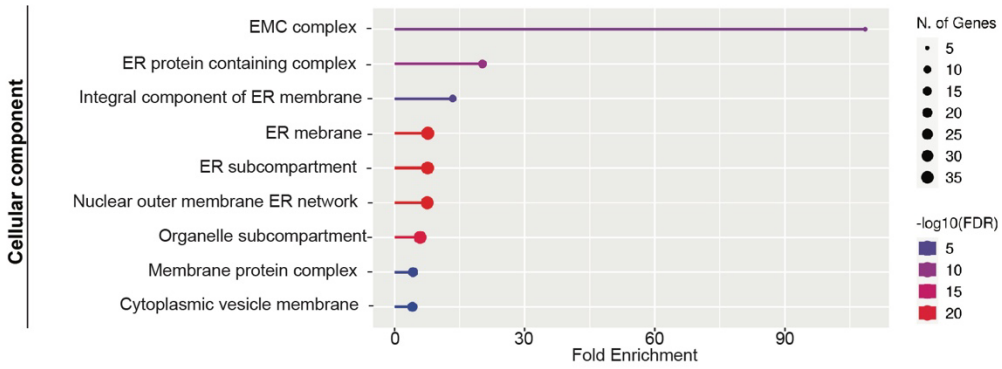
A


N-BirA-RHBDL4
RHBDL4-BirA-C
GORS2
EMC4
IRS4
PGAM5
MAN1
CLCC1
COPG2
COPD
COPA
COPB2
COPB
ESYT1
NDC1
SNP29
AAAS
TERA

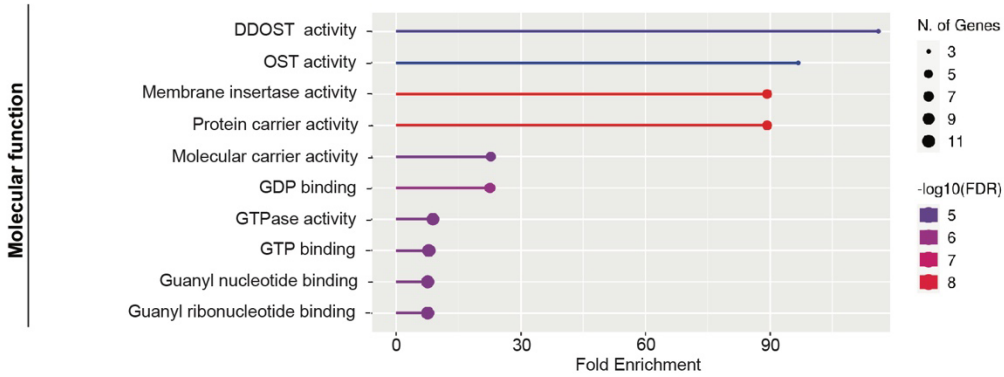
B


Biological process
Protein insertion into ER by STMA seq.
TAM protein insertion into ER
General protein insertion into ER
ER organization
ER to Golgi vesicle mediated transport
Endomembrane system organization
Membrane organization
Secretion
Establishment of protein localization
Cellular protein localization
0
25
50
75
100
Fold Enrichment
N. of Genes
10
15
20
25
30
-log10(FDR)
6
8
10
12

C


Cellular component
EMC complex
ER protein containing complex
Integral component of ER membrane
ER mebrane
ER subcompartment
Nuclear outer membrane ER network
Organelle subcompartment
Membrane protein complex
Cytoplasmic vesicle membrane
0
30
60
90
Fold Enrichment
N. of Genes
5
10
15
20
25
30
35
-log10(FDR)
5
10
15
20

D


Molecular function
DDOST activity
OST activity
Membrane insertase activity
Protein carrier activity
Molecular carrier activity
GDP binding
GTPase activity
GTP binding
Guanyl nucleotide binding
Guanyl ribonucleotide binding
0
30
60
90
Fold Enrichment
N. of Genes
3
5
7
9
11
-log10(FDR)
5
6
7
8

E

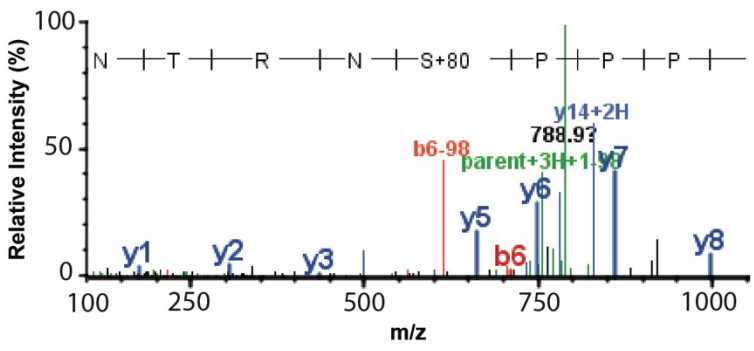

Relative Intensity (%)
100
50
0
N T R N S+80 P P P
y14+2H
788.97
b6-98
parent+3H+1
y7
y6
y5
y1
y2
y3
b6
y8
100
250
500
750
1000
m/z

Figure 3: Analysis of RHBDL4 interactors and putative phosphorylation site from mass spectrometry data. A: Venn diagram of stringent RHDBL4 interactome analysis. Using N-BirA-RHBDL4 and RHBDL4-BirA-C fusion proteins, a total of 16 proteins were identified and of which only 3 are in common. GORS2: golgi reassembly-stacking protein 2. EMC4: ER membrane protein complex subunit 4. PGAM5: serine/threonine-protein phosphatase. IRS4: insulin receptor substrate 4. MAN1: inner nuclear membrane protein Man1. CLCC1: chloride channel CLIC-like protein 1. COPG2: coatomer subunit gamma-2. COPD: coatomer subunit delta. COPA: coatomer subunit alpha. COPB2: coatomer subunit beta'. COPB: coatomer subunit beta. NDC1: nucleoporin NDC1. ESYT1: extended synaptotagmin-1. SNP29: synaptosomal-associated protein-29. AAAS: aladin. TERA: transitional endoplasmic reticulum ATPase (VCP). **B-D:** Analysis of broader RHBDL4 interactome using ShinyGO software. Fold enrichment of GO terms stratified by biological process (B), cellular component (C) and molecular function (D). Enrichment data was obtained using HEK293T transcriptome as the reference expression set. FDR cut-off for analysis is 0.05, and only the most significant groups are shown. The number of genes per group is shown by circle size while FDR significance is shown as a colour gradient from blue (least) to red (most). STMA: stop transfer membrane-anchor; TAM: tail anchored membrane; DDOST: dolichyl-diphosphooligosaccharide-protein glycoltransferase; OST: oligosaccharyl transferase. **E:** MS/MS spectrum with ion annotation of the phosphorylated RHBDL4 parent peptide (789.89 m/z + $3H^+$; deconvoluted 2,366.04 Da) showing a phosphorylation at serine residue 291 (S+80).

Pathway analysis of RHBDL4 BioID hits

We then conducted a second, less stringent approach in analysing the mass spectrometry data to identify relevant pathways that RHBDL4 may be involved in. In this analysis, we identified proteins as potential interactors if the TSCs were at least one and were observed in at least two out of the three independent pull-downs. This resulted in a large overlap between the BioID hits derived from the RHBDL4 N- and C-terminal fusion proteins; thus, we combined the data from both termini (see supplemental information 1). The resulting list of potential interactors was compared to a HEK293T gene expression set (available in the Human Protein Atlas [19] from https://www.proteinatlas.org/search/NOT+celline_category_rna%3AHEK+293%3BNot+detected) using ShinyGO 0-76.1 annotations [20]. Gene Ontology (GO) assigns genes to functional categories, which allows the identification of an enrichment in such groupings within a given gene or protein set. The RHBDL4 BioID putative interactome revealed GO Biological Processes on protein insertion into the ER and membrane organisation with at least a 10-fold enrichment from the HEK293T gene expression set (Figure 3B). A list comprising the proteins assigned to each GO category can be found in the supplemental information. GO Cellular Components identified a more than 100-fold enrichment for the EMC complex, which is an ER protein complex that facilitates the insertion of tail-anchored and polytopic membrane proteins (Figure 3C) [21, 22]. Amongst the GO Molecular Functions, OST-related activities were more than 90-fold enriched, which is likely related to the nature of OST acting as a substrate of RHBDL4 [10]. Functions relating to GTP binding and hydrolysis were found to be approximately 9-fold enriched as well (Figure 3D). One kinase that frequently appeared in these analyses was the serine/threonine protein kinase

(VRK2, also Vaccinia-related kinase 2), which plays a role in stress, membrane association, and signalling. As an incidental observation, we discovered several phosphorylated RHBDL4 peptides by mass spectrometry (without the use of phosphatase inhibitors, Figure 3E) with a prominent phosphorylation site at S291. To evaluate the potential kinases that could phosphorylate this site, we utilized GPS 5.0, which is a kinase prediction software[23], and VRK2 was proposed to be one of several kinases. Thus, we included VRK2 in our interaction validation assays.

BioID validation

To validate some of the BioID hits with an alternative method, we used PLA. This technique utilizes antibody stains with DNA-tag amplification and validates the close proximity between two proteins in a cell. As a negative control, we utilised HEK293T cells expressing an inducible shRNA against RHBDL4 [7] (Figure 4B). Our PLA analyses indeed support the interactions of RHBDL4 with VCP/p97 and COPB, which are hits from the stringent analysis (Figure 4C). Additionally, we validated the proximity of RHBDL4 and VRK2 (Figure 4C). The quantifications of PLA signals per nucleus reveal a statistical significant difference between PLA signals in wild type cells versus RHBDL4 knockdown cells (Figure 4C).

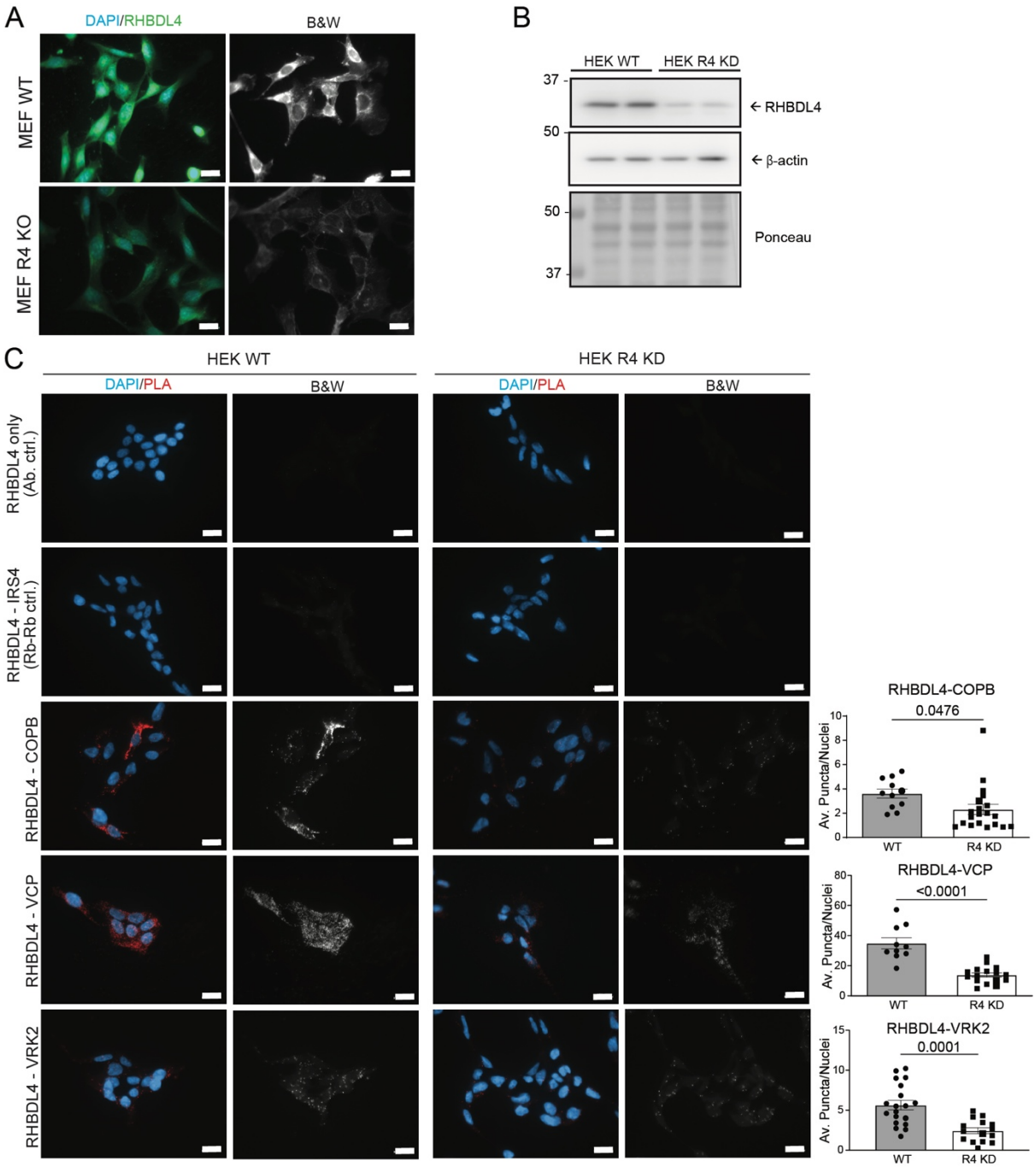


Figure 4: Validation of BioID hits using PLA. A: Confirmation of in-house RHBDL4 rabbit polyclonal antibody specificity in RHBDL4 WT and KO MEFs. The nucleus is stained with DAPI (blue) while RHBDL4 is represented by the green signal in the merged images. Black and white images show RHBDL4 signal intensity only. The scale bar is 20 µm. **B:** RHBDL4 protein expression in HEK293T RHBDL4 knockdown cells in comparison to WT after 4-day doxycycline treatment by Western blot. **C**: PLA experiments in wild type HEK293T and doxycycline-inducible RHBDL4 knockdown cells. The nucleus is in blue (DAPI) while the red dots indicate one proximity ligation of RHBDL4 with a putative interactor. Black and white images show the PLA signal intensity only. The reliability of the technique is tested by either using only the in-house rabbit RHBDL4 antibody or combining it with a rabbit antibody against an interacting protein (IRS4), since two rabbit antibodies

should not produce a PLA signal (top two rows). Shown are representative images of at least two biological replicates. The scale bar is 20 μm. Each data point of the quantification represents the number of PLA signals per nucleus per one image; significance was tested with the unpaired student t-test.

Discussion

The rhomboid protease RHBDL4 is an intriguing protein since its increased expression is clinically associated with several types of cancer [2, 12, 24]. In addition, we identified that RHBDL4 cleaves APP [14] within the cell, which alludes to a potential implication in Alzheimer's disease. In fact, cancer and Alzheimer's disease appear to affect similar pathways because it has been shown that cancer patients have an 11% decreased risk of getting Alzheimer's disease [25]. While its function in these potentially shared pathways is speculative, understanding the exact biological mechanisms of RHBDL4 should be a priority for both diseases. As such, we conducted BioID experiments to identify proteins putatively in proximity to RHBDL4. In our assays, we profoundly controlled the expression levels and stabilities of all our fusion proteins to minimize false positive interactions. We also conducted the proteomics assays with two negative control proteins and, not only one, but two distinct RHBDL4 fusion proteins with BirA attached to either the N- or C-terminus and thus identified novel interactors of RHBDL4. With these critical steps that allowed for the production of high-quality experimental data, a putative list of RHBDL4 protein interactors was thus identified. The resulting proteins were then analyzed via a stringent method as well as a pathway analysis carried out by ShinyGO.

Thus far, RHBDL4 has mostly been discussed in the context of ERAD. Interestingly, the results of our study as analyzed by ShinyGO suggests a novel role of RHBDL4 in general membrane and vesicle trafficking, as well as protein membrane insertion (see supplemental information). Moreover, we validated some of the identified interaction partners of RHBDL4 using PLA, which increases the confidence of these interactions. For example, while the VCP/p97 binding site of RHBDL4 was already identified in 2012 and supported by co-immunoprecipitations[7], we herein provide the first experimental data confirming their physical interaction and subcellular localization using PLA. VCP/p97 has been implicated in several functions, such as the clearance of stress granules, regulation in response to DNA double-stranded breaks, and mediating endolysosomal sorting of ubiquitylated cargo; however, it is predominantly known to be involved in the complex that participates in the retro-translocation of proteins from the ER to the cytoplasm for clearance by ERAD[26, 27]. We also validated the interaction of VRK2 and RHBDL4 by PLA, since VRK2 is of interest for two reasons: First, it is implicated in membrane biology as per ShinyGO annotations. Second, serine residue at S291 of RHBDL4 was found phosphorylated in our BioID analyses (Fig. 3E), which is potentially being carried out by VRK2 according to the kinase-specific prediction software GPS 5.0. S291 lies between the ubiquitin and VCP/p97 binding motifs of RHBDL4, so that S291 phosphorylation could potentially modulate the interaction of RHBDL4 with ubiquitin and/or VCP/p97. Lastly, we validated the interaction of RHBDL4 and COPB by PLA. Amongst the hits of the stringent analysis were several COPI coatomer subunits. These proteins mediate the retrograde vesicle transport

from the cis-Golgi to the ER and the trafficking of proteins between the Golgi cisternae. Thus, their functions may explain the prominence of protein and membrane trafficking pathways from the ShinyGO pathway analysis.

It should be noted that RHBDL4 BioID analyses have been previously published by Ikeda et. al [28]. Because our interactome is congruent with many of their hits,[28] this correspondence further supports the validity of the interactors. For example, our analysis determined PTP1B as an interactor of RHBDL4, and Ikeda et al.[28] suggested that this protein may mediate dephosphorylation of tyrosine residues in RHBDL4 and thereby influence the ability of RHBDL4 to bind to ubiquitin. Further overlap between our data sets exists for VCP/p97 (*TERA*), VRK2, MAN1, and COPG2 in relation to our stringent analysis. In addition, protein interactors were investigated with pull down assays by the Lemberg lab, who discovered, for example, protein disulfide isomerases (PDI) [9], which we could confirm in our BioID assay as well (PDIA6, S1). Conversely, there are also remarkable differences between our BioID analyses and those of Ikeda et al.. For example, IRS4 was a hit in our stringent analysis, but not identified by Ikeda et al[28]. IRS4, in contrast to the related proteins IRS1 and IRS2, is silent in adult tissues but is transcribed in some forms of breast cancer, suggesting a mechanism of interest for the involvement of RHBDL4 in cancer [29]. Further, we found PGAM5 as a putative RHBDL4 interactor. This protein partner is a mitochondrial serine/threonine phosphatase that regulates mitochondrial biogenesis [30, 31], suggesting the involvement of RHBDL4 at ER-mitochondria contact sites and in mitochondrial integrity. The Golgi reassembly-stacking protein 2 (GORS2/GRASP55) regulates the intracellular trafficking of certain transmembrane proteins, including the TGFα precursor [32], that disrupts trafficking in the absence of RHBDL4 as previously discussed[4]. IRS4, PGAM5, and GORS2 were identified in the pull-downs of our N-BirA-RHBDL4 fusion protein, which may explain some of the differences in our BioID analyses as compared to the interactors identified by Ikeda et al. [28], who solely used a C-terminal RHBDL4 fusion protein. We also identified several subunits of the EMC complex, which are putative novel RHBDL4 interactors. The EMC physically interacts with several ERAD components, such as derlin-2 [33], suggesting that RHBDL4 could be involved in ERAD through these protein interactions.

It has been suggested that the EMC is required to promote the biogenesis of sterol-sensing enzymes and was thus implicated in cellular cholesterol homeostasis [34]. Similarly, the EMC is involved in the maturation of proteins required for the biogenesis of lipid droplets [35]. Interestingly, we previously found that RHBDL4 binds to cholesterol, and this event affects its catalytic activity [36]. Together, this suggests that RHBDL4 and the EMC may coordinate in lipid dynamics. However, it should be noted that the EMC facilitates the protein membrane insertion of tail-anchored and multi-pass proteins, and because RHBDL4 is such a protein, the putative interaction with the EMC could be derived from the nascent RHBDL4 fusion proteins. Conversely, in line with the potential role of RHBDL4 in protein membrane insertion, an unbiased approach identified several dozens of new RHBDL4 substrates, including CCDC47, which is part of the ER membrane insertion complex for the biogenesis of multi-pass membrane complexes [13]. Overall, we did not identify many substrates of RHBDL4 using the BioID approach, except for OST subunits. This observation may be related to RHBDL4 possessing both catalytic and non-catalytic functions [37], and perhaps its catalytic activity is relatively low under standard cell conditions and may increase when cells are challenged.

Materials and Methods

Plasmids: C-terminal pcDNA5-TMEM214-BirA-FLAG, C-terminal pcDNA5-TMCO1-BirA-FLAG, pDONR223-RHBDL4, and N-terminal and C-terminal tagged pcDNA5-pDEST-BirA-FLAG Gateway vectors (kindly provided by Dr. David Y. Thomas, McGill University) and pOG44 Flp-Recombinase expression vector (Invitrogen) and pcDNA3.1 empty vector for mock control were used. pDONR223-RHBDL4 was used with the Gateway cloning system to insert the RHBDL4 sequence into N-terminal and C-terminal tagged pcDNA5-pDEST-BirA-FLAG plasmids to generate N-terminal tagged pcDNA5-BirA-FLAG-RHBDL4 (N-BirA-RHBDL4) and C-terminal tagged pcDNA5-RHBDL4-BirA-FLAG (RHBDL4-BirA-C) constructs. Expression vector inserts were verified by DNA sequencing (McGill University and Génome Québec Sequencing Center).

Antibodies: Mouse anti-β-actin (8H10D10, Cell Signaling), rabbit anti-RHBDD1 (HPA013972, Sigma-Aldrich) for western blots, and rabbit anti-FLAG (D6W5B, Cell Signaling) were used. Horseradish peroxidase-coupled secondary antibodies against mouse or rabbit IgG (Promega) were used for chemiluminescent detection using the ImageQuant LAS 500 or 600 system (GE Healthcare). Blots were quantified using ImageJ and Image Studio Lite (Licor).

Cell culture and lysis: Flp-In T-REx-293 wild type or Flp-In T-REx-293 cells with stably genome-integrated N-BirA-RHBDL4, RHBDL4-BirA-C, TMCO1-BirA-FLAG, or TMEM214-BirA-FLAG, were generated using the Flp-In system. 8 x10^5 Flp-In T-Rex-293 cells per 6-well were seeded 24 hours prior to transfection. The Flp recombinase, pOG44 (Invitrogen) and pcDNA5-Gene Of Interest (GOI)-BirA-FLAG were co-transfected with a total of 3 µg DNA with a ratio of 2:1 of pOG44 to pcDNA5-GOI-BirA-FLAG with 8 µL Lipofectamine 2000 (Invitrogen) in 1 mL Opti-MEM (Gibco) per well. Five hours post- transfection, the cell culture media was replaced with DMEM (Wisent: 4.5 g/L glucose, 0.584 g/L L-glutamine, 0.11 g/L sodium pyruvate) supplemented with 10% FBS (Wisent). 24 hours after transfection, the cells were re-seeded in 10 cm dishes. 48 hours after transfection, 200 µg/mL hygromycin B (Invitrogen) was used for selection. Stable Flp-In T-REx cells were cultured in DMEM supplemented with 10% FBS , 5 µg/mL blasticidin S, HCl (Bioshop), and 100 µg/mL hygromycin B at 37°C and 5% CO_2. 2 x 10^5 cells per 12-well or 5 x 10^5 cells per 6-well were seeded 48 hours prior to treatments and before lysis with TNE lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% Nonidet P-40, and cOmplete protease inhibitor (Roche)). 6X SDS sample buffer (300 mM Tris-HCl pH 6.8, 6% SDS, 24% glycerol, and bromophenol blue) was added to samples for a final concentration of 1X. Samples were analysed by standard SDS-PAGE and western blot procedures.

For PLA assays, HEK293T cells with doxycycline-inducible expression of a shRNA against RHBDL4 were kindly provided by Dr. Marius Lemberg, University of Cologne [7]. Cells were maintained in DMEM + 10% FCS at 37°C and 5% CO_2. The medium was supplemented with doxycycline or a vehicle control for at least 4 days before cells were seeded for PLA assays. Wild type or RHBDL4-knockout mouse embryonic fibroblasts (MEFs) were kindly provided by Dr. Matthew Freeman, University of Oxford. MEFs were maintained in DMEM supplemented with 10% FBS at 5% CO_2. Cells were fixed, permeabilized, stained with an in-house generated rabbit polyclonal serum against the N terminus of RHBDL4 (amino acids 4-14 RRSRGINTGL), and imaged using an Axiovert epifluorescent microscope.

BioID and LC-MS: BioID experiments were modified based on published procedures[17]. 2 x 10^6 Flp- In T-REx-293 cells expressing inducible N-BirA-RHBDL4, RHBDL4-BirA-C, or TMCO1-BirA-FLAG and TMEM214-BirA-FLAG were seeded onto two 10 cm dishes. 72 hours later, a four-hour induction of tetracycline at 10 ng/mL or 50 ng/mL plus 50 µM biotin was used to express and biotinylate the fusion proteins. Cells were washed twice with D-PBS and then incubated with 50 µM biotin (Sigma-Aldrich) alone for an additional four hours. The plates were washed with D-PBS, scraped in pre-chilled D-PBS, and centrifuged at 1000 x g for 10 minutes at 4°C. The pellet was lysed with RIPA lysis buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1% Triton X- 100, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 1 mM PMSF) for 10 minutes on ice. The lysate was centrifuged at 20,000 x g for 10 minutes. The supernatant was collected, and the protein concentration was diluted to 1-2 mg/mL using a wash buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl). The sample was pre-cleared with Sephadex G-25 Coarse beads (Amersham Pharmacia Biotech), rotating for 30 minutes at 4°C, prior to incubation with 15 µL of packed neutravidin agarose resin (Thermo Fisher) that was washed twice with RIPA buffer prior to their use, for two hours at 4°C. The resin was washed three times with wash buffer before being completely drained. 6 µL of Trypsin Gold, Mass Spectrometry Grade (Promega) was added overnight at 30°C. The digested peptides were washed from the resin with a 50:50 ratio of water to acetonitrile. The digest supernatant was collected through a C-18 ZipTip, dried in a SpeedVac and resolubilized in 20 µL water with 0.1% formic acid (Fluka Analytical). Using a nanoflow liquid chromatography system (nLC1000, Thermo Scientific) 10 µL of the digests were injected onto a 2 cm 2µm C18 trapping column (PepMap, Thermo Scientific) and separated by reverse phase chromatography through a 25 cm reverse phase column (2 µm C18, Thermo Scientific) running an acetonitrile (ACN) gradient of 0-40% ACN in 100 minutes at a flow rate of 350 nL per minute into an Orbitrap mass spectrometer (Q-Exactive HF, Thermo Scientific). The mass spectrometer was set to the following data acquisition duty cycle parameters: MS scan at 240,000 resolution, followed by 25 MSMS isolations and fragmentations at 8000 resolution, with a dynamic time exclusion of 4 seconds.

BioID bioinformatics and databases for protein identification

All MS/MS raw data files were converted into peak lists using Mascot Peak Distiller (Matrix Science). Mgf files were searched against the curated Human Proteome database (Uniprot: Homo Sapiens) using the Mascot search engine (Matrix Science) with the following parameters: Enzyme: Trypsin. Missed cleavages: 1. #C13:1. Fixed modifications: none. Variable modifications: oxM (and, when expected, phospho-T and phospho-S). Precursor mass accuracy: 6 ppm. MSMS mass accuracy: 50 mmu. Instrument type ICRFT. Mascot files were imported into Scaffold v3.4 (Proteome Software Inc), and a second search engine (X!Tandem) was requested to search the data again, after which the combined data was validated through the trans proteomics pipeline. Data were visualized as total spectral counts, passing peptide and protein thresholds of 95%.

In the first analysis, the identified proteins and peptides from Scaffold were simultaneously filtered to identify unique BioID hits for the N-terminus and C-terminus of RHBDL4, independently, which were only accepted if the TSC from the MS/MS raw data were at least five for RHBDL4-containing samples and simultaneously at least five-fold higher than both the TSCs from TMCO1- and TMEM214-containing samples in at least two independent pull-downs. In the second analysis, the

identified proteins and peptides from Scaffold were simultaneously filtered to identify general BioID hits for RHBDL4 C-terminus or N-terminus, which were only accepted if the TSC from the MS/MS raw data was at least one in at least two independent pull-down replicates for each terminus.

Proximity ligation assays: $8x10^4$ cells/24-well were seeded on coverslips. Cells were fixed with 4% paraformaldehyde, quenched with ammonium chloride, and permeabilized using 0.1% triton-X100 in PBS for 10 minutes. Following the manufacturer's instructions (Sigma-Aldrich), cells were blocked for 1 hour in a humidity chamber and then incubated with the primary antibodies overnight at 4°C. Detection antibodies were incubated for 1 hour at 37°C in a humidity chamber, followed by the ligation reaction for 30 minutes, and then the amplification reaction for 100 minutes. Slips were mounted on slides and imaged with an Axiovert 3 fluorescent microscope (ABIF core facility). The antibodies used in PLA are in-house generated rabbit polyclonal serum against the N terminus of RHBDL4 amino acids 4-14 (RRSRGINTGL), mouse monoclonal anti-VCP antibody (MA3-004 Invitrogen), rabbit polyclonal anti-IRS4 antibody (PA5-53596 Invitrogen), mouse monoclonal anti-COPB antibody (M3A5 Millipore Sigma), and mouse monoclonal anti-VRK2 antibody (WH0007444M1 Millipore Sigma). For quantifications, fluorescent intensities were adjusted between WT and RHBDL4 knockdown cells and PLA signals per nucleus counted using Image J. Each data point represents the quantification from one image. Statistical significance was analyzed using the GraphPadPrism software using the unpaired student t-test.

Conclusions

In conclusion, we have identified a putative proteome map of RHBDL4 that not only confirms previously reported protein interaction partners, but more importantly, our analyses revealed several novel interactors, some of which we further validated. Our results indicate that the cellular functions of RHBDL4 include ER quality control, protein membrane insertion, and membrane organisation.

Acknowledgements: Dr. Jason Young was the primary co-supervisor of Jacqueline Melissa Hsiao along with Dr. Lisa Munter as her secondary co-supervisor. Dr. Young's guidance, initiative, and contributions to this project have been highly valuable, and we are grateful to have had the pleasure to work with him. He had previously agreed to the publication of the presented data but passed away in February 2022 prior to the completion of the PLA validation data.

Images were collected and analysed in the McGill University Advanced BioImaging Facility (ABIF), RRID:SCR_017697. Mass spectrometry data were obtained in the Biochemistry Core Facility through Kurt Dejgaard. We thank Dr. Marius Lemberg for providing the HEK293T RHBDL4-knockdown cell lines. We thank Sherilyn Recinto for supportive experiments, which were not included in this manuscript.

Sample data availability statement: The data underlying this study are available in the published article, its online supplementary material, and are openly available in the MassIVE data repository at https://massive.ucsd.edu/ as dataset MSV000091174.

Conflict of Interest: The authors declare no conflict of interests.

Supporting Information:

Supporting information 1: XLS data set of RHBDL4 BioID hits and assigned genes from the GO annotations.

Funding: This work was funded by the Alzheimer Society of Canada, Research Grant 20-02, the Scottish Rite Charity Foundation 16112, the Canadian Institute of Health Research CIHR-PJT-175306, Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants RGPIN-2015-04645 and supported by the Fonds de recherche du Québec - santé FRQS research allocation FRQ-S 36571 and the Canada Foundation of Innovation Leaders Opportunity Fund Grant 32565. JMH received a studentship through the Canada First Research Excellence Fund, awarded to McGill University for the Healthy Brains for Healthy Lives initiative. YP received the Davis fellowship through McGill's Faculty of Medicine.

Accession Code:

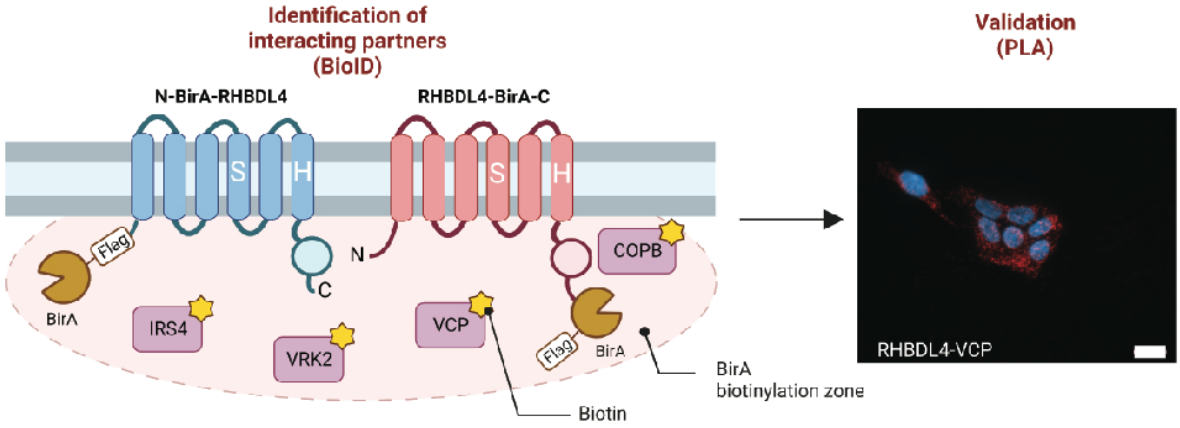
RHBDL4: Uniprot accession number Q8TEB9

References:

[1] Freeman, M. (2014) The rhomboid-like superfamily: molecular mechanisms and biological roles, *Annu Rev Cell Dev Biol 30*, 235-254.
[2] Song, W., Liu, W., Zhao, H., Li, S., Guan, X., Ying, J., Zhang, Y., Miao, F., Zhang, M., Ren, X., Li, X., Wu, F., Zhao, Y., Tian, Y., Wu, W., Fu, J., Liang, J., Wu, W., Liu, C., Yu, J., Zong, S., Miao, S., Zhang, X., and Wang, L. (2015) Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway, *Nat Commun 6*, 8022.
[3] Zhang, X., Zhao, Y., Wang, C., Ju, H., Liu, W., Zhang, X., Miao, S., Wang, L., Sun, Q., and Song, W. (2018) Rhomboid domain-containing protein 1 promotes breast cancer progression by regulating the p-Akt and CDK2 levels, *Cell Commun Signal 16*, 65.
[4] Wunderle, L., Knopf, J. D., Kuhnle, N., Morle, A., Hehn, B., Adrain, C., Strisovsky, K., Freeman, M., and Lemberg, M. K. (2016) Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGFalpha, *Sci Rep 6*, 27342.
[5] Huang, C., Ji, X., Peng, Y., Wu, M., Wu, W., Luo, Y., Cheng, G., and Zhu, Y. (2018) Silencing of rhomboid domain containing 1 to inhibit the metastasis of human breast cancer cells in vitro, *Iran J Basic Med Sci 21*, 1161-1166.
[6] Xu, Z., Wang, R., Li, X., Yang, L., Peng, H., Wang, Y., and Wang, P. (2021) RHBDD1 silencing inhibited cell growth and invasion of non-small cell lung cancer by mediating ZEB1/PI3K/AKT signaling pathway, *J Mol Histol 52*, 503-510.
[7] Fleig, L., Bergbold, N., Sahasrabudhe, P., Geiger, B., Kaltak, L., and Lemberg, M. (2012) Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins., *Mol Cell 47*, 558-569.
[8] Lim, J. J., Lee, Y., Ly, T. T., Kang, J. Y., Lee, J. G., An, J. Y., Youn, H. S., Park, K. R., Kim, T. G., Yang, J. K., Jun, Y., and Eom, S. H. (2016) Structural insights into the interaction of p97 N-terminus domain and VBM in rhomboid protease, RHBDL4, *Biochem J 473*, 2863-2880.
[9] Bock, J., Kuhnle, N., Knopf, J. D., Landscheidt, N., Lee, J. G., Ye, Y., and Lemberg, M. K. (2022) Rhomboid protease RHBDL4 promotes retrotranslocation of aggregation-prone proteins for degradation, *Cell Rep 40*, 111175.

[10] Knopf, J. D., Landscheidt, N., Pegg, C. L., Schulz, B. L., Kuhnle, N., Chao, C. W., Huck, S., and Lemberg, M. K. (2020) Intramembrane protease RHBDL4 cleaves oligosaccharyltransferase subunits to target them for ER-associated degradation, *J Cell Sci 133*, jcs243790.
[11] Shibuya, K., Ebihara, K., Ebihara, C., Sawayama, N., Isoda, M., Yamamuro, D., Takahashi, M., Nagashima, S., and Ishibashi, S. (2022) AAA-ATPase valosin-containing protein binds the transcription factor SREBP1 and promotes its proteolytic activation by rhomboid protease RHBDL4, *J Biol Chem 298*, 101936.
[12] Wan, C., Fu, J., Wang, Y., Miao, S., Song, W., Wang, L., and opaptosis. (2012) Exosome-related multi-pass transmembrane protein TSAP6 is a target of rhomboid protease RHBDD1-induced proteolysis., *PLoS One 7*, e37452.
[13] Tang, S., Beattie, A. T., Kafkova, L., Petris, G., Huguenin-Dezot, N., Fiedler, M., Freeman, M., and Chin, J. W. (2022) Mechanism-based traps enable protease and hydrolase substrate discovery, *Nature 602*, 701-707.
[14] Paschkowsky, S., Hamze, M., Oestereich, F., and Munter, L. M. (2016) Alternative Processing of the Amyloid Precursor Protein Family by Rhomboid Protease RHBDL4, *J Biol Chem 291*, 21903-21912.
[15] Roux, K. J., Kim, D. I., and Burke, B. (2013) BioID: a screen for protein-protein interactions, *Curr Protoc Protein Sci 74*, 19 23 11-19 23 14.
[16] Go, C. D., Knight, J. D. R., Rajasekharan, A., Rathod, B., Hesketh, G. G., Abe, K. T., Youn, J. Y., Samavarchi-Tehrani, P., Zhang, H., Zhu, L. Y., Popiel, E., Lambert, J. P., Coyaud, E., Cheung, S. W. T., Rajendran, D., Wong, C. J., Antonicka, H., Pelletier, L., Palazzo, A. F., Shoubridge, E. A., Raught, B., and Gingras, A. C. (2021) A proximity-dependent biotinylation map of a human cell, *Nature 595*, 120-124.
[17] Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells, *J Cell Biol 196*, 801-810.
[18] Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z. Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A. C., and Nesvizhskii, A. I. (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data, *Nat Methods 10*, 730-736.
[19] Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., Sanli, K., von Feilitzen, K., Oksvold, P., Lundberg, E., Hober, S., Nilsson, P., Mattsson, J., Schwenk, J. M., Brunnstrom, H., Glimelius, B., Sjoblom, T., Edqvist, P. H., Djureinovic, D., Micke, P., Lindskog, C., Mardinoglu, A., and Ponten, F. (2017) A pathology atlas of the human cancer transcriptome, *Science 357*, eaan2507.
[20] Ge, S. X., Jung, D., and Yao, R. (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants, *Bioinformatics 36*, 2628-2629.
[21] Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S., and Hegde, R. S. (2018) EMC Is Required to Initiate Accurate Membrane Protein Topogenesis, *Cell 175*, 1507-1519 e1516.
[22] Guna, A., Volkmar, N., Christianson, J. C., and Hegde, R. S. (2018) The ER membrane protein complex is a transmembrane domain insertase, *Science 359*, 470-473.
[23] Wang, C., Xu, H., Lin, S., Deng, W., Zhou, J., Zhang, Y., Shi, Y., Peng, D., and Xue, Y. (2020) GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins, *Genomics Proteomics Bioinformatics 18*, 72-80.
[24] Zhao, C., Ling, X., Li, X., Hou, X., and Zhao, D. (2019) MicroRNA-138-5p inhibits cell migration, invasion and EMT in breast cancer by directly targeting RHBDD1, *Breast Cancer 26*, 817-825.
[25] Ospina-Romero, M., Glymour, M. M., Hayes-Larson, E., Mayeda, E. R., Graff, R. E., Brenowitz, W. D., Ackley, S. F., Witte, J. S., and Kobayashi, L. C. (2020) Association Between Alzheimer Disease and Cancer With Evaluation of Study Biases: A Systematic Review and Meta-analysis, *JAMA Netw Open 3*, e2025515.
[26] Ritz, D., Vuk, M., Kirchner, P., Bug, M., Schutz, S., Hayer, A., Bremer, S., Lusk, C., Baloh, R. H., Lee, H., Glatter, T., Gstaiger, M., Aebersold, R., Weihl, C. C., and Meyer, H. (2011) Endolysosomal sorting of ubiquitylated caveolin-1 is regulated by VCP and UBXD1 and impaired by VCP disease mutations, *Nat Cell Biol 13*, 1116-1123.
[27] van den Boom, J., and Meyer, H. (2018) VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling, *Mol Cell 69*, 182-194.
[28] Ikeda, K. N., and Freeman, M. (2019) Spatial proteomics reveal that the protein phosphatase PTP1B interacts with and may modify tyrosine phosphorylation of the rhomboid protease RHBDL4, *J Biol Chem 294*, 11486-11497.
[29] Ikink, G. J., and Hilkens, J. (2017) Insulin receptor substrate 4 (IRS4) is a constitutive active oncogenic driver collaborating with HER2 and causing therapeutic resistance, *Mol Cell Oncol 4*, e1279722.
[30] Bernkopf, D. B., Jalal, K., Bruckner, M., Knaup, K. X., Gentzel, M., Schambony, A., and Behrens, J. (2018) Pgam5 released from damaged mitochondria induces mitochondrial biogenesis via Wnt signaling, *J Cell Biol 217*, 1383-1394.
[31] Sekine, S., Kanamaru, Y., Koike, M., Nishihara, A., Okada, M., Kinoshita, H., Kamiyama, M., Maruyama, J., Uchiyama, Y., Ishihara, N., Takeda, K., and Ichijo, H. (2012) Rhomboid protease PARL mediates the mitochondrial membrane potential loss-induced cleavage of PGAM5, *J Biol Chem 287*, 34635-34645.
[32] Kuo, A., Zhong, C., Lane, W. S., and Derynck, R. (2000) Transmembrane transforming growth factor-alpha tethers to the PDZ domain-containing, Golgi membrane-associated protein p59/GRASP55, *EMBO J 19*, 6427-6439.
[33] Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W., and Kopito, R. R. (2011) Defining human ERAD networks through an integrative mapping strategy, *Nat Cell Biol 14*, 93-105.

[34] Volkmar, N., Thezenas, M. L., Louie, S. M., Juszkiewicz, S., Nomura, D. K., Hegde, R. S., Kessler, B. M., and Christianson, J. C. (2019) The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis, *J Cell Sci 132*, jcs223453.
[35] Leznicki, P., Schneider, H. O., Harvey, J. V., Shi, W. Q., and High, S. (2022) Co-translational biogenesis of lipid droplet integral membrane proteins, *J Cell Sci 135*, jcs259220.
[36] Paschkowsky, S., Recinto, S. J., Young, J. C., Bondar, A. N., and Munter, L. M. (2018) Membrane cholesterol as regulator of human rhomboid protease RHBDL4, *J Biol Chem 293*, 15556-15568.
[37] Lastun, V. L., Levet, C., and Freeman, M. (2021) RHBDL4 protects against endoplasmic reticulum stress by regulating the morphology and distribution of ER sheets, *bioRxiv*, 2021.2006.2015.448480.



For Table of Contents Only