

Title	An integrated study of the evolution, molecular function, expression and ubiquitome of the ligand of NUMB Protein X family
Authors	Erwin, Leah G.
Publication date	2026-03-31
Original Citation	Erwin, L. G. 2026. An integrated study of the evolution, molecular function, expression and ubiquitome of the ligand of NUMB Protein X family. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2026, Leah Gabrielle Erwin. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-10-26 00:59:00
Item downloaded from	https://hdl.handle.net/10468/19309

Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork



**An Integrated Study of the Evolution, Molecular Function,
Expression and Ubiquitome of the Ligand of NUMB Protein X
Family.**

Thesis submitted to the National University of Ireland, Cork, in fulfilment of the
requirements for the degree of

Doctor of Philosophy

by

Leah Gabrielle Erwin, BSc, MSc.

School of Biochemistry & Cell Biology

University College Cork, Ireland

Supervisor: Dr. Paul Young

Head of School: Prof. Justin McCarthy

March 2026

Table of Contents

Declaration	7
Acknowledgements.....	8
Contributions	9
Publications and Other Scientific Contributions	10
Abstract	11
Abbreviations.....	14
Chapter 1	16
General Introduction.....	16
1.1 Ubiquitination	17
1.1.1 The discovery and molecular structure of ubiquitin	17
1.1.2 The mechanism of ubiquitination.....	19
1.1.3 Ubiquitin activation	20
1.1.4 Ubiquitin conjugation	20
1.1.5 Ubiquitin ligation	20
1.1.5.1 HECT domain E3 ligases	22
1.1.5.2 RING finger domain E3 ligases.....	23
1.1.5.3 RING-Between-RING E3 ligases.....	24
1.1.6 Types of ubiquitination.....	25
1.1.6.1 The power of one - monoubiquitination.....	25
1.1.6.2 The infinite complexity of polyubiquitination	26
1.1.7 Deubiquitinating enzymes	29
1.1.8 Ubiquitination in the context of cancer.....	30
1.1.8.1 Druggability of the ubiquitin system in cancer	31
1.2 PDZ domains	33
1.2.1 Classification of PDZ domains.....	33
1.2.1.1 Class I	35
1.2.1.2 Class II	35
1.2.1.3 Class III	35
1.3 Ligand of NUMB Protein X family	37
1.3.1 NUMB proteins	37
1.3.2 The discovery of the Ligand of Numb Protein X family	39
1.3.3 Structure of LNX1/2 proteins.....	40
1.3.4 Evolution of LNX proteins	43

1.3.5 Expression of LNX1/2 mRNA and protein in mammals	46
1.3.5.1 LNX1 mRNA expression.....	46
1.3.5.2 <i>LNX2</i> mRNA expression.....	46
1.3.5.3 Protein expression of LNX1 and LNX2.....	47
1.3.6 Functions of LNX1/2	48
1.3.6.1 Role of LNX proteins in neurogenesis and behaviour	48
1.3.6.2 Role of LNX2 in osteoclastogenesis.....	50
1.3.7 Disease associations of LNX1 and LNX2.....	51
1.3.7.1 LNX1/2 protein involvement in infectious disease.....	51
1.3.7.2 LNX1/2 in colorectal cancer	51
1.3.7.3 LNX1 involvement in brain tumours.....	53
1.4 Overall summary or aims? To tie the introduction together	54
Chapter 2	55
Materials and Methods.....	55
2.1 Molecular biology	56
2.1.1 Plasmid construction	56
2.1.2 Polymerase chain reaction (PCR).....	57
2.1.3 Reverse transcriptase PCR (RT-PCR)	59
2.1.4 Annealing oligonucleotides	60
2.2 Mammalian cell culture and DNA transfections	63
2.2.1 Mammalian cell culturing conditions	63
2.2.2 HEK 293T and MCF-7 cell lines	64
2.2.3 Colorectal carcinoma cell lines	64
2.2.4 Glioblastoma and Neuroblastoma cell lines	65
2.2.5 Cell pellets and lysates	66
2.2.6 Maintenance of established cell lines	66
2.2.7 Cryopreservation and recovery of cell lines	66
2.2.8 Seeding and transfection of cell lines	67
2.3 Immunocytochemistry and cell lysate preparation	67
2.3.1 Immunocytochemistry	67
2.3.2 Cell lysis	68
2.3.3 Tissue homogenisation and lysis	68
2.3.4 Determination of protein concentration	69
2.4 Luciferase reporter assay.....	69
2.5 Protein expression, purification and immunoprecipitation.....	70

2.5.1 Expression of LNX fusion proteins	70
2.5.2 Nickel-Nitrilotriacetic Acid (Ni-NTA) Purification.....	71
2.5.3 Amylose pull down of MBP fusion proteins	71
2.5.4 Dialysis of purified proteins	72
2.5.5 Pull down of LNX interacting proteins using GFP-Trap-M beads.....	72
2.5.6 Ubiquitination assays	73
2.6 SDS-PAGE and immunoblotting	73
2.6.1 SDS-PAGE	73
2.6.2 Coomassie blue staining	74
2.6.3 Western immunoblotting	74
2.7 Yeast-based protein interactions assays.....	76
2.7.1 Transformation of yeast and yeast-2-hybrid analysis	76
2.7.2 Yeast-two hybrid analysis	77
2.7.3 Protein interaction analysis using yeast GFP-nanobody display and flow cytometry evaluation	77
2.8 Identification of LNX2 ubiquitination substrates <i>via</i> proximity labelling.....	78
2.8.1 Cloning, transient transfection and biotin labelling of LNX-BirA substrates	78
2.8.2 Neutravidin pull down of biotinylated proteins	79
2.8.3 Liquid chromatography - mass spectrometry (LC-MS/MS)	80
2.8.4 Data processing and protein identification	81
2.9 Computational analysis.....	81
2.9.1 Building a phylogenetic tree of LNX orthologous proteins.....	81
Chapter 3	83
A Multi-Modal Approach to Identify and Validate Novel LNX1 and LNX2 Protein Interactions.....	83
3.1 Introduction	84
3.2 Results.....	87
3.2.1 Identifying possible interactors of LNX1/2 from ProfAff and	87
<i>in silico</i> predictions	87
3.2.2 Yeast-two hybrid analysis of selected candidates	91
3.2.3 Validation of full-length protein interactions by GFP and MBP pull down (PD) assays.....	94
3.2.4 Exploration of colocalization between LNX and novel interacting proteins.....	97
in HeLa cells	97
3.2.5 Characterisation of the ZNF24 – LNX1 interaction	103
3.2.5.1 Validating anti-GFP-Nanobody displaying yeast as a method of.....	103
co-immunoprecipitation.	103
3.2.5.2 The ZNF24 interaction is mapped to PDZ2 of LNX1	105

3.2.5.3 ZNF24 is a ubiquitin substrate of LNX1	108
3.3 Discussion	111
3.3.1 The importance of verifying protein domain interactions	111
3.3.2 Non-traditional methods of identifying LNX interactors	113
Chapter 4	116
Novel LNX2 Substrates Revealed Through Proximity-Dependent Biotin-Labeling of Substrates and Proteomic Analysis.....	116
4.1 Introduction	117
4.2 Results.....	120
4.2.1 Modifying the BioE3/Ub-POD method	120
4.2.2 Proteomic identification of LNX2 substrates in HEK 293T cells.....	124
4.2.3 Functional Classification and grouping of novel LNX2 substrates by Gene Ontology (GO) enrichment	130
4.2.4 Examination of proteins identified by proteomics as LNX2 interactors	133
4.2.5 Validation of LNX2 substrates by direct in vitro ubiquitination assays	136
4.3 Discussion	141
4.3.1. The importance of properly validating novel methods	141
4.3.2 Interpreting proteomic results	142
Chapter 5	144
Molecular Evolution of LNX Proteins: Structural Conservation and Interaction Networks	144
5.1 Introduction	145
5.2 Results.....	148
5.2.1 Many metazoan species contain a well conserved LNX ortholog	148
5.2.2 LNX protein orthologs have a well conserved domain architecture.	154
5.2.3 Phylogenetic analysis of LNX orthologs reveals sequence clustering to three major clades ..	156
5.2.4 The second PDZ domain of LNX is predicted to be spatially more distant from other PDZ domains across metazoan orthologs.....	158
5.2.5 Conservation of intramolecular interaction of carboxyl terminal PBM and first PDZ domain of LNX2.....	163
5.2.6 Homodimerization of RING domains is a conserved feature	168
of LNX orthologs	168
5.2.7 “What’s in a name?” – analysing the ancestral LNX-NUMB interaction	171
5.3 Discussion	173
Chapter 6	177
Characterisation of LNX Expression and Promoter Responsiveness in Colorectal and Glial Cancers ...	177
6.1 Introduction	178
6.2 Results.....	181

6.2.1 Analysis of chromosomal amplification and gene expression at the LNX1 and LNX2 loci in GBM and CRC	181
6.2.2 In silico and in vitro analysis of LNX1/2 expression	187
6.2.3 LNX1, LNX2 and LNX3 promoter analysis	194
6.3 Discussion	204
6.3.1 Analysis of LNX1 and 2 expression and regulation	204
6.3.2 Brief investigation of LNX3 expression	206
Chapter 7	208
General Discussion and Future Perspectives	208
7.1 Expanding the LNX1 and 2 interactomes	209
7.2 The ubiquitome of LNX2	211
7.3 Structural and functional insights of LNX protein evolution.....	212
7.4 Analysis of LNX expression in the context of cancer.....	214
7.5 Identified limitations and future directions	215
Bibliography.....	218
Supplementary Figures	237
Chapter 3	238
Chapter 4	239
Chapter 5	243
Chapter 6	249

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

Signed:

A handwritten signature in black ink, appearing to read 'Leah Erwin', is written over a horizontal line.

Leah Erwin

Acknowledgements

Completing this thesis marks the end of a long and transformative journey, and I there are so many people without whom this would not have been possible, along with my sponsors, Eli Lilly and SEFS University College Cork

Firstly, I would like to sincerely thank my supervisor, Paul, for his patience, guidance and support over the past five years. From my time as a MSc student through to the completion of this PhD, your mentorship has been a constant source of encouragement and direction. I am truly grateful for the time and knowledge you have shared with me throughout this journey.

To Laura, Grazie mille, amore mio for being by my side through it all. Sharing the highs and lows, especially throughout the infamous LNX project, made even the most challenging moments easier. I couldn't have asked for a better partner in both science and survival.

Nicola and George, thank you both so much for your friendship over these years, both in and out of the lab. You have always been there for me, both scientifically and emotionally.

To all the members of our Supergroup, past and present – Kellie, Tom, Justin, Elena, Joan, Vanja, Oza, Olivia, Abha, Shaun, James, Magdalina and Aideen, thank you for creating such a fun, collaborative and supportive environment. It has been a privilege to work with you all. I would also like to acknowledge the technical staff, Jenny, Trish, Orla and Patrycja, whose assistance and expertise have been invaluable.

To the extended BCCB lab members and friends who have helped along the way - Gary, Katja, Rosemary, Basti, Andrew and Fabio thank you!

To my partner, Lee, thank you for your endless patience, love, support and understanding especially over the last few months. You were there even in the most stressful, hair-pulling, times, and I am so grateful for your support.

To my new and old friends, Killian, Ellen, Arina, Kristiana, Niamh, Claudia, Emma, thank you for your encouragement!

And to my family, I love you all. This is what I've been up to the last 4 years! Your support and belief in me have carried me through every stage. And to Honor, I hope this inspires you to follow your curiosity and pursue STEM!

Contributions

Throughout this project, several people contributed to various aspects both conceptual and experimental. Dr. Paul Young participated and supervised in all elements of this thesis.

In Chapter 3 Section 3.2.1, ProfAff database searching was carried out by Lucy Holmes, an undergraduate student I supervised throughout her Final Year Project. Lucy curated a list of LNX1 and LNX2 PDZ domain interacting candidates and carried out some molecular cloning and yeast-two hybrid assays. In Section 3.2.4, Dr. Laura Cioccarrelli assisted me in fluorescent microscopy imaging.

In Chapter 4 Section 4.2.5, Laura analysed the expression of LNX2 ubiquitin substrate candidates and endeavoured to perform additional validation. Paul carried out the anti-GFP nanobody displaying yeast pull downs of candidates and analysed the flow cytometry outputs.

In Chapter 6 Section 6.2.1, George Merrin facilitated the creation of *LNX1* and *LNX2* expression heatmaps using the UCSC Xena browser and RStudio.

Publications and Other Scientific Contributions

Jurić, V., Erwin, L.G., O’Riordan, N.M., Maher, E., Holmes, J.D., Young, P.W., 2025. Yeast MoClo secretion and surface display toolkit 2.0: improvements and applications for analysis of protein-protein interactions and whole-cell biocatalysis. bioRxiv 2025.08.19.671047. <https://doi.org/10.1101/2025.08.19.671047>

Note:

Chapter 3 of this thesis incorporates data/figures (Figure 3.6 & 3.7) from the above paper that I generated or contributed to. Figures or data from the paper that I did not contribute to are not included in this thesis.

Cioccarelli L, Lenihan JA, Erwin LG, Young PW. Differential neuronal functions of LNX1 and LNX2 revealed by behavioural analysis in single and double knockout mice. Behav Brain Funct. 2025 Apr 23;21(1):13. doi: 10.1186/s12993-025-00276-z. PMID: 40269869; PMCID: PMC12020136.

Note:

This thesis does not include figures or data from this publication.

Conferences

- Virtual institute of Bioinformatics and Evolution (VIBE), Trinity College Dublin. 8th-9th December 2025. Presented components of Chapter 5 in a short talk.
- Ubiquitin and Friends Symposium, Universität Wien, Vienna. 8th-9th May 2025. Travel grant awardee. Presented a poster based on Chapter 4.
- European Association for Cancer Research (EACR), Athlone. 22nd-24th February 2023. Poster presentation detailing components of Chapter 6.

Abstract

The LNX (Ligand of Numb Protein-X) / PDZRN (PDZ and RING) family are evolutionarily conserved E3 ubiquitin ligases characterised by a catalytic RING domain and multiple PDZ (protein binding) domains, originally identified through their interaction with the cell fate determinant NUMB. While there are several studies describing LNX proteins and their roles in vertebrate development, especially neuronal development in mice and dorso-ventral patterning in zebrafish, their broader physiological functions remain poorly understood. This is largely due to their extremely low endogenous expression levels and the complexity of their interaction networks.

This thesis aims to provide a comprehensive depiction of LNX proteins, integrating bioinformatic, biochemical and experimental approaches. Evolutionary analyses in Chapter 5, demonstrate that LNX proteins maintain a highly conserved domain architecture across metazoan lineages, suggesting strong functional conservation. Structural modelling reveals that the second PDZ domain adopts an exposed conformation, explaining its promiscuity in mediating protein interactions, as confirmed in Chapter 3. However, specific interactions, such as the LNX-NUMB association mapped to the NPAY/F motif, appear to be vertebrate specific, indicating evolutionary neofunctionalisation.

To further characterise the LNX protein interactome, a combination of yeast-two hybrid assays and affinity-purification based approaches were employed. While high-throughput datasets like ProfAff predict hundreds of potential interactors, experimental

validation is necessary to reveal a more restricted set of physiologically relevant binding partners (Chapter 3 and Chapter 4). As an outcome of this approach, several novel interactions were identified, including neuronal associated proteins SDK2 and NRXN1 β along with the transcriptional regulator ZNF24.

To address challenges associated with low-abundance E3 ligases, new methodologies such as BioE3/Ub-POD were modified to identify LNX2 substrates. While proteomic analyses revealed diverse ubiquitination targets, interpretation of this data is complicated by apparent residual catalytic activity in the mutant LNX2 protein used as a control. Nevertheless, a list of putative novel LNX2 substrates was identified in Chapter 4. Several of these candidates were confirmed as binders of LNX PDZ domains, though further rigorous validation of their status as direct substrates of LNX2 ubiquitination is required.

Lastly in Chapter 6, investigation of *LNX1* and *LNX2* expression showed that genomic amplification of LNX loci in colorectal and glial malignancies, does not appear to result in detectable protein levels in relevant cancer cell lines. *LNX1* appears to act as a passenger gene in gliomas, whereas *LNX2* displays expression patterns consistent with oncogenic drivers in colorectal cancer. Despite detectable mRNA and cell line transcriptomic data (Cancer Cell Line Encyclopaedia) suggesting RNA expression, endogenous protein levels remain low, suggesting tight post-transcriptional regulation. Promoter activity is generally weak but can be induced modestly under specific conditions like nutrient starvation, protein kinase C activation and drug treatments, indicating a role for LNX proteins in cellular stress and signalling adaptation.

Overall, the data presented in this thesis provide novel functional and evolutionary insights regarding the enigmatic LNX1 and LNX2 ubiquitin ligases.

Abbreviations

(i)ptm- (interfaced) predicted template model

AP - affinity purification

BCR - breakpoint cluster region

BOZ - Bozozok

BSA – bovine serum albumin

CNS - Central Nervous System

Co-IP - co-immunoprecipitation

CRC - colorectal cancer

DAPI - 4',6-diamidino-2-phenylindole

DMSO - dimethyl sulphoxide

DUB - deubiquitinating enzyme

ECL – enhanced chemiluminescence

EDTA - ethylenediamine tetra-acetic acid

EMT - epithelial to mesenchyme transition

FBS - foetal bovine serum

GBM – glioblastoma multiforme

GFP - green fluorescent protein

GST - glutathione-S-transferase

GWAS - genome-wide association study

HECT - homologous to E6-AP carboxy-terminus

HEK - human embryonic kidney cells

HRP – horseradish peroxidase

IP - immunoprecipitation

IPTG - isopropyl- β -D-thiogalactopyranoside

KO - knockout

LNK - Ligand of numb protein X

MBP – maltose binding protein

MS - mass spectrometry

NEM - N-ethylmaleimide

PAE – predicted error alignment

PBK - PDZ binding kinase

PBS - phosphate buffer saline

PCR - polymerase chain reaction

PDB - Protein Data Bank

PDZ - PSD-95, dlga, ZO-1

PDZRN - PDZ and RING

PM - plasma membrane

PRR – Proline rich region

PTB - phosphotyrosine binding domain

RING - Really Interesting New Gene

SDS - sodium dodecyl sulfate

SNP - single nucleotide polymorphism

TBS(T) – tris buffered saline (Tween-20)

Ub - ubiquitin

UBC - ubiquitin-conjugating

uORF – upstream open reading frame

UTR – untranslated region

WB - Western blot

WT - wild-type

Y2H – yeast-two hybrid

Znf - zinc finger motif

Chapter 1

General Introduction

1.1 Ubiquitination

1.1.1 The discovery and molecular structure of ubiquitin

Ubiquitin was originally isolated from a bovine thymus in 1975. The Goldstein group were investigating thymopoietin hormone and discovered a polypeptide chain, 76 (first thought to be 74) amino acids in length that induced lymphocyte differentiation and assigned it - Ubiquitous Immunopoietic Polypeptide or UBIP (Goldstein *et al.*, 1975). The group later concluded that bovine ubiquitin is identical to the human sequence and is indeed found in not only thymic cells, but all other animal cells along with yeasts and plants (Schlesinger and Goldstein, 1975). These discoveries lead to the appreciation of how essential ubiquitin is to all living cells.

The tertiary structure of the ubiquitin molecule was resolved by X-ray crystallography and NMR spectroscopy. The structure is composed of five-stranded β -sheet ($\beta 1$ – $\beta 5$), a single α -helix ($\alpha 1$) and short loops connecting secondary structural elements (see Figure 1.1) (Vijay-kumar, Bugg and Cook, 1987). This compact and highly thermostable and pH resistant conformation gives rise to the ubiquitin fold, which provides the basis for a multitude of protein – protein interactions (Agrata and Komander, 2025).

Ubiquitin contains many distinct surface patches that mediate interactions with ubiquitin-binding domains (UBDs). One of the most prominent is a hydrophobic region, centred around residues Leu8, Ile44, and Val70. Hydrophobicity is critical for recognition by many ubiquitin receptors and enzymes involved in the ubiquitination system (Sloper-Mould *et al.*, 2001).

Additionally, ubiquitin contains seven lysine residues (K6, K11, K27, K29, K33, K48, and K63) and the first methionine, each of which can serve as a site for polyubiquitin chain elongation. The C-terminal glycine residue (G76) is essential for ubiquitin's function as it is the site of linkage with the substrate or another ubiquitin moiety (Callis, 2014) (Kwon and Ciechanover, 2017).

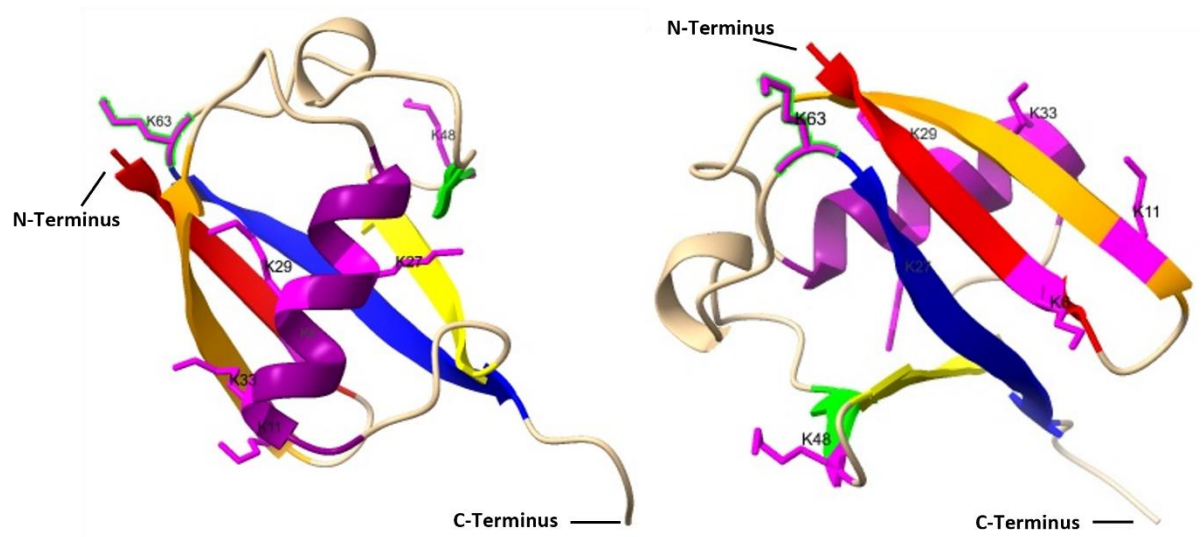


Figure 1.1: Ubiquitin tertiary structure. The alpha helix is positioned and stabilised among the beta sheets (left) and contains 4 lysine residues (in magenta). The 3D structure is flipped (right) to highlight possible polyubiquitin chain sites - K48 and K63 (magenta) (right). Amino and carboxyl termini are indicated. $\beta 1$ =red, $\beta 2$ =orange, $\beta 3$ =yellow, $\beta 4$ =green, $\beta 5$ =blue, $\alpha 1$ =purple, lysine residues=magenta. PDB ID: 1UBQ structures annotated using ChimeraX 1.10 (Meng *et al.*, 2023).

1.1.2 The mechanism of ubiquitination

The dynamic state of cellular proteins has been recognised for almost 85 years (Schoenheimer, 1942). We now understand that the mechanism of attaching ubiquitin molecule or molecules to a protein is a method to signal degradation in a three-step enzymatic cascade known as ubiquitination. This discovery was awarded the 2004 Nobel Prize in Chemistry to Irwin Rose, Avram Hershko and Aaron Ciechanover, their work is nicely summarised in (Wilkinson, 2005).

Ubiquitination is a highly regulated, ATP-dependent, post-translational modification (PTM) in which one or many ubiquitin molecules are enzymatically attached to a substrate protein (see Figure 1.2). Ubiquitin proteins are added to a lysine residue on the substrate protein. The isopeptide bond is formed between the free carboxyl group (COO^-) of glycine 76 on ubiquitin and the ϵ -amino group ($\epsilon\text{-NH}_3^+$) on the substrate's lysine residue (Pickart, 2001). Depending on location or linkage pattern, lysine residues on substrate proteins or on ubiquitin itself can be modified during chain formation, thereby contributing to the structural and functional diversity of ubiquitin signals (Callis, 2014) (Kwon and Ciechanover, 2017). The entire process requires three classes of enzymes, E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugating enzymes and finally E3 ubiquitin-ligating enzymes, allowing specific and efficient ubiquitination within a cell.

1.1.3 Ubiquitin activation

There are only two ubiquitin activating enzymes (E1) in the human genome: *UBA1* and *UBA6* (Groettrup *et al.*, 2008). The process begins with an E1 enzyme binding both, inactive ubiquitin and an ATP molecule, sequentially catalysing the acyl-adenylation of the carboxyl terminal of ubiquitin, forming an active ubiquitin-AMP intermediate. Next, a thioester linkage is formed between the terminal (G76) COO⁻ group of ubiquitin and the E1's cysteine sulphydryl group with the release of AMP (Schulman and Wade Harper, 2009) (Kwon and Ciechanover, 2017).

1.1.4 Ubiquitin conjugation

Once the ubiquitin molecule is attached to an E1 enzyme, the ubiquitin-conjugating enzyme (E2) can interact with both the E1 and activated ubiquitin, catalysing the trans-thioesterification reaction of ubiquitin to a cysteine residue on the E2 enzyme (Jin *et al.*, 2007). E2 enzymes are more common in the human genome, totalling 35 unique genes.

1.1.5 Ubiquitin ligation

Penultimately, the ligation of ubiquitin to the substrate protein's lysine residue is completed using E3 enzymes. These class of enzymes are the most populous of the three with hundreds of genes encoding ubiquitin ligases. An isopeptide bond is formed between the lysine residue of the substrate and glycine 76 of the ubiquitin. E3 ligases act as substrate recognition modules and can distinguish between thousands of E2 and substrate combination (Pickart and Eddins, 2004). Finally, several individual ubiquitin molecules (monoubiquitination) can be added to the same substrate on different lysine residues or even a chain of multiple ubiquitin proteins (polyubiquitination) can be

formed. This enormous complexity in the ubiquitome allows for extreme specificity of substrate fates in the cell (Komander, 2009). I will further discuss three main groups of E3 ligases – RING (Really Interesting New Gene) finger domain ligases, RING-between-RING (RBR) ligases and HECT E3 ligases (Homologous to E6-AP Carboxyl Terminus).

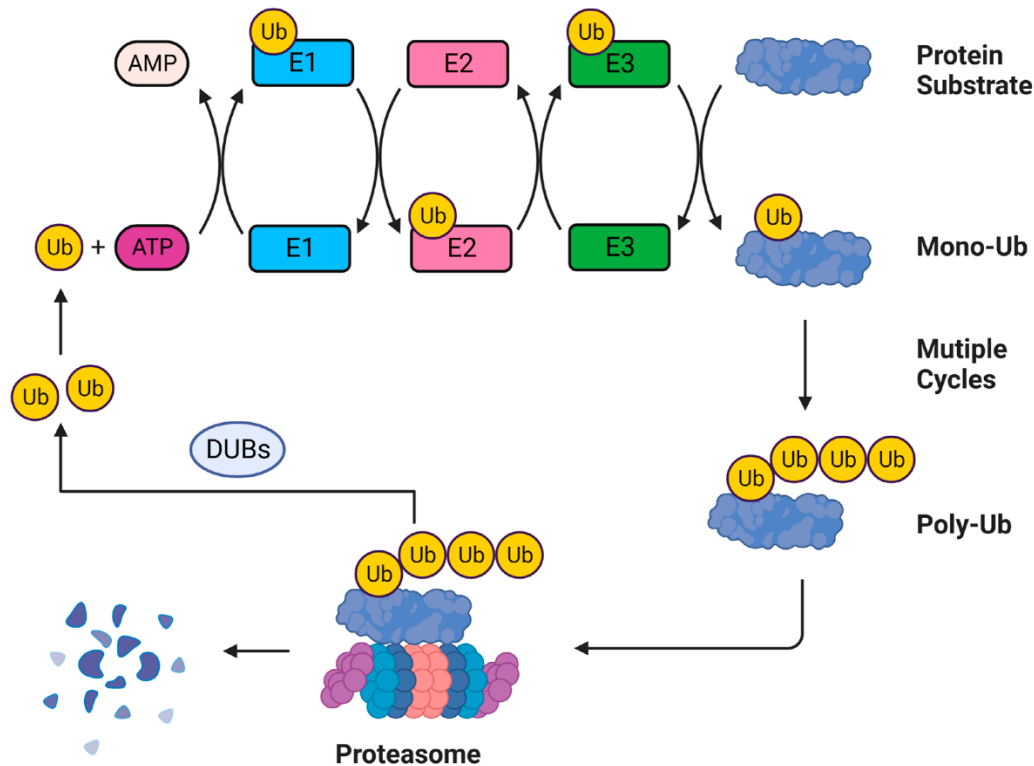


Figure 1.2: Schematic representation of the ubiquitination mechanism. The process is initiated by E1 ubiquitin-activating enzymes when ATP is used to activate the ubiquitin molecule (Ub), forming the E1-Ub (AMP) intermediate. Subsequently, the transfer of active ubiquitin to the sulfhydryl site of a cysteine residue on the E2 conjugating enzyme occurs, forming the E2-Ub conjugate. The E3 ligase catalyses the covalent interaction between the E2-Ub and the substrate, effectively “labelling” the substrate’s lysine with ubiquitin. Monoubiquitination occurs with the addition of one ubiquitin molecule per lysine residue or polyubiquitination when several cycles of ubiquitination occur. The fate of the substrate is determined based on the pattern of ubiquitination one of which is degradation by the proteasome. Ubiquitin molecules can also be removed by deubiquitylating enzymes (DUBs) and recycled. Made using BioRender.

1.1.5.1 HECT domain E3 ligases

The mechanism of how ubiquitin ligases operate was first detailed in the HECT E3 ligase family from research carried out on tumours arising from human papilloma virus (HPV) infections (Scheffner, Nuber and Huibregtse, 1995). Oncogenic strains of HPV possess a viral protein known as early protein 6 (E6). Human cells are hijacked by HPV following infection, the E6 viral protein sequesters the ubiquitin ligase, E6-associated protein (E6-AP), and leads to the ubiquitination of p53 and subsequent degradation (Talis, Huibregtse and Howley, 1998). E6-AP contains a catalytic domain at the C-terminus that is 350 amino acids long. Thereafter, bioinformatical analysis revealed a family of E3 ligases in the human proteome that contain this conserved domain and the family was named “homologous to E6-AP carboxyl-terminus” (HECT) to communicate their origin (Huibregtse *et al.*, 1995).

The amino terminal of E6-AP mediates substrate recognition, while the catalytically active carboxyl terminal interacts with the E2-Ub complex and accepts the ubiquitin at the active cysteine residue located in the HECT domain. The formation of a thioester interaction with ubiquitin is unseen in other classes of E3 ligases. The E2 conjugation enzyme is released and the HECT E3 ligase forms an isopeptide between the ϵ -amino group of the substrate lysine and the ubiquitin's glycine 76 (Rotin and Kumar, 2009).

Crystal structures of HECT E3 ligases have been solved revealing a clear bilobular, L-shape conformation. The larger amino terminal lobe recognises the substrate while the smaller carboxyl terminal lobe, contains the conserved, catalytic cysteine residue of the HECT domain (Huang *et al.*, 1999). The hinge between the two lobes is flexible and is required

for the juxtaposition of E2 and E3 cysteine residues during ubiquitin transfer. The importance of this area makes it a hotspot for clinically significant mutations, one being Angelman syndrome. A mutation from isoleucine to lysine (I827K) in the catalytically active area of the E6-AP E3 ligase, disrupting E6-AP dependent ubiquitination (Beasley *et al.*, 2020).

To date, 28 HECT E3 ligases have been described which can be further subcategorised based on the substrate specificity of the amino terminal extensions: Nedd4 family – contains tryptophan-tryptophan (WW) domains, the HERC family – these contain RLDs (RCC1-Like Domains), and other HECT ligases that do not contain either (Rotin and Kumar, 2009).

1.1.5.2 RING finger domain E3 ligases

This class of E3 ligases, amounting to over 600 genes in humans, share a common 40-60 amino acid region known as a RING domain (Li *et al.*, 2008). RING domains consist of an octet of spatially conserved seven cysteine residues and one histidine residue that bind two atoms of zinc. The basic sequence of the canonical RING is Cys-X₂-Cys-X₍₉₋₃₉₎-Cys-X₍₁₋₃₎-His-X₍₂₋₃₎-Cys-X₂-Cys-X₍₄₋₄₈₎-Cys-X₂-Cys (where X is any amino acid). These zinc binding domains or zinc fingers border each side of the RING domain and endow the RING domain with a globular formation centralised by an α -helix (Freemont *et al.*, 1991). RING finger zinc binding domains are unlike zinc finger transcription factor domains due to the difference in Cys/His spacing resulting in 3D conformational variances (Callis, 2014). In contrast to HECT type E3 ligases, RING finger ligases do not form a thioester intermediate with a ubiquitin molecule, instead they transfer the ubiquitin directly from the E2

conjugating enzyme to the substrate protein (Raymond J. Deshaies and Joazeiro, 2009). There are also RING-type E3 ligases that are present and function as a multisubunit complex. The anaphase promoting complex/cyclosome (APC/C) is a large multisubunit E3 ligase comprised of 13 core subunits in humans. The APC/C complex contains a cullin like scaffold protein and a small RING domain containing protein. Activity is regulated by two interchangeable co-activators - CDC20 and CDH1, which recognise distinct substrate proteins and function at various stages of the cell cycle (Schreiber *et al.*, 2011).

1.1.5.3 RING-Between-RING E3 ligases

RING-between-RING (RBR) E3 ubiquitin ligases represent a distinct family of 14 E3 enzymes that combine structural and mechanistic features of both RING and HECT E3 ligases (Wang *et al.*, 2023). RBR proteins are characterised by a conserved tripartite domain architecture comprising a RING1 domain, an In-Between-RING (IBR) domain, and a catalytic RING2 domain. Within the ubiquitin system, the RING1 domain recruits the ubiquitin-loaded E2 conjugating enzyme, while the catalytic cysteine residue within RING2 accepts ubiquitin via a transthiolation intermediate before transferring it to the substrate. This two-step mechanism distinguishes RBR ligases from canonical RING E3 ligases, which catalyse direct ubiquitin transfer from the E2 enzyme to the substrate (Lechtenberg *et al.*, 2016). In humans, the best-characterised RBR E3 ligase is Parkin (PRKN), which plays a central role in mitochondrial quality control through mitophagy, and whose dysfunction is associated with autosomal recessive Parkinson's disease (Gladkova *et al.*, 2018). More recently, structural studies have demonstrated that this

catalytic mechanism is conserved across the RBR family, providing a unifying model for ubiquitin transfer by these enzymes.

1.1.6 Types of ubiquitination

Ubiquitination is an extremely intricate system of interactions and signalling cascades that has been accurately named “The Ubiquitin Code” by David Komander and Micheal Rape (Komander and Rape, 2012). Researchers began to crack the code with the arrival of proteomic mass spectrometry for PTM analysis (Gupta *et al.*, 2007). Trypsin digestion, commonly used for mass spectrometry of proteins, cleaves peptides after an arginine or lysine residue. Conveniently, the carboxyl terminus of ubiquitin reads leucine-arginine-glycine-glycine. After tryptic digest of polypeptides, di-glycine (GG) tails remain on a substrate leaving a recognisable fingerprint in the proteome (Komander, 2009).

1.1.6.1 The power of one - monoubiquitination

Monoubiquitination refers to the addition of a singular ubiquitin on a lysine residue of a substrate protein while multi-monoubiquitination (see Figure 1.3) is similar but more than one substrate lysine is ubiquitinated (Kwon and Ciechanover, 2017). Monoubiquitination is usually employed by the cell when proteolysis is not the end fate for a substrate. More commonly, monoubiquitination signals for DNA repair, transcription and cellular trafficking depending on the location of the PTM (Passmore and Barford, 2004). In the case of p53, multi-monoubiquitination by MDM2 results in nuclear export (Li *et al.*, 2003). Modification of the p53 C-terminus leads to a conformational change which exposes the nuclear export signal (NES) and also aids

dissociation from MDM2 (Carter *et al.*, 2007). Monoubiquitination of histone proteins is necessary for the control of transcriptional activation and regulation. Ubiquitination can lead to cross talk between other PTMs, one of which is promoting the addition or removal of the second modification by increasing the binding affinity of an enzyme for the nucleosome (Kriegesmann and Brik, 2023). An example of this is the monoubiquitination at K120 in human histone H2B which allows for subsequent acetylation of H2BK120-Ub and affecting transcriptional regulation (Zhu *et al.*, 2005).

1.1.6.2 The infinite complexity of polyubiquitination

The formation of ubiquitin polymers on a substrate protein can label it for a variety of processes within the cell – proteolysis (proteasomal or lysosomal), DNA repair, coordinating cellular localisation, signal transduction, activation and deactivation and modulating protein interactions (Hershko and Ciechanover, 1998) (Mukhopadhyay and Riezman, 2007)(Wilkinson, 2005)(Komander, 2009). The molecular function can be determined by the topology of ubiquitin chain and the location of the lysine residue (see Figure 1.3).

In cells, polyubiquitination side chains are found on all seven lysine residues along with the first methionine of ubiquitin (Peng *et al.*, 2003) (Komander and Rape, 2012). The function of all ubiquitin chain moieties is not completely understood to date, nonetheless, I will attempt to give an overview of several well studied examples.

Polyubiquitination through K48 side chains generally directs substrates for 26S proteasomal degradation. This fact was determined very early on by the Nobel Laureates Hershko and Ciechanover in 1998 (Hershko and Ciechanover, 1998). K48 polymers are

the most common chain type in eukaryotic cells and thought to be the canonical variation. If at least four ubiquitin moieties are linked in a K48 chain they can arrange in a tetrameric conformation and proteins are usually degraded within minutes of ubiquitination (Thrower *et al.*, 2000). Similarly, K11 chains also serve as degradation signals. In yeast cells, K11 chains synthesised on substrates by the E2 conjugating enzyme Ubc6, have been implicated in endoplasmic-reticulum-associated-degradation (ERAD) (Xu *et al.*, 2009). In mammalian cells, K11 linked chains have a role in the cell cycle control. The E2 enzyme UBE2C functions alongside the E3 ligase APC/C, to assemble short K11 chains on substrates requiring degradation during this process (Jin *et al.*, 2008).

K63 ubiquitin chains can serve non-proteasomal degradative purposes in the cell. One example of this is the redistribution of CD8 α away from the plasma membrane mediated by ubiquitination by LNX1 or 2. LNX1/2 can mediate non-K48 chain formation on CD8 α promoting endocytosis and lysosomal degradation (D'Agostino *et al.*, 2011). These types of homotypic chains serve in DNA damage response and signalling. One of the most elegant examples of this is cytokine signalling (Komander, 2009). In response to proinflammatory cytokines, RING domain E3 ligase, TRAF6 (TNF receptor-associated factor 6), forms K63 polyubiquitin chains on MAP3K7/TAK1 (TGF β ((transforming growth factor β)-activated kinase 1) triggering it's activation by phosphorylation and allosteric changes. Subsequently, MAP3K7 is responsible for the activation of IKK (I κ B kinase) complexes and eventually to the activation of the NF- κ B transcription factor (Deng *et al.*, 2000).

K6, K27 and K33-linked polyubiquitin chains are less characterised and generally not linked as degradative signals. Below, I will give a brief overview of each here. K6 chains are implicated in DNA repair and mitochondrial quality control. The RBR E3 ligase Parkin produces K6 chains during mitophagy to facilitate the recruitment of autophagy receptors to damaged mitochondria (Ordureau *et al.*, 2014). K27 ubiquitin chains are linked to regulation of innate immune signalling and inflammatory responses. K33 chains are associated with intracellular trafficking and kinase regulation (Zhou and Zhang, 2022). All in all, polyubiquitination is dynamic and rarely simple in which cell processes often require multiple different homotypic and heterotypic ubiquitin signals (Komander, 2009).

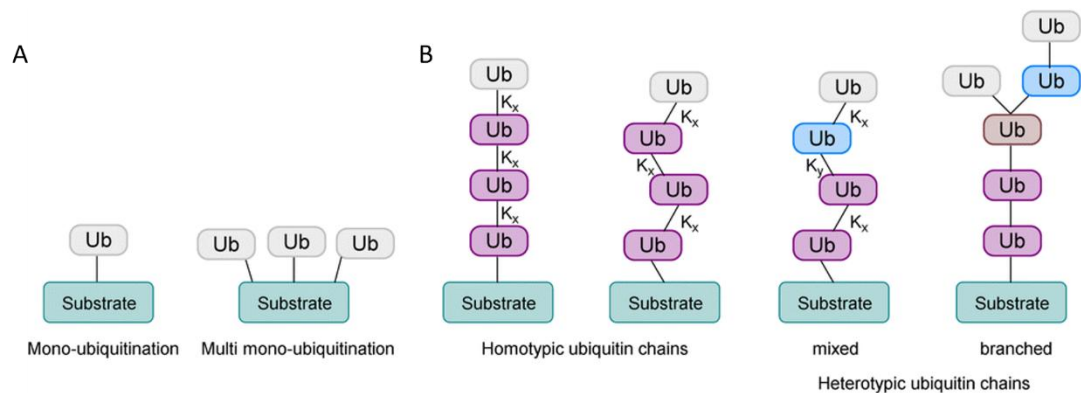


Figure 1.3: Diagram illustrating types of substrate ubiquitination. A. Monoubiquitination refers to a singular lysine having one ubiquitin molecule attached while multi-monoubiquitination is where multiple lysine residues on the same substrate have one ubiquitin attached. **B.** Polyubiquitination describes the formation of Ub-Ub chains, joined together by isopeptide bonds between one of the seven lysines (K6, K11, K27, K29, K33, K48, K63) or M1 residues of ubiquitin and the G76 of another. If the same residue is modified during elongation, the Ub chains are called homotypic, whereas different linkages lead to mixed chains – heterotypic. If a single Ub is modified with multiple Ubs, this leads to branched chains. K_x and K_y= different lysine residues. This figure is adapted from (Kriegesmann and Brik, 2023), under the terms of a Creative Commons 4.0 (CC BY) attribution license

1.1.7 Deubiquitinating enzymes

Ubiquitination is a reversible post translational modification. As molecules of ubiquitin are required so frequently in cellular processes, they are usually not degraded with the substrate protein, they are detached by a superfamily of deubiquitinating enzymes (DUBs) and recycled. Classically, the protease family of DUBs contains around 100 members and can be subdivided into five classes – UCHs (ubiquitin C-terminal hydrolases), USPs (ubiquitin specific proteases), OTUs (ovarian tumour proteases), Josephins and JAMMs/MPN+ (JAB1/MPN/Mov34 metalloenzymes) (Nijman *et al.*, 2005) (Reyes-Turcu, Ventii and Wilkinson, 2009). Recently, two new, albeit small, classes of DUBs have been discovered – the MINDY family of K48-specific deubiquitinates (Abdul Rehman *et al.*, 2016) and a ZUFSP class which only contains the *ZUP1* gene (Kwasna *et al.*, 2018).

Most DUBs carry out their deubiquitinating functions by hydrolysing the isopeptide linkages between the lysine ϵ -amino group and a carboxyl group glycine 76 of ubiquitin (Komander, Clague and Urbé, 2009). Members of the cysteine protease group; USPs, UCHs, OTUs and Josephins contain two or three crucial amino acids (Cys, His and Asn/Asp) that are responsible for launching a nucleophilic attack on isopeptide bonds (Storer and Ménard, 1994).

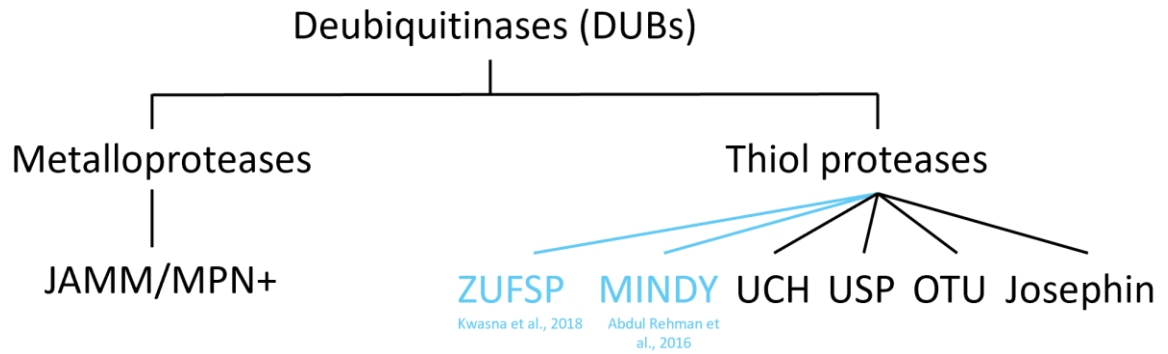


Figure 1.4: Schematic outlining the classes of deubiquitinase (DUBs) enzymes. Only one group of metalloproteases are known to date - JAMMs/MPN+ (JAB1/MPN/Mov34 metalloenzymes). Thiol or Cys containing proteases are a bigger class with six members - UCHs (ubiquitin C-terminal hydrolases), USPs (ubiquitin specific proteases), OTUs (ovarian tumour proteases), Josephins, MINDY and ZUFSP. MINDY and ZUFSP are in blue to illustrate recently identified classes.

1.1.8 Ubiquitination in the context of cancer

Uncontrolled cellular proliferation, evading immune destruction and resisting cell death are but a few of the hallmarks of cancer (Hanahan and Weinberg, 2000). There have been multiple studies on how ubiquitination and its dysregulation can affect these principles, a few examples of roles played by ubiquitination in cancer are highlighted below.

As aforementioned, expression of the tumour suppressor p53 is controlled by the E3 ligase MDM2 along with many other E3 ligases. Without a doubt there are many factors at play regarding p53 in cancer, but control of its ubiquitination status is a valuable subject. MDM2 works as a ubiquitin ligase to label p53 for proteasomal degradation when MDM2 levels are high, and as an inhibitor of transcriptional activation by directly binding to p53 (Brooks and Gu, 2006). *MDM2* is amplified in multiple tumour types at an incidence rate of 4-20% suggesting it's an important player in cancer (Chen *et al.*, 2026).

Promoting ubiquitination of tumour suppressive proteins is one method of facilitating tumorigenic progression. Another method is deubiquitinating oncogenic factors that have been targeted for degradation. The transcription factor c-MYC, a potential oncogenic factor of the WNT signalling pathway, is commonly dysregulated in most colorectal cancers (CRC) (Muzny *et al.*, 2012). The DUB enzyme USP28 antagonises ubiquitination of c-MYC by “piggybacking” on the E3 ligase FBW7 and a cycle of ubiquitination and deubiquitination controls c-MYC stability (Popov *et al.*, 2007). Overexpression of USP28 in human and murine intestinal tumours is frequent. Depletion of USP28 in human CRC cell lines and in a murine CRC model significantly reduces not only c-MYC levels, but also c-JUN and NICD1 which are other oncogenic transcription factors of the WNT signalling pathway. *Usp28^{ΔG}* mice had fewer intestinal tumours which were smaller in size than WT mice suggesting a prominent role of DUBs in CRC (Diefenbacher *et al.*, 2014).

1.1.8.1 Druggability of the ubiquitin system in cancer

The ubiquitination system in cancers is often dysregulated, making components of ubiquitination high value targets for therapeutic interventions. PROTACs (PROteolysis TArgeting Chimeras) are a class of small, heterobifunctional molecules that can induce the proteasomal degradation of specific target proteins. PROTACs consist of two ligands connected by a linker, one ligand can target an E3 ligase the other attaches simultaneously, to a protein of interest, bringing both enzyme and target in proximity and allowing ubiquitination (Békés, Langley and Crews, 2022). PROTACs were first synthesised in 2001, where “Protac-1” was used to recruit the Skp1-Cullin-F box (SCF)

complex to methionine aminopeptidase-2 (MetAP2). MetAP2 promotes angiogenesis and cell proliferation and is commonly targeted by the angiogenesis inhibitors - ovalicin and fumagillin (Sakamoto *et al.*, 2001). The E3 ligase Cereblon (CRBN) is a target of immunomodulatory imide drugs (IMiDs) like thalidomide, pomalidomide and lenalidomide (Ito *et al.*, 2010). Lenalidomide has been approved as a therapy for multiple myeloma as it modulates CRBN to degrade Ikaros family transcription factors, IKZF1 and IKZF3, which are essential in lymphocyte development (Lu, 2014)(Kronke, 2014). CRBN is used in many PROTACs and has been widely characterised in terms of structure and ubiquitome – interacting with versatile substrates. Expression of CRBN is constitutive across a wide range of cell and tissue types making it an ideal E3 ligase for PROTAC development (Thapa *et al.*, 2024).

Occasionally, E3 ligases can be overactive in cancers and studies on inhibiting them have been extremely promising in the context of potential therapeutic strategies. Overexpression of *HUWE1* is observed in multiple myeloma cell lines (Crawford *et al.*, 2016). *HUWE1* is a large E3 ligase involved in modulating c-MYC activity depending on cellular context. *HUWE1* catalyses K48 polyubiquitination destining c-MYC for degradation but can also stabilise it through K63 linked chains (Myant *et al.*, 2017). Knockdown or small molecule inhibition (BI8622 and BI8626) of *HUWE1* in multiple myeloma models shows significant decrease in tumour burden and cell viability, respectively, especially in combination with lenalidomide (Crawford *et al.*, 2020). Very recently research was published showing that tumours with MDM2 amplification or overexpression can be selectively targeted using PROTAC approach, by employing the

compound A1874 (JQ1-Idasanutlin chimera targeting BRD4). A1874 increases cytotoxicity in cells with MDM2 overexpression allowing p53 to function normally in cells without high levels of MDM2 (Chen *et al.*, 2026). The concept of targeted protein degradation (TPD) is being explored at great lengths and there are many approved therapeutic modalities as well as treatments in clinical trials (Békés, Langley and Crews, 2022).

Deubiquitinating enzymes can also be useful targets for cancer treatments. As we know, polyubiquitin chains can be protective or destructive depending on the linkage site (Komander and Rape, 2012). USP7 is a DUB that targets both p53 and MDM2 which contributes to loss of stabilisation. The small molecule autophagy inhibitor, Spautin-1, also inhibits USP10/13 activity, allowing p53 levels to increase and execute tumour suppressive activities (Liu *et al.*, 2011).

1.2 PDZ domains

1.2.1 Classification of PDZ domains

The crystal structure of a PDZ (PDB: 1BE9) domain, originally known as a DHR domain or GLGF (Gly-Leu-Gly-Phe) repeat (Cho, Hunt and Kennedy, 1992), was first determined from a protein domain of the postsynaptic density protein PSD-95 (Doyle *et al.*, 1996). Subsequently, a second structure was solved from the human homolog of *Drosophila discs-large* tumour suppressor, DlgA (Morais Cabral *et al.*, 1996). Following the identification of similar domains in the tight junction protein - Zonula Occludens-1, the GLGF repeat was renamed the PSD-95, DlgA, and ZO-1 (PDZ) domain to better reflect its origin and broader distribution (Kennedy, 1995).

PDZ domains are generally 80 to 110 amino acids in length and conform into six β strands and two α helices forming a carboxyl terminal peptide binding pocket as seen in Figure 1.5. The binding groove is characterized by the presence of a highly conserved lysine or arginine residue that interacts with the carboxylate group at the end of the peptide. In addition to this electrostatic interaction, three main-chain amide protons of the Gly-Leu-Gly-Phe motif form hydrogen bonds with the peptide leading to the more general binding signature R/K-XXX-GLGF (Doyle *et al.*, 1996)(Morais Cabral *et al.*, 1996)(Daniels *et al.*, 1998). There are approximately 180 unique human PDZ domain containing proteins that are known, with that many of them containing multiple PDZ domains. Nevertheless, the PDZ domains can be classified based on the consensus sequence of the PDZ binding motif (PBM) that is preferred (Songyang *et al.*, 1997) (Bezprozvanny and Maximov, 2001). The ultimate ~10 amino acids of a protein are regarded as comprising the PBM however, it is believed that the last 4 amino acids are the most crucial (Saras and Heldin, 1996). The final amino acid is denoted as 0, -1 is the second last and so on. Position 0 of PDZ binding motifs is commonly occupied by a hydrophobic residue. Variations in the conformation of the PDZ binding pocket determine residue specificity, as exemplified by PDZ1 and PDZ2 of PDS-95, which preferentially recognise carboxyl terminal sequences terminating in valine, leucine or isoleucine (Songyang *et al.*, 1997)(Cohen *et al.*, 1996). The residue at position -2 is the determinant of the PDZ classification and is crucial for binding specificity (Nardella *et al.*, 2021). For simplicity, I will discuss the three main classes of PDZ domains albeit, there are further subclass divisions that bring the total 16.

1.2.1.1 Class I

PDZ domains are assigned to class I if they prefer a peptide ligand with a serine or threonine at position -2 and either a valine or leucine as the last residue - X-S/T-X-V/L. Serine or threonine at position-2 is vital to enable hydrogen bonding with a highly conserved histidine residue in the $\alpha 2$ helix of the PDZ domain binding pocket (Bezprozvanny and Maximov, 2001).

1.2.1.2 Class II

Class II PDZ domains have a binding signature of X- Φ -X- Φ , X being any residue and Φ being a hydrophobic residue (Songyang *et al.*, 1997).

1.2.1.3 Class III

The presence or absence of an aspartate at position -2 of a binding ligand determines the class category as class III as shown here - X-D-X-V. The possession of an aspartate residue provides access to a second carboxylate in addition to the C-terminal one (Zimmermann *et al.*, 2002). This characteristic offers PDZ domains 2 available carboxylate sites to recognise and generally have higher binding affinities than class II (Lee and Zheng, 2010). Class III PBMs are infrequent in the human proteome, although the important signalling protein, Dishevelled (DVL) contains one

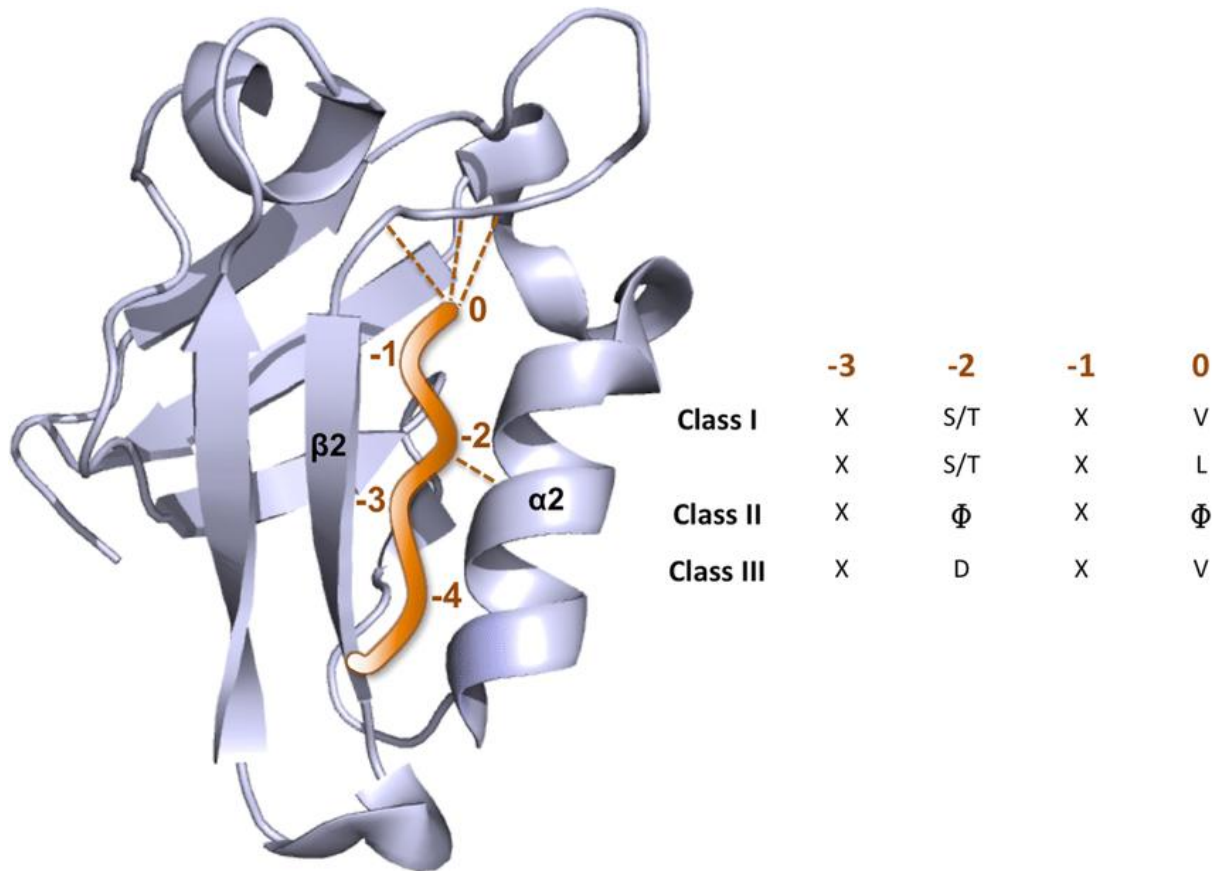


Figure 1.5: 3D structure of a PDZ domain in complex with a generic C-terminus of a ligand (in orange). The ligand (orange) is located between the $\alpha 2$ and the $\beta 2$ of the PDZ binding pocket. PDZ Binding Motif positions are highlighted as 0, -1, -2, -3 and -4. The conventional three binding classes are shown in the table in which X corresponds to an unspecified amino acid. This figure is reproduced from (Nardella *et al.*, 2021), under the terms of a Creative Commons 4.0 (CC BY) attribution license.

1.3 Ligand of NUMB Protein X family

1.3.1 NUMB proteins

NUMB was first identified in *Drosophila melanogaster* as an important determinant of cell fate in sensory organ development (Uemura *et al.*, 1989). NUMB is highly evolutionarily conserved and two homologous proteins, NUMB and NUMBL, have been found in many different metazoan lineages including nematodes (Zhong *et al.*, 1997)(Dho *et al.*, 1999). The mammalian ortholog of NUMB has the same role in development but has evolved more complex functions. The expression pattern of NUMB is indicative of an important role in neurogenesis. In adult stages, NUMB is also expressed in most tissues, suggesting a role beyond early embryonic development (Gulino, Di Marcotullio and Screpanti, 2010). NUMB is known to exhibit its function by negatively regulating NOTCH signalling, which can have many consequences in both development but also in cell proliferation making it an interesting focus for cancer research (Guo, Jan and Jan, 1996) (Gulino, Di Marcotullio and Screpanti, 2010) (Pece *et al.*, 2011).

The PTB (Phosphotyrosine Binding) motif and PRR (Proline Rich Region) of NUMB are affected by alternative splicing exons 3 and 9, respectively. The presence or absence of exon 3 results in an 11 amino acid insert (PTBi) in the PTB domain and either a long or a short PRR (PRR^{L/S}) domain if exon 9 is present (Verdi *et al.*, 1999)(Pedersen *et al.*, 2002).

These splicing events give rise to the four known isoforms of NUMB protein in mammals. The longest isoform, NUMB1 (p72), has an 11 amino acid insert in the PTB domain along with a PRR domain. NUMB2 (p66) lacks this inserted PRR domain, while NUMB3 (p71)

lacks the PTB insert. Lastly, NUMB4 (p65) is the shortest isoform with both PTBi and PRR domains absent (Dho *et al.*, 1998).

The PTB domain is required for the binding of NPXpY motifs however, it is now understood a phosphorylated tyrosine is not compulsory for binding as no change in binding affinity was observed when tyrosine was substituted for phenylalanine (Dho *et al.*, 1998). Conversely, isoform specificity for ubiquitination does exist. Alternative splicing of the PTB domain insert does affect the ubiquitination status of NUMB.

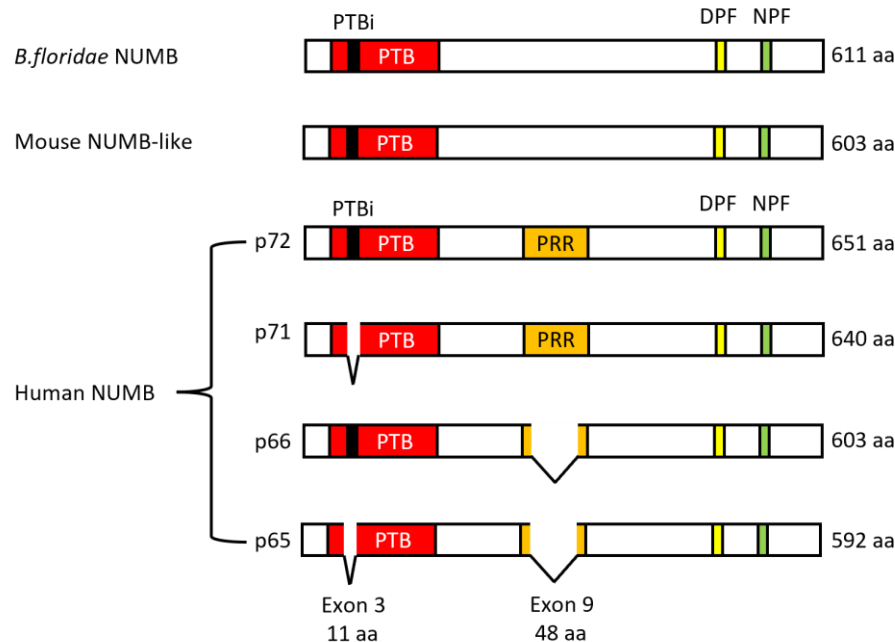


Figure 1.6. Human NUMB protein isoforms and ancestral NUMB orthologs. The diagram shows *Branchiostoma floridae*/Florida lancelet NUMB and mouse NUMB-like protein lacking a PRR domain but containing the 11 amino acid insert in the PTB. The four main isoforms of human NUMB that are relevant in LNX1/2 binding are shown below. NUMB is subjected to alternative splicing of exon 3 and exon 9 depending on tissue type and developmental stage. PTB (i) - phosphotyrosine binding domain (exon 3 insert), PRR – proline rich region (long or short), DPF/NPF – Eps15 homology domains.

1.3.2 The discovery of the Ligand of Numb Protein X family

The discovery of a novel protein that interacts with NUMB was carried out by a yeast two hybrid screen (Dho *et al.*, 1998). The term, Ligand of Numb protein X1 (LNX1) was coined due to the novel protein's ability to recognise and attach to the PTB of Numb, *via* its NPAY amino acid sequence (Dho *et al.*, 1998). LNX1 was first discovered in mice, however, the *Homo sapiens'* ortholog was revealed soon after in the foetal brain. Human LNX1 has 88% sequence identity with the mouse sequence (Xie *et al.*, 2001). LNX1 is now the first of a family of LNX proteins. LNX2 protein was uncovered soon after also by utilising a

yeast two hybrid assay, which demonstrated binding to NUMB through the LNX2 NPAF motif (Rice, Northcutt and Kurschner, 2001). The more distant relatives of the family; LNX3 and LNX4 (PDZRN3 and PDZRN4) have a RING domain and two PDZ domains. LNX3 was again discovered around the same time by a yeast two hybrid screen using Semaphorin (SEMA4C) as prey from which the acronym SEMCAP3 originates (Wang *et al.*, 1999). LNX4 and LNX5 proteins were uncovered by *in silico* analysis of LNX1 homologs in 2004. While LNX3 and LNX4 are very similar to each other, LNX5 (PDZK4 or PDZRN4L) lacks a RING domain. Nevertheless, LNX5 is still regarded as a family member as it shares a high sequence homology with LNX3 and LNX4 (figure 1.6) All three of these family members were first identified from mouse sequences (Katoh and Katoh, 2004). In human sequences, *PDZRN4L* was identified as an upregulated gene in synovial sarcomas (Nagayama *et al.*, 2004).

1.3.3 Structure of LNX1/2 proteins

In humans, LNX1 is present in two isoforms – the CNS specific isoform LNX1^{p70} and the non-CNS isoform LNX1^{p80}. LNX1^{p80} and LNX2 contain four PDZ domains and a RING domain as seen in Figure 1.7. To date, there are no full-length solved structures for any of the LNX proteins, however there are multiple PDZ domain and RING domain structures of mouse and human LNX proteins. Specifically, structures of mouse LNX1 PDZ2 (PDB ID: 3VQF, 3VQG), human LNX1 PDZ3 (PDB ID: 3B76) and human LNX2 PDZ2 (PDB ID: 2VWR, 5E1Y, 5E11, 5E21, 5E22) have been submitted to the PDB. The catalytic RING domain of LNX1 was solved in complex with the E2 conjugating enzyme Ubc13 (PDB ID: 5H7R, 5H7S), an enzyme that specifically catalyses the formation of K63 linked ubiquitin chains

(Branigan *et al.*, 2015). This provides evidence of LNX1's ability to form K63 polyubiquitin chains and that zinc finger motifs at either side of the RING domain are required for E2 enzyme interaction (Nayak and Sivaraman, 2018). The RING domain of LNX2 (PDB ID: 5DIN) was solved by the same group and in a similar fashion. They determined UbcH5b was the preferential E2 conjugating enzyme and all seven ubiquitin polyubiquitin chains could be formed. LNX2 is present as a dimer in solution and crystal structure that undergoes autoubiquitination and N-terminal polyubiquitination (Nayak *et al.*, 2015).

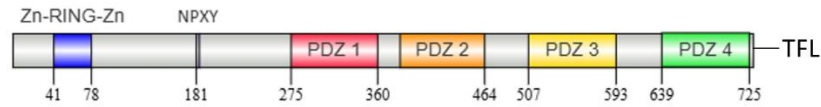
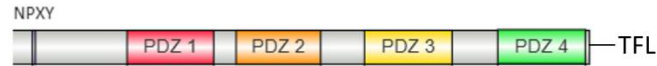
LNK1^{p80} 728aa**LNK1^{p70} 628aa****LNK2 690aa****LNK3/PDZRN3 1066aa****LNK4/PDZRN4 1036aa****LNK5/PDZRN4L 679aa**

Figure 1.7: Domain schematic of human LNK protein family. LNK1^{p80} and LNK2 share identical domain layout. LNK1^{p70} has a truncated amino terminal domain leading to the loss of the RING domain. LNK3 and LNK4 are approximately 60% identical and have a similar structure both having two PDZ domains and a RING domain. LNK5 is more distant than the rest of the family – lacking a RING domain and only having a single PDZ domain. LNK3/4/5 have a LNK3 homology domain (LNK3H) which is a sequence of significant similarity ~50% at the amino acid level, between the three proteins and is mapped to the last 150 amino acids. All family members have a class I PDZ binding motif at their carboxyl terminus as indicated by 3 amino acids. NPXY=NPAY (LNK1) or NPAF (LNK2). Figure created using IBS 2.0 Biological sequence illustrator (Xie *et al.*, 2022).

1.3.4 Evolution of LNX proteins

The *LNX1/2* or *LNX-like* gene has clear orthologs in many invertebrate lineages. The most evolutionarily distant clade being the sponges or poriferans. *Amphimedon queenslandica* has a clear *LNX1/2* ortholog that encompasses four PDZ domains and a RING domain. This would suggest that the ancestral *LNX* gene arose early in the metazoan divergence (Flynn, Saha and Young, 2011). The authors also describe a loss of a *LNX1/2*-like ortholog in multiple evolutionarily recent lineages like urochordates, arthropods, nematodes and molluscs. *Drosophila melanogaster* and *Caenorhabditis elegans* being among the model organisms lacking an obvious protein containing four PDZ domains and RING domain. However, cephalochordates and platyhelminths possess a conserved *LNX1/2* like ortholog (see Figure 1.8). All vertebrate lineages have highly conserved *LNX* proteins thus, suggesting that *LNX* has an essential function. The ancestral function of *LNX1/2* appears to be non-essential in the many invertebrate species (Young, 2018).

Throughout vertebrate evolution, the original *LNX1/2-like* gene underwent a duplication event, giving rise to three paralogous genes – *LNX1*, *LNX2* and *LNX2b* (Flynn, Saha and Young, 2011). Only eutherian mammals lack the *LNX2b* gene due to a pseudogenisation event that enabled the donation of several exons to the emerging *Xist* gene (Young, 2018). *LNX2b* lost its protein coding ability however, parts of its sequence were repurposed within *Xist*. This aided the functionality of the long non-coding RNA, *Xist*, to coat and inactivate one X chromosome in placental mammals (Duret *et al.*, 2006).

As mentioned, there are several invertebrate lineages that have a *LNX1/2*-like protein. However, none of these orthologous proteins contain an NPXY motif which would allow

NUMB protein binding (Flynn, Saha and Young, 2011). In vertebrate LNX proteins, this motif is constantly present, and the LNX-NUMB binding and regulation is deemed an essential characteristic of neuronal and central nervous system development (Rice, Northcutt and Kurschner, 2001)(Nie, Li and McGlade, 2004)(Dho *et al.*, 1998) (Won, Ro and Dawid, 2015). Interestingly, virtually all metazoan species have a NUMB or NUMBL ortholog (Dho *et al.*, 1999). Despite the presence of conserved orthologs of both LNX and NUMB, this interaction is absent, indicating that this canonical interaction likely arose in vertebrates and is dispensable for dorsal hollow nerve cord formation in invertebrates. Thus, despite their name LNX1/2 proteins are likely to have significant functions unrelated to their interaction with Numb.

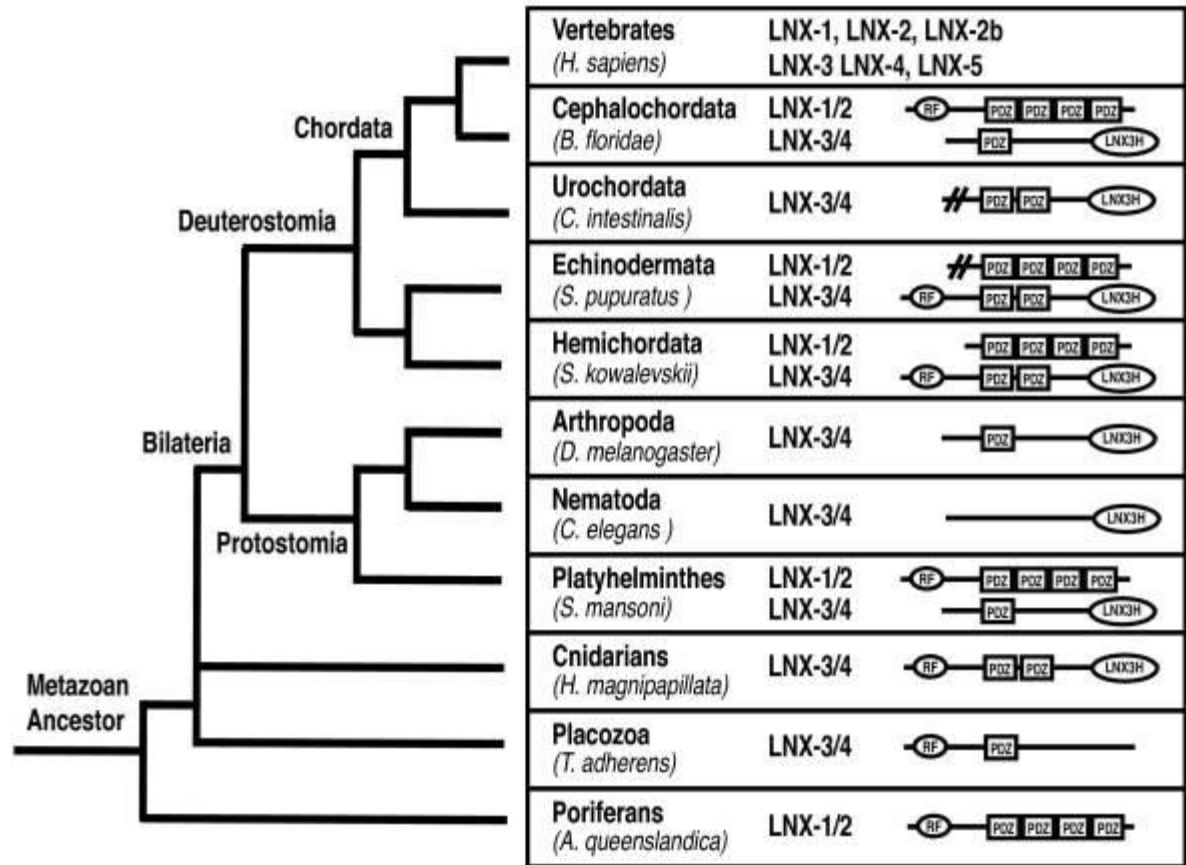


Figure 1.8: Metazoan phylogenetic tree of LNX orthologs. The domain structures of LNX protein orthologs is laid out here. A representative species has been chosen to illustrate evolutionary conservation. Proteins are classified as being orthologs of vertebrate LNX1 and LNX2 (LNX1/2) or vertebrate LNX3 and LNX4 (LNX3/4). LNX3/4 orthologs were identified based on the presence of both a RING and PDZ domain, or the presence of the LNX3H domain. At the time of publishing the *C. intestinalis* and *S. purpuratus* LNX3/4 sequences were incomplete as indicated by a double hash line. This figure is reproduced from (Flynn, Saha and Young, 2011), under the terms of a Creative Commons 4.0 (CC BY) attribution license.

1.3.5 Expression of LNX1/2 mRNA and protein in mammals

1.3.5.1 LNX1 mRNA expression

Ln timer mRNA expression is detected in most adult mouse tissues – brain, spinal cord, lung, heart, skin, skeletal muscle, kidney and testis have *Ln timer* expression (Dho *et al.*, 1998) (Mirza *et al.*, 2006). Interestingly, the same group observed a different splice variant in mouse brain and spinal cord corresponding to a smaller 2.6 kb signal by Northern blot. This signal differs from the canonical non-central nervous system (CNS) signal of 2.8 kb, which encodes the LNX1^{p80} protein. The CNS specific isoform has been determined to encode an N-terminally truncated LNX1^{p70} protein that lacks the catalytic Zn-RING-Zn domain. Later, a second *Ln timer* isoform was detected in mouse CNS corresponding to *Ln timer*^{p62} (Lenihan *et al.*, 2014). A similar expression pattern is seen for human LNX1 with the addition of pancreas and placenta. Much lower levels of human LNX1 were detected in the lung, liver and skeletal muscle (Xie *et al.*, 2001). The additional CNS isoform encoding LNX1^{p62} in mice has not been reported in human brain.

1.3.5.2 LNX2 mRNA expression

In comparison to *Ln timer*, *Ln timer*2 is present as a single isoform signature and is seen extensively across adult mouse tissues – heart, lung, spleen, kidney, thymus, skeletal muscle, liver and in the brain throughout embryonic stage E11.5 to E15.5 and declines into adulthood (Rice, Northcutt and Kurschner, 2001). *Ln timer*2 expression is also seen in mouse testis (Mirza *et al.*, 2006). Is there a similar decline in brain expression? In humans, LNX2 expression is broadly seen across most tissues including liver, pancreas, lung, breast, retina, testis and heart (see Figure 1.9). Tissues from the digestive system appear to have the highest levels of LNX2

expression– stomach, duodenum, small intestine, colon and rectum (The Human Protein Atlas) (Uhlén *et al.*, 2015).

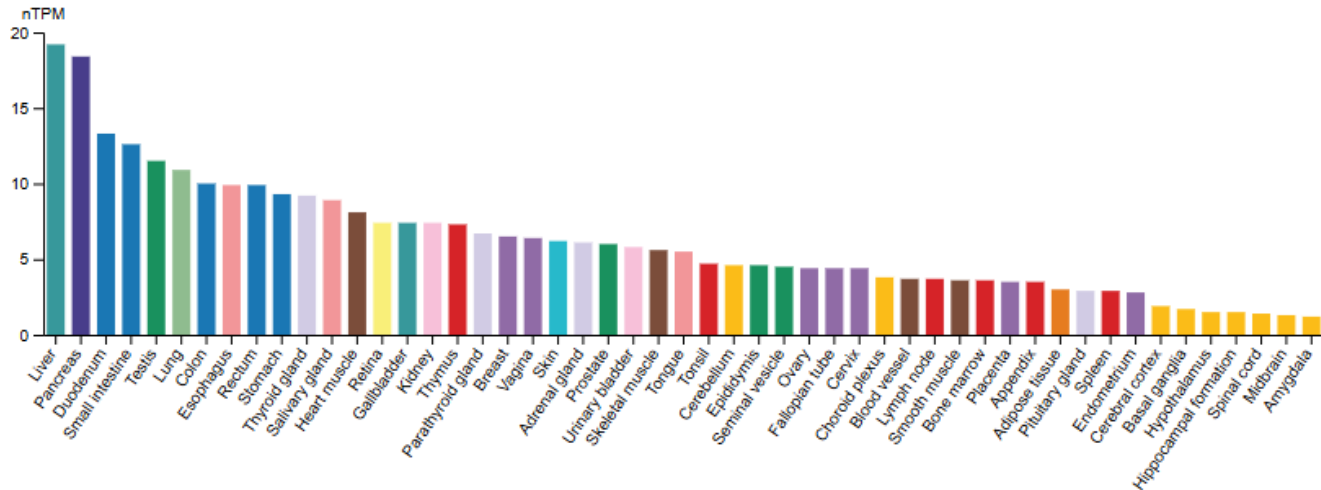


Figure 1.9: Expression of *LNX2* mRNA across normal human tissues. *LNX2* expression is shown in transcripts per million (nTPM) for a consensus dataset created by combining normalised transcriptomics datasets from GTEx and HPA. This graph is available from (Tissue expression of *LNX2* - The Human Protein Atlas, accessed 2026-01-21), under the terms of a Creative Commons 4.0 (CC BY) attribution license, link - www.proteinatlas.org/ENSG00000139517-LNX2/tissue#rna_expression.

1.3.5.3 Protein expression of *LNX1* and *LNX2*

Despite prevalent expression of *LNX1/2* mRNAs their corresponding proteins are expressed at minute levels. Endogenous *LNX1/2* proteins generally require immunoprecipitation from lysates to allow visualisation by Western blot (Young *et al.*, 2005)(D'Agostino *et al.*, 2011) (Lenihan *et al.*, 2014)(Mirza *et al.*, 2006)(Won, Ro and Dawid, 2015). *LNX2* can be detected directly from E.14.5, E.18.5 and up to postnatal day 7 in mouse brain, but not in adults (Lenihan *et al.*, 2017). Immunohistochemistry staining of *LNX2* protein in E16.5 mouse embryos, revealed widespread and uniform expression

in most organs with eye, brain and digestive system giving the strongest signal (Mirza *et al.*, 2006).

Endogenous expression LNX1 of can be detected by immunoprecipitation from mouse brain lysates (LNX1^{p70/p62}) (Lenihan *et al.*, 2014) and in perisynaptic Schwann cells of neuromuscular junctions by IHC staining (Young *et al.*, 2005). Both LNX1/2 proteins can be visualised by staining murine spermatozoa (Mirza *et al.*, 2006). It is important to note that endogenous LNX protein levels are very low. LNX1^{p70/p62} protein expression appears to be strictly associated with developmental stages which could be explained by upstream open reading frames (ORFs) or other regulatory elements in the 5' untranslated region (UTR). LNX1^{p80} and LNX2 do not have the same regulatory mechanisms (Lenihan *et al.*, 2014). Upstream ORFs can divert translation away from the main coding sequence of LNX1 and be responsible for tight regulation in the CNS (Wethmar, Smink and Leutz, 2010).

1.3.6 Functions of LNX1/2

The four PDZ domains of LNX1 and 2 allow binding to hundreds of other proteins with many of them substrates of ubiquitination. It is through this function that LNX1/2 proteins can exhibit their molecular functions.

1.3.6.1 Role of LNX proteins in neurogenesis and behaviour

NUMB is a negative regulator of the NOTCH signalling pathway, which plays a critical role in cell fate determination during early neurogenesis and nervous system development (Mumm and Kopan, 2000). During neurogenesis, the asymmetric localisation of NUMB in dividing neural progenitor cells results in its preferential segregation into one daughter

cell, where it antagonises NOTCH signalling and specifying a distinct cell fate (Gulino, Di Marcotullio and Screpanti, 2010). This significance can also be seen in zebrafish, where the NUMB and *Ln2b* (LNX ortholog) interaction has shown to be essential for the development of the nervous system (Won, Ro and Dawid, 2015). LNX proteins can control NUMB levels by ubiquitination and therefore indirectly regulate NOTCH signalling.

GLI3 is a downstream component of the Sonic hedgehog signalling pathway. While studying *Gli3* knockout mice Wang *et al* noticed an increase in LNX2 expression especially in the subventricular zone (SVZ) along with dramatically decreased NUMB levels. The SVZ is a neurogenic region in adult and embryonic mammals and contains neural progenitor cells. *Gli3*^{-/-} mice have defects in the specification of ependymal cells suggesting the LNX2 regulation of NUMB in this region is necessary to proper development (Wang *et al.*, 2014).

LNX1/2 proteins can function as molecular scaffolds stabilising multiprotein complexes. In 2018, Lui *et al* determined that LNX1 is expressed and localised post synaptically in hippocampal CA3 pyramidal neurons and stabilises EphB receptors, which are required for proper synapse formation and mossy fiber (MF) pruning. In their study, *Ln1*^{-/-} mice had abnormal synaptic arrangement and it was determined that LNX1 is required for MF axon targeting during the postnatal period (X.-D. Liu *et al.*, 2018). This group also revealed that the NMDA type glutamate receptor (GluN2B) interacts with LNX1 as part of a molecular multiplex. They explained that social memory was defective in mice that lacked LNX1 due to the disruption of the LNX1-EphB2-GluN2B complex in the CA3 region (Liu *et al.*, 2019).

Interestingly, mice lacking LNX1^{p70/p62} and LNX2 exhibit reduced anxiety-related behaviours in a set of behavioural tests like elevated plus maze and open field compared to WT animals (Lenihan *et al.*, 2017). Recently, further investigation of these double knockout and single knockout animals was undertaken. *Ln timer*^{-/-} and *Ln timer*^{-/-}; *Ln timer*^{-/-} knockout pups had more complex and significantly higher power of ultrasonic vocalisations upon maternal separation. LNX2 was attributed to an increase in risk-taking and exploratory behaviours in the wire beam bridge and light-dark emergence paradigms (Cioccarelli *et al.*, 2025).

Overall, these studies show that LNX1 and LNX2 proteins have a role in the development of the CNS which, upon deletion, causes alterations in synaptogenesis. These mechanisms are likely to underly the phenotypic variations displayed by knockout mice.

1.3.6.2 Role of LNX2 in osteoclastogenesis

A non-neuronal function of LNX2 involvement in osteoclast differentiation was described by Zhou *et al* in 2016, they found *LNX1* and *LNX2* expression is upregulated upon osteoclast differentiation *in vitro* (Zhou *et al.*, 2015). It is known that NOTCH signalling is significant in maintaining bone homeostasis and control of osteoclast fate (Engin and Lee, 2010). LNX2 was determined as a modulator of NOTCH2 signalling through the regulation of NUMB. The authors describe knockdown of LNX2 by lentiviral transduction of short hairpin RNAs (shRNAs) leads to decreased osteoclast formation and increased cell death of pre-osteoclasts. Upon LNX2 knockdown, reduced activity of important osteoclastogenic signalling pathways, namely M-CSF induced ARK and AKT activation, as

well as RANKL induced NF- κ B and JNK activity was reported (Zhou *et al.*, 2015). This work shows a promising mechanism for LNX2 involvement in formation in bone cells.

1.3.7 Disease associations of LNX1 and LNX2

As previously mentioned in section 1.1.8., E3 ubiquitin ligases are involved and frequently dysregulated in various ways in disease states. The same is true for LNX1 and LNX2 proteins.

1.3.7.1 LNX1/2 protein involvement in infectious disease

Kawasaki disease (KD) is an inflammatory paediatric condition that occurs after an excessive immune response to an infection in which some children are genetically predisposed to. KD leads to damage of the coronary artery in 25% of untreated cases (McCrindle *et al.*, 2017). A GWAS study found single nucleotide polymorphisms (SNPs) in the intronic region of *LNX1* was revealed as susceptibility markers for KD (Burgner *et al.*, 2009). Reports from another short communication revealed *LNX1* and *LNX2* mRNAs were significantly elevated in chronic Q fever, an illness triggered by *Coxiella burnetii* infection (Mehraj *et al.*, 2012). Similarly, both Q fever and KD aetiologies develop because of an elevated immune response. LNX1 and LNX2 modulate the ubiquitination of the T cell receptor CD8 α (D'Agostino *et al.*, 2011), this fact and the relevance of LNX1/2 expression in Q fever and KD suggest a mechanism of LNX proteins and the immune system in terms of affecting the ubiquitin status of key immunological components.

1.3.7.2 LNX1/2 in colorectal cancer

There have been several reports of LNX1 and LNX2 having a direct or indirect role in the progression and therapeutic resistance of colorectal cancer (CRC), which I attempt to

summarise below (Camps *et al.*, 2013)(Ma *et al.*, 2017a) (Yan *et al.*, 2021). *LNK2* is mapped to chromosome 13q12.2 in humans, a region which is frequently amplified in CRC. Camps *et al.* (Camps *et al.*, 2013) examined this region of interest and found expression of *LNK2* mRNA is upregulated in CRC and CRC cell lines. The mRNA levels of *LNK2* correlated with protein expression allowing endogenous *LNK2* protein to be detected in CRC cell lines without immunoprecipitation. Knockdown of *LNK2* expression using siRNAs induced apoptosis in SW480 and DLD-1 cell lines. Upon *LNK2* knockdown in SW480 cells, 680 dysregulated genes were observed using genome-wide expression profiling, many of which are involved in the WNT/ β -catenin and NOTCH signalling pathways. TCF7L2 a transcription factor in WNT signalling and NOTCH ligand – JAG1 were both downregulated upon *LNK2* silencing (Camps *et al.*, 2013). Surprisingly, decreased NUMB protein levels were observed upon *LNK2* knockdown—the opposite of what one might expect for NUMB as a substrate for *LNK2* ubiquitination.

Dysregulation in the Hedgehog (Hh) signalling pathway has been implicated with tumorigenesis (Briscoe and Thérond, 2013). The tumour suppressor SuFu negatively regulates Hh signalling by restricting the nuclear translocation of the Gli transcription factor. SuFu negating protein 1 (SNEP1) is highly expressed in CRC and targets SuFu for degradation thus promoting Hh signalling. In this study the authors find that SuFu is a novel substrate of *LNK1* ubiquitination but is dependent on SNEP1. In HCT116 and CaCo2 cells, overexpression of SNEP1 led to increased degradation of SuFu and decreased sensitivity to vismodegib, a Hh pathway targeting drug (Yan *et al.*, 2021). This activity in CRC, reveals *LNK1* as oncogenic as it mediates degradation of a tumour suppressor.

On the other hand, Ma *et al.* (Ma *et al.*, 2017a) deduce that LNX1 has the opposite effect in CRC. The authors describe a downregulation of *LNX1* mRNA in cancer stem cells (CSCs), isolated from HT-29 cells. LNX1 knockdown in HT-29 cells increased the number of CSCs, allowed the formation of more colonospheres *in vitro* and more tumours in mice that were immunised with these HT-29, LNX1 knockdown cells. This suggested that LNX1 is a negative regulator of cancer stemness in CRC. Upon treatment with tamoxifen, *LNX1* gene expression was exponentially increased and colonosphere formation was halted (Ma *et al.*, 2017a). Overall, the role of LNX1 in colorectal cancer remains controversial, with evidence suggesting both tumour-suppressive functions through inhibition of cancer stemness and oncogenic activity by promoting SNEP1-dependent ubiquitination and degradation of the Hh pathway suppressor SuFu, thereby contributing to Hedgehog signalling and resistance to vismodegib. LNX2 is frequently overexpressed in colorectal cancer, where it promotes tumour cell survival by supporting WNT/ β -catenin and NOTCH signalling, and its knockdown induces apoptosis, consistent with an oncogenic role.

1.3.7.3 LNX1 involvement in brain tumours

In a subset of 15% of glioblastoma multiform (GBM) tumours, amplification of chromosomal region 4q12 is seen. It is important to note while *LNX1* is mapped to this region so are several other potentially cancer promoting receptor tyrosine kinase encoding genes (*c-KIT*, *PDGFR- α* and *VEGFR2*). This amplicon is also seen in 23% of GBM cell lines, examples of these cell lines are U-251MG and U-87MG (Holtkamp *et al.*, 2007). LNX1 has been tied to cell cycle regulation and cisplatin resistance in lung cancer (A549, H460) and breast cancer (MCF7) cell lines. The authors provide evidence that knocking

out *LNK1* in A549 cells delays cell cycle progression (Jang *et al.*, 2021). Baisiwala *et al.* demonstrate similar results in patient derived xenograft (PDX) models of glioblastoma and in U-251MG cells. Upon temozolomide treatment, *LNK1* expression (qPCR) and cell number increases as does symmetric stem cell divisions as compared to DMSO treatment (Baisiwala *et al.*, 2020). The authors suggest *LNK1* acts through the NUMB-NOTCH signalling pathway to exhibit these effects, by observing increased expression of NOTCH downstream markers – p21 and Hes1. Data to this point is suggestive of a *LNK1* role in the tumorigenesis of GBM, but more work is needed to confirm the observations are not circumstantial.

1.4 Overall summary

Overall, the current literature suggests that *LNK1* and *LNK2* participate in diverse biological processes, including neurodevelopment, cell signalling, immunity and cancer. However, despite their evolutionary conservation and their identification more than two decades ago, relatively little is known about the physiological functions, substrates and molecular mechanisms of *LNK* proteins compared with other RING-type E3 ubiquitin ligases. Much of the available evidence is derived from individual studies in specific cellular contexts, and in several cases, such as the role of *LNK1* in colorectal cancer, the findings are conflicting. Consequently, there remains a substantial gap in our understanding of how *LNK* proteins regulate cellular signalling and contribute to disease. Addressing these knowledge gaps is a central aim of this thesis, which seeks to identify and characterise novel functions and substrates of *LNK* proteins and to define the molecular mechanisms underlying their biological activity.

Chapter 2

Materials and Methods

2.1 Molecular biology

2.1.1 Plasmid construction

Desired vector backbones were digested using required restriction enzymes (Table 2.1) and purified using the Trans gel purification kit (EasyPure®). Inserts were either digested from existing plasmids amplified by PCR from appropriate cDNA or obtained constructs and digested using the complementary restriction enzymes and size was verified using agarose gel electrophoresis. Ligation of insert into vector was preformed using 1 µL vector backbone and 3 µL of purified insert and allowed to progress for 2 hrs at 16° C. Ligation reactions were transformed into DH5α *E. coli* cells and plated on selective antibiotic agar plates. Colonies were screened using colony PCR and plasmid DNA was purified using Trans DNA purification kit (EasyPure®). Confirmation was carried out by appropriate restriction enzyme digestion and size determining agarose gel electrophoresis and verified by Sanger sequencing (Eurofins).

Table 2.1: Vector backbones and restriction enzymes used in plasmid construction.
Kan=Kanamycin resistance, Amp=Ampicillin resistance

Relevant assay	Vector	Antibiotic selection	Restriction Enzymes	Source
Yeast-2 hybrid vectors	pLEX-MG	Amp	KpnI-BamHI BamHI-XhoI	P.Y (University College Cork)
	pGAD10	Amp	N/A	P.Y (University College Cork)
Mammalian cell expression	pCMV-GF	Kan	KpnI-BamHI BamHI-XhoI HindIII	P.Y (University College Cork)
	pCMV-N-HA	Kan	KpnI-BamHI BamHI-XhoI HindIII	P.Y (University College Cork)
	pCMV-IL2ss-N-HA	Kan	EcoRV- NheI	L.E (University College Cork)
	pCMV-N-Flag	Kan	XhoI-HindIII	P.Y (University College Cork)
	pEGFP-C2	Kan	XhoI-HindIII XhoI-EcoRI	P.Y (University College Cork)
	pEGFP-N1	Kan	XhoI-HindIII XhoI-EcoRI	P.Y (University College Cork)
LNx interacting proteins	HA-tagged BCR	Amp	N/A	(Hastie <i>et al.</i> , 2015)
	pCMV mouse sidekick2	Kan	XhoI-EcoRI	(Yamagata and Sanes, 2008)
	pCAG-HA-Nrxn1beta AS4(+)	Amp	N/A	(Chih, Gollan and Scheiffele, 2006)
	pCI-SEP GluR2(R)	Amp	XhoI-HindIII	(Kopec <i>et al.</i> , 2006)
Luciferase reporter assay	pGL4.18	Kan	NheI-HindIII	G.L (University College Cork)
	pDluc-Renilla	Amp	N/A	G.L (University College Cork)
Bacterial expression	pET24d-N-His-MBP-TEV	Kan	EcoRI-BamHI	P.Y (University College Cork)
Proximity labelling assay (BioE3)	Lenti-EFS-BirAopt-GSQ-RBXN-P2A-BLAST	Amp	BglII-XhoI- HindIII-EcoRI	(Barroso-Gomila <i>et al.</i> , 2023)
	TRIPZ-bioGEF-UBnc-PURO-addgene	Amp	XhoI-HindIII EcoRI-BglII	(Barroso-Gomila <i>et al.</i> , 2023)

2.1.2 Polymerase chain reaction (PCR)

Amplification reactions for molecular cloning were carried out using high fidelity Taq polymerases (PfuUltra II Fusion HotStart DNA Polymerase, Agilent #600670 or FIREPol® DNA Polymerase, Solis Biodyne #01-01-KIT-00500) according to the manufacturer's guidelines. Unless otherwise stated primer annealing temperatures were designed at 60 °C and 30 PCR cycles were performed. Primers are indicated in Table 2.

Table 2.2: Primers in 5' to 3' direction used for amplification of outlined inserts and required restriction enzyme for cloning.

Relevant assay	Primer name	Primer sequence
LNX protein related	hs-mm-LNX2-R-screen	gtggtgatctctccttctggtaaactaag
	hs-mm-LNX1-R-screen	gacacgtaggcagggttgctag
	hsLNX1p80-F-Sall	ttgtcgacatgaaccagccagagtctgcc
	hsLNX1-R-EcoRI	ttgaattcctctgacccatcattgattc
	hsLNX2-F-XhoI	ttctcgagcatgggaacaacaagtgtgatgatg
	hsLNX2-R-EcoRI	ttgaattcctatacaaggctgccaggcc
LNX interacting proteins	EphB2-FLAG-insert-F-SpeI	gactacaaagacgatgacgacaagactagtgtggaagaaaccctgatggactctacg
	EphB2-FLAG-insert-R-SpeI	cacactagtcttgcgtcatcgtctttagtccacggcggctagcagcgccag
	HSZNF24-F-XhoI	ttctcgagaatgtctgcacagtcatggg
	HSZNF24-R-HindIII	ttaagcttattaaactttcacaacattcagaag
	MmAmigo_F_XhoI	ttctcgagcatgcaacccagcgtgacc
	MmAmigo_R_EcoRI	ttgaattcatcacaccacaatgggtgtatc
	MmAmigo1-noSS-F-XhoI	ttctcgagaggccgatctgtggttagttg
	msSidekick-F-XhoI	gatcctcgagcgcccaagatgatgtaccgccatac
	msSidekick-R-EcoRI	tcgagaattctcaaacgaaggatgagaaccagc
	MmSidekick-Cyto-F-XhoI	ttctcgagaggtcagagcaagaagtactcc
	MmSidekick2-int-R-PY	tctcatgggtcacactttggc
	IL2ss-HA-BOT-EcoRV	atagggatatcctgcggaattagtgtgacaagtgcagacttagtgcaatgcaagacaggag
	IL2ss-HA-TOP-NheI	tccgctagccaccatgtacaggatgcaactcctgtcttgattgcactaagtcttgact
	EXOC4-F-XhoI	ttctcgagaatggcggcagaagcagctg
	EXOC4-R-EcoRI	ttgaattcactaaacggtagtattcttctgt
	FLOT2-F-XhoI	ttctcgagcatgggcaattgccacacgg
	FLOT2-R-EcoRI	ttgaattctcacacctgcacaccagt
	PKN2-F-Sall	ttgtcgaccatggcgccaaccccgaaac
	PKN2-R-EcoRII	ttgaattcttaacaccaatcagcaatgtagt
	RHOA-F-XhoI	ttctcgagaatggctgccatccggaaga
	RHOA-R-EcoRI	ttgaattcatcacaagacaaggcaccca
	CNP-R-MfeI	gtccaattgtcatatgatgtgtgcaggactgc
	CNP-F-Sall	gtcgtcgaccatgaacagaggcttctccgaaaa
	DCUN1D1-R-EcoRI	gtcgaattcctacactgttacttttgtccc
	DCUN1D1-F-XhoI	gtcctcgagcatgatcttcacacaatctagtgaag
	GNL3-R-HindIIIMfeI	gtccaattgaagctgttacacataatctgtactgaagtc
	GNL3-F-XhoI	gatcctcgagcatgacctgccataagcggataaaa
Proximity labelling (BioE3) related	bioGEF-Ub-F-XhoI-HindIII	gatctcgagctcaagcttgccaccatgggtttgaatgac
	BirA-F-Seq	ggaatcatcaagccctggatgg
	BirAlink-R-BglII-XhoI-HindII-EcoRI	gcagaattcgaagcttgagctcgagatctgcccttacttgcgtaagaaatttg
	BirA-FKoZ--NheI/Agel	tccgctagcgctaccggtcgccaccatgaaggacaacaccgtg
	bioGEF-Ub-R-EcoRI-BglII	ctcagatctgaattcttaaccaccttaggacggag
	hsLNX1 C44A-R	cagcaaagcctgcagagcgtatggcagatgag
	hsLNX1 C44A-F	ctcatctgccacatcgctctgcaggctttgtctg
	hsLNX2 C53A-R	ctgcagcagaggttgaagggaatatggcagactaggtc
	hsLNX2 C53A-F	gacctagtctgcatattgcccttcaacctctgctgcag
Evolution of LNX-	hsLNX1-PDZ4-RdelFL-EcoRI	gatagaattcttaagtgcaggccaagaaacaatag

NUMB interaction related	hsLNX2-PDZ4-R-delLV- EcoRI	gatcgaattcttagctgccaggccaacaaataac
	Bf-NUMB-F1	ctttcgctcagactaaacgagc
	Bf-LNX-F1	ctggtgtccatgattgtacagg
	Bf-Numb internal- F2	cttcatggccatcaaggattc
	Bf-LNX-R1	ctttcttcagctgaccgtttcc
	bf LNX PDZ1-R Bam Eco	tttggatccgaattcttactccaccgtaggtatcctgtac
	bf LNX PDZ1-F xhoI	tttctcgagcgttctggagggtgaatcacc
	Bf LNX PDZ4-R del TML eco bam	tttggatccgaattcttagcccgccaggacaccacagc
LNX promoters (Luciferase)	hsLNX1-201-prom-F-NheI	gatccgctagccaacatttatttctgaagg
	hsLNX1-201-prom-R- HindIII	gatcaagcttgggtgaaccgagcagctccttg
	hsLNX1-202-prom-F-NheI	gatcgctagcgttcatctgcactttaaactgtag
	hsLNX1-202-prom-R- HindIII	gatcaagcttatgtctctgattctcagatgag
	hsLNX2-prom-R-HindIII	gatcaagcttcgcggatttctggcaaaaagc
	hsLNX2prom-F-NheI	gatcgctagcctcaggtggtgctggcggtac
	hsLNX2-prom-pGL4.18-F- Ass	ggcctaactggccgttacctgagctcgctagcctcaggtggtgctggcggtac
	HsLNX2-prom-pGL4.18-R- Ass	ggctttaccaacagtaccggattgccaagcttcgcggatttctggcaaaaagc
	HsLNX2-Prom-IntR-BsmBI	gatccgtctcactgcaagctccccggtgaaac
	Hs-LNX2prom-intF-BsmBI	gatccgtctctgagttgcattttcaactggg
	hsLNX3-201-prom-F-NheI	gatcgctagctggatcgccagggtttcc
	hsLNX3-201-prom-R- HindIII	gatcaagcttcagacaggccggctacgc
	pGL4.18 seq-R	gtcttcgagtggttagaatggc
	pGL4.18 seq F	ataggctgtccccagtgcaag

2.1.3 Reverse transcriptase PCR (RT-PCR)

Indicated cell lines were harvested in TRIzol™ (Invitrogen) at 70% confluency and total RNA was extracted following the manufacturer guidelines. After RNA quality checks by agarose gel electrophoresis and quantification by nano-drop spectrophotometry, 0.3 µg was treated with DNaseI (Thermo Scientific™) and reverse transcribed into cDNA using RevertAid RT Reverse Transcription Kit (Thermo Scientific™). Negative controls for each sample were reactions that didn't include the reverse transcription enzyme. Real Time quantitative PCR was performed on StepOnePlus™ real-time PCR system (Life

Technologies) using PowerTrack™ SYBR Green Master Mix for qPCR and 400 nM primers. The thermocycling conditions were as follows: 10 minutes at 95 °C, 40 cycles of 15 seconds at 95 °C followed by 1 minute at 60 °C, and a melt curve stage of 0.3 °C increase of temperature every 15 seconds until the reach of 95 °C. Fluorescence was recorded during the 60 °C stage and the melting curve stage. Three biological replicates and two technical replicates were tested for each sample. The relative changes in gene (*LNX1/LNX2*) expression were compared to the expression of the housekeeping gene (*β-actin*) and presented using the $2^{(-\Delta Ct)}$ method (Livak and Schmittgen, 2001).

Table 2.3: RT-PCR primers and targets

Primer target	Primer sequence	Reference
LNX1	F-Exon4: 5' gaaggacaaagagcgggag 3'	(Yan <i>et al.</i> , 2021)
	R-Exon5: 5' actcgattgatcttgatgctgg 3'	
LNX1	F-Exon8: 5' tgagcccgaggagtcataa 3'	(Jang <i>et al.</i> , 2021)
	R-Exon9: 5' attccagccacatgacccag 3'	
LNX2	F-Exon2: 5' atgcaacgttgatctgga 3'	(Camps <i>et al.</i> , 2013)
	R-Exon3: 5' ccaaacagtctgcttctgga 3'	
β-actin	F-Exon3: 5' ccaaccgcgagaagatga 3'	(González-Bermúdez <i>et al.</i> , 2019)
	R-Exon4: 5' ccagaggcgtacagggatag 3'	

2.1.4 Annealing oligonucleotides

Oligonucleotides of approximately 30 base pairs in length were designed using Benchling and a restriction site was included at either 5' or 3' end to allow for cloning. Oligonucleotides were synthesised by Integrated DNA technologies (IDT). Pellets were dissolved in 1 X TE (10 mM Tris-HCl and 0.1 mM EDTA) buffer to 100 μM stocks. Oligonucleotides were phosphorylated using T4 polynucleotide kinase ligated using T4 ligase, heated to 95°C and allowed to cool slowly to room temperature over a few hours.

Table 2.4: Oligonucleotides in 5' to 3' direction, used for yeast-two hybrid analysis in the indicated chapters

Chapter 3	CTNNA2-10AA-TOP-KpnI-BAMHI	ctattaagtgaattcaaagcaatggattccttctagg
	CTNNA2-10AA-BOT-KpnI-BAMHI	gatccctagaaggaatccattgctttgaattcacttaataggtac
	CTNNB1-10AA-TOP-KpnI-BAMHI	ctaaatcagctggcctggttgatactgacctgtaag
	CTNNB1-10AA-BOT-KpnI-BAMHI	gaccttacaggtcagtatcaaaccaggccagctgatttaggtac
	CTNND2-10AA-TOP-KpnI-BAMHI	ctacactaccggcctccccgactcctgggtgtgag
	CTNND2-10AA-BOT-KpnI-BAMHI	gacctcacacccaggagtcggggaggccgggtagtgtaggtac
	RYK-10AA-TOP-KpnI-BAMHI	ctagagtttcagtcagccctggggcctacgtctgag
	RYK-10AA-BOT-KpnI-BAMHI	gacctcagacgtaggccccagggtgcatgaaactctaggtac
	SMG1-10AA-TOP-KpnI-BAMHI	ctacagctgtatgaaggttgacagcctgggtgtgag
	SMG1-10AA-BOT-KpnI-BAMHI	gacctcacacccaggctgtccaaccttcatacagctgtaggtac
	YES1-10AA-TOP-KpnI-BAMHI	ctagagccacagtaccagccaggagaaaatttataag
	YES1-10AA-BOT-KpnI-BAMHI	gaccttataaattttctcctggctgtactgtggctctaggtac
	ITCH-10AA-TOP-KpnI-BAMHI	ctaataagaagaacagaaggatttgacaagagtaag
	ITCH-10AA-BOT-KpnI-BAMHI	gaccttactctgtcctaaatccttctgttcttctattaggtac
	MARCHF9-10AA-TOP-KpnI-BAMHI	ctacgagaggtgtcatgagggtcactacagctctgag
	MARCHF9-10AA-BOT-KpnI-BAMHI	gacctcagactgtagtaccctcatgacaacctctcgtaggtac
	RNF146-10AA-TOP-KpnI-BAMHI	ctacctgatggacagtgcacagtaactgaagttaag
	RNF146-10AA-BOT-KpnI-BAMHI	gaccttaaacctcagttactgtgactgtccatcaggttaggtac
	IGSF11-10AA-TOP-KpnI-BAMHI	ctaccagcccagagtcgggcccgggtccttggtatagg
	IGSF11-10AA-BOT-KpnI-BAMHI	gacccctataccaaggacccggcccgactctgggtggttaggtac
	SDK2-10AA-TOP-KpnI-BAMHI	ctagctcccatagcaggattttcatcattgtttgag
	SDK2-10AA-BOT-KpnI-BAMHI	gacctcaaacaatgatgaaatcctgctatgggagctaggtac
	Nrxn1-10AA-TOP-KpnI-BamHI	ctaaagaaaaacaaggataaggagtattacgtctag
	Nrxn1-10AA-BOT-KpnI-BamHI	gaccttagacgtaataactccttatcctgtttttcttaggtac
	Amigo-10AA-TOP-KpnI-BamHI	ctatcggtcttttctgataccaccattgtggtgtag
	Amigo-10AA-BOT-KpnI-BamHI	gacctacaccacaatgggtgtatcagaaaagaccgataggtac
	Srr-10AA-TOP-KpnI-BamHI	ctacggccagctcctaccagacggtttctgtttag
	Srr-10AA-BOT-KpnI-BamHI	gacctaacagaaaccgtctggttaaggagctggccgtaggtac
	CYP4X1-10AA-TOP-KpnI-BAMHI	ctatatttgcacctgaagaaactctctgaattgttagg
	CYP4X1-10AA-BOT-KpnI-BAMHI	gacccctaacattcagagagtttcttcaggtgcaaataataggtac
	GRIA2-10AA-TOP-KpnI-BAMHI	ctaaacgtatatggcatcgaaagtgttaaaatttagg
	GRIA2-10AA-BOT-KpnI-BAMHI	gacccctaaattttaacactttcgtatgcatatatacgttaggtac

	ZNF24-10AA-TOP-KpnI-BamHI	ctagcagaaaaacttctgaatgttgtaaagtttag
	ZNF24-10AA-BOT-KpnI-BamHI	gacctaactttcacacattcagaagttttctgctaggtac
	PPFIA1-10AA-TOP-KpnI-BAMHI	ctagattctgctacagtcaggacttactctgctaag
	PPFIA1-10AA-BOT-KpnI-BAMHI	gaccttagcaggagtaagtcctgactgtacagaatctaggtac
	ORC1-10AA-TOP-KpnI-BAMHI	ctattttcatttctgcttttgtaagccattgctgag
	ORC1-10AA-BOT-KpnI-BAMHI	gacctcagcaatggcttacaaaaagcagaaatgaaataggtac
	NDRG3-10AA-BOT-KpnI-BamHI	gaccttagcaagacacttccatggctggtgtgtctaggtac
	NDRG3-10AA-TOP-KpnI-BamHI	ctagacagacaccagaccatggaagtgtcttgctag
	Prkar1a-10AA-TOP-KpnI-BamHI	ctacagtacaacagcttcgtgtccctgtccgtctag
	Prkar1a-10AA-BOT-KpnI-BamHI	gaccttagacggacaggacacgaagctgtgtactgtaggtac
	BCR-10AA-TOP-KpnI-BAMHI	ctaagacagagcatcctgttctccaccgaagtctaag
	BCR-10AA-BOT-KpnI-BAMHI	gaccttagacttcggtggagaacaggatgctctgtcttaggtac
	DVL2-10AA-TOP-KpnI-BAMHI	ctaaatcccagcgagttcttgtggatgttatgtagg
	DVL2-10AA-BOT-KpnI-BAMHI	gacccctacataacatccacaagaactcgctgggatttaggtac
	KRTAP4-3-10AA-TOP-KpnI-BAMHI	ctacaccccatctgtcgggcagttcttctgctgtag
	KRTAP4-3-10AA-BOT-KpnI-BAMHI	gacctcagcagcaagaactgccgcagcagatgggggttaggtac
Chapter 4	ABRAXAS2 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	ABRAXAS2 10AACT Top XhoI	-tcgagcgaccccaggaacactcagacctccagatttaag
	CNP 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	CNP 10AACT TOP XhoI	tcgagggggggcgcttgcagtcctgcacatcatatgag
	DCUN1D1 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	DCUN1D1 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	EXOC4 10AACT BOT BamHI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	EXOC4 10AACT TOP XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	FERMT2 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	FERMT2 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	FLOT2 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	FLOT2 10AACT TOP XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	GNL3 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	GNL3 10AACT TOP XhoI	-tcgagcctacactgtgtacttttgtcccagcaatttggc
	HEATR1 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	HEATR1 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	KARS1 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	KARS1 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	KPNA1 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	KPNA1 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	PAFAH1B3 10AACT Bot BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	PAFAH1B3 10AACT Top XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	PKN2 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	PKN2 10AACT TOP XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	PKN2 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	PKN2 10AACT TOP XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc
	PKN2 10AACT BOT BamHI	gacctaataatctgggaggtctgagtggtcctgggggtcgc
	PKN2 10AACT TOP XhoI	tcgagcctacactgtgtacttttgtcccagcaatttggc

	PKN2 10AACT Top XhoI	tcgagcagagattttgactacattgctgattggtgtaag
	RAB18 10AACT BOT BamHI	gatccttataacacagagcaataaccaccacaggctccgc
	RAB18 10AACT Top XhoI	tcgagcggagcctgtggtggtattgctctgtgtataag
	RHOA 10AACT BOT BamHI	-gatcctcacaagacaaggcaccagatttttcttccgc
	RHOA 10AACT Top XhoI	tcgagcgggaagaaaaaatctgggtgccttgcttctgtgag
	TOM1 10AACT Bot BamHI	gatcctcataaggcaaacagcatgtcatcatcttctcgc
	TOM1 10AACT Top XhoI	tcgagcgagaaagatgatgacatgctgtttgccttatgag
	USO1 10AACT Bot BamHI	gatccctagatatgatctagatccttgccaggatcttcgc
	USO1 10AACT Top XhoI	tcgagcgaagatcctggcaaggatctagatcatatctagg
Chapter 5	msLNX1-10AA-BOT-KpnBam	gatccttataaaaaagtaccaggccaagaagcaatggtgcctccgctgtac
	msLNX1-10AA-TOP-KpnBam	agcggaggcaccattgcttcttgccctggtactttttataag
	msLNX2-10AA-BOT-KpnBam	ggatccctatacaggctgcctggccagcagatgaccgtgcctccgctgtac
	msLNX2-10AA-TOP-KpnBam	agcggaggcacggtcatctgctggccaggcagcctcgtatagg
	bfLNX-10AA-BOT-KpnBam	gatccttacagcatagtgccggccaggacaccacagcgctccgctgtac
	bfLNX-10AA-TOP-KpnBam	agcggaggcgctgtggtgtcctggccgggcactatgctgtaag
	bfNumb-10AA-BOT-KpnBam	gatccttataattgcaattcaaaagtcttttctactgcgcctccgctgtac
	bfNumb-10AA-TOP-KpnBam	agcggaggcgcgtagataaaagacttttgaattgcaattataag

2.2 Mammalian cell culture and DNA transfections

2.2.1 Mammalian cell culturing conditions

Cell culture was performed under sterile conditions using biological safety class II vertical laminar flow cabinets. All the cells were maintained at 100% humidity, 37°C and 5% CO₂. Sub cultivation of all cells was carried out when dishes were 70-90% confluent. Dishes were gently washed once with warm 1 X PBS to remove old culture medium. Trypsin 0.25% w/v solution (Merck) (2.5 g/L porcine trypsin and 0.2 g/L EDTA in Hank's Balanced Salt Solution with phenol red) was used to detach cells. Cells were plated again to the desired concentration. Cells were regularly tested by PCR for the presence of *Mycoplasma*.

2.2.2 HEK 293T and MCF-7 cell lines

Human Embryonic Kidney 293T (CRL-3216) and MCF-7 (HTB-22) cells were originally purchased from the American Type Culture Collection (ATCC) and were kindly received from Prof. Rosemary O' Connor (University College Cork, Ireland). Both MCF-7 and HEK 293T cells were maintained in Dulbecco's Modified Eagle Media (Sigma-Aldrich) supplemented with 10% foetal bovine serum (FBS) and antibiotics 1 X penicillin/streptomycin (P/S) (100 units/mL penicillin and 50 µg/mL streptomycin).

2.2.3 Colorectal carcinoma cell lines

All cell lines listed below were originally purchased from the American Type Culture Collection (ATCC) and were kindly received from Dr. Ken Nally (University College Cork, Ireland)

HT-29 (ATCC HTB-38) and HCT-116 (ATCC CCL-247) cells were maintained in McCoy's 5A Medium (Sigma #M9309-500ML) supplemented with 10% FBS, 1X P/S.

DLD-1 (ATCC CCL-221) cells were maintained in RPMI-1640 (Sigma #R8758-500ML), 2 mM L-glutamine, 10mM HEPES, 1mM sodium pyruvate, 10% FBS and 1X P/S.

Colo-205 (ATCC CCL-222) cells have a mixed, adherent and suspension type attachment. Regular cell culture dishes or attachment plus cell culture dishes can be used. Cells were maintained in RPMI-1640, 2 mM L-glutamine, 2g/L sodium bicarbonate (Sigma #R8758-500ML), 10 mM HEPES (Gibco #15630056), 1 mM sodium pyruvate (Sigma #S8636-100ML), 10% FBS and 1X P/S. When harvesting and passaging these cells, media was first collected followed by the attached cell layer to ensure no cells were discarded.

HCT-15 (ATCC CCL-225) cells were cultured in RPMI-1640, 2 mM L-glutamine, 1 mM sodium pyruvate, 10% FBS and 1X P/S.

RKO (ATCC CCL-2577) and LS-174T (ATCC CL-188) cells were maintained and cultured in MEM α , GlutaMAX™, no nucleosides (Gibco #32561029) 10% FBS and 1XP/S.

SW-480 (ATCC CCL-228) and SW-48 (ATCC CCL-231) cell lines were maintained in Leibovitz's L-15 Medium with 2 mM L-glutamine (Gibco #11415049) 10% FBS and 1X P/S.

2.2.4 Glioblastoma and Neuroblastoma cell lines

U-118MG (Cyton / Cell Lines Services catalogue number #300362), U-251MG (Cyton / Cell Lines Services catalogue number #300385) and U-87MG (Cyton / Cell Lines Services catalogue number #300367) cells were originally obtained from Cyton formally known as Cell Lines Services and then generously gifted by Prof. Justin McCarthy (University College Cork, Ireland).

These Glioblastoma cells were cultured as a monolayer in Eagles Minimum Essential Medium (Sigma #M5650) supplemented with 2 mM L-glutamine, 0.1 mM nonessential amino acids (Sigma #S8636-100mL), 1 mM sodium pyruvate, 10% FBS and 1X P/S.

U-87MG (Cyton / Cell Lines Services catalogue number #300367) and A172 (Cyton / Cell Lines Services catalogue number #300108) cells were all sourced from Prof. Justin McCarthy (University College Cork, Ireland). SH-SY5Y (Cyton / Cell Lines Services catalogue number #300154) cells were kindly gifted from Prof. Cora O'Neill (University College Cork, Ireland). These cell lines were cultured in Dulbecco's Modified Eagle Media

(Sigma-Aldrich) supplemented with 2 mM L-glutamine, 10% foetal bovine serum (FBS) and antibiotics 1 X penicillin/streptomycin.

2.2.5 Cell pellets and lysates

Cell pellets and/or cell lysates from LoVo (ATCC CCL-229) and SW-948 (ATCC CCL-237) cells were also gifted from Dr. Ken Nally's lab. While Calu-3 (HTB-55), Caco-2 (HTB-37) and Huh-7 (Cytion / Cell Lines Services catalogue #300156) cell lysates were gifts from Mr. Jame Harte. U2-OS (HTB-96) cell pellets were available in the Young lab.

2.2.6 Maintenance of established cell lines

All cells were cultured as a monolayer in their respective maintenance media. To prevent overgrowth and loss of surface contact with the culture dish, cells were routinely passaged at approximately 80–90% confluence. For passaging, cells were gently washed with 1 X PBS and detached using trypsin, a volume of maintenance media four times that of the trypsin was added to inactivate the enzyme. The cell suspension was then gently resuspended to obtain a single-cell solution. Cells were counted using a hemacytometer and plated at the desired density.

2.2.7 Cryopreservation and recovery of cell lines

Cells from an 80-90% confluent 100mm culture dish were collected as described in section 2.2.6. The cell suspension was then centrifuged at 1,000 g for 5 minutes, resuspended in their maintenance media supplemented with 10% of dimethyl sulphoxide (DMSO) to a concentration of 5×10^6 / mL and 1 mL was aliquoted per cryopreservation vial. Vials were stored at -80°C for short term storage or liquid nitrogen for long term storage.

For cell line recovery, the cryopreservation vial was thawed at 37°C, and the cell suspension was added dropwise to a 100 mm culture dish containing warm maintenance media. The following day, the media was replaced with fresh media. Prior to plating the cells for experiments, they were passaged at least once.

2.2.8 Seeding and transfection of cell lines

As shown in table 2.5, transfection was performed in different size wells using the calcium phosphate precipitation method (calcium chloride (CaCl₂) and HEPES buffered saline (HBS)) (Graham and van der Eb, 1973) or using Lipofectamine 2000 with cells at approximately 70-80% confluence. Lipofectamine 2000 transfections were carried out as per the manufacturers' recommendations or otherwise indicated (ThermoFisher Scientific). Media was changed 6-14 hours post-transfection. Cells were harvested after 48 hours for lysate preparation or fixed with 4% paraformaldehyde (PFA) in PBS (pH 7.5) for immunocytochemistry.

Table 2.5: Calcium phosphate transfection method and concentration for indicated well sizes.

Plate/well size	DNA (µg)	2 M CaCl ₂ (µL)	H ₂ O (µL)	2 X HBS (µL)
12 well	1	6	44	50
6 well	2	12	86	100
10cm ² dish	8	64	446	500
15cm ² dish	16	128	872	1000

2.3 Immunocytochemistry and cell lysate preparation

2.3.1 Immunocytochemistry

HeLa cells were seeded onto coated 12mm coverslips at a cell density of 50,000 in 12 well tissue culture dish and allowed to adhere overnight, prior to transfection. 24 h post-

transfection cells were gently washed once in PBS and were fixed with 4% paraformaldehyde (PFA) for 10 mins at room temperature. After fixation, cells were washed three times in PBS, permeabilised and blocked in blocking solution (PBS containing 0.2% Triton X-100, 5 % normal goat serum and 2% bovine serum albumin) for 1 hr at room temperature. Coverslips were then incubated with the indicated primary antibody for 3 hr at room temperature, washed in PBS three times and incubated with the appropriate secondary antibody for 1 hr, in the dark, at room temperature. For the last 10 minutes of secondary antibody, DAPI stain was added at a concentration of 1 ug/mL. All antibodies were diluted in blocking solution lacking Triton X-100. Following three washes with PBS for 5 min each, the coverslips were mounted onto glass slides using Fluoromount™ (Merck / Sigma-Aldrich) and dried overnight before imaging. Cells were imaged using an EVOS FL Auto microscope (Life Technologies Corporation, Bothell, WA)

2.3.2 Cell lysis

Cells were washed twice and then harvested with ice-cold PBS pH 7.5. Cells were pelleted for 3 minutes 5,000 g and then lysed in lysis buffer (10 mM TrisHCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP-40 and 1x cOmplete protease inhibitors (Roche Applied Sciences)) for 30 minutes under agitation. Post incubation, lysates were cleared by 30 minutes centrifugation at 16,000 g at 4 °C.

2.3.3 Tissue homogenisation and lysis

Mice brains were weighted and gently homogenised on ice using a Dounce homogeniser in lysis buffer (20 mM, Tris pH 7.5, 10 mM NaCl, 1% NP40, 0.1% sodium deoxycholate,

1mM EDTA and 1× Complete protease inhibitors (Roche Applied Sciences)) in a volume 9 times the weight of the tissue. Samples were then incubated for 30 minutes at 4 °C under agitation and then centrifuged at 16,000 g for 30 minutes at 4 °C and pellet was discarded.

2.3.4 Determination of protein concentration

The protein concentration of cells and brain lysates was obtained using the Pierce™ bicinchoninic acid assay (BCA) Protein Assay Kit (Thermo-Fisher Scientific), a colorimetric detection and quantification protocol, as per manufacturer's instructions. A protein standard curved was generated using BSA at concentrations of 2000, 1500, 1000, 750, 500, 250, 125 and 25 µg/mL. 100 µL of BCA working reagent was added to 5 µL of cell lysate or standard, in triplicate in a 96-well plate. The plate was then mixed using an orbital shaker before being incubated at 37 °C for 30 min in the dark. The absorbance was measured at 562 nm on a microplate reader (Tecan Group Ltd., Männedorf, Switzerland).

2.4 Luciferase reporter assay

Cells were cultured and maintained in DMEM lacking Phenol red indicator and antibiotics and supplemented with 10% FBS. For each transfection, 25 µl Opti-Mem medium, 0.2 µl Lipofectamine 2000 (Invitrogen) plus 12.5 ng pGL4.18-RLuc and 12.5 ng pGL4.18-Fluc plasmids were added to clear bottom – full area 96-well dishes. DNA complexes were overlayed with 5×10^4 cells in 50 µL DMEM w/o Phenol red + 10% FBS at 37°C in 5% CO₂ for 24hrs. Un-transfected cells were used to account for background luminescence.

If cells were subjected to drug treatments of phorbol 12-myristate 13-acetate (PMA) (10 nM, 50 nM) or 4-hydroxy tamoxifen (4-OH-T) (0.1 μ M, 0.5 μ M, 1 μ M,) or vehicle controls, media was changed and supplemented 24 hours post transfection and allowed to progress for indicated time. If the effect of serum starvation was being analysed, media was replaced 24 hours after transfection with 0.1% or 10% FBS (control) and analysed after 18 hours.

Media was aspirated and cells were lysed in 20 μ L passive lysis buffer (40 mM Tris-HCl pH 7, 25 mM NaCl, 1 mM $MgCl_2$, 0.25% (v/v) Triton-X) and incubated at RT on an orbital shaker at 500 g for 30 minutes. Lysate was transferred to white opaque 96-well plates and assayed for Firefly and Renilla luciferase activities using a Veritas Microplate Luminometer (Turner Biosystems) and reagents as described by Baker and Boyce, 2014 (Baker and Boyce, 2014).

The relative level of reporter gene expression was calculated as the ratio of firefly luciferase activity to co-transfected control plasmid expressing Renilla luciferase.

2.5 Protein expression, purification and immunoprecipitation

2.5.1 Expression of LNX fusion proteins

Expression constructs encoding LNX1^{p80} and LNX2 described in section 2.1 were transformed into *Escherichia coli* BL21 cells. Cells were grown overnight at 37°C before being diluted and allowed to reach OD₆₀₀ - 0.6. Trial, small scale protein expression was induced at a range of temperatures from 18-30°C, and 18°C provided the most soluble proteins. Induction was initiated by adding of 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and cells were harvested 18 h post induction. Cells pellets

were resuspended in PBS, 0.2% Triton X-100, 20 mM β -mercaptoethanol, 500 mM NaCl and 1 mM PMSF. Cells were lysed by addition of 0.1 mg/ml lysozyme for 30 min on ice, followed by 0.1 mg/ml DNase I for a further 10 min on ice, before finally sonicating. Lysates were cleared by centrifugation at 10,200 g for 30 min at 4°C.

2.5.2 Nickel-Nitrilotriacetic Acid (Ni-NTA) Purification

For His tag purifications, the cleared supernatant was loaded on 500 μ L Ni-NTA bead columns (Cytiva). To equilibrate the columns, typically 3 x 10 mL washes were performed using equilibration/wash buffer containing 20 mM Tris pH 8, 500 mM NaCl, and 5 mM Imidazole. The equilibrated beads were incubated with the supernatant for 15 minutes to allow binding. 3 x 10 mL washes using the equilibration/wash buffer and one 20 mM imidazole wash/elution was typically performed. Elution was performed using 5 mL wash buffer supplemented with 200 mM imidazole and the elutant was re loaded onto the column 3 times to ensure total elution.

2.5.3 Amylose pull down of MBP fusion proteins

50 μ L of amylose beads (NEB) were used for most small-scale MBP purifications. The beads were first equilibrated by washing 3 times with wash buffer which consisted of 20 mM Tris HCl pH 7.5 and 150 mM NaCl. All centrifugation steps were performed at 1,000 g for 1 minute. The beads were typically incubated with the sample to be purified for 30 mins to 1 hour at 4 °C with head-over-heels rotation. If co-immunoprecipitation was the goal, beads were centrifuged and supernatant discarded. Cleared brain or cell lysate was then added to the prebound beads and incubated for a further 1 hour. The beads were washed 3 x 1 mL with wash buffer and elutions were performed by incubating the beads

for 5 minutes with elution buffer (10 mM Maltose, 20 mM Tris pH 7.5, and 150 mM NaCl) or for western blot analysis – boiling for 5 mins in 2X Laemmli buffer (final loading concentration: 50 mM Tris-HCl pH 6.8, 2% (v/v) SDS, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 5% (v/v) β -mercaptoethanol).

2.5.4 Dialysis of purified proteins

To remove high concentrations of imidazole and NaCl, purified and eluted proteins were loaded into SpectraPor dialysis membrane with a 10 kDa molecular weight cut off to allow buffer exchange. The proteins were floated on top of 1 L dialysis buffer per 5mL of protein eluant (5 mM β -mercaptoethanol, 20 mM Tris-HCl pH 7.5 and 150 mM NaCl) and incubated overnight at 4°C while constantly stirred.

2.5.5 Pull down of LNX interacting proteins using GFP-Trap-M beads

Confluent HEK 293T cells, that were transfected with one (negative control) or two plasmids were harvested using phosphate buffered saline (PBS) and pelleted at 1000rpm for 5 minutes. Cells were lysed using the ChromoTek lysis buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % Nonidet™ P40 Substitute, 1X cOmplete protease inhibitor (Roche Applied Science)). Lysates were diluted 4-fold with lysis buffer lacking NP-40 and 10 μ L of GFP-Trap Agarose beads were added to the lysate and incubated for 2 hours at 4°C with head over heels rotation. Beads were washed 5 times at 4°C with 1mL wash buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.05 % Nonidet™ P40 Substitute, 0.5 mM EDTA). Proteins were eluted from the beads by boiling for 5 minutes in 2X Laemmli Buffer.

2.5.6 Ubiquitination assays

Twenty-four hours post transfection; cells were incubated in fresh medium containing 10 μ M MG132 (Merck Millipore). Cells were then washed twice with ice-cold PBS and pelleted by centrifugation at 3,000 g for 5 min at 4 °C. Pellets were resuspended in 100 μ l 1% (w/v) Sodium Dodecyl Sulfate (SDS) and bursts with 15mM N-ethylmaleimide (NEM) and 1 $\times$ cOmplete protease inhibitors (Roche Applied Sciences), sonicated at 20 kHz for 2 x 5 sec and boiled for 5min. Following cooling on ice, samples were diluted with 900 μ l of cold buffer that contained 50 mM Tris (pH 8.0), 150 mM NaCl, 1% Triton X, 0.5% sodium deoxycholate, 50 mM sodium fluoride, 0.1 mM sodium orthovanadate and 1 $\times$ cOmplete protease inhibitors (Roche Applied Science). Lysates were cleared by centrifugation at 14,000 g for 30min at 4 °C. The supernatant was collected and protein concentrations were determined. Equivalent amounts (800 μ g) of lysates were immunoprecipitated overnight at 4 °C, with rotation, using 0.5 μ g of the indicated antibody, followed by incubation with 10 μ L Magnetic Protein A/G agarose beads (Thermo Scientific Pierce) for 4 h. Protein A/G agarose beads were washed 3 x 5mins in the dilution buffer. Proteins were eluted by boiling the Protein A/G agarose in 2X Laemmli sample buffer for 5 min and then analysed by western blotting.

2.6 SDS-PAGE and immunoblotting

2.6.1 SDS-PAGE

Sodium dodecyl sulphate–polyacrylamide gels of required acrylamide concentrations were made in house. After protein concentration determination, samples were prepared by adding Laemmli buffer to the samples and then heated at 95°C for 5 minutes to

denature the proteins. Equal protein content was loaded on SDS-PAGE gels and electrophoresed in running buffer (25 mM Tris, 192 mM glycine, 0.1% (v/v) SDS). After electrophoresis SDS-PAGE gels underwent either Coomassie blue staining (section 2.6.2) or Western blotting (section 2.6.3).

2.6.2 Coomassie blue staining

SDS gels were washed with MilliQ purified H₂O and stained with Coomassie blue stain containing 0.1% Coomassie, 45% (v/v) methanol, 10% (v/v) acetic acid at room temperature for 30 minutes while gentle shaking. Background binding for the stain was removed with destaining solution consisting of 5% (v/v) methanol, 7.5% (v/v) acetic acid, leaving the gels in solution until clear bands were shown and imaged using a BioRad Gel-doc.

2.6.3 Western immunoblotting

SDS-PAGE gels were transferred to Immobilon® membrane (Millipore) by electroblotting at 100 V in cold transfer buffer 25 mM Tris, 192 mM glycine, 20% (v/v) methanol for 70 minutes. The membranes were then incubated in 4% (w/v) fat free milk powder dissolved in TBS-T (10 mM Tris, 150 mM NaCl, 0.05% (v/v) Tween 20, pH 7.4) for an hour at room temperature to block unspecific binding of the antibodies. Primary antibody incubation was performed at 4°C overnight or 2 hours at room temperature in a humidified chamber. The membranes were washed three times with TBS-T for 5 minutes each under bland agitation and then incubated for 1 hour (or 2 hours in case of HRP-conjugated antibodies) at room temperature under bland agitation. Membranes were then washed three times with TBS-T for 5 minutes each under bland agitation. Primary and secondary

antibodies were diluted in 4% (w/v) fat free milk powder dissolved in TBS-T. The membranes were visualised using the Odyssey infra-red imaging system (LI-COR Biosciences, Cambridge, UK), where a fluorescent secondary antibody was used, or by enhanced chemiluminescence (ECL), where HRP-conjugated secondary antibodies were used, following the manufacturer's instructions (Thermo Scientific / Pierce).

Table 2.6: Application and required antibodies.

	Supplier and product code	Antibody
Immunocytochemistry	Sigma #F3165	Mouse anti-Flag M2
	Covance #MMS-101R	Mouse anti-HA
	Jackson Laboratories, INC., 711-545-152	Alexa Fluor® 488 anti-rabbit IgG
	Jackson Laboratories, INC., 715-545-150	Alexa Fluor® 488 anti-mouse IgG
Flow Cytometry	Bioledgend #901509	Alexa Fluor 488 mouse IgG1a anti-HA
	Bioledgend #637323	PE-Cyanine7 rat IgG2a anti-Flag (DYKDDDDK)
	Bioledgend #626810	Alexa Fluor 647 mouse IgG1a anti-c-Myc
	Jackson ImmunoResearch	Alexa Fluor 488 conjugated donkey anti-rabbit
Western blot primary	Abcam #ab290	Rabbit anti-GFP
	Sigma #F3165	Mouse anti-Flag M2
	New England Biolabs #E8032S	Mouse anti-MBP
	Sigma #A5441	Mouse anti- β Actin
	Covance #MMS-101R	Mouse anti-HA
	Sigma #V9131	Mouse anti-Vinculin
	R&D Systems #AF467-SP	Goat anti-EPHB2
	Abcam #ab14140	Rabbit anti-NUMB
	Novus Biologicals NB500-178 (Young <i>et al.</i> , 2005)	Rabbit anti-NUMB all isoforms Rabbit anti-LNX2 8B
	Wuxi Abgent #AP13071b (Young <i>et al.</i> , 2005)	Rabbit anti-LNX2 Guinea pig anti-LNX A
	(Dho <i>et al.</i> , 1998)	Guinea pig anti-LNX NPAY
	Abcam #ab76716	Mouse anti-LNX1
	Abcam #ab140793	Mouse anti-LNX1 (1H9)
	Genscript #A00674	Rabbit anti-Avi tag
	Genscript #A02252	Rabbit anti-Biotin-HRP
	Abcam #ab26192	Chicken anti-Liprin α 1

Western blot secondary	Invitrogen #A11374	AlexaFluor 790 donkey anti rabbit IgG
	Invitrogen #A11371	AlexaFluor 790 donkey anti mouse IgG
	Invitrogen #A10043	AlexaFluor 680 donkey anti rabbit IgG
	Invitrogen #A10038	AlexaFluor 680 donkey anti mouse IgG
	LiCor #926-68075	IRDye 680 Donkey anti-chicken
	Sigma #A5420	HRP Rabbit anti-goat IgG
	eBioscience #18-8817	HRP TrueBlot ULTRA anti-mouse IgG
	Jackson ImmunoResearch #771-035-152	HRP donkey anti-rabbit IgG
	Jackson ImmunoResearch #106-035-003	HRP goat anti-guinea pig IgG
	Rockland #18-8816-31	HRP TrueBlot mouse anti Rabbit IgG
	Rockland #206-1150	HRP Goat anti-guinea pig IgG

2.7 Yeast-based protein interactions assays

2.7.1 Transformation of yeast and yeast-2-hybrid analysis

1 mL of an overnight yeast culture per co-transformation was centrifuged at 12,000 g for 30 seconds. 0.5 µg of each plasmid was added to 10mg/mL boiled and cooled single stranded carrier DNA (Roche) and mixed with the pelleted yeast cells. 500 µL of 100 mM LiAc, 1X TE buffer, 40% (v/v) Polyethylene glycol 4000 was added to the pellet and incubated at room temperature for 15 minutes with occasional mixing. DMSO 10% (v/v) was added and immediately incubated at 42°C for a further 15 minutes. Transformants were plated on appropriate synthetic dropout (SD) (complete supplement mix - Sigma) medium.

2.7.2 Yeast-two hybrid analysis

Saccharomyces cerevisiae (L40 strain (*MATa his3Δ200 trp1–901, leu2–3, -112 ade2 LYS2::(lexAop)₄-HIS3 URA3::(lexAop)₈-lacZ*) was transformed with bait (pLexMG – *Trp1* selectable marker) and prey (pGAD10 – *Leu2* selectable marker) plasmids, identified in section 3.2.1.

Transformants were plated on SD media lacking Leucine, Tryptophan, Histidine and Uracil, 2% (w/v) dextrose, 0.7% (w/v) Yeast nitrogen base) media supplemented with 45 mg/L L-Histidine and 200 mg/L Uracil (Thermo Scientific) and incubated at 30°C for 3 days. To analyse activation of *His3* gene, selected colonies were then re-streaked on SD media lacking L-Histidine.

To confirm interaction and investigate if there is activation of the *LacZ* reporter gene, a β -galactosidase filter lift assay was performed. The colonies were transferred to Hybond-N membrane (Amersham Biosciences GE Healthcare) and lysed twice using liquid nitrogen. Z-buffer (130 mM Na₂HPO₄/NaH₂PO₄, pH 7, 10 mM KCl, 1mM MgSO₄·7H₂O), 0.0001% β -mercaptoethanol and 0.001M X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactospyranoside) were added to the membrane and incubated at 37°C for 2-6 hours.

2.7.3 Protein interaction analysis using yeast GFP-nanobody display and flow cytometry evaluation

Saccharomyces cerevisiae strain BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) was transformed with pYSD069 (MoClo YSD toolkit modified to display an anti-GFP nanobody) (Juric *et al.*, 2026 manuscript under review) plasmid and plated on minimal SD media lacking leucine (Uracil, 2% (w/v) dextrose, 0.7% (w/v) Yeast nitrogen base

media supplemented with 45 mg/L L-Histidine and 200 mg/L Uracil (Thermo Scientific)). 2 mL yeast cultures were grown in 13 mL tubes, in selective synthetic media lacking leucine at 30 °C overnight, shaking at 200 g. Yeast cell density was measured. Approximately 1×10^6 cells (assuming $1 \text{ OD}_{600} \approx 1.5 \times 10^7$ cells/mL) were pipetted into each microcentrifuge tube and centrifuged for 3 minutes at 3500 g. The supernatant was removed, and cells were resuspended in 100 μ L of flow cytometry buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.1% (w/v) bovine serum albumin, 5 mM maltose). 30 μ L of mammalian cell lysate which was expressing GFP-LNX1/2 and the protein interactor of interest, was then added and cells incubated for 30 minutes on ice. Then 0.5 μ g of anti-epitope tag antibody, as indicated, when using only antibodies or 4 °C when using cell lysates. After incubation, yeast cells were washed once using 500 μ L of flow cytometry buffer and pelleted. A further 500 μ L of flow cytometry buffer was used to resuspend the yeast and analysed by flow cytometry on an Accuri C6 instrument (BD Biosciences). Analysis of flow cytometry data was performed using Floreada (<https://floreada.io>).

2.8 Identification of LNX2 ubiquitination substrates *via* proximity labelling

This method was modified from the BioE3 (Barroso-Gomila *et al.*, 2023) and Ub-POD (Mukhopadhyay *et al.*, 2024) ubiquitin-specific proximity biotin labelling methods.

2.8.1 Cloning, transient transfection and biotin labelling of LNX-BirA substrates

To utilize transient transfections rather than creation of stable cell lines the TRIPZ-bioGEF-UBnc-PURO and Lenti-EFS-BirAopt-GSQ-RBXN-P2A-BLAST which were gifts from Rosa Barrio & James Sutherland (Addgene plasmid #208044 and #208048, respectively) were modified as follows. BioGEF-UBnc was digested out of the original TRIPZ-bioGEF-UBnc-

PUR0 using XhoI and EcoRI and ligated into pCMV-GF backbone (available in Young lab). Forward primer- 5' tccgctagcgctaccggctcgccaccatgaaggacaacaccgtg 3' and reverse primer - 5' gcagaattcgaagcttgagctcgagatctgccctctacttgcgtaagaaattg 3' were used to amplify the BirA-GSQ link insert from the Lenti-EFS-BirAopt-GSQ-RBXN-P2A-BLAST vector. BirA-GSQ link was then inserted into pCMV-GF and the human WT LNX1/LNX2 and RING mutants LNX1 C48A and LNX2 C53A were cloned C-terminally to the BirA tag. Confirmed by Sanger sequencing (Eurofins).

Before transfection, cells were seeded in desired dish size (6 well for trials to 15cm² for large scale experiments to analyse by mass spectroscopy) in fully supplemented DMEM (Section 2.2.6). Cells were allowed to grow overnight, and media was changed 6 hours before transfection to remove endogenous biotin (DMEM supplemented with 10% FBS was dialysed in Pur-A-lyzer tubes MWCO 3.5 kDa (Merck #PURG35020) to remove endogenous biotin). HEK 293T cells were co-transfected following section 2.2.8 protocol. After 24 hours, media was replaced again with DMEM supplemented with 50 µM biotin (Merck) and 10 µM MG132 for 6 hours to allow biotin labelling of substrates.

2.8.2 Neutravidin pull down of biotinylated proteins

Cells are rinsed twice in ice cold 1X PBS to remove waste media and excess biotin and pelleted for 5 minutes at 1500 g. Cell pellets are lysed in RIPA buffer (50 mM Tris-HCl pH8, 150 mM NaCl, 5 mM EDTA, 0.5% (w/v) Sodium deoxycholate, 0.1%(v/v) SDS, 1% Triton X-100, 1xPI tablet), sonicated at 20 kHz for 3 x 5 sec and cleared by centrifugation at 4°C for 30 mins at 16,000 g. Protein concentrations were determined and equivalent amounts of lysate were brought forward to neutravidin pull down.

Special care was taken to avoid Keratin contamination from this point on. All wash buffers were made fresh immediately before use.

Lysates were diluted by a factor of 4 in binding buffer (RIPA buffer without Triton X100). 50 μ L of Pierce NeutrAvidin Agarose beads (Fisher Scientific) per condition were equilibrated in binding buffer, added to the lysate and allowed to incubate for 4 hours with constant rotation. Beads were removed by centrifugation 5 mins at 2,000 g and supernatant was discarded. Beads were washed 6 times in highly stringent washes as the neutravidin affinity for biotin is very high. Wash 1+2 (RIPA with Triton X), wash 3 (1 M KCl), wash 4 (0.1 M Na_2CO_3), wash 5 (2 M urea in 10 mM Tris-HCl pH8) and wash 6 in RIPA again. Beads were either washed a further 3 times in 1X PBS and sent to be analysed by mass spectroscopy or eluted in 4X Laemmli buffer supplemented with 15 mM biotin and boiled for analysis by Western blot.

2.8.3 Liquid chromatography - mass spectrometry (LC-MS/MS)

Experiments sent for LC-MS were performed in triplicate. Each condition involved lysate from 2x15 cm^2 dishes, negative control, WT and mutant. Biotinylated proteins conjugated to neutravidin beads were sent to King's College London MS facility to be analysed. Samples were reduced with dithiothreitol, alkylated with iodoacetamide, and digested overnight with trypsin at 37 °C. Resulting peptides were purified using C18 spin columns, dried by vacuum centrifugation, and stored at -20 °C prior to analysis.

Peptides were resuspended in 0.05% formic acid / 2% acetonitrile and analysed by nano-LC MS/MS using a Vanquish Neo UHPLC system coupled to an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific). Peptides were separated on a 75 μm $\times$ 50 cm

C18 column using a 60-min reversed-phase gradient. Data were acquired in data-dependent acquisition mode with higher-energy collisional dissociation (HCD) fragmentation.

2.8.4 Data processing and protein identification

Raw data were processed using Proteome Discoverer v2.5 and searched with the Sequest HT algorithm against the UniProt human database. Protein identification and relative quantification were used to assess comparative fold-changes between experimental groups and quantification was carried out in Scaffold v5 proteomics tool (www.proteomesoftware.com/products/scaffold-5).

2.9 Computational analysis

2.9.1 Building a phylogenetic tree of LNX orthologous proteins

Orthologous protein sequences representing ancestral lineages of the gene family of interest were compiled into a FASTA-formatted file. Phylogenetic reconstruction was performed using amino acid sequences to reduce substitutional saturation and improve inference of deeper evolutionary relationships.

Multiple sequence alignment was conducted using MAFFT v7 (Katoh and Standley, 2013) executed within a Windows Subsystem for Linux (WSL) Ubuntu environment. Default alignment parameters were applied, which employ an FFT-accelerated progressive alignment strategy optimized for protein sequences.

```
mafft input_LNX_seq.fasta > aligned_LNX_sequences.fasta
```

Phylogenetic relationships were inferred using FastTree v2 (Price, Dehal and Arkin, 2010), which implements an approximately maximum-likelihood (ML) method optimized for large alignments. For protein sequences, FastTree applies the WAG substitution model with a discrete gamma model of rate heterogeneity across sites by default.

```
FastTree aligned_LNX_sequences.fasta > tree.nwk
```

The Newick-formatted tree was visualised and annotated using Interactive tree of life (iTOL) v6 (Letunic and Bork, 2024). Branch labels and taxon groups were manually colour coded.

Chapter 3

A Multi-Modal Approach to Identify and Validate Novel LNX1 and LNX2 Protein Interactions

3.1 Introduction

Ligand of NUMB protein X1 and 2 (LNX1/2) proteins were first identified based on their ability to interact with NUMB protein (Dho *et al.*, 1998) (Rice, Northcutt and Kurschner, 2001). Later structural analysis of LNX1 and LNX2 revealed four PDZ (PSD-95, DlgA, ZO-1) domains, an N-terminal NUMB binding motif (NPAY/NPAF) and a RING (Really Interesting New Gene) domain (Figure 1.7). RING domain containing proteins are classed as the largest family of E3 ubiquitin ligase enzymes (Li *et al.*, 2008). The longer, non-neuronal LNX1 isoform, LNX1^{p80}, can ubiquitinate NUMB p72 and p66 isoforms through its catalytic RING domain, and target it for proteasomal degradation (Nie *et al.*, 2002a)(Nie, Li and McGlade, 2004). The shorter, central nervous system (CNS) specific LNX1^{p70} isoform does not have a RING domain nevertheless evidence may suggest LNX1^{p70} does interact and sequester proteins for ubiquitination perhaps indirectly (Lenihan, Saha and Young, 2017).

PDZ domains are short ~80 amino acid domains that function as protein-protein interaction modules. PDZ domains are characterised by the type of carboxyl terminal PDZ binding motif (PBM) of other proteins that they favour (Songyang *et al.*, 1997) (Bezprozvanny and Maximov, 2001). Both LNX1 and LNX2 proteins have four PDZ domains each allowing for a plethora of protein-protein interactions. Wolting *et al.* (Wolting *et al.*, 2011) compiled a list of 220 mammalian LNX1/2 interacting proteins and the Guo group (Guo *et al.*, 2012) added 30 more interactors to this list. Both publications used protein arrays or yeast-two hybrid screens of individual LNX PDZ domains. More recently (Lenihan, Saha and Young, 2017) used affinity purification of full length GFP-

LNK1 from a stably expressing cell line followed by proteomic analysis to identify 76 more interactors. From these lists of interactors only a few have been fully validated by *in vitro* assays and even less have been confirmed as LNK ubiquitination substrates.

A survey of some of these known interactions reveals the potential diversity of LNK protein function, for example, neuronal development, control of tight junctions and regulation of cellular signalling. LNK1/2 PDZ domain interactions include protein(s) that make up components of epithelial tight junctions - CLAUDINS (Takahashi *et al.*, 2009), junction adhesion molecules (JAMs) (Kansaku *et al.*, 2006) and CXADR/CAR (Sollerbrant *et al.*, 2003). CLAUDIN-1-2 and -4 are substrates of LNK1 ubiquitination leading to endocytosis and subsequent lysosomal degradation (Takahashi *et al.*, 2009). JAM4 and CAR have not been identified as substrates yet, but compelling evidence supports interaction with LNK1 and LNK2. These interactions might serve as evidence for a role in modulating formation or remodelling of epithelial tight junctions and could potentially be relevant in the context of epithelial to mesenchymal transition (EMT) in metastasis (Kansaku *et al.*, 2006).

Interactions with transcriptional regulators and components of intracellular signalling pathways have suggested a function for LNK1/2. LNK1 localises at least in part, to the nucleus (Weiss *et al.*, 2007)(Zheng *et al.*, 2010) and contains a putative nuclear localisation signal (NLS) (Cioccarelli, L. manuscript in preparation, 2026. P.Y. Lab). Ski-interacting protein (SKIP), a component of NOTCH transcriptional activation complexes, interacts with PDZ1 of LNK1 (Vasquez-Del Carpio *et al.*, 2011) (Chen *et al.*, 2005). In addition, LNK1 ubiquitinates c-SRC and the RHO GTPase activating protein BCR, leading

to reduced cellular levels of these proteins; this effect is reversed upon proteasome inhibition (Weiss *et al.*, 2007) (Guo *et al.*, 2012). Collectively, these findings indicate that the LNX1 interactome plays a significant role in intracellular signalling pathways.

Although the list of interacting partners for LNX1 and LNX2 is extensive, it is not yet complete. Many of the reported interactions were identified using individual PDZ domains of LNX1/2, while the majority have not been validated using full-length proteins *in vitro*. In this chapter, I aim to expand the currently known interactome by integrating *in silico* predictions and publicly available databases, and to validate selected candidates using a combination of domain-based screening approaches, affinity purification assays, yeast display based flow cytometry assays and co-localisation microscopy.

Aims of Chapter 3

- Gather information on potential LNX1 and 2 interactors from PDZ to PBM interactome databases (ProfAff) and machine learning predictions.
- Experimentally validate the ability of candidate PBMs to bind full-length or near full-length LNX proteins in yeast two hybrid assays.
- Confirm these interactions using full-length versions of both LNX1 or LNX2 and the interacting proteins in a mammalian cell context using pull-down assays.
- Examine subcellular localization of candidate interactors and possible colocalization with LNX proteins.

3.2 Results

3.2.1 Identifying possible interactors of LNX1/2 from ProfAff and

***in silico* predictions**

ProfAff (Profiling Affinities) is a publicly available interactomic database of 5 million human protein affinities. This enormous effort was developed by the Université Côte d'Azur and contains 65,000 PDZ-PBM affinities along with 14-3-3 motif interactions (Gogl *et al.*, 2022). The holdup assay is a quantitative, high-throughput method for measuring protein–peptide interactions. An immobilised protein domain is incubated with a peptide ligand, and the amount of peptide retained after equilibrium binding is determined from the depletion of peptide in the flow-through, allowing the affinity of the interaction to be quantified. Hold-up assays between PDZ domains and 10-mer PBM peptide fragments were used to determine a pK_d (steady-state dissociation constants or the negative of the base 10 logarithm of the dissociation constant) i.e. the relative strength or likelihood of interacting. Only pK_d values of >3.4 are considered as confident interactors. The ProfAff database determined 267 interactors for LNX1 and 262 for LNX2. Individual PDZ interaction breakdown and the specific consensus sequence preferred is given in Figure 3.1. LNX1 PDZ1 and 2 and LNX2 PDZ2, 3 and 4 prefer class I PDZ binding motifs X-S/T-X-V/L while LNX1 PDZ3 prefers to bind class II PBMs having a hydrophobic residue (W/V) at position -2 and 0. LNX2 PDZ4 only has 6 binding PBMs out of the 44 tested so, a the weighted affinity logo is not as confident as the other PDZ domains. The binding affinity given from ProfAff is not definitive and further verification is needed, especially since domain only individual PDZ domains and 10 amino acid PBM peptides are utilised in the hold-up assays.

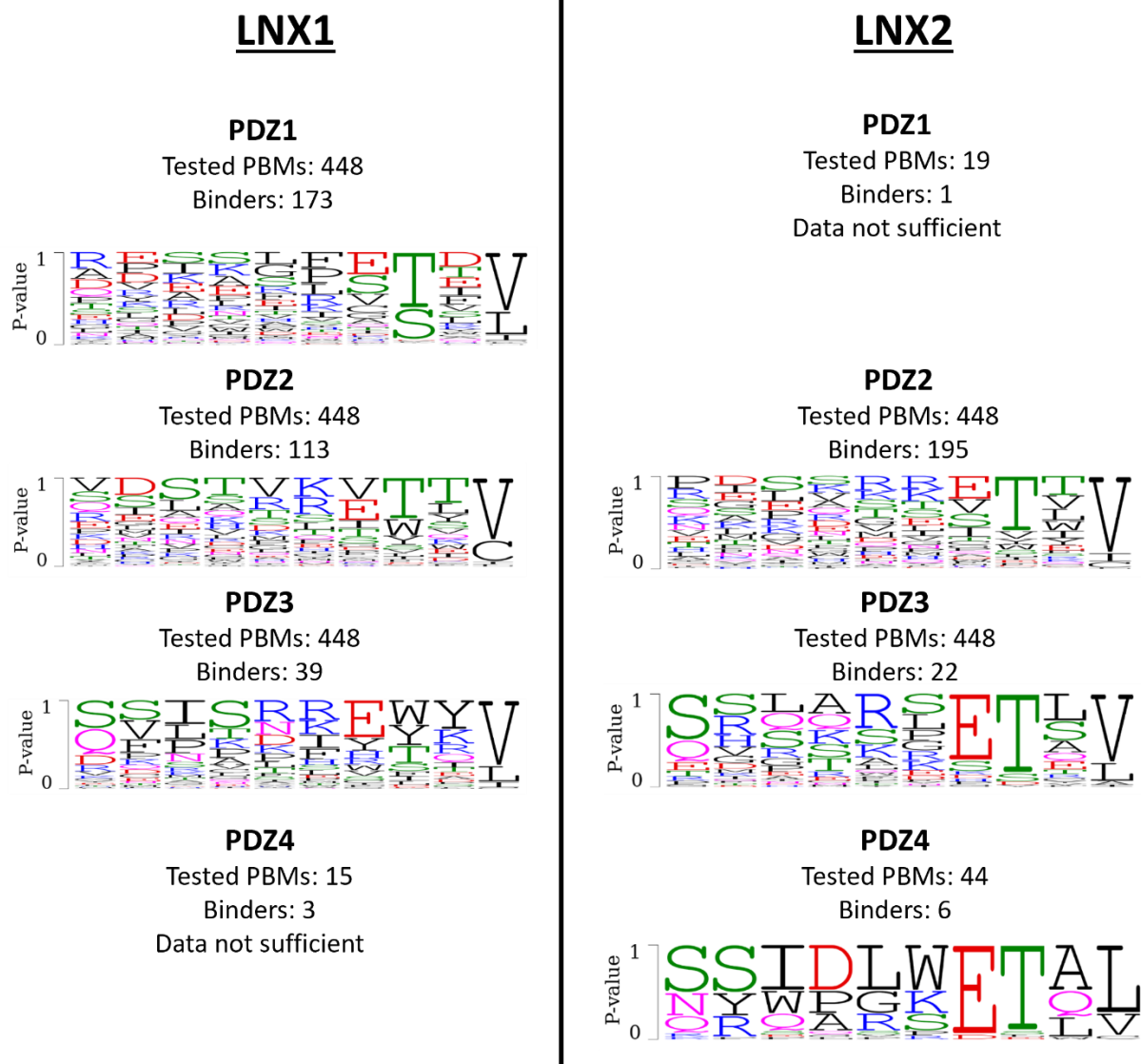


Figure 3.1: ProfAff-generated affinity-weighted logos for each LNK1/2 PDZ domains consensus binding motif. Number of PBMs tested against each PDZ domain versus how many interact above a threshold of $pK_d=3.5$ is indicated (Gogl *et al.*, 2022). LNK1 PDZ4 and LNK2 PDZ1 were not tested against enough PBM sequences to generate a reliable consensus motif.

Previously validated and published full-length LNX1/2 interactors reported by Guo *et al.*, Wolting *et al.*, and Lenihan *et al.*, were excluded from further analysis. In collaboration with Lucy Holmes, a final-year undergraduate student whom I supervised, PBM predictions from the ProfAff database were cross-referenced with machine learning-based predictions generated by MSc student Lee Kavanagh (L.K). A comprehensive list of LNX1/2 interactors from the outlined sources compared to ProfAff data is described in Figure S1. Candidates were also investigated based on their function; as LNX proteins have a role in neuronal development and tumorigenic roles in cell signalling, CNS-related proteins (NRXN1 β , GRIA2) and components of signalling pathways were included (YES1, RYK, SMG1, NDRG3). From this integrated analysis, a list of 24 unique candidate LNX1/2 interactors was selected for further examination (Table 3.1).

Table 3.1: A list of potential LNX1 and/or LNX2 interactors that contain PBMs at their carboxyl terminus. Liprin alpha 1 (PPFIA1) is included as a positive control for the assay as it is a well-known and verified interactor of LNX1. The source where the interaction was identified is cited. Only interactions that had not yet been validated with full-length LNX1 and/or LNX2 proteins were selected.

Gene symbol	Protein name	Carboxyl terminus	Source
CTNNA2	Alpha catenin-2	LSEFKAMDSF	(Lenihan <i>et al.</i> , 2017) ProfAff
RYK	Related to receptor tyrosine kinase	EFHAALGAYV	L.K. Machine learning
YES1	Tyrosine-protein kinase YES	EPQYQPGENL	L.K. Machine learning
ITCH	Itchy E3 ubiquitin protein ligase	IEETEGFGQE	L.K. Machine learning
MARCHF9	Membrane-associated RING-CH-Type finger 9	REVVMRVTTV	L.K. Machine learning ProfAff
RNF146	<i>RING finger protein 146</i>	PDGQCTVTEV	L.K. Machine learning ProfAff
IGSF11	Immunoglobulin superfamily member 11	PAQSRAGSLV	L.K. Machine learning

SDK2	Sidekick-2	APIAGFSSFV	L.K. Machine learning ProfAff
NRNX1β	Neurexin-1-alpha	KKNKDKEYYV	(Lenihan <i>et al.</i> , 2017) ProfAff
AMIGO1	Adhesion molecule with Ig like domain-1	SVFSDTPIVV	(Lenihan <i>et al.</i> , 2017)
SRR	Serine racemase protein	RPAPYQTVSV	(Lenihan <i>et al.</i> , 2017)
CYP4X1	Cytochrome P450 family X member 1	YLHLKKLSEC	L.K. Machine learning
GRIA2	Glutamate Ionotropic receptor AMPA type subunit 2	NVYGIESVKI	(Lenihan <i>et al.</i> , 2017) ProfAff
ZNF24	Zinc finger protein 24	AEKLLNVVKV	(Lenihan, Saha and Young, 2017)
PPFIA1	Liprin-alpha-1-protein	DSATVRTYSC	(Lenihan <i>et al.</i> , 2017) ProfAff
ORC1	Origin recognition complex-1	FSFLLFVSHC	ProfAff
EPHB2	Erythropoietin-producing hepatocellular carcinoma B2	QMNQIQSVEV	(X. D. Liu <i>et al.</i> , 2018)
BCR	Breakpoint cluster region protein	RQSILFSTEV	(Wolting <i>et al.</i> , 2011) ProfAff
CTNNB1	Catenin beta-1	NQLAWFDIDL	(Lenihan <i>et al.</i> , 2017) ProfAff
CTNND2	Catenin delta-2	HYPASPDWV	(Wolting <i>et al.</i> , 2011) ProfAff
SMG1	Serine/threonine-protein kinase SMG1	QLYEGWTAWV	L.K. Machine learning
PRKAR1α	cAMP-dependent protein kinase type I- alpha regulatory subunit	QYNSFVSLSV	(Lenihan <i>et al.</i> , 2017)
NDRG3	N-myc downstream-regulated gene 3	DRHQTMVSC	(Lenihan <i>et al.</i> , 2017)
DVL2	Segment polarity protein dishevelled homolog DVL-2	NPSEFFVDVM	L.K. Machine learning
KRTAP4-3	Keratin-associated protein 4-3	HPICCGSSCC	L.K. Machine learning

3.2.2 Yeast-two hybrid analysis of selected candidates

The selected candidates from Table 3.1 were screened against LNX1 PDZ1-4, LNX2 and empty vector control, available in the lab. We previously used full length LNX1 as a prey, but it was found to be inefficient in this assay context. Individual PDZ domains could be used here instead, however, the predicted 3D folding of LNX proteins reveals PDZ1, 3 and 4 are less accessible than PDZ2. Full length LNX2 and the N-terminally truncated LNX1 PDZ1-4 constructs are likely to reflect the natural 3-dimensional conformation and conserved intramolecular interactions of the LNX PDZ domains (Chapter 5) and were thus expected to provide insights regarding the physiological relevance of candidate interacting proteins. As PDZ binding motifs are short, the C-terminal 10 amino acids of each of these proteins was cloned into pLEX bait vector. A Yeast-2-hybrid assay was chosen as it is a quick and inexpensive method of identifying binding partners (Fields and Song, 1989). Activation of reporter genes *His3* and *LacZ* indicate interaction between the LNX protein (prey-pGAD) and the 10-mer (bait-pLEX). Initially, a screen against a negative control prey was carried out to ensure the bait candidates did not cause autoactivation of the reporter genes. The candidates that autoactivated (PRKAR1 α , NDRG3, DVL2 and KRTAP4-3) were not taken further in the yeast-2-hybrid analysis, but they cannot be excluded as interactors of LNX1 or LNX2. CTNNB1, CTNND2 and SMG1 also showed signs of autoactivating (Figure 3.2).

From machine learning-based PBM-LNX interactions, only RYK and IGSF11 10-mer peptides gave positive screens in the Y2H. Interestingly Sidekick-2 (SDK2), which was seen in the ProfAff database was correctly predicted by machine learning, had an

obviously higher reporter gene activation for LNX2 rather than LNX1 (Table 3.2 and Figure 3.2).

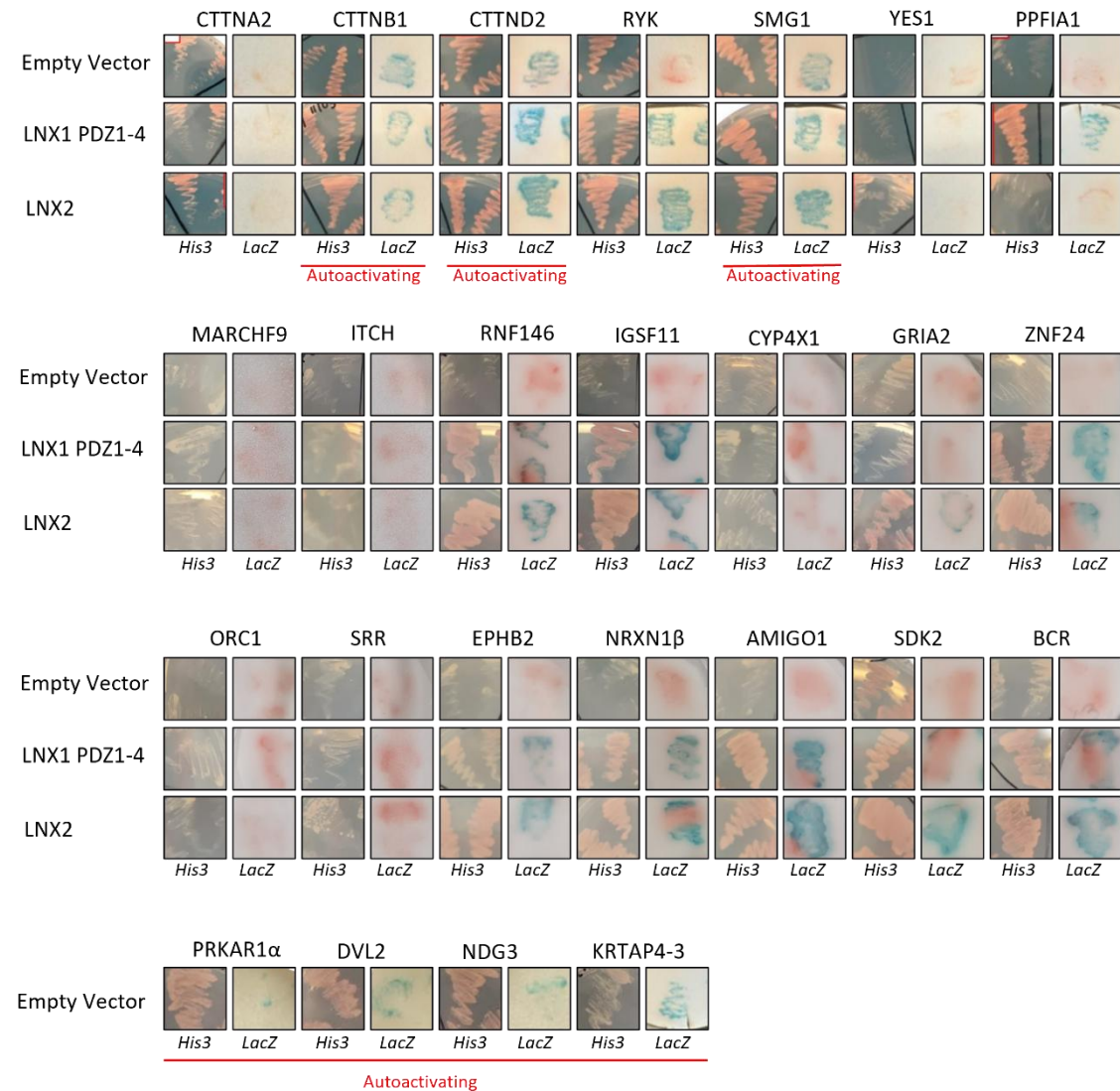


Fig 3.2: Evaluation of candidate LNX interactions by yeast-2-hybrid assay. Co-transformed yeast colonies were streaked on media lacking leucine, tryptophan and histidine, growth of reddish/pink cells indicates the bait and prey proteins are interacting resulting in activation of the *His3* reporter gene. A positive β -galactosidase lift assay showing blue colonies when subjected to X-gal substrate indicates activation of *LacZ*. Co-transformation of an empty prey vector with each bait serves as a negative control which should show no growth or colour change. When this is not the case the bait is causing autoactivation of the reporter gene(s) and the presence or absence of an interaction cannot be determined.

Table 3.2: Qualitative results of Yeast-2-hybrid assays. +, ++, +++ or - is representative of the strength of the β -galactosidase colour change or the amount of yeast growth in histidine lacking conditions, as seen in Figure 3.1. “Auto” indicates autoactivation of the reporter gene(s) by the bait in the negative control, and an interaction cannot be determined. * Indicates further analysis was performed.

Gene symbol	LNK1 PDZ1-4		LNK2	
	<i>His3</i>	<i>LacZ</i>	<i>His3</i>	<i>LacZ</i>
CTNNA2	+	-	+	-
RYK	++	++	++	++
YES1	+	-	+	-
ITCH	-	-	-	-
MARCHF9	-	-	-	-
RNF146	++	+	++	++
IGSF11	+++	++	+++	++
SDK2*	++	+	+++	+
NRNX1 β *	+++	+	+++	++
AMIGO1	++	++	++	++
SRR	-	-	-	-
CYP4X1	-	-	-	-

Gene symbol	LNK1 PDZ1-4		LNK2	
	<i>His3</i>	<i>LacZ</i>	<i>His3</i>	<i>LacZ</i>
GRIA2	+	-	++	++
ZNF24*	+++	+	+	+
PPFIA1	+++	++	++	+
ORC1	-	-	-	-
EPHB2*	+++	++	+++	++
BCR*	+++	+	++	+
CTNNB1	Auto	Auto	Auto	Auto
CTNND2	Auto	Auto	Auto	Auto
SMG1	Auto	Auto	Auto	Auto
PRKAR1 α	Auto	Auto	Auto	Auto
NDRG3	Auto	Auto	Auto	Auto
DVL2	Auto	Auto	Auto	Auto
KRTAP4-3	Auto	Auto	Auto	Auto

3.2.3 Validation of full-length protein interactions by GFP and MBP pull down (PD) assays

While yeast-two hybrid screens are efficient the results must be interpreted with caution as only the 10 amino acid peptide is used. To further validate the interactions observed in section 3.2.2, we decided on five proteins that were most relevant to our investigation and were not fully characterised previously.

LNK1 has been shown to interact with and ubiquitinate BCR (Guo *et al.*, 2012) but an interaction between full-length LNK2 and BCR has not been seen. Similarly, with ZNF24, a kruppel like transcription factor (Li *et al.*, 2013), LNK1 interaction has been identified but not LNK2 (Lenihan, Saha and Young, 2017). NRXN1 β was immunoprecipitated by PDZ2 of LNK2 from mouse brain lysate but an interaction with LNK1 was not seen (Lenihan *et al.*, 2017). Interactions between SDK2 10-mer carboxyl terminal sequence and PDZ1 of LNK1 and PDZ2 of LNK2 were identified from the ProfAff database, but full-length interactions were not ascertained. Lui *et al.* characterised the receptor tyrosine kinase EPHB2, and LNK1 interaction as part of a multicomplex scaffold on the surface of CA3 hippocampal neurons, but an interaction with LNK2 has not been studied (Liu *et al.*, 2018).

The specificity of the interaction of these proteins with LNK1 and LNK2 was first assessed. GFP-tagged LNK proteins and either HA or Flag tagged interacting proteins were co-transfected into HEK 293T cells and interactions were assessed by GFP-trap (methods section 2.5.5). Results are shown as Western blots in Figure 3.3.

Strong enrichment for the GFP tagged LNK proteins and GFP alone was seen, and no interacting protein specific signal was detected for the negative control. NRXN1 β , BCR

and ZNF24 were successfully pulled down by both LNX1 and LNX2. As suggested by the yeast-two hybrid assay, SDK2 bound efficiently and specifically to LNX2 versus LNX1.

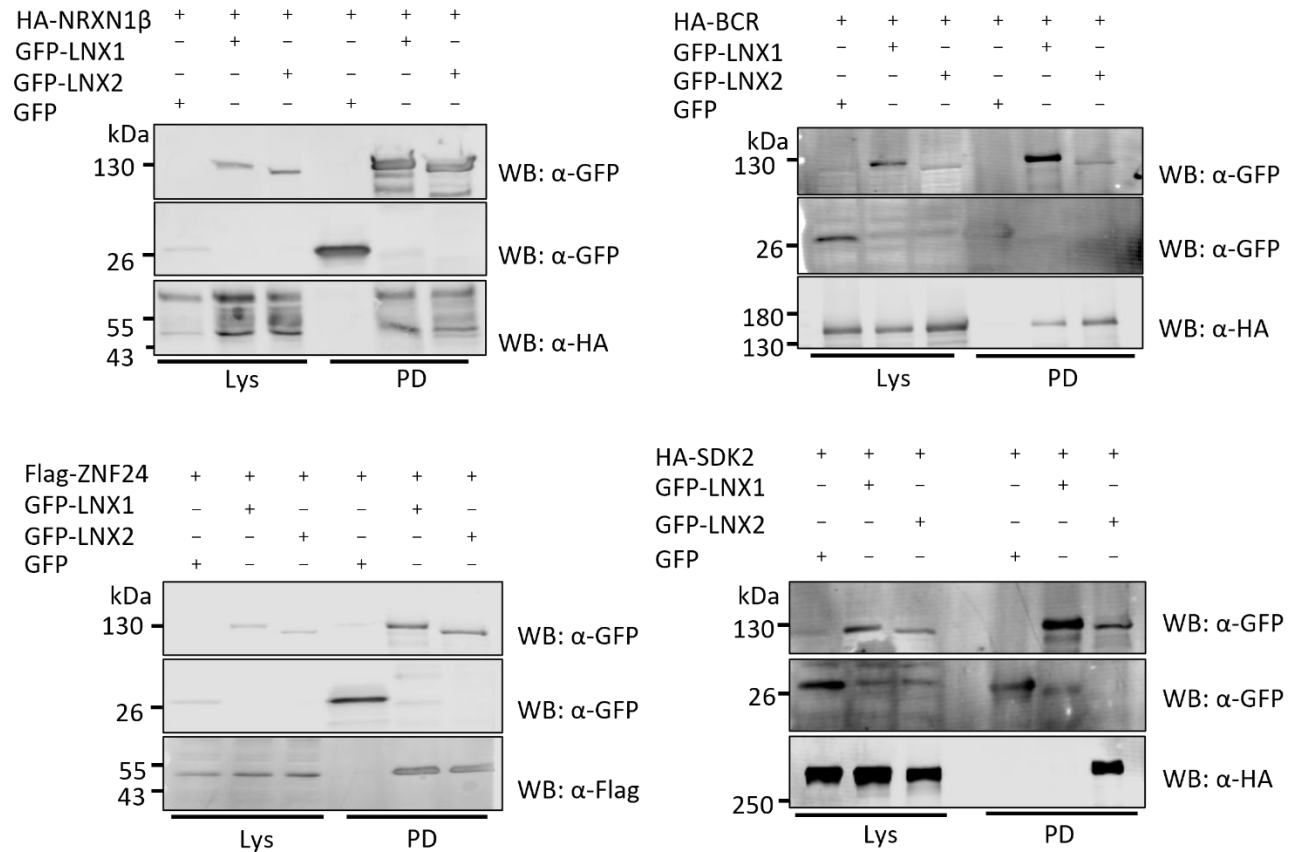


Figure 3.3: Specificity of the interaction of full-length NRX1β, BCR, ZNF24 and SDK2 proteins for LNX1 versus LNX2. The ability of the indicated proteins N-terminally tagged with either a HA or Flag epitope to interact with transfected GFP-tagged LNX constructs was assessed in HEK 293T cells. For each interacting protein NRXN1β, BCR, ZNF24 and SDK2, Western blots of cell lysates (Lys) confirmed expression of the indicated construct, and the output of a GFP pull-down (PD) assay is shown by indicated antibodies. Successful expression and pull-down of GFP-tagged LNX proteins was verified in all assays by Western blot using an anti-GFP antibody. Blots are representative of three independent experiments.

From our lab, verified *LnX1* and *LnX2* double knock out mouse brain lysates were available. EPHB2 is well expressed in mouse brain, and from Lui *et al.* the LNX1 interaction was described but no interaction with LNX2 was investigated (Liu *et al.*, 2018). Due to the size of EPHB2 (110 kDa) it proved difficult to differentiate between GFP-LNX and EPHB2 by Western blot (data not shown). Our solution was to amino terminally tag LNX2 with MBP and express and purify the fusion protein from BL21 *E. coli* cells. The purified protein was mixed with brain lysate and amylose beads were used to pull down LNX2 and any associated proteins (Methods section 2. 5.3). Pull down (PD) samples were analysed by Western blot and probed for endogenous EPHB2 (Figure 3.4).

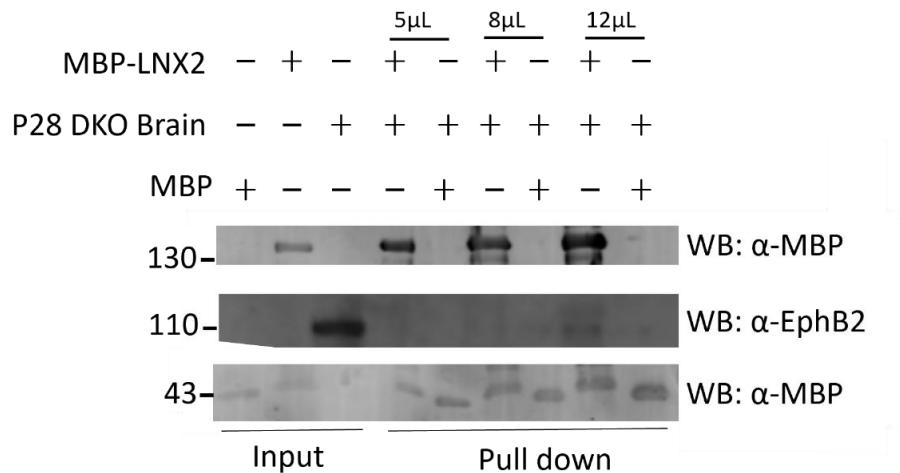


Figure 3.4: Endogenous EPHB2 can be pulled down from mouse brain lysate by purified MBP-LNX2. Purified MBP-LNX2 protein was mixed with postnatal day 28 *LnX1^{-/-}; LnX2^{-/-}* mouse brain lysate and allowed to rotate end over end for 4 hours. Amylose beads were utilised as a method for purifying MBP-LNX2, beads were boiled in 2x Laemmli buffer and different amounts (5, 8 and 12 μL) of sample were analysed by Western blotting technique. Endogenous EPHB2 was observed in the MBP-LNX2 pull-down fractions (most noticeably with 12 μL of the sample was loaded). Blot is representative of three independent experiments.

3.2.4 Exploration of colocalization between LNX and novel interacting proteins in HeLa cells

To further confirm an *in vivo* association between LNX and interacting proteins, we examined their colocalization in cultured cells. Immunofluorescence microscopy was utilised to determine the distribution pattern of both proteins in a cellular system and visualise any colocalization. HeLa cells were chosen for this as they are larger and more adherent than HEK 293T cells and are easy to transfect.

GFP tagged LNX protein was co-transfected with HA or Flag tagged interacting proteins and nuclei were stained using DAPI (methods section 2.3.1). LNX1^{p80} exhibits a diffuse

nuclear and cytosolic pattern while LNX2 is primarily in the cytosol with very little signal in the nucleus (Figure 3.5A).

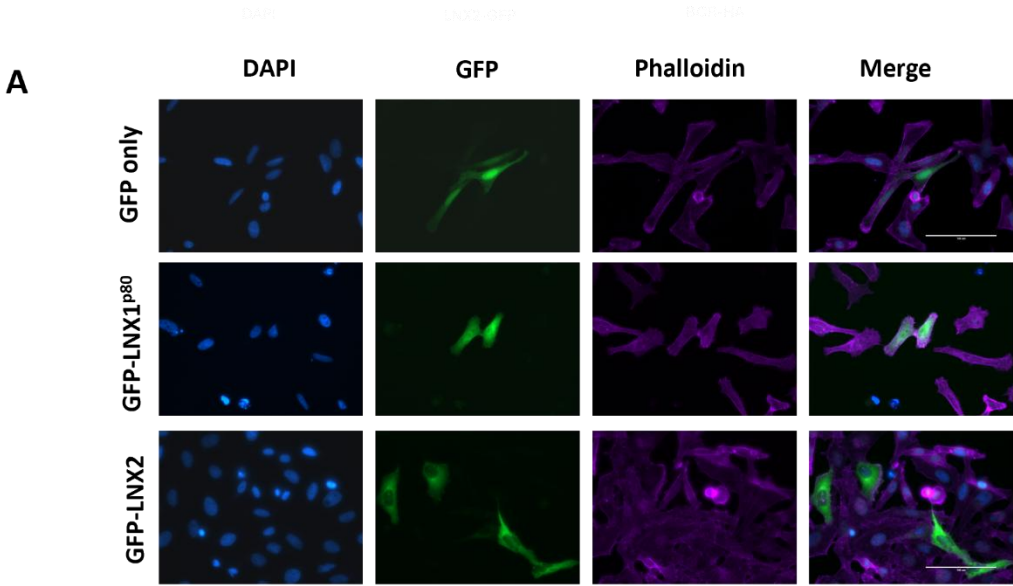
Zinc finger protein 24 (ZNF24) as the name suggests is a DNA binding transcription factor which localises exclusively to the nucleus (Jia *et al.*, 2015). Upon co-transfection, small intranuclear structures corresponding to LNX1^{p80} formed which were not present with GFP alone or LNX2 (Figure 3.5B). DAPI stain did not reflect the reorganising of DNA into these punctuae. The formation of these dot-like structures and strong nuclear ZNF24 signal supports the interaction with LNX1 and ZNF24.

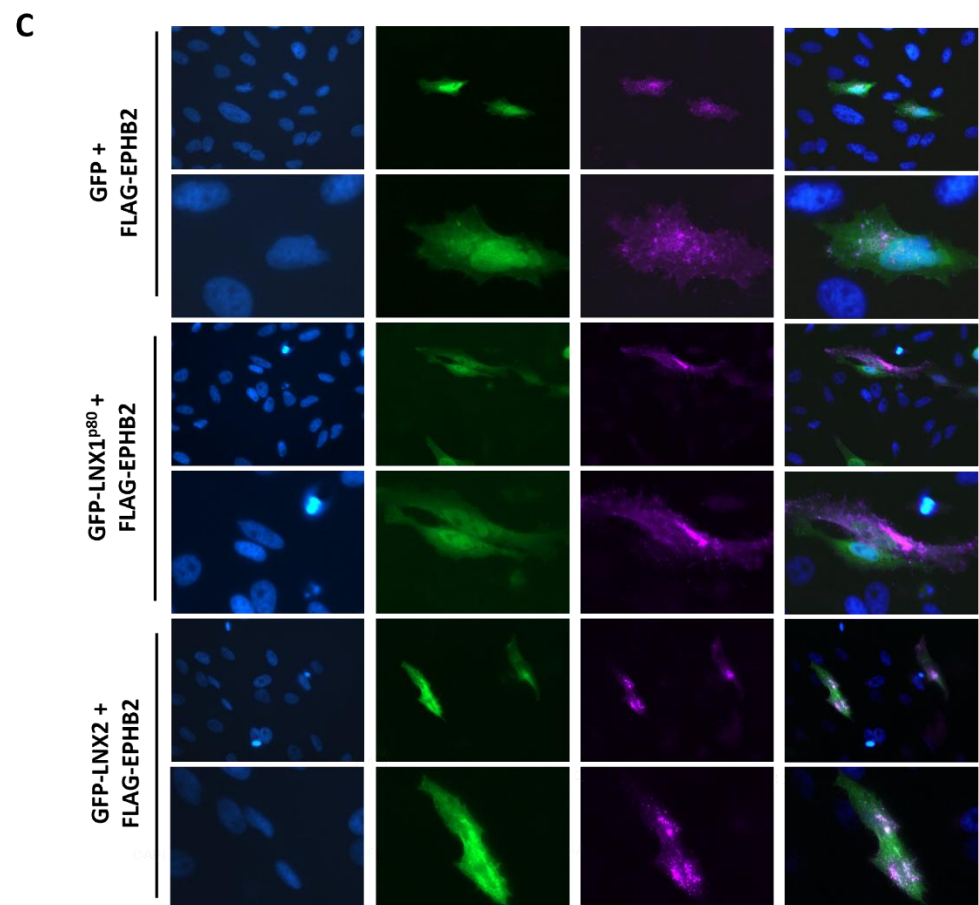
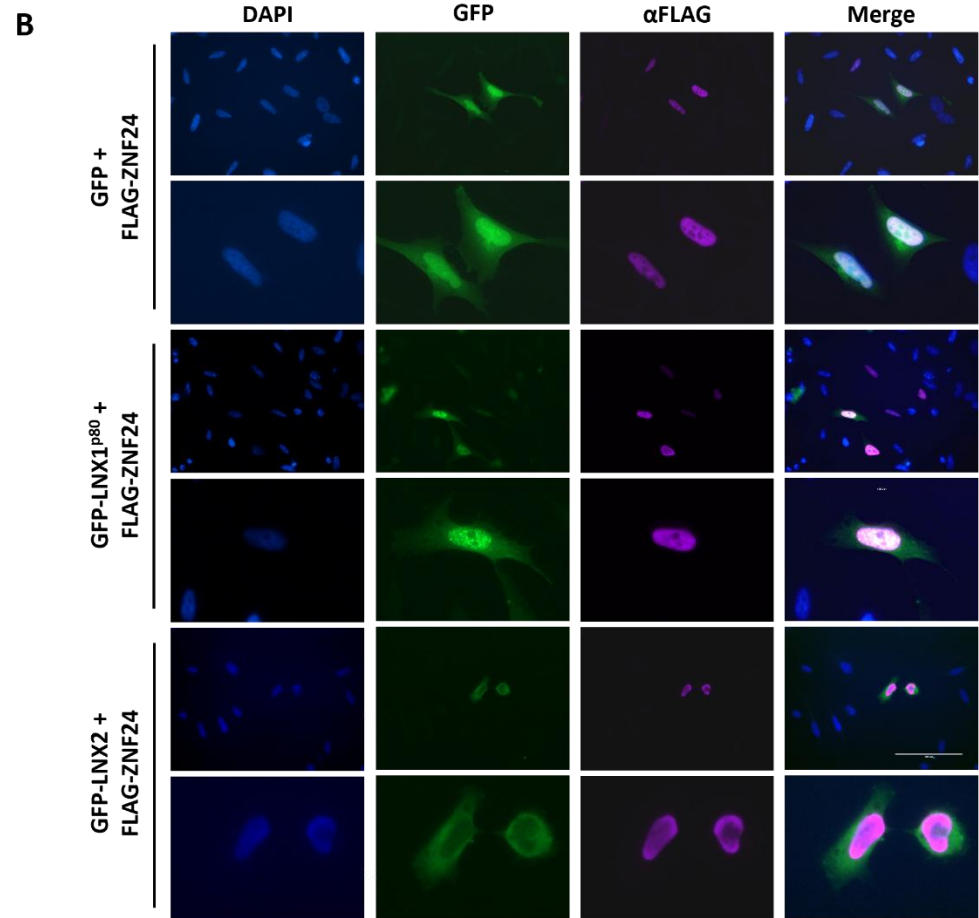
Flag-EPHB2 displays a mostly punctuate distribution across the cytosol and nucleus of HeLa cells. When LNX1 or LNX2 is co-transfected, EPHB2 forms dense focal points around the nucleus that seem to co-localise with LNX2. There is no obvious formation of dense localisation of LNX1 around the nucleus (Figure 3.5C).

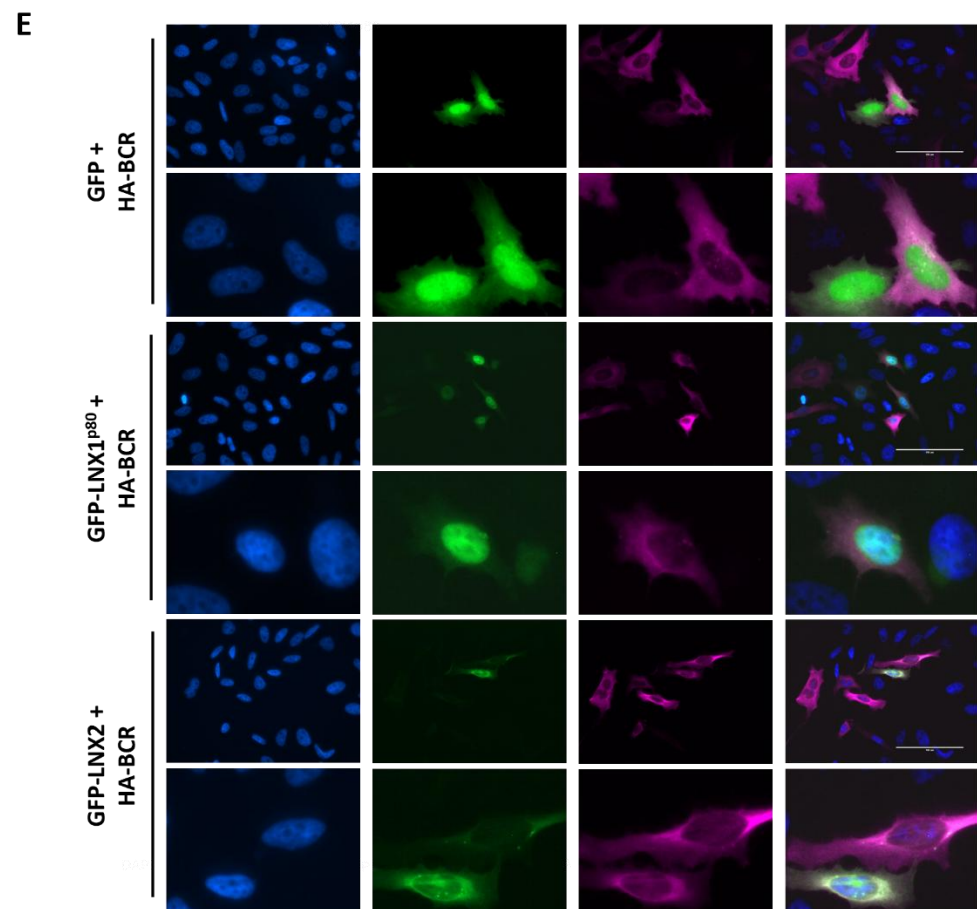
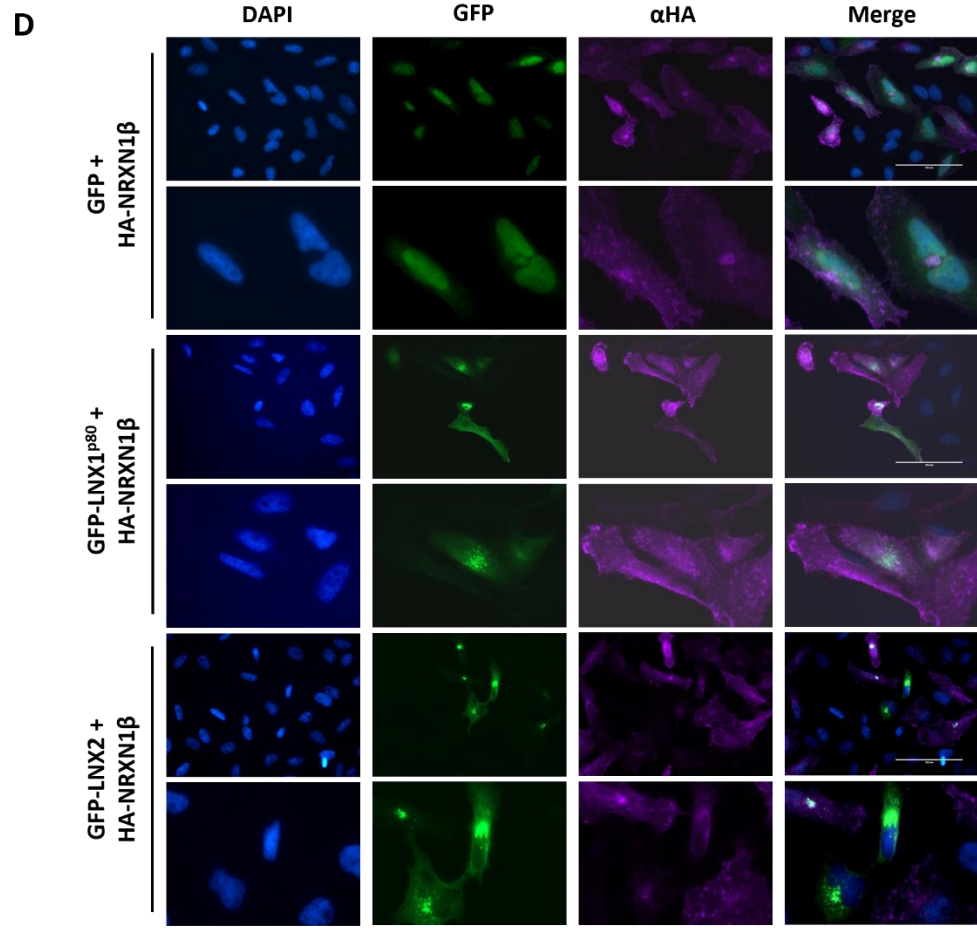
Neurexin-1- β is a neuronal surface protein involved in cell adhesion and formation of synaptic junctions (Bang and Owczarek, 2013). When transfected it inclines towards the plasma membrane except upon the addition of LNX protein, where it forms punctuae around the nuclear envelope. This is more obvious upon co-transfection with LNX2, as LNX2 forms a similar localisation pattern in which slight overlap can be seen. Both LNX1^{p80} and LNX2 change their cellular distribution when NRXN1 β is overexpressed in cells (Figure 3.5D). BCR is expressed primarily in the cytosol in a diffuse pattern which would support the already known interaction with LNX1^{p80}. Upon the addition of GFP-LNX2, BCR localises around the nucleus perhaps in the endoplasmic reticulum in an

almost filamentous pattern spanning outwards which then co-localises with LNX2 as seen in the merged image in yellow (Figure 3.5E).

Similarly to NRXN1 β , Sidekick-2 is also a synaptic adhesion protein containing a transmembrane domain (Yamagata, 2020). SDK2 displays a filamentous structure in HeLa cells which condenses into a compact punctate structure localising near the endoplasmic reticulum when LNX2 is also present (Figure 3.5F). There is clear evidence for co-localisation of LNX2, and SDK2 but not of LNX1, at least in HeLa cells, thus supporting the results from section 3.2.3.







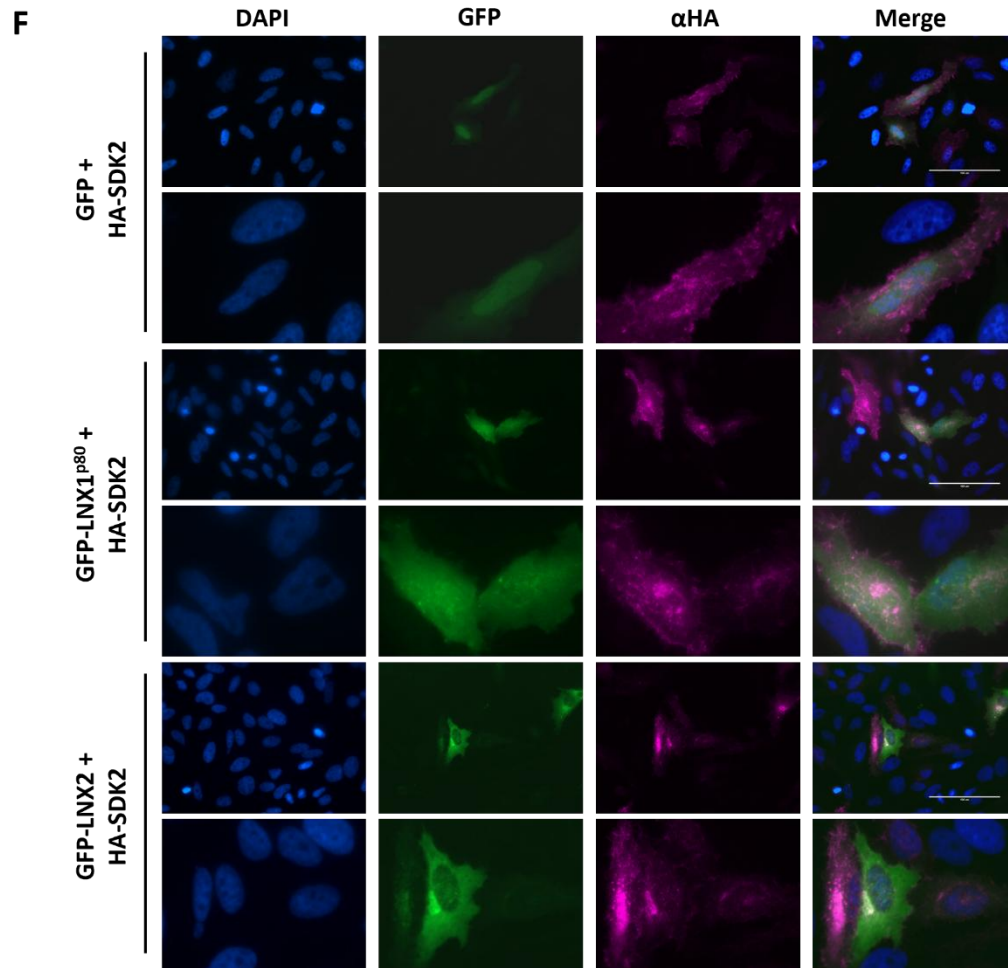


Figure 3.5: Cellular localisation of LNX candidate interacting proteins upon transient expression of LNX1^{p80} or LNX2. Empty GFP-expressing vector, GFP-LNX1^{p80} or GFP-LNX2 were transfected alone **(A)** or together with the indicated candidates tagged with FLAG **(B-C)** or HA **(D-F)** epitope tags in HeLa cells for 24 hr. Cells were then stained with DAPI (blue) and probed for Phalloidin (F-actin), HA or FLAG tags (magenta) using antibodies. For each condition, images are presented with a 40X magnification (top) following by an enlarged 60X magnification (bottom). Scale bar indicates 100 μ m. Images are representative of two independent experiments.

3.2.5 Characterisation of the ZNF24 – LNX1 interaction

3.2.5.1 Validating anti-GFP-Nanobody displaying yeast as a method of co-immunoprecipitation.

ZNF24 was identified as an interacting partner of full length LNX1 protein from a large proteomics screen but the domain responsible for this interaction was not described (Lenihan, Saha and Young, 2017). As commercially available GFP-trap beads are expensive, I aimed to develop an alternative, cheaper and more convenient strategy to assess such interactions. For this I utilised anti-GFP nanobody displaying *S. cerevisiae* which were developed in our lab (O’Riordan *et al.*, 2024) (Jurić *et al.*, 2025 Manuscript in press) to immunoprecipitate GFP tagged LNX1 PDZ domains and ZNF24 and analyse the binding efficiency using flow cytometry.

To first validate the possibility of using this system to study LNX interactions, two well-known LNX interactions – NUMB and Liprin α 1, were assessed (Dho *et al.*, 1998)(Lenihan *et al.*, 2017). Flag tagged NUMB or Liprin α 1 proteins were co-expressed with GFP or GFP-LNX1 in HEK293T cells, these lysates were incubated with 1×10^6 GFP-nanobody yeast cells for 30 minutes. PE-Cy7-conjugated anti-FLAG antibody was used to detect Flag-NUMB/Liprin α 1 (FL3 channel), while GFP is picked up by the FL1 channel by flow cytometry. Strong labelling signals were picked up in the GFP channel >77% of cells bound GFP or GFP tagged protein. For GFP-LNX1 + Flag-Liprin- α 1, 30% of cells were labelled while only 0.96% of cells were labelled using lysates containing GFP + Flag-Liprin- α 1 (Figure 3.6A&B). This suggests background signal is minimal and false negatives for labelling the GFP-nanobody are low. Flag-NUMB showed similar results however the proportion of yeast cell labelling positive for Flag was much lower ~10% (Figure 3.6C).

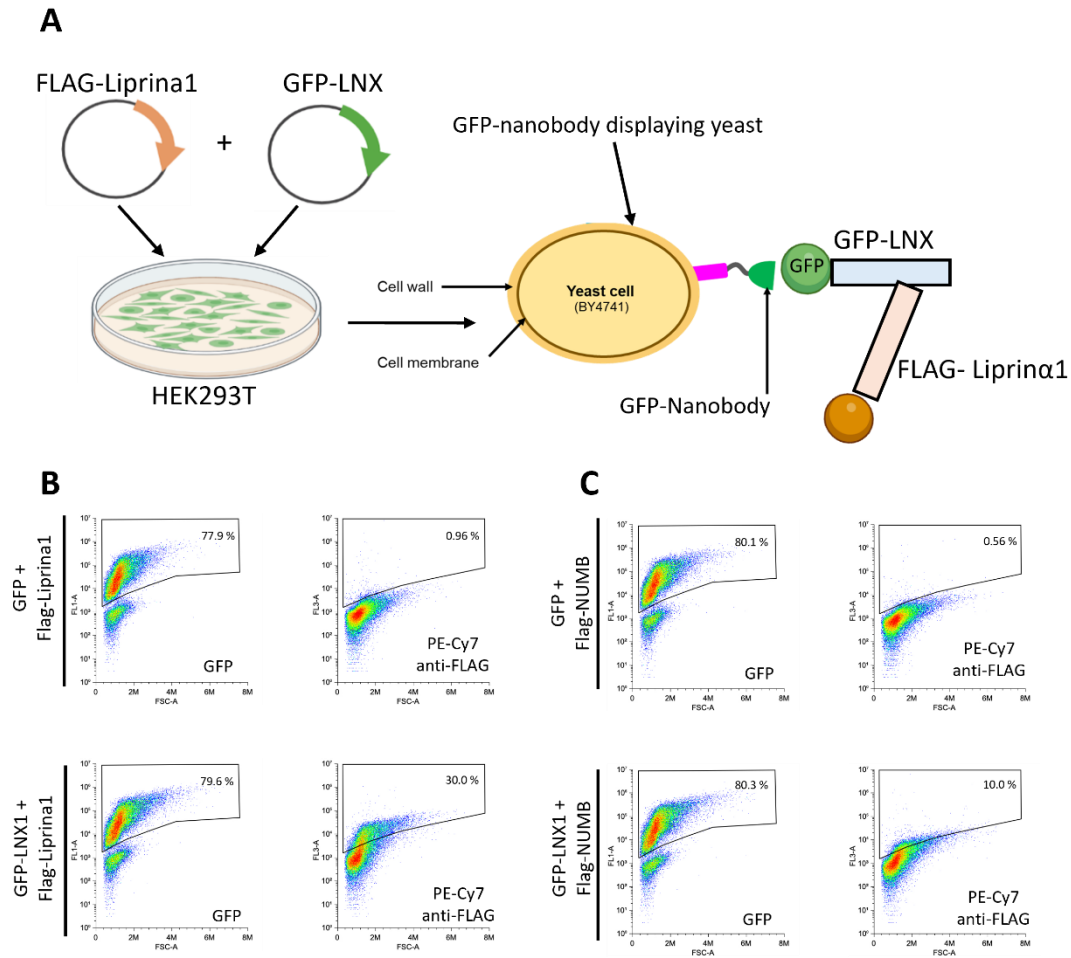


Figure 3.6: Validation of the novel yeast displaying an anti-GFP nanobody to analyse protein: protein interactions by co-immunoprecipitation and flow cytometry. (A) Schematic of the Liprin- α 1 interaction with LNX1 by GFP-Nanobody displaying yeast cells and detection by flow cytometry. 1×10^6 anti-GFP nanobody displaying yeast cells were incubated with lysates from HEK 293T cells expressing the proteins indicated on the left and analysed by flow cytometry using a PE-Cy7-anti-Flag epitope tag antibody. **(B)** Flag tagged Liprin- α 1 are detected bound to anti-GFP nanobody yeast specifically in the presence of GFP-LNX1. **(C)** NUMB protein was also immunoprecipitated in the same way by GFP-LNX1 and detected by PE-Cy7-anti-Flag antibody. Data shown is representative of two independent experiments.

3.2.5.2 The ZNF24 interaction is mapped to PDZ2 of LNX1

As shown in Figure 3.3, Zinc finger protein 24 interacts with full length LNX1 using commercial GFP-Trap beads. From yeast-two-hybrid analysis it was discovered that the interaction is between the carboxyl terminal PBM of ZNF24 and one of the PDZ domains of LNX1 (Figure 3.1). After validating the GFP-nanobody displaying yeast for determining protein - protein interactions (Figure 3.6), it was then used to identify the specific LNX1 region responsible for the ZNF24 interaction. GFP-epitope tagged full-length LNX1, and a series of truncated constructs, were employed and their interaction with ZNF24 tested following co-transfection into HEK 293T cells. Amino-terminal (RING-NPAY motif) and carboxyl-terminal (PDZ1-PDZ4) fragments, as well as constructs encoding the individual LNX1 PDZ domains were used.

Full length GFP-LNX1 bound to ZNF24 with a very high percentage of yeast cells labelled ~53%. However, LNX2 had much lower ability to bind to ZNF24, showing only 18% of cells labelled, this still being above the level of non-specific binding. The full carboxyl-terminal fragment (PDZ1-4) showed an interaction with ZNF24 and out of these PDZ domains, PDZ2 showed labelling of both GFP (79%) and Flag (5.65%). The signal for the other PDZ domains was not above the negative control (GFP alone, Figure 3.7A) and thus was interpreted as no interaction (Figure 3.7B).

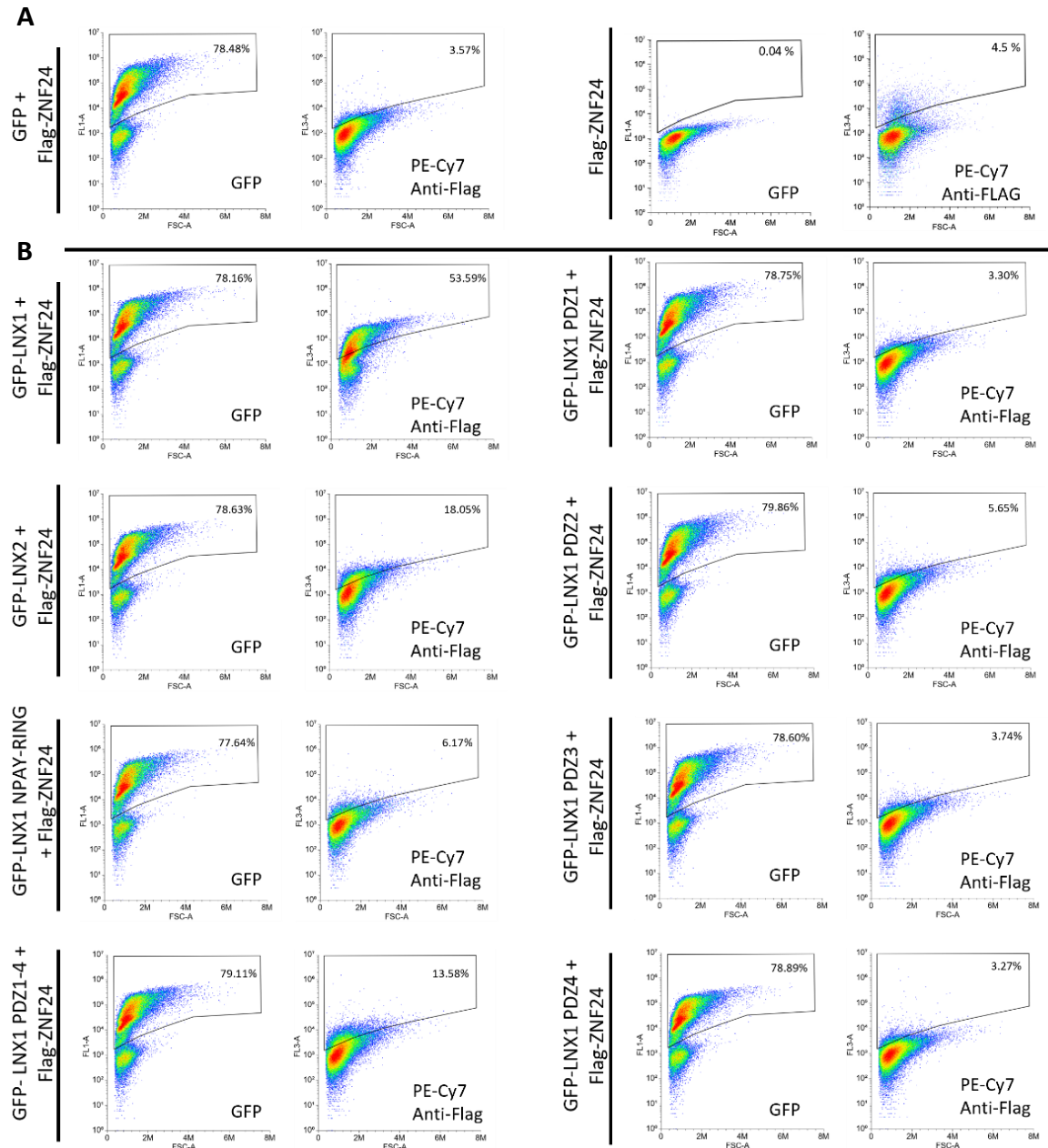


Figure 3.7: Flow cytometry analysis reveals co-immunoprecipitation of LNX1 and ZNF24 using an anti-GFP nanobody displayed by *S. cerevisiae*. For each independent experiment 1×10^6 yeast cells were mixed with HEK 293T lysate expressing the indicated (left) proteins. Analysis was carried out using an Accuri Flow cytometer and detected using a GFP fluorescence filter and PE-Cy7-anti-Flag antibody. **(A)** Negative controls for background signal. GFP alone mixed with Flag-ZNF24 (left) and Flag-ZNF24 alone (right). **(B)** Full length GFP-LNX1/LNX2 immunoprecipitated Flag-ZNF24, as did LNX1 N-terminally truncated mutant – GFP-PDZ1-4. No strong signal for Flag labelling was present with the amino terminal region of LNX1 GFP-RING-NPAY (left). Finally, individual LNX1 PDZ domains tagged with GFP were tested for ZNF24 binding revealing PDZ2 as the domain with strongest binding (right).

To further confirm this single domain interaction and to confirm the results of the flow cytometry anti-GFP nanobody, the same GFP fused constructs were again co-transfected into HEK 293T cells with Flag-ZNF24, but this time commercially available GFP-Trap beads were employed. Immunoprecipitation was carried out as per methods 2.5.5 and equal amounts of protein lysate were used as in Figure 3.7. Interacting proteins or protein fragments were separated on SDS-PAGE gels, transferred to PDVF membrane and visualised using Western blot detection antibodies (Figure 3.8). Again, full length LNX1 showed strong affinity with ZNF24. Contrary to the nanobody pull down, LNX2 showed strong binding also, this could be due to increased sensitivity of Western blotting detection. As seen in Figure 3.7B, the four PDZ domains of LNX1 could pull down ZNF24, while PDZ2 was the only singular domain that interacted with ZNF24. Outputs from both methods of co-immunoprecipitation revealed largely the same results.

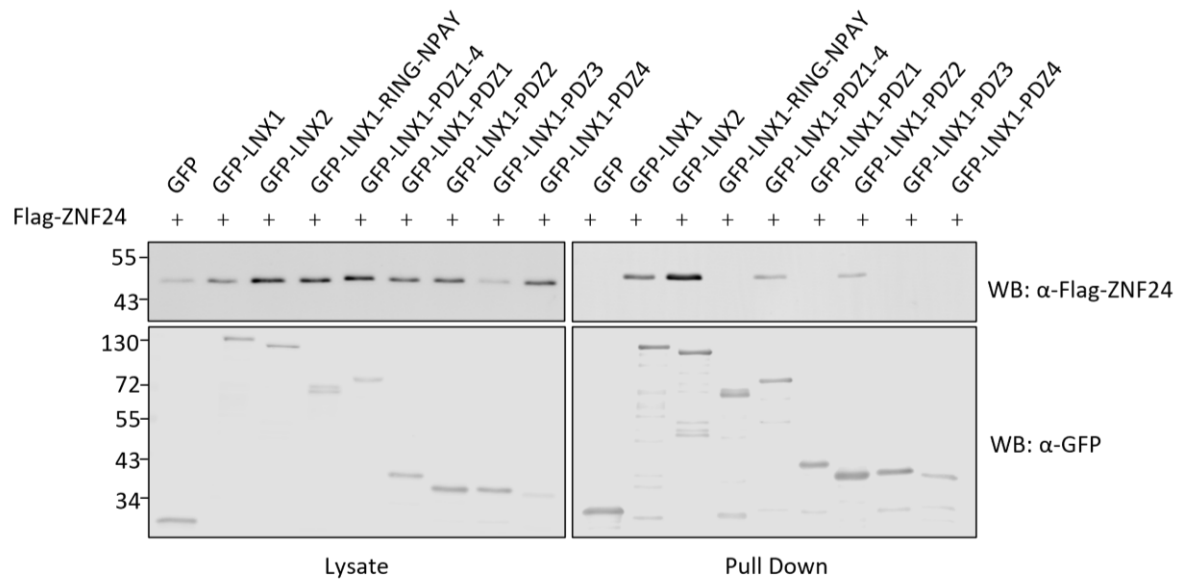


Figure 3.8: The ZNF24-LNX1 interaction can be mapped to PDZ2. Flag-ZNF24 and LNX1 domain specific interaction confirmed using GFP-trap assay of GFP fused LNX1 and LNX2 as well as individual PDZ domains of LNX1 and truncation mutants or the GFP moiety alone as a control. Indicated proteins were co-transfected into HEK 293T cells with Flag tagged ZNF24 protein and the lysate (input sample, left) was mixed with 5uL Chromotek GFP-Trap beads per condition and washed extensively. The pull-down fraction (right) was eluted using 2x Laemmli buffer and boiling and analysed by Western blot.

3.2.5.3 ZNF24 is a ubiquitin substrate of LNX1

An examination on whether LNX1^{p80} was able to ubiquitinate ZNF24, was carried out to supplement its characterisation as an interactor (Figure 3.9). For this purpose, GFP-ZNF24, HA-tagged ubiquitin and WT or RING mutant (C48A) LNX1^{p80} were co-expressed in HEK 293T cells. 24 hours post transfection; cells were treated for a further 6 hours with the potent proteasome inhibitor MG132 prior to harvesting. To detect ubiquitination on ZNF24 specifically, rather than on non-covalent interacting proteins, the cell lysates were boiled in the presence of SDS to disrupt protein-protein interactions and then immunoprecipitated with anti-HA using protein A/G beads. Equal amounts of protein

from cell lysates and A/G beads were used for each condition which can be seen by the β -actin signal.

To quantify the ubiquitination status of ZNF24, densitometry was used to determine the ubiquitination smear normalised to the level of protein in the input. A high molecular weight smear corresponding to ubiquitinated ZNF24 is seen in the immunoprecipitated (IP) fraction for WT and mutant LNX1^{p80}, but not in the negative control lacking an E3 ligase (Figure 3.9). The ability of the RING mutant to ubiquitinate ZNF24 could provide evidence of LNX1's ability to indirectly enhance ubiquitination which has been previously observed by Lenihan *et al.*, when the RING domain lacking isoform, LNX1^{p70}, also displayed ubiquitination signatures on Liprin- α 1 and KLHL11 (Lenihan, Saha and Young, 2017). The authors speculated that LNX1^{p70} could be acting as a scaffold protein allowing enhancement of target protein ubiquitination independently of its catalytic domain.

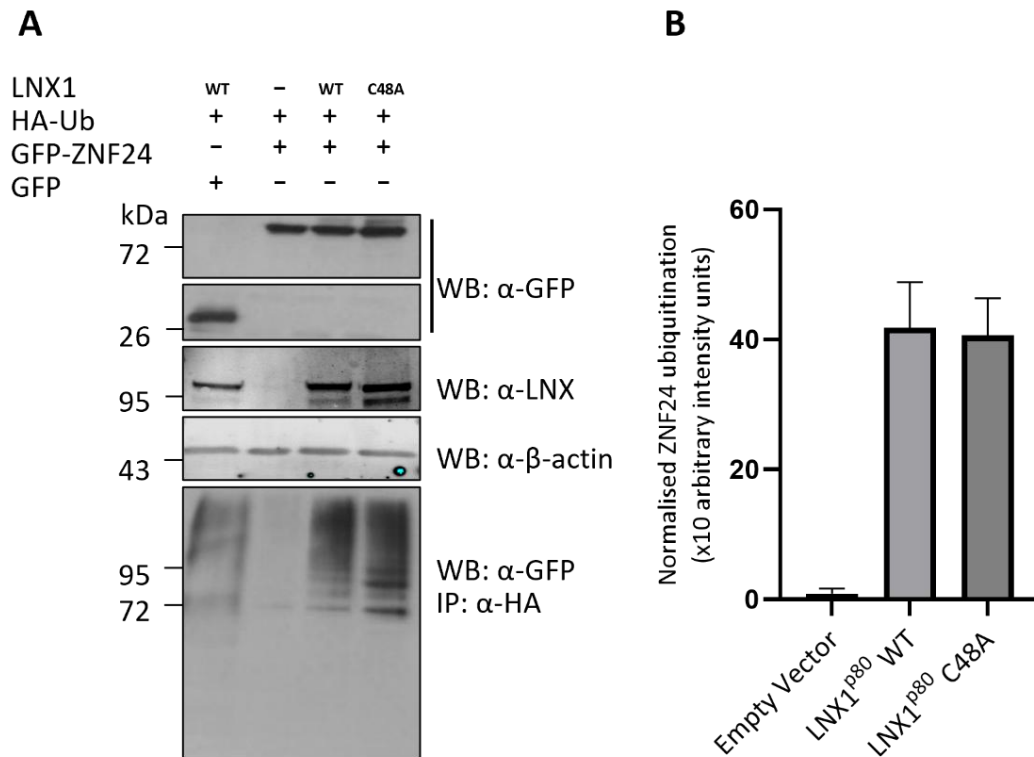


Figure 3.9: ZNF24 is a substrate for LNK1-mediated ubiquitination. HEK 293T cells were transiently transfected with the indicated constructs. 24 hrs post-transfection cells were treated with 10 μ m of the proteasomal inhibitor MG132 for 6 h. Under stringent SDS-denaturing conditions, cell lysates were then prepared and ubiquitinated proteins were immunoprecipitated from cell lysates using an anti-HA antibody. The ubiquitination status of ZNF24 was then revealed by Western blot analysis using an anti-Flag antibody. **(A)** A smear corresponding to ubiquitinated ZNF24 was detected, at varying levels, by Western blot (WB) for the wild-type and mutant LNK1 proteins. GFP alone was used to screen non-specific ubiquitination by LNK1 WT. Western blot of whole cell lysates confirmed expression of all constructs. **(B)** Levels of ubiquitinated ZNF24 (GFP ubiquitination was considered as “noise” or non-specific protein signals and subtracted) were quantified by densitometry and normalised for ZNF24/ β -actin expression in whole cell lysates. Data are expressed as standard deviation SD. Blot is representative of two independent experiments.

3.3 Discussion

3.3.1 The importance of verifying protein domain interactions

The availability of PDZ interactome databases has been vital for enriching what is known about LNX proteins. The ProfAff (protein affinities) database screened 512 PDZ binding motifs against each individual PDZ domain of LNX1 and LNX2 and found over 200 carboxyl terminal interactors with a pKd value above the threshold (Gogl *et al.*, 2022). This enormous number of interactors is intriguing considering how tightly regulated LNX protein expression is (Lenihan *et al.*, 2014). Determining whether interactions are physiologically relevant or just circumstantial is another major challenge.

As ProfAff only assigns dissociation affinities based on fragmentomic hold-up assays, one must be cautious on how one interprets these results. Similarly, machine learning predictions are excellent for quickly screening thousands of proteins, but are not fully comprehensive (Zhang *et al.*, 2025). From the list of twenty-four LNX1/2 protein interactors that were screened by yeast-two hybrid assays, only 12 were confirmed. Again, these results are not definitive as the yeast-two hybrid system has limitations. Firstly, yeast cells have different molecular machinery to mammalian cells which may not enable proper post-translational processing of proteins and allow correct folding (Koegl and Uetz, 2007). Secondly, yeast do not have the same subcellular structures such as neuronal membrane and synapses which would be the natural location of some of the proteins. Lastly, the yeast-two hybrid screen only accounts for the final 10 amino acids of the bait protein of interest.

The most promising proteins from the yeast-two hybrid analysis were brought forward for full length analysis in a mammalian cellular system. Several proteins that interacted with LNX1/LNX2 in the Y2H screen could not be expressed in HEK 293T cells despite many efforts – AMIGO1, GRIA2 and IGSF11. Endogenously, these transmembrane proteins are preferentially expressed in neurons and are essential to neuronal development (Peltola *et al.*, 2016)(Salpietro *et al.*, 2019)(Jang *et al.*, 2016). Therefore, in embryonic kidney cells their expression may not be favoured, for example expressing a single subunit of the AMPA receptor (GRIA2) may not be stable. Considering the behavioural defects observed in *Ln timer*^{p70} and *Ln timer* knockout mice (Lenihan *et al.*, 2017), it would be beneficial to continue the study of these neuronal proteins' interactions with LNX, especially as GRIA2 defects have been associated with neurodevelopmental disorders (Salpietro *et al.*, 2019). ZNF24, BCR and NRXN1 β showed the same result in the GFP trap assay as in the yeast-two hybrid screen, interacting with both LNX1 and LNX2. It is important to note that the interaction between BCR or ZNF24 and full length LNX2 was never identified previously, neither was LNX1's ability to ubiquitinate ZNF24. In future studies, specific ubiquitin linkages modulated by LNX1 could be determined by employing anti-K48/K63/K6 etc. antibodies. Determining which ubiquitin molecule is modified during ubiquitination could suggest whether ZNF24 is degraded by LNX1 or not. Interestingly, SDK2 a synaptic protein (Yamagata, 2020), showed an affinity for LNX1 and LNX2 but when subjected to full length co-immunoprecipitation analysis LNX2 could interact more convincingly. The LNX2 interactome is less characterised than LNX1 (Wolting *et al.*, 2011) (Guo *et al.*, 2012)

(Lenihan *et al.*, 2017) (Lenihan, Saha and Young, 2017), so an interaction specific to LNX2 could reveal a valuable function in the context of neuronal development.

Confirmation of the *in vivo* significance of interactions typically demands co-localization of the two proteins in the same cellular compartment which we confirm in the case of ZNF24 and LNX1^{p80} in the nucleus and BCR with LNX1^{p80} and LNX2 in the cytoplasm. Regulation of actin stress fibres in keratinocytes due to the GTPase activity BCR has on the regulation of RhoA, has been observed (Dubash *et al.*, 2013). This could explain the filamentous patterning when BCR is overexpressed in HeLa cells (Figure 3.5E). Both SDK2 and LNX2 have altered patterning upon co-expression that isn't seen for LNX1, validating the LNX2 specific interaction. It is noteworthy to consider that antibodies to detect these proteins endogenously were unavailable so transient overexpression was utilized, a fact which could be seen as artificial. The EPHB2-LNX1 interaction was identified in mouse hippocampal neurons which would explain the absence of subcellular co-localization in HeLa cells (Liu *et al.*, 2018). The novel interaction of EPHB2 and LNX2 would be compelling to investigate further as *EphB2* KO mice display depression-like behaviours and impairment in sociability (Zhen *et al.*, 2018), which is comparable to the phenotypes observed in *Ln timer* KO mice (Liu *et al.*, 2019).

3.3.2 Non-traditional methods of identifying LNX interactors

The biotechnology industry has always been dependent on the use of yeast from the brewing industry to the production of heterologous proteins for pharmacological advancements (Nielsen, 2025). *S. cerevisiae* is the mainly involved in these processes as it is fast growing, easy to modify genetically and generally recognised as safe (GRAS) by

the FDA (“Microorganisms & Microbial-Derived Ingredients Used in Food (Partial List) | FDA,” date accessed 11/02/2026). Yeast are frequently used in research to benefit protein expression but occasionally heterologous proteins fail to enter the secretory pathway to allow easy purification. A novel toolkit – MoClo, has recently been developed to modify foreign proteins to be secreted and displayed on the surface of yeast cells (Lee *et al.*, 2015) (O’Riordan *et al.*, 2024). We utilised this toolkit to achieve surface display of an anti-GFP nanobody (Kubala *et al.*, 2010) as an alternative to expensive, commercially available GFP-Trap agarose beads (Juric *et al.*, 2026 manuscript in preparation), in which we found that 1×10^8 yeast cells was equivalent to 1.2uL of GFP-Trap beads.

The anti-GFP nanobody yeast were employed to ascertain which PDZ domain of LNX1 is required for ZNF24 binding. LNX1 was fused to a GFP and immunoprecipitated by the anti-GFP nanobody yeast. The nanobody pull down was first validated using known LNX1 interactors which showed clear results. Liprin- α 1 is a well characterised interactor of LNX1 (Lenihan, Saha and Young, 2017) which gave a strong result when analysed by flow cytometry with an anti-Flag signal of 30% (Figure 3.6B). The NUMB binding domain is mapped to the NPAY motif of LNX1 (Dho *et al.*, 1998). When subjected to the anti-GFP nanobody pull down with LNX1, NUMB was less efficiently immunoprecipitated with an anti-Flag signal of 10%. The low binding of NUMB to LNX1 in this system could be due to the large GFP tag on the N-terminal of LNX1 effectively “hiding” or covering the short 4 amino acid NPAY binding motif. When full length ZNF24 was subjected to binding with full length LNX1, LNX2 and truncated protein versions in Figure 3.7, it can be determined that the interaction between LNX2 and ZNF24 is less robust (which could reinforce the

lack of co-localisation), and successful co-immunoprecipitation could be an artefact of overexpressing the two proteins. LNX1 PDZ2 gave the highest pull down of Flag-ZNF24 with a positive Flag signal of ~6.5%, which could be interpreted as low, but confirmation of this interaction was seen very clearly by Western blot, using commercial GFP-Trap beads (Figure 3.8). The effectiveness of the novel anti-GFP nanobody displaying yeast is yet to be fully validated however, one could argue the low cost of producing these yeast cells could outweigh the any potential loss of sensitivity providing good controls are used.

Chapter 4

Novel LNX2 Substrates Revealed Through Proximity-Dependent Biotin-Labeling of Substrates and Proteomic Analysis

4.1 Introduction

A considerable number of proteins have been reported to interact with LNX1 and LNX2; however, the list of experimentally confirmed substrates is markedly smaller, particularly in the case of LNX2 (Guo *et al.*, 2012) (Lenihan, Saha and Young, 2017) (Young, 2018). While LNX1 has been more thoroughly characterized and widely studied within the research community, LNX2 remains relatively underexplored.

NUMB p72 and p66 proteins are substrates of both LNX1 and LNX2 mediated ubiquitination (Dho *et al.*, 1998) (Nie *et al.*, 2002a) (Nie, Li and McGlade, 2004) (Won, Ro and Dawid, 2015). LNX proteins can target NUMB for proteasomal degradation, which in turn has been associated with influencing NOTCH signalling. Luciferase reporter assays show an increase of ~30% in NOTCH signalling upon overexpression of LNX1 by quantifying the activity of *Hes-1*, a NOTCH target gene (Nie *et al.*, 2002a). Regarding osteoclastogenesis, *Hes1* expression is decreased upon LNX2 knockout in bone-derived macrophages, indicating a decrease in NOTCH2 activity (Zhou *et al.*, 2015). In the same study the authors determine that LNX2 ubiquitination activity is essential in regulating M-CSF/RANKL signalling in the formation of osteoclasts. *Lnx2b* (LNX2 ortholog in fish, birds, marsupials and amphibians (Flynn, Saha and Young, 2011)) targets BOZ (bozozok) protein for degradation. BOZ is essential to the correct dorso-ventral patterning in zebrafish embryos. When *Lnx2b* was overexpressed in zebrafish, there were defects in the forebrain and incorrect neural patterning, consistent with BOZ knockout zebrafish (Ro and Dawid, 2009). From these studies one can be convinced that LNX proteins are critical for normal development and cellular function. However, LNX2 is only known to

play a part in very few pathways, simply because LNX1 has been of more interest to researchers and LNX2 is understudied.

To address this gap in knowledge, we employed a modified proximity-dependent biotinylation technique known as BioE3 (Barroso-Gomila *et al.*, 2023), which had been shown to enable the identification of direct substrates of RNF4, RNF214, MARCH5 and NEDD4 E3 ubiquitin ligases within a cellular context. Another group put forward their Ub-POD proximity labelling method of identifying ubiquitination substrates of RAD18 and TRAF6, only a few months later (Mukhopadhyay *et al.*, 2024). Both methods employ almost identical approaches – fusing BirA, a biotin ligase, to the E3 ligase of interest. Barroso-Gomila *et al.* utilised extreme wash conditions including a wash with 8M urea. Mukhopadhyay *et al.*, also use harsh washes but less so than the BioE3 paper, and still show very specific results. These papers supplement what is already known about well expressed E3 ligases. However, LNX2 is not a well expressed or characterised enzyme (Lenihan *et al.*, 2014). In this chapter, we overcome this challenge by implementing both BioE3/Ub-POD methods but with a slightly different approach in terms of set up.

The BioE3/Ub-POD approach that I employed involves generating a fusion construct in which BirA, a biotin ligase, is covalently linked by a GSQ sequence to the E3 ubiquitin ligase of interest - in this instance, LNX2. Cells are co-transfected with another construct encoding an Avi-tagged ubiquitin, which serves as a substrate for BirA-mediated biotinylation. BirA-Avi Tag technology and has been used widely across molecular biology and is beneficial for specific labelling of targets (Fairhead and Howarth, 2015). When BirA-LNX2 catalyses the ubiquitination of target proteins, the proximity between the E3

ligase and its substrates permits the biotinylation of the Avi-tag on ubiquitin (Figure 4.1). This selective biotin labelling effectively marks bona fide LNX2 substrates within the cellular environment. The biotinylated proteins can be enriched *via* neutravidin pull-down and identified by mass spectrometry, providing a means of uncovering authentic LNX2 substrate profiles.

Aims of Chapter 4

- Optimise and validate novel, proximity-based biotinylation methods for identifying ubiquitin substrates for use on LNX2.
- Discover LNX2 ubiquitin substrates by proteomic mass spectroscopy and annotate gene molecular function and cellular compartment.
- Further investigate these candidate proteins as interactors and/or substrates of LNX2.

4.2 Results

4.2.1 Modifying the BioE3/Ub-POD method

As mentioned in section 4.1, the methods of proximity labelling of ubiquitin substrates proposed by Barroso-Gomila *et al.* and Mukhopadhyay *et al.* have seen successful identification of targets from their respective well known E3 ligases (Barroso-Gomila *et al.*, 2023) (Mukhopadhyay *et al.*, 2024). In the interest of time, we opted to modify the plasmids - TRIPZ-bioGEF-UBnc-PURO and Lenti-EFS-BirAopt-GSQ-RBXN-P2A-BLAST which were gifts from Rosa Barrio & James Sutherland as per methods section 2.8.1, so that we could utilise transient transfections instead of creating stably expressing cell lines. BirA tagged E3 ligases; LNX1 and LNX2 along with catalytic mutants in the RING domain LNX1 C48A and LNX2 C53A were created. Our focus was directed on characterising the ubiquitinome of LNX2, but we created and validated the same system for LNX1, as a future project.

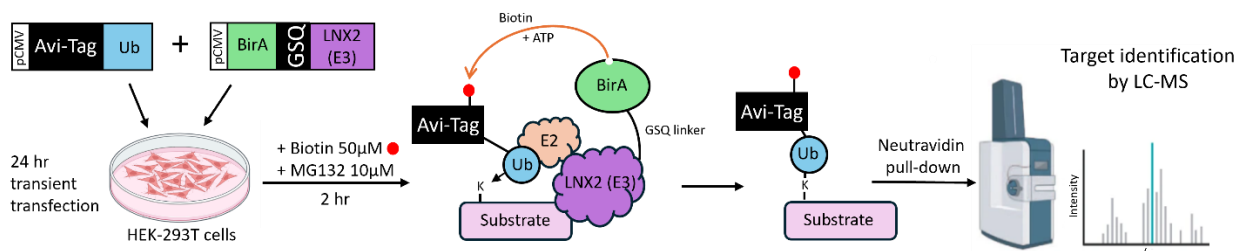


Figure 4.1: Workflow schematic of the modified BioE3/Ub-POD method. Transient transfection of pCMV-AviTag-Ubiquitin and pCMV-BirA-GSQlinker-LNX2 into HEK 293T cells for 24 hours followed by biotin labelling and MG132 treatment for 2 hours. A process of biotinylating the Avi tagged substrates occurs as pictured. Only the artificial ubiquitin that has been ligated to the substrate by the BirA-LNX2 E3 ligase should be biotinylated. The targeted substrates are purified from lysate using NeutrAvidin beads and sent for LC-MS. Image created using BioRender.

We then tested our modified constructs, experimental controls and compared the ubiquitination ability of the catalytic inactivating mutants (LNx1 C48A and LNx2 C53A) *in vitro*. Also included, was a BirA enzyme empty control that was not shown in Barroso-Gomila *et al.* We also acknowledged the presence of abundant naturally biotinylated proteins – Acetyl-CoA carboxylase alpha (ACACA, 114kDa), Acetyl-CoA carboxylase 2 (ACACB, 276kDa), Pyruvate carboxylase (PC, 129kDa) and Propionyl-CoA carboxylase alpha chain (PCCA, 80kDa) that were present during biotin detection (Dumrongprechachan *et al.*, 2021). The expression of endogenously biotinylated proteins does not increase significantly upon biotin treatment of HEK 293T cells (Sun *et al.*, 2022).

From analysis of BirA/Ub expressing HEK 293T whole cell lysate, it was evident that a lot of background biotin signal was produced by the unfused BirA enzyme. In the empty vector and BioGEF-Ub only controls only endogenously biotinylated proteins appeared, as expected. The BirA fused LNx1/2 provided a detectable smear corresponding to ubiquitinated and biotinylated substrates and was less obvious in the RING mutants (Figure 4.2). Equal amounts of protein for each condition were examined, and LNx1/2 proteins themselves were expressed at the correct size allowing us to quantify and compare the levels of biotinylated material.

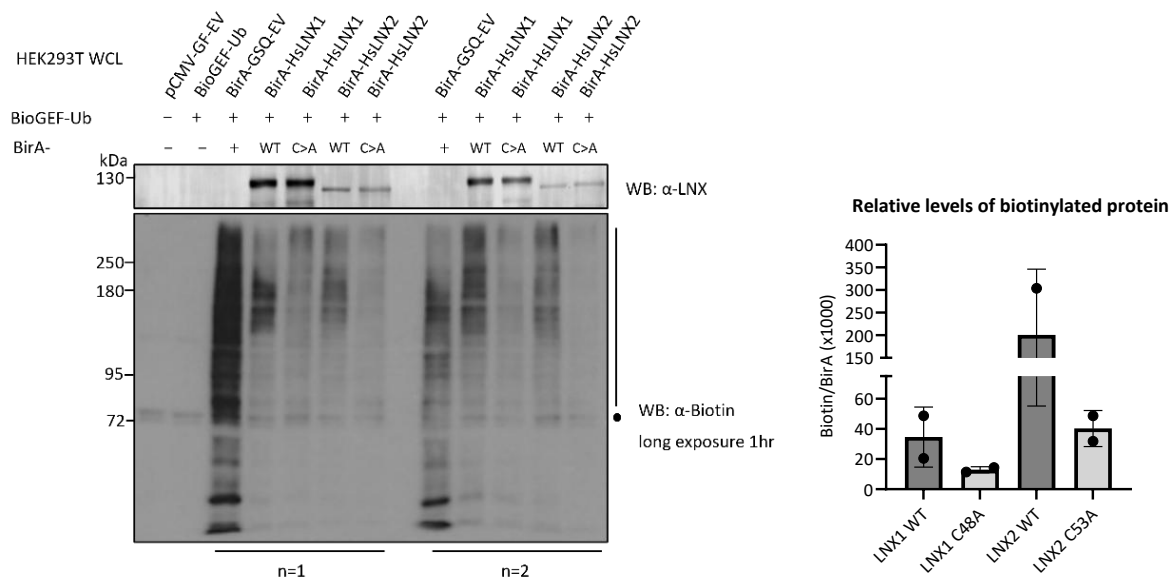


Figure 4.2: Validation of modified BioE3 method and determining catalytic activity of LNX proteins with a functional or mutant catalytic RING domain. HEK 293T cells were transfected with equal amounts of indicated plasmids and 24 hours after transfection cells were treated with 50 μ M biotin and 10 μ M MG132 for 2 hours, to block proteasomal activity. Cells were lysed in WB5 (8 M urea, 1% SDS, 1x PI, 50 μ M NEM in 1x PBS) and protein quantity was determined before Western blot analysis. Whole cell lysate (no pull down preformed) of LNX1/2 WT and mutants were probed for biotin and LNX using indicated antibodies (Chapter 2, table 2.6). Smears representing biotinylated proteins were observed in all BirA-LNX transfected samples but both LNX1 C48A and LNX2 C53A samples had less intense smears. HEK 293T cells expressing the empty BirA vector showed intense non-specific smearing, but only endogenous biotinylated material was detected in the pCMV-GF-empty and BioGEF-Ub-nc lysates. Smear intensity was determined by ImageJ and normalised against LNX-BirA expression; n=2, error bars represent SEM.

Before carrying out large scale experiments, the effectiveness of the NeutrAvidin - biotin pull down was assessed. We exploited the high expression of naturally occurring biotinylated proteins (ACACA, ACACB, PCCA and PC) to investigate how efficient biotin immunoprecipitation would be in the wash buffers described by Barroso-Gomila *et al.* HEK 293T cells were lysed in WB5 and protein content was subjected to NeutrAvidin binding overnight. The binding affinity was examined after each wash step and as seen

in Figure 4.3, and very little biotinylated material remained attached to the beads after the final wash in WB6. The vast majority of endogenous biotinylated proteins remained in the unbound fraction, suggesting that 8M urea and 1% SDS is not a suitable binding buffer. When a small-scale trial to detect the biotinylated substrates of BirA-LNX2 was carried out in Figure S2, no enrichment in the pull-down fraction was observed and most of the material remained unbound.

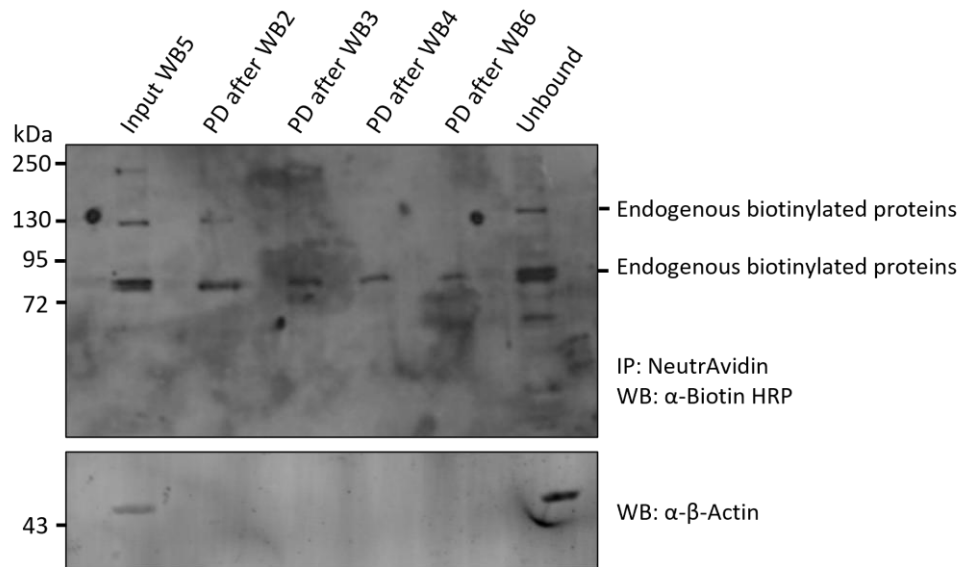


Figure 4.3: Low NeutrAvidin binding efficiency of endogenously biotinylated HEK 293T proteins in stringent washes. HEK 293T cells were lysed in WB5 (8 M urea, 1% SDS, 1x PI, 50 μ M NEM in 1x PBS) and 20 μ L of NeutrAvidin beads were added and allowed to interact overnight at room temperature. An unbound sample was taken before removing the supernatant and washing with WB2 (6 M Guanidine-HCl), WB3 (6.4 M urea, 1 M NaCl, 0.2% SDS), WB4 (4 M urea, 1 M NaCl, 10% isopropanol, 10% ethanol, and 0.2% SDS) and finally WB6 (2% SDS). A 5 μ L sample of beads was taken after each washing set and eluted by 2 rounds of boiling and cooling in 4X Laemmli buffer supplemented with 15 mM biotin. Blocking and immunoblotting was carried out in 1% BSA in TBST and detected using ECL. β -actin was only present in the input and unbound fractions as it is not naturally biotinylated.

4.2.2 Proteomic identification of LNX2 substrates in HEK 293T cells

After unsuccessful biotin pull down attempts using extremely stringent buffers, a review of previously published literature regarding proximity dependent biotinylation was undertaken. Two methods papers describing the use of APEX2 peroxidase to covalently tag proximal endogenous proteins with biotin were useful (Hung *et al.*, 2016) (Tan *et al.*, 2020). The overall goal of these papers was like ours – perform purifications of biotinylated proteins using NeutrAvidin beads. The wash steps from these methods were described in detail and were easily adopted to our own protocol as described in methods section 2.8.2. In Figure 4.4, we employ these new wash strategies, and we can see strong enrichment of biotinylated material in the pull-down fraction compared to the input. The dark smears indicate ubiquitinated substrates of various molecular weights with possibly, a range of different ubiquitination signatures. The lighter smear (Figure 4.4, right) corresponds to material ubiquitinated and subsequently biotinylated by the catalytic inactivating mutant, LNX2 C53A. There is evidently less biotinylating/ubiquitinating activity, but it could suggest that a single cysteine to alanine residue in the RING domain is not sufficient to completely halt ubiquitin ligating activity (Lenihan *et al.*, 2017).

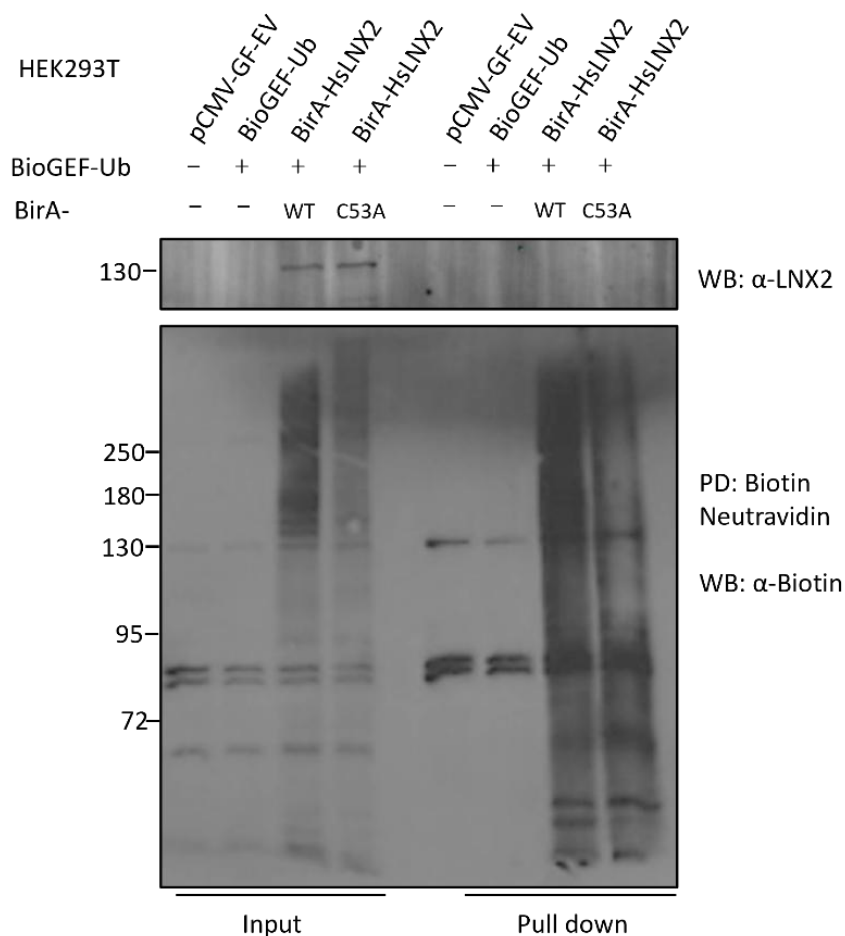


Figure 4.4: LNX2 ubiquitin substrates identified using modified proximity dependent biotin labelling method and biotin affinity purifications. Two 15cm² plates of HEK 293T cells were co-transfected with BioGFF-Ub along with BirA tagged LNX2 WT or mutant plasmids. HEK 293T cells were precultured using DMEM supplemented with 10% FBS (that had been dialysed to remove biotin). 24 hours after transfection, 50 μ M biotin was added back along with 10 μ M MG132 for 2 hours. Cells were washed three times in ice cold 1x PBS to remove excess biotin and lysed in RIPA buffer, sonicated and cleared by centrifugation. Protein concentrations were quantified (8,500 μ g protein for each condition) before diluting lysates 4-fold to decrease detergent content. 50 μ L of NeutrAvidin beads were added to each sample and allowed to bind for 4 hours before extensive washing steps were performed as per methods section 2.8.2. After the last wash step, beads were rinsed 5 times in 1 x PBS and a 5 μ L sample of beads was taken for testing by Western blot analysis (above). Enrichment of biotinylated material is seen in the pull-down fraction and only endogenous biotinylated proteins are evident as background noise in the negative control lanes. Blot is a representative of three independent experiments.

After ensuring experimental success of the LNX2 substrate proximity labelling in figure 4.4, we sent the remaining NeutrAvidin beads that were bound to biotinylated material to the LC-MS facility in Kings College London for proteomic identification. Peptides were treated with trypsin and fragments processed by Neo UHPLC coupled to an Orbitrap Astral mass spectrometer (Methods section 2.8.3). The raw data spectra were sorted into peak list files using proteome discoverer and searched against the Uniprot Human Taxonomy database (Taxon ID: 9606).

From three biological replicates of WT LNX2, mutant LNX2 and negative control processed by mass spectrometry, 1,293 proteins were identified. Manual data processing was required to filter out contaminants like keratins, albumin, endogenous biotinylated proteins (PC, PCCA and ACACA) and other abundant interferences. Once these were filtered out, we compiled the total amount of spectra of each peptide from each condition across the three replicates, and excluded proteins that were only detected in the negative control or LNX2 mutant samples as these were classified as non-specific hits. We then eliminated any proteins identified based on only one unique peptide. The remaining proteins were ranked based on the ratio of peptides detected in the WT versus mutant samples and those with a WT: Mutant ratio of less than 5 were eliminated. The list of 96 peptides identified (Table 4.1) were considered as putative LNX2 ubiquitination substrates. The list is considerably shorter than the original list, but thorough filtering ensures only credible proteins are further investigated as substrates.

Table 4.1: Mass spectrometry identified LNX2 substrates from proximity dependent biotinylation assays. Carboxyl-terminal PBMs are underlined. PDZ binding motifs were predicted by Pfam/InterPro (EMBL-EBI) (Madeira *et al.*, 2024). Protein list is first sorted by ratio of the number of peptides detected for WT versus mutant LNX2 and then by total number of peptides. Asterisk (*) indicates proteins selected for further validation.

Protein symbol	Protein name	Carboxyl terminus	WT: Mutant ratio
TAB1	TGF-beta-activated kinase 1 and MAP3K7-binding protein 1	GEQSVVTAP	24:0
*KARS1	Lysine--tRNA ligase	LESTTVGTSV	15:0
MAP3K7	Mitogen-activated protein kinase kinase kinase 7	SQQQKRQGT	15:0
TAB2	TGF-beta-activated kinase 1 and MAP3K7-binding protein 2	CEQCEMPRHF	13:0
BABAM2	BRISC and BRCA1-A complex member 2	QEAAFANGKL	13:0
*FLOT2	Flotillin-2	LIKATGVQV	7:0
*SRR	Serine racemase	RPASYQSVSV	7:0
NCAPG	Condensin complex subunit 3	NLAQFLNEDLS	7:0
*EXOC4	Exocyst complex component 4	ATKDKKITTV	6:0
UBAC2	Ubiquitin-associated domain-containing protein 2	NVATNFLLQH	6:0
RBBP4	Histone-binding protein RBBP4	GSVDPEGQGS	6:0
*TRIP6	Thyroid receptor-interacting protein 6	ELSATVTDC	5:0
MYO6	Unconventional myosin-VI	ATAMLQSLLK	5:0
BRCC3	Lys-63-specific deubiquitinase BRCC36	ELMQELSSLE	5:0
ILVBL	2-hydroxyacyl-CoA lyase 2	TDFRDGSI	4:0
PPP3CA	Protein phosphatase 3 catalytic subunit alpha	SNGSNSSNIQ	4:0
GLN3	Guanine nucleotide-binding protein-like 3	DAYDFSTDYV	4:0
SEC62	Translocation protein SEC62	ETPKSSHEKS	4:0
DCUN1D1	DCN1-like protein 1	QIAGTKSTTV	4:0
HBS1L	HBS1-like protein	AAGVVTEIKE	4:0
*USO1	General vesicular transport factor p115	EDPGKDL	4:0
*RAB18	Ras-related protein Rab-18	GACGGYCSVL	4:0
*PAFAH1B3	Platelet-activating factor acetylhydrolase IB subunit alpha1	QGAPLLEPAP	4:0
ATP1B3	Sodium/potassium-transporting ATPase subunit beta-3	RVMFKITARA	4:0
SLC12A2	Solute carrier family 12 member 2	NHQSVLTFYS	4:0
CNN3	Calponin-3	YQYSDQGIDY	4:0
KRT36	Keratin, type I cuticular Ha6	SREHVQSRPL	3:0
RTL8C	Retrotransposon Gag-like protein 8C	VFGWEEDEDF	3:0
TUBG1	Tubulin gamma-1 chain	DYISWGTQEQ	3:0
BRAP	BRCA1-associated protein	SRKGRSKRGK	3:0
MSH6	DNA mismatch repair protein Msh6	HKLLTLIKEL	3:0
SQSTM1	Sequestosome-1	IQYSKHPPPL	3:0
SRP68	Signal recognition particle subunit SRP68	YIKGIFGFRS	3:0
MINDY3	Ubiquitin carboxyl-terminal hydrolase MINDY-3	WTTDRSPSLN	3:0
COPG1	Coatomer subunit gamma-1	PVDIILASVG	3:0

UBL4A	Ubiquitin-like protein 4A	TETMEKGF ^{SK}	3:0
RFC2	Replication factor C subunit 2	CQKTMAPVAS	3:0
BABAM1	BRISC and BRCA1-A complex member 1	DEAIEVEATV	3:0
TNRC6A	Trinucleotide repeat-containing gene 6A protein	DHLGGGGESM	3:0
FBXO45	F-box/SPRY domain-containing protein 1	LVYLGKPLDG	3:0
RFC4	Replication factor C subunit 4	TVMQQLSQNC	3:0
NOP2	28S rRNA (cytosine(4447)-C(5))-methyltransferase	SRGNSQLLLS	3:0
*PKN2	Serine/threonine-protein kinase N2	RDFDYIADWC	2:0
*ZNF24	Zinc finger protein 24	AEKLLNVVKV	2:0
*KPNA1	Importin subunit alpha-5	CEAPMEGF ^{QL}	2:0
ZFYVE16	Zinc finger FYVE domain-containing protein 16	LVFFIIEHLF	2:0
UBE2D2	Ubiquitin-conjugating enzyme E2 D2	AREWTQKYAM	2:0
OS9	Protein OS-9	GTGDLDEF ^{DF}	2:0
DNAJB2	DnaJ homolog subfamily B member 2	EEKASRCLIL	2:0
NUDT5	ADP-sugar pyrophosphatase	KPF ^{FE} VVPFL ^{KF}	2:0
SLC7A5	Large neutral amino acids transporter small subunit 1	KLMQVVPQET	2:0
KLHL9	Kelch-like protein 9	SPLSAPSDHS	2:0
TRIP12	E3 ubiquitin-protein ligase TRIP12	REGQQSFHLS	2:0
EIF3M	Eukaryotic translation initiation factor 3 subunit M	KNSLLSLSDT	2:0
EIF4G2	Eukaryotic translation initiation factor 4 gamma 2	EEEESEEEAD	2:0
*HEATR1	HEAT repeat-containing protein 1	VLGEPLQSYF	2:0
SUPT5H	Transcription elongation factor SPT5	LRFLGKLLEA	2:0
NKAP	NF-kappa-B-activating protein	YRKTKGKDDK	2:0
DCAKD	Dephospho-CoA kinase domain-containing protein	LLTHYLLPYA	2:0
*ABRAXAS2	BRISC complex subunit Abraxas 2	DPRNTQT ^{SQI}	2:0
EIF5B	Eukaryotic translation initiation factor 5B	VELKKVFEII	2:0
*CNP	2',3'-cyclic-nucleotide 3'-phosphodiesterase	GGALQ ^{SCTII}	2:0
MATR3	Matrin-3	AEERRQKKET	2:0
UBFD1	Ubiquitin domain-containing protein UBFD1	DTVLGK ^{WQYF}	2:0
TIPRL	TIP41-like protein	ADSQKSTQVE	2:0
PTPN11	Tyrosine-protein phosphatase non-receptor type 11	GLMQQ ^{QKSFR}	2:0
AZIN1	Antizyme inhibitor 1	SQEDSFSAEA	2:0
RNF122	RING finger protein 122	NIGILLDELV	2:0
NDUFA10	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial	EVGDKWIWLK	2:0
SAFB	Scaffold attachment factor B1	PSDARFTRRY	2:0
SMAD2	Mothers against decapentaplegic homolog 2	SPSVRCSSMS	2:0
NUP85	Nuclear pore complex protein Nup85	IIREGSLEGS	2:0
RWDD4	RWD domain-containing protein 4	LSKTGSKDDE	2:0
RPS19	Small ribosomal subunit protein eS19	GQVAAANKKH	2:0
RFC3	Replication factor C subunit 3	MEDGLE ^{GMMF}	2:0
RTN4	Reticulon-4	KIPGLKRKAE	2:0
TMOD3	Tropomodulin-3	RKRRVEGDHQ	2:0
*FERMT2	Fermitin family homolog 2	MFYKLTSGWV	27:2
*KLHL11	Kelch-like protein 11	RRVPSSQ ^{IEC}	10:1

RBBP7	Histone-binding protein RBBP7	TTSELEGQGS	10:1
CDC34	Ubiquitin-conjugating enzyme E2 R1	DEDDSGTEES	9:1
*RHOA	Transforming protein RhoA	GKKKSGCLVL	9:1
SMC4	Structural maintenance of chromosomes protein 4	PKEIASKGLC	20:3
ZNF330	Zinc finger protein 330	GYAHYEEQEN	13:2
WRNIP1	ATPase WRNIP1	GVDFFKQRR	19:3
HECTD1	E3 ubiquitin-protein ligase HECTD1	ATMEKGFHLN	16:3
ANKRD13A	Ankyrin repeat domain-containing protein 13A	QVLQLSLTDK	10:2
NCAPH	Condensin complex subunit 2	LSDVLVRQGD	7:1
*TOM1	Target of Myb1 membrane trafficking protein	EKDDDMLEFAL	7:1
TTLL12	Tubulin--tyrosine ligase-like protein 12	PGGCHVTCLV	7:1
PFKM	ATP-dependent 6-phosphofructokinase, muscle type	TRKRSGEAAV	7:1
PLS3	Phospholipid scramblase 3	GGAGPSAVTS	6:1
UBE2R2	Ubiquitin-conjugating enzyme E2 R2	DDDDSGNEES	6:1
USP9X	Ubiquitin carboxyl-terminal hydrolase 9X	EVSPPTKQDQ	34:7
PAFAH1B2	Platelet-activating factor acetylhydrolase IB subunit alpha2	TPEEKQTTIA	5:1
GNAI3	Guanine nucleotide-binding protein G(i) subunit alpha-3	KNNLKECGLY	5:1

From the list in Table 4.1, some of our most confident peptide hits like KLHL11, EXOC4, SRR and ILVBL were previously identified as LNX2 PDZ2 interacting proteins, but were never investigated as being LNX2 substrates (Lenihan *et al.*, 2017). A search of the ProfAff database further revealed KARS1, PFKM and PKN2 as LNX2 PDZ2 interactors (Gogl *et al.*, 2022). Kelch-like protein 11 (KLHL11) functions as a substrate adaptor in the Cullin-3 E3 ligase complex and has been confirmed as a substrate of LNX1 (Lenihan, Saha and Young, 2017). In chapter 3 (Figure 3.9), I demonstrate ZNF24, the kruppel-like transcription factor that represses VEGF transcription in breast cancer (Jia *et al.*, 2015), as a LNX1 ubiquitin substrate.

Interestingly, known LNX2 substrates like NUMB and NUMB-like proteins were identified in the list of mass spectrometry hits. The amount of peptide fragments present across the three replicates for NUMB was 5 in LNX2 WT and 2 in LNX2 Mutant (data not shown).

The ratio of WT: Mutant was not high enough for our cutoff point. A similar result for LNX2 itself was observed from the list, with a WT (149) to Mutant (75) ratio of 1.9, again under our threshold. E3 ligases are known to mediate their own ubiquitination as a method of regulation and this is true for LNX1 (Vittal *et al.*, 2015) (Lenihan *et al.*, 2014). Our results from Chapter 5, RING: RING dimerization (Rice, Northcutt and Kurschner, 2001)) and here, could indicate LNX2 to be a self-regulating E3 ligase.

4.2.3 Functional Classification and grouping of novel LNX2 substrates by Gene Ontology (GO) enrichment

Gene Ontology was developed in a massively collaborative project, to provide easy access to annotation of gene function and cellular position in humans, mice, bacteria, yeast, plants, fruit flies, fish and nematodes (Harris *et al.*, 2007). ShinyGO v0.85.1 (Ge *et al.*, 2020) is a server that can quickly categorise a large set of genes into their biological function, cellular compartment and biological process and statistically determine how enriched a set of genes are in a particular group. Caution must be taken when interpreting GO enrichment results as false discovery rate (FDR) value (adjusted p-value) of 0.01 or 0.001 can often reflect noise due to massive numbers of terms tested. A more significant value of 1×10^{-4} or 1×10^{-5} is expected for gene lists of reasonable size and should be considered alongside the fold enrichment (FE) value. For reader simplicity I have underlined the proteins that appear in our table of identified substrates.

The list of identified putative LNX2 substrates in table 4.1 was subjected to GO enrichment analysis of cellular compartment (Table 4.2) and biological process (Table S1). The BRISC and BRCA1-a complexes function as deubiquitinating complexes in the

removal of K63-linked ubiquitin chains. Both complexes have an overall similar catalytic function with the DUB enzyme, BRCC36 at the core of both BRISC and BRCA1-A but carry out physiologically distinct tasks. BRCA1-A is involved in DNA repair by removing ubiquitin chromatin signals, sequestering BRCA1 away from the damage site and suppressing homologous recombination (Rabl, 2020). The BRISC complex consisting of five genes (RAP80, BRCC36, NBA1, BRCC45 and ABRAXAS1), is involved in control of mitosis by regulating Aurora B kinase activity by removing K63-linked chains (Qin Li *et al.*, 2022). From our list both complexes have high fold enrichment and significant p-values (Table 4.2) and could be suggestive of a regulatory role for LNX2 in these pathways.

DNA replication factor C (RFC) complex had a fold enrichment of 145 with three (RFC2, RFC3, RFC4) out of five components identified on our list. The RFC functions as a sliding clamp for tethering DNA polymerases like PCNA to the template during DNA replication, repair and modification (Majka and Burgers, 2004). Supportingly, these genes are enriched (FE: 80.75) in the GO biological processes of regulation of DNA-directed DNA polymerase activity and positive regulation of DNA-directed DNA polymerase activity (Table S1). Functionally similar complexes with impressive fold enrichment scores for our LNX2 substrate lists are the Condensin complex (FE: 90) and the Ctf18 rfc-like complex (FE: 90).

1-alkyl-2-acetylglycerophosphocholine esterase or PAF-AH complex is a potent platelet activating factor involved in inflammatory system regulation and contains catalytic subunits from our list, PAFAH1B3 and PAFAH1B2 (Prescott *et al.*, 2000). The PAF-AH complex along with myelin sheath abaxonal region complex (FE: 80.75 gene hits CNP and

EXOC4) have roles in neuronal development and stability as both have lipid membrane functions.

Table 4.2: Cellular compartment breakdown of LNX2 substrates using GO enrichment analysis. ShinyGo v0.85.1 (Ge *et al.*, 2020) was used to analyse the 96 protein hits from proteomic identification. Searching stringency set to 0.05 FDR cutoff, 50 pathways with redundancy removed, selected by FDR and then by fold enrichment. Only cellular compartments with a significant *p* value < 0.05, are shown here.

<i>p</i> -value	No. of genes	Genes in pathway	Fold enrichment	Cellular compartment	Genes
8.1x10 ⁻⁹	5	7	173.0506	GO:0070552 Brisc complex	BABAM1 BRCC36 BABAM2 ABRAXAS2
0.0000101	4	10	96.90833	GO:0070531 Brca1-a complex	BABAM1 BABAM2 BRCC36 ABRAXAS2
0.0000312	9	214	10.18896	GO:0030496 Midbody	GNAI3 PKN2 ZNF330 SAFB ABRAXAS2 KLHL9 RHOA TTLL12 EXOC4
0.0000335	41	4670	2.127003	GO:0005654 Nucleoplasm	TNRC6A HECTD1 RBBP7 BABAM1 KPNA1 TRIP12 RBBP4 SUPT5H DCUN1D1 GNAI3 PKN2 FERMT2 BRAP CDC34 TAB1 NKAP UBL4A RPS19 ZNF330 SMC4 MSH6 NCAPH NUP85 MINDY3 SAFB SQSTM1 RFC4 ZNF24 BRCC3 MYO6 RFC2 TAB2 KARS1 NOP2 HEATR1 UBE2D2 RFC3 PPP3CA BABAM2 SMAD2 PTPN11
0.0000489	3	5	145.3625	GO:0005663 Dna replication factor c complex	RFC2 RFC3 RFC4
0.0001939	3	8	90.85156	GO:0000796 Condensin complex	NCAPH NCAPG SMC4
0.0001939	3	8	90.85156	GO:0031390 Ctf18 rfc-like complex	RFC2 RFC3 RFC4
0.0020234	2	3	161.5139	GO:0008247 1-alkyl-2-acetylglucophosphocholine esterase complex	PAFAH1B3 PAFAH1B2
0.0020234	21	1990	2.556627	GO:1902494 Catalytic complex	DCUN1D1 GNAI3 KARS1 RBBP7 ILVBL NDUFA10 UBE2D2 PPP3CA PFKM RBBP4 FBXO45 KLHL9 BRAP MAP3K7 DNAJB2 BABAM2 RFC4 BRCC3 PAFAH1B3 PAFAH1B2 OS9
0.0030091	8	340	5.70049	GO:0000151 Ubiquitin ligase complex	DCUN1D1 UBE2D2 FBXO45 KLHL9 BRAP BABAM2 BRCC3 OS9
0.0069816	13	990	3.181334	GO:1990234 Transferase complex	DCUN1D1 RBBP7 ILVBL UBE2D2 PFKM RBBP4 FBXO45 KLHL9 BRAP MAP3K7 BABAM2 BRCC3 OS9
0.0075266	2	6	80.75694	GO:0035748 Myelin sheath abaxonal region	EXOC4 CNP

0.0077121	16	1450	2.673333	GO:0140513 Nuclear protein-containing complex	<i>RBBP7 BABAM1 MSH6 HEATR1 NUP85 BABAM2 RBBP4 SMAD2 BRCC3 SUPT5H RFC2 RFC3 RFC4 ABRAXAS2 KPNA1</i>
0.0088050	4	83	11.6757	GO:0000794 Condensed nuclear chromosome	<i>NCAPG SMC4 NCAPH TUBG1</i>
0.0088050	2	7	69.22024	GO:0016589 Nurf complex	<i>RBBP4 RBBP7</i>
0.0088050	13	1052	2.993841	GO:0140535 Intracellular protein-containing complex	<i>DCUN1D1 KARS1 EIF4G2 UBE2D2 PPP3CA FBXO45 KLHL9 BRAP MAP3K7 DNAJB2 BABAM2 BRCC3 OS9</i>
0.0124933	6	249	5.837851	GO:0000228 Nuclear chromosome	<i>NCAPG SMC4 NCAPH TUBG1 RBBP7 RBBP4</i>
0.0162662	12	996	2.918926	GO:0070161 Anchoring junction	<i>RHOA FERMT2 TRIP6 RPS19 EIF4G2 CNN3 FLOT2 TMOD3 PKN2 SRP68 RTN4 PPP3CA</i>
0.0173689	4	104	9.318109	GO:0005902 Microvillus	<i>SLC7A5 EXOC4 CNP MYO6</i>
0.0200108	8	501	3.868596	GO:0045177 Apical part of cell	<i>SLC12A2 SLC7A5 PFKM ATP1B3 RAB18 TUBG1 SRR MYO6</i>
0.0243757	5	199	6.087207	GO:0005912 Adherens junction	<i>EIF4G2 CNN3 TMOD3 FERMT2 FLOT2</i>
0.0243757	2	14	34.61012	GO:0031010 Iswi-type complex	<i>RBBP4 RBBP7</i>
0.0294450	8	556	3.485911	GO:0005635 Nuclear envelope	<i>NUP85 BRAP RTN4 DNAJB2 MINDY3 MYO6 MATR3 KPNA1</i>
0.0294450	8	551	3.517544	GO:0005911 Cell-cell junction	<i>EIF4G2 CNN3 TMOD3 PKN2 RHOA FERMT2 FLOT2 PPP3CA</i>
0.0294450	2	17	28.50245	GO:0016581 Nurd complex	<i>RBBP7 RBBP4</i>
0.0294450	2	17	28.50245	GO:0035098 Esc/e(z) complex	<i>RBBP7 RBBP4</i>
0.0307129	7	442	3.836868	GO:0098687 Chromosomal region	<i>RBBP7 ZNF330 WRNIP1 NUP85 RBBP4 NCAPG SMC4</i>
0.0412614	6	350	4.153214	GO:0031965 Nuclear membrane	<i>BRAP DNAJB2 MINDY3 MYO6 NUP85 MATR3</i>
0.0446634	3	75	9.690833	GO:0005657 Replication fork	<i>RFC2 RFC3 RFC4</i>
0.0469463	2	23	21.06703	GO:0070822 Sin3-type complex	<i>RBBP7 RBBP4</i>
0.0473869	7	490	3.461012	GO:0005925 Focal adhesion	<i>RHOA FERMT2 TRIP6 RPS19 CNN3 FLOT2 SRP68</i>

4.2.4 Examination of proteins identified by proteomics as LNX2 interactors

To be a ubiquitin substrate, there would usually be a demand for a physical interaction between the E3 ligase enzyme and target protein (Buetow and Huang, 2016), but not always. PDZ binding motif to PDZ domain interaction are the most common mode of LNX2 protein association, apart from the NPAF-PTB domain interaction with NUMB (Weiss *et al.*, 2007) (Zheng *et al.*, 2010)(Wolting *et al.*, 2011) (Guo *et al.*, 2012)(Lenihan

et al., 2017)(Lenihan, Saha and Young, 2017) (Gogl *et al.*, 2022). As many of the identified peptide hits are components of highly enriched protein complexes, we considered whether their association with LNX2 might be direct or indirect. Although the modified BioE3/Ub-POD assay is designed to specifically biotinylate LNX2 ubiquitin substrates, it is possible that proteins in proximity within tightly organized complexes could also become labelled due to their spatial association with a true substrate. Using the InterPro/Pfam (EMBL-EBI) server, we identified conserved PDZ binding motifs in some candidates and not in others (Table 4.1). Carboxyl terminal PBMs can be efficiently investigated for LNX2 interaction by yeast-two hybrid screens.

From enriched complexes we brought forward the PBM containing proteins, ABRAXAS2, CNP, EXOC4 and DCUN1D1 forward for analysis. These 10-mer carboxyl terminal PBMs successfully interacted with LNX2 (Figure 4.5). PAFAH1B3 from the PAF-AH complex did not have a conventional PBM and failed to activate *His3* and *LacZ* reporter genes in the yeast-two hybrid assay (Figure 4.5). In Lenihan *et al.*, 2017, LNX1 pulled down members of the multi amino acetyl tRNA synthase complex like MARS, IARS and LARS (Lenihan, Saha and Young, 2017) and here we identify KARS (lysine-tRNA ligase) as a LNX2 interactor. In this research, the group also concluded the scaffold protein, Flotillin-2 (FLOT2) as a LNX1 interactor and here we see the same for LNX2.

Other interesting putative LNX2 substrates identified from proteomics included Ras related GTPase protein RAB18, membrane trafficking protein TOM1, transcription factor GLN3, HEATR1 a HEAT repeat containing protein and nuclear import protein KPNA1. These proteins encoded PBMs and all showed interaction with full length LNX2

protein. We cannot discount protein kinase N (PKN2, 112 kDa) as a LNX2 interactor due to autoactivation of the assay, because ProfAff determines the binding affinity as $pK_a=4.59$ to PDZ2 of LNX2 (Gogl *et al.*, 2022). Weak activation of *His3* and *LacZ* by USO1 carboxyl terminus and LNX2 was observed. USO1 has a type III PBM (D-H-I) which is not usually favoured by PDZ1/2 of LNX2.

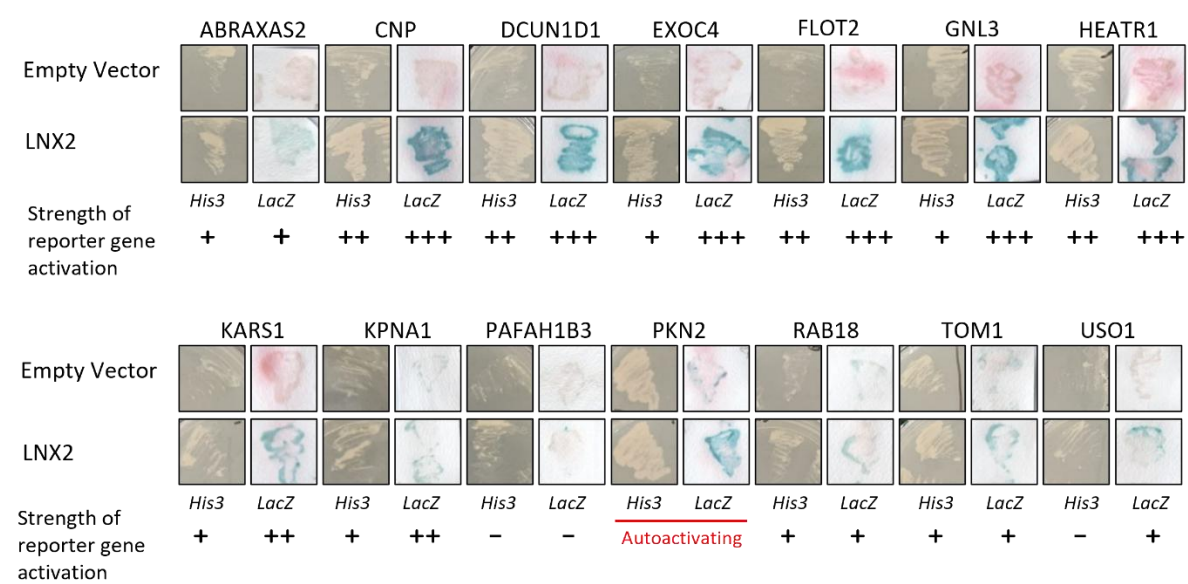


Fig 4.5: Evaluation of LNX2 substrates for the functionality of their PDZ binding motifs by yeast-2-hybrid assay.

Co-transformed yeast colonies were streaked on media lacking leucine, tryptophan and histidine, growth of reddish/pink cells indicates the bait (10-mer PBM of indicated protein) and prey (LNX2) proteins are interacting resulting in activation of the *His3* reporter gene. A positive β -galactosidase lift assay showing blue colonies when subjected to X-gal substrate indicates activation of *LacZ*. Co-transformation of an empty prey vector with each bait serves as a negative control which should show no growth or colour change. When this is not the case the bait is causing autoactivation of the reporter gene(s) and the presence or absence of an interaction cannot be determined, like in the case of PKN2. +, ++, +++ or – indicate strength of reporter gene activation and are merely observative, not quantitative.

4.2.5 Validation of LNX2 substrates by direct in vitro ubiquitination assays

LNX2 was previously experimentally determined to be an interactor of a few candidates from the list of identified protein substrates in table 4.1, like KLHL11 (Lenihan *et al.*, 2017) and ZNF24 (Chapter 3, Figure 3.3). In Figure 4.5, we see LNX2 interactions with PDZ binding motifs of candidate substrates *via* yeast-two hybrid (Y2H) screens. Strong activation of reporter genes by LNX2 and CNP (47 kDa), DCUN1D1 (30 kDa), FLOT2 (47 kDa), GNL3 (60 kDa) and ABRAXAS2 (47 kDa) was observed so these genes were amplified from HEK 293T cDNA and cloned into pCMV-N-Flag as per methods section 2.1. Although the 10 amino acid carboxyl terminal of PKN2 displayed autoactivation of *His3* and *LacZ* in the Y2H, we opted to analyse it further due to the confidence from the WT: Mut ratio from the proteomics screen and the ProfAff affinity score to LNX2 PDZ2 ($pK_d=4.59$) (Gogl *et al.*, 2022).

RHOA (21 kDa), is a small GTPase associated with cytoskeletal organisation. RHOA binds to a variety of effector proteins and its ubiquitination by SCF can influence cell migration in lung epithelial cells (Wei *et al.*, 2013) Another promising LNX2 substrate identified by mass spectrometry is the serine/threonine kinase MAP3K7 (67 kDa) also known by its alias, TAK1. MAP3K7 is a major mediator of the MAP kinase signal transduction pathway and can form as part of a ubiquitination complex with TAB1 and TAB2 (TGF β -activated kinase 1/2 or TAK1 binding proteins 1 and 2), which were also identified from our list (Thiefes *et al.*, 2006).

The first steps towards our validation of these proteins as LNX2 ubiquitin substrates is complete. Successful cloning and expression of these amino-terminally Flag tagged proteins is seen in Figure 4.6.

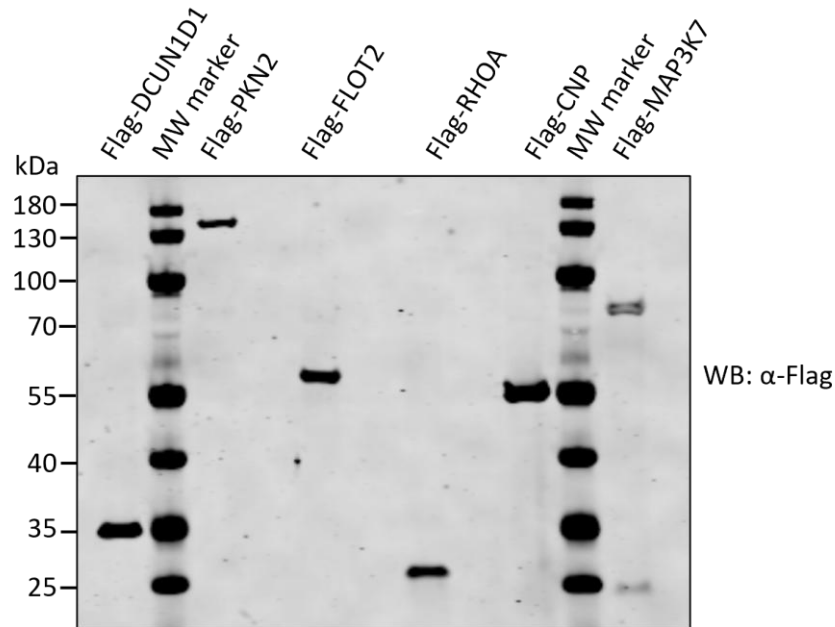


Figure 4.6: Successful cloning and expression of candidate LNX2 substrates. N-terminally Flag tagged proteins were transfected into HEK 293T cells and examined for expression by Western Blot. Indicated proteins correspond to their corresponding molecular weight and are all easily detected by anti-Flag antibody.

As mentioned before, for an E3 ligase to modify its substrate with ubiquitin(s) a direct association would usually be required (Buetow and Huang, 2016). The first step in validating LNX2 substrates is to confirm their direct, full-length interaction in mammalian systems. From the yeast-two hybrid screen in Figure 4.5 and other methods of selecting the most probable LNX2 substrates we opted to use the anti-GFP nanobody expressing yeast cells to perform GFP-LNX2 pull downs and analyse using the BD Accuri C6 Flow

Cytometer (BD Biosciences) (O’Riordan *et al.*, 2024). After full length candidate proteins were cloned fused to a Flag tag and expression was verified (Figure 4.6), we co-expressed them in HEK 293T cells with GFP-LNX2. Yeast cells were employed as a method of purifying and analysing the interactions.

The NUMB-LNX2 interaction is a strong positive control where we can see a big shift in the anti-Flag signal (x-axis) for Flag-NUMB binding to GFP-LNX2 (orange) compared to GFP-only (green) (Figure 4.7A). In comparison, Flag-tagged ABRAXAS2, CNP, DCUN1D1 and PKN2 have a very slight shift away from the GFP-only signal when GFP-LNX2 is co-expressed suggesting a very weak or no interaction. In particular, CNP seems to exhibit substantial non-specific binding in this assay. The level of FL3 signal is almost comparable to the negative control in Figure 4.7I. Flag-tagged GNL3 and MAP3K7 have slightly more of a change in FL3 signal in the presence of GFP-LNX2 versus GFP only. mediated purification but confidence of interaction is not definite. The most promising interaction of LNX2 is with Flotillin-2 (FLOT2) in Figure 4.7E. However, these experiments cannot be accepted with certainty as the difference in signal for GFP-LNX2 versus the GFP only control is very slight, especially in comparison to well characterised interactor NUMB. These results are thus merely an initial investigation of LNX2 interaction of these candidates. Repeat experiments must be performed and it could be possible that the anti-GFP nanobody displaying yeast is not the optimal method for confirming subtle interactions such as these.

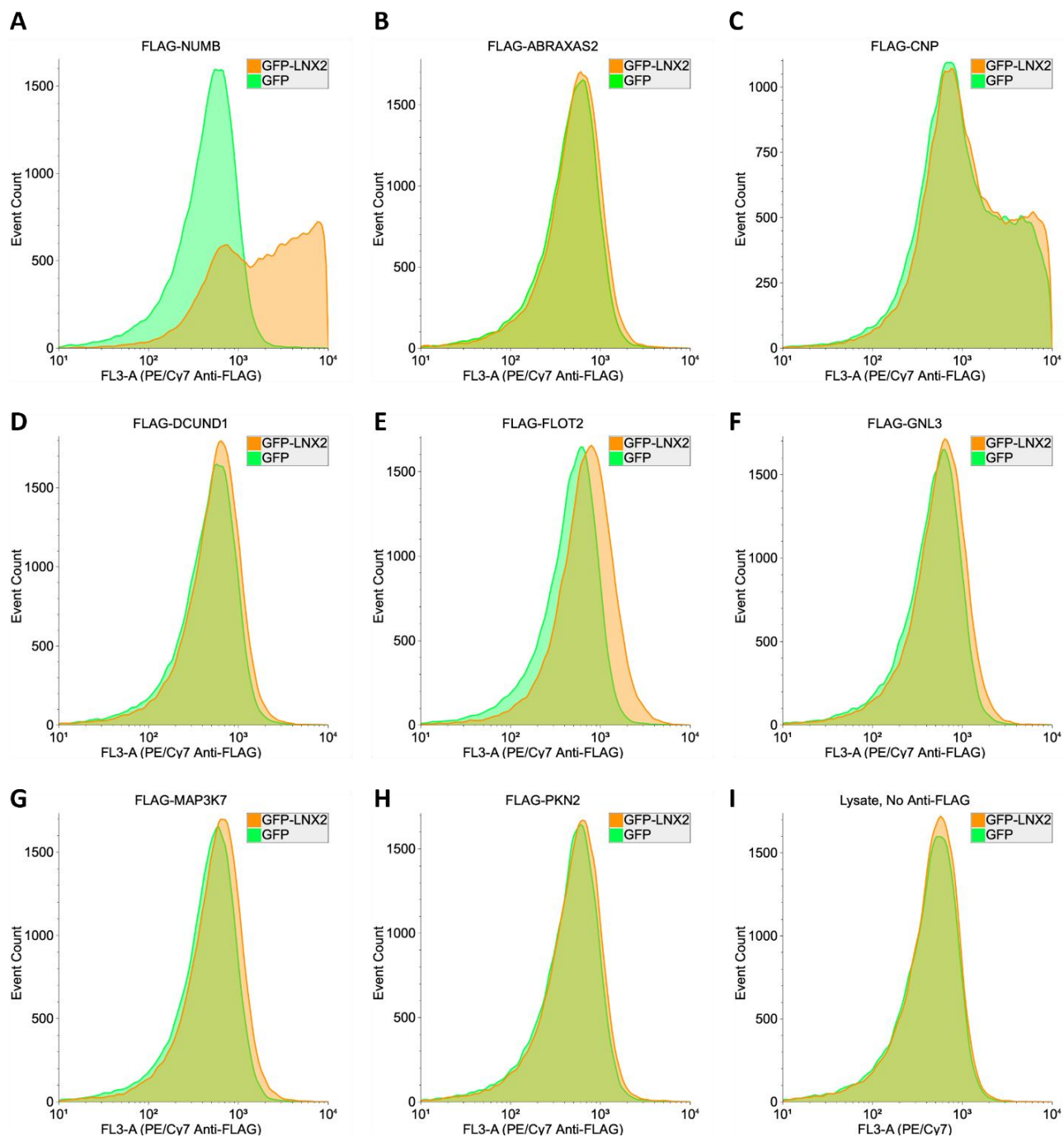


Figure 4.7: Flow cytometry analyses of GFP-LNX2 interaction with Flag tagged candidates identified from proteomic analyses. Anti-GFP nanobody displaying yeast cells were incubated with lysates from HEK 293T cells co-expressing the proteins indicated and analysed by flow cytometry. Bound Flag-tagged putative interacting proteins are detected using a PE-Cy7-anti-Flag epitope tag (FL3 , x-axis). Histograms for lysates co-expressing GFP and GFP-LNX2 are presented in green and orange respectively. **(A)** As a positive control for the assay, Flag tagged NUMB is detected bound to anti-GFP nanobody yeast specifically in the presence of GFP-LNX2. **(B-H)** Candidate proteins were examined in the same way and interactions, if any, with GFP-LNX2 are represented by a shift in the orange peaks away from the green ones **(I)** GFP only lysate with no antibody is used as a negative control to correct for signal spill over from FL1 (GFP signal) channel into the FL3 (anti-Flag signal) channel. Data was corrected by compensation for the spectral spill over on the Accuri Software by subtracting 0.8% of the FL1 signal from FL3. Data shown is representative of one experiment.

”

4.3 Discussion

4.3.1. The importance of properly validating novel methods

In this chapter, we further modified and simplified the BioE3/Ub-POD (Barroso-Gomila *et al.*, 2023) (Mukhopadhyay *et al.*, 2024) methodology by removing the doxycycline-inducible expression system used in the original protocol. Instead, we utilised transient transfection of HEK293T cells, which significantly reduced both the experimental time and overall cost, while maintaining the specificity and sensitivity of the system. While optimising the proximity dependent labelling of LNX substrates, it was clear that the extremely harsh washing conditions (e.g. 8M urea or 6M guanidine-HCl) were eliminating nearly all the ubiquitinated material from the beads (Figure 4.2 and S1). While these studies developed their protocols based on well-expressed and characterised E3 ubiquitin ligases, LNX2 is presumed to have a more subtle effect *in vivo*. In that case we opted to use less stringent washes that allowed efficient binding of material to NeutrAvidin beads and enrichment in the pull down fraction (Tan *et al.*, 2020) (Dumrongprechachan *et al.*, 2021).

We also noticed from the BioE3 paper there was no BirA unfused to an E3 ligase control. When this control was transfected with artificial ubiquitin (BioGEF-Ub) we noticed a strong background biotin signal (Figure 4.2). Because ubiquitin is incorporated into substrates by numerous endogenous E3 ligases, expression of Avi-tagged ubiquitin together with free BirA is expected to generate widespread biotinylation of ubiquitinated proteins. This creates a substantial background signal, highlighting the importance of appropriate controls when interpreting proximity-dependent ubiquitination labelling

approaches. This could be considered a limitation of the assay and a reason for the highly stringent washes employed (Barroso-Gomila *et al.*, 2023).

4.3.2 Interpreting proteomic results

One of the first things we noticed from the raw data list of peptide spectra was the considerable number of proteins ubiquitinated by the LNX2 C53A mutant. A cysteine-to-alanine mutation within the RING domain disrupts Zn^{2+} coordination and supposedly abolishes E3 ligase activity, generating a catalytically inactive mutant used as a control for ubiquitination-dependent effects (Raymond J Deshaies and Joazeiro, 2009) (Garcia-Barcena *et al.*, 2020). A well-characterised ubiquitin substrate of LNX2 is NUMB protein (Nie, Li and McGlade, 2004)(Rice, Northcutt and Kurschner, 2001). As mentioned, NUMB peptide fragments were present in both WT and C53A mutant samples, which could suggest the single amino acid substitution is not severe enough to abolish the complete catalytic ability of LNX2. The biotin smearing signals from mutant and WT LNX protein activity was suggestive of this, due to convincing smear intensities present for the mutant (Figure 4.2 & 4.4). Lenihan *et al.*, also observed this when demonstrating that Liprin- α 1 is a LNX1 substrate (Lenihan, Saha and Young, 2017) and also LNX1 itself (Lenihan *et al.*, 2014). One concern when evaluating these findings is whether the threshold used to define LNX2 substrates was overly stringent, which may have resulted in the exclusion of valid candidate targets.

On the other hand, could the presence of known ubiquitin targets in the catalytic mutant of LNX2 suggest an ability of LNX2 to sequester other ubiquitin ligases and mediate their

regulation, without binding to them directly? We see from the GO analysis of our proteomic hits for cellular compartment enrichment that many ubiquitin complexes are significantly represented. BRISC and BRCA1-A complex members do not all have PDZ binding motifs at their C-terminus for example BABAM1 and BRCC3, even though they were pulled down in the WT LNX2 biotinylation assay. Similarly, in the PAF-AH complex PAFAH1B2 has a PBM (T-T-I-A) but the other catalytic subunit PAFAH1B3 does not. We can either assume LNX2 is acting as a scaffold for the mediation of external ubiquitinating factors, or the nature of the proximity assay is flawed and non-specific ubiquitinated substrates are pulled down. Only further examination of these interactions *in vitro* would clarify.

Furthermore, even the proteins containing PBMs that we observed having a direct interaction with LNX2 in the Y2H screens (GNL3, ABRAXAS2, DCUN1D1 FLOT2 and GNL3), are not confidently confirmed when subjected to full-length interaction analyses. The sensitivity or efficiency of such a method may not be suitable for determining weak interactions. Nevertheless, the preliminary steps have been taken with the ambition to authenticate true substrates and supplement the LNX2 ubiquitome.

Chapter 5

Molecular Evolution of LNX Proteins: Structural Conservation and Interaction Networks

5.1 Introduction

The LNX/PDZRN family contains five members and for clarity we refer to them as LNX1-LNX5. (Dho *et al.*, 1998)(Wang *et al.*, 1999) (Rice, Northcutt and Kurschner, 2001)(Katoh and Katoh, 2004). LNX1 and LNX2 have the same domain architecture consisting of a RING domain, an NPAY/F motif and four PDZ domains (Figure 1.7, Chapter 1) (Rice, Northcutt and Kurschner, 2001). Across chordate evolution, LNX 1/2 proteins have retained this highly conserved domain architecture. These PDZ domains are closely related to the four carboxy-terminal domains from multiple PDZ domain containing protein-1 (MUPP1). LNX and MUPP1 are proposed to have diverged from a common ancestral gene that contained PDZ domains (Flynn, Saha and Young, 2011). Many studies report the impact of each of these domains in cellular processes and organism development in vertebrates (Young, 2018) (Hong, Won and Ro, 2020)The LNX interaction with NUMB has been well documented, and as far as I am aware it is the only interactor mapped to the NPAY/F motif, suggesting the importance of these four amino acids. In vertebrates, LNX1/2 play critical roles in neurogenesis and dorso-ventral patterning, largely through regulation of Notch signalling *via* NUMB (Dho *et al.*, 1998) (Rice, Northcutt and Kurschner, 2001) (Nie *et al.*, 2002b) (Nie, Li and McGlade, 2004). Flynn *et al.* (2011) reported that many invertebrate LNX orthologs lack this NPAY/F motif, suggesting the LNX-NUMB interaction was not present in such species. However, this characteristic was only studied in the invertebrate genomes available at the time.

LNX proteins carry out their roles largely through the ubiquitination of substrate proteins (Guo *et al.*, 2012). The presence of a catalytic RING domain renders LNX as an E3

ubiquitin ligase, and is a conserved domain across all vertebrate and many invertebrate species granting its importance throughout evolution (Flynn, Saha and Young, 2011). Substrates of LNX1/2 proteins first interact with one or many of the PDZ domains or in the case of NUMB, the NPAY/F motif and then can be ligated to a ubiquitin moiety. These PDZ domains have also been conserved in the metazoan lineage but some LNX orthologs are more similar to LNX3-5 than to LNX1/2 therein having less than four PDZ domains (Flynn, Saha and Young, 2011).

Conservation of domain architecture is important to imply protein function, but the quaternary folding and oligomerisation is necessary to understand how this occurs in cells. Rice *et al* revealed that PDZ1 of both LNX1 and LNX2 proteins can intramolecularly bind their own carboxyl terminal PDZ binding motif (PBM). The propensity to self-binding could essentially “hide” or cover certain domains within the protein so they are inaccessible in certain dynamic states. LNX2 was also found to dimerise between the RING domains of two separate LNX2 moieties. This work carried out by yeast-two hybrid analysis suggests that LNX proteins may exist as dimers or higher order multimeric proteins (Rice, Northcutt and Kurschner, 2001). There are only partial LNX1 and LNX2 crystal structures solved to date, these include RING domains of LNX1 (PDB ID: 5H7R) and LNX2 (PDB ID: 5DIN) (Nayak and Sivaraman, 2018) (Nayak *et al.*, 2015). Individual LNX PDZ domains have mostly been solved in complex with an interacting protein - LNX1 PDZ2 (PDB ID: 3VQG), LNX1 PDZ3 (PDB ID: 3B76), LNX2 PDZ2 (PDB ID: 2VWR). The lack of available full-length crystal structures provides challenges for researchers to infer the 3D conformation of LNX proteins, however, protein modelling predictions have become

advanced recently and AlphaFold3 is an excellent tool for *in silico* conformational analysis (Abramson *et al.*, 2024).

In 2011, Flynn *et al* compared the sequences of several metazoan species containing LNX1/2 like protein homologs - *Schistosoma mansoni*, *Branchiostoma floridae*, *Saccoglossus kowalevskii*, *Strongylocentrotus pupuratus*, *Amphimedon queenslandica* as well as human, mouse, chicken, frog and zebrafish. The group used bioinformatical methods like BLASTp (NCBI Protein Database) and PubMed to identify sequences (Flynn, Saha and Young, 2011). At the time of publishing there were a limited number of annotated genomes especially those of primitive species, so obtaining LNX sequences was challenging. From the completion of the Human Genome Project in 2004 to the newest Ensembl release 116, an astounding 37,546 genomes have been annotated (Abdellah *et al.*, 2004)(Yates *et al.*, 2026). In this chapter, we aim to identify and characterise more LNX orthologs to further the understanding of LNX evolution.

Aims of Chapter 5

- Identify metazoan LNX orthologs to enhance the list of what is known to date by using increased availability of annotated genomes.
- Determine the level of domain and structural conservation in these orthologs to possibly elucidate function.
- Predict inter- and intramolecular interactions using AlphFold3 modelling.
- Evaluate the evolution of the NUMB and LNX protein interaction.

5.2 Results

5.2.1 Many metazoan species contain a well conserved LNX ortholog

From initial research on LNX1/2 protein evolution in 2011, only a few genomes of members of the lesser metazoans were available to be investigated (Flynn, Saha and Young, 2011). The authors determined that an early *LNX1/2-like* ancestral gene underwent duplication that gave rise to *LNX1*, *LNX2* and *LNX2b*. *LNX2b* is not found in placental mammals but is found in all other vertebrates and deemed interesting in the evolution of *Xist* RNA in X-chromosome inactivation (Duret *et al.*, 2006). *Branchiostoma floridae* (Florida lancelet) LNX shares higher sequence homology with vertebrate LNX2 ~41%, rather than LNX1 (Flynn, Saha and Young, 2011). The lancelet has been deemed as a proxy for the last invertebrate ancestor of chordates (Garcia-Fernàndez and Benito-Gutiérrez, 2009). We exploited this knowledge and searched for LNX2-like proteins among currently annotated metazoan species genomes using BLASTp (NCBI Protein Database).

Clear LNX2 orthologs from 21 representative species from major phyla were selected for further investigation (Table 5.1). From the Chordata clade, species from mammals (human and mouse), amphibians (frog), Actinopterygii (eel and zebrafish), Cephalochordates (lancelet), tunicates (sea squirt) and hemichordates (acorn worm) were selected. Three Echinodermata species were chosen – common star fish, feather star and purple sea urchin. From the Arthropoda phylum, six species with LNX2 orthologs were found - American lobster, barnacle (Thecostraca), woodlouse (Malacostraca), calanoid copepod (Copepod), termite and leafhopper (insect). Lastly, three

representative Platyhelminthes – flatworm, blood fluke and lung fluke and a single Poriferae (sea sponge) LNX2 orthologous sequences were gathered. All identified orthologs were investigated individually and the corresponding RefSeq (NCBI-NIH) accession code was determined.

Table 5.1: Scientific and common names of selected metazoan species and their orthologous LNX protein RefSeq accession codes.

Scientific name	Common name	RefSeq accession
<i>Homo sapiens</i>	Human	NP_699202.1
<i>Mus musculus</i>	Mouse	AAK94476.1
<i>Danio rerio</i>	Zebrafish	NP_001106696.2
<i>Anguilla anguilla</i>	European eel	XP_035266888.1
<i>Xenopus tropicalis</i>	Frog	NP_001072606.1
<i>Branchiostoma floridae</i>	Amphioxus/Florida lancelet	XP_035682952.1
<i>Styela clava</i>	Sea squirt	XP_039264061.1
<i>Ptychodera flava</i>	Acorn worm	XP_070566843.1
<i>Asterias rubens</i>	Common starfish	XP_033639041.1
<i>Anneissia japonica</i>	Japanese feather star	XP_033121042.1
<i>Strongylocentrotus purpuratus</i>	Purple sea urchin	XP_030848543.1
<i>Homarus americanus</i>	American lobster	XP_042221237.1
<i>Amphibalanus amphitrite</i>	Striped barnacle	XP_043218263.1
<i>Armadillidium nasatum</i>	Pill woodlouse	KAB7496441.1
<i>Eurytemora carolleeae</i>	Calanoid copepod	XP_023321993.1
<i>Cryptotermes secundus</i>	Termite	XP_023709112.1
<i>Macrosteles quadrilineatus</i>	Aster leafhopper	XP_054280717.1
<i>Macrostomum lignano</i>	Flatworm	PAA46561.1
<i>Schistosoma japonicum</i>	Japanese blood fluke	TNN17880.1
<i>Paragonimus heterotremus</i>	Lung fluke	KAF5399768.1
<i>Amphimedon queenslandica</i>	Sea sponge	XP_019848881.1

A multiple sequence alignment (MSA) was constructed using Clustal Ω (EMBL-EBI) (Madeira *et al.*, 2024) in which the conserved domains and motifs sequences (RING, NPAY/F, PDZ and PBM) of LNX could be easily visualised and assessed (Figure 5.1). Certain LNX orthologs have variations in spacing between the domains but Clustal Ω employs UPGMA/neighbour joining method to generate a distance matrix and produces the best fit MSA for the input sequences (Madeira *et al.*, 2024). A high level of conservation can be seen for the zinc finger RING domain, all four PDZ domains and the carboxyl terminal PBMs. The NPAY/F and surrounding sequence (E-P-G-L-V/D-NPAF-E-E-S) is well conserved in the vertebrate LNX2 orthologs - *Homo sapiens*, *Mus musculus*, *Danio rerio*, *Anguilla Anguilla* and *Xenopus tropicalis*. This region does not align with the other species as depicted in figure 5.1. A search of every LNX sequence encoded by the selected metazoan species' genomes was carried out by Wregex v3.0, to which no similar or conserved NPAY/F motif was discovered across the entire protein sequence.

		10	20	30	40	50	60	70																					
1	<i>A.queenslandica</i> /1-700																												
2	<i>S.clava</i> /1-898	MEYSYQA	P	DCL	PGQNDL	IDFSSDSSV	G	VHNQEE	G	E	ESSSTLT	TKNRNQ	INRDVVAVKS	G	T	G	D	I	G	FVAV	G	R	P	T	S				
3	<i>S.japonicum</i> /1-1060	---	MQT	G	S	C	---	---	DDST	G	SVGG	D	---	---	SETNTTV	---	---	ASS	P	G	F	D	P	P	V	T	K	S	S
4	<i>A.anguilla</i> /1-698	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
5	<i>D.reio</i> /1-737	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
6	<i>X.tropicalis</i> /1-685	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
7	<i>H.sapiens</i> /1-690	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
8	<i>M.musculus</i> /1-687	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
9	<i>A.amphitrite</i> /1-680	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
10	<i>E.carolleeae</i> /1-860	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
11	<i>C.secundus</i> /1-744	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
12	<i>M.quadrilineatus</i> /1-616	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
13	<i>H.americanus</i> /1-846	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
14	<i>A.nasatum</i> /1-726	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
15	<i>P.heterotremus</i> /1-876	---	---	---	---	---	---	---	---	---	---	---	---	---	---	MN	S	S	---	---	---	---	---	---	---	---	---	---	
16	<i>M.lignano</i> /1-799	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
17	<i>B.floridae</i> /1-652	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
18	<i>P.flava</i> /1-601	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
19	<i>A.japonica</i> /1-596	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
20	<i>A.rubens</i> /1-598	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
21	<i>S.purpuratus</i> /1-573	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	

RING domain

RING domain

		320		330		340		350		360		370		380		390																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
1	E	V	V	D	E	D	L	C	A	M	C	H	K	P	L	I	E	P	V	E	T	V	C	G	H	T	F	C	Y	K	C	L	F	G	R	I	R	T	S	P	Y	C	P	V	E	N	H	P	I	S	R	I	D	G	D	I	Q	Q	T	S	F	L	V	R	K	L	L	D	K	M	K	V	I	C	P	N	T	A	Y	C	D	L	V																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
2	D	E	V	D	E	D	L	M	C	I	D	I	C	L	Q	P	L	V	D	P	D	T	K	C	G	H	T	F	C	S	I	C	K	N	Y	L	R	V	K	H	I	C	P	V	D	R	E	E	L	T	N	T	A	T	D	C	W	Q	S	S	L	I	R	K	L	V	N	K	L	T	V	C	P	H	E	E	Y	Q	D	K	K																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
3	Q	T	V	D	A	L	C	T	L	C	R	Q	P	L	V	D	P	L	D	T	K	C	G	H	T	F	C	S	P	C	L	N	H	L	A	V	Q	A	L	C	P	V	D	R	Q	I	N	-	Y	L	E	C	Q	Q	A	S	N	I	V	R	K	L	L	D	K	L	L	V	C	P	N	S	A	C	D	S	I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
4	D	E	V	D	D	L	V	C	H	I	C	L	Q	P	L	L	R	P	M	D	T	P	C	G	H	T	F	C	F	Q	L	S	S	F	L	R	E	Q	D	F	C	P	V	D	R	Q	R	L	Q	-	-	L	H	Q	C	R	S	S	S	L	V	R	N	L	L	D	K	L	T	V	L	C	P	F	H	A	C	Q	Q	Q																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
5	L	E	V	D	D	L	V	C	H	I	C	L	Q	P	L	V	Q	P	L	D	T	P	C	G	H	T	F	C	A	R	C	L	R	S	F	L	Q	E	R	D	F	C	P	L	D	R	A	H	L	Q	-	-	L	Q	V	C	R	S	S	S	L	V	H	K	L	L	D	K	L	S	V	T	C	P	L	T	P	S	C	S	L	S																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
6	N	E	V	D	D	L	V	C	H	I	C	L	Q	P	L	L	Q	P	L	D	T	P	C	G	H	T	F	C	F	K	L	R	N	F	L	Q	E	K	D	F	C	P	L	D	R	K	R	L	H	-	-	F	K	L	C	K	K	S	S	L	V	H	K	L	L	D	K	L	V	C	P	N	S	A	C	K	E	T																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
7	N	E	V	D	D	L	V	C	H	I	C	L	Q	P	L	L	Q	P	L	D	T	P	C	G	H	T	F	C	Y	K	L	R	N	F	L	Q	E	K	D	F	C	P	L	D	R	K	R	L	H	-	-	F	K	L	C	K	K	S	S	L	V	H	K	L	L	D	K	L	V	L	C	P	F	S	S	V	C	K	D	V																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
8	G	E	V	D	D	L	V	C	H	I	C	L	Q	P	L	L	Q	P	L	D	T	P	C	G	H	T	F	C	H	K	C	L	R	N	F	L	Q	E	K	D	F	C	P	L	D	R	K	R	L	H	-	-	F	K	L	C	K	K	S	S	L	V	H	K	L	L	D	K	L	V	L	C	P	F	S	P	V	Q	D	V																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
9	D	E	V	D	E	D	L	T	C	H	I	C	L	Q	P	L	V	G	S	D	T	P	C	G	H	T	F	C	Q	A	C	L	V	P	L	N	S	H	Q	N	C	P	I	C	R	A	A	L	-	-	T	Q	A	C	R	N	S	P	L	V	L	Q	I	E	K	L	L	I	K	C	P	N	A	D	V	C	K	E	T																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
10	E	D	V	D	E	D	L	M	C	Q	I	C	L	Q	P	F	V	N	P	T	D	T	P	C	S	H	T	F	C	Q	A	C	M	T	A	V	L	R	V	N	C	T	P	I	C	R	K	P	L	N	-	-	Q	L	T	C	S	P	S	N	L	V	R	K	L	L	D	K	L	P	V	V	C	P	N	S	D	C	S	E	T																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
11	E	E	I	D	E	D	L	I	C	Y	I	C	L	Q	P	F	V	T	P	L	D	I	P	C	G	H	T	F	C	Q	A	C	L	S	T	F	L	R	V	Q	Q	L	C	P	I	D	R	K	P	L	S	-	-	A	N	L	C	F	P	S	S	I	V	L	R	K	L	L	E	K	L	P	V	Q	C	P	N	T	T	I	C	D	K	I																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
12	Q	D	V	D	E	D	L	M	C	Q	I	C	L	Q	P	F	V	C	P	D	L	L	T	P	C	G	H	T	F	C	L	P	C	L	H	K	L	Q	V	Q	K	L	C	P	I	D	R	K	P	L	T	-	-	L	K	E	S	L	P	S	S	L	V	L	K	L	I	D	K	L	L	V	C	P	N	S	A	D	C	Q	T																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
13	D	D	V	D	E	D	V	T	C	H	I	C	L	Q	P	L	V	T	P	L	D	T	P	C	G	H	T	F	C	R	P	C	L	L	S	F	T	V	Q	Q	N	C	P	V	D	R	K	P	L	S	-	-	Q	Q	L	C	S	P	S	S	L	V	L	K	L	L	E	K	L	K	V	R	C	P	N	V	D	C	H	T	M																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

NPAY domain

		480	490	500	510	520	530	540	550	
1	---	---	---	---	---	---	---	---	---	---
2	---	---	---	---	---	---	---	---	---	---
3	---	---	---	---	---	---	---	---	---	---
4	---	---	---	---	---	---	---	---	---	---
5	---	---	---	---	---	---	---	---	---	---
6	---	---	---	---	---	---	---	---	---	---
7	---	---	---	---	---	---	---	---	---	---
8	---	---	---	---	---	---	---	---	---	---
9	---	---	---	---	---	---	---	---	---	---
10	---	---	---	---	---	---	---	---	---	---
11	---	---	---	---	---	---	---	---	---	---
12	---	---	---	---	---	---	---	---	---	---
13	---	---	---	---	---	---	---	---	---	---
14	---	---	---	---	---	---	---	---	---	---
15	---	---	---	---	---	---	---	---	---	---
16	---	---	---	---	---	---	---	---	---	---
17	---	---	---	---	---	---	---	---	---	---
18	---	---	---	---	---	---	---	---	---	---
19	---	---	---	---	---	---	---	---	---	---
20	---	---	---	---	---	---	---	---	---	---
21	---	---	---	---	---	---	---	---	---	---

PDZ1

560 570 580 590 600 610 620 630

1 K R Q I K E G V P T V V E I P K G T D Q L G I G I K K C G R S . . . G S I V I D D I Y V T G N V Y K D G R I L P G D E I I E I D D V L F R H K L L D Q V R I A

2 . . L H P F S E E I V V L F R P E G T A L G L L I V G G S D T P L H Y I I V D D I L P Q S L A F Q D G R I H V G D I I M E I N G E P M F S V M H A E A R A A

3 S P A I Y E G V P V S I L I Y R P A Y C M D L G F T V G G I D T P L T C I L I Q E I Y L D G L V A T D G R L R P G D Q L L E V N G N L L T H V H S E A R S L

4 . . S S L V A E A S V V E I Y R D E P E E L G V R I V G G K D T P L G N I V I Q E V L R D S L A A R D G K L A P G D Q I L E V N D V N V A C V P H R R A I A V

5 L P S L P E G E I T T I E V H R T N P Y S E M G I S I V G G N E T P L I N V V I Q E V Y R D G V I A R D G R L L A G D Q I L Q V N N V D I S N V P H N F A R S T

6 P P S L P E G E I T T I E I H R S N P Y I E L G I S I V G G N E T P L I N V V I Q E I Y R D G V I A R D G R L L A G D Q I L Q V N N F D I S N V S H N Y A R W V

7 P L S L P E G E I T T I E I H R S N P Y I Q L G I S I V G G N E T P L I N I V I Q E V Y R D G V I A R D G R L L A G D Q I L Q V N N Y I S N V S H N Y A R A V

8 L P S L P E G E I T T I E I H R S N P Y I Q L G I S I V G G N E T P L I N I V I Q E V Y R D G V I A R D G R L L A G D Q I L Q V N N Y D I S N V S H N H A R A V

9 R L P I I E G E I S H I E I P R . . T A T N L G M T I V G G A D T P L R C V V I Q E V F N D S L V A Q D G R L Q A G D Q I V E V N G V D M T N A S H K M I C Q A

10 R L P I L E G E V S H V E I P R . . T S S C L G I T I V G G S D T P L R C V V V Q E V F P D G L I S G D G R L Q P G D Q L I E I N G V D M T C A S H A Q V C A A

11 R L P V L E G E I S H I E I P R . . V A T N L G I T I V G G A D T A L R C V V V Q E V F P E G L V S Q D G R L R P G D Q L I E I N G I D M T T A S H Q T C Q A

12 R L P V L E G E V S H V E I P R . . I S P S L G I T V G G S D T A L R C I V V Q E V F P D G L V A Q D G R L K P G D Q L V E I N G V D M S T A S H H Q A C Q A

13 R P P I V E G E I T H L E I P R . . T H T A L G I T I V G G S D T P L R C V V V Q E V F P D G L V A Q D G R L Q P G D Q I I E V D G V D M T S A T H Q V V C Q A

14 R N P V L E G E I T F I E I P R . . T Q T A L G I A I V G G A D T P L R C V V I Q E V F P D G L I A H D G R L Q A G D Q I I E I E G S D M T N A S H H V C T A

15 P . V V A P S N I A S I D L K L D . G R T E L G I S L V G G W D T P L L C I I Q E I F L D G L A A K D G R L R P G D Q I L E V N G E D L S Q A T H L Q A C H I

16 Q Q V V S G D V S T I E F G R . . G R H D L G V S F V G G S D T P L S C I V I Q E I Y M D G L A I D G R L R P G D Q I L E V N G E D V T R A S H S E A R R L

17 R L P V L E G E I T T I E V R K S S Q S E H L G I S V V G G C E T P L V S M I V Q E V F S S G A I A T D G R I G P G D Q I L E V N G E D L R E L T H H Q A C R V

18 P T Q I L E G V V T T I E I N R . . T E E L G I S L V G G S D T P L V A I V V Q E I F P E G L I A K D G R I N P G D Q I L E I N G E D L R D V S H A Q A R A A

19 V V A I K T G G L T T I E I R R N . T D E F G V S L V G G N E T P L I V T I V Q E I Y P E G V I A K D G R I M P G D Q I V E V N G K L R D V S H Y T A R E T

20 G V P I K D G T L T T I E I L R . R E E L G I S I V G G N E T P L I G I V I Q E I T A G G I A R D G R I A A G D Q I V Q V N D V D L K D V S H Q V A R M A

21 G A P I R H G Q L T S I E I N R K . P N R D L G I Y F V G G N E T P L I G V V V Q E V I P G E V V A E D A R I L P G D Q I V Q V N S R D L R D I P H S I A R E A

1 LSSGKPLLIKFTIVRVHDEEYDPSSSVASLVKPKADIEEIHVLRDECRPLGIQI
 2 LTKP.SNLMRFMLLRDVONEDPSDSVASLVKPKPEEVLGIKPPLEGLGIRI
 3 LTIPLSSVQLTVFREPVSINLRHQFQPOYQQEMFYVKLFKRQGVKVLGIKL
 4 LRRP.CPLLRLTVMQEKGFSTRNRPPEHHPPAGTNATPLSLGPAPNNNNSSGPGPQPGSTVTQVMLVKRDRSEPLGIKL
 5 LARP.CATLQLTLVLRERRCSRPPAAATASPKGSPASIRITLHKRESSEGLGIKL
 6 LSQP.CTVLHLLTVLRERRFGSRAHGHGSGSQRDESFHITLNKRGSNEGLGIKL
 7 LSQPCNTLHLLTVLRERRFGNRAHHSDNSPREEIFQVALHKRDSGEGLGIKL
 8 LSQPCSTLQLTLVLRERRFGSRANSHADGSAPRDEVFQVLLHKRDSTEGLGIKL
 9 LKVQT.PVLQLGVYREKIDPPPPSDDPELHTVTLEKMPGQTLGIKL
 10 LKKMS.SVLR LGVYREIEKYGIYNSGSKSGSGSGSGTRGSGVGGGQTSEQILTVTLERQGLHLGIRI
 11 LRKLS.PVLR LGVYREIEAYSARSSAHYIEHTVTLERQTNQGLVRL
 12 LRKLKSQLR LGVYREIEYNGPNDALQVETVTVTIERESQALGIRL
 13 LRRPQ.TVLR LGVYREIEAQRTLSTTTPAPPATTHVPQPGEAMMVTLQKESSRQLGLKL
 14 LRRIS.PVLR LDELVQVTLNREAGRLGIKL
 15 LSSLVGPVCLWTIYREQPFQAGTSCLGPCLPLKFDEKCLIRIDLKSEHKRLGIKL
 16 LSQG.SQFRLTVYREACEFDAQQQQQQQFLQH.HHQQQQRRLLQDERDEVFPIILSKRSKTLGIKL
 17 LRKP.CPLLRLTILRERVAEERN.TGAEVQDTIRVELTKTSGEGLGIRI
 18 FSRQ.SSVARLTVLREKVEDRGLPDREELIKITLTVPGVQLGIKI
 19 IMRA.SSPIRMTIYREKAEERNIDLLNVTLRKVKGKQLGIKI
 20 LHRA.RSPVRLTIYREKAE.NTSKLEMFKINLTKVPGVPLGIKI
 21 LCSA.ETPLRLTIYREKQESKAPVFTTTLVQVSLTKVTGVPLGIKI

PDZ1

PDZ2

PDZ2

	1120	1130	1140	1150	1160	1170	1180	1190					
1	TND	ELLGLILGGGYH	AVNTLT	GDRIDLP	IFVGG	LNNDCA	IEIANKV	QPGDII	VSSIN	NLSLAKK	SLNEAL	FILESI	KIG
2	SPKEN	LGIIVSGG	-----	SPRG	NLP	IFVQD	VLMSV	LGGK	DKRIR	RGGDLL	SEINS	IRLTGL	SHTAM
3	SP	EELG	ISIA	GGRH	-----	SQR	GD	TPI	VYTNIS	ECV	LVGR	SKLIR	GD
4	EPRL	SLGIT	ITIA	GGRD	-----	SRS	RVP	VYIT	SVQ	PGV	CLHR	DTIK	TD
5	EPH	ESL	GMT	VAGGR	-----	SKS	GELP	IFV	TSV	QPH	GC	LSR	DG
6	EP	QESL	GMT	VAGGR	-----	SKS	GELP	IFV	TSV	QPH	GC	LSR	DG
7	EPH	ESL	GMT	VAGGR	-----	SKS	GELP	IFV	TSV	QPH	GC	LSR	DG
8	EPH	ESL	GMT	VAGGR	-----	SKS	GELP	IFV	TSV	QPH	GC	LSR	DG
9	VAK	ESL	G	LRI	GGG	I	H	-----	CIE	GDT	PI	YIAN	I
10	NNR	ESL	G	MRI	GGG	I	G	-----	SN	E	GDT	PI	YIAN
11	EP	NESL	G	MRI	GGG	L	G	-----	SN	E	GDT	PI	YIAN
12	EP	NESL	G	MRI	GGG	L	G	-----	SN	E	GDT	PI	YIAN
13	GGN	ESL	G	MRI	GGG	V	G	-----	SN	E	GDT	PI	YIAN
14	DAS	ESL	G	MRI	GGG	V	N	-----	SN	E	G	N	I
15	Q	S	E	S	L	G	M	S	V	A	G	G	V
16	S	P	G	E	S	L	G	M	S	V	A	G	G
17	NAR	ESL	G	I	S	V	A	G	G	R	-----	GTR	G
18	GS	T	E	S	L	G	M	S	V	S	G	G	L
19	PP	S	E	F	L	G	I	S	V	C	G	G	V
20	G	K	E	S	L	G	I	S	V	S	G	G	R
21	G	N	K	E	C	M	G	I	S	V	S	G	G

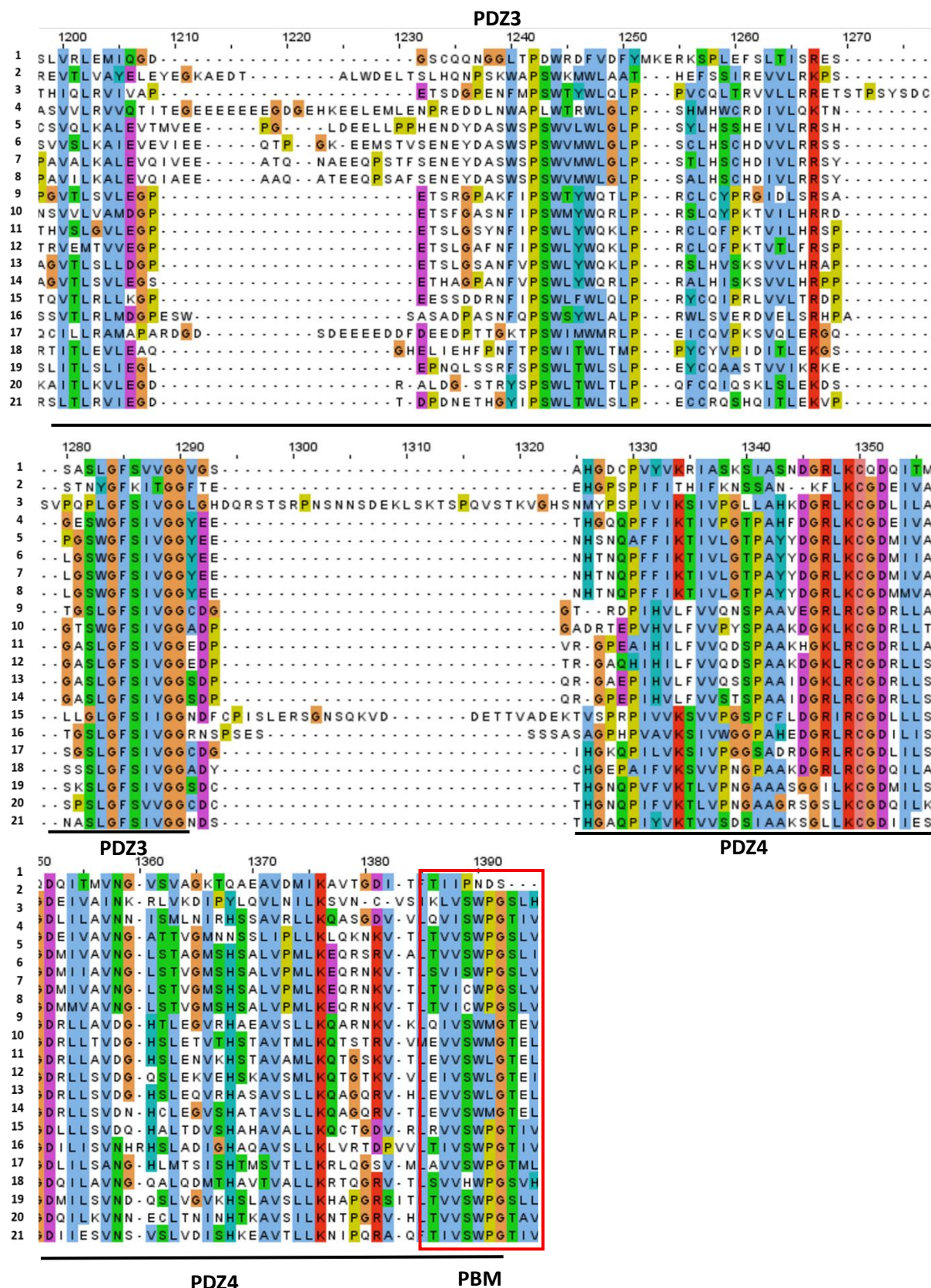


Figure 5.1: Partial LNX protein multiple sequence alignment consisting of twenty-one metazoan species. LNX orthologs from indicated species (numbered 1-21 for space) were aligned by ClustalΩ and amino acids were coloured accordingly. Conserved domains were boxed – first red box shows good conservation of the catalytic RING domain, the second box is the NPAF motif, PDZ domains 1-4 are boxed in black and carboxyl terminal ten amino acids (PBM) is also highlighted in red. For MSA of entire protein sequences see Figure S3.

5.2.2 LNX protein orthologs have a well conserved domain architecture.

The protein sequences of vertebrate LNX2 are between 680-740 amino acids in length. Some orthologous LNX2 sequences are much longer – *Styela clava*, *Homarus americanus*, *Eurytemora carolleeae* and *Paragonimus heterotremus* have proteins consisting of over 800 amino acids and *Schistosoma japonicum* over 1000. From the multiple sequence alignment in figure 5.1 we observed large gaps in the alignment that were not conserved with any other protein. In figure 5.2, each LNX2 ortholog structure was visualised using NCBI CD (conserved domain) search tool (J. Wang *et al.*, 2023).

Invertebrate LNX protein architectures were considerably different to those from vertebrate clade. In the case of the tunicate species, *Styela clava* and the platyhelminth, *Schistosoma japonicum* LNX a 250 amino acid N-terminal space occurs before the RING domain is encoded. This is unusual when comparing the rest of the metazoan LNX proteins as the RING domain typically occurs close to the amino terminal within the first 100 amino acids. Another noticeable difference between certain members of metazoan lineage is the spacing between PDZ domains. Vertebrates along with echinoderms and chordates encode LNX proteins with four evenly spaced PDZ domains, whereas arthropods, platyhelminths and poriferans have a large gap between PDZ2 and PDZ3

(Figure 5.2). The LNX protein encoded by these species consists of PDZ1 and PDZ2 situated very close to the catalytic RING domain a feature that is not conserved in vertebrate LNX.

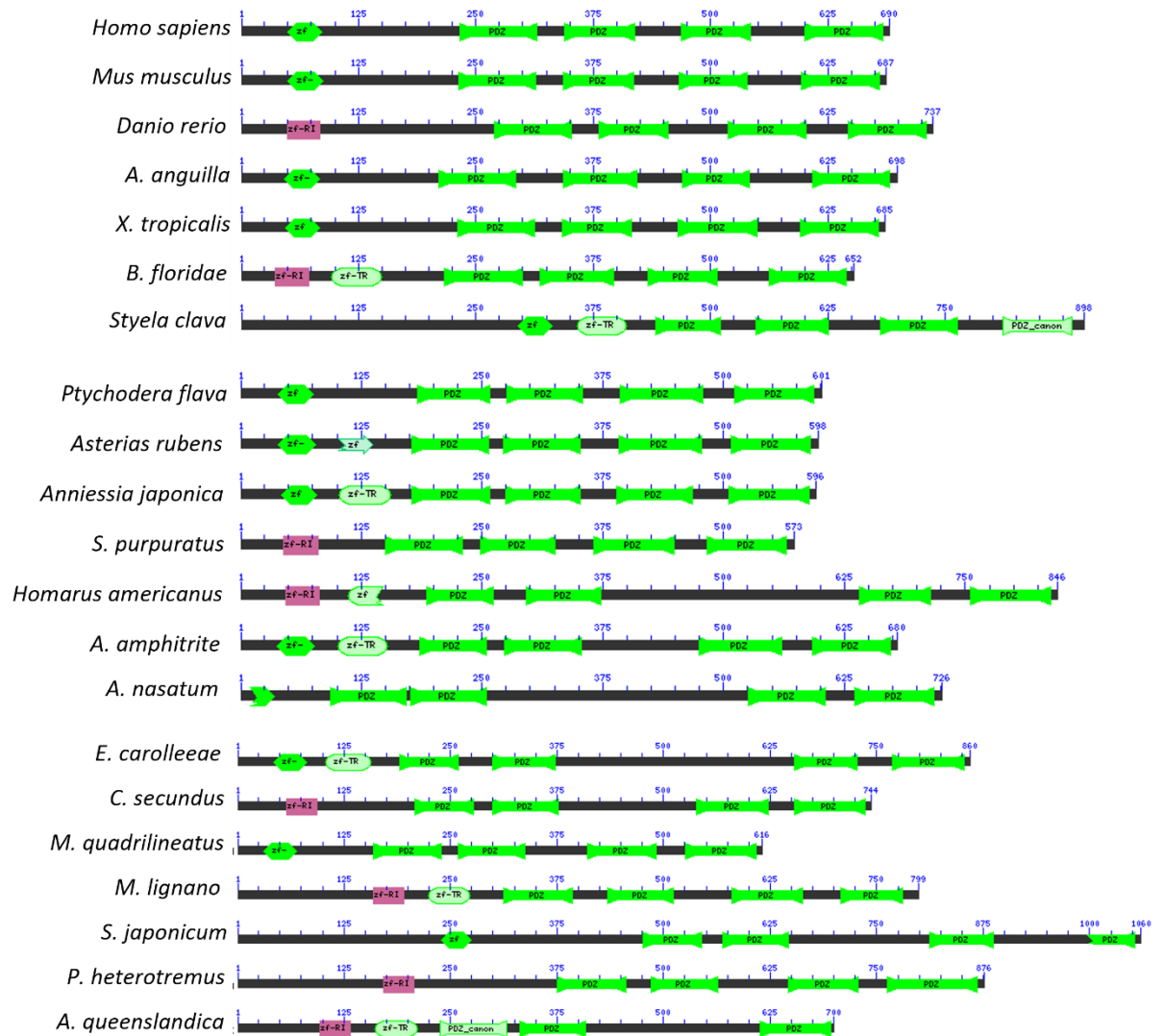


Figure 5.2: Domain architectures of LNX orthologs. Conserved domains (CD) of LNX orthologous proteins found in various metazoan species. Four PDZ domains (green) and a RING type catalytic domain (green=Zinc finger, pink=RING domain) are found across all selected species. *A. queenslandica* did not have four predicted PDZ domains, but other software and publications (Flynn, Saha and Young, 2011) have confirmed there are four. Image created using the NCBI CD-search tool (J. Wang *et al.*, 2023).

5.2.3 Phylogenetic analysis of LNX orthologs reveals sequence clustering to three major clades

After analysing the sequence alignments (Figure 5.1) and domain structures (Figure 5.2) of our chosen 21 LNX orthologous species, we decided to infer a maximum-likelihood phylogenetic relationship. MAFFT v7 (Kato and Standley, 2013) aligned sequences were used to infer phylogeny using FastTree software (Price, Dehal and Arkin, 2010), which implements an approximate maximum-likelihood algorithm optimised for large protein alignments. The resulting topology generated by the interactive tree of life (iTOL) v6 represents the evolutionary relationships among the orthologous proteins and provides an estimate of divergence among the ancestral lineages represented in the dataset (Letunic and Bork, 2024).

Three major clades were revealed from the analysis (Figure 5.3). Clade I consisted of cephalochordates and echinoderms (lancelet, acorn worm, feather star, starfish, and sea urchin), which formed a well-supported group with bootstrap values between 0.946 and 1.00. Clade II was comprised of two separate sub taxa – platyhelminths (blood fluke, flatworm, and lung fluke) and arthropods (American lobster, barnacle, woodlouse, calanoid copepod, termite, and leafhopper). These two groups had strong internal node support implying sequence similarity which reinforces results seen in section 5.2.2. Within clade II, the poriferan (sea sponge) exhibited the longest branch length (1.374), indicating substantial sequence divergence relative to other taxa. Finally, clade III entailed the vertebrate LNX orthologous sequences (human, mouse, frog, eel, and zebrafish) with an expected, closely related sister pair in the mouse and human

sequences. Peculiarly in this clade, a split node resulted in two branches resulting in a divergence between the sea squirt and vertebrate lineages.

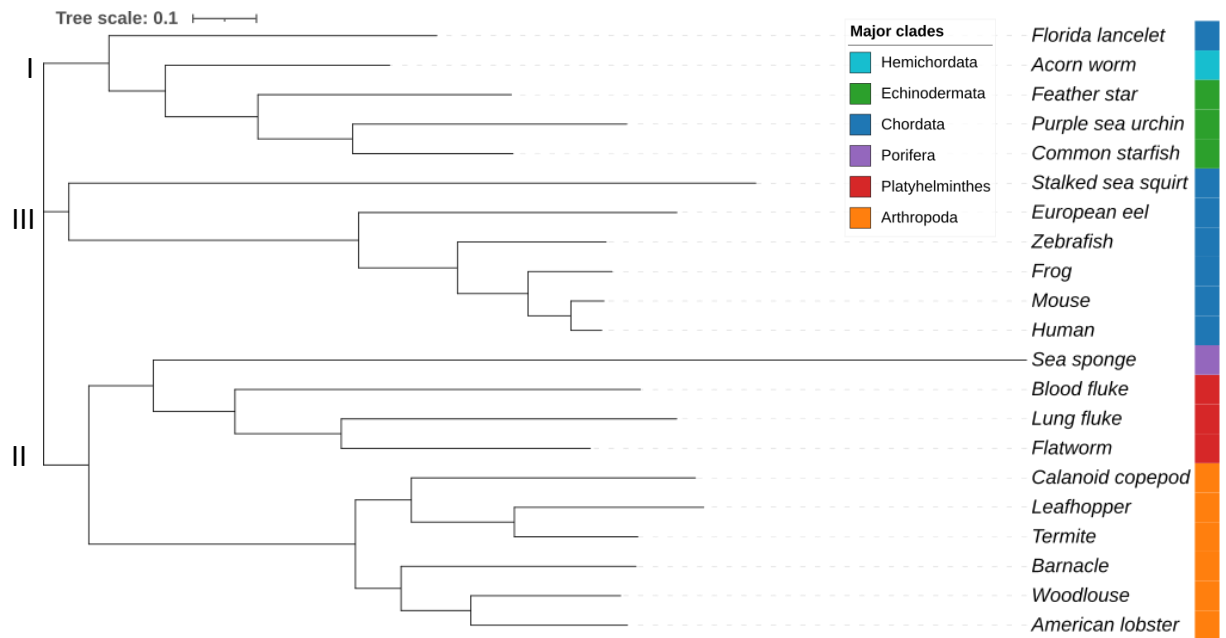


Figure 5.3: Phylogenetic tree illustrating LNX protein evolution throughout the metazoan divergence. Phylogenetic reconstruction was performed using orthologous LNX protein sequences (RefSeq-NCBI) from ancestral lineages. Sequences were aligned with MAFFT using iterative refinement to optimise homologous domain alignment. A maximum-likelihood phylogeny was inferred with FastTree (Price, Dehal and Arkin, 2010) under the WAG+Γ evolutionary model, the output was loaded into iTOL (interactive tree of life) v6 software (Letunic and Bork, 2024) to generate a correctly distanced tree with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The major clades are labelled I, II and III and coloured coded.

5.2.4 The second PDZ domain of LNX is predicted to be spatially more distant from other PDZ domains across metazoan orthologs

From the analysis of LNX2 orthologous sequences thus far, we have determined that all selected proteins contain four well conserved PDZ domains and a RING domain (Figure 5.1 and 5.2). There are no solved crystal structures of LNX proteins to date so structural prediction tools like AlphaFold 3 are an effective way to visualise the 3D confirmation of these domains (Jumper *et al.*, 2021)(Abramson *et al.*, 2024).The predicted confidence of AlphaFold3 interaction models is determined by:

1. pLDDT or per-residue confidence score. This output gives a range from 0-100 and outlines how reliable each residue's position is. Values of >70-80 are suggestive of a meaningful interaction.
2. ipTM or interface predicted template modelling score. AlphaFold outputs this value based on how confident it is on the orientation of the two proteins, a value close to 1 is considered a strong prediction of relative positioning (Zhang and Skolnick, 2004) (Xu and Zhang, 2010).
3. The predicted alignment error (PAE) evaluates how the proteins dock together and their relative positioning, this is determined along the entirety of both chains and between chains. This output is also graphed as a 2D heatmap in which you can visualise the conformation. A PAE score of <1 Å suggests very low predicted error between chains and is a confident interaction.

4. Steric clash or has clash relates to the unnatural overlap of residues in a model. If AlphaFold predicts no clash there is confidence that the conformation of the residues is correct (Abramson *et al.*, 2024).

AlphaFold 3 predicts that PDZ2 of vertebrate (human) LNX2 is spatially more distant from the main body of the protein whereas PDZ1, 3 and 4 are tightly packed together (Figure 5.4A). AlphaFold 3 gives a predicted template modelling (pTM) score which measures the accuracy of the entire structure, a pTM>0.5 is a highly confident model. Along with that predicted error alignment plots (PAE) can be used to two-dimensionally depict spatial arrangements, the darker the green colour, the higher the confidence of the relative positioning of two residues. A light green or white colour (> 5-10 Å) is suggestive of prediction uncertainty but could be illustrating flexible linker regions within the protein.

With a pTM=0.55, AlphaFold 3 is confident in the predicted structure of vertebrate LNX2 and the PAE very nicely illustrates the close association of PDZ1, 3 and 4 in comparison to PDZ2. Similarly, in figure 5.4B *Branchistoma floridae* LNX2 is assigned a pTM=0.51, indicating that PDZ2 is distanced further away from the bulk of the protein. LNX2 orthologs of *Paragonimus heterotremus* (Figure 5.4C) and *Amphimedon queenslandica* (Figure 5.4D) have less confident prediction scores, 0.46 and 0.31 respectively and a clear PAE map for *A. queenslandica* LNX2 is not produced.

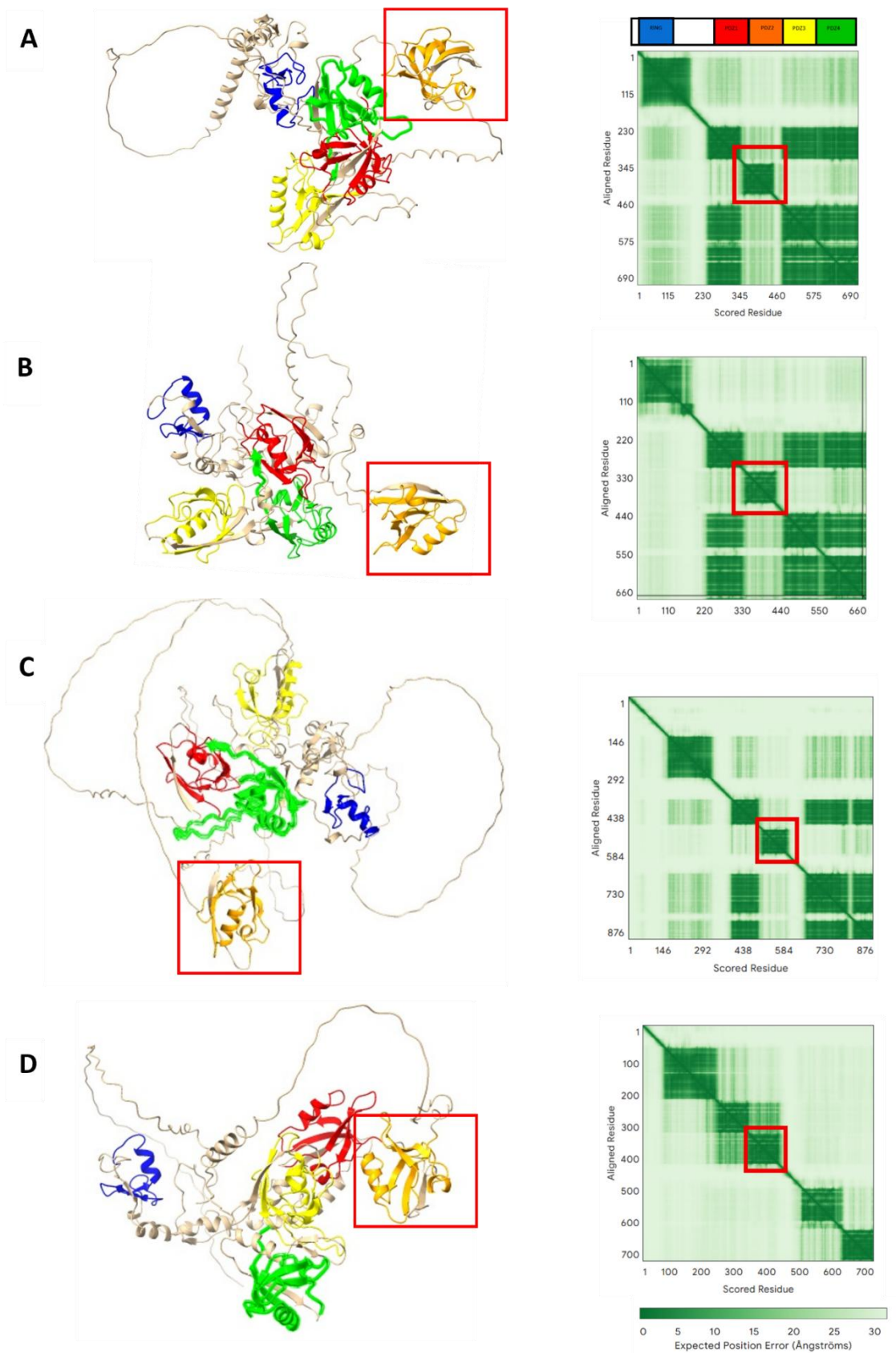


Figure 5.4: AlphaFold 3 predicts PDZ2 as spatially estranged from the body of LNX. LNX proteins were modelled using AlphaFold 3 and domains are colour coded using ChimeraX, RING domain = blue, PDZ1 = red, PDZ2 = orange, PDZ3 = yellow and PDZ4 - C terminus = green. **A.** *Homo sapiens* pTM=0.55 **B.** *Branchistoma floridae* LNX2, pTM= 0.51. **C.** *Paragonimus heterotremus* LNX2, pTM= 0.46. **D.** *Amphimedon queenslandica*, pTM = 0.3.

AlphaFold 3 modelling was carried out on all 21 selected LNX orthologs and prediction TM scores were assigned to each. Representative PAE 2D maps generated by the full-length modelling are also shown to give a clarity of the LNX spatial conformation in relation to itself (Figure 5.5). Most of the predictions were confident giving pTM values over 0.5, while those for *Eurytemora carolleeae*, *Macrostomum lignano*, *Schistosoma japonicum* and *Styela clava* predictions were somewhat less confident. Notably the PAE 2D maps all strikingly support the close association of PDZ1, 3 and 4 while suggesting that PDZ2, along with the RING domain, are more estranged from the other PDZ domains. The observation of this pattern across the PAE 2D maps indicates that this is an evolutionarily conserved characteristic of LNX proteins.

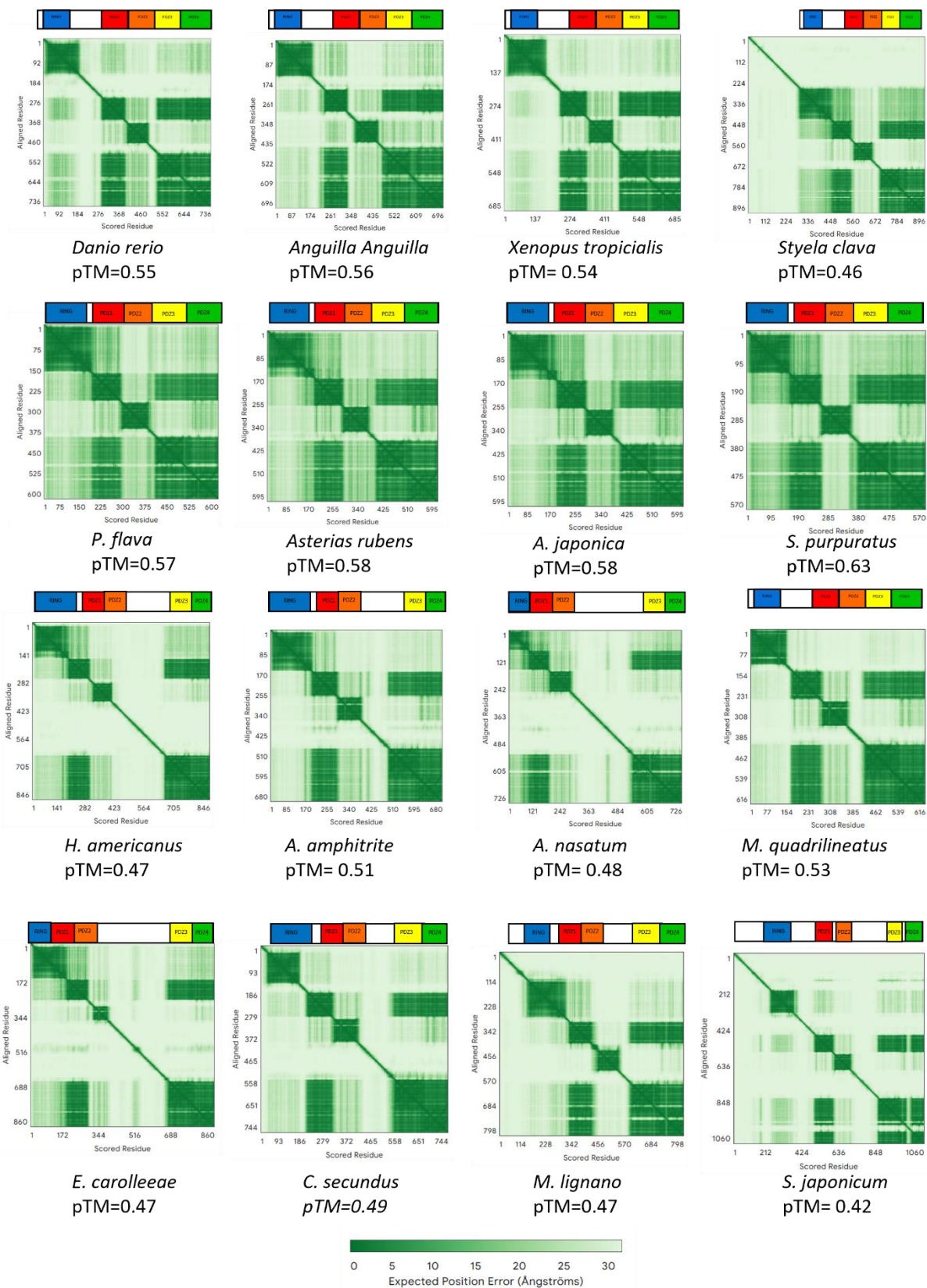


Figure 5.5: AlphaFold 3 predicted error alignment graphs for full length LNX2 orthologs. Protein domains are emphasised above each PAE graph. LNX2 ortholog species name and predicted template modelling (pTM) scores are exhibited beneath.

5.2.5 Conservation of intramolecular interaction of carboxyl terminal PBM and first PDZ domain of LNX2

The ability of the first PDZ domain of LNX proteins to bind to the carboxyl terminal PDZ motif of a second LNX moiety was first noticed by Rice *et al* in 2001. The group published work that revealed PDZ1 of LNX1 and LNX2 could dimerise *via* their own carboxyl terminus and also described RING to RING domain homodimerization of LNX2 (Rice, Northcutt and Kurschner, 2001). To investigate the evolutionary preservation of these features we employed AlphaFold 3 to run double chain interaction modelling.

Firstly, each LNX2 orthologous sequence was screened through Wregex v3.0 server to identify short linear motifs (SLiMs) that may be present (Prieto, Rodríguez and Fullaondo, 2024). A class I PBM was identified at the C-terminus of every ortholog except *P. flava*, *Styela clava* and *A. queenslandica* (Table 5.2). Next, the sequence of the first PDZ domain was determined by SMART (EMBL) from each ortholog (Letunic and Bork, 2025). The first PDZ domain, usually 70-80 amino acids in length, was then modelled against the carboxyl terminal of that sequence and the confidence of interaction score was evaluated. AlphaFold 3 prediction is very strict on shorter sequences so only PDZ1 and the final 10 amino acids were inputted to give accurate results.

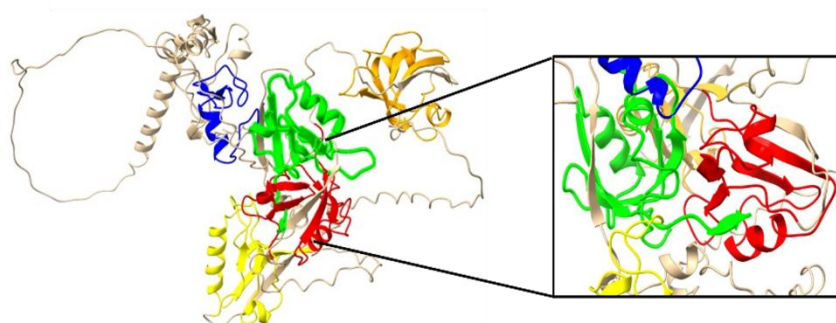
Predicted error alignments, iPTM and overall ranking scores of the pairs were generally quite high for all sequences except *P. flava* (0.57) and *A. queenslandica* (0.58), which was as expected due to the absence of a PBM consensus, suggesting in most tested LNX2 orthologs, PDZ1 can bind their own carboxyl terminus.

Table 5.2. Auto binding of PDZ1 and carboxyl terminal PBMs of LNX orthologs. AlphaFold3 prediction scores. LNX2 carboxyl terminal tails of each species are outlined. Those that lack a consensus PBM (according to Wregex v3.0) have their last three residues highlighted in red text. AlphaFold 3 outputs are colour coded based on lowest (0.5) to highest (1.0) confidence. 0.5  1.0

Species	C-terminal tail	chain_pair_pae_min	chair_pair_iptm	has_clash	ranking_score
<i>Homo sapiens</i>	TVICWPGSLV	[0.76,1.23],[1.06,0.76]	[0.02,0.68],[0.68,0.81]	0	0.76
<i>Mus musculus</i>	TVICWPGSLV	[0.76,0.82],[0.88,0.76]	[0.87,0.88],[0.81,0.02]	0	0.88
<i>Danio rerio</i>	TVVSWPGSLI	[0.76,1.07],[1.13,0.76]	[0.88,0.79],[0.79,0.03]	0	0.87
<i>Anguilla anguilla</i>	TVVSWPGSLV	[0.76,1.33],[1.4,0.76]	[0.53,0.8],[0.8,0.02]	0	0.85
<i>Xenopus tropicalis</i>	SVISWPGSLV	[0.76,0.81],[0.86,0.76]	[0.89,0.83],[0.83,0.03]	0	0.9
<i>Branchiostoma floridae</i>	AVVSWPGTML	[0.76,1.16],[1.1,0.76]	[0.02,0.68],[0.68,0.76]	0	0.75
<i>Styela clava</i>	KLVSWPGLH	[0.76,2.09],[1.73,0.76]	[0.86,0.71],[0.71,0.02]	0	0.8
<i>Ptychodera flava</i>	SVVHWPGSVH	[0.76,4.8],[5.22,0.76]	[0.82,0.44],[0.44,0.02]	0	0.57
<i>Asterias rubens</i>	TVVSWPGTAV	[0.76,3.23],[3.25,0.76]	[0.87,0.59],[0.59,0.02]	0	0.7
<i>Anneissia japonica</i>	TVVSWPGSLL	[0.76,0.98],[0.96,0.76]	[0.87,0.79],[0.79,0.03]	0	0.87
<i>Strongylocentrotus purpuratus</i>	TIVSWPGTIV	[0.76,0.87],[0.9,0.76]	[0.82,0.77],[0.77,0.02]	0	0.85
<i>Homarus americanus</i>	EVVSWLGTEL	[0.76,1.13],[1.28,0.76]	[0.78,0.71],[0.71,0.02]	0	0.8
<i>Amphibalanus amphitrite</i>	QIVSWMGTEV	[0.76,0.99],[1.0,0.76]	[0.75,0.69],[0.69,0.02]	0	0.76
<i>Armadillidium nasatum</i>	EVVSWMGTEL	[0.76,1.03],[1.1,0.76]	[0.75,0.67],[0.67,0.02]	0	0.75
<i>Eurytemora carolleae</i>	EVVSWMGTEL	[0.76,1.34],[1.26,0.76]	[0.8,0.72],[0.72,0.02]	0	0.79
<i>Cryptotermes secundus</i>	EVVSWLGTEL	[0.76,1.05],[1.1,0.76]	[0.8,0.74],[0.74,0.02]	0	0.82
<i>Macrosteles quadrilineatus</i>	EIVSWLGTEI	[0.76,1.19],[1.36,0.76]	[0.75,0.68],[0.86,0.02]	0	0.75
<i>Macrostomum lignano</i>	TIVSWPGTIV	[0.76,0.95],[0.95,0.76]	[0.82,0.77],[0.77,0.02]	0	0.84
<i>Schistosoma japonicum</i>	QVISWPGTIV	[0.76,0.91],[0.99,0.76]	[0.86,0.77],[0.77,0.02]	0	0.85
<i>Paragonimus heterotremus</i>	RVVSWPGTIV	[0.76,0.91],[0.9,0.76]	[0.87,0.81],[0.81,0.02]	0	0.88
<i>Amphimedon queenslandica</i>	ITFTIIPNDS	[0.76,4.44],[3.98,0.76]	[0.02,0.48],[0.48,0.65]	0	0.58

A subset of sequences representing the highest confidence scores (*Homo sapiens*, *Branchiostoma floridae*, and *Paragonimus heterotremus*) and the lowest confidence score (*Amphimedon queenslandica*) were selected for structural modelling of full-length LNX proteins using ChimeraX (Figure 5.6). These taxa were chosen as representative species from major metazoan clades, including vertebrates, cephalochordates, platyhelminths, and poriferans. Structural visualisation revealed that in high confidence models (Figure 5.6A-C), the carboxyl terminal PBM (shown in green) is positioned within the binding groove of the PDZ1 domain (shown in red). The observation of this interaction in the full-length protein models indicates that the carboxyl-terminal PDM can bind to PDZ1 within a single LNX molecule in an intramolecular manner. In contrast, this interaction was not observed in the *A. queenslandica* model (Figure 5.6D), where the carboxyl-terminus, which lacks a consensus PBM, did not localise within the PDZ1 binding pocket.

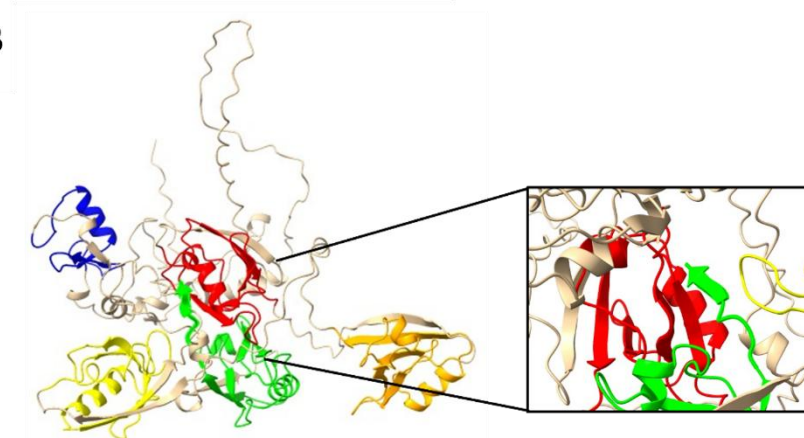
A



Homo sapiens

Score: 0.76

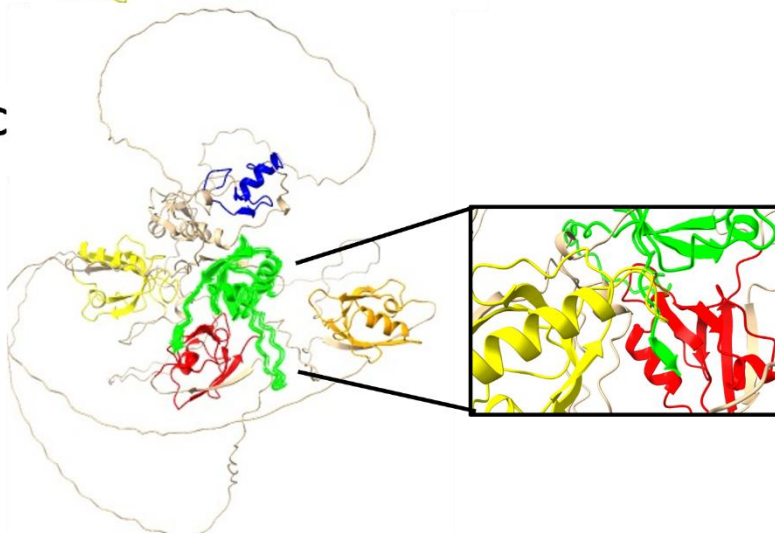
B



B. floridae

Score: 0.75

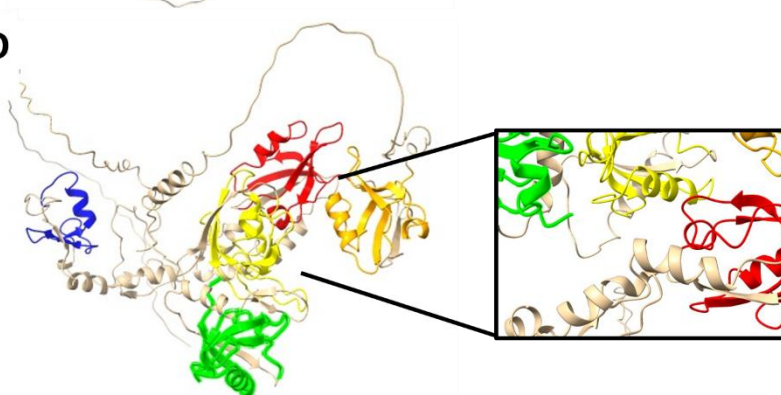
C



P. heterotremus

Score: 0.88

D



A. queenslandica

Score: 0.58

Fig 5.6: AlphaFold3 models of LNX proteins predict intramolecular binding. Predicted 3D structure of LNX2 PDZ1 (red) and carboxyl terminal PBM (green) of highlighted orthologous sequences. ChimeraX software was used to annotate the AlphaFold 3 outputs. **A.** *Homo sapiens*, **B.** *Branchiostoma floridae*, and **C.** *Paragonimus heterotremus* show PBM interaction within the binding pocket of PDZ1. **D.** *Amphimedon queenslandica* PDZ1 is not spatially close to its own C-terminus suggesting no interaction.

Next, *Branchiostoma floridae* and *Mus musculus* LNX proteins were verified for self-binding using experimental methods. Yeast two hybrid assays were conducted (Methods section 2.7.2) using truncated LNX proteins to investigate if activation of the *His3* and *LacZ* reporter genes occurs due to an interaction between PDZ1 and carboxyl terminus PDZ3-4. The entire carboxyl end of LNX (PDZ3-4) was used during this analysis, as longer truncated mutants are more physiologically relevant in interaction studies.

Yeast cells were co-transformed with *M. musculus* (*Mm*) LNX PDZ3-4 (prey) and either PDZ1 or RING-PDZ2 (bait) as a positive control, and *B. floridae* (*Bf*) LNX PDZ1 was transformed with its own carboxyl terminus (PDZ3-4). Both vertebrate (Figure 5.7A) and invertebrate (Figure 5.7B) LNX proteins showed self-binding, with no background interaction in the negative control.

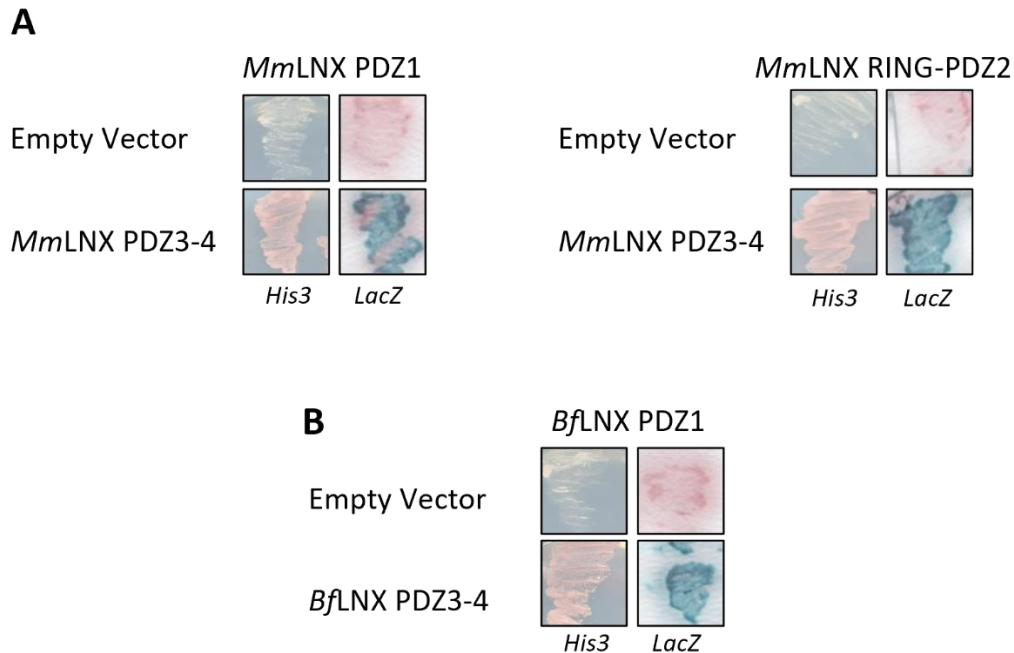


Figure 5.7: Experimental validation of *B. floridae* LNx auto binding by way of yeast-two hybrid analyses. **A.** *Mus musculus* (*Mm*) LNx PDZ1 or RING-PDZ2 were co-transformed into L40 strain *S. cerevisiae* with PDZ3-4 or empty vector and showed interaction only in the presence of the PDZ3-4 domain. **B.** The same was carried out for *B. floridae* (*Bf*) LNx PDZ1 with PDZ3-4, in which activation of both *His3* and *LacZ* reporter genes was observed. Pink growth on histidine (*His3*) lacking media and a blue colour change when yeast were cells were subjected to X-gal substrate (*LacZ*).

5.2.6 Homodimerization of RING domains is a conserved feature of LNx orthologs

In the ubiquitination system, E3 ligases often auto ubiquitinate themselves as a method of regulating their catalytic activity. E3 ligases can do this by self-ubiquitination or by sequestering external ligases (De Bie and Ciechanover, 2011). *Mus musculus* LN2 was discovered to form homomeric structures by the oligomerization of their RING domains (Rice, Northcutt and Kurschner, 2001) and confirmed in crystal structures (Nayak *et al.*, 2015)(Nayak and Sivaraman, 2018). In the attempt to provide some clarity to the

evolutionary function of LNX proteins we again, adopted AlphaFold 3 to predict metazoan LNX RING dimerization.

From the ipTM, PAE and ranking scores produced from AlphaFold 3 in Table 5.3, we can determine that LNX2 orthologs from the studied species can interact with themselves *via* their RING domains except for *Macrosteles quadrilineatus*, *Amphimedon queenslandica* orthologs. Remarkably, *Homo sapiens* LNX2 RING domain dimerization was not confidently predicted (ranking score of 0.36), even though mouse and human RING domain sequences only differ by a few amino acids.

Table 5.3: Prediction of RING: RING dimerization in LNX2 orthologs. RING domain sequences of each LNX2 protein were found from SMART (EMBL). RING domain dimerization was determined by AlphaFold 3, and output scores were colour coded based on the confidence likelihood of interaction. 0.2  1

Species	chain_pair_pae_min	chair_pair_iptm	has_clash	ranking_score
<i>Homo sapiens</i>	[0.76,9.98],[9.9,0.76]	[0.48,0.27],[0.27,0.49]	0	0.36
<i>Mus musculus</i>	[0.76,3.99],[4.0,0.76]	[0.52,0.44],[0.44,0.52]	0	0.67
<i>Danio rerio</i>	[0.76,4.49],[4.83,0.76]	[0.57,0.49],[0.49,0.56]	0	0.7
<i>Anguilla anguilla</i>	[0.76,4.62],[4.65,0.76]	[0.49,0.41],[0.41,0.48]	0	0.63
<i>Xenopus tropicalis</i>	[0.76,5.33],[5.35,0.76]	[0.58,0.42],[0.42,0.58]	0	0.5
<i>Branchiostoma floridae</i>	[0.76,5.88],[6.11,0.76]	[0.53,0.42],[0.42,0.53]	0	0.65
<i>Styela clava</i>	[0.76,1.38],[1.39,0.76]	[0.67,0.74],[0.74,0.67]	0	0.91
<i>Ptychodera flava</i>	[0.76,4.73],[4.88,0.76]	[0.54,0.45],[0.45,0.54]	0	0.67
<i>Asterias rubens</i>	[0.76,3.0],[3.08,0.76]	[0.58,0.6],[0.6,0.58]	0	0.75
<i>Anneissia japonica</i>	[0.76,6.6],[6.81,0.76]	[0.4,0.26],[0.26,0.39]	0	0.55
<i>Strongylocentrotus purpuratus</i>	[0.76,4.03],[4.04,0.76]	[0.54,0.48],[0.48,0.54]	0	0.69
<i>Homarus americanus</i>	[0.76,3.66],[3.68,0.76]	[0.47,0.43],[0.43,0.47]	0	0.57
<i>Amphibalanus amphitrite</i>	[0.76,8.14],[8.05,0.76]	[0.41,0.29],[0.29,0.42]	0	0.47
<i>Armadillidium nasatum</i>	[0.76,2.85],[2.82,0.76]	[0.47,0.51],[0.51,0.47]	0	1.02
<i>Eurytemora carolleeae</i>	[0.58,0.54],[0.54,0.58]	[0.76,3.29],[3.15,0.76]	0	0.74
<i>Cryptotermes secundus</i>	[0.76,3.82],[3.78,0.76]	[0.52,0.51],[0.51,0.52]	0	0.71
<i>Macrosteles quadrilineatus</i>	[0.76,7.67],[7.89,0.76]	[0.46,0.23],[0.23,0.46]	0	0.27
<i>Macrostomum lignano</i>	[0.76,5.63],[5.61,0.76]	[0.5,0.37],[0.37,0.51]	0	0.58
<i>Schistosoma japonicum</i>	[0.76,6.34],[6.2,0.76]	[0.5,0.37],[0.37,0.49]	0	0.51
<i>Paragonimus heterotremus</i>	[0.76,7.29],[7.43,0.76]	[0.51,0.25],[0.25,0.51]	0	0.47
<i>Amphimedon queenslandica</i>	[0.76,8.84],[8.83,0.76]	[0.43,0.15],[0.15,0.42]	0	0.33

5.2.7 “What’s in a name?” – analysing the ancestral LNX-NUMB interaction

Ligand of NUMB protein X1 and X2 (LNX1/2) were originally characterised based on their ability to interact with the cell fate determinant NUMB via its conserved NPAY/F motif (Dho *et al.*, 1998). This interaction led to the identification and naming of the LNX protein family. Since its initial discovery, LNX1/2 proteins have been implicated in multiple cellular processes, most notably in neuronal development, where NUMB plays a critical regulatory role (Young, 2018).

The lack of evolutionary conservation of the NUMB-binding NPAY/F motif in invertebrate LNX sequences (Figure 5.1) suggests that the NUMB interaction may be unique to vertebrates. However, I wanted to test this experimentally. Therefore, amylose pull-down assays (Methods section 2.5.3) were performed using purified LNX proteins from a vertebrate species (*Mus musculus*) and an invertebrate species (*Branchiostoma floridae*), together with their corresponding NUMB proteins. Analysis of the pull-down results demonstrated a clear interaction between mouse MBP-LNX and Flag-NUMB proteins and not with the MBP moiety alone. Dissimilarly, no detectable interaction was observed between the invertebrate LNX and NUMB homologs under the same experimental conditions (Figure 5.8).

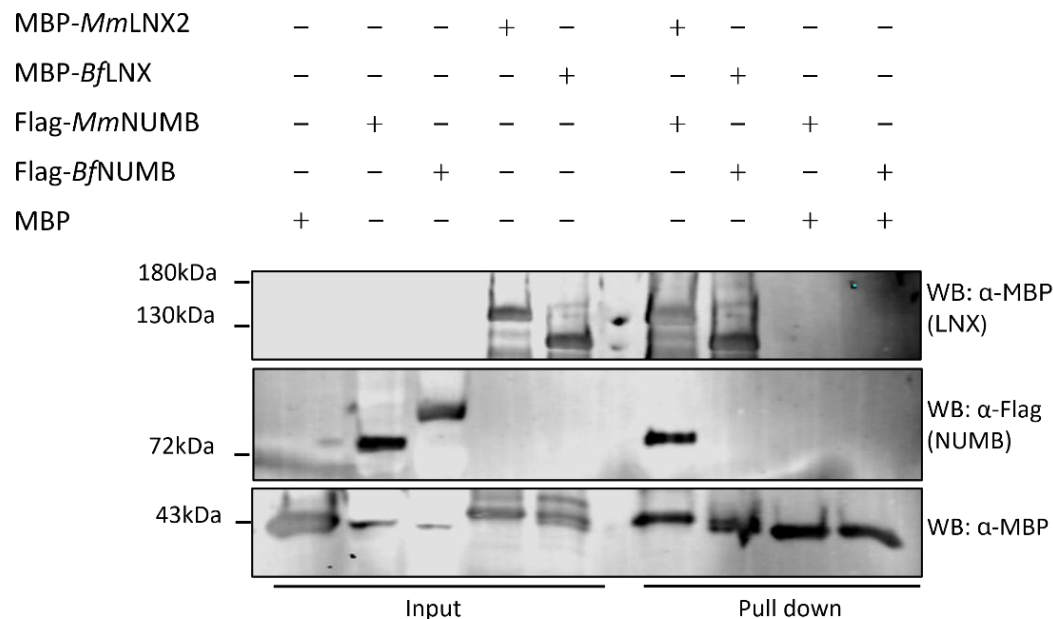


Figure 5.8: *Branchiostoma floridae* LNK does not interact with NUMB protein *in vitro*. MBP tagged LNK proteins of both *Mus musculus* (*Mm*) and *Branchiostoma floridae* (*Bf*) were expressed in BL21 *E. coli* cells, purified and dialysed to remove excess salt as per methods section 2.5. Flag tagged NUMB proteins from both *Mm* and *Bf* were expressed in HEK 293T cells. Protein concentrations were determined and equal amounts of each protein were mixed along with amylose beads. Binding was allowed to occur for 4 hrs at 4°C before being washed and eluted using 2X Laemmli buffer and boiling. Pull downs were analysed by Western and immunoblotting with the indicated antibodies. Clear pull down of *Mm*NUMB by *Mm*LNK2 can be observed at the correct molecular weight but no band is seen for *Bf*NUMB or the MBP moiety alone. Blot is representative of three independent experiments.

5.3 Discussion

The evolution of LNX1/2 proteins has only been studied at a very basic level to date. The first study of LNX protein conservation was conducted in 2011 in which the first LNX1/2 orthologs were identified in early metazoan lineages (Flynn, Saha and Young, 2011). As vertebrate LNX1 and 2 proteins contain a RING domain an NPAY/F motif and four nearly equally spaced PDZ domains, we attempted to supplement what was already known about LNX1/2 evolution and delve into characterising structurally conserved features (Rice, Northcutt and Kurschner, 2001).

The availability of annotated genomes has rapidly increased since the annotation of the human genome in 2004 (Abdellah *et al.*, 2004). Several early metazoan ancestral species have become accessible for researchers such as those from platyhelminth, arthropod, poriferans and echinoderm taxa. Well conserved LNX orthologs have been selected from 21 species in these lineages (Figure 5.1, 5.2 & S2). The overall domain architecture of these orthologs has been predicted to be almost identical to the vertebrate proteins in species like *Homo sapiens*, *Mus musculus*, *Danio rerio*, *Anguilla Anguilla* and *Xenopus tropicalis* along with some invertebrate species like *Branchiostoma floridae*, *Ptychodera flava*, *Asterias rubens* and *Strongylocentrotus purpuratus*. However, variation in splicing and translational selection is common throughout protein evolution. Disordered regions evolve three times faster than buried domains and can independently from other domains/regions in the same sequence (Dubreuil and Levy, 2021). In species such as *Styela clava* and *Schistosoma japonicum*, larger interprotein spaces between domains compared to vertebrate LNX2 orthologs are seen suggesting a possible splicing variation

or the loss of intrinsically disordered regions in higher chordate species in comparison to hemichordates or platyhelminths. Predominantly, the study of LNX2 orthologs from species selected revealed comparable domain structures. Conservation of structural architecture across metazoans suggests strong evolutionary constraint on LNX function, supporting an ancient and potentially conserved regulatory role (Flynn, Saha and Young, 2011).

We also demonstrate that the three-dimensional structure of LNX proteins has been well maintained throughout evolution. From an extensive review of the literature and ProfAff database searches, PDZ2 of both human and mouse LNX1/LNX2 are responsible for the majority of interactions (Wolting *et al.*, 2011)(Guo *et al.*, 2012) (Lenihan *et al.*, 2017)(Lenihan, Saha and Young, 2017)(Gogl *et al.*, 2022). In this chapter, we investigated a potential reason behind this. It was discovered that the predicted 3D conformation of *Homo sapiens* LNX2 included a spatially distant PDZ2 domain (Figure 5.4A). The distance of PDZ2 from the main protein body could make it more accessible to interacting proteins, compared to the more condensed PDZ1,3 and 4. This structural design is conserved across all the examined lineages, however some AlphaFold 3 predictions are not highly confident. We cannot confirm that the 3D arrangements of *Eurytemora carolleeae*, *Macrostomum lignano*, *Schistosoma japonicum*, *Styela clava* or *Amphimedon queenslandica* are conserved based on their low pTM scores (Figure 5.4D & 5.5). A protruding PDZ2 is evident in all other LNX2 orthologs which can be considered as functionally important to LNX function throughout evolution. Evidence of preserved LNX

PDZ2 protein interactions among these orthologs could confirm the importance of this structural feature throughout evolution.

Similarly, intramolecular binding between PDZ1 and the carboxyl terminal PBM is present in vertebrates, cephalochordates, platyhelminths, echinoderms and arthropods but not poriferans. The self-binding characteristic of LNX proteins imply a method of maintaining protein structural rigidity, which we experimentally confirmed in the invertebrate proxy, *B. floridae* (Figure 5.7). As an E3 ubiquitin ligase enzyme, this could a method of auto-regulation as seen with SMURF1 forming a head-to-tail dimer (Vittal *et al.*, 2015). Another enzymatic approach to self-regulating is self-catalysed ubiquitination which is frequently employed by CDH1 during the G₀ and G₁ phase of the cell cycle within the context of the APC/C complex (De Bie and Ciechanover, 2011). RING dimerization is predicted in nearly all the studied metazoan LNX orthologs (Table 5.3). These intramolecular interactions may be necessary for controlling the ubiquitination of LNX substrates in most metazoan lineages, but not in *A. queenslandica*.

From the multiple sequence alignment in Figure 5.1, all vertebrate LNX2 orthologs that were tested contained a well-maintained NUMB binding, NPAF motif. None of the invertebrate metazoan LNX orthologs contained such a motif (Flynn, Saha and Young, 2011). Again, the functionality of this interaction was demonstrated in *M. musculus* and *B. floridae* (Figure 5.8) and it was determined that the LNX-NUMB association is not present in the closely related lancelet species. While this result was expected based on the lack of an NPAF motif, it was possible that the LNX: NUMB interaction might be mediated by some of the other domains within these two proteins. These experimental

results indicate that the NUMB–LNX interaction is not conserved in the invertebrate species tested, suggesting that this regulatory association may have arisen or been strengthened during vertebrate evolution.

These findings suggest that the LNX-Numb interaction may have been specifically required for central nervous system development in vertebrates (Nie *et al.*, 2002b), whereas it is dispensable for dorsal hollow nerve cord formation in invertebrates. Results suggest a lineage specific neofunctionalisation of the NUMB interaction with LNX. However, further studies are needed to elucidate the evolutionary pressures maintaining LNX structure and its broader functional significance.

Chapter 6

Characterisation of LNX Expression and Promoter Responsiveness in Colorectal and Glial Cancers

6.1 Introduction

Expression of *Ln timer*1/2 mRNAs is widely distributed across multiple adult mouse tissues, with early embryonic expression of both genes detected in the central nervous system (Dho *et al.*, 1998) (Rice, Northcutt and Kurschner, 2001). Contrastingly, LNX1 and LNX2 protein levels remain poorly defined. Endogenous LNX1 protein was first identified in perisynaptic Schwann cells at neuromuscular junctions (Young *et al.*, 2005). Additionally, both LNX1 and 2 proteins have been detected in spermatozoa and LNX2 protein has been reported to be expressed in a subset of blood vessels (Mirza *et al.*, 2005)(Mirza *et al.*, 2006). Generally for LNX1/2 protein detection by Western blotting, prior immunoprecipitation is required (Mirza *et al.*, 2005)(Weiss *et al.*, 2007) (Young *et al.*, 2005) suggesting that despite widespread presence of *LNX1/2* mRNAs, corresponding proteins are expressed at very low levels.

The presence of regulatory elements in the untranslated regions (UTRs) can divert away from translation of the main coding sequence (Wethmar, Smink and Leutz, 2010). Studies have shown that conserved upstream open reading frames (uORFs) are present in the 5'UTR of *LNX1/2* genes in vertebrate species (Lenihan *et al.*, 2014). The presence of multiple uORFs in the *LNX1^{p70}* gene serves as a method to negatively regulate translation. Upon AUG to AAA mutation in the five uORFs in the 5'UTR of *LNX1^{p70}*, luciferase activity was reported to be 2-3-fold higher (Lenihan *et al.*, 2014). This mechanism for controlling translation was not observed in *LNX1^{p80}* or *LNX2*. However, LNX1^{p80} protein levels appear to double following proteasomal inhibition using MG132. Similarly, LNX1p80 protein levels decline notably after the inhibition of protein synthesis *via* cycloheximide

treatment, suggesting LNX1^{p80} protein is turned over within 8 hours (Lenihan *et al.*, 2014).

There have been multiple studies that claim LNX1/2 proteins have a role in the development of glioblastoma (LNX1) and colorectal cancers (LNX1 and LNX2) (Camps *et al.*, 2013)(Ma *et al.*, 2017)(Baisiwala *et al.*, 2020)(Yan *et al.*, 2021)(Jang *et al.*, 2021). *LNX2* is mapped to chromosome 13q12, a region which is frequently amplified in cancers of the colon and rectum (Camps *et al.*, 2013). In the same study, the authors demonstrate how this amplified region correlates with increased *LNX2* mRNA expression and subsequently protein expression in CRC cell lines. Surprisingly, there was no requirement for immunoprecipitation of LNX2 from these cell lines to detect endogenous protein by Western blotting. Multiple genes (680) were also found to be dysregulated upon siRNA *LNX2* silencing, many of which were involved in the WNT/ β -catenin and NOTCH signalling pathways. A separate group suggested that LNX1 protein can indirectly control the ubiquitination/degradation of the tumour suppressor protein SuFu, thus promoting activation of Hedgehog signalling in CRC cell lines (Yan *et al.*, 2021). This is contradictory to what is reported by Ma *et al.*, in which it was reported that LNX1 serves as a negative regulator for the formation of cancer stem cells and thus is important to halt CRC progression (Ma *et al.*, 2017). See Chapter 1, Section 1.3.7.2 for more information.

In a subset of 15% of glioblastoma multiforme (GBM) tumours, the human chromosomal region where *LNX1* is encoded (4q12), is amplified (Holtkamp *et al.*, 2007). On inspection, 10% of 90 patient glial tumour tissue samples had *LNX1* gene amplification (Blom *et al.*, 2008). However, conflicting reports suggest *LNX1* expression is downregulated in human

gliomas (Chen *et al.*, 2005). Upon temozolomide therapy in patient derived xenograft mouse models, *LNK1* mRNA expression increases as does symmetric stem cell divisions (Baisiwala *et al.*, 2020). A similar study in lung cancer cell lines revealed knocking out *LNK1* delays cell cycle progression (Jang *et al.*, 2021). A more detailed literature review is described in Chapter 1, Section 1.3.7.3.

LNK3/PDZRN3 is less studied in a tumorigenic setting compared to its family members *LNK1* and *LNK2*. One study demonstrates how lower expression of *LNK3* in endometrial carcinomas allows increased proliferative ability and metastasis compared to those with lower expression. The group suggest this function is based on activation of the WNT signalling pathway in the absence of the E3 ligase *LNK3* (Qihong Li *et al.*, 2022). From reviewing the literature, there is little known about *LNK3* relevance in CRC except for its ability to ubiquitinate *TBX20*, a tumour suppressor that inhibits non-homologous DNA end joining (NHEJ) mediated DNA repair in CRC (Luo *et al.*, 2022).

Aims of Chapter 6

- Analyse new publicly available expression datasets, to link amplification of *LNK1* and *LNK2* genetic loci to increased expression.
- Investigate both *LNK* mRNA and protein expression in cancer cell lines.
- Potentially understand how *LNK* transcription might be regulated by studying promoter region activity.

6.2 Results

6.2.1 Analysis of chromosomal amplification and gene expression at the LNX1 and LNX2 loci in GBM and CRC

GBM is has a poor survival rate even with aggressive radio- and chemotherapy in patients, so knowledge of prognostic markers and drug targets is necessary. *LNK1* is mapped to chromosome 4q12 in humans (Figure 6.1A), a region known to be dysregulated in subset of GBM tumours (Holtkamp *et al.*, 2007). Among the *LNK1* gene, there are several other tyrosine kinase genes like platelet-derived growth factor receptor (PDGFR α) and KIT which have been previously reported to be amplified in GBM tumours (Joensuu *et al.*, 2005). The insulin-like growth factor-binding protein 7 (IGFBP7) is also adjacent to these genes and has been assessed as a biomarker for GBM tumour vessels (Iqbal *et al.*, 2010). The transcription factor homeodomain-only protein (HOPX) has recently been identified as an orchestrator of growth and differentiation in invading GBM cells (Doroszko *et al.*, 2025).

Further assessment of the amplification of Chr4q12 and if it leads to increased gene expression in gliomas (GBM and LGG-lower grade glioma) was carried out. RNA-seq data was analysed using the Xena Browser platform, specifically drawing from The Cancer Genome Atlas (TCGA) datasets ("The Cancer Genome Atlas Program (TCGA) - NCI," date accessed 20-03-2026.) The gene expression data was subsequently processed in RStudio to generate comparative heatmaps illustrating the relative enrichment of genes located within the same chromosomal locus. For visualisation, genes were clustered based on their expression levels, while patient samples were grouped according to shared

expression patterns across the selected gene set. This approach enabled direct comparison of expression profiles within the locus across different tumour samples.

As expected, in Figure 6.1B we see extremely high expression of *IGFBP7*, *HOPX* and *PDGFR α* genes. *LNX1* did not have increased expression in GBM or LGG compared to normal tissues relative to other genes in the locus, suggesting that *LNX1* is a bystander in this genomic region amplification and is not likely to influence the progression of GBM or LGG. Interestingly, the antisense-1 and -2 *LNX1* genes had notably decreased expression compared to the other genes in this locus. But a role or effect this has on glial cancers is yet to be investigated.

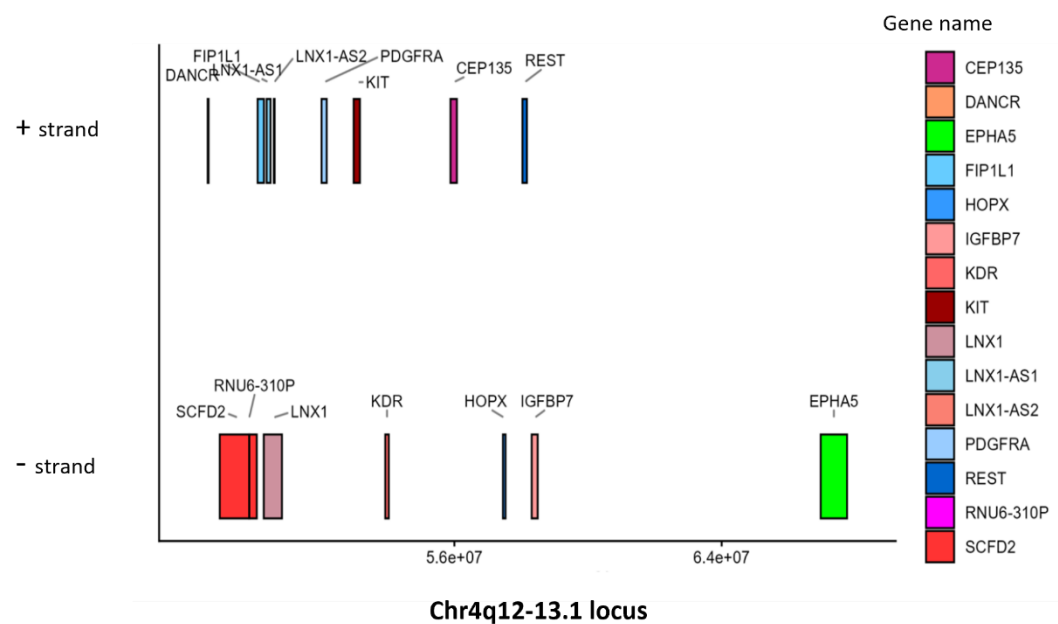


Figure 6.1A: Locus map of chromosomal region 4q12. A schematic of the genes mapped to the Chr 4q12-13 locus, and their relative position to each other, which was created using locuszoomr R package (Lewis and Wang, 2024). *LNX1* is encoded on the – strand while *LNX1-AS1* and *AS2* are located on the + strand.

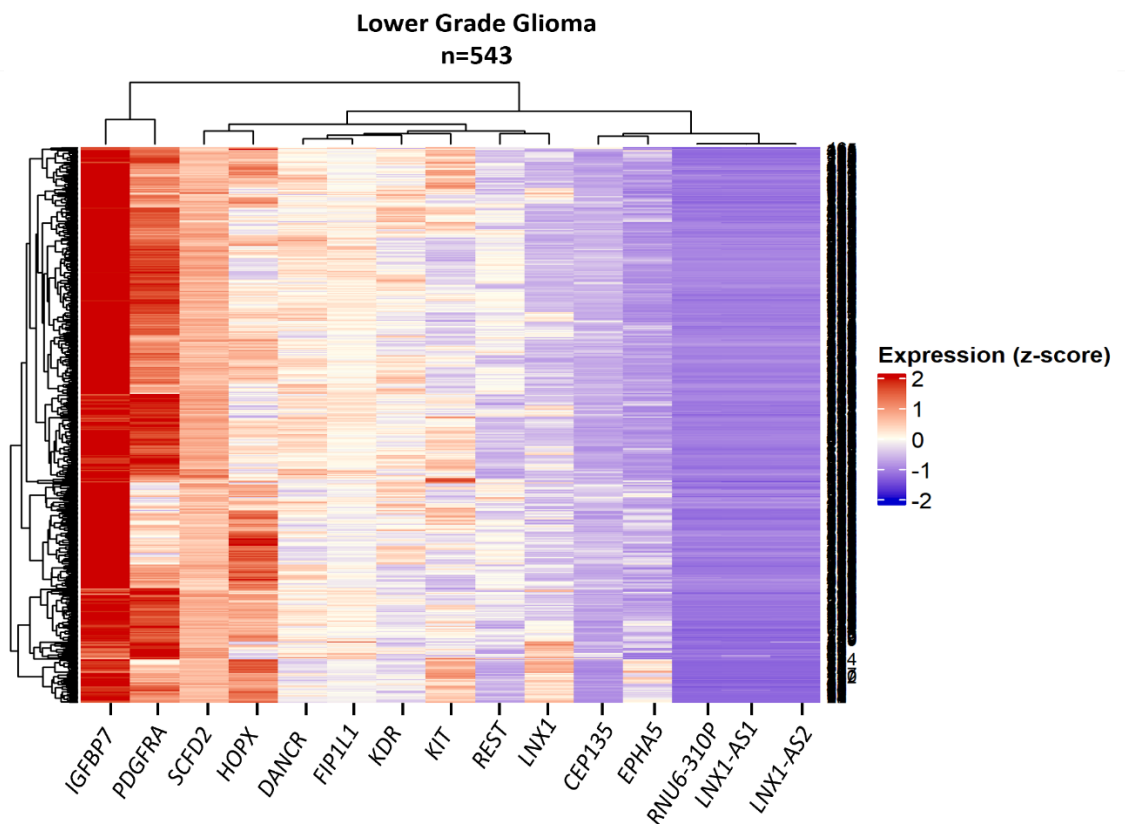
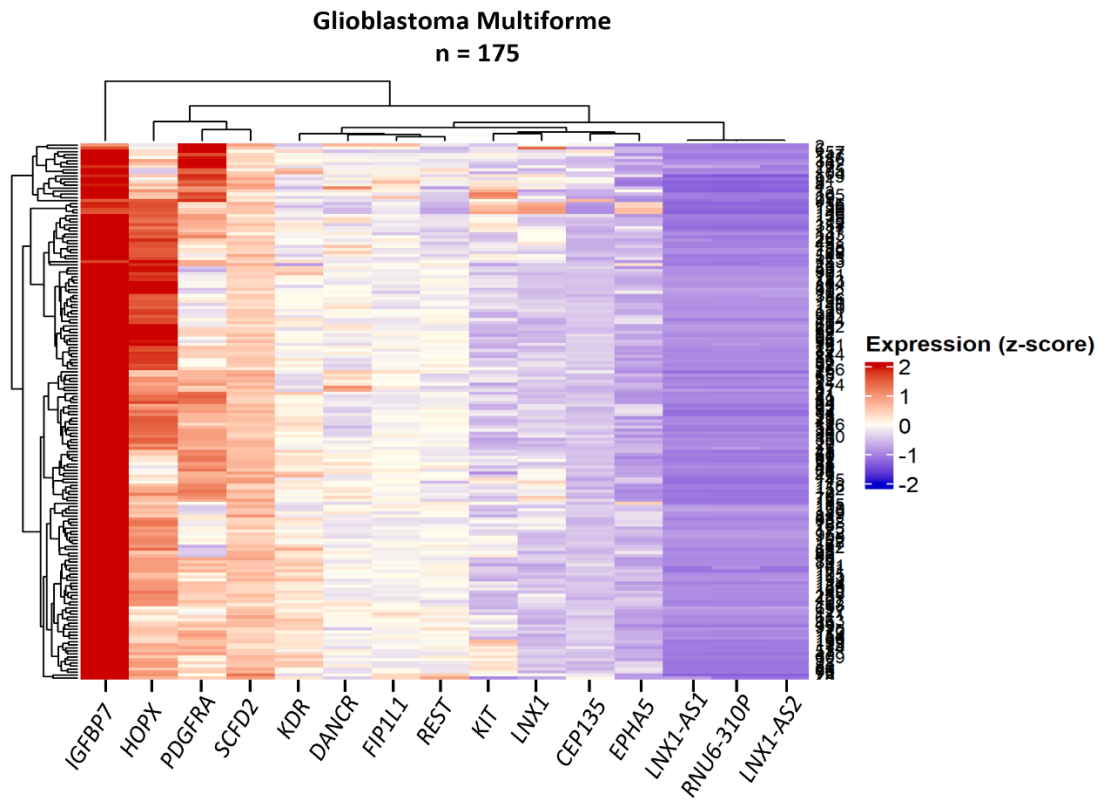


Figure 6.1B: Analysis of the amplified chromosomal locus 4q12-13 and the relative expression of genes in this locus in GBM and LGG. Heatmaps of gene expression across a chromosomal locus were generated using the UCSC Xena Browser. Each row represents an individual sample, and each column represents a gene within the locus. Gene expression values are displayed as Z-scores, indicating the number of standard deviations from the mean expression of each gene across all samples (red = higher-than-average expression; blue = lower-than-average expression). Hierarchical clustering of both samples and genes is shown by dendrograms, revealing groups of co-expressed genes and distinct sample clusters (Goldman *et al.*, 2020).

Chromosome 13q12 has been identified as a “driver cluster” of genes in CRC (Camps *et al.*, 2013), and the amplicon is especially associated with late-stage clinical features of CRC (Zhou *et al.*, 2020) (Figure 6.1A). We employed the same method as Figure 6.1B to investigate the expression of genes within this locus. Angiogenesis is one of the hallmarks of cancer (Hanahan and Weinberg, 2000) which has an important role in the progression of CRC. The vascular endothelial growth factor receptor (VEGFR1) is encoded by the *FLT1* gene and is expressed in 50% of CRCs with little to no expression in normal colonic mucosa or adenomas (Hu *et al.*, 2007). RNA polymerase I and III subunit, *POLR1D* is frequently overexpressed in CRC and upon its silencing in CRC cells, proliferation and migration was reduced (Wang *et al.*, 2019). Another study found that *POLR1D* is associated with upregulation of VEGF α and *EREG* in CRC (Zhou *et al.*, 2020).

In the same locus the expression of the palmitoyltransferase enzyme *ZDHHC20*, appears to be highly upregulated in colorectal adenocarcinoma and rectal adenocarcinoma (Figure 6.2B). The PTM, palmitoylation is key in modulating pathways such as AKT-mTOR and WNT/ β -catenin especially in liver cancer, CRC and gliomas (Yang *et al.*, 2026). Transcription factor IIIA (*GTF3A*) and mitochondrial translation initiation factor 3 (*MTIF3*)

genes are highly expressed in CRCs with *GTF3A* being deemed as a critical driver of CRC tumorigenesis. *GTF3A* increases cystatin (*CSTA*) expression by binding to the *CSTA* promoter and regulating epithelial to mesenchymal transition (Wang *et al.*, 2022). The elevated expression of the *LNX2* gene is on par with that of notable oncogenic drivers such as *POLR1D*, suggesting that elevated expression of *LNX2* protein as a result of gene amplification could have a role in the development of CRC (Figure 6.2B).

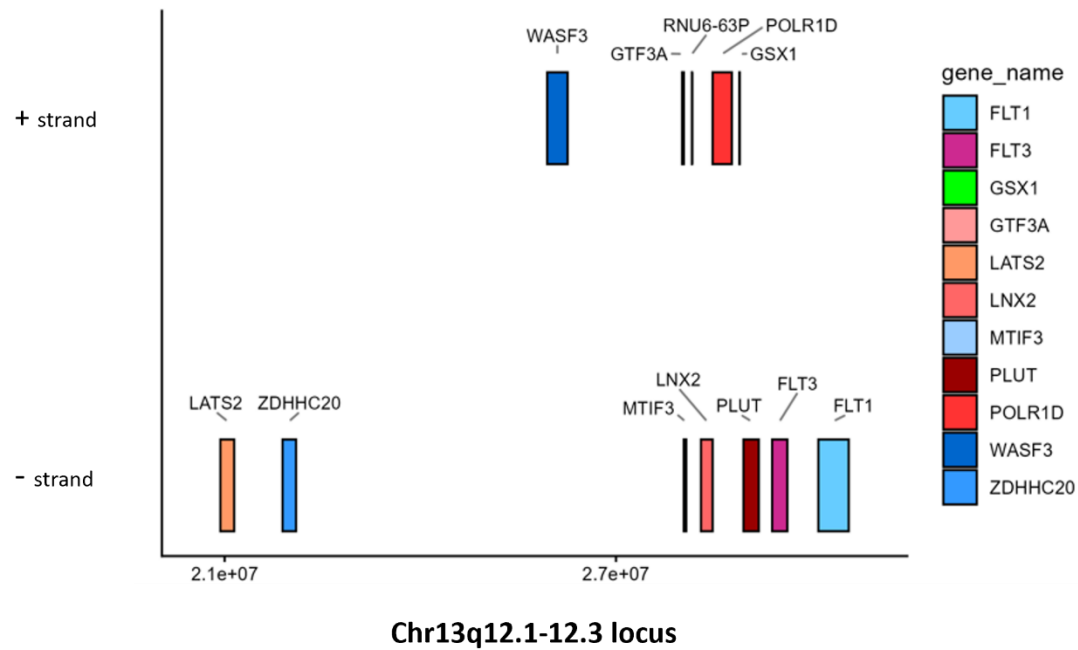


Figure 6.2A: Schematic of the chromosome 13q12 amplicon commonly present in CRC. Genes mapped to the Chr 13q12.1-12.3 locus, and their relative position to each other, which was created using the locuszoomr Rstudio package (Lewis and Wang, 2024).

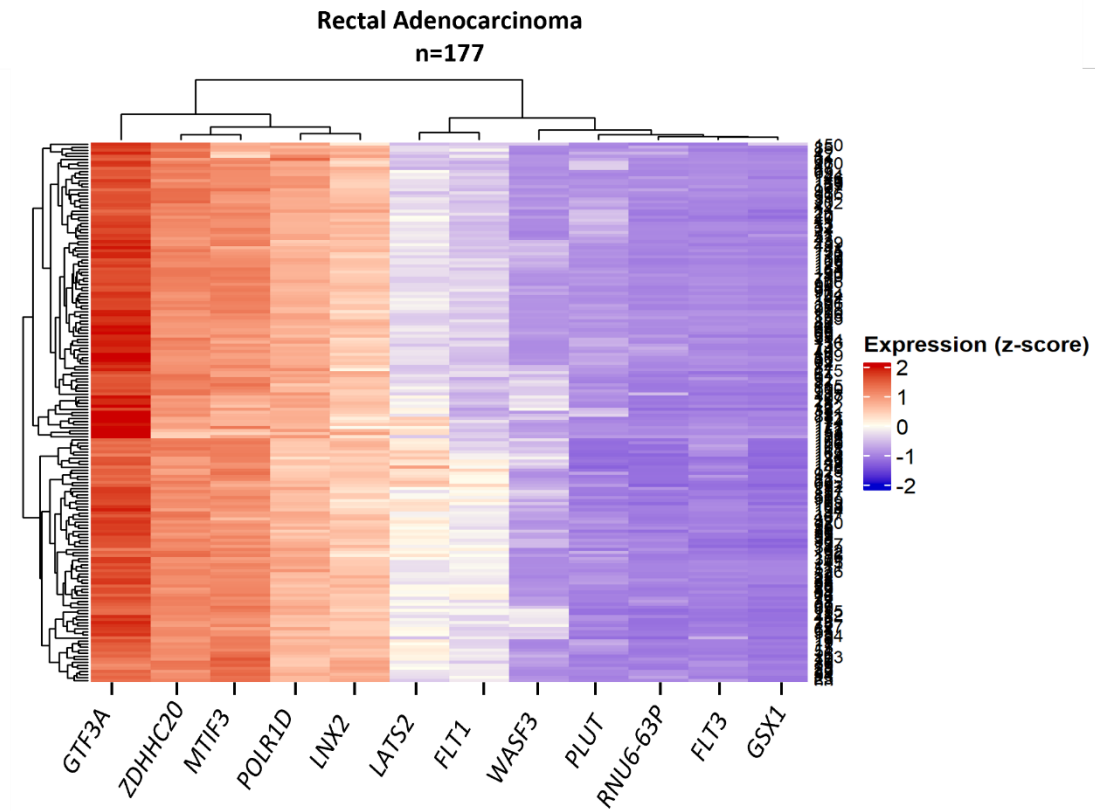
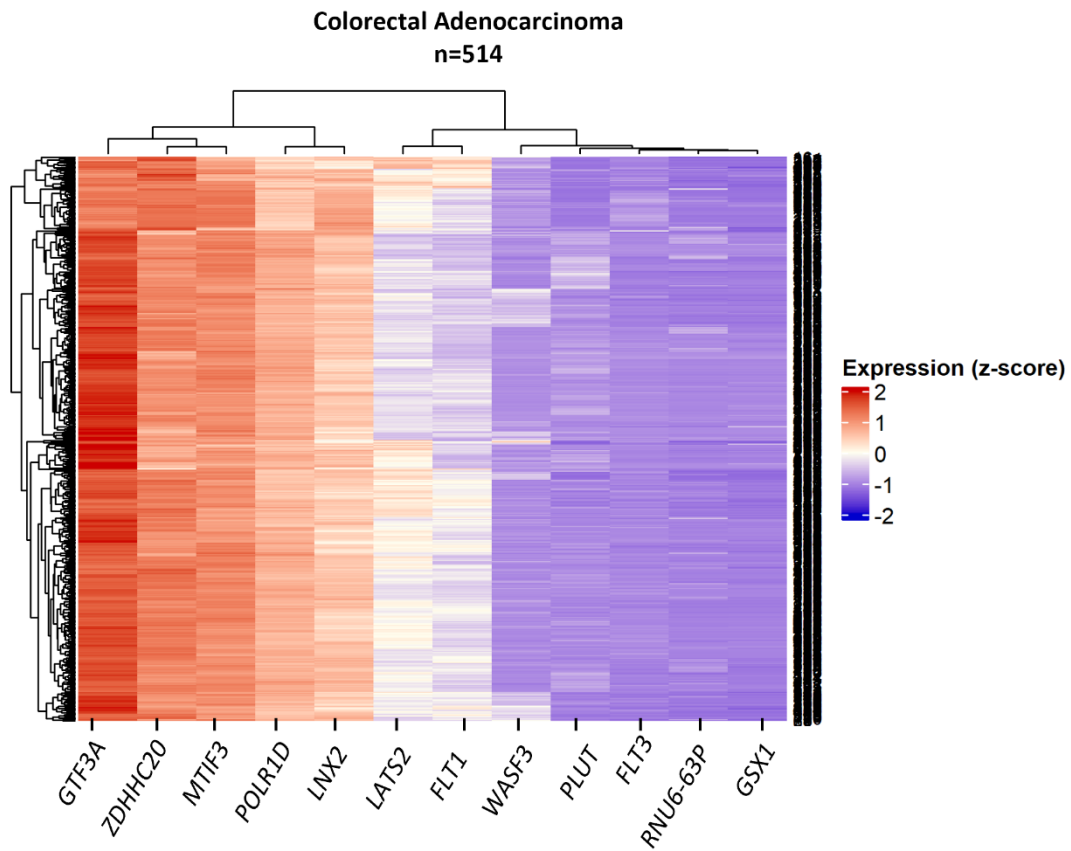


Figure 6.2B: Analysis of the chromosome 13q12 amplicon and its relevance to colorectal and rectal adenocarcinomas regarding gene expression. Heatmaps of gene expression across amplicon Chr13q12 were generated using the UCSC Xena Browser. Each row represents an individual sample, and each column represents a gene within the locus. Gene expression values are displayed as Z-scores, indicating the number of standard deviations from the mean expression of each gene across all samples (red = higher-than-average expression; blue = lower-than-average expression). Hierarchical clustering of both samples and genes is shown by dendrograms, revealing groups of co-expressed genes and distinct sample clusters (Goldman *et al.*, 2020).

6.2.2 In silico and in vitro analysis of LNX1/2 expression

From analysing the genomic loci of *LNX1* and *LNX2* genes in Figure 6.1A & B and 6.2A & B, expression seemed to correlate with regional amplification, at least for *LNX2*. One drawback of TCGA datasets from the Genomic Data Commons is that although isoform-level estimates are available from RNA-seq pipelines, specific isoform level results are not part of the standard TCGA output. Partly due to transcript-level quantification from short-read RNA-seq (Trincado *et al.*, 2016). Differentiating between *LNX1*^{p80} and *LNX1*^{p70} is not possible by investigating TCGA data alone, so GTEx data was also implemented by using the GEPIA3 server to analyse isoform specific expression of LNX mRNAs in various cancers (Kang *et al.*, 2025). All analyses were based on the human reference genome hg38. GEPIA3 uses RNA-seq data from TCGA and GTEx, with expression values calculated as transcripts per million (TPM) and expected counts using RSEM (UCSC Toil pipeline). Default settings were applied to all searches, Log₂FC cutoff set to 1 and *q*-value was set to 0.05. As graphs compiled by GEPIA3 are large with lots of cancer types compared on each one statistical information is difficult to portray in a tidy manner. Table S2 details the t-test *p*-values of each comparison and full list of abbreviations

From Figure 6.3 (top), *LNK1^{p70}* is only expressed in GBM and LGG tumours which is expected as *LNK1^{p70}* is a CNS specific isoform. Interestingly, there is higher expression in the normal tissue compared to the tumorigenic samples. This could be due to a mix of cell types taken during the normal adjacent tissue biopsy or it could be supporting the expression pattern seen from Figure 6.1. We cannot determine from these bioinformatical graphs alone. On the contrary, *LNK1^{p80}* has little to no expression in GBM or LGG tumours or normal tissue. Instead *LNK1^{p80}* expression seems to be expressed at a higher level in lung squamous cell carcinoma and thyroid carcinoma. Normal tissue isolated beside esophageal and kidney carcinomas seems to have much higher expression of *LNK1^{p80}*. As expected *LNK2* expression is vastly upregulated in colon, rectal and stomach carcinomas. A curious observation seen from GEPIA3 was the *LNK2* expression profile is also increased in breast invasive and esophageal carcinomas. *LNK3* expression was also investigated, and a general level of downregulation is seen across the cancers analysed, but especially in uterine, ovarian and GI tract tumours (Figure S4).

Something to note when analysing the RNA-seq data is the TPM scales used for each individual gene in Figure 6.3 and S3. The *LNK1^{p70/p80}* and *LNK2* expression scale only reaches 25-60 TPM while *LNK3* has much higher overall expression values with 210 TPM. We can only compare the gene expression levels relative to itself across these tumour types.

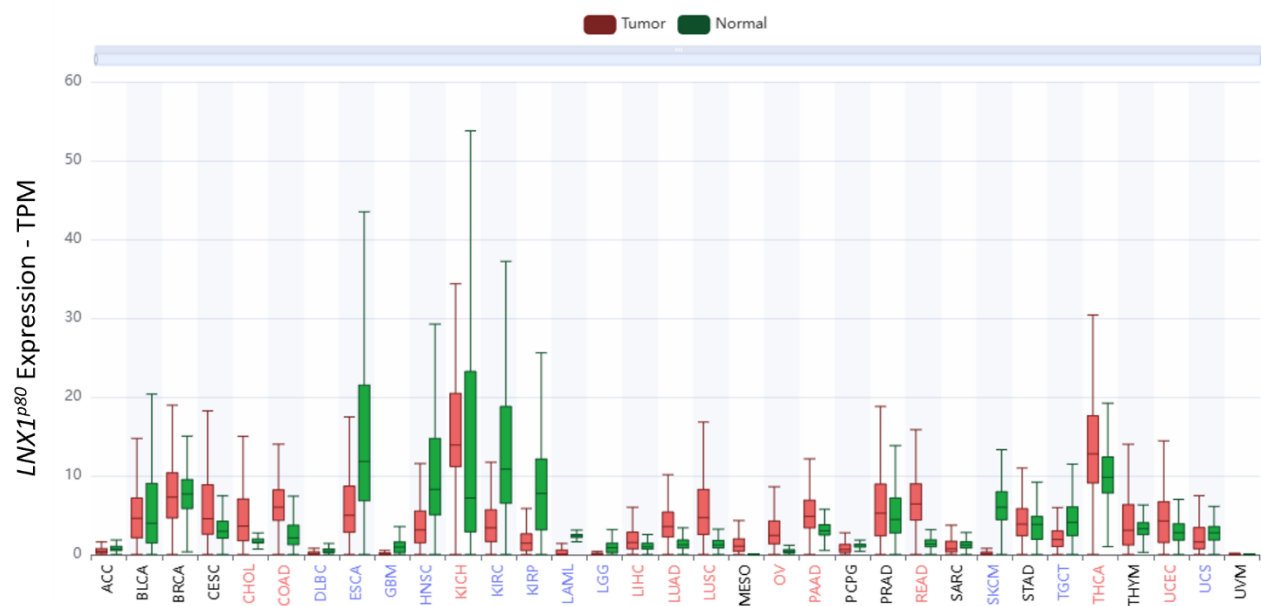
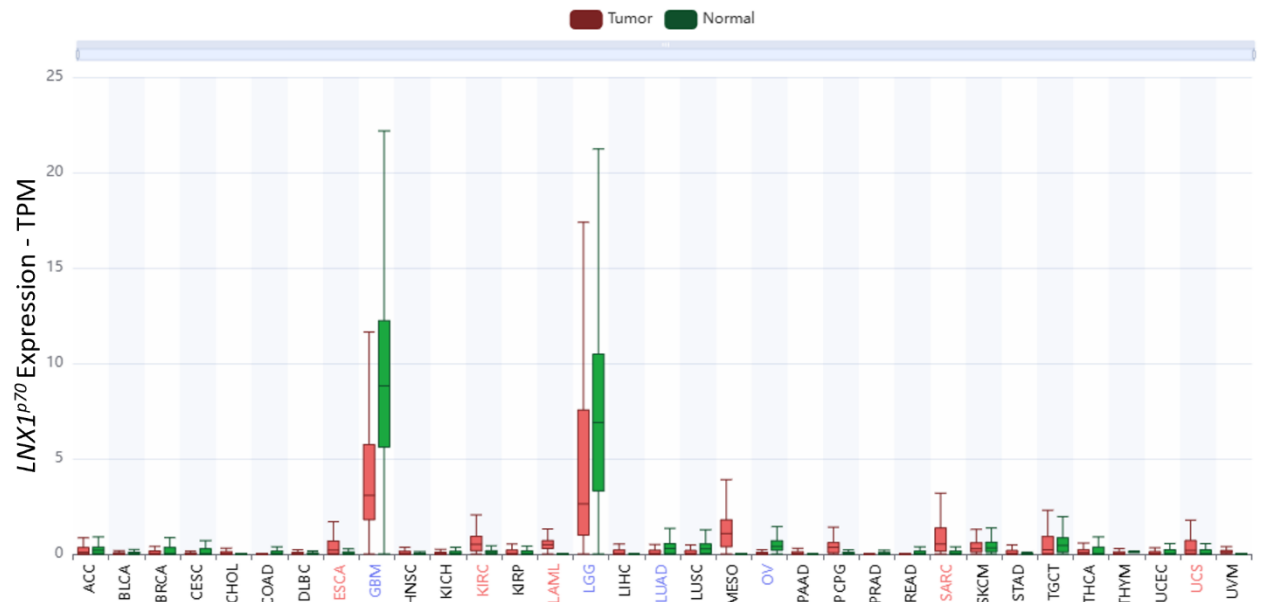


Fig 6.3: *LNK1^{p70/p80}* and *LNK2* expression in normal versus tumorigenic tissues from GTEx and TCGA datasets and compiled by GEPIA3. Normal tissue samples analysed were biopsied from adjacent tumorigenic samples RNA expression is given in transcripts per million (TPM) as $\text{Log}_2(\text{TPM}+1)$ scale. Cancers highlighted in blue have reduced expression compared to normal tissues of the indicated gene and red is elevated expression. Full list of statistical information from these analyses is available in Table S2 (Kang *et al.*, 2025).

Many of the tumorigenic studies concerning the relevance of LNK proteins was carried out in cancer cell lines (Chen *et al.*, 2005)(Camps *et al.*, 2013)(Ma *et al.*, 2017) (Park *et al.*, 2019)(Jang *et al.*, 2021) *LNK1* and *LNK2* expression at both RNA and protein levels have not been determined in most of these cell lines to the best of our knowledge. We were kindly donated a panel of GBM and CRC cell lines from Dr. Ken Nally and Dr. Cora O'Neill (University College Cork) and they were cultured as described in Methods section 2.2.1. Firstly, mRNA levels were quantified by qPCR analysis using primers from Table 2.3 Section 2.1.3. Extremely low endogenous expression was seen for both *LNK1* and *LNK2* across the cell lines using *ACTB* as a housekeeping gene for normalization, as indicated in Figure 6.4. As a comparison, HEK 293T cells were transfected with *LNK1^{p70/p80}* and *LNK2* expressing plasmids to determine the approximate levels of *LNK* mRNA expression that would be needed to detect protein *via* Western blot (Data not shown. Cells from the same transfection were analysed by qPCR in the same way as indicated cell lines. Transfected *LNK1* (both isoforms) gave a relative expression level of 31.26 and *LNK2* was 13.58 by the $2^{-\Delta\text{CT}}$ method. This suggests, in the case of *LNK2*, that endogenous *mRNA* expression in the HT29 cell line with a $2^{-\Delta\text{CT}}$ value of ~ 0.005 , for example, is ~ 2600 times lower than *LNK* mRNA in the transfected HEK which have detectable LNK protein levels.

Such a determination cannot be quantified for LNX1 as primers used amplify both isoforms of LNX1, but *mRNA* levels appear to be a few thousandfold lower in GBM and CRC cell lines compared to the protein detectable level. This is a very diverse panel of cell lines and the expression of the house keeping gene used, *ACTB*, is not consistent across the entire panel, making accurate comparisons between cell lines difficult. Other housekeeping gene like *TBP*, *GAPDH* and *GUSB* were trialled across the cell lines, but the CT values were more variable than that of *ACTB*. CT values for β -actin were between 18-24, so slightly less consistent than what would be desired. Nevertheless, the endogenous mRNA levels for both *LNX1* and *LNX2* appear to be dramatically lower than in transfected HEK cells in which LNX proteins were just above the limit of detection by Western blotting.

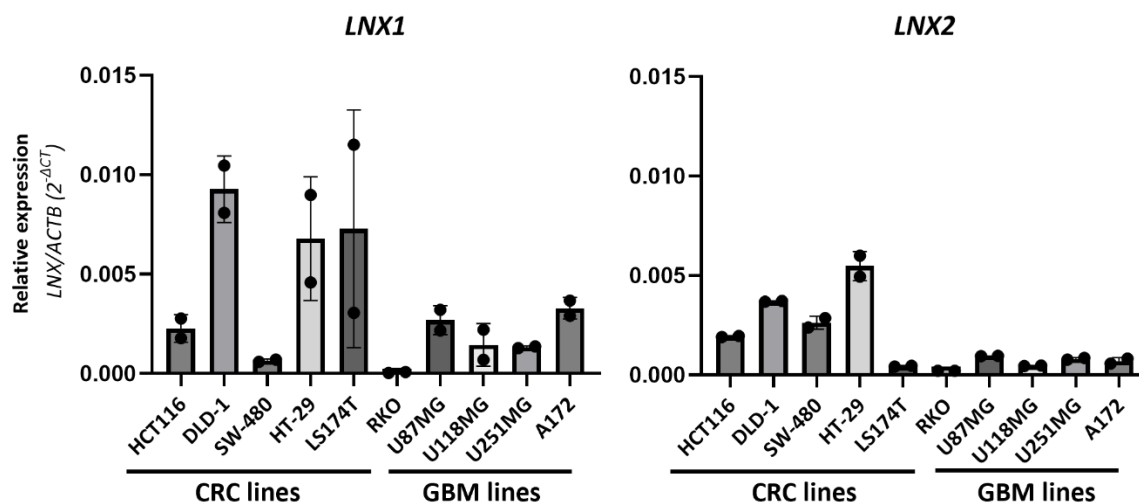


Figure 6.4: *LNX1* and *LNX2* mRNA expression levels in the indicated cell lines. RNA was extracted and reverse transcribed to cDNA as per Methods Section 2.1.3. *LNX1* and *LNX2* specific primers were used to quantify *mRNA* levels and compared to *ACTB* for normalisation. Data points represent mean values of technical triplicates from two independent experiments.

Secondly, we attempted to detect LNX1/2 protein levels across the same panel of cell lines with a few extra cell lines that were not quantified for RNA levels (LoVo, Colo-205, HCT-15, CaCo2, HuH7 and Calu3 – only harvested cell pellets available). As mentioned before different cell lines express housekeeping genes at different levels so a one size fits all loading control could not be used. Instead, extra caution was taken to ensure equal protein quantities were loaded (30 µg). Transfected LNX1 and LNX2 proteins were loaded alongside the panel of cell lines to validate the efficiency of the antibodies. Five antibodies were first examined (Table 2.6, Section 2.6) to determine the specificity to human LNX1 and/or LNX2. Only one commercial antibody (α-LNX2 (Wuxi/Abgent) and two antibodies (α-LNX1/2 PDZ3-4 and α-LNX1/2 NPAY-RING) previously manufactured by

Dr. Paul Young/ Dr. Jane McGlade (Young *et al.*, 2005) detected transfected human LNX protein.

In the CRC cell lines (Figure 6.5, left), no band corresponding to LNX1^{p80} (80kDa) or LNX2 (76kDa) protein was seen. A non-specific band was detected at ~130kDa by tested antibodies. A similar result was seen for the GBM lines (Figure 6.5, right) in which no LNX2 or LNX1^{p70} (70kDa) band was detected. A signal corresponding to a protein of around 75-80kDa (suggesting LNX1^{p70} or LNX2 protein) was observed in the top right panel of Figure 6.5, however, this band was not detected by other antibodies and therefore it cannot be confirmed as LNX protein.

Further efforts were made in the attempt to detect protein in these cell lines, but no signal was detected. MG132 was used to inhibit the proteasome in HCT-116, HT-29 and Colo-205 cells to investigate if LNX proteins are recycled before expression can be examined (Figure S5A). Immunoprecipitation using an antibody proven to pull-down LNX1^{p70/p80} and LNX2 was carried out in U-251MG, HCT-116, HT-29 and SW-480 cells, and again, no LNX proteins were purified (Figure S5B).

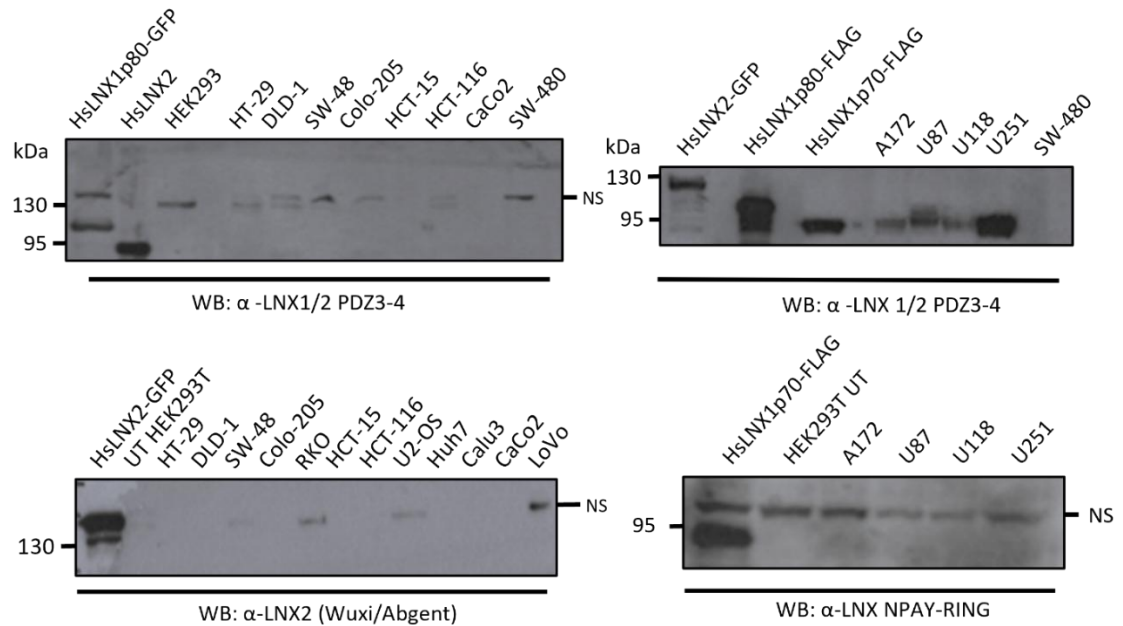


Figure 6.5: LNX1 and LNX2 proteins are undetectable by Western blot across a panel of GBM and CRC cell lines. Indicated cells were lysed and a BCA assay was performed as per section 2.3.4. 30 µg of each lysate was loaded along with transfected LNX1/2 expressing HEK 293T cells for antibody comparison. GFP-LNX2 and GFP-LNX1^{p80} run at 130kDa, while endogenous/untagged LNX2 commonly appears at 75-80kDa. LNX1^{p80} is slightly larger and is endogenous protein is seen to run at ~95kDa, with a flag tag, only slightly higher. LNX1^{p70} without a tag runs around 75kDa. Commercial antibody α-LNX2 was purchased from Wuxi Abgent (#AP13071b) the other antibodies were specifically designed to target indicated epitopes of LNX proteins (Young *et al.*, 2005). All blots were developed using manual ECL technique at various exposure lengths using secondary antibodies tagged with HRP (Jackson ImmunoResearch Table 2.6).

6.2.3 LNX1, LNX2 and LNX3 promoter analysis

LNX1/2 mRNA and protein levels are very low, and we questioned whether this is due to low expression of the mRNA due to weak promoter activity. Alternatively, the low expression and detection levels could be high turnover or instability of the mRNA and or protein. Previous work revealed the presence of conserved uORFs in the untranslated region of LNX1^{p70}, which decreased translation by 2-3 fold (Lenihan *et al.*, 2014). The

same mechanisms do not control LNX1^{p80} or LNX2 translation. However, one study suggests LNX1 transcription and translation is enormously increased upon addition of tamoxifen into HT-29 cells (Ma *et al.*, 2017). The regulatory mechanisms controlling LNX transcription and translation is far from complete and here, we aim to potentially understand how LNX transcription might be regulated by studying promoter region activity in a variety of environments under different conditions.

To identify promoter regions of LNX1^{p70/p80}, LNX2 and LNX3 the analysis focused on the region immediately upstream of the transcription start sites. The UCSC Human Genome Browser (GRCh38/hg38) together with the ENCODE Candidate Cis-Regulatory Elements (cCREs) resource provides a comprehensive map of regulatory regions in the human genome, indicating potential regulatory activity and promoter signatures (Casper *et al.*, 2026) (Moore *et al.*, 2024). The eukaryotic promoter database (EPDnew) human database (version 006) enables identification of experimentally validated promoters (Dreos *et al.*, 2013). The ORegAnno database contains a curated set of experimentally validated regulatory elements with supporting evidence of transcription factor binding sites (Lesurf *et al.*, 2015). With this information putative promoter regions were identified as follows: LNX1^{p70} – Chr4:53,459,301-53,558,217, LNX1^{p80} – Chr4:53,459,301-53,591,486, LNX2 – Chr13:27,545,913-27,620,529 and LNX3 – Chr3:73,382,431-73,624,941 and cloned into firefly luciferase (FLuc) reporter vectors. Initially, we explored how LNX promoters would behave in different tumourigenic environments. Co-transfection of Fluc (LNX promoters) and renilla luciferase (RLuc as a control for transfection efficiency) plasmids was performed into cell lines, activity expressed as

relative luciferase units (RLU), as indicated in Figure 6.6. Overall, LNX promoter activity was very low (LNX1^{p70}: <0.1 RLU, LNX1^{p80}: <1 RLU, LNX2: <0.8 RLU, LNX3: <1 RLU) in comparison to the SV40 (2.5 RLU) and elongation factor 1 α (EF-1 α , 170 RLU) (Figure 6.6 A). Both LNX2 and LNX3 promoter activity was stronger than LNX1^{p70}. The CNS isoform LNX1^{p70} promoter is more active in neuronal and glial tumour backgrounds than the non-CNS isoform, LNX1^{p80}, while almost no luciferase activity is observed in CRC, breast cancer or HEK 293T cells. Despite CRC cancer having increased *LNX2 mRNA* expression (Figure 6.2B & 6.3), LNX2 promoter activity appears to be quite low across the examined CRC cell lines, with highest activity in SW480 cells (Figure 6.6E). The LNX1^{p80} promoter activity is again low across the panel of cell lines tested, with highest activity in the breast cancer derived cell lines, MCF-7 (1 RLU). The low activity of LNX1 and LNX2 promoters may explain the low mRNA and lack of detectable LNX protein level by Western blot seen in Figure 6.5 and S4.

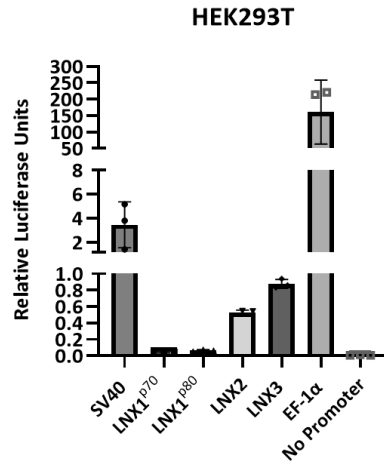
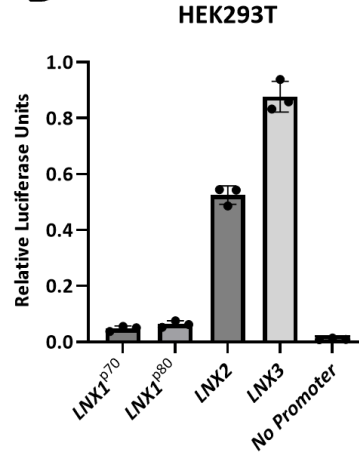
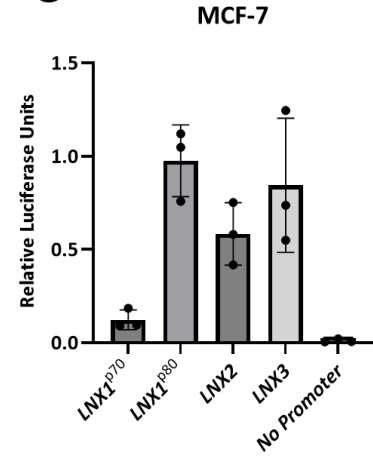
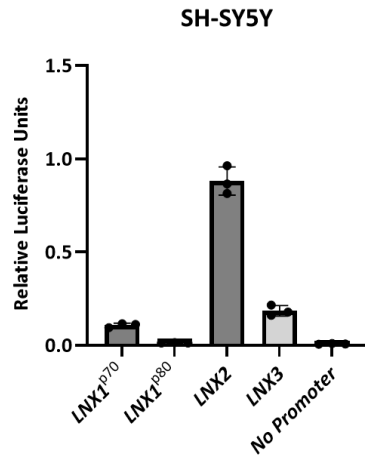
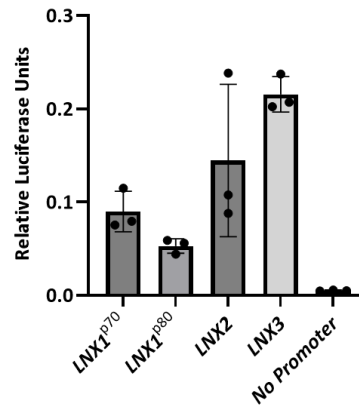
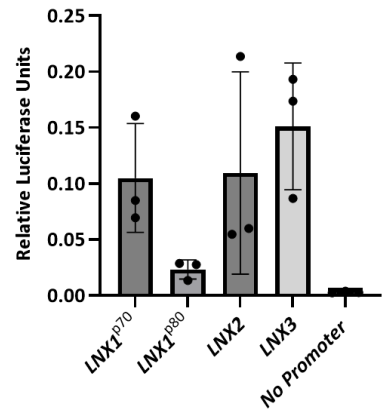
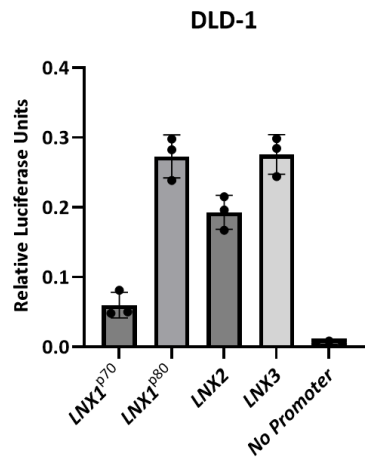
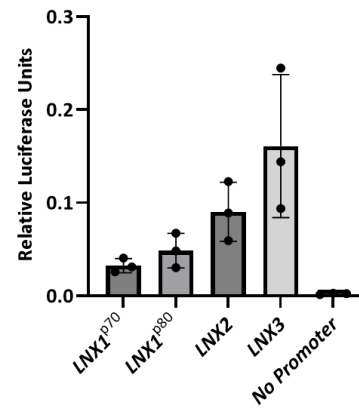
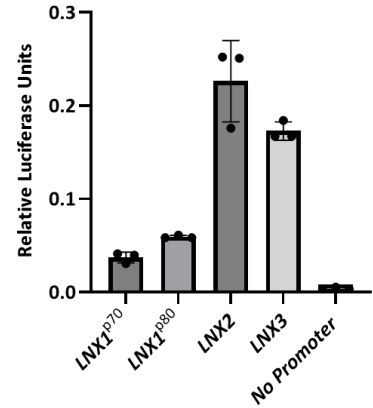
A**B****C****D****U-251 MG****U-87 MG****E****HCT 116****SW480**

Figure 6.6: LNX promoter activity in a panel of cancer cell lines. Cells were co-transfected with equal amounts of indicated the LNX promoter sequence or no promoter upstream of a Firefly luciferase (FLuc) reporter gene along with a Renilla luciferase (RLuc) to control for transfection and to normalise reporter activity. Activity was analysed after 24 hours as per Methods Section 2.4 **A-B**. HEK 293T as a reference cell line from a non-tumorigenic background **A** includes the strong SV40 and EF1 α promoters, while **B** presents the same data excluding these strong promoters to facilitate comparison with the other panels. **C**. MCF-7 cells derived from a breast cancer origin. **D**. Neuronal (SH-SY5Y) and glial carcinoma (U-251MG, U-87MG) derived cell lines **E**. CRC cell lines. Each biological replicate consisted of five technical replicates, and the means were plotted. Each graph is representative of three independent experiments, error bars represent SD.

Ma *et al.*, 2017, reported that treatment of the HT-29 CRC-derived cell line with tamoxifen (concentrations or details used were not disclosed) drastically increased LNX1 promoter activity, as well as elevating LNX1 protein levels (Ma *et al.*, 2017). We sought to replicate and build on this observation, however, in our hands, HT-29 cells were extremely difficult to transfect, even though we did attempt multiple methods including, electroporation. Other CRC cell lines were successfully transfected (HCT-116, SW-480) but tamoxifen treatment at as low as 0.1 μ M decreased cell viability before 24 hours. As MCF-7 cells could withstand the tamoxifen treatment for 24 hours and had a good level of transfection, we could determine that LNX1^{p80}, LNX2 and LNX3 promoters were slightly more active upon introduction of tamoxifen but this result was not significant (Figure 6.7).

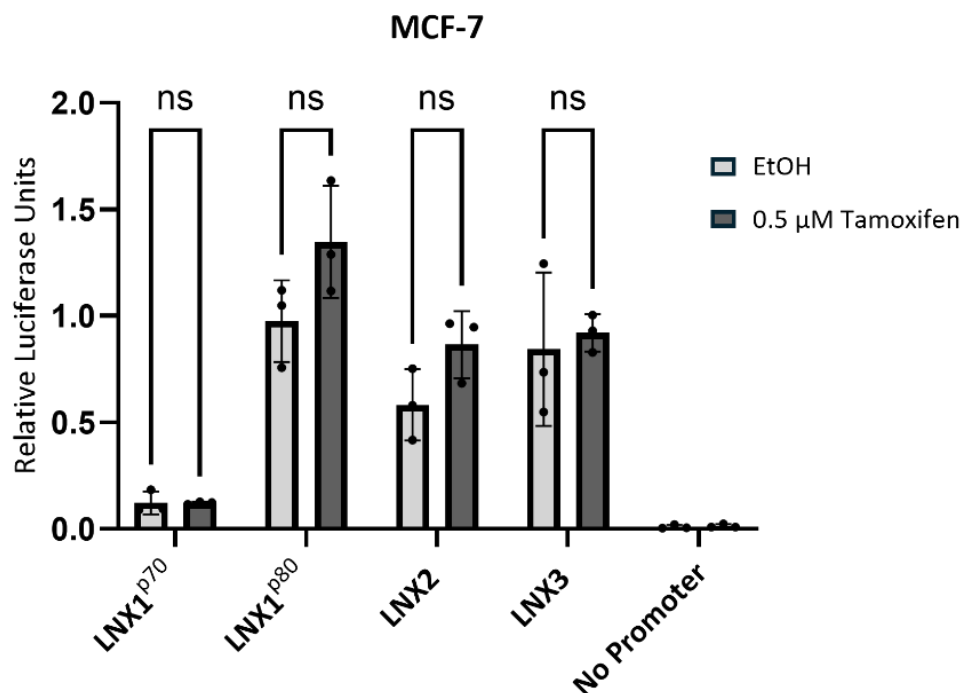


Figure 6.7: Tamoxifen treatment does not significantly increase LNX1^{p80} promoter activity in MCF-7 cells. Cells were co-transfected with the indicated FLuc and LNX promoter plasmids along with Rluc as a transfection control in DMEM containing 10% FBS and no phenol red. 24 hours post transfection media was changed to either contain 0.5 μ M of 4-hydroxy tamoxifen (Sigma) or an equal volume of vehicle control (100% EtOH) and allowed to continue for 16 hours before being analysed for promoter activity by luciferase assay. Each of the three biological replicates consisted of five technical replicates, and the means were plotted. TWO-Way ANOVA followed by Tukey's Ad Hoc multiple comparisons test was employed. Error bars represent SD.

Protein kinase C (PKC) is a very important serine/threonine kinase involved in a plethora of cellular functions like, regulating transcription, modulating membrane structure, regulating cell growth (Newton, 2018) and mediating pathways involving memory and leaning in neurons (termed memory kinases) (Sun and Alkon, 2014). PKC isoforms carry out their functions by phosphorylating protein substrates depending on cell type and

cellular pressure (Newton, 2018). PKC signalling can also modulate expression and activity of the E3 ligase MDM2 leading to increased or decreased stability of p53 and in turn influencing tumour progression (Min *et al.*, 2019). Phorbol 12-myristate 13-acetate (PMA) is a diacylglycerol analogue that directly activates PKC.

When LNX1 and 3 promoters were treated with 10 nM of PMA there was a prominent increase in activity. Difficulty in cloning resulted in the absence of the LNX2 promoter from this assay, it would be interesting to repeat the PMA treatment to observe the effects on LNX2 promoter. In the case of the LNX3 promoter, activity continued to increase with a higher (50 nM) concentration of PMA (Figure 6.8). PMA activation of PKC in cells can lead to the phosphorylation of many different effectors and trying to imply a specific mechanism for LNX promoters is not possible without further investigation.

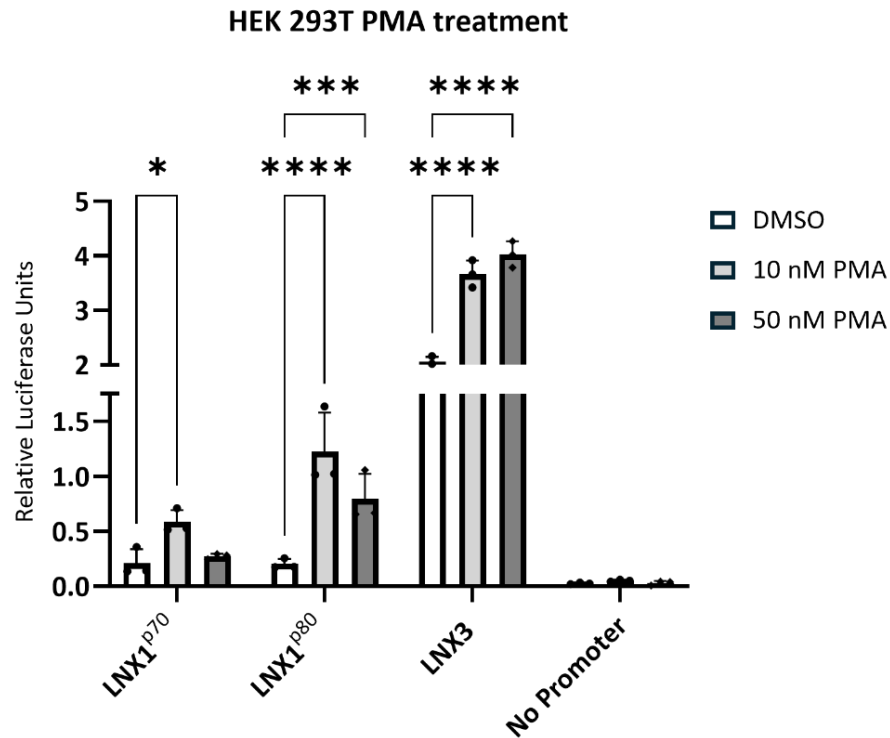


Figure 6.8: Activation of the PKC signalling pathway increases transcription of *LNX1* and *LNX3* genes. HEK 293T were co-transfected with the indicated FLuc and LNX promoter plasmids along with Rluc as a transfection control in DMEM containing 10% FBS and no phenol red. 24 hours post transfection media was changed to either contain DMSO or 10 nM or 50 nM of PMA (Sigma) and allowed to continue for 24 hours before being analysed for promoter activity by luciferase assay. Each of the three biological replicates consisted of five technical replicates, and the means were plotted. TWO-Way ANOVA followed by Tukey's Ad Hoc multiple comparisons test was employed, *p<0.05, ***p<0.001, ****p<0.0001. Error bars represent SD.

The ubiquitination system is generally more active and dynamically regulated during cellular stress, hypoxia, DNA damage and starvation, with E3 ubiquitin ligases playing central, context-specific roles in reshaping the proteome (Sheng *et al.*, 2023). During nutrient starvation, several E3 ligases (TRAF6, RNF152, SCF complex) can promote selective autophagy by ubiquitinating proteins and organelles for degradation mainly through the inhibition of mTORC1 (Dikic, 2017). Similarly, upon chemical treatment using high concentrations of potassium chloride or glutamate in mouse primary neuronal cell cultures, *LNK2* mRNA expression increases (Cioccarelli. L, Manuscript in preparation, 2026).

Following on from these facts, we attempted to induce a starvation response in HEK 293T cells by nutrient deprivation (FBS) and determine the LNK promoter activity. In general, phosphorylation activity decreases in cells under starvation conditions (Krech *et al.*, 2010). We see a reduction of phosphorylated AKT compared to total AKT when serum is reduced from 10% to 0.1% (Figure 6.9B). Following on from this *LNK2* promoter is significantly more active suggesting a possible role of *LNK2* activity under nutrient starvation at least in HEK 293T cells. The same change in promoter activity is not reflected for *LNK1* or *LNK3* (Figure 6.9A).

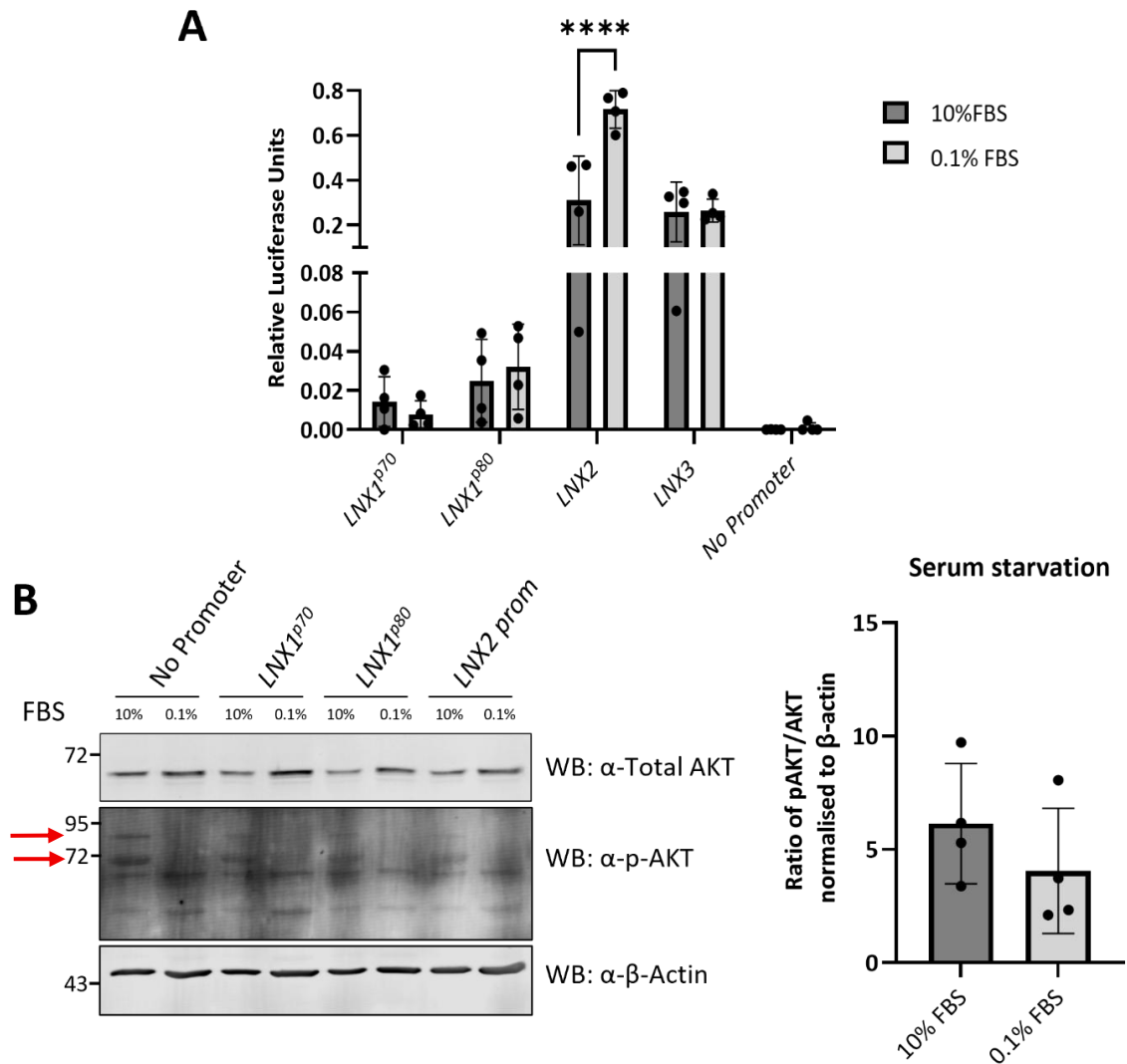


Figure 6.9: LNX2 promoter activity is increased upon serum starvation. HEK 293T cells were co-transfected with the indicated FLuc and LNX promoter plasmids along with Rluc as a transfection control in DMEM containing 10% FBS. 24 hours post transfection media was changed to either 10% FBS or 0.1% containing DMEM and allowed to continue for 16 hours before being analysed for **A**. LNX promoter activity assessed by luciferase assay or by **B**. Western blot to verify effective serum starvation. HEK 293T cells showed a decrease in phosphorylated AKT (p-AKT) under starvation conditions (0.1% FBS) compared to normal conditions. Intensities of p-AKT (red arrows) to total AKT signals were normalised to β-actin. Luciferase assays were analysed by TWO-way ANOVA followed by Tukey's Ad Hoc multiple comparisons **** $p < 0.0001$, graph is representative of four independent experiments, error bars represent SD of the mean. Western blot intensity quantification carried out using ImageJ and analysed using Students t-test n.s $p > 0.05$, Blot is representative of four independent experiments, error bars represent SD of the mean.

6.3 Discussion

6.3.1 Analysis of LNX1 and 2 expression and regulation

The overall aim of this chapter was to provide some clarity regarding LNX expression particularly from a cancer perspective. From the analysis of gene expression within the *LNX1* and *LNX2* loci it was revealed that genetic amplification of the loci does not correspond to increased expression in the case of *LNX1*, at least in glial cancers. We confirmed this by observing markedly low *LNX1* mRNA expression in GBM cell lines. It is more likely that *LNX1* is a passenger gene within this amplified region rather than a contributor to tumour progression, and that the elevated expression of established oncogenes *IGFBP7*, *HOPX* and *PDGFRα* are the real drivers of glial tumorigenesis (Holtkamp *et al.*, 2007)(Iqbal *et al.*, 2010) (Harwood *et al.*, 2025). In addition, *IGFBP7* and *PDGFRα* have been identified as prognostic biomarkers correlated with immune infiltrations of gliomas (Wang *et al.*, 2023). Interestingly, the reduced expression of the antisense transcripts (Figure 6.1) associated with *LNX1* highlights a potentially overlooked regulatory layer, as antisense RNAs are increasingly recognised for their roles in gene regulation and epigenetic control; however, their specific contribution to glial cancers remains to be elucidated (Rehman *et al.*, 2023).

Extending this analysis to chromosome 13q12 in colorectal cancer (CRC), a region previously characterised as a driver cluster, reveals a similarly complex landscape of co-amplified genes (*POLR1D*, *MTIF3* and *ZDHHC20*) and with diverse oncogenic functions (Tian *et al.*, 2020) (Wang *et al.*, 2022) (Yang *et al.*, 2026). Notably, the expression of *LNX2* is comparable to established oncogenic drivers within this locus, suggesting a potential

functional role in CRC that contrasts with the apparent lack of involvement of *LNK1* in glioma. Interestingly however, *LNK1* mRNA expression appears to be higher than *LNK2* in DLD-1, HT-29 and LS174T CRC cell lines. Nevertheless, in Figure 6.5 and S4, there is no detectable LNK1 or 2 proteins, reinforcing the minuscule levels of *LNK* mRNA expression in these cell lines, even though, Camps *et al.* detected endogenous LNK2 protein without immunoprecipitation, using a commercially available α -LNK2 antibody from Abcam (Camps *et al.*, 2013). This led us to question the specificity of this antibody for LNK2. Furthermore in the same publication, the authors observed that LNK2 depletion reduced the levels of both NUMB and NOTCH1 proteins without affecting the levels of *NUMB* transcripts in SW480 cells, which was counterintuitive, considering that Lnk2b in zebrafish and LNK2 in mammals is commonly known to degrade NUMB protein (Won, Ro and Dawid, 2015)(Nayak *et al.*, 2015). Similarly, another study saw endogenous LNK1 protein expression in HT-29, HCT-116, CaCo-2 and LS174T cells and confirmed antibody specificity using transient LNK1 knockdowns in HT-29 cells (Ma *et al.*, 2017), a result that we could not reproduce.

The low levels of *LNK1^{p70/p80}* and *LNK2* transcription could be due to the activity of their respective promoters. We already know that uORFs in the UTR of *LNK1^{p70}* mRNA explain the tightly controlled translation of LNK1p70 protein (Lenihan *et al.*, 2014), but knowledge of the regulation mechanisms regarding *LNK1^{p80}* and *LNK2* remain elusive. Figure 6.6, we observe that LNK1/2 promoters are only active at a very low level when compared to SV40 and EF-1 α promoter signatures in HEK-293T cells (Figure 6.6A) Even when LNK1/2 promoters are subjected to different tumorigenic environments, activity

remains low. The effect of external stimulation was also investigated. In Ma *et al.*, *LNK1* transcription was apparently increased enormously, upon tamoxifen treatment in HT-29, leading to suppression of cancer stemness (Ma *et al.*, 2017). We could not replicate the level of increase that was previously observed, at least in MCF-7 cells (Figure 6.7). It was evident however that activation of PKC signalling by PMA did increase *LNK1* promoter activity in both isoforms. Comparably, when HEK 293T cells are starved of serum for 16 hours, the transcription of *LNK2* is significantly increased. Together these facts suggest a role of *LNK* E3 ligases in cellular stress and signalling adaptation, particularly in response to nutrient deprivation and PKC-mediated pathways.

6.3.2 Brief investigation of *LNK3* expression

While the third family member *LNK3*/*PDZRN3*, is largely outside the scope of this project it warranted an investigation of its expression pattern and regulation. There are a few studies that indicate a potential tumorigenic role for *LNK3* (Hong, Won and Ro, 2020) including reduced expression in endometrial cancer (QiuHong Li *et al.*, 2022), which we also observe from analysing GTEx and TCGA databases (Figure S4 and Table S2). Promoter analyses were not carried out in an endometrial cancer cell line, but future work could be intriguing. Luo *et al.* describe the ubiquitination of *TBX20* by *LNK3* in CRC, preventing NHEJ in DNA repair (Luo *et al.*, 2022). However, from analysing the GTEx/TCGA transcript data for *LNK3*, we observed a substantial decrease in expression in colon and rectal carcinomas. Similarly, *LNK3* promoter activity is not markedly higher in HEK 293T cells versus any of the CRC lines tested, but this cannot be confirmed without further quantification and control measures. Like the *LNK1* promoter, when PKC is activated in

HEK 293T cells there is a significant increase in LNX3 promoter activity. It would be interesting to investigate how supposedly decreased *LNX3* expression impacts tumorigenicity, and other influential factors or stimuli modulate transcription or translation, as LNX3 translation was not studied over the course of this project.

Chapter 7

General Discussion and Future Perspectives

A central challenge in elucidating the biological role of the Ligand of NUMB Protein X family (LNX) lies in accurately defining their interaction networks, functional outputs, regulatory and structural organisation across diverse cellular and evolutionary contexts. The main aim of this thesis was to increase knowledge for each of these elements towards the goal of further understanding these elusive E3 ubiquitin ligases.

7.1 Expanding the LNX1 and 2 interactomes

While high-throughput approaches, including PDZ-interactome databases like ProfAff have identified hundreds of candidate LNX interactors (Gogl *et al.*, 2022), caution must be taken if depending solely on binding affinities. The abundance of candidate interactors contrasts sharply with the tightly regulated expression of LNX proteins, emphasising the necessity of rigorous validation (yeast-two hybrid and co-purification) to distinguish physiologically relevant interactions. Only a subset of candidate interactors identified through Y2H screening were validated in mammalian systems, illustrating the importance of cellular context. We were unable to efficiently express several neuronal proteins including AMIGO1, GRIA2 and IGSF11 in HEK 293T cells, suggesting that it may be necessary to examine these interactions in the context of a neuronal cell type or in brain or other tissue types where these synaptic and membrane-associated proteins are expressed endogenously. Nevertheless, SDK2 and NRXN1 β were successfully expressed upon manipulation of the signal peptide domain, and association with LNX proteins was described. Notably, LNX2-specific interactions, such as SDK2, could advocate functional divergence within the LNX family and highlight the comparatively underexplored nature of the LNX2 interactome. In addition, ProfAff determines that SDK2 interacts with LNX2

PDZ2 and LNX1 PDZ1 when only domain or motif fragments are investigated. When full length proteins are examined *in vitro* studies, we do not observe LNX1 interacting with SKD2, possibly due to intramolecular association between PDZ1 to the carboxyl terminal (See Chapter 5), effectively covering the PDZ1 binding pocket. This result further reinforces the importance of validating results in mammalian cell systems. Subcellular co-localisation studies further supported the physiological relevance of selected interactions, although the dependence on overexpression systems due to limited antibody availability remains an important caveat.

Novel methodological approaches like a yeast surface display of an anti-GFP nanobody, provide a cost and time-effective alternative to commercial immunoprecipitation reagents (O’Riordan *et al.*, 2024). Although we did identify PDZ2 as the mediator of ZNF24 binding to LNX1, supporting previous research that PDZ2 is a promiscuous interaction domain (Wolting *et al.*, 2011) (Guo *et al.*, 2012) (Lenihan *et al.*, 2017) (Lenihan, Saha and Young, 2017), and confirmed well-known interactions (NUMB and Liprin- α 1) using this innovative method, sensitivity of detection remains a concern. Less robust protein-protein interactions may be overlooked. Even though, domain-specific interrogation of protein interactions using ProfAff, Y2H and pull-down methods can condense the potential list of interactors, validation much be completed using full-length proteins.

7.2 The ubiquitome of LNX2

Proximity-based biotin labelling is a fashionable and elegant technique (Tan *et al.*, 2020)(Dumrongprechachan *et al.*, 2021) that has been adapted for use in identifying E3 ubiquitin ligases substrates (Barroso-Gomila *et al.*, 2023) (Mukhopadhyay *et al.*, 2024). By simplifying the BioE3/Ub-POD system to a transient transfection-based approach, experimental efficiency was improved without compromising effectiveness. Harsh wash conditions eliminated a substantial proportion of ubiquitinated material particularly for an E3 ligase with modest catalytic activity such as LNX2. Improvements made by us, improved substrate recovery as seen by Western blot but also accentuated the importance of carefully interpreting background signals including the presence of naturally occurring biotinylated proteins.

When interpreting proteomic data from both WT and RING domain mutant (C53A) LNX2, it was revealed that catalytic activity was not totally diminished. This either raises the question about the ability of a single amino acid substitution in abolishing ubiquitin ligase ability or that LNX2 could also perform as a scaffold for other components of the ubiquitination machinery as is suggested by the RING domain lacking isoform, LNX1^{p70} (Lenihan *et al.*, 2017) A topic that warrants further investigation is whether LNX proteins function as components of larger ubiquitin-regulating protein complexes. The enrichment of multi-protein ubiquitin complexes lacking canonical PDZ-binding motifs, including BABAM2 and BRCC36 of the BRISC/BRCA1-A complexes, provides further support for this hypothesis, although the possibility of non-specific proximity labelling cannot be excluded. Ultimately, distinguishing between these possibilities will require

targeted in vitro and in vivo validation. The formation of higher-order E3 ligase complexes is a well-established mechanism for regulating ubiquitination, as complex assembly can influence substrate specificity, catalytic activity, ubiquitin chain topology, and subcellular localisation. For example, the E3 ligase MDM2 functions as an active homodimer but also forms a heterodimer with MDMX (MDM4), which enhances its ligase activity towards p53 and modulates substrate recognition (Brooks and Gu, 2006) (Ortega-Campos and García-Heredia, 2023). Similarly, many RING E3 ligases, including members of the cullin-RING ligase family, require assembly into multi-subunit complexes to achieve efficient substrate recruitment and ubiquitin transfer. Therefore, it is plausible that LNX proteins similarly associate with regulatory protein complexes that expand or refine their substrate repertoire beyond direct PDZ-mediated interactions. Defining these potential interactions will be important for understanding the molecular mechanisms governing LNX-dependent ubiquitination.

7.3 Structural and functional insights of LNX protein evolution

From an evolutionary standpoint, good conservation of LNX protein domain structure across metazoan lineages underscores their fundamental biological importance. Preservation of key features including RING domain, multiple PDZ domains suggests an ancient and conserved functional role. Overall structural organisation supports the fact that divergence of *LNX* and *MUPP1* from a common, ancestral multiple PDZ domain containing gene early in the metazoan lineage (Flynn, Saha and Young, 2011).

Notably structural modelling indicates that the second PDZ domain adopts a spatially exposed conformation, potentially explaining its dominant role in mediating protein interactions (also seen in Chapter 3 and 4). This feature appears to be widely conserved across the metazoan orthologs tested giving support to its functional significance. Caution is advised when interpreting modelling techniques like AlphaFold 3 as predictive *in silico* outputs, no matter how confident the score is, do not guarantee the same protein behaviour *in vivo* (Abramson *et al.*, 2024). In terms of characterising full length LNX 3D structure there is a large knowledge gap in relation to vertebrate LNX proteins and nothing is known invertebrate species beyond AlphaFold predictions.

The canonical NPAF/Y motif is absent in all invertebrate LNX orthologs. The LNX-NUMB interaction is not present in the cephalochordate and most recent invertebrate proxy of vertebrates, *Branchioma floridae*. The occurrence of NPAY/F motifs is present in all other vertebrates examined (fish, amphibians, mammals and birds), suggest that the NUMB-LNX interaction emerged to support increased complexity in vertebrate systems, especially in the context of nervous system development. Organisms with dorsal hollow nerve cord like structures like *Branchioma floridae*, appear not have a developmental requirement for such an interaction. Despite this fact, the *in vivo* significance of the LNX-NUMB interaction in mammals is not clear. Functions carried out by the LNX-NUMB association could occur possibly due to indirect regulation of NOTCH signalling (Nie *et al.*, 2002) (Zhou *et al.*, 2015) (Lenihan *et al.*, 2017) (Young, 2018). In zebrafish development, the role of the LNX2 ortholog, Lnx2b, has been extensively studied. Lnx2b mediates the ubiquitination of the homeobox domain containing protein Bozozok (BOZ).

Bozozok promotes embryonic dorsalization and is important for anterior-posterior neural patterning in developing embryos (Ro and Dawid, 2009). Dysregulation of the Lnx2b/Lnx2-NUMB interaction also disrupts endocrine cell differentiation in the early zebrafish pancreas, a phenotype resulting due to NOTCH signalling malfunction (Won, Ro and Dawid, 2015). Such functions do not translate to mammals as they lack a Lnx2b protein.

7.4 Analysis of LNX expression in the context of cancer

Despite genomic amplification of the Chr4q12 region (Holtkamp *et al.*, 2007) , the *LNX1* locus, in a subset of gliomas, *LNX1* expression remains low supporting its classification as a passenger gene, in contrast to neighbouring oncogenic drivers that are well described. Conversely, *LNX2* expression in colorectal cancer is comparable to adjacent oncogenes within the Chr13q12 amplicon, which could suggest a functional role. However, discrepancies between transcript and protein detection, as well as inconsistencies and conflicting arguments with previously published studies, raise concerns regarding antibody specificity and experimental reproducibility in the field.

Promoter analyses also demonstrated that LNX1 and LNX2 are subject to tight transcriptional control, with generally low basal level activity. The higher activity of the CNS isoform, LNX1^{p70} promoter in glial and neuronally derived cell lines does indicate that there must be some context specific signals driving transcription. Promoter activity of LNX1 and LNX3 is increased under specific conditions like PKC activation, similarly LNX2 transcription is upregulated when subjected to cellular stresses like serum starvation. These findings suggest that LNX proteins may function in adaptive cellular

responses to environmental or signalling cues rather than as constitutively active regulators. While LNX3 is considerably less studied than LNX1 and LNX2 in the context of cancer, the initial results from literature and our preliminary findings here hint at a role in cancer, especially the reduction of expression in certain cancers like endometrial cancers.

Collectively, this project highlights the multifaceted nature of LNX protein biology. From supplementing the LNX interactomes and substrate networks to unveiling the importance of structure and conservation throughout metazoan evolution, I hope a clearer picture is drawn regarding the Ligand of NUMB Protein X family. A key unifying theme identified throughout this project is the need for rigorous validation and controls – both experimental and computational. Future studies integrating physiologically relevant models, improved detection tools and orthogonal validation strategies will be required to further elucidate the functions of LNX protein in normal biology and disease.

7.5 Identified limitations and future directions

Although this study expands current knowledge of LNX protein biology, several technical and biological limitations should be acknowledged. Many experiments relied on transient overexpression in HEK293T cells, which may alter protein abundance, localisation and interaction dynamics compared with endogenous expression. Likewise, proximity-based biotinylation identifies proteins within a defined spatial radius rather than exclusively direct substrates, meaning some candidate proteins may represent neighbouring complex members instead of bona fide ubiquitination targets. Furthermore, the persistence of ubiquitination following mutation of the catalytic

cysteine (C53A) suggests either incomplete loss of ligase activity or compensation through interactions with endogenous ubiquitination machinery, highlighting the importance of generating catalytically inactive mutants with experimentally verified loss of activity. Finally, the limited availability and specificity of antibodies against endogenous LNX proteins restricted validation in physiologically relevant systems, necessitating the use of tagged constructs and overexpression models throughout much of this work.

Future studies should focus on validating candidate substrates under endogenous expression conditions using approaches such as CRISPR-mediated epitope tagging or gene knockout in physiologically relevant cell types. Determining whether LNX proteins directly ubiquitinate candidate substrates or instead function as components of larger ubiquitin-regulatory complexes, will require the reconstitution of purified proteins *in vitro*, together with ubiquitination assays and quantitative mass spectrometry to define ubiquitin linkage specificity. Identifying the cognate E2 conjugating enzymes will further improve our understanding of LNX catalytic activity and substrate selectivity, while structural studies of full-length LNX proteins using cryo-electron microscopy or X-ray crystallography could provide valuable insight into domain organisation and the proposed intramolecular regulation of PDZ accessibility. Given the highly restricted expression patterns of LNX proteins, extending these studies to neuronal organoids, primary neuronal cultures and genetically engineered animal models will be essential for establishing the physiological significance of LNX-mediated ubiquitination in tissue development and disease. Collectively, these complementary biochemical, structural

and functional approaches will provide a comprehensive framework for defining the molecular mechanisms governing this relatively understudied family of E3 ubiquitin ligases and clarifying their roles in normal physiology and human disease.

Bibliography

- Abdellah, Z. *et al.* (2004) 'Finishing the euchromatic sequence of the human genome', *Nature* 2004 431:7011, 431(7011), pp. 931–945. doi: 10.1038/nature03001.
- Abdul Rehman, S. A. *et al.* (2016) 'MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes', *Molecular Cell*, 63(1), pp. 146–155. doi: 10.1016/j.molcel.2016.05.009.
- Abramson, J. *et al.* (2024) 'Accurate structure prediction of biomolecular interactions with AlphaFold 3', *Nature* 2024 630:8016, 630(8016), pp. 493–500. doi: 10.1038/s41586-024-07487-w.
- Agrata, R. and Komander, D. (2025) 'Ubiquitin-A structural perspective', *Molecular cell*, 85(2), pp. 323–346. doi: 10.1016/J.MOLCEL.2024.12.015.
- Baisiwala, S. *et al.* (2020) 'LNX1 Modulates Notch1 Signaling to Promote Expansion of the Glioma Stem Cell Population during Temozolomide Therapy in Glioblastoma', *Cancers*, 12(12), pp. 1–19. doi: 10.3390/CANCERS12123505.
- Baker, J. M. and Boyce, F. M. (2014) 'High-throughput functional screening using a homemade dual-glow luciferase assay', *Journal of visualized experiments : JoVE*, (88). doi: 10.3791/50282.
- Bang, M. L. and Owczarek, S. (2013) 'A Matter of Balance: Role of Neurexin and Neuroligin at the Synapse', *Neurochemical Research* 2013 38:6, 38(6), pp. 1174–1189. doi: 10.1007/s11064-013-1029-9.
- Barroso-Gomila, O. *et al.* (2023) 'BioE3 identifies specific substrates of ubiquitin E3 ligases', *Nature Communications* 2023 14:1, 14(1), pp. 1–19. doi: 10.1038/s41467-023-43326-8.
- Beasley, S. A. *et al.* (2020) 'An Angelman syndrome substitution in the HECT E3 ubiquitin ligase C-terminal Lobe of E6AP affects protein stability and activity', *PLoS ONE*, 15(7), p. e0235925. doi: 10.1371/JOURNAL.PONE.0235925.
- Békés, M., Langley, D. R. and Crews, C. M. (2022) 'PROTAC targeted protein degraders: the past is prologue', *Nature Reviews Drug Discovery*, 21(3), pp. 181–200. doi: 10.1038/S41573-021-00371-6;SUBJMETA=154,61,631,92;KWRD=BIOTECHNOLOGY,CHEMICAL+BIOLOGY,DRUG+DISCOVERY.
- Bezprozvanny, I. and Maximov, A. (2001) 'Classification of PDZ domains', *FEBS Letters*, 509(3), pp. 457–462. doi: 10.1016/S0014-5793(01)03214-8.
- De Bie, P. and Ciechanover, A. (2011) 'Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms', *Cell Death and Differentiation*, 18(9), p. 1393. doi: 10.1038/cdd.2011.16.

- Blom, T. *et al.* (2008) 'Mutation and copy number analysis of LNX1 and Numb1 in nervous system tumors', *Cancer Genetics and Cytogenetics*, 186(2), pp. 103–109. doi: 10.1016/j.cancergencyto.2008.07.003.
- Branigan, E. *et al.* (2015) 'Structural basis for the RING catalyzed synthesis of K63 linked ubiquitin chains', *Nature structural & molecular biology*, 22(8), p. 597. doi: 10.1038/NSMB.3052.
- Briscoe, J. and Théron, P. P. (2013) 'The mechanisms of Hedgehog signalling and its roles in development and disease', *Nature reviews. Molecular cell biology*, 14(7), pp. 418–431. doi: 10.1038/NRM3598.
- Brooks, C. L. and Gu, W. (2006) 'p53 ubiquitination: Mdm2 and beyond', *Molecular Cell*, 21(3), pp. 307–315. doi: 10.1016/j.molcel.2006.01.020.
- Buetow, L. and Huang, D. T. (2016) 'Structural insights into the catalysis and regulation of E3 ubiquitin ligases', *Nature Reviews Molecular Cell Biology*, 17(10), pp. 626–642. doi: 10.1038/nrm.2016.91.
- Burgner, D. *et al.* (2009) 'A Genome-Wide Association Study Identifies Novel and Functionally Related Susceptibility Loci for Kawasaki Disease', *PLoS Genetics*, 5(1), p. e1000319. doi: 10.1371/JOURNAL.PGEN.1000319.
- Callis, J. (2014) 'The Ubiquitination Machinery of the Ubiquitin System', *The Arabidopsis Book / American Society of Plant Biologists*, 12, p. e0174. doi: 10.1199/TAB.0174.
- Camps, J. *et al.* (2013) 'Genetic amplification of the NOTCH modulator LNX2 upregulates the WNT/ β -catenin pathway in colorectal cancer', *Cancer research*, 73(6), p. 2003. doi: 10.1158/0008-5472.CAN-12-3159.
- Carter, S. *et al.* (2007) 'C-terminal modifications regulate MDM2 dissociation and nuclear export of p53', *Nature Cell Biology* 2007 9:4, 9(4), pp. 428–435. doi: 10.1038/ncb1562.
- Casper, J. *et al.* (2026) 'The UCSC Genome Browser database: 2026 update', *Nucleic acids research*, 54(D1), pp. D1331–D1335. doi: 10.1093/nar/gkaf1250.
- Chen, J. *et al.* (2005) 'Characterization of human LNX, a novel ligand of Numb protein X that is downregulated in human gliomas', *The international journal of biochemistry & cell biology*, 37(11), pp. 2273–2283. doi: 10.1016/J.BIOCEL.2005.02.028.
- Chen, J. *et al.* (2026) 'MDM2 amplification enables selective PROTAC targeting of tumor cells', *Molecular Cancer Therapeutics*. doi: 10.1158/1535-7163.MCT-25-0990.
- Chih, B., Gollan, L. and Scheiffele, P. (2006) 'Alternative splicing controls selective trans-synaptic interactions of the neuroligin-neurexin complex', *Neuron*, 51(2), pp. 171–178. doi: 10.1016/J.NEURON.2006.06.005.
- Cho, K. O., Hunt, C. A. and Kennedy, M. B. (1992) 'The rat brain postsynaptic density

- fraction contains a homolog of the drosophila discs-large tumor suppressor protein', *Neuron*, 9(5), pp. 929–942. doi: 10.1016/0896-6273(92)90245-9.
- Cioccarelli, L. *et al.* (2025) 'Differential neuronal functions of LNX1 and LNX2 revealed by behavioural analysis in single and double knockout mice', *Behavioral and Brain Functions*, 21(1). doi: 10.1186/S12993-025-00276-Z.
- Cohen, N. A. *et al.* (1996) 'Binding of the inward rectifier K⁺ channel Kir 2.3 to PSD-95 is regulated by protein kinase A phosphorylation', *Neuron*, 17(4), pp. 759–767. doi: 10.1016/S0896-6273(00)80207-X.
- Crawford, L. J. *et al.* (2016) 'Identification of the APC/C co-factor FZR1 as a novel therapeutic target for multiple myeloma', *Oncotarget*, 7(43), pp. 70481–93. doi: 10.18632/oncotarget.12026.
- Crawford, L. J. *et al.* (2020) 'The E3 ligase HUWE1 inhibition as a therapeutic strategy to target MYC in multiple myeloma', *Oncogene* 2020 39:27, 39(27), pp. 5001–5014. doi: 10.1038/s41388-020-1345-x.
- Cstorer, A. and Ménard, R. (1994) 'Catalytic mechanism in papain family of cysteine peptidases', *Methods in Enzymology*, 244(C), pp. 486–500. doi: 10.1016/0076-6879(94)44035-2.
- D'Agostino, M. *et al.* (2011) 'Ligand of Numb proteins LNX1p80 and LNX2 interact with the human glycoprotein CD8 α and promote its ubiquitylation and endocytosis', *Journal of cell science*, 124(Pt 21), pp. 3545–3556. doi: 10.1242/JCS.081224.
- Daniels, D. L. *et al.* (1998) 'Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition', *Nature structural biology*, 5(4), pp. 317–325. doi: 10.1038/NSB0498-317.
- Deng, L. *et al.* (2000) 'Activation of the I κ B kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain', *Cell*, 103(2), pp. 351–361. doi: 10.1016/S0092-8674(00)00126-4.
- Deshaies, Raymond J. and Joazeiro, C. A. P. (2009) 'RING domain E3 ubiquitin ligases', *Annual Review of Biochemistry*, 78(Volume 78, 2009), pp. 399–434. doi: 10.1146/ANNUREV.BIOCHEM.78.101807.093809/1.
- Deshaies, Raymond J and Joazeiro, C. A. P. (2009) 'RING Domain E3 Ubiquitin Ligases'. doi: 10.1146/annurev.biochem.78.101807.093809.
- Dho, S. E. *et al.* (1998) 'The Mammalian Numb Phosphotyrosine-binding Domain: CHARACTERIZATION OF BINDING SPECIFICITY AND IDENTIFICATION OF A NOVEL PDZ DOMAIN-CONTAINING NUMB BINDING PROTEIN, LNX *', *Journal of Biological Chemistry*, 273(15), pp. 9179–9187. doi: 10.1074/JBC.273.15.9179.
- Dho, S. E. *et al.* (1999) 'Characterization of Four Mammalian Numb Protein Isoforms: IDENTIFICATION OF CYTOPLASMIC AND MEMBRANE-ASSOCIATED VARIANTS OF THE

PHOSPHOTYROSINE BINDING DOMAIN', *Journal of Biological Chemistry*, 274(46), pp. 33097–33104. doi: 10.1074/JBC.274.46.33097.

Diefenbacher, M. E. *et al.* (2014) 'The deubiquitinase USP28 controls intestinal homeostasis and promotes colorectal cancer', *The Journal of Clinical Investigation*, 124(8), p. 3407. doi: 10.1172/JCI73733.

Dikic, I. (2017) 'Proteasomal and Autophagic Degradation Systems', *Annual review of biochemistry*, 86, pp. 193–224. doi: 10.1146/annurev-biochem-061516-044908.

Doroszko, M. *et al.* (2025) 'The invasion phenotypes of glioblastoma depend on plastic and reprogrammable cell states', *Nature Communications* 2025 16:1, 16(1), pp. 6662-. doi: 10.1038/s41467-025-61999-1.

Doyle, D. A. *et al.* (1996) 'Crystal structures of a complexed and peptide-free membrane protein- binding domain: Molecular basis of peptide recognition by PDZ', *Cell*, 85(7), pp. 1067–1076. doi: 10.1016/S0092-8674(00)81307-0.

Dreos, R. *et al.* (2013) 'EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era', *Nucleic Acids Research*, 41(D1), pp. D157–D164. doi: 10.1093/nar/gks1233.

Dubash, A. D. *et al.* (2013) 'The GEF Bcr activates RhoA/MAL signaling to promote keratinocyte differentiation via desmoglein-1', *Journal of Cell Biology*, 202(4), pp. 653–666. doi: 10.1083/jcb.201304133.

Dubreuil, B. and Levy, E. D. (2021) 'Abundance Imparts Evolutionary Constraints of Similar Magnitude on the Buried, Surface, and Disordered Regions of Proteins', *Frontiers in Molecular Biosciences*, 8, p. 626729. doi: 10.3389/fmolb.2021.626729.

Dumrongprechachan, V. *et al.* (2021) 'Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum', *Nature Communications* 2021 12:1, 12(1), pp. 4855-. doi: 10.1038/s41467-021-25144-y.

Duret, L. *et al.* (2006) 'The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene', *Science*, 312(5780), pp. 1653–1655. doi: 10.1126/SCIENCE.1126316;REQUESTEDJOURNAL:JOURNAL:SCIENCE;CTYPE:STRING:JOURNAL.

Engin, F. and Lee, B. (2010) 'NOTCHing the bone: insights into multi-functionality', *Bone*, 46(2), pp. 274–280. doi: 10.1016/J.BONE.2009.05.027.

Fairhead, M. and Howarth, M. (2015) 'Site-specific biotinylation of purified proteins using BirA', *Methods in molecular biology (Clifton, N.J.)*, 1266, pp. 171–184. doi: 10.1007/978-1-4939-2272-7_12.

Fields, S. and Song, O. K. (1989) 'A novel genetic system to detect protein–protein interactions', *Nature* 1989 340:6230, 340(6230), pp. 245–246. doi: 10.1038/340245a0.

- Flynn, M., Saha, O. and Young, P. (2011) 'Molecular evolution of the LNX gene family', *BMC Evolutionary Biology*, 11(1), p. 235. doi: 10.1186/1471-2148-11-235.
- Freemont, P. *et al.* (no date) 'A novel cysteine-rich sequence motif', *cir.nii.ac.jp*. Available at: <https://cir.nii.ac.jp/crid/1364233270599444992> (Accessed: 12 January 2026).
- Garcia-Barcena, C. *et al.* (2020) 'How to Inactivate Human Ubiquitin E3 Ligases by Mutation', *Frontiers in Cell and Developmental Biology*, 8, p. 39. doi: 10.3389/fcell.2020.00039.
- Garcia-Fernández, J. and Benito-Gutiérrez, È. (2009) 'It's a long way from amphioxus: Descendants of the earliest chordate', *BioEssays*, 31(6), pp. 665–675. doi: 10.1002/bies.200800110.
- Ge, S. X. *et al.* (2020) 'ShinyGO: A graphical gene-set enrichment tool for animals and plants', *Bioinformatics*, 36(8), pp. 2628–2629. doi: 10.1093/BIOINFORMATICS/BT2931.
- Gladkova, C. *et al.* (2018) 'Mechanism of parkin activation by PINK1', *Nature* 2018 559:7714, 559(7714), pp. 410–414. doi: 10.1038/s41586-018-0224-x.
- Gogl, G. *et al.* (2022) 'Quantitative fragmentomics allow affinity mapping of interactomes', *Nature Communications* 2022 13:1, 13(1), pp. 5472-. doi: 10.1038/s41467-022-33018-0.
- Goldman, M. J. *et al.* (2020) 'Visualizing and interpreting cancer genomics data via the Xena platform', *Nature biotechnology*, 38(6), p. 675. doi: 10.1038/s41587-020-0546-8.
- Goldstein, G. *et al.* (1975) 'Isolation of a polypeptide that has lymphocyte differentiating properties and is probably represented universally in living cells', *Proc. natn. Acad. Sci. U.S.A.*, 72(1), p. 11. doi: 10.1073/pnas.72.1.11.
- González-Bermúdez, L. *et al.* (2019) 'Identification of reference genes for RT-qPCR data normalisation in aging studies', *Scientific Reports* 2019 9:1, 9(1), pp. 13970-. doi: 10.1038/s41598-019-50035-0.
- Graham, F. L. and van der Eb, A. J. (1973) 'A new technique for the assay of infectivity of human adenovirus 5 DNA', *Virology*, 52(2), pp. 456–467. doi: 10.1016/0042-6822(73)90341-3.
- Groettrup, M. *et al.* (2008) 'Activating the ubiquitin family: UBA6 challenges the field', *Trends in Biochemical Sciences*, 33(5), pp. 230–237. doi: 10.1016/j.tibs.2008.01.005.
- Gulino, A., Di Marcotullio, L. and Screpanti, I. (2010) 'The multiple functions of Numb', *Experimental Cell Research*, 316(6), pp. 900–906. doi: 10.1016/J.YEXCR.2009.11.017.
- Guo, M., Jan, L. Y. and Jan, Y. N. (1996) 'Control of daughter cell fates during asymmetric division: interaction of Numb and Notch', *Neuron*, 17(1), pp. 27–41. doi: 10.1016/S0896-6273(00)80278-0.

- Gupta, N. *et al.* (2007) 'Whole proteome analysis of post-translational modifications: Applications of mass-spectrometry for proteogenomic annotation', *Genome Research*, 17(9), pp. 1362–1377. doi: 10.1101/GR.6427907.
- Hanahan, D. and Weinberg, R. A. (2000) 'The hallmarks of cancer', *Cell*, pp. 57–70. doi: 10.1016/S0092-8674(00)81683-9.
- Harris, M. A. *et al.* (2007) 'The Gene Ontology project in 2008', *Nucleic Acids Research*, 36(Database issue), p. D440. doi: 10.1093/nar/gkm883.
- Harwood, D. S. L. *et al.* (2025) 'Genomic heterogeneity drives distinct infiltration patterns in glioblastoma', *Acta Neuropathologica Communications* 2025 14:1, 14(1), pp. 5-. doi: 10.1186/s40478-025-02106-9.
- Hastie, E. *et al.* (2015) 'An unexpected inhibition of antiviral signaling by virus-encoded tumor suppressor p53 in pancreatic cancer cells', *Virology*, 483, pp. 126–140. doi: 10.1016/J.VIROL.2015.04.017.
- Hershko, A. and Ciechanover, A. (1998) 'The ubiquitin system', *Annual review of biochemistry*, 67, pp. 425–479. doi: 10.1146/ANNUREV.BIOCHEM.67.1.425.
- Holtkamp, N. *et al.* (2007) 'Characterization of the amplicon on chromosomal segment 4q12 in glioblastoma multiforme', *Neuro-Oncology*, 9(3), p. 291. doi: 10.1215/15228517-2007-009.
- Hong, J., Won, M. and Ro, H. (2020) 'The Molecular and Pathophysiological Functions of Members of the LNX/PDZRN E3 Ubiquitin Ligase Family', *Molecules (Basel, Switzerland)*. NLM (Medline). doi: 10.3390/molecules25245938.
- Hu, W. G. *et al.* (2007) 'Vascular endothelial growth factors C and D represent novel prognostic markers in colorectal carcinoma using quantitative image analysis', *European Surgical Research*, 39(4), pp. 229–238. doi: 10.1159/000101855.
- Huang, L. *et al.* (1999) 'Structure of an E6AP-UbcH7 complex: Insights into ubiquitination by the E2-E3 enzyme cascade', *Science*, 286(5443), pp. 1321–1326. doi: 10.1126/SCIENCE.286.5443.1321;WEBSITE:WEBSITE:AAAS-SITE;JOURNAL:JOURNAL:SCIENCE;WGROU:STRING:PUBLICATION.
- Huibregtse, J. M. *et al.* (1995) 'A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase', *Proceedings of the National Academy of Sciences of the United States of America*, 92(7), pp. 2563–2567. doi: 10.1073/PNAS.92.7.2563.
- Hung, V. *et al.* (2016) 'Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2', *Nature protocols*, 11(3), p. 456. doi: 10.1038/nprot.2016.018.
- Iqbal, U. *et al.* (2010) 'Molecular imaging of glioblastoma multiforme using anti-insulin-like growth factor-binding protein-7 single-domain antibodies', *British Journal of Cancer* 2010 103:10, 103(10), pp. 1606–1616. doi: 10.1038/sj.bjc.6605937.

- Ito, T. *et al.* (2010) 'Identification of a primary target of thalidomide teratogenicity', *Science (New York, N.Y.)*, 327(5971), pp. 1345–1350. doi: 10.1126/SCIENCE.1177319.
- Jang, M. *et al.* (2021) 'LNX1 contributes to cell cycle progression and cisplatin resistance', *Cancers*, 13(16). doi: 10.3390/CANCERS13164066/S1.
- Jang, S. *et al.* (2016) 'Synaptic adhesion molecule IgSF11 regulates synaptic transmission and plasticity', *Nature neuroscience*, 19(1), pp. 84–93. doi: 10.1038/nn.4176.
- Jia, D. *et al.* (2015) 'The endogenous zinc finger transcription factor, ZNF24, modulates the angiogenic potential of human microvascular endothelial cells', *FASEB Journal*, 29(4), pp. 1371–1382. doi: 10.1096/FJ.14-258947/-/DC1.
- Jin, J. *et al.* (2007) 'Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging', *Nature*, 447(7148), pp. 1135–1138. doi: 10.1038/NATURE05902.
- Jin, L. *et al.* (2008) 'Mechanism of Ubiquitin-Chain Formation by the Human Anaphase-Promoting Complex', *Cell*, 133(4), pp. 653–665. doi: 10.1016/j.cell.2008.04.012.
- Joensuu, H. *et al.* (2005) 'Amplification of genes encoding KIT, PDGFRalpha and VEGFR2 receptor tyrosine kinases is frequent in glioblastoma multiforme', *The Journal of pathology*, 207(2), pp. 224–231. doi: 10.1002/path.1823.
- JS, M. and R, K. (2000) 'Notch signaling: from the outside in', *Developmental biology*, 228(2), pp. 151–165. doi: 10.1006/DBIO.2000.9960.
- Jumper, J. *et al.* (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature* 2021, pp. 1–11. doi: 10.1038/s41586-021-03819-2.
- Jurić, V. *et al.* (2025) 'Yeast MoClo secretion and surface display toolkit 2.0: improvements and applications for analysis of protein-protein interactions and whole-cell biocatalysis', *bioRxiv*, p. 2025.08.19.671047. doi: 10.1101/2025.08.19.671047.
- Kang, Y. J. *et al.* (2025) 'GEPIA3: Enhanced drug sensitivity and interaction network analysis for cancer research', *Nucleic Acids Research*, 53(W1), p. W283. doi: 10.1093/NAR/GKAF423.
- Kansaku, A. *et al.* (2006) 'Ligand-of-Numb protein X is an endocytic scaffold for junctional adhesion molecule 4', *Oncogene*, 25(37), pp. 5071–5084. doi: 10.1038/SJ.ONC.1209468.
- Katoh, K. and Standley, D. M. (2013) 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Molecular Biology and Evolution*, 30(4), pp. 772–780. doi: 10.1093/molbev/mst010.
- Koegl, M. and Uetz, P. (2007) 'Improving yeast two-hybrid screening systems', *Briefings in Functional Genomics*, 6(4), pp. 302–312. doi: 10.1093/BFGP/ELM035.
- Komander, D. (2009) 'The emerging complexity of protein ubiquitination', *Biochemical*

Society transactions, 37(Pt 5), pp. 937–953. doi: 10.1042/BST0370937.

Komander, D., Clague, M. J. and Urbé, S. (2009) 'Breaking the chains: structure and function of the deubiquitinases', *Nature Reviews Molecular Cell Biology* 2009 10:8, 10(8), pp. 550–563. doi: 10.1038/nrm2731.

Komander, D. and Rape, M. (2012) 'The ubiquitin code', *Annual Review of Biochemistry*, 81(Volume 81, 2012), pp. 203–229. doi: 10.1146/ANNUREV-BIOCHEM-060310-170328/CITE/REFWORKS.

Kopeck, C. D. *et al.* (2006) 'Glutamate receptor exocytosis and spine enlargement during chemically induced long-term potentiation', *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(7), pp. 2000–2009. doi: 10.1523/JNEUROSCI.3918-05.2006.

Krech, T. *et al.* (2010) 'Characterization of AKT independent effects of the synthetic AKT inhibitors SH-5 and SH-6 using an integrated approach combining transcriptomic profiling and signaling pathway perturbations', *BMC Cancer* 2010 10:1, 10(1), pp. 287-. doi: 10.1186/1471-2407-10-287.

Kriegesmann, J. and Brik, A. (2023) 'Synthesis of ubiquitinated proteins for biochemical and functional analysis', *Chemical Science*, 14(37), pp. 10025–10040. doi: 10.1039/D3SC03664B.

Kronke, J. (2014) 'Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells', *Science*, 343, pp. 301–305.

Kubala, M. H. *et al.* (2010) 'Structural and thermodynamic analysis of the GFP:GFP-nanobody complex', *Protein science : a publication of the Protein Society*, 19(12), pp. 2389–2401. doi: 10.1002/pro.519.

Kwasna, D. *et al.* (2018) 'Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability', *Molecular Cell*, 70(1), p. 150. doi: 10.1016/J.MOLCEL.2018.02.023.

Kwon, Y. T. and Ciechanover, A. (2017) 'The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy', *Trends in Biochemical Sciences*, 42(11), pp. 873–886. doi: 10.1016/j.tibs.2017.09.002.

Lechtenberg, B. C. *et al.* (2016) 'Structure of a HOIP/E2~ubiquitin complex reveals RBR E3 ligase mechanism and regulation', *Nature* 2016 529:7587, 529(7587), pp. 546–550. doi: 10.1038/nature16511.

Lee, H. J. and Zheng, J. J. (2010) 'PDZ domains and their binding partners: structure, specificity, and modification', *Cell communication and signaling : CCS*, 8. doi: 10.1186/1478-811X-8-8.

Lee, M. E. *et al.* (2015) 'A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly', *ACS Synthetic Biology*, 4(9), pp. 975–986. doi: 10.1021/sb500366v.

- Lenihan, J. A. *et al.* (2014) 'Tight, cell type-specific control of LNX expression in the nervous system, at the level of transcription, translation and protein stability', *Gene*, 552(1), pp. 39–50. doi: 10.1016/J.GENE.2014.09.011.
- Lenihan, J. A. *et al.* (2017) 'Decreased Anxiety-Related Behaviour but Apparently Unperturbed NUMB Function in Ligand of NUMB Protein-X (LNX) 1/2 Double Knockout Mice', *Molecular Neurobiology*, 54(10), pp. 8090–8109. doi: 10.1007/s12035-016-0261-0.
- Lenihan, J. A., Saha, O. and Young, P. W. (2017) 'Proteomic analysis reveals novel ligands and substrates for LNX1 E3 ubiquitin ligase', *PLoS ONE*, 12(11). doi: 10.1371/journal.pone.0187352.
- Lesurf, R. *et al.* (2015) 'OREGAnno 3.0: a community-driven resource for curated regulatory annotation', *Nucleic Acids Research*, 44(Database issue), p. D126. doi: 10.1093/nar/gkv1203.
- Letunic, I. and Bork, P. (2024) 'Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool', *Nucleic Acids Research*, 52(W1), pp. W78–W82. doi: 10.1093/nar/gkae268.
- Letunic, I. and Bork, P. (2025) 'SMART v10: three decades of the protein domain annotation resource', *Nucleic Acids Research*. doi: 10.1093/NAR/GKAF1023.
- Lewis, M. J. and Wang, S. (2024) 'locuszoomr: an R package for visualizing publication-ready regional gene locus plots', *Bioinformatics Advances*, 5(1). doi: 10.1093/bioadv/vbaf006.
- Li, J. Z. *et al.* (2013) 'Identification of a Functional Nuclear Localization Signal Mediating Nuclear Import of the Zinc Finger Transcription Factor ZNF24', *PLOS ONE*, 8(11), p. e79910. doi: 10.1371/journal.pone.0079910.
- Li, M. *et al.* (2003) 'Mono- versus polyubiquitination: differential control of p53 fate by Mdm2', *Science (New York, N.Y.)*, 302(5652), pp. 1972–1975. doi: 10.1126/SCIENCE.1091362.
- Li, Qihong *et al.* (2022) 'Lower expression of PDZRN3 induces endometrial carcinoma progression via the activation of canonical Wnt signaling', *Oncology Letters*, 23(3), p. 98. doi: 10.3892/ol.2022.13218.
- Li, Qin *et al.* (2022) 'The deubiquitinating enzyme complex BRISC regulates Aurora B activation via lysine-63-linked ubiquitination in mitosis', *Communications Biology* 2022 5:1, 5(1), pp. 1335–. doi: 10.1038/s42003-022-04299-4.
- Li, W. *et al.* (2008) 'Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling', *PLOS ONE*, 3(1), p. e1487. doi: 10.1371/JOURNAL.PONE.0001487.
- Liu, J. *et al.* (2011) 'Beclin1 controls the levels of p53 by regulating the deubiquitination

- activity of USP10 and USP13', *Cell*, 147(1), pp. 223–234. doi: 10.1016/j.cell.2011.08.037.
- Liu, X.-D. *et al.* (2018) 'Retrograde regulation of mossy fiber axon targeting and terminal maturation via postsynaptic Lnx1', *Journal of Cell Biology*, 217(11), pp. 4007–4024. doi: 10.1083/JCB.201803105.
- Liu, X. D. *et al.* (2018) 'Retrograde regulation of mossy fiber axon targeting and terminal maturation via postsynaptic Lnx1', *The Journal of Cell Biology*, 217(11), p. 4007. doi: 10.1083/JCB.201803105.
- Liu, X. D. *et al.* (2019) 'Hippocampal Lnx1–NMDAR multiprotein complex mediates initial social memory', *Molecular Psychiatry*, pp. 1–14. doi: 10.1038/s41380-019-0606-y.
- Livak, K. J. and Schmittgen, T. D. (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method', *Methods*, 25(4), pp. 402–408. doi: 10.1006/meth.2001.1262.
- Lu, G. (2014) 'The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins', *Science*, 343, pp. 305–309.
- Luo, J., Chen, J.-W., *et al.* (2022) 'TBX20 inhibits colorectal cancer tumorigenesis by impairing NHEJ-mediated DNA repair', *Cancer Science*. doi: 10.1111/CAS.15348.
- Luo, J., Chen, J. W., *et al.* (2022) 'TBX20 inhibits colorectal cancer tumorigenesis by impairing NHEJ-mediated DNA repair', *Cancer science*, 113(6), pp. 2008–2021. doi: 10.1111/cas.15348.
- M, K. and M, K. (2004) 'Identification and characterization of PDZRN3 and PDZRN4 genes in silico', *International journal of molecular medicine*, 13(4), pp. 607–613. Available at: <https://pubmed.ncbi.nlm.nih.gov/15010864/> (Accessed: 15 July 2021).
- Ma, L. *et al.* (2017a) 'Suppression of cancer stemness by upregulating Ligand-of-Numb protein X1 in colorectal carcinoma', *PLOS ONE*, 12(11), p. e0188665. doi: 10.1371/JOURNAL.PONE.0188665.
- Ma, L. *et al.* (2017b) 'Suppression of cancer stemness by upregulating Ligand-of-Numb protein X1 in colorectal carcinoma', *PLOS ONE*, 12(11), p. e0188665. doi: 10.1371/JOURNAL.PONE.0188665.
- Madeira, F. *et al.* (2024) 'The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024.', *Nucleic Acids Research*, 52(W1), pp. W521–W525. doi: 10.1093/nar/gkae241.
- Majka, J. and Burgers, P. M. J. (2004) 'The PCNA-RFC Families of DNA Clamps and Clamp Loaders', *Progress in Nucleic Acid Research and Molecular Biology*, 78, pp. 227–260. doi: 10.1016/S0079-6603(04)78006-X.
- McCrindle, B. W. *et al.* (2017) 'Diagnosis, Treatment, and Long-Term Management of

- Kawasaki Disease: A Scientific Statement for Health Professionals From the American Heart Association', *Circulation*, 135(17), pp. e927–e999. doi: 10.1161/CIR.0000000000000484.
- Mehraj, V. *et al.* (2012) 'The ligands of Numb proteins X1 and X2 are specific markers for chronic Q fever', *FEMS Immunology & Medical Microbiology*, 64(1), pp. 98–100. doi: 10.1111/J.1574-695X.2011.00860.X.
- Meng, E. C. *et al.* (2023) 'UCSF ChimeraX: Tools for structure building and analysis', *Protein science : a publication of the Protein Society*, 32(11). doi: 10.1002/PRO.4792.
- Microorganisms & Microbial-Derived Ingredients Used in Food (Partial List) | FDA* (no date). Available at: <https://www.fda.gov/food/generally-recognized-safe-gras/microorganisms-microbial-derived-ingredients-used-food-partial-list> (Accessed: 11 February 2026).
- Min, X. *et al.* (2019) 'Mdm2-mediated ubiquitination of PKC β II in the nucleus mediates clathrin-mediated endocytic activity', *Biochemical Pharmacology*, 170. doi: 10.1016/j.bcp.2019.113675.
- Mirza, M. *et al.* (2005) 'The cell surface protein coxsackie- and adenovirus receptor (CAR) directly associates with the Ligand-of-Numb Protein-X2 (LNx2)', *Experimental Cell Research*, 309(1), pp. 110–120. doi: 10.1016/j.yexcr.2005.05.023.
- Mirza, M. *et al.* (2006) 'Coxsackievirus and adenovirus receptor (CAR) is expressed in male germ cells and forms a complex with the differentiation factor JAM-C in mouse testis', *Experimental Cell Research*, 312(6), pp. 817–830. doi: 10.1016/J.YEXCR.2005.11.030.
- Moore, J. E. *et al.* (2024) 'An Expanded Registry of Candidate cis-Regulatory Elements for Studying Transcriptional Regulation', *bioRxiv*, p. 2024.12.26.629296. doi: 10.1101/2024.12.26.629296.
- Morais Cabral, J. H. *et al.* (1996) 'Crystal structure of a PDZ domain', *Nature*, 382(6592), pp. 649–652. doi: 10.1038/382649A0.
- Mukhopadhyay, D. and Riezman, H. (2007) 'Proteasome-independent functions of ubiquitin in endocytosis and signaling', *Science*, 315(5809), pp. 201–205. doi: 10.1126/SCIENCE.1127085;ISSUE:ISSUE:DOI.
- Mukhopadhyay, U. *et al.* (2024) 'A ubiquitin-specific, proximity-based labeling approach for the identification of ubiquitin ligase substrates', *Science Advances*, 10(32), p. eadp3000. doi: 10.1126/SCIADV.ADP3000.
- Muzny, D. M. *et al.* (2012) 'Comprehensive molecular characterization of human colon and rectal cancer', *Nature*, 487(7407), pp. 330–337. doi: 10.1038/NATURE11252.
- Myant, K. B. *et al.* (2017) 'HUWE1 is a critical colonic tumour suppressor gene that prevents MYC signalling, DNA damage accumulation and tumour initiation', *EMBO Mol*

Med, 9(2), pp. 181–97. doi: 10.15252/emmm.201606684.

Nagayama, S. *et al.* (2004) 'Identification of PDZK4, a novel human gene with PDZ domains, that is upregulated in synovial sarcomas', *Oncogene*, 23(32), pp. 5551–5557. doi: 10.1038/SJ.ONC.1207710.

Nardella, C. *et al.* (2021) 'Targeting PDZ domains as potential treatment for viral infections, neurodegeneration and cancer', *Biology direct*, 16(1). doi: 10.1186/S13062-021-00303-9.

Nayak, D. *et al.* (2015) 'Structural basis for the indispensable role of a unique zinc finger motif in LNX2 ubiquitination', *Oncotarget*, 6(33), pp. 34342–34357. doi: 10.18632/ONCOTARGET.5326.

Nayak, D. and Sivaraman, J. (2018a) 'Structure of LNX1:Ubc13 ~ Ubiquitin Complex Reveals the Role of Additional Motifs for the E3 Ligase Activity of LNX1', *Journal of Molecular Biology*, 430(8), pp. 1173–1188. doi: 10.1016/J.JMB.2018.02.016.

Nayak, D. and Sivaraman, J. (2018b) 'Structure of LNX1:Ubc13~Ubiquitin Complex Reveals the Role of Additional Motifs for the E3 Ligase Activity of LNX1', *Journal of molecular biology*, 430(8), pp. 1173–1188. doi: 10.1016/J.JMB.2018.02.016.

Newton, A. C. (2018) 'Protein Kinase C: Perfectly Balanced', *Critical reviews in biochemistry and molecular biology*, 53(2), p. 208. doi: 10.1080/10409238.2018.1442408.

Nie, J. *et al.* (2002a) 'LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation', *The EMBO Journal*, 21(1–2), p. 93. doi: 10.1093/EMBOJ/21.1.93.

Nie, J. *et al.* (2002b) 'LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation', *The EMBO Journal*, 21(1–2), p. 93. doi: 10.1093/EMBOJ/21.1.93.

Nie, J., Li, S. S.-C. and McGlade, C. J. (2004) 'A Novel PTB-PDZ Domain Interaction Mediates Isoform-specific Ubiquitylation of Mammalian Numb *', *Journal of Biological Chemistry*, 279(20), pp. 20807–20815. doi: 10.1074/JBC.M311396200.

Nielsen, J. (2025) 'The Power of Yeast', *Yeast (Chichester, England)*, 42(12), p. 303. doi: 10.1002/yea.70009.

Nijman, S. M. B. *et al.* (2005) 'A genomic and functional inventory of deubiquitinating enzymes', *Cell*, 123(5), pp. 773–786. doi: 10.1016/J.CELL.2005.11.007.

O'Riordan, N. M. *et al.* (2024) 'A Yeast Modular Cloning (MoClo) Toolkit Expansion for Optimization of Heterologous Protein Secretion and Surface Display in *Saccharomyces cerevisiae*', *ACS Synthetic Biology*, 13(4), pp. 1246–1258. doi: 10.1021/acssynbio.3c00743.

Ordureau, A. *et al.* (2014) 'Quantitative proteomics reveal a feedforward mechanism

for mitochondrial PARKIN translocation and ubiquitin chain synthesis', *Molecular Cell*, 56(3), pp. 360–375. doi: 10.1016/j.molcel.2014.09.007.

Ortega-Campos, S. M. and García-Heredia, J. M. (2023) 'The Multitasker Protein: A Look at the Multiple Capabilities of NUMB', *Cells*, 12(2), p. 333. doi: 10.3390/CELLS12020333.

Park, R. *et al.* (2019) 'LNX1 contributes to tumor growth by down-regulating p53 stability', *The FASEB Journal*, 33(12), pp. 13216–13227. doi: 10.1096/FJ.201900366R.

Passmore, L. A. and Barford, D. (2004) 'Getting into position: the catalytic mechanisms of protein ubiquitylation', *Biochemical Journal*, 379(Pt 3), p. 513. doi: 10.1042/BJ20040198.

Pece, S. *et al.* (2011) 'NUMB-ing down cancer by more than just a NOTCH', *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1815(1), pp. 26–43. doi: 10.1016/J.BBCAN.2010.10.001.

Pedersen, W. A. *et al.* (2002) 'Numb isoforms containing a short PTB domain promote neurotrophic factor-induced differentiation and neurotrophic factor withdrawal-induced death of PC12 Cells', *Journal of neurochemistry*, 82(4), pp. 976–986. doi: 10.1046/J.1471-4159.2002.01036.X.

Peltola, M. A. *et al.* (2016) 'AMIGO-Kv2.1 Potassium Channel Complex Is Associated With Schizophrenia-Related Phenotypes', *Schizophrenia bulletin*, 42(1), pp. 191–201. doi: 10.1093/schbul/sbv105.

Peng, J. *et al.* (2003) 'A proteomics approach to understanding protein ubiquitination', *nature.com* J Peng, D Schwartz, JE Elias, CC Thoreen, D Cheng, G Marsischky, J Roelofs, D Finley *Nature biotechnology*, 2003 • *nature.com*. Available at: https://idp.nature.com/authorize/casa?redirect_uri=https://www.nature.com/articles/nbt849&casa_token=Jxi-y9_4m60AAAAA:DzNlwXEN1SQ17faflXRqQek6vq9QtkWlj-ECDnvoZ-xm0KvkNGfjY5Fq0SDOEae60FEQh40n5zFeaoJX0Q (Accessed: 14 January 2026).

Pickart, C. M. (2001) 'Mechanisms underlying ubiquitination', *Annual Review of Biochemistry*, 70(Volume 70, 2001), pp. 503–533. doi: 10.1146/ANNUREV.BIOCHEM.70.1.503/CITE/REFWORKS.

Pickart, C. M. and Eddins, M. J. (2004) 'Ubiquitin: Structures, functions, mechanisms', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1695(1–3), pp. 55–72. doi: 10.1016/j.bbamcr.2004.09.019.

Popov, N. *et al.* (2007) 'The ubiquitin-specific protease USP28 is required for MYC stability', *Nature cell biology*, 9(7), pp. 765–774. doi: 10.1038/NCB1601.

Prescott, S. M. *et al.* (2000) 'Platelet-activating factor and related lipid mediators', *Annual review of biochemistry*, 69, pp. 419–445. doi: 10.1146/annurev.biochem.69.1.419.

- Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) 'FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments', *PLOS ONE*, 5(3), p. e9490. doi: 10.1371/journal.pone.0009490.
- Prieto, G., Rodríguez, J. A. and Fullaondo, A. (2024) 'Enhancing prediction of short linear protein motifs with Wregex 3.0', *Computational and Structural Biotechnology Journal*, 23, pp. 2978–2984. doi: 10.1016/j.csbj.2024.07.013.
- Rabl, J. (2020) 'BRCA1-A and BRISC: Multifunctional Molecular Machines for Ubiquitin Signaling', *Biomolecules* 2020, Vol. 10, Page 1503, 10(11), p. 1503. doi: 10.3390/biom10111503.
- Rehman, S. U. *et al.* (2023) 'Recent insights into the functions and mechanisms of antisense RNA: emerging applications in cancer therapy and precision medicine', *Frontiers in Chemistry*, 11, p. 1335330. doi: 10.3389/fchem.2023.1335330.
- Reyes-Turcu, F. E., Ventii, K. H. and Wilkinson, K. D. (2009) 'Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes', *Annual review of biochemistry*, 78, pp. 363–397. doi: 10.1146/ANNUREV.BIOCHEM.78.082307.091526.
- Rice, D. S., Northcutt, G. M. and Kurschner, C. (2001) 'The Lnx family proteins function as molecular scaffolds for Numb family proteins', *Molecular and Cellular Neuroscience*, 18(5), pp. 525–540. doi: 10.1006/mcne.2001.1024.
- Ro, H. and Dawid, I. B. (2009) 'Organizer restriction through modulation of Bozozok stability by the E3 ubiquitin ligase Lnx-like', *Nature cell biology*, 11(9), p. 1121. doi: 10.1038/NCB1926.
- Rotin, D. and Kumar, S. (2009) 'Physiological functions of the HECT family of ubiquitin ligases', *Nature reviews. Molecular cell biology*, 10(6), pp. 398–409. doi: 10.1038/NRM2690.
- Sakamoto, K. M. *et al.* (2001) 'Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation', *Proceedings of the National Academy of Sciences of the United States of America*, 98(15), p. 8554. doi: 10.1073/PNAS.141230798.
- Salpietro, V. *et al.* (2019) 'AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders', *Nature communications*, 10(1). doi: 10.1038/s41467-019-10910-w.
- Saras, J. and Heldin, C. H. (1996) 'PDZ domains bind carboxy-terminal sequences of target proteins', *Trends in Biochemical Sciences*, 21(12), pp. 455–458. doi: 10.1016/S0968-0004(96)30044-3.
- Scheffner, M., Nuber, U. and Huibregtse, J. M. (1995) 'Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade', *Nature* 1995 373:6509, 373(6509), pp. 81–83. doi: 10.1038/373081a0.

- Schlesinger, D. H. and Goldstein, G. (1975) 'Molecular conservation of 74 amino acid sequence of ubiquitin between cattle and man', *Nature* 1975 255:5507, 255(5507), pp. 423–424. doi: 10.1038/255423a0.
- Schoenheimer, R. (1942) 'The dynamic state of body constituents.' doi: 10.5555/19442200657.
- Schreiber, A. *et al.* (2011) 'Structural basis for the subunit assembly of the anaphase-promoting complex', *Nature*, 470(7333), pp. 227–235. doi: 10.1038/NATURE09756.
- Schulman, B. A. and Wade Harper, J. (2009) 'Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways', *Nature reviews. Molecular cell biology*, 10(5), pp. 319–331. doi: 10.1038/NRM2673.
- Sheng, X. *et al.* (2023) 'The ubiquitin codes in cellular stress responses', *Protein & Cell*, 15(3), p. 157. doi: 10.1093/procel/pwad045.
- Sloper-Mould, K. E. *et al.* (2001) 'Distinct functional surface regions on ubiquitin', *The Journal of biological chemistry*, 276(32), pp. 30483–30489. doi: 10.1074/JBC.M103248200.
- Sollerbrant, K. *et al.* (2003) 'The Coxsackievirus and adenovirus receptor (CAR) forms a complex with the PDZ domain-containing protein Ligand-of-Numb protein-X (LNX)', *Journal of Biological Chemistry*, 278(9), pp. 7439–7444. doi: 10.1074/jbc.M205927200.
- Songyang, Z. *et al.* (1997) 'Recognition of unique carboxyl-terminal motifs by distinct PDZ domains', *Science*, 275(5296), pp. 73–77. doi: 10.1126/science.275.5296.73.
- Sun, M. K. and Alkon, D. L. (2014) 'The "Memory Kinases": Roles of PKC Isoforms in Signal Processing and Memory Formation', *Progress in Molecular Biology and Translational Science*, 122, pp. 31–59. doi: 10.1016/B978-0-12-420170-5.00002-7.
- Sun, X. *et al.* (2022) 'Deep Single-Cell-Type Proteome Profiling of Mouse Brain by Nonsurgical AAV-Mediated Proximity Labeling', *Analytical Chemistry*, 94(13), pp. 5325–5334. doi: 10.1021/acs.analchem.1c05212.
- Takahashi, S. *et al.* (2009) 'The E3 ubiquitin ligase LNX1p80 promotes the removal of claudins from tight junctions in MDCK cells', *Journal of cell science*, 122(Pt 7), pp. 985–994. doi: 10.1242/JCS.040055.
- Talis, A. L., Huibregtse, J. M. and Howley, P. M. (1998) 'The role of E6AP in the regulation of p53 protein levels in human papillomavirus (HPV)-positive and HPV-negative cells', *Journal of Biological Chemistry*, 273(11), pp. 6439–6445. doi: 10.1074/jbc.273.11.6439.
- Tan, B. *et al.* (2020) 'An Optimized Protocol for Proximity Biotinylation in Confluent Epithelial Cell Cultures Using the Peroxidase APEX2', *STAR Protocols*, 1(2), p. 100074. doi: 10.1016/j.xpro.2020.100074.
- Thapa, R. *et al.* (2024) 'CRBN-PROTACs in Cancer Therapy: From Mechanistic Insights to

Clinical Applications', *Chemical Biology and Drug Design*, 104(5), p. e70009. doi: 10.1111/CBDD.70009;JOURNAL:JOURNAL:13993011B;PAGE:STRING:ARTICLE/CHAPTER.

The Cancer Genome Atlas Program (TCGA) - NCI (no date). Available at: <https://www.cancer.gov/ccg/research/genome-sequencing/tcga> (Accessed: 20 March 2026).

Thiefes, A. *et al.* (2006) 'The *Yersinia enterocolitica* effector YopP inhibits host cell signalling by inactivating the protein kinase TAK1 in the IL-1 signalling pathway', *EMBO reports*, 7(8), pp. 838–844. doi: 10.1038/sj.embor.7400754.

Thrower, J. S. *et al.* (2000) 'Recognition of the polyubiquitin proteolytic signal', *The EMBO journal*, 19(1), pp. 94–102. doi: 10.1093/EMBOJ/19.1.94.

Tian, Y. *et al.* (2020) 'Expression and Clinical Significance of POLR1D in Colorectal Cancer', *Oncology*, 98(3), pp. 138–145. doi: 10.1159/000504174.

Tissue expression of LNX2 - Summary - The Human Protein Atlas (no date). Available at: https://www.proteinatlas.org/ENSG00000139517-LNX2/tissue#rna_expression (Accessed: 21 January 2026).

Trincado, J. L. *et al.* (2016) 'The prognostic potential of alternative transcript isoforms across human tumors', *Genome Medicine* 2016 8:1, 8(1), pp. 85-. doi: 10.1186/s13073-016-0339-3.

Uemura, T. *et al.* (1989) 'numb, a gene required in determination of cell fate during sensory organ formation in *Drosophila* embryos', *Cell*, 58(2), pp. 349–360. doi: 10.1016/0092-8674(89)90849-0.

Uhlén, M. *et al.* (2015) 'Tissue-based map of the human proteome', *Science*, 347(6220). doi: 10.1126/SCIENCE.1260419;SUBPAGE:STRING:ABSTRACT;WEBSITE:WEBSITE:AAAS-SITE;JOURNAL:JOURNAL:SCIENCE;WGROU:STRING:PUBLICATION.

Vasquez-Del Carpio, R. *et al.* (2011) 'Assembly of a Notch transcriptional activation complex requires multimerization', *Molecular and cellular biology*, 31(7), pp. 1396–1408. doi: 10.1128/MCB.00360-10.

Verdi, J. M. *et al.* (1999) 'Distinct human NUMB isoforms regulate differentiation vs. proliferation in the neuronal lineage', *Proceedings of the National Academy of Sciences of the United States of America*, 96(18), pp. 10472–10476. doi: 10.1073/PNAS.96.18.10472.

Vijay-kumar, S., Bugg, C. E. and Cook, W. J. (1987) 'Structure of ubiquitin refined at 1.8 Å resolution', *Journal of molecular biology*, 194(3), pp. 531–544. doi: 10.1016/0022-2836(87)90679-6.

Vittal, V. *et al.* (2015) 'Regulating the Regulators: Recent Revelations in the Control of E3 Ubiquitin Ligases', *The Journal of Biological Chemistry*, 290(35), p. 21244. doi: 10.1074/jbc.R115.675165.

- Wang, H. *et al.* (2014) 'Gli3 repressor controls cell fates and cell adhesion for proper establishment of neurogenic niche', *Cell Reports*, 8(4), pp. 1093–1104. doi: 10.1016/j.celrep.2014.07.006.
- Wang, H. *et al.* (2023) 'Co-amplified with PDGFRA, IGFBP7 is a prognostic biomarker correlated with the immune infiltrations of glioma', *Cancer Medicine*, 12(4), pp. 4951–4967. doi: 10.1002/cam4.5187.
- Wang, J. *et al.* (2022) 'Transcriptional factor IIA promotes colorectal cancer progression by upregulating cystatin A', *World Journal of Gastrointestinal Oncology*, 14(10), p. 1918. doi: 10.4251/wjgo.v14.i10.1918.
- Wang, J. *et al.* (2023) 'The conserved domain database in 2023', *Nucleic acids research*, 51(D1), pp. D384–D388. doi: 10.1093/nar/gkac1096.
- Wang, L. H. *et al.* (1999) 'A PDZ protein regulates the distribution of the transmembrane semaphorin, M-SemF', *Journal of Biological Chemistry*, 274(20), pp. 14137–14146. doi: 10.1074/jbc.274.20.14137.
- Wang, M. *et al.* (2019) 'POLR1D promotes colorectal cancer progression and predicts poor prognosis of patients', *Molecular carcinogenesis*, 58(5), pp. 735–748. doi: 10.1002/mc.22966.
- Wang, X. S. *et al.* (2023) 'The unifying catalytic mechanism of the RING-between-RING E3 ubiquitin ligase family', *Nature Communications* 2023 14:1, 14(1), pp. 168-. doi: 10.1038/s41467-023-35871-z.
- Wei, J. *et al.* (2013) 'A new mechanism of RhoA ubiquitination and degradation: Roles of SCFFBXL19 E3 ligase and Erk2', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(12), pp. 2757–2764. doi: 10.1016/j.bbamcr.2013.07.005.
- Weiss, A. *et al.* (2007) 'c-Src is a PDZ interaction partner and substrate of the E3 ubiquitin ligase Ligand-of-Numb protein X1', *FEBS letters*, 581(26), pp. 5131–5136. doi: 10.1016/J.FEBSLET.2007.09.062.
- Wethmar, K., Smink, J. J. and Leutz, A. (2010) 'Upstream open reading frames: molecular switches in (patho)physiology', *BioEssays : news and reviews in molecular, cellular and developmental biology*, 32(10), pp. 885–893. doi: 10.1002/BIES.201000037.
- Wilkinson, K. D. (2005) 'The discovery of ubiquitin-dependent proteolysis', *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), pp. 15280–15282. doi: 10.1073/PNAS.0504842102.
- Wolting, C. D. *et al.* (2011) 'Biochemical and Computational Analysis Of LNX1 Interacting Proteins', *PLoS ONE*. Edited by A. K. Tyagi, 6(11), p. e26248. doi: 10.1371/journal.pone.0026248.
- Won, M., Ro, H. and Dawid, I. B. (2015) 'Lnx2 ubiquitin ligase is essential for exocrine

- cell differentiation in the early zebrafish pancreas', *Proceedings of the National Academy of Sciences of the United States of America*, 112(40), p. 12426. doi: 10.1073/PNAS.1517033112.
- Xie, Y. *et al.* (2022) 'IBS 2.0: an upgraded illustrator for the visualization of biological sequences', *Nucleic Acids Research*, 50(W1), pp. W420–W426. doi: 10.1093/NAR/GKAC373.
- Xu, J. and Zhang, Y. (2010) 'How significant is a protein structure similarity with TM-score = 0.5?', *Bioinformatics*, 26(7), pp. 889–895. doi: 10.1093/BIOINFORMATICS/BTQ066.
- Xu, P. *et al.* (2009) 'Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation', *Cell*, 137(1), pp. 133–145. doi: 10.1016/j.cell.2009.01.041.
- Y, X. *et al.* (2001) 'Identification of a human LNX protein containing multiple PDZ domains', *Biochemical genetics*, 39(3–4), pp. 117–126. doi: 10.1023/A:1010269908398.
- Yamagata, M. (2020) 'Structure and Functions of Sidekicks', *Frontiers in Molecular Neuroscience*, 13, p. 139. doi: 10.3389/fnmol.2020.00139.
- Yamagata, M. and Sanes, J. R. (2008) 'Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina', *Nature*, 451(7177), pp. 465–469. doi: 10.1038/NATURE06469.
- Yan, Z. *et al.* (2021) 'Positive feedback of SuFu negating protein 1 on Hedgehog signaling promotes colorectal tumor growth', *Cell death & disease*, 12(2). doi: 10.1038/S41419-021-03487-0.
- Yang, Y. *et al.* (2026) 'Targeting palmitoylation: A novel frontier in cancer biology and immunotherapy', *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1881(1), p. 189509. doi: 10.1016/j.bbcan.2025.189509.
- Yates, A. D. *et al.* (2026) 'Ensembl 2026', *Nucleic acids research*, 54(D1), pp. D1053–D1060. doi: 10.1093/nar/gkaf1239.
- Young, P. *et al.* (2005) 'LNX1 is a perisynaptic Schwann cell specific E3 ubiquitin ligase that interacts with ErbB2', *Molecular and Cellular Neuroscience*, 30(2), pp. 238–248. doi: 10.1016/J.MCN.2005.07.015.
- Young, P. W. (2018) 'LNX1/LNX2 proteins: functions in neuronal signalling and beyond', *Neuronal Signaling*, 2(2). doi: 10.1042/ns20170191.
- Z, G. *et al.* (2012a) 'Proteomics strategy to identify substrates of LNX, a PDZ domain-containing E3 ubiquitin ligase', *Journal of proteome research*, 11(10), pp. 4847–4862. doi: 10.1021/PR300674C.
- Z, G. *et al.* (2012b) 'Proteomics strategy to identify substrates of LNX, a PDZ domain-containing E3 ubiquitin ligase', *Journal of proteome research*, 11(10), pp. 4847–4862.

doi: 10.1021/PR300674C.

Zhang, J. *et al.* (2025) 'Predicting protein-protein interactions in the human proteome', *Science*, 390(6771). doi: 10.1126/science.adt1630.

Zhang, Y. and Skolnick, J. (2004) 'Scoring function for automated assessment of protein structure template quality', *Proteins: Structure, Function and Genetics*, 57(4), pp. 702–710. doi: 10.1002/PROT.20264.

Zhen, L. *et al.* (2018) 'EphB2 deficiency induces depression-like behaviors and memory impairment: Involvement of NMDA 2B receptor dependent signaling', *Frontiers in Pharmacology*, 9(AUG), p. 400968. doi: 10.3389/fphar.2018.00862.

Zheng, D. *et al.* (2010) 'LNX (Ligand of Numb-protein X) interacts with RhoC, both of which regulate AP-1-mediated transcriptional activation', *Molecular biology reports*, 37(5), pp. 2431–2437. doi: 10.1007/S11033-009-9754-5.

Zhong, W. *et al.* (1997) 'Differential expression of mammalian Numb, Numlike and Notch1 suggests distinct roles during mouse cortical neurogenesis', *Development*, 124(10), pp. 1887–1897. doi: 10.1242/DEV.124.10.1887.

Zhou, J. *et al.* (2015) 'Ubiquitin E3 Ligase LNX2 is Critical for Osteoclastogenesis in vitro by Regulating M-CSF/RANKL Signaling and Notch2', *Calcified tissue international*, 96(5), p. 465. doi: 10.1007/S00223-015-9967-7.

Zhou, Q. *et al.* (2020) 'Cell-free DNA analysis reveals POLR1D-mediated resistance to bevacizumab in colorectal cancer', *Genome Medicine*, 12(1), p. 20. doi: 10.1186/s13073-020-0719-6.

Zhou, Q. and Zhang, J. (2022) 'K27-linked noncanonic ubiquitination in immune regulation', *Journal of Leukocyte Biology*, 111(1), pp. 223–235. doi: 10.1002/JLB.4RU0620-397RR;WGROU:STRING:PUBLICATION.

Zhu, B. *et al.* (2005) 'Monoubiquitination of human histone H2B: The factors involved and their roles in HOX gene regulation', *Molecular Cell*, 20(4), pp. 601–611. doi: 10.1016/j.molcel.2005.09.025.

Zimmermann, P. *et al.* (2002) 'PIP(2)-PDZ domain binding controls the association of syntenin with the plasma membrane', *Mol Cell*, 9(6), pp. 1215–1225. doi: 10.1016/s1097-2765(02)00549-x.

Supplementary Figures

Chapter 3

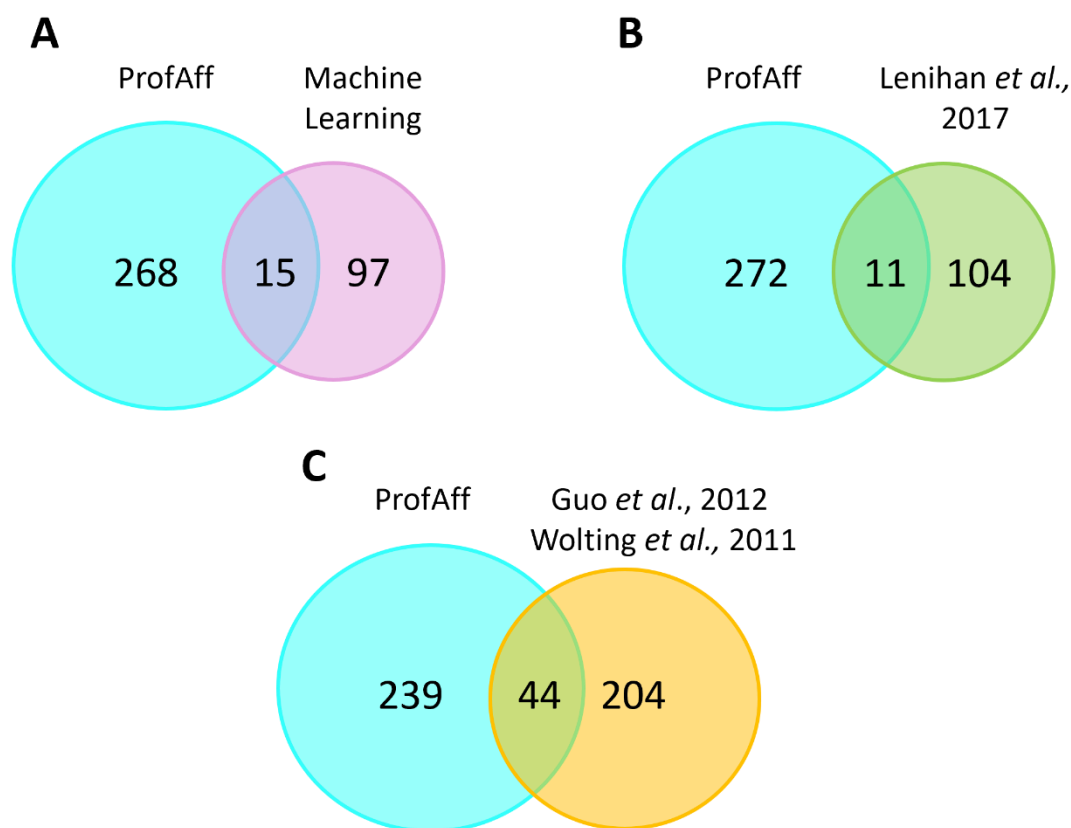


Figure S1: Similarity of known or predicted LNX1 and LNX2 interacting proteins

between various sources. ProfAff database determines there are 283 unique LNX1 and LNX2 interacting PBMs. **(A)** These peptide interactors from ProfAff were compared to machine learning algorithm predicted interactors, carried out by Lee Kavanagh MSc. 15 interactors were common to both sets. **(B)** Lenihan *et al.*, 2017 experimentally validated a total of 115 interactions with LNX1 or LNX2 or both. ProfAff databased determined 11 of these were common. **(C)** The Wolting and Guo research groups validated a total of 248 proteins which could interact with LNX1/2 by protein arrays or yeast-two hybrid screens of individual LNX PDZ domains, of which 44 were also determined by ProfAff's hold-up assay method.

Chapter 4

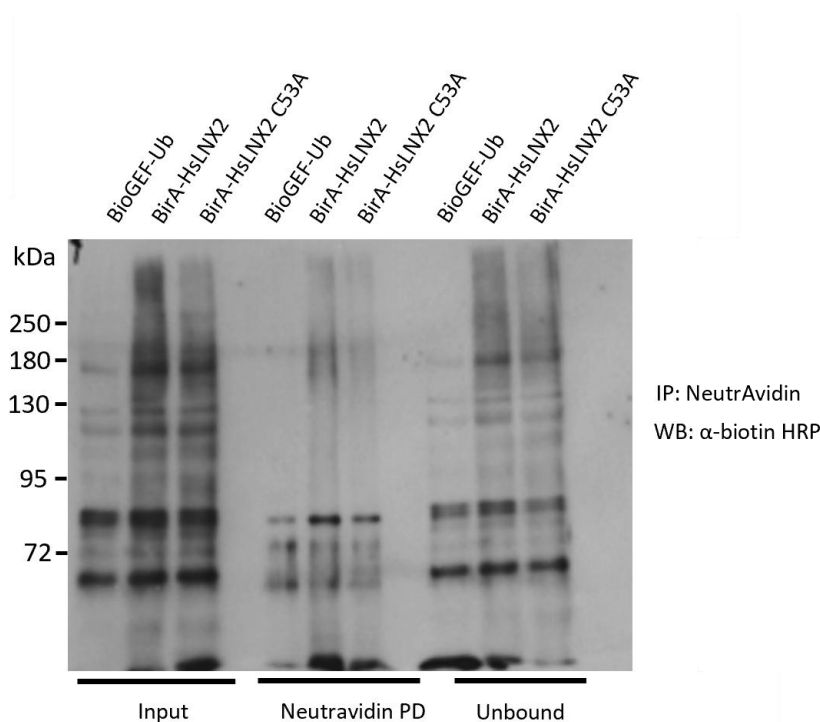


Figure S2: Small scale trial pull-down of biotinylated LNX2 substrates in BioE3 buffers.

Transfected HEK 293T cells (with indicated plasmids) were lysed in WB5 (8 M urea, 1% SDS, 1x PI, 50 μ M NEM in 1x PBS) and 20 μ L of NeutrAvidin beads were added and allowed to interact overnight at room temperature. An unbound sample was taken before removing the supernatant and washing with WB2 (6 M Guanidine-HCl), WB3 (6.4 M urea, 1 M NaCl, 0.2% SDS), WB4 (4 M urea, 1 M NaCl, 10% isopropanol, 10% ethanol, and 0.2% SDS) and finally WB6 (2% SDS). Samples were eluted by 2 rounds of boiling and cooling in 4X Laemmli buffer supplemented with 15 mM biotin. Blocking and immunoblotting was carried out in 1% BSA in TBST and detected using ECL.

Table S1: GO enrichment of biological processes of 96 LNX2 substrates as determined by ShinyGo V0.85.1 (Ge *et al.*, 2020).

p-value	No. of genes	Genes in pathway	Fold enrichment	GO Biological process	Genes
0.000831	16	979	3.959482465	GO:0000278 Mitotic cell cycle	<i>BABAM1, NCAPG NCAPH TUBG1 ABRAXAS2 SMC4 TMOD3 PPP3CA BABAM2 PTPN11 BRCC3 PKN2 RHOA TTLL12 CDC34 EXOC4</i>
0.000831	14	728	4.659054487	GO:0006886 Intracellular protein transport	<i>SEC62 ZFYVE16 RAB18 KPNA1 NUP85 OS9 USO1 USP9X SQSTM1 PPP3CA PTPN11 COPG1 UBAC2 MYO6</i>
0.000831	30	2888	2.516663781	GO:0009056 Catabolic process	<i>TNRC6A HECTD1 CDC34 UBE2R2 UBE2D2 OS9 DNAJB2 PFKM TRIP12 AZIN1 SQSTM1 FBXO45 UBL4A ILVBL HBS1L NUDT5 PAFAH1B3 SLC7A5 MAP3K7 PAFAH1B2 CNP GNAI3 TOM1 EIF4G2 USP9X RNF122 UBAC2 SUPT5H TAB2 EXOC4</i>
0.000831	24	2047	2.84049829	GO:0033554 Cellular response to stress	<i>RFC2 BABAM1 MSH6 WRNIP1 RFC3 OS9 DNAJB2 TRIP12 BABAM2 RFC4 BRCC3 RBBP7 USP9X MAP3K7 FBXO45 MYO6 TNRC6A SLC7A5 PTPN11 SQSTM1 UBAC2 TIPRL RTN4</i>
0.000831	18	1159	3.762618637	GO:0043687 Post-translational protein modification	<i>DCUN1D1 BRAP HECTD1 CDC34 UBE2R2 USP9X UBE2D2 TRIP12 ABRAXAS2 KLHL9 BABAM2 DNAJB2 SQSTM1 FBXO45 BRCC3 MINDY3 RNF122 OS9</i>
0.000831	18	1102	3.957236842	GO:0070647 Protein modification by small protein conjugation or removal	<i>DCUN1D1 BRAP HECTD1 CDC34 UBE2R2 USP9X UBE2D2 TRIP12 ABRAXAS2 KLHL9 BABAM2 DNAJB2 SQSTM1 FBXO45 BRCC3 MINDY3 RNF122 OS9</i>
0.000831	6	90	16.15138889	GO:0070972 Protein localization to endoplasmic reticulum	<i>SEC62 UBL4A SRP68 RTN4 UBAC2 OS9</i>
0.000906	35	3988	2.126248537	GO:0009893 Positive regulation of metabolic process	<i>TNRC6A HEATR1 NUP85 DNAJB2 SMAD2 RFC2 TAB2 RHOA TAB1 RPS19 RFC3 MAP3K7 PPP3CA SQSTM1 RFC4 ZNF24 SUPT5H DCUN1D1 KARS1 BABAM1 EIF4G2 BABAM2 PTPN11 BRCC3 GNAI3 FERMT2 TOM1 SLC7A5 PFKM PAFAH1B2 HECTD1 RTN4 RBBP7 RBBP4</i>
0.000906	10	394	6.149005922	GO:0010506 Regulation of autophagy	<i>SQSTM1 SLC7A5 MAP3K7 GNAI3 TOM1 EIF4G2 PAFAH1B2 SUPT5H TAB2 EXOC4</i>
0.000906	19	1394	3.302113223	GO:0022402 Cell cycle process	<i>BABAM1 NCAPG NCAPH TUBG1 ABRAXAS2 KLHL9 RHOA SMC4 USP9X TMOD3 PPP3CA BABAM2 PTPN11 BRCC3 PKN2 TIPRL MSH6 CDC34 EXOC4</i>
0.000906	13	694	4.538214457	GO:0051301 Cell division	<i>KLHL9 GNAI3 RHOA PKN2 BABAM1 NCAPG SMC4 NCAPH USP9X BABAM2 ABRAXAS2 BRCC3 EXOC4</i>
0.000906	3	9	80.75694444	GO:1900262 Regulation of dna-directed dna polymerase activity	<i>RFC2 RFC3 RFC4</i>

0.000906	3	9	80.75694444	GO:1900264 Positive regulation of dna-directed dna polymerase activity	<i>RFC2 RFC3 RFC4</i>
0.000906	3	8	90.8515625	GO:1905821 Positive regulation of chromosome condensation	<i>NCAPG SMC4 NCAPH</i>
0.001537	4	33	29.36616162	GO:0070536 Protein k63-linked deubiquitination	<i>ABRAXAS2 BABAM2 USP9X BRCC3</i>
0.001856	27	2767	2.364044995	GO:0008104 Protein localization	<i>SEC62 ZFYVE16 RAB18 UBL4A KPNA1 NUP85 FLOT2 OS9 USO1 SRP68 COPG1 ATP1B3 FERMT2 ANKRD13A RTN4 USP9X UBAC2 SQSTM1 TOM1 EXOC4 PPP3CA PFKM SMAD2 PTPN11 MYO6 DNAJB2 SLC12A2</i>
0.001856	3	12	60.56770833	GO:0060623 Regulation of chromosome condensation	<i>NCAPG SMC4 NCAPH</i>
0.001856	27	2779	2.353836812	GO:0070727 Cellular macromolecule localization	<i>SEC62 ZFYVE16 RAB18 UBL4A KPNA1 NUP85 FLOT2 OS9 USO1 SRP68 COPG1 ATP1B3 FERMT2 ANKRD13A RTN4 USP9X UBAC2 SQSTM1 TOM1 EXOC4 PPP3CA PFKM SMAD2 PTPN11 MYO6 DNAJB2 SLC12A2</i>
0.001873	19	1549	2.971688724	GO:0015031 Protein transport	<i>SEC62 ZFYVE16 RAB18 KPNA1 NUP85 OS9 USO1 COPG1 USP9X SQSTM1 TOM1 EXOC4 PPP3CA PFKM SMAD2 PTPN11 MYO6 UBAC2 SLC12A2</i>
0.001873	14	887	3.823891394	GO:0016567 Protein ubiquitination	<i>BRAP HECTD1 CDC34 UBE2R2 UBE2D2 TRIP12 KLHL9 DNAJB2 SQSTM1 FBXO45 DCUN1D1 RNF122 OS9 USP9X</i>
0.001966	15	1023	3.55235826	GO:0006974 Dna damage response	<i>RFC2 BABAM1 MSH6 WRNIP1 RFC3 TRIP12 BABAM2 RFC4 BRCC3 USP9X FBXO45 MYO6 PTPN11 TIPRL</i>
0.001966	21	1860	2.73531586	GO:0045184 Establishment of protein localization	<i>SEC62 ZFYVE16 RAB18 UBL4A KPNA1 NUP85 OS9 USO1 SRP68 COPG1 USP9X SQSTM1 TOM1 EXOC4 PPP3CA PFKM SMAD2 PTPN11 MYO6 UBAC2 SLC12A2</i>
0.002455	6	139	10.45773381	GO:0016579 Protein deubiquitination	<i>USP9X ABRAXAS2 BABAM2 DNAJB2 BRCC3 MINDY3</i>
0.003063	13	819	3.845568783	GO:1903047 Mitotic cell cycle process	<i>BABAM1 NCAPG NCAPH TUBG1 ABRAXAS2 SMC4 TMOD3 PPP3CA BABAM2 BRCC3 RHOA CDC34 EXOC4</i>
0.003155	7	218	7.779338685	GO:0007179 Transforming growth factor beta receptor signaling pathway	<i>SMAD2 TAB1 FERMT2 USP9X RBBP7 RBBP4 MAP3K7</i>
0.003359	15	1089	3.33706382	GO:0006259 Dna metabolic process	<i>RFC2 BABAM1 MSH6 WRNIP1 RFC3 BABAM2 RFC4 BRCC3 USP9X RBBP7 TRIP12 RBBP4 CDC34 KPNA1</i>
0.003359	10	496	4.884492608	GO:0031329 Regulation of cellular catabolic process	<i>SQSTM1 SLC7A5 MAP3K7 GNAI3 TOM1 EIF4G2 PAFAH1B2 SUPT5H TAB2 EXOC4</i>

0.003366	31	3659	2.052581534	GO:0010604 Positive regulation of macromolecule metabolic process	<i>TNRC6A HEATR1 NUP85 DNAJB2 SMAD2 RFC2 TAB2 RHOA TAB1 RFC3 MAP3K7 PPP3CA RFC4 ZNF24 SUPT5H DCUN1D1 KARS1 BABAM1 EIF4G2 BABAM2 PTPN11 BRCC3 FERMT2 SLC7A5 PFKM HECTD1 RTN4 RBBP7 RBBP4 SQSTM1</i>
0.003366	28	3125	2.170746667	GO:0036211 Protein modification process	<i>DCUN1D1 BRAP HECTD1 CDC34 UBE2R2 USP9X UBE2D2 TRIP12 ABRAXAS2 KLHL9 BABAM2 TAB2 PKN2 RHOA TAB1 MAP3K7 DNAJB2 PPP3CA SQSTM1 FBXO45 PTPN11 BRCC3 TTLL12 MINDY3 FERMT2 RNF122 OS9 UBL4A</i>
0.003392	12	726	4.004476584	GO:0006511 Ubiquitin-dependent protein catabolic process	<i>HECTD1 CDC34 UBE2R2 UBE2D2 DNAJB2 TRIP12 FBXO45 UBL4A USP9X RNF122 OS9 SQSTM1</i>
0.003392	14	973	3.485911271	GO:0032446 Protein modification by small protein conjugation	<i>DCUN1D1 BRAP HECTD1 CDC34 UBE2R2 UBE2D2 TRIP12 KLHL9 DNAJB2 SQSTM1 FBXO45 RNF122 OS9 USP9X</i>
0.003398	5	95	12.75109649	GO:0016239 Positive regulation of macroautophagy	<i>GNAI3 TOM1 PAFAH1B2 SUPT5H MAP3K7</i>
0.003398	12	736	3.950067935	GO:0019941 Modification-dependent protein catabolic process	<i>HECTD1 CDC34 UBE2R2 UBE2D2 DNAJB2 TRIP12 FBXO45 UBL4A USP9X RNF122 OS9 SQSTM1</i>
0.003398	29	3331	2.109232713	GO:0043412 Macromolecule modification	<i>DCUN1D1 BRAP HECTD1 CDC34 UBE2R2 NOP2 USP9X UBE2D2 TRIP12 ABRAXAS2 KLHL9 BABAM2 TAB2 PKN2 RHOA TAB1 MAP3K7 DNAJB2 PPP3CA SQSTM1 FBXO45 PTPN11 BRCC3 TTLL12 MINDY3 FERMT2 RNF122 OS9 UBL4A</i>
0.003398	6	158	9.200158228	GO:0070646 Protein modification by small protein removal	<i>USP9X ABRAXAS2 BABAM2 DNAJB2 BRCC3 MINDY3</i>
0.003452	12	741	3.923414305	GO:0043632 Modification-dependent macromolecule catabolic process	<i>HECTD1 CDC34 UBE2R2 UBE2D2 DNAJB2 TRIP12 FBXO45 UBL4A USP9X RNF122 OS9 SQSTM1</i>
0.003452	4	50	19.38166667	GO:0090329 Regulation of dna-templated dna replication	<i>RFC2 WRNIP1 RFC3 RFC4</i>
0.003772	8	325	5.963589744	GO:0051347 Positive regulation of transferase activity	<i>RFC2 TAB2 RHOA TAB1 RFC3 MAP3K7 RFC4 FERMT2</i>
0.003803	3	19	38.25328947	GO:0007076 Mitotic chromosome condensation	<i>NCAPG NCAPH SMC4</i>

Chapter 5

1 *A.queenslandica*/1-700
2 *S.clava*/1-898 MEYSYQAPDCLPGQNDLIDFSSDSSVGVHNQEEGEIESSSLTKNRNQINRDVVAVKSGTGDIGFVAVGRPTTS
3 *S.japonicum*/1-1060 - - - - - MQTGSC - - - - - D DSTG SVGGD - - - - - SETNTTV - - - - - ASSPGFGDPPVTKSS
4 *A.anguilla*/1-698
5 *D.erio*/1-737
6 *X.tropicalis*/1-685
7 *H.sapiens*/1-690
8 *M.musculus*/1-687
9 *A.amphitrite*/1-680
10 *E.carolleeae*/1-860
11 *C.secundus*/1-744
12 *M.quadrilineatus*/1-616
13 *H.americanus*/1-846
14 *A.nasatum*/1-726
15 *P.heterotremus*/1-876 - - - - - MNSS - - - - - MMASTTSAS
16 *M.lignano*/1-799 - - - - - MA P - - - - -
17 *B.floridae*/1-652
18 *P.flava*/1-601
19 *A.japonica*/1-596
20 *A.rubens*/1-598
21 *S.purpuratus*/1-573

1
 2 ENPITKQKHKRTRSNIEKEGYHSDSSPRCPNGGIDNLGYQTISRNESYSYDDTGLSYCSELERWGNTEFL
 3 PLFPEPPEDIAATQP-----ATYDNL-----ATFKKGVCKMENVICSSLSNTTTIAICDSSTPKSSSKVACD
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15 LLPLSMPAGT-----SAFRAI-----HTSGDSL-----SSSATANCTLTFP-----APPPECLLLTENSK-----
 16 PPPQTQPPQQTGTATAAASFWKSQ-----KFAAAI-----DTFEQGGDDADFP-----PPPPCATAGGVKSSSAAGV
 17
 18
 19
 20
 21

150 160 170 180 190 200 210 220

1 MAA NKLARSWSYSHG TIDG TKT

2 NTNLSSSQPEYVSVEGERYTTID YSNATI QHVTTNQPR . NDRGGQFVELPETSGQRTQISDTNRS

3 SSTPKSSKVACDASVSNLFSNPPSF . EPPNLRYSYSENISPGLLNCLNLPATSTPSSFPSPRSAFTELHRKVD PCNKF

4

5

6

7

8

9

10

11

12

13

14

15 LLTENS NEFLVREKLF . QRQPPSAAKSDGSGEG SQLNHTTQIHSRL

16 AGGVKSSSAAGVNQQQALFPSPTAF . SLEPTPPPPP PPPP

17

18

19

20

21

230 240 250 260 270 280 290 300 310

```

1 SKTGGAI PGPQEEGERKERPSYKPGRATIPH...SRTFYFPSAKSSPLQVCDACGQCHPPTKNHVVYDYA
2 SRQGRQTPSSDRES...SYSSHSSRGNSPVENEPESTEMRTSSVENNVNETHKSTSPPVFKSPPLCKCKCGQNHGVTENHEYNVY
3 V-QPSIVPLMSP-K...SISKECI PPHHPVSTSVRSQ...NEVSSCNPRNIVIKKYLCRS CGQSHDPTSPHIYNYT
4 .....MT...EARAP...SAPAGLSWAQVCKECGQGGGGQEGHLYEYL
5 .....MGS-LGEGVDP...GMVSEALLEALCSECGQIHRTWENHLYNYR
6 .....MTA-TTEELVLM...DDTSLNPLCFECGQQHWTRFNHLYNYQ
7 .....MGT-TSDEMVSVEQTS-SSSLNPLCFECGQQHWTRFNHLYNYQ
8 .....MGT-TSDEMVPVEQASSTSSLDPLCFECGQQHWARENHLYNYQ
9 .....MDTLRSWGRPCRTCGQHHARHEPHVYDYV
10 .....MEVRSSGRKSVTRRRVCRNCGQHHSQHEPHLYDYS
11 .....MEKQGHSGSDGRNGSGAGGGIGGSSGGTSVCRSCGQSHAAHEPHVYDYV
12 .....MEGVCRTCGQVHGGGDPHLYDYA
13 .....MENNE...DVG...GSVRSGRRRNCRTCGQHQHPHEPHLYDYI
14 .....
15 .....SGPPTLTGTHEHDCCLKQSHLVLEQTSNVFSNTGTVPGDSEVLWCGACGVIHELSSAHKYKYC
16 .....PPPPQAPTSAAVRFDQRSSGGGGAGGSLSSL-TGGSGRRLELCRACGQFHEPGGAHLYDYN
17 .....MSVRKSEACRSCGQVHEGNGNHLYDYV
18 .....MSARSRKQVLLCAVCGQLHEMGTNHLFDY
19 .....MSCFGVMAKFNQTVLCQTCGQVHDIGQNHLYDY
20 .....MSTRSRHRAELCKTCGQVHEAHGNHLYDYC
21 .....MTESSTPAPAPRNPIILCRTCGQRHQLHNNHVVYDYC

```

RING domain

310 320 330 340 350 360 370 380 390

```

1 EVVDEDFLCAMCHKPLIEPVETVCGHTFCYKCLFGRI RTS PYCPVENHPISRIDGDIQQTSFLVRKLLDKMKVICPNTAYQDLV
2 DEVEDLMDICLQPLVDTPYDTKCGHTFCSICINLYLRVKHICPVDRREELNTATDCWQSSILRLKLVNKLTVCPHEEYQDKK
3 QTVADALCCTLCRQPLVDPLDTKCGHTFCSICLNHLAVQALCPVDRQIIN-YLECQQAENIVKRLDKLLVACPNSPACDSI
4 DEVDDLELCHICLQPLLRPMDFPCGHTYCFQCLSSFLREQDFCPVDRQLQ-LHQCRSSLLVRNLLDKLTCLPFHAHQQLQ
5 LEVDDDLVCHICLQPLVQPLDTPCGHTFCARCLRSFLQERDFCPLDRAHLQ-LQVCRSSILVHKLLDKLSTVTCPLTPSCSLS
6 NEVDDDLVCHICLQPLVQPLDTPCGHTFCFKCLRNFLQEKDFCPDRKRLH-FKLCKKSSILVHKLLDKLVVSCPYSAACKET
7 NEVDDDLVCHICLQPLVQPLDTPCGHTFCYKCLRNFLQEKDFCPDRKRLH-FKLCKKSSILVHKLLDKLVLCPPSSVQKDV
8 GEVDDDLVCHICLQPLVQPLDTPCGHTFCFKCLRNFLQEKDFCPDRKRLH-FKLCKKSSILVHKLLDKLVLCPPSPVQGDV
9 DEVEDLTCICLQPLVGLDTPCGHTFCGACLVFYLNSHQNCPCRAALT-TQACRNSPLVLQKILEKLLIKCPNADVCKET
10 EDVDEDLMDICLQPLVNPDTTPCGHTFCQACMTAYLRVNCPCICRKPLN-QLTCSPLNVLKRLDKLPPVCPNSDSQSET
11 EEIDEDLCICLQPLVFTPLDTPCGHTFCGACSTFLRVQQLCPIDRKPLS-ANLCFSSIVLRKLLKLVQCPNTTIDDKI
12 QDVDEDLMDICLQPLVCPVDLLCGHTFCLPLHKLQVQKLCPIDRKPLD-LKESLPSLVLKKLIDKLLVKCPPSADCGQT
13 DDVDEEDVTICLQPLVTPDTPCGHTFCRPLLSFLTQVQNCPCIDRKPLS-QQLCSPSSVLKLLKLVKRCPNVDHGTME
14 .....MGLVTPVDTPCGHTFCRPLLSFLTQVQNCPCIDRKPLS-QQLCSPSSVLKLLKLVKRCPNVDHGTME
15 DIVDIDLCLCRICKQPLVDPLDTKCGHTFCSSCLNHLAVQALCPEDKQIIN-YLECQQAENIVKRLDKLVLVCPNSEHNEV
16 QPIDDDLCLCRICKQPLVDPLDTKCGHTFCTCLRVHLARQALCPIDKSIIN-FLECQQAENIVKRLDKLVLVCPNVQGEDV
17 NEVDEELTCICLQPLVNPDTTPCGHTFCARCLRSFLQERDFCPDRKRLH-FKLCKKSSILVHKLLDKLVVSCPYSAACKET
18 EAVDEELVQVCLQPLVNPMDTTPCGHTYCAKCLSVLRVHRCQVDRSPLT-SKQCKQASIMVQKLLDKLVLCPPSTYQTDV
19 EAVDEELTCICLQPLVNPDTTPCGHTYCAKCLSVLRVHRCQVDRSPLT-AQDCRPSLIKRLDKLVLVCPNTTYGDDA
20 ESVDEELTCICLQPLVSPDTTPCGHTFCIKCIRSTVRIHKMCPIDRKPLT-SQDCRPSFIVKRLDKLVLVCPNTSYGAT
21 DIVDEELTCICLQPLVNPMDTTPCGHTFCMRCLSVLNLHKKFCPIDRKPLT-QQDCRASSLIVRRMLDLRLVLCNPTDFGTSK

```

400 410 420 430 440 450 460 470

```

1 MMRSELEPHLKRTCPGTYSCLPRR...ERGCL...HL...SPRCQLMEH...LWSCMYGEDCIMK...
2 IGRCDLKGHLKEICPGAVSRCPKS...MYGCD...HK...GPRAGIEAH...LKICAFKDNNDK...
3 VYRSNLEHLREWCVG VHSSAIKPNRSRASQHNQFKNIFLSTFIGTEPVYSNM...LQPIV...TDSTIY...PEIMNS
4 MRCLELQPHLHNRCPAFRKLREEAEKRKRPAWNEKGRAE...GEPGADAKH...P...ATQS
5 MPRCDLEAHLKHRCPGTQSQRTKLERTQMEGSE...DRVT...ATN...PPKSPQTEIKE-DTPS-PPAGSNA
6 MRCDEMEAHLKHRCPGASHRRVALLERRKTSRTQ...AEIE...GENGSVLIDHOGFVSPESDKTR-SGSA-LPERNVT
7 MRCDELEAHLKHRCPGASHRRVALLERRKTSRTQ...AEIE...NENGPTLLDPAGTLSPEADCLG-TGAV-PVERHLT
8 MRCDELEAHLKHRCPGASHRRVALLERRKTSRTQ...TQIE...GETGSTVIDPPGTLPPEPDCSG-TV-PGERNLT
9 AHPRGDLEDHLNRRCQANFIACRYA...AHGCT...VR...GPQKAVHKH...QADCECRNNGI...
10 NORGLEDHLKYRCCEGTLVACQYA...GAGCE...HR...GPAKKMKDH...QTDQCFKKEGI...
11 LQRGDLEDHLKRCPGAGAVAAVPPRNSSTPP...VS...SQRQSSPSK...HGEVTSRNGKI...
12 VQRGDLEDHLKRCPGAVLKSPLRVQPV...SPPNS...VSVDSSPKIGI...
13 MRCDELEAHLKYRCPGNWSCHHT...TSQCD...YR...GSQRSITTH...LTTQTYKDHGV...
14 LPRGDLEDHLKYS...
15 LARCDLESHLAYWCRGAVIACVNH...NLGCT...YR...GLRAHQPAH...RWTGCFQLSATRATTTSDRSCD
16 LPRCDLESHLLHRCRGAVTRCAKA...HLGCT...YQ...GPRSAHSH...LLWECAYRTEGEVPLLP
17 MRCDELEAHLIRCKGTTVRCLRA...PAGCR...FK...GTPTALKEH...MQKQTYKHTISIDGTK-NRLYR
18 MRCDELEAHLNHRCPGTMTYCVNA...PVGCR...FR...GPRVELDTH...IWECFYRDSQVV...
19 MRCDELEAHLANRCKGNMSHCIIHA...AHGCT...FR...GSRTALES...IRECPHAKQK...
20 MRCLEPHLTSRCPGNAAHCSKE...AQGCT...WH...GPAEAKETH...VWECFYRSPV...
21 LPRSELEPHLQNRCPGRI...V...

```

480 490 500 510 520 530 540 550

1 YN KK
2 LS D
3 SQQVMKNNLI AHQRRANF HSSDNESDFT - - VVYPNQYMKEF SHIYPSVTQIEMNK - FS - KTNLNITL PSVEDSSDS
4 SRTAAPRGPAEPGLVNPFAFESEDEAPLR
5 ACNNVPTWTEEPGIDNPFAFEESTEEDSVHGLECALPRVK - - RPLSNPCIHLLRTGSSASSGWDVFETL PLSAEEGCVKL
6 PPSVMPAWMEEHGFNDNPAFEESTVADATH PP
7 PSASLSTWSEEPGLDNPFAFEESAGADTTQ QP
8 PASLPVWTEEPGLDNPFAFEESAAADSVQ QP
9 TR
10 TR
11 TR
12 SR
13 SR
14 R
15 GGRASLEWNLS SPSPRPKSSS - NINTNS - ELRII - GN - - SS PADVCRDGSRSTSR SATSWSGN PP
16 PEPDPTGGSAG KQ
17 RARGHLR LA - QG - - DDSNCASIRDMRQVRT PVTCITAI PR
18 P EP
19 G GV
20 P VG
21 G P

560 570 580 590 600 610 620 630

1 KRQIKEGVPTIVVEIPKGT DQLGIGIKKCGRS...GSIVIDDIIYVTGNVYKDGRI LFGDEIIEID DVLF RHKLLDQVRIA
2 --LHPFSEGEIVVELFRPEGTAEGLGLLVGGSDTPLHYIIVQDILPQS LAFQDGRIRHVGDII MEINGEPMFSVMHAEARAA
3 SPAIYEGVVPVSLIRPAYCMLDGLGTFVGGIDTPLTCLIQEIIYLDGLVATDGRRLRPGDQLLEVNGNLLTHVTHSEARSL
4 --SSLVAEASVVEIYRDEPEEELGVRIVGGKDTPLGNIVIQEVLRDSLAAARDGKLAPGDHILEVNDVNVACVPHRRAI AV
5 LPSSLPEGEITTTIEVHRNTNPYSEMGISIVGGNETPLINVVIIQEVYRDGVIARDGRLLAGDQILQVNVVDISNVPHNFARST
6 PPSLPEGEITTTIEHRSNPYIELGISIVGGNETPLINVVIIQEVYRDGVIARDGRLLAGDQILQVNVVDISNVPHNFARST
7 PLSLPEGEITTTIEHRSNPYIQLGISIVGGNETPLINIVIQEVYRDGVIARDGRLLAGDQILQVNNYNISNVSHNYARAV
8 PLSLPEGEITTTIEHRSNPYIQLGISIVGGNETPLINIVIQEVYRDGVIARDGRLLAGDQILQVNNYD ISNVSHNHARAV
9 RLPPIEGEIVSHIEIPR--TATNLGMTIVGGADTPLRCVVIIQEVFNDSLVAQDGRLLQAGDQIIEVNGVDMTNASHKMICQA
10 RLPPIEGEIVSHIEIPR--TSSCLGITIVGGSDTPLRCVVVQEVFPDGLISGDGRLLQPGDQILEINGVDMTCASHAQVCA
11 RLPVLEGEVSHIEIPR--VATNLGITIVGGSDTALRCVVVQEVFPEGLVSGDGRLLRPGDQILEINGIDMTTASHHQTCA
12 RLPVLEGEVSHIEIPR--ISPSLGITIVGGSDTALRCVVVQEVFPDGLVAQDGRLLKPGDQILEINGVDMSTASHHACQA
13 RNPPIEGEIVTHLEIPR--THTALGITIVGGSDTALRCVVVQEVFPDGLVAQDGRLLQPGDQIIEVGDVMTSATHQVVCQA
14 RNPVLEGEITFIEIPR--TQTALGIALVGGADTPLRCVVIIQEVFPDGLIAHGDGRLLQAGDQIIEEGSDMTNASHHGVCTA
15 R--VVAPSNIASIDLKLD--GRTELGISLVGGWDTPLLCLIIQEIFLDGLAAKDGRLRPGDQILEVNGEDLSQATHLQACH
16 QQQVVSQDVSTIEFGR--GRHDLGVSVFVGGSDTPLSCVIIQEIYMDGLVAIDGRLLRPGDQILEVNGEDVTRASHSEARRL
17 RLPVLEGEITTTIEVRKSSQSEHLGISVGGCETPLVMIVQEVFSSGAIATDGRIRPGDQILEVNGEDLRELTHHQAACRV
18 PTLQILEGVVTTIEINR--TEEELGISLVGGSDTPLVAIVVQEIFPEGLIAKDGRLNPGDQILEINGEDLRDVSAAHQAARAA
19 VVAIKTGGLTTIEIRRN--TDEDFGVSLVGGNETPLIVTIQVEIYPEGVIAKDGRLIMPQDQIEVNGKDLRDVSHYTARET
20 GVPKIDGTLTTIEILRR--REEELGISIVGGNETPLIGIVIQEVTAGGIVARDGRIAAGDQIVQVNDVLDKDVSHQVARMARMA
21 GAPIRHGQLTSIEINRK--PNRDLGIYFVGGNETPLIGVVVQEVIPG EVVAEDARILPGDQIVQVNSRLRLDRIPHSIAAREA

640 650 660 670 680 690 700 710

1 LSSIT--KPLIKFTIVRVHDEEYDP-----ADIEIHLVR--DECRPLGIGI
2 LTKP--SNLMRFMLLRVDVNEPDSDSVASLVRKPE-----SSSAYY--EECVGLIK--PPLELGLIRI
3 LTIP--SSVQQLTVFREPVSLNTPHQPP-----QPQQYQEMFYVKLFK--RQGKVLGRIKL
4 LRRP--CPLLRLTVMQEKGFSTRNPRPEHHPPAGTNATPLSLGPAFNNNNSSGPGPQGSGTIVTQVMLVKRDRSEPLGIKL
5 LARP--CATLQLTVLRERRCSARPPAA-----TASPKSGSPASIRITLHKRESSEDLGIKL
6 LSQP--CTVLHLTVLRERRFGSRAHNG--GEGGSQRDSFHTITLNRKSGNEELGIKL
7 LSQP--CNTLHLTVLRERRFGNRAHNN-----SDSNSPREEIFQVALHKRDSGEDLGIKL
8 LSQP--CSTLQLTVLRERRFGSRANSH-----ADGSAPRDVEVFQVLLHKRDSTERDLGIKL
9 LKVQT--PVLQLGVYREKIDPPPPS-----DDPELHTVTLK--MPGQTLGIKL
10 LKKMS--SVRLRGVYRERIEKYGIYNSGSKSGSGS-----GSGTRGSGVGGGGTQSEDLTVTLTER--QGLHLGIRI
11 LRKLS--PVLRLGVYRERIEAYSARSSA-----HYIEIHTVTTLER--QTNQGLGVRL
12 LRKLKSQLRLGVYRERIEAYNGPNDA-----QHV--TIVTVTLER--ESQALGIRL
13 LRRPQ--TVLRLGVYRERIEAQRTLSTSTTPPAPP-----ATTHVPQPGAEAMVVTQK--ESSRLGLKL
14 LRRIS--PVLRLDE-----LVQVTLNR--EAGRLGIKL
15 LSSLVGPVCWLTIIYREQPFAGTSCL-----GPCLPLKFDEKCELI RIDLVK--SEHKRLGIKL
16 LSQG--SQFCRLTVYREACEPDAQQQQQQFLQH-----HHQQQQRRRLQLLERDEVFPI TLK--RSQKTLGIRI
17 LRKP--CPLLRLTILRERVAEERN-----TGAEVQDTIRVELTK--TSGEGLGIRI
18 FSRQ--SSVARLTVLREKVE--DRG-----LPPDREEILKITLTR--VPGVOLGIKI
19 IMRA--SSPIRMTIYREKAE--ERN-----IDLLNVTLRK--VKGKDLGIKI
20 LHRA--RSPVRLTIYREKAE--NTS-----Q-----LEMFKNILTK--VPGVPLGIKI
21 LCSA--ETPLRLTIYREKQE--SKA-----P--VFITITLVQVSLTK--VTGVPGLGIRI

PDZ2

PDZ2

	720	730	740	750	760	770	780	790	
1	HNYKS	...PGIYVYQLVPGSPAYDHGSLRVGDRILEVNYTDVRLASLDKIANLMIESYGIQLQLLVGHRYHPEI	...YDP						
2	SSSETG	...AKGVYIIEIPGGQVAESDGRLLQVGDRLIMKIGNTVEGDSSTPEQAARVIGSCDKEVKILVRHPLVKPTIV	...						
3	VGKKH	...LPGLYVLELVAGCEADIDGRMKKDDRIEINGIDLENGTQEQAAQIINSVSDYVIFKISRNRSDT	...P						
4	IRKSE	...EPGVFILDLLAGGVAAKDGKLRNNDKVLAINGHDLRHGTPESAAQIIQASEVRVNFVVMQAEVPEEGGT	...						
5	VRRTD	...EAGVFILDLLLEGGLAAQDGRLLCSNDRVLAVNEHDLRHGTPELAAQIIQASGERVNLLISRSSKQTMVHT	...						
6	VRRTD	...ESGVFILDLLLEGGLAAQDGRLLHSNDRVLAINGHDLKHGTPELAAQIIQSGSERVNLTISRPGKHQPGNIT	...						
7	VRRTD	...EPGVFILDLLLEGGLAAQDGRLLSSNDRVLAINGHDLKHGTPELAAQIIQASGERVNLTISRPGKPPGNTI	...						
8	VRRTD	...EPGVFILDLLLEGGLAAQDGRLLSNDRVLAINGHDLKHGTPELAAQIIQASGERVNLTISRPGKPPSPNGS	...						
9	VSRRCSAG	DGVYVAEVISGSAASQDSRLRRDRILYINGRDVRHGRLLQACELIQSNS	...TVSLVVSRRSPARG	...					
10	SGSCP	...APGIYIVEILPESITALDARIKFFDRLLSINGTDVRDCKINNEK	...VVLVVARKKTKPAIDEP						
11	SGWWM	...EQGIFILEIIPESPAAKSGKLEPYDRILSVNGEEVVSCLDQASRLIQSNE	...KVTLVVARKRL	...					
12	SGSWM	...ESGIYITEIVPGSPAALTNLCPYDRILYVNDEDVANSQIQVSTLLQSSE	...KVTLVVARKR	...					
13	GGRR	...EPGIFVMEVLEGSVAADGRLLHPRDRILAINGADVRYARLDHASRLIQQRHHSVLSVSRGAATSV	...VD						
14	GGRRS	...ETGIFIMELMAGSVAEQDGRLLQPNDRILSINGGDVRCARLDHAFRLIKQSHGHLISLVS	...						
15	AGKKN	...CLGLYLGLVQDGAFLDGRLLRKDDRIEINSVNLNGTQEQAAQLIADAKDHVTFLVARMTFPQT	...P						
16	VGKKS	...FPGLYVLGLVEGGLAQLDGRLLRRDDRIEINGVDIRGGTQAEAAEIQASADKVSFLVSRRCRPET	...P						
17	AGLQG	...TTGVCVRDLVEGGLAATDGRLLRPDRIEINGQDMRQATSEQAELIQAASVEKVDVVSRRRRPQP	...P						
18	VGKKT	...EPGVYILDIEDGTACRDGRLLCPDDKILEINHQDVRDCTPEMAAAHAIKTSGEKVKFLICRTIRPT	...P						
19	IAKKH	...SAGIYILDLDVHSIAYQDGRLLSIDDRILQINGEDMTEGSPQAASIKHSDKTIVRLAISRLSRST	...P						
20	AAKRN	...ELGVYVVEVLEDSVAWCRRLLKADDLVVEVNGQDLRQATPDRAAHAINDAGDRILKLVILRHARPPEDSQDS	...						
21	AAKQS	...EQGIYIMEVLESSLARRDGRLLKADDRLVLIQNGFNVEDEMPERVARIRDCGDRVEMVVS	...RQVQVTT	...P					
	800	810	820	830	840	850	860	870	
1	ATVSI	SNCSASS	...	ANSSQ	PANPA	...			
2	...	CNEWA	...	HDS	...	IIQ	...		
3	DILRAT	TDSEKL	...	ENTLLAVESLQLSSNNHRQHSES	...	NEINN	NFDN	...	
4	...	STEG	PGRGS	...					
5	...	GSTL	TRDIWS	...	H	...			
6	...	RDLGM	NQ	...	H	...			
7	...	REAGN	HSSS	...	Q	...			
8	...	REAG	AHSS	...					
9	...	PSDDV	PDGLFRRAG	PG	LAVSHARTKSA	PERLSQS	GGGGSVA	...	VPAAP
10	DGVAAE	VGENGSGHES	SQSGY	PGPQSH	PGVQLSHTRSKSA	PVPYSK	GESD	TD	TGSEFRQSFTDDGDEDEYDGGEGRG
11	...	NRNLQSRD	GSALP	VCHGR	TKSAPEALGSSV	...	QPSVALVSGGM	QQTES	INNS
12	...	SSLP	VSHSR	TKSAPESHDS	SDS	...	SPSLSSN	...	
13	...	SSGPP	PEYQH	LFRFP	PSLPVHHWR	TKSAPDRLV	PRL	...	DCHGKDINKSLG
14	...	NTK	DEYNH	LFKKPHPEL	PVSHGR	TKSAPERLVEQNMDRNDGR	DRFNRSF	GNNKDL	IEG
15	DI	IRTAS	GENVSK	...	CVIF	PG	...		
16	DL	IRTAS	GENLAS	...	LGLSAG	...			
17	DL	IK	MLEADT	...	A	...	K	...	
18	DL	IK	SMAQEFS	...	FG	ISN	...		
19	EIM	PSVAQELA	...	A	...	AMVP	...		
20	EAMQ	ALLNEFP	...	S	...	EAI	...		
21	EVIQV	PPAELP	...	T	...	EP	FTL	...	
	880	890	900	910	920	930	940	950	
1	...				QVVN	...		SSP	
2	...				RQKSTNDAST	...			
3	...	RDSSNE	...	SQNF	AWSSLDVQDTNQSTS	...	AST	...	TTHSTAELIHCSKGQ
4	...								KSTNF
5	...								
6	...								
7	...								
8	...								
9	...			DEPSGTDE	...	PD	...		
10	...	GRDLLTSS	DDMLAE	PFYDV	PEGEDYLAI	...	YETIDRKRLEAREFKNKSASMEELET	PTSSLSKSF	NI
11	N	...	IDPMLHSSLSQS	...	QTWQA	KLQQIS	...		PDRNL
12	...		PPLS	...					
13	GLS	ASCESLDDRG	SVTRREGSLLHSNGSLST	PGEPY	SGSLGSLN	...	GVN	...	THYQSNGLG
14	DRS	ASCESLGR	DRSVRTSRVN	RFHSNGSVMSNGKYSNANSLS	SVT	...	DFEG	...	IHRNKSCSLG
15	...								SVTSGF
16	...								
17	...								
18	...								
19	...								
20	...								
21	...								

960 970 980 990 1000 1010 1020 1030

1 NSSDECQKSTETRS.....

2 SMSK.....

3 NLSSTAATT-VDEDL.....R.....LKQMNHSCNNELT.....

4 RRV.....

5 DHI-PPLPSTATPS.....

6 HHT.....

7 HHT.....

8 NHA.....

9 RAAISRQ.....ARGESMDHLSAPDPPSV.....

10 NLETAGAKAE PQVRDKRRHRTKS SRTEHRRSMHGSNPPALYGGAPQISEPLGGDTSRDPLPLPRDSI.....HIRDSS

11 NNV-TKALC-Y.....ERSHSHSCDSLQNEAK.....

12 HKSVSCDSLTPPA.....

13 NQS-SGLASSAVRGG.....GPSTVEE.....EGGPPALP.....PRRHALARTSSDQLDPPDNASDS.....LREKL

14 NKS-SRSLISYSQTL.....NEDDEER.....ENIPVLP.....PRSRQPLGQNHSDHLNDSLSHLSSESYLNESN

15 SSSQQCDSEVSPSTASTSVQSTNSR.....

16 GSA-SCQHLADAAAAASATT.....

17

18

19

20

21

1040 1050 1060 1070 1080 1090 1100 1110

1 EL-V.....QSK.....LNERVMTPTGYHLGEGK

2 STNPYS.....QQKKLSKSSHTKK

3 NLSN.....DSAAANSRTR.....NDDKQLVLMSSGSSCQER

4 PE.....TQYFRRSTY.....VKDPPGGFSSQEK

5 PV.....PSLHLARSST.....QRDLSQCVNCKEKI

6 QVYHHRPSA.....HKDLAQCVCKQEKI

7 PPPYYSRPSS.....HKDLTQCVTQEKI

8 QPPSHSRPGS.....HKDLTRCVTQEKI

9 APFP.....LPAKPGDP.....RRRSKEARRSLRIEGDVLQKK

10 RDIPINIRDPGLPRDL.....PPVPREPK-EQDSNGESMAQFLKKALRIEGAQLTQK

11 PC.....QMPSPPCTRNV.....PLPVIRSSKEVAVEDLALTINRNLOIEGILQHK

12 PMP.....PHHAATINLNRDFSSELQHK

13 KETPNLDVSV.....QSVGSALSGGGPGVDADGSTSRMRGRSSRDSSDMSDLVHGFRRLRIESTQLHQK

14 VANDPNFNDSSSREALDDFDSKESKAGGTALPASGCAEMVDASVVGKTKVRRSRDCTDIADLAIGMRKSLRLEGSQHLQK

15 TRITPTFS.....NRPLEI.....ACTSGVRVCQER

16 GA.....PKQQQL.....P.....QPQYGTVPQER

17 T.....PEDSRPPSFQDKI

18 A.....KGG-HVLGPKHK

19 S.....RNVSAVRVRRER

20 S.....RESSALKTRQEKI

21 L.....RHALAPKVRAEK

1120 1130 1140 1150 1160 1170 1180 1190

1 TNDPELLGLILGGGYHAVNTLTGDRIPLIFVGGGLNNDCAIEIANKVGPQDIIVSINNLSLAKKSLNEALFILESI-KIG

2 SPKENLGLIIVSGNG.....SPRGNLPIFVQDVLSSMSVLGKDKRIIRGDDLSEINSIRLTGLSHTAMNALRES-AQQ

3 SPEEPLGLISAGGRH.....SQRGDTPIYVTNISSECVLGRSKLIQRGDIILEVNGVGLVGLTHHEAVEVLRQISLMQ

4 EPRLSLGITIAGGRD.....SRS-RVPVYITSVQPVGCLHRDGTITGQDVLLSINGVDLTQLTYSEAVSVLKAQ-TAQ

5 EPHESLGMTVAGGRG.....SKSGELPIFVTSVQPHGCLSRDGRIRKQDVLLSINGQDLTYLSHSEAVGTLKSS-ATS

6 EPQESLGMTVAGGRG.....SKSGELPIFVTSVQPHGCLARDSRIRKQDVLLSINGIDLTNLSHSEAVAMLKAS-ATS

7 EPHESLGMTVAGGRG.....SKSGELPIFVTSVPPHGCLARDGRIRKQDVLLNINGIDLTNLSHSEAVAMLKAS-AAS

8 EPHESLGMTVAGGRG.....SKSGELPIFVTSVPPHGCLARDGRIRKQDVLLNINGIDLTNLSHSEAVAMLKAS-AAS

9 VAKESLGLRIGGGIH.....CIEGDTPIYIANINHPGCVGKSKQIKKGDILLSVNGRSLVGLTHAQAVALKST-AEL

10 NNRESLGMRIIGGGIG.....SNEGDTPIYIANIHHPGCVGKSKQIKKGDILLSVNGRSLVGLTHAQAVATLKAT-ISM

11 EPNESLGMRIIGGGLG.....SNEGDTPIYIANIHHPGCVGKTQQILKGDILLSVNGTSLMGLTHAQAVATLKAT-VEN

12 EPNESLGMRIIGGGLG.....SNEGDTPIYIANIHHPGCVGKTQQIWKGDILIGVNGNSLVGLTHGQAVATLKAT-VET

13 GGNESLGMRIIGGGV.....SNEGDTPIYIANINHPGCVGKSKQVKKGDILLSVNGQSLGLTHGQAVALLKAT-AEL

14 DASESLGMRIIGGGVN.....SNEGNIPYIANIHPPGCVGKSKQVKKGDILLSVNGHSLGLTHGTQ.....Q-PEL

15 QSGESLGMVAGGVA.....SQRGDPVPIYVTNLHANGIAAQSGRVFRGDILLAVNEFELLGLSHERAVAALLNARDSC

16 SPGESLGMVAGGVG.....SQRGDPVYVTSLRPDGPAGR-TVTKGDLILLAVQDSKLLGLTHEQAVRLLKSA-SAD

17 NARESGLISVAGGRA.....GTRGDPVVFVTNVQPDGCLGGNQQLKKGDILVSNGTTLVGLTHEAVGVLRAS-ALS

18 GSTESLGMVSVGGK.....SGRGDIPVYVSDIQPHGVLGDGELCRGQDVLLTINGTSLVKLTHAQAVAILKAC-AGS

19 PPSEFLGLISVCGGVK.....GKRGDTPIFISNVLPGGLADRKGHVVRGDILLAVNDTSLNLTHTAQAVTALKSA-Q-G

20 GSKEGLISVSGGRG.....GPRGDPVVFVTNVQADGCVGRQGLKKGDILLTVNGTSLWNMSHTDAVNNLKAN-ANS

21 GNKECMGLISVSGGK.....QRKGDPVPIFITGIQADGCVARHGQLKKGDILLSVNGTSLLDLPHTAEVVKLKEK-ANA

PDZ3

PDZ3

	1200	1210	1220	1230	1240	1250	1260	1270																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
1	SLVRL	EMI	QGD		G	S	Q	Q	N	G	G	L	T	P	D	W	R	F	V	D	F	Y	M	K	E	R	K	S	P	L	E	F	S	L	T	I	S	R	E	S																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
2	REV	T	L	V	A	Y	E	L	E	Y	E	G	K	A	E	D	T																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				

	1280	1290	1300	1310	1320	1330	1340	1350									
1	--SASLGF	SVVGGVGS				AHGDC	PVYVRIASKSIASNDGRLKQDQITM										
2	--STNYGFKIT	TGGFTE				EHGSP	SPIFITHIFKNSSAN--KFLKCGDEIVA										
3	SVFPQLGFS	IVGG	LGHQDRSTSR	PNSNNSDEKLSKTS	PQVSTKVG	HSNMYPS	PIVIKSIVPGLLAHKDGR	LKCGDILILA									
4	--GESWGF	IVGGYEE				THGQ	PPFIKTIVPGTPAHFDGR	LKCGDEIVA									
5	--PGSWGFS	IVGGYEE				NHSNQ	AFFIKTIVLGT	PAYYDGR	LKCGDMIVA								
6	--LGSWGF	IVGGYEE				NHTNQ	PPFIKTIVLGT	PAYYDGR	LKCGDMIIA								
7	--LGSWGF	IVGGYEE				NHTNQ	PPFIKTIVLGT	PAYYDGR	LKCGDMIVA								
8	--LGSWGF	IVGGYEE				NHTNQ	PPFIKTIVLGT	PAYYDGR	LKCGDMMVA								
9	--TGS LGFS	IVGGCDG				GT--	RDPIHVL	FVQNSPA	AVEGR	LKCGDRLLA							
10	--GTSWGF	IVGGADP				GADRT	EPVHVL	FVVRYS	SPAAK	DGK	LKCGDRLLT						
11	--GASLGF	IVGGEDP				VR--	GPEA	IHLF	FVVDSP	AAKHG	KLR	CGDRLLA					
12	--GASLGF	IVGGEDP				TR--	GAQH	IHLF	FVVDSP	AAK	DGK	LRCGDRLLS					
13	--GASLGF	IVGGS DP				QR--	GAEP	IHLV	FVQSSP	AAIDG	KLR	CGDRLLS					
14	--GASLGF	IVGGS DP				QR--	GPEP	IHLV	FVVS	TSPAAIDGR	LRCGDRLLS						
15	--LLGLGF	II GGND	FCPIS	LERSG	NSQKVD					DETTVADEKTV	SPRP	IVVKSVV	PGSP	CFLDGR	IRCGD	LLLS	
16	--TGS LGFS	IVGGRN	SPSES							SSSAS	AGHP	VAVK	SI	VWGG	PAHEDGR	LRCGD	ILIS
17	--SGSLGF	IVGGCDG								IHGKQ	PILVKS	IVPGSS	ADR	DGR	LRCGD	LILS	
18	--SSSLGF	IVGGADY								CHGE	PAIFVKS	VVPNG	PAAKDGR	LRCGD	QILA		
19	--SKSLGF	IVGGSDC								THGNQ	PFVKT	LV	PNGA	AASSG	ILK	CGDM	ILS
20	--SPSLGF	SVVGG	CDG							THGNQ	PIFVKT	LV	PNGA	AAGRS	GLK	CGD	ILK
21	--NASLGF	IVGGNDS								THGAQ	PIYVKT	VVSDS	IAAKS	G	LLK	CGD	ILIE

	50	1360										1370										1380										1390													
1		D	Q	I	T	M	V	N	G	-	V	S	V	A	G	K	T	Q	A	E	A	V	D	M	I	A	V	T	G	D	I	-	T	F	T	I	I	P	N	D	S	-	-	-	
2		D	E	I	V	A	I	A	N	-	R	L	V	K	D	I	P	Y	L	Q	V	L	N	I	L	K	S	V	N	-	C	-	V	S	I	K	L	V	S	W	P	G	S	L	H
3		D	L	I	L	A	V	A	N	-	I	S	M	L	N	I	R	H	S	S	A	V	R	L	L	K	Q	A	S	G	D	V	-	L	K	V	S	W	P	G	T	I	-	-	
4		D	E	I	V	A	V	N	G	-	A	T	T	V	G	M	N	N	S	S	L	I	P	L	L	K	L	Q	K	N	K	V	-	T	L	T	V	V	S	W	P	G	S	L	
5		D	M	I	V	A	V	N	G	-	L	S	T	A	G	M	S	H	S	A	L	V	P	M	L	K	E	Q	R	S	R	V	-	A	L	T	V	V	S	W	P	G	S	L	
6		D	M	I	V	A	V	N	G	-	L	S	T	V	G	M	S	H	S	A	L	V	P	M	L	K	E	Q	R	N	K	V	-	T	L	S	V	I	S	W	P	G	S	L	
7		D	M	I	V	A	V	N	G	-	L	S	T	V	G	M	S	H	S	A	L	V	P	M	L	K	E	Q	R	N	K	V	-	T	L	T	V	I	C	W	P	G	S	L	
8		D	M	M	V	A	V	N	G	-	L	S	T	V	G	M	S	H	S	A	L	V	P	M	L	K	E	Q	R	N	K	V	-	T	L	T	V	I	C	W	P	G	S	L	
9		D	R	L	L	A	V	D	G	-	H	T	L	E	G	V	R	H	A	E	A	V	S	L	L	K	Q	A	R	N	K	V	-	K	L	Q	I	V	S	W	M	G	T	E	
10		D	R	L	L	T	V	D	G	-	H	S	L	E	T	V	T	H	S	T	A	V	T	M	L	K	Q	T	S	T	R	V	-	V	M	E	V	S	W	M	G	T	E		
11		D	R	L	L	A	V	D	G	-	H	S	L	E	N	V	K	H	S	T	A	V	A	M	L	K	Q	T	G	S	K	V	-	T	L	E	V	V	S	W	L	G	T	E	
12		D	R	L	L	S	V	D	G	-	Q	S	L	E	K	V	H	S	K	A	V	S	M	L	K	Q	T	G	T	Q	V	-	V	L	E	I	V	S	W	L	G	T	E		
13		D	R	L	L	S	V	D	G	-	H	S	L	E	G	V	R	H	A	S	A	V	S	L	L	K	Q	A	Q	R	V	-	H	L	E	V	S	W	L	G	T	E			
14		D	R	L	L	S	V	D	N	-	H	C	L	E	G	V	S	H	A	T	A	V	S	L	L	K	Q	A	Q	R	V	-	T	L	E	V	S	W	M	G	T	E			
15		D	L	L	L	S	V	D	Q	-	H	A	L	T	D	V	S	H	A	H	A	V	A	L	L	K	Q	C	T	G	D	V	-	R	L	R	V	V	S	W	P	G	T	I	
16		D	I	L	S	V	N	H	R	H	S	L	A	D	I	G	H	A	Q	A	V	S	L	L	K	L	V	R	T	D	P	V	V	-	L	T	I	V	S	W	P	G	T	I	
17		D	L	I	L	S	A	N	G	-	H	L	M	T	S	I	S	H	T	M	S	V	T	L	L	K	R	L	Q	S	V	-	M	L	A	V	S	W	P	G	T	M	L		
18		D	Q	I	L	A	V	N	G	-	Q	A	L	Q	D	M	T	H	A	V	T	V	A	L	L	K	R	T	Q	G	R	V	-	T	L	S	V	V	H	W	P	G	S	V	H
19		D	M	I	L	S	V	N	G	-	Q	S	L	E	G	V	K	H	S	A	V	S	L	L	K	H	A	P	G	R	S	I	-	T	L	T	V	V	S	W	P	G	S	L	
20		D	Q	I	L	K	V	N	-	E	C	L	T	N	I	N	H	T	K	A	V	S	L	L	K	N	I	P	T	P	G	R	V	-	H	L	T	V	S	W	P	G	T	A	
21		D	I	I	E	S	V	N	S	-	V	S	L	V	D	I	S	H	K	E	A	V	T	L	L	K	N	I	P	Q	R	A	-	Q	F	T	I	V	S	W	P	G	T	I	

PDZ4

PDZ4

PBM

Figure S3: Full LNX protein multiple sequence alignment consisting of twenty-one metazoan species. LNX orthologs from indicated species (numbered 1-21 for space) were aligned by ClustalΩ and amino acids were coloured accordingly. Conserved domains were boxed – first red box shows good conservation of the catalytic RING domain, the second box is the NPAF motif, PDZ domains 1-4 are boxed in black and carboxyl terminal ten amino acids (PBM) is also highlighted in red.

Chapter 6

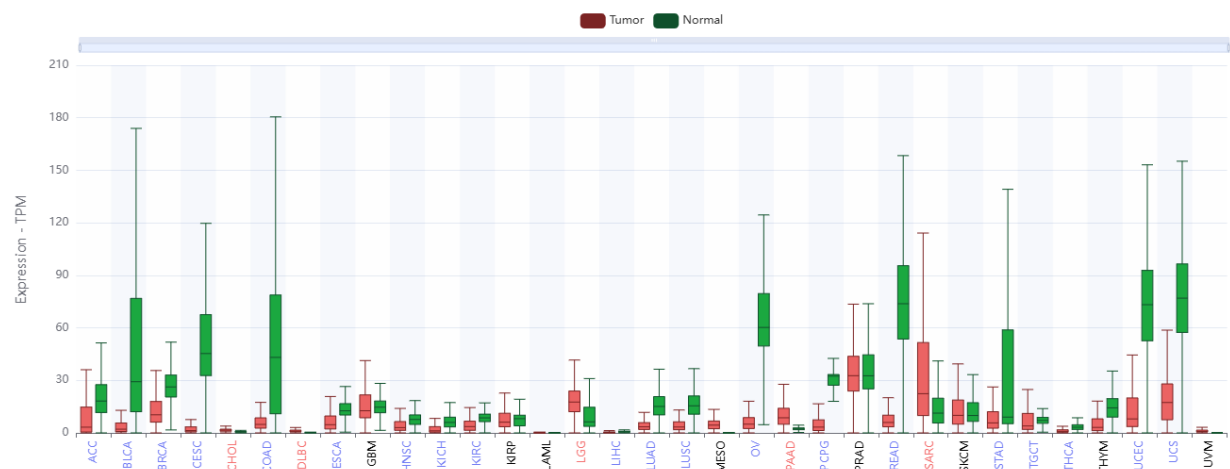


Figure S4: *LNX3/PDZRN3* expression in normal versus tumorigenic tissues from GTEx and TCGA datasets and compiled by GEPIA3. Normal tissue samples analysed were biopsied from adjacent tumorigenic samples RNA expression is expressed in transcripts per million (TPM) as Log₂ (TPM+1) scale. Cancer types in blue indicate the patient sample expression is lower in cancer versus normal tissue, red indicates higher *LNX3* expression in cancer compared to normal tissue and black is no statistical significance. For full statistical analyses and cancer abbreviations see Table S2 (Kang *et al.*, 2025).

Table S2: GEPIA3 statistical analyses of LNX1,2 and 3 mRNA expression in indicated cancer tissues compared to adjacent normal tissues. GEPIA3 carries out t-tests based on median values with a Log₂FC cutoff set to 1 and *q*-value threshold of 0.05 to give indicated *p*-value scores. Nan suggests that a sample size for a particular gene or cancer type was insufficient for statistical analyses (Kang *et al.*, 2025).

Cancer type	Abbrev.	<i>p</i> -value			
		<i>LNX1</i> ^{p70}	<i>LNX1</i> ^{p80}	<i>LNX2</i>	<i>LNX3</i>
Adrenocortical carcinoma	ACC	0.909	0.0409	1.92E-10	2.41E-15
Bladder Urothelial Carcinoma	BLCA	0.823	0.562	0.776	9.01E-30
Breast invasive carcinoma	BRCA	8.86E-10	0.0307	2.19E-50	7.31E-72
Cervical squamous cell carcinoma and endocervical adenocarcinoma	CESC	0.0372	0.118	7.95E-06	6.16E-35
Cholangiocarcinoma	CHOL	0.12	0.0229	0.409	0.0205
Colon adenocarcinoma	COAD	0.0446	3.88E-63	3.72E-82	6.40E-88
Lymphoid Neoplasm Diffuse Large B-cell Lymphoma	DLBC	0.122	6.21E-06	2.21E-18	1.25E-24
Esophageal carcinoma	ESCA	5.98E-20	2.09E-23	2.36E-35	1.55E-31
Glioblastoma multiforme	GBM	9.34E-16	9.58E-15	6.36E-09	0.307
Head and Neck squamous cell carcinoma	HNSC	0.0621	6.90E-11	0.255	2.25E-07
Kidney Chromophobe	KICH	0.158	0.000232	5.79E-05	3.20E-08
Kidney renal clear cell carcinoma	KIRC	7.17E-18	8.80E-31	0.956	5.26E-15
Kidney renal papillary cell carcinoma	KIRP	0.171	7.29E-27	0.286	0.868
Acute Myeloid Leukemia	LAML	1.74E-29	4.60E-50	8.29E-15	0.0336
Brain Lower Grade Glioma	LGG	3.3E-08	3.86E-44	7.75E-14	3.54E-50
Liver hepatocellular carcinoma	LIHC	7.8E-09	2.61E-08	1.03E-05	0.00017
Lung adenocarcinoma	LUAD	6.92E-28	2.53E-77	0.0258	2.31E-150
Lung squamous cell carcinoma	LUSC	4.71E-15	2.55E-92	2.07E-09	6.32E-127
Mesothelioma	MESO	nan	nan	nan	nan
Ovarian serous cystadenocarcinoma	OV	1.49E-22	3.02E-39	1.22E-10	2.60E-108
Pancreatic adenocarcinoma	PAAD	0.00597	4.86E-17	0.00131	4.36E-53
Pheochromocytoma and Paraganglioma	PCPG	0.097	0.624	0.00728	0.000224
Prostate adenocarcinoma	PRAD	0.0503	0.146	0.546	0.35
Rectum adenocarcinoma	READ	0.000281	1.74E-42	2.34E-56	1.76E-64
Sarcoma	SARC	1.46E-49	0.00351	1.18E-06	1.49E-16
Skin Cutaneous Melanoma	SKCM	0.0228	nan	0.223	0.00871

Stomach adenocarcinoma	STAD	3.08E-05	0.00656	7.97E-56	4.37E-19
Testicular Germ Cell Tumors	TGCT	0.787	3.66E-16	2.37E-30	0.00346
Thyroid carcinoma	THCA	0.0194	4.16E-13	0.0337	4.78E-67
Thymoma	THYM	0.847	0.969	0.613	0.233
Uterine Corpus Endometrial Carcinoma	UCEC	0.15	0.00746	8.13E-09	7.36E-49
Uterine Carcinosarcoma	UCS	1.97E-05	0.00156	2.58E-05	4.55E-28
Uveal Melanoma	UVM	nan	nan	nan	nan

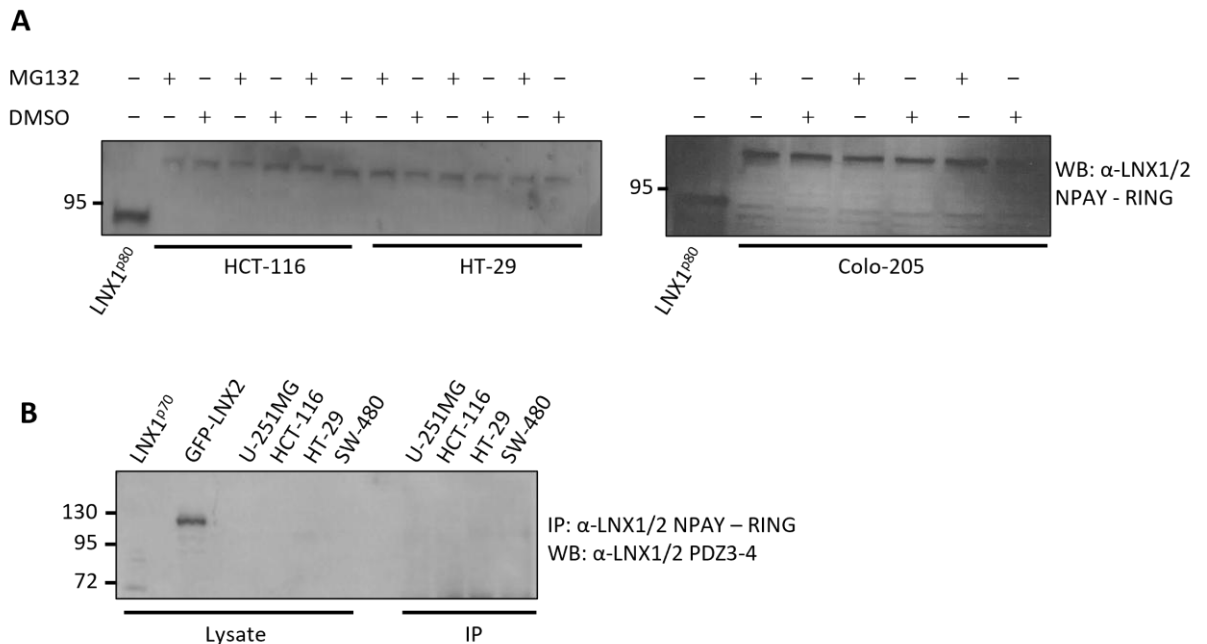


Figure S5: LNX1 or LNX2 proteins cannot be detected in a panel of CRC and GBM cell lines by proteasomal inhibition or immunoprecipitation. **A.** HCT-116, HT-29 and Colo-205 cells were subjected to proteasomal inhibition using 10 μ M of MG132 or vehicle only control (DMSO). HEK 293T cells transfected with LNX1 (LNX2 is smaller) for size comparison and antibody efficiency are loaded also. Only high molecular weight, non-specific bands are present across each cell line tested. **B.** Immunoprecipitation (IP) using 0.5 μ g of α -LNX1/2 NPAY-RING antibody (Methods Section 2.5) was carried out in GBM line, U-251MG and CRC lines, HCT-116, HT-29 and SW-480. The antibody was previously validated in immunoprecipitating LNX1^{p70/p80} and LNX2 from mouse brain lysates (Lenihan *et al.*, 2014). No bands corresponding to any LNX1/2 protein are seen.