



Kent Academic Repository

Cowell, Alana Romain (2020) *Investigating the Structure and Function of the Novel Protein Talin Rod Domain-Containing Protein 1 (TLNRD1)*. Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from

<https://kar.kent.ac.uk/82792/> The University of Kent's Academic Repository KAR

The version of record is available from

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Investigating the Structure and Function of the novel protein Talin Rod Domain- Containing Protein 1 (TLNRD1)

By Alana Romain Cowell

A thesis submitted to the University of Kent for the degree
of Doctor of Philosophy in Biochemistry

The University of Kent
School of Biosciences

August 2020

Declaration

No part of this thesis has been submitted in support of an application for any other degree or other qualification at the University of Kent, or any other University or Institution of learning.

Acknowledgements

First off, I would like to thank Dr Ben Goult for all the support and guidance over the years and providing me with the opportunity to work on such an awesome project, it has been a fantastic journey of discovery and development as a researcher in the field. Thanks also must go to all members of the Goult lab past and present: Rosie Gough, Marie Anderson, Karen Baker, Abhi Singh, Gabriella Laker, Austin Whitewood, Rejina Khan and Lorena Álvarez. Special thanks go to Austin and Rejina for the copious amounts of much needed coffee and cake breaks (the true fuel of science) and pancake movie nights. Thanks also go to Abhi for his training in x-ray crystallography experimental setup and data analysis, as well as his steady hand for picking those fragile protein crystals! Along with this I would like to thank Dave Brown for allowing us to borrow crystallography equipment and helping with reagents.

I'd like to thank Gary Thompson for all the NMR training, help with running all the NMR experiments and NMR analysis, as well as Ian Brown for his help and assistance in preparing samples and imaging with the electron microscope.

A big thank you goes to our collaborators Guillaume Jacquemet and Ilkka Paatero from Johanna Ivaska's group for all the support on the TLNRD1 project, and to Johanna Ivaska for allowing me to have the opportunity to do some cell biology work in her lab for two months.

Finally the biggest thank you goes to my partner Dan, who has always been there to lift me up when full of self-doubt and has supported me throughout this PhD.

Abstract

Talin is a large mechanosensitive linker which couples integrin activation with mechanical cues from the cells cytoskeleton to regulate adhesion complex assembly. Talin is heavily implicated in cancer progression, with alterations in expression or interaction with ligands leading to larger scale changes in cell behaviour. It has multiple ligands which regulate different pathways, coordinating cellular processes in response to the mechanical environment. A large proportion of these ligands interact with the R7R8 region, a 'signalling hub' at the core of the rod domain. Talin rod domain-containing protein 1 (TLNRD1) shares 22% sequence homology with this R7R8 region of talin. Overexpression of TLNRD1 in cancerous cells is associated with reduced patient survival but little is known about the proteins function. The research presented in this thesis utilises a range of structural, biophysical and biochemical techniques to explore the relationship between TLNRD1 and talin R7R8.

X-ray crystallography and small angle x-ray scattering (SAXS) reveals the unique structure of TLNRD1 which constitutively adopts a novel antiparallel dimer but retains the same structural topology of talin R7R8. We went on to show that not only is TLNRD1 structurally similar to talin R7R8, but it has retained the ability to interact with numerous R7R8 ligands. Using fluorescence polarisation and nuclear magnetic resonance (NMR) spectroscopy, we showed that TLNRD1 interacts with RIAM, KANK1 and CDK1. Both RIAM and CDK1 interact with TLNRD1 4-helix domain, the equivalent of R8 in talin, and KANK1 interacts with the 5-helix domain, the equivalent of R7. Furthermore, previous research demonstrated TLNRD1's ability to interact with actin filaments, however, with the new knowledge of TLNRD1's existence as a constitutive dimer we were able to reveal that TLNRD1 is actually an actin-bundling protein. Using actin cosedimentation assays and electron microscopy we show that TLNRD1 forms tight bundles of actin filaments and that this bundling behaviour is primarily mediated by the 4-helix domain of TLNRD1. Finally, to explore the connection between TLNRD1 and cancer survival, confocal microscopy and 2D migration assays were used to analyse GFP labelled TLNRD1 in U2OS cells. This revealed that TLNRD1 promotes filopodia formation, localises to filopodial tips and increases cellular migration. Overall, this thesis describes the discovery of TLNRD1's intriguing structure, identifies a number of novel ligands which allow the protein to behave in a similar manner to talin R7R8 and provides new evidence that TLNRD1 is a novel actin-bundling protein which promotes filopodia formation.

Table of Contents

Chapter 1. Introduction	17
1.1. The extracellular matrix	18
1.2. Cell migration and adhesion	18
1.3. The cytoskeleton	19
1.3.1. Microtubules	19
1.3.2. Intermediate filaments	20
1.3.3. Actin	20
1.3.4. Actin bundling proteins	22
1.4. Actin in cell migration and adhesion	23
1.4.1. Filopodia	24
1.4.2. Lamellipodia	25
1.5. Cell-matrix adhesion	25
1.6. Adhesion complexes	25
1.6.1. Nascent adhesions	26
1.6.2. Focal complexes and focal adhesion	27
1.6.3. Reticular adhesions	28
1.6.4. Fibrillar adhesion	28
1.6.5. Podosomes	28
1.7. Talin	29
1.7.1. Talin evolution and isoforms	30
1.8. Talin-1 ligands	32
1.8.1. Integrins	32
1.8.2. Vinculin	34
1.8.3. Actin	35
1.8.4. MRL family of proteins	35
1.8.5. The KANK family of proteins	37
1.8.6. Cyclin dependent kinase 1	41
1.9. Talin Rod Domain Containing Protein 1 (TLNRD1)	44
1.9.1. TLNRD1 in Cancer	44
1.10. Project aims	46
Chapter 2. Materials and Methods	48
2.1. Chemicals and Reagents	49
2.2. Buffers and Media	49
2.3. Plasmid Constructs- Biochemistry	51
2.4. Plasmid Constructs- Cell Culture	52

2.5. Molecular Biology	53
2.5.1. DNA constructs	53
2.5.2. Calcium chloride treatment for making competent E. coli.....	53
2.5.3. DNA transformations	53
2.5.4. Primer design and Polymerase Chain Reaction	54
2.5.5. Double restriction digest and ligation.....	54
2.5.6. Site-directed mutagenesis	55
2.6. Protein Expression and Purification.....	55
2.6.1. Protein expression	55
2.6.2. Cell lysis.....	56
2.6.3. Protein purification by Ni-NTA column affinity	56
2.6.4. Measuring protein concentration.....	57
2.6.5. SDS-PAGE	57
2.7. Biochemical Characterisation	57
2.7.1. Circular Dichroism (CD).....	57
2.7.2. Oligomerisation analysis with SEC-MALS.....	58
2.7.3. Nano Differential Scanning Fluorimetry (nanoDSF).....	58
2.7.4. Fluorescence Polarisation	59
2.7.5. Microscale thermophoresis (MST).....	60
2.7.6. Nuclear magnetic resonance spectroscopy (NMR).....	61
2.7.7. Actin purification and polymerisation	63
2.7.8. Actin co-sedimentation assay	64
2.7.9. Negative stain transmission electron microscopy	64
2.8. Structural Biology.....	65
2.8.1. Small angle x-ray scattering (SAXS).....	65
2.8.2. X-ray crystallography	67
2.9. Cell Culture	70
2.9.1. Antibodies.....	70
2.9.2. Cell lines and culture maintenance.....	70
2.9.3. Cell plasmid transfections.....	70
2.9.4. siRNA silencing.....	71
2.9.5. Cell lysis.....	71
2.9.6. Immunoprecipitation	71
2.9.7. Western blot	72
2.9.8. Immunofluorescence- fixed confocal microscopy.....	72
2.9.9. 2D Random Migration Assays	73
Chapter 3. Structural Characterisation of TLNRD1	74

3.1. Sequence analysis and secondary structure prediction	75
3.1.1. TLNRD1 shares sequence similarity with talin R7R8 and has a similar predicted secondary structure	75
3.1.2. TLNRD1 has an intriguing evolutionary history	77
3.1.3. TLNRD1 N-terminal unstructured region	80
3.1.4. Predicted TLNRD1 phosphorylation sites	81
3.1.5. Designing TLNRD1 constructs for structural and biochemical experiments.....	82
3.2. Stability and secondary structure analysis.....	84
3.2.1. Circular Dichroism theory	84
3.2.2. Secondary structure and stability analysis of TLNRD1 constructs.....	85
3.3. TLNRD1 crystallisation and structure determination	88
3.3.1. TLNRD1-FL crystallisation.....	88
3.3.2. TLNRD1-4H crystallisation.....	89
3.3.3. Data collection and analysis for TLNRD1 constructs.....	90
3.3.4. X-ray crystal structure of TLNRD1 full-length protein	92
3.3.5. X-ray crystal structure of TLNRD1 4-helix domain.....	94
3.3.6. TLNRD1 is a symmetric antiparallel dimer	95
3.3.7. Designing a TLNRD1 dimerisation knockout mutant	97
3.4. Solution analysis of TLNRD1 dimerisation	98
3.4.1. SEC-MALS analysis of TLNRD1 oligomeric state.....	98
3.4.2. MST analysis confirms TLNRD1 as a high affinity dimer	100
3.5. Small Angle X-ray Scattering (SAXS) of TLNRD1.....	101
3.5.1. SAXS analysis of the TLNRD1 dimer	105
3.5.2. Ab initio shape determination of the TLNRD1 dimer	107
3.6. Discussion.....	109
Chapter 4. Exploring TLNRD1 Interactions with Talin R7R8 Ligands	112
4.1. Overview	113
4.2. Leucine-Aspartic acid motifs.....	113
4.3. Talin interactions with LD motifs	114
4.4. Results section 1: TLNRD1 interacts with RIAM and Lpd via the 4-helix domain	115
4.4.1. Structural comparison of TLNRD1 and the talin R8 RIAM binding site.....	115
4.4.2. Fluorescence polarisation reveals TLNRD1 interaction with RIAM	117
4.4.3. HSQC titration of TLNRD1 4-helix domain with RIAM peptide.....	118
4.4.4. TLNRD1 4-helix domain binds the RIAM paralog Lamellipodin	122
4.4.5. Designing a TLNRD1 mutant to disrupt RIAM binding.....	123
4.5. Results section 2: TLNRD1 interaction with KANK proteins.....	127
4.5.1. Structural comparison of talin R7 KANK binding site with TLNRD1.....	127

4.5.2. TLNRD1 interaction with KANK1 LD motif	130
4.5.3. KANK1-TLNRD1 interaction is mediated by the 5-helix domain	131
4.5.4. TLNRD1 interacts with KANK2	131
4.6. Results section 3: TLNRD1 interacts with newly identified CDK1 LD motif	133
4.6.1. Structural comparison of talin R8 CDK1 binding site with TLNRD1	133
4.6.2. TLNRD1 interacts with CDK1 via the 4-helix domain	134
4.6.3. NMR analysis of TLNRD1-CDK1 interaction	135
4.7. Results section 4: TLNRD1 is an actin bundling protein.....	137
4.7.1. Structural comparison of R7R8 actin binding region and TLNRD1	137
4.7.2. Both TLNRD1 domains can interact with F-actin	138
4.7.3. TLNRD1 2E mutant stops actin filament engagement.....	140
4.7.4. TLNRD1 as an actin bundling protein.....	141
4.7.5. Visual inspection of TLNRD1 bundling activity using electron microscopy	143
4.8. Discussion.....	145
Chapter 5. Investigating TLNRD1 function.....	148
5.1. Introduction	149
5.2. Results.....	149
5.2.1. TLNRD1 zebrafish morpholino leads to developmental defects	149
5.2.2. TLNRD1 localises to the tips of filopodia	153
5.2.3. TLNRD1 promotes filopodia formation.....	155
5.2.4. Identifying TLNRD1 endogenous protein.....	156
5.2.5. TLNRD1 4-helix domain shows increased F-actin localisation	158
5.2.6. TLNRD1 influences 2D random cell migration	161
5.3. Discussion.....	163
Chapter 6. Exploration of cancer associated point mutations in talin	166
6.1. Introduction	167
6.1.1. Talin and its association with cancer	167
6.1.2. Cancer-associated databases.....	167
6.1.3. Chapter overview.....	168
6.2. Identification and evaluation of talin-1 single point mutations	169
6.3. Phenotypic comparison of talin-1 point mutations.....	173
6.3.1. Most talin-1 point mutations have little impact on cell morphology.....	173
6.3.2. Cancer-associated talin-1 point mutations alter cell behaviour.....	174
6.3.3. Cancer-associated talin-1 point mutations change adhesion composition.....	176
6.4. Biochemical analysis of talin-1 cancer associated mutations.....	179
6.4.1. Mutation L1539P disrupts the R8 helical structure	179

6.4.2. R7 domain mutant Y1389C reduces protein thermostability	180
6.4.3. NMR analysis of R1368W and Y1389C.....	181
6.4.4. R7 domain mutant Y1389C influences vinculin binding	183
6.4.5. Binding to KANK and RIAM appears unaffected by R1368W and Y1389C	184
6.4.6. C-terminal L2509P mutation may impair talin dimerisation	187
6.4.7. <i>L2509P leads to minor perturbation of ABS3 actin binding</i>	188
6.5. Discussion.....	189
Chapter 7. Conclusions and future directions	193
7.1. TLNRD1 as a third talin gene.....	194
7.2. Limitations and future directions.....	198
7.1. Pipeline development for assessment of cancer associated mutations.....	200
Chapter 8. References	202

List of Abbreviations

Δ	Delta, gene deletion
$^{\circ}\text{C}$	Degrees Celsius
μl	Microlitre
μM	Micromolar
2M9	Minimal media
ABS	Actin binding site
Amp	Ampicillin
ATP	Adenosine 5'-triphosphate
BSA	Bovine Serum Albumin
CaCl_2	Calcium Chloride
CD	Circular dichroism
CDK1	Cyclin Dependent Kinase 1
C-terminus	Carboxyl terminus
DD	Dimerisation domain
dH ₂ O	distilled water
DLC1	Deleted in Liver Cancer 1
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECM	Extracellular matrix
FA	Focal adhesion
F-actin	Filamentous actin
FAK	Focal adhesion kinase
FERM	4.1 protein, ezrin, radixin, moesin
FP	Fluorescence polarisation
G-actin	Globular actin (monomeric)
GFP	Green fluorescent protein
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum spectroscopy
IP	Immunoprecipitation
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Kan	Kanamycin
KANK	Kidney Ankyrin Repeat-containing protein
kDa	kilodalton
Kdr1	Vascular endothelial growth factor receptor kdr-like
LB	Lysogeny broth

LD	Leucine Aspartate
Lpd	Lamellipodin
MALS	Multi-angle light scattering
MES	2-(N-Morpholino)ethanesulfonic acid hemisodium salt
MESDC1	Mesoderm development candidate 1
MgCl ₂	Magnesium Chloride
miRNA	Micro RNA
mRNA	Messenger RNA
MST	Microscale thermophoresis
MT	Microtubules
NaCl	Sodium Chloride
NaN ₃	Sodium Azide
nanoDSF	Nanoscale differential scanning fluorimetry
NMR	Nuclear magnetic resonance
N-terminus	Amino-terminus
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase Chain Reaction
pI	Isoelectric point
RIAM	Rap1-interacting adapter molecule
SAXS	Small angle x-ray scattering
SDM	Site-directed mutagenesis
SEC	Size exclusion chromatography
sgRNA	Single guide RNA
TEM	Transmission electron microscopy
TEV	Tobacco Etch Virus
Tln1	Talin 1
Tln2	Talin 2
TLNRD1	Talin rod domain containing protein 1
TROSY	Traverse relaxation optimised spectroscopy
VBS	Vinculin binding site
WT	Wild-type

List of Figures

Figure 1.1. Architecture of cell cytoskeleton networks	19
Figure 1.2. Dynamic actin structures in a migrating cell (Blanchoin <i>et al.</i> 2014). 21	
Figure 1.3. Structure of monomeric G-actin	22
Figure 1.4. Schematic of filament bundling, adapted from Winkleman <i>et al.</i> (2016)	23
Figure 1.5. Filopodia architecture, adapted from Jacquemet <i>et al.</i> (2015)	24
Figure 1.6. General principles of cell migration and adhesion formation	26
Figure 1.7. Nanoscale organisation of focal adhesion complexes (Kanchanawong <i>et al.</i> 2010)	27
Figure 1.8. Schematic representations of the talin molecule	29
Figure 1.9. Schematic representation of talin-2 splice variants (Debrand <i>et al.</i> 2009).....	31
Figure 1.10. Schematic of talin interactors.....	32
Figure 1.11. Diagram of all integrin α and β pairings	33
Figure 1.12. Integrin activation states	34
Figure 1.13. Schematic representation of MRL protein composition.....	36
Figure 1.14. KANK Family of Proteins.....	38
Figure 1.15. Schematic of a cortical microtubule stabilising complex (Bouchet <i>et al.</i> 2016)	40
Figure 1.16. KANK isoforms KN domain sequence alignment.....	41
Figure 1.17. Schematic of the cell cycle with CDK and cyclin stages	42
Figure 1.18. CDK1 talin binding region.....	43
Figure 1.19. TLNRD1 overexpression in cancer and association with poor survival	45
Figure 2.1. Schematic of fluorescence polarisation theory	59
Figure 2.2. Representation of peak selection in a TROSY experiment.....	63
Figure 2.3. Representation of a typical setup for SEC-SAXS.....	66
Figure 3.1. Sequence alignment of TLNRD1 with talin-1 and talin-2	76
Figure 3.2. Talin R7R8 and TLNRD1 domain structure comparison	76
Figure 3.3. Evolutionary tree of TLNRD1	79
Figure 3.4. Amino acid sequence alignment of TLNRD1	80
Figure 3.5. TLNRD1 PrDOS disorder prediction	81
Figure 3.6. TLNRD1 construct design and purification	83
Figure 3.7. Circular Dichroism secondary structure spectra for proteins.....	84

Figure 3.8. CD secondary structure and stability analysis	86
Figure 3.9. Thermal stability analysis of TLNRD1-FL	87
Figure 3.10. TLNRD1-FL screen hits and optimisation	89
Figure 3.11. TLNRD1-4H screen hit and optimisation	90
Figure 3.12. TLNRD1 crystal structure reveals an R7R8-like topology and antiparallel dimer	93
Figure 3.13. TLNRD1 4-helix domain crystal structure	94
Figure 3.14. TLNRD1 dimerisation interface.....	95
Figure 3.15. TLNRD1 dimerisation interface is highly conserved	96
Figure 3.16. F250 mutation to aspartic acid.....	97
Figure 3.17. SEC-MALS analysis of TLNRD1 oligomer formation and F250D mutation	99
Figure 3.18. TLNRD1 is a high affinity dimer	100
Figure 3.19. Typical SAXS scattering curves adapted from Mertens and Svergun (2010).....	102
Figure 3.20. Typical Guinier analysis for R_g estimation adapted from Mertens and Svergun (2010).....	102
Figure 3.21. Typical Kratky plot adapted from Mertens and Svergun (2010)	103
Figure 3.22. $P(r)$ distance distribution curves adapted from Mertens and Svergun (2010).....	104
Figure 3.23. TLNRD1 SAXS data.....	105
Figure 3.24. Normalised Kratky plot and SIBYLS flexibility plot.....	106
Figure 3.25. <i>Ab initio</i> shape determination of the TLNRD1 dimer	108
Figure 4.1. LD binding mechanism.....	113
Figure 4.2. LD motif comparison of talin R7R8 ligands	114
Figure 4.3. Comparison of RIAM binding sites between TLNRD1 and talin	116
Figure 4.4. Full-length TLNRD1 binding with RIAM peptide spanning residues 4-30	117
Figure 4.5. TLNRD1 4-helix domain interacts with RIAM peptide spanning residues 4-30.....	118
Figure 4.6. NMR exchange rates and chemical shift patterns	119
Figure 4.7. ^1H ^{15}N HSQC spectra of TLNRD1-4H	120
Figure 4.8. ^1H ^{15}N HSQC TLNRD1-4H titration against RIAM peptide 4-30	121
Figure 4.9. LD motifs of RIAM and Lamellipodin.....	122

Figure 4.10. TLNRD1 4-helix domain interacts with Lpd peptide spanning residues 20-46.....	123
Figure 4.11. Identifying TLNRD1 residues which mediate TLNRD1-RIAM interaction	124
Figure 4.12. Comparison of surface electrostatics between wild-type and 2E TLNRD1.....	125
Figure 4.13. TLNRD1-2E no longer interacts with RIAM peptide	126
Figure 4.14. Comparison of talin R7 KANK binding domain with TLNRD1 5-helix domain.....	128
Figure 4.15. Helix arrangement for talin R7 vs TLNRD1.....	128
Figure 4.16. Structural modelling of KANK1 LD with talin R7 and TLNRD1	129
Figure 4.17. Full-length TLNRD1 interacts with KANK1 LD peptide spanning residues 30-68.....	130
Figure 4.18. TLNRD1 4-helix domain does not interact with KANK1 peptide	131
Figure 4.19. TLNRD1 interacts with KANK2 LD motif peptide spanning residues 31-69	132
Figure 4.20. CDK1 207-222 modelling with TLNRD1 4-helix domain.....	133
Figure 4.21. TLNRD1 interaction with CDK1 LD motif peptide spanning residues 206-223.....	134
Figure 4.22. ¹ H ¹⁵ N HSQC overlaid spectra of CDK1 206-223 titration against TLNRD1.....	136
Figure 4.23. R8 actin binding site comparison with TLNRD1	138
Figure 4.24. TLNRD1 actin co-sedimentation assay.....	139
Figure 4.25. TLNRD1 4-helix wild-type vs 2E mutant	140
Figure 4.26. TLNRD1 bundles actin filaments in a dimerisation dependent manner	142
Figure 4.27. Transmission electron microscopy images of TLNRD1 actin bundling	144
Figure 5.1. TLNRD1 zebrafish morpholino.....	151
Figure 5.2. CRISPR/Cas9 mediated <i>tlnd1</i> knock-out leads to severe developmental defects	152
Figure 5.3. Spinning disk confocal microscopy of GFP-TLNRD1 imaged by Guillaume Jacquemet.....	154
Figure 5.4. TLNRD1 expression and filopodia quantification from Guillaume Jacquemet.....	155

Figure 5.5. Immunoprecipitation test of TLNRD1 antibodies.....	157
Figure 5.6. Identification of endogenous TLNRD1 expression	158
Figure 5.7. GFP-TLNRD1 actin stress fibre localisation as imaged by Guillaume Jacquemet.....	159
Figure 5.8. GFP-TLNRD1 4-helix domain localisation with ventral stress fibres	160
Figure 5.9. Spider plots of random migration in 2D	161
Figure 5.10. Migration analysis of wild-type vs. GFP-TLNRD1 and siRNA knockdowns	162
Figure 6.1. Location of selected mutations along the talin-1 molecule.....	170
Figure 6.2. Structural comparison of wild-type and mutated residues.....	172
Figure 6.3. Comparison of wild-type vs. talin-1 mutant phenotypes (Latifeh Azizi)	174
Figure 6.4. Comparison of cell migration, division and invasion in wild-type vs. mutants (Latifeh Azizi)	175
Figure 6.5. Comparison of adhesion component expression and localisation with talin (Latifeh Azizi)	177
Figure 6.6. CD spectroscopy of wild-type R7R8 vs L1539P	180
Figure 6.7. CD spectroscopy of wild-type R7R8 vs R1368W and Y1389C.....	181
Figure 6.8. NMR ¹ H ¹⁵ N HSQC TROSY spectra of R7R8 and mutants R1368W and Y1389C	182
Figure 6.9. Analysis of vinculin binding to wild-type R7R8 and mutants R1368W and Y1389C	184
Figure 6.10. Mutations R1368W and Y1389C have no effect on ligand interaction. Cell imaging by Latifeh Azizi	186
Figure 6.11. Talin L2509P mutation stops talin dimerisation	187
Figure 6.12. L2509P reduces actin binding.....	188
Figure 7.1. Model of TLNRD1 interactions.....	195
Figure 7.2. Schematic representation of TLNRD1 actin bundling	197
Figure 7.3. Schematic summary of mutation impact on talin functions	200

List of Tables

Table 1. Chemicals and reagents	49
Table 2. Buffer composition	50
Table 3. Biochemistry construct list.....	51
Table 4. Cell biology construct list.....	52
Table 5. Primary and Secondary Antibodies	70
Table 6. Data collection and refinement statistics for TLNRD1-FL and 4H.....	91
Table 7. COSMIC talin-1 mutations identified for further investigation (Latifeh Azizi).....	170
Table 8. Summary of observations for talin-1 mutations	189

Chapter 1. Introduction

1.1. The extracellular matrix

All multicellular organisms have a molecular non-cellular scaffold, the extracellular matrix (ECM), which provides dynamic support for the development and maintenance of tissues. The ECM is composed of a variable meshwork of proteins including fibronectin, collagen, laminin and versican, and is continuously rebuilt and remodelled by interacting fibroblasts (Humphrey, Dufresne and Schwartz 2014). The complex interaction between cells and the ECM regulates all major biological processes from cell division and migration, to wound repair, tissue development and immunological responses to pathogens. Disruption to the ECM or cellular communication with the ECM is a key driver for pathological diseases such as osteoarthritis, vascular fibrosis and metastatic cancer (Zhen and Cao 2014; Venning, Wullkopf and Eler 2015; Harvey et al. 2016).

1.2. Cell migration and adhesion

The coordinated and controlled mechanisms behind cellular migration and adhesion are fundamental to correct tissue development in multicellular organisms. With an integral role in the development of all tissues, disruptions to the regulation and coordination of cell migration or adhesion processes can lead to a number of pathological diseases. The migration of immune cells to key locations in the body is crucial for the regulation of their function, with abnormal localisation leading to exacerbation of diseases such as rheumatoid arthritis or atherosclerosis (Worbs, Hammerschmidt and Förster 2017). Disruptions to neural cell migration and adhesion through mutation in key regulatory proteins contributes to a number of devastating neurological diseases such as Alzheimer's disease and Polymicrogyria (Valiente and Marín 2010). Finally, aberration in cell adhesion machinery and subsequent abnormal cell migration is one of the main contributors of cancers progression through to metastatic cancer in a range of different cell types (Bendas and Borsig 2012; Hamidi and Ivaska 2018).

Cell migration and adhesion is regulated in response to certain mechanical cues exerted by the ECM composition and chemical signals released by nearby cells into the extracellular environment. Cells can either migrate individually or coordinate migration as a collective group which is important for processes such as wound healing or angiogenesis (Mayor and Etienne-Manneville 2016). In order to migrate, cells need to establish polarity, a process defined by an asymmetric organisation of cell shape and protein distribution (Nelson 2003). Through this mechanism, migrating cells form a front

facing leading edge which determines the direction of movement, and a rear facing trailing edge which continuously retracts during migration. This process of migration is regulated both through the formation of adhesion contact sites between the cell and the ECM, and the structural remodelling of the intracellular cytoskeleton.

1.3. The cytoskeleton

The cells cytoskeleton is an incredibly complex piece of cell architecture that is continuously changing in response to cellular processes. It regulates cell shape and protrusion formation as cells migrate, acts as a transport network for the movement of proteins from one location to another within the cell and organises spindle formation for the splitting of chromosomes during cell division. The cell cytoskeleton is composed of three core types of filamentous protein networks, the actin cytoskeleton, the microtubule cytoskeleton and the intermediate filament cytoskeleton, each with a different protein composition, size and structure (Figure 1.1).

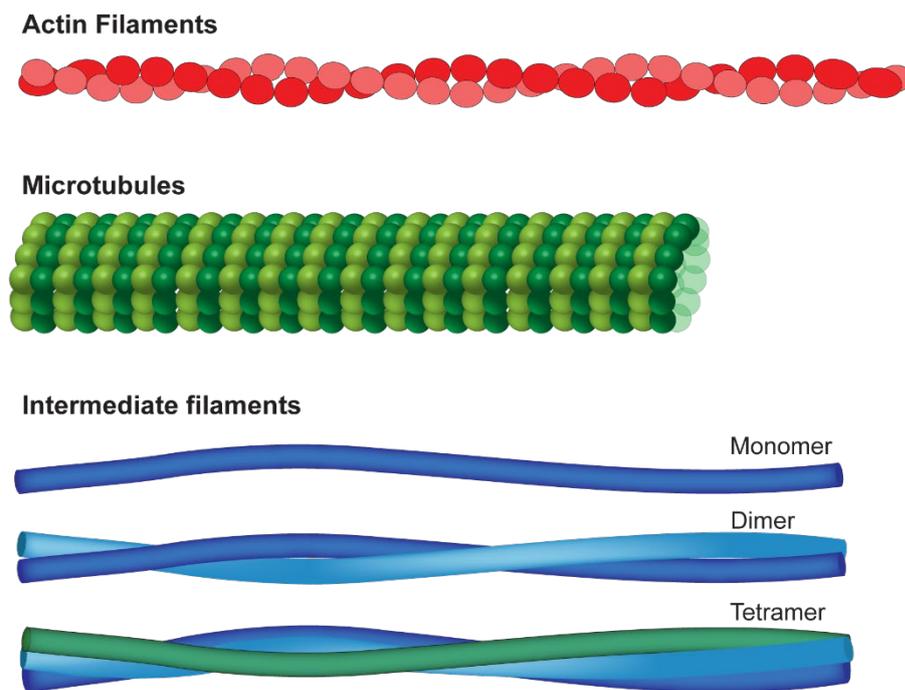


Figure 1.1. Architecture of cell cytoskeleton networks

Schematic of cytoskeletal components. Polymers of globular actin monomers create contractile actin filaments which can be intertwined and bundled, microtubules consist of tubulin monomers which form hollow tube-like structures and intermediate filaments consist of rope-like structures which can be coiled to form thicker filaments.

1.3.1. Microtubules

Microtubule filaments perform a number of crucial roles in the cell, with one example being cell division whereby arrays of microtubules form the mitotic spindle and trigger

cleavage furrow formation during cell division (Dechant and Glotzer 2003). Microtubule filaments are built up from a series of α and β tubulin dimers, which polymerise end to end to form large hollow 25 nm diameter protofilaments. Structurally these are the widest types of filaments which form the cytoskeleton and have higher rigidity than the other filaments allowing them to withstand much greater forces (Brangwynne *et al.* 2006). The initiation of microtubule filament formation, i.e. nucleation, starts with γ -tubulin complexes located in microtubule organising centres such as the centrosome (Moudjou *et al.* 1996). Here the minus end of the filament, which is the slow growing α -tubulin exposed end, is anchored ready for filament assembly from the plus end of the filament, which is the fast growing β -tubulin exposed end (Sulimenko *et al.* 2017).

1.3.2. Intermediate filaments

Intermediate filament proteins comprise a family of approximately 70 genes, which can be classified into five different groups depending on localisation, structure and sequence. Groups I and II include the well-known keratins which are mainly expressed in the epithelium, group III comprises vimentin, desmin and other related proteins which are more widely expressed, group IV includes proteins such as α -internexin and neurofilaments which are specifically expressed in neurons and group V comprises those found ubiquitously in the cell nucleus such as lamin (Lowery *et al.* 2015). Intermediate filaments typically consist of 2 or more α -helical rod shaped homopolymers or heteropolymers which interact laterally to form a coiled-coil complex (Figure 1.1). Unlike microtubules and actin, intermediate filaments do not have polarity which determines the direction of polymerisation. These filaments play a number of supportive roles in the cell and are important for directionality and persistence during cell migration and maintaining cell-cell interaction during collective migration (Leduc and Etienne-Manneville 2015; Gan *et al.* 2016; De Pascalis *et al.* 2018).

1.3.3. Actin

The actin network is perhaps one of the most well studied components of the cell cytoskeleton, with ubiquitous high expression and an integral role in maintaining cell shape and regulating associated migratory processes such as adhesion and protrusion formation. Actin is a highly conserved 42 kDa protein with 6 highly similar mammalian isoforms; α 1-skeletal (ACTA1), α 2-smooth (ACTA2), α -cardiac (ACTC1), β -actin (ACTB), γ 1-actin (ACTG1) and γ -smooth (ACTG2). Actin either exists as monomeric globular actin (G-actin) or polymerises into filamentous actin (F-actin), forming a

number of different structures in the cell with different mechanical properties (Figure 1.2).

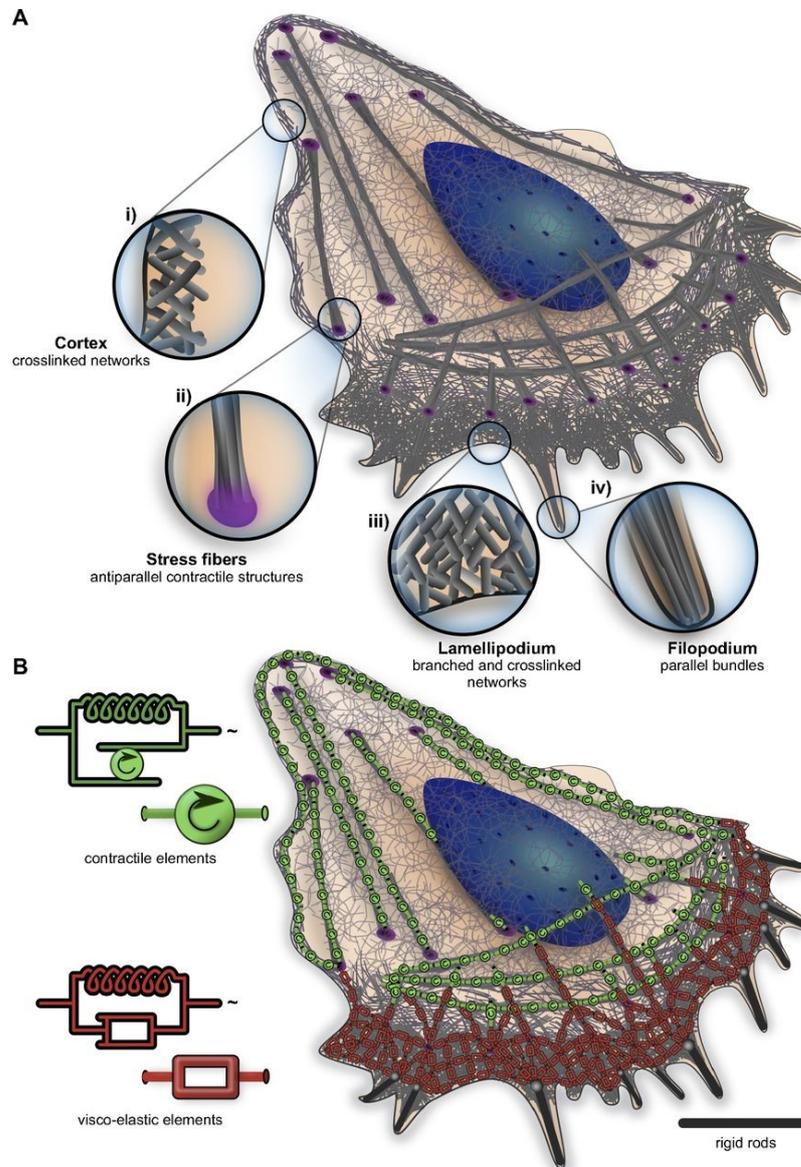


Figure 1.2. Dynamic actin structures in a migrating cell (Blanchoin *et al.* 2014)

(A) Representation of different actin-based structures in a migrating cell. Highlighted components include crosslinked actin networks in the cell cortex (i), large antiparallel contractile stress fibers (ii) and protrusion associated networks (iii and iv). **(B)** Mechanical properties of the different actin-based structures, with contractile structures in green and more dynamic elastic structures in red.

Actin itself is a conformationally flexible protein consisting of two major domains, which can be further split into four subdomains, with an ATPase site nestled in a centrally located cleft (Figure 1.3). Subdomain II contains a highly dynamic DNase I binding loop which is thought to be important for stabilising actin monomer-monomer contacts during filament nucleation (Khaltina and Strzelecka-Gołaszewska 2002; Wawro *et al.*

2005). The domain arrangement of actin creates a pointed minus end (-) encompassing domains II and IV, while the opposing end is termed the barbed plus end (+) with domains I and III. This polarised arrangement determines the direction of filament assembly, with addition of actin monomers largely occurring at the barbed end in the presence of ATP.

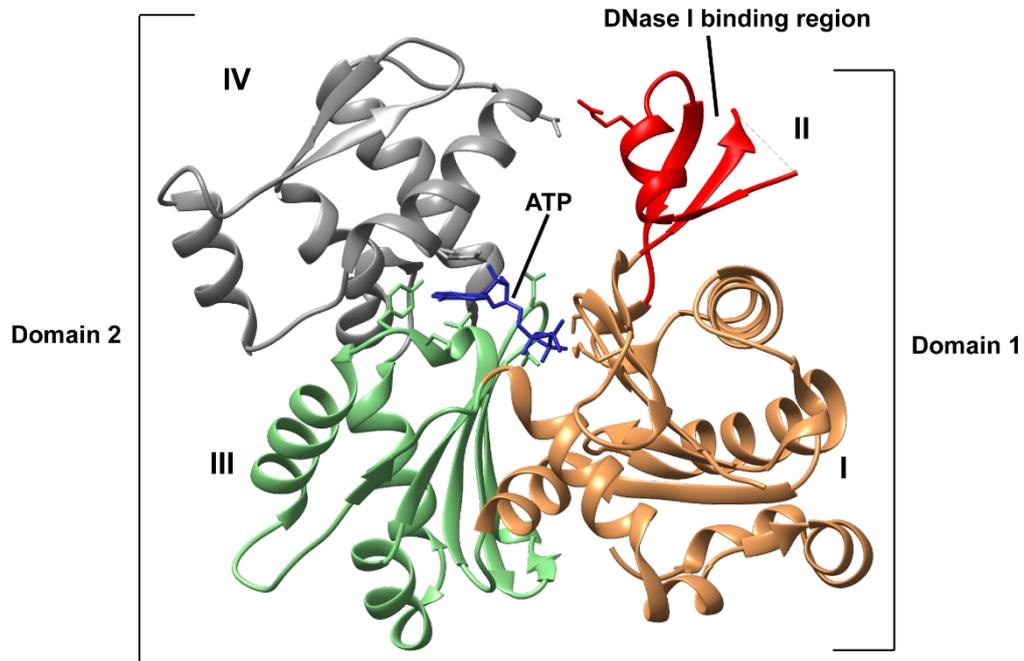


Figure 1.3. Structure of monomeric G-actin

X-ray crystal structure of native G-actin (PDB ID: 3HBT) (Wang, Robinson and Burtnick 2010) in complex with a single ATP molecule (blue). G-actin consists of four subdomains organised into two major domain regions. Subdomain II (red) contains a DNase I binding loop not visible on this structure.

The dynamic assembly of F-actin begins with nucleation, which requires a trimeric complex of actin monomers or nucleating factors such as the Arp2/3 complex which mimic an actin trimer. This is followed by polymerisation and elongation of the filament which is regulated and coordinated by a vast number of directly interacting proteins such as Ena/VASP family members and formins (Bear and Gertler 2009; Breitsprecher and Goode 2013). Final assembled single F-actin filaments contain two intertwined protofilaments which form a double helix-like structure (Figure 1.1).

1.3.4. Actin bundling proteins

Actin filaments are brought together by bundling and cross-linking proteins to form large superstructures, including larger linear bundled filaments and cross-linked actin mesh-like structures which regulate cell rigidity and protrusion formation during migration and invasion (Stevenson, Veltman and Machesky 2012). The polarised nature of actin

filaments allows selective bundling of filaments in different orientations and with different filament distances. Some proteins such as fascin, which has two actin binding sites per monomer, only bundle in a parallel formation with a tight ~8 nm filament distance (Jansen *et al.* 2011; Winkelman *et al.* 2016) (Figure 1.4). Others such as α -actinin, which bundles through dimerisation via a single actin binding domain per monomer, can bundle filaments in both antiparallel and parallel orientations with a greater filament distance of ~35 nm.

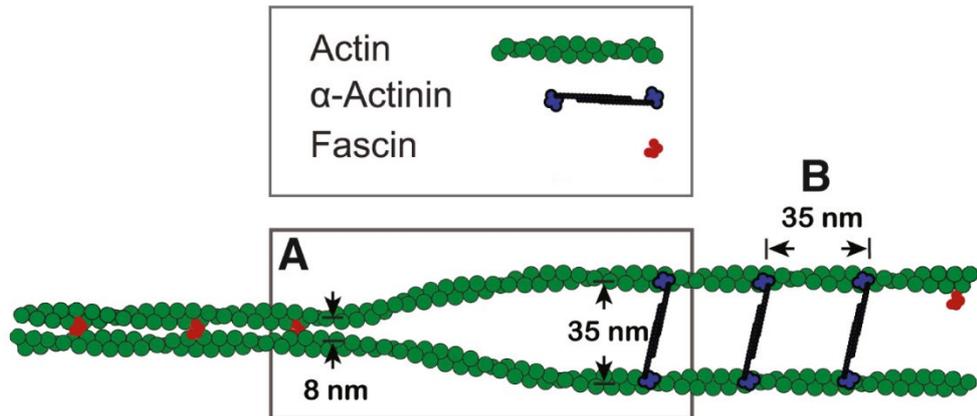


Figure 1.4. Schematic of filament bundling, adapted from Winkelman *et al* (2016)

(A) Fascin bundles filaments with tight 8 nm spacing whereas α -actinin bundles with wider 35 nm spacing. (B) Distances between actin-bundling protein binding along filaments are approximately 35 nm.

1.4. Actin in cell migration and adhesion

For cells to migrate, they must be polarised, a process regulated by the antagonistic behaviour of GTPases Rho and Rac which orchestrate remodelling of the actin cytoskeleton. Collectively these GTPases control the mode of cell migration which can be either mesenchymal with an elongated morphology, amoeboid with a more rounded morphology or collective whereby cells maintain cell-cell interactions during migration. Aberrant regulation or expression of either Rac or Rho, and subsequent abnormal regulation of the actin cytoskeleton, is a key feature in cancers leading to increased invasion and promoting metastasis by destabilising cell-cell interactions (Parri and Chiarugi 2010). GTPases such as Rac1 and Cdc42 are also regulators of linear actin filament polymerisation and cross-linking at the leading edge of the cell, promoting actin nucleation and protrusion formation via WAVE complex activation (Warner, Wilson and Caswell 2019; Mehidi *et al.* 2019). This actin polymerisation at the leading edge of a cell is crucial for the formation of filopodia, small finger-like projections which enable cells to probe the extracellular environment, and lamellipodia, thin protruding sheet-like

structures which promote cell migration (Krause and Gautreau 2014; Jacquemet, Hamidi and Ivaska 2015).

1.4.1. Filopodia

Filopodia are elongated finger-like protrusions which form at the leading edge of a migrating cell (Figure 1.5). These actin rich structures contain adhesion regulating integrin receptors at their tips and allow the cell to sense changes in the ECM such as substrate stiffness (Wong, Guo and Wang 2014). Cells using filopodia to sense the extracellular environment is crucial for developmental processes such as angiogenesis (Wakayama *et al.* 2015; Zhang *et al.* 2018) as well as immune surveillance (Ward 2009) and wound healing (Wood *et al.* 2002) throughout an organisms lifespan.

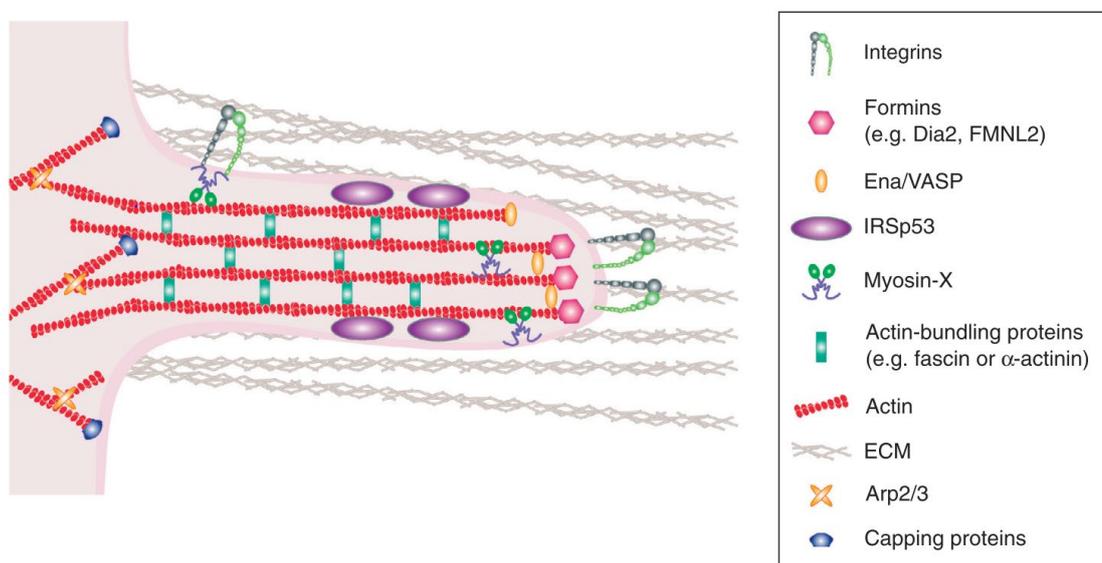


Figure 1.5. Filopodia architecture, adapted from Jacquemet *et al* (2015)

Bundles of parallel actin filaments within filopodia enable protrusion for ECM sensing and to provide a platform for myosin X transport of adhesion signalling receptors such as integrins to the filopodia tips.

Filopodia formation is promoted by linear actin polymerisation by formins such as Dia2 which act downstream of Rho family GTPases (Arjonen, Kaukonen and Ivaska 2011). Polymerised actin within filopodia is stabilised in a parallel formation by actin bundling proteins such as fascin which provides rigidity to the thin protrusions (Vignjevic *et al.* 2006; Van Audenhove *et al.* 2016). Talin-mediated integrin activation at the tips of filopodia promotes assembly of distinct filopodial adhesion complexes, with enrichment of proteins such as kindlins, lamellipodin, integrin beta-1-binding protein 1 (ITGB1BP1) and myosin X (Jacquemet *et al.* 2019).

1.4.2. Lamellipodia

Lamellipodia are wide fan-shaped forward driving protrusions which form at the leading edge of the cell, typically on 2D surfaces. The growth of lamellipodia is primarily driven by the continuous polymerisation of actin filaments, with the actin nucleating Arp2/3 complex acting downstream of Rac1 (Mehidi *et al.* 2019) being primarily responsible for the complex dendritic actin branching which gives them their structure (Figure 1.2A). The direction of lamellipodial growth is driven by cues from filopodial protrusions and filopodial adhesion complexes, with fascin containing F-actin bundles from filopodia creating a template for lamellipodia formation (Johnson *et al.* 2015). During migration, actin is continuously polymerised and cross-linked at the tip of the lamellipodium, with myosin motors at the lamellipodium base/rear (lamella) generating force by retracting actin filaments (Sayyad *et al.* 2015).

1.5. Cell-matrix adhesion

Cellular attachment to the ECM is mediated by large multi-protein adhesion complexes which at their core contain transmembrane cell adhesion receptors. Cell adhesion receptors can be distinguished into five core groups with distinguishable functions. The immunoglobulin superfamily members such as nectins and mucins form both cell-cell and cell-ECM interactions, selectins are vascular adhesion molecules important for leukocyte trafficking, the cadherin superfamily mediates cell-cell adhesion, and integrins coordinate cell-ECM adhesion in most cell types (Läubli and Borsig 2010; Wai Wong, Dye and Coombe 2012; Brasch *et al.* 2012; Kechagia, Ivaska and Roca-Cusachs 2019).

1.6. Adhesion complexes

Not only are there different adhesion receptors, but there is also variety in the type of adhesion complexes which can form within cells. Adhesion complexes are essential for force transmission during migration and enable cells to sense the rigidity of the extracellular environment which can vary from tissue to tissue. Their formation can vary in location and composition depending on whether the cell is migrating, invading or forming cell-cell contacts. Currently known adhesion complexes include integrin-based nascent, focal, fibrillar and reticular adhesions (Wehrle-Haller 2012; Lock *et al.* 2018), newly documented fibronectin dependent 'stitch' adhesions (Pankov *et al.* 2019) and

actin-rich podosomes (Linder and Wiesner 2016). Figure 1.6 below shows a typical representation of adhesion complex formation during cell migration.

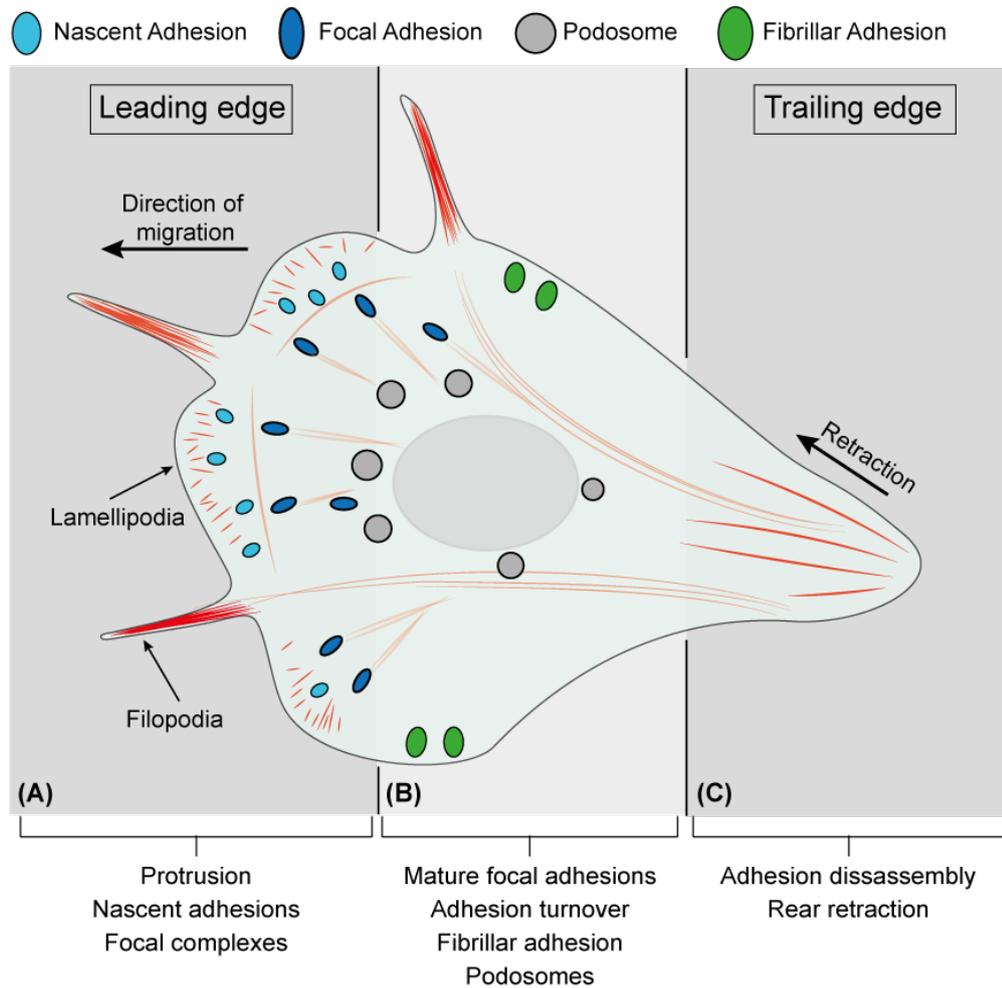


Figure 1.6. General principles of cell migration and adhesion formation

(A) At the leading edge, nascent adhesions form which either rapidly disassemble or mature into focal complexes. Lamellipodial protrusion and filopodia formation in this region is a classic characteristic of migrating cells. **(B)** In the central region of the cell, focal complexes mature into stable focal adhesions or progress to fibrillar adhesion formation. Here is also where ECM degrading podosomes form. **(C)** At the trailing edge adhesions disassemble releasing the cell-ECM adhesion and allowing rear retraction.

1.6.1. Nascent adhesions

Nascent adhesions, first described as nascent focal contact sites (Jockusch *et al.* 1995), are some of the first adhesions to rapidly form at the leading edge of the cell below the protrusive front. Integrin nanoclusters trigger nascent adhesion formation through binding to ECM ligands and subsequent signalling through talin (Changede *et al.* 2015). This initial process has been shown to occur independently of the mechanical

environment experienced by the cell (Changede *et al.* 2015; Baade *et al.* 2019), whereas the subsequent maturation of nascent adhesions towards focal complex formation is highly dependent on mechanical forces and myosin II-actin crosslinking (Choi *et al.* 2008). An integrin-talin-actin complex forms the core of early nascent adhesion composition with few other proteins such as paxillin being recruited as the adhesion stabilises (Pasapera *et al.* 2010).

1.6.2. Focal complexes and focal adhesion

Nascent adhesions either disassemble or begin to mature into focal complexes as the cell experiences mechanical force, eventually becoming highly stable integrin-mediated focal adhesions (FA). These are large multi-protein structures that connect the extracellular matrix to the intracellular cytoskeleton primarily through a mechanically sensitive integrin-talin-actin core (Elosegui-Artola *et al.* 2016). FA's translate mechanical forces into intracellular signals for the regulation of cell behaviour. They contain a vast number of different proteins including scaffolding proteins, kinases and phosphatases, all organised in unique nanoscale 'layers' of signalling components (Kanchanawong *et al.* 2010; Liu *et al.* 2015) (Figure 1.7).

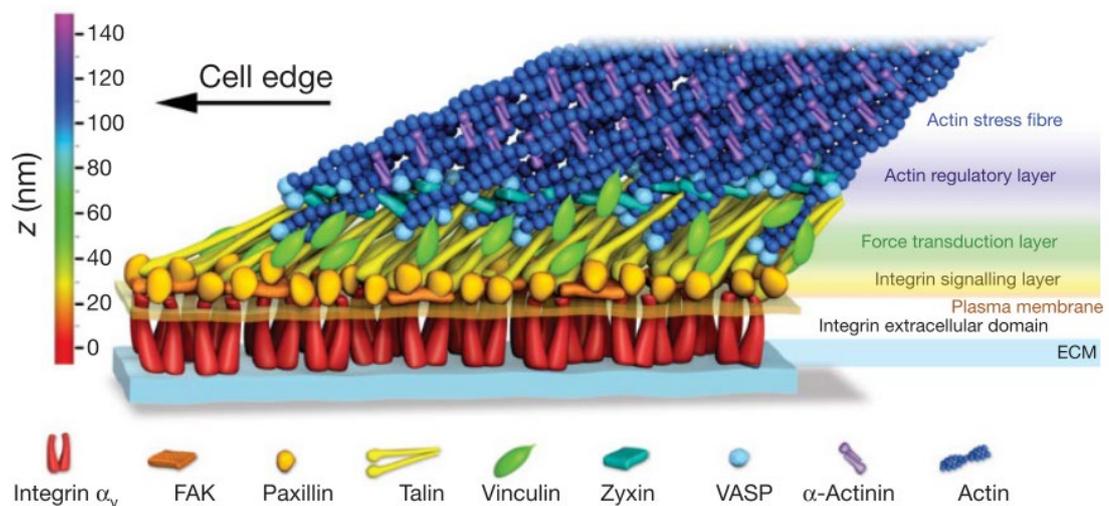


Figure 1.7. Nanoscale organisation of focal adhesion complexes (Kanchanawong *et al.* 2010)

Schematic showing the organisation of focal adhesion signalling components into distinct layers. At the membrane is the integrin signalling layer, followed by the force transduction layer with talin as the core platform for adhesion protein recruitment, and ending with the actin regulatory area where actin filament attachment to talin provides connection to the cell cytoskeleton.

1.6.3. Reticular adhesions

Reticular adhesions are newly characterised distinct adhesions which mediate cell-ECM interaction during mitotic cell division (Lock *et al.* 2018). During mitosis, cells disassemble focal adhesions and round up for division, at which point reticular adhesions continue to aid residual adhesion and guide re-spreading of the cell. Unlike focal adhesions, these do not require talin mediated coupling to the actin cytoskeleton and are dependent on $\alpha\beta 5$ integrin attachment to the ECM during interphase (Lock *et al.* 2018; Lock *et al.* 2019).

1.6.4. Fibrillar adhesion

Fibrillar adhesions are specialised elongated complexes specifically found in fibroblasts, typically developing from established focal complexes and located towards the centre of the cell (Zamir *et al.* 2000). Fibrillar adhesions promote remodelling of fibronectin rich ECM and fibril formation which guides deposition of other ECM components such as fibrillin and collagen (Sabatier *et al.* 2009; Sottile and Hocking 2002; Saunders and Schwarzbauer 2019). The composition of fibrillar adhesions are dominated by integrin $\alpha 5\beta 1$ and the 220 kDa phosphoprotein tensin, and are regulated in a mechanoresponsive manner through interaction with the paxillin related scaffold protein Hic-5 (Goreczny, Forsythe and Turner 2018).

1.6.5. Podosomes

Podosomes are unique actin-rich ring structures surrounded by classic adhesion proteins that are largely found in monocytic cell types and osteoclasts (Linder and Wiesner 2016). These podosome structures are distinctly different and more cell type specific than other adhesion complexes, and are typically found clustered near the lamellipodia of cells undergoing rapid turnover (Murphy and Courtneidge 2011). Podosome formation is crucial during development, with research showing importance in neural crest migration during embryonic development and aiding angiogenic sprouting by promoting proteolysis of the ECM (Murphy *et al.* 2011; Spuul *et al.* 2016). Coordination of ECM degradation with migration has been suggested as the core role for podosome formation in cells, with podosomes recruiting a number of different proteases to aid ECM degradation (Jevnikar *et al.* 2012; Xiao *et al.* 2013).

1.7. Talin

Talin is a large 270 kDa cytoplasmic protein which acts as a mechanosensitive signalling platform for adhesion regulation. At the N-terminus, the talin head region contains a FERM domain which directly interacts with the plasma membrane and binds to the tail of β -integrins, promoting integrin activation for adhesion to the ECM (Elliott *et al.* 2010). At the C-terminus, talin has a dimerisation domain which is essential for adhesion related functions and actin filament binding (Gingras *et al.* 2008). The central core of talin, the rod domain, consist of a series of 4 and 5 α -helical bundles split into 13 domains (R1-R13) (Figure 1.8). These domains are able to unfold and stretch in response to mechanical force in a step-wise manner which regulates binding site availability for different ligands like a mechanochemical switch (Goult, Yan and Schwartz 2018; Yao *et al.* 2016). Some ligands such as RIAM only interact with folded talin rod domains and are unable to interact when talin unfolds under force (Goult *et al.* 2013a). On the other hand, this unfolding of talin also exposes a number of cryptic binding sites for vinculin along the entire rod-domain which promotes maturation of adhesion complexes to stable focal adhesions.

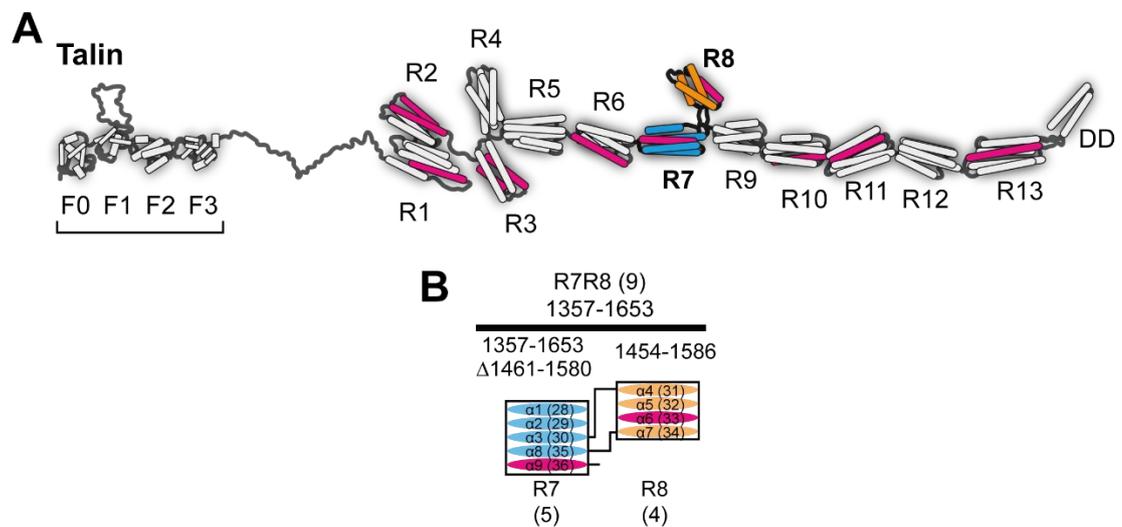


Figure 1.8. Schematic representations of the talin molecule

(A) Schematic of talin N-terminal FERM domain connected to a series of 4 and 5-helix bundles and C-terminal dimerisation domain. The R7 (blue) and R8 (orange) domains of talin are highlighted. Vinculin binding sites are highlighted in pink. **(B)** Schematic representation of talin R7R8 domain topology, with vinculin binding sites highlighted in pink.

Towards the central rod region of talin lies the domains R7 and R8 which have a unique domain organisation compared to other helical structures of talin (Figure 1.8). The 4-helix module of the R8 domain lies between two helices (30 and 35) of the R7 domain, which allows the R8 domain to sit outside the force transmission pathway, with R7 protecting the region from force induced unfolding (Yao *et al.* 2016). This region of talin is an important signalling platform for binding of ligands containing LD-motifs, many of which are discussed in **section 1.8** below.

1.7.1. Talin evolution and isoforms

There are only two known talin isoforms, talin-1 and talin-2, with few differences between the two both in amino acid sequence and protein structure. Talin-1 is located on chromosome 9p13 with a gene product of ~30 kb due to relatively small intron sizes and talin-2 is located on chromosome 15q22 with a considerably larger ~190 kb product with larger introns (Debrand *et al.* 2009). The extent of conservation between the two isoforms is relatively high, with the C-terminal R13 domain, dimerisation domain and N-terminal FERM F3 domain showing the highest level of conservation (Gough and Goult 2018). These two domains are the most crucial for talins adhesive function, with the F3 domain binding and activating integrins, and R13-DD promoting talin dimerisation for actin filament engagement.

Each isoform exhibits a different pattern of expression in the human body. Inspection of expression data from the EBI expression atlas (Petryszak *et al.* 2016) reveals that the expression of talin-1 is widespread, with ubiquitous expression across all tissues and cell types. Talin-2 on the other hand shows a lot more tissue specific expression patterns, with high expression in cells of the cerebral cortex, the kidneys and the cardiovascular system, particularly in the skeletal muscle of the heart. The subtle differences between the two isoforms likely affects the mechanical properties and dynamic interactions with ligands, however, complete knock-down of talin-1 can be rescued by expression of talin-2, suggesting that their behaviour is not mutually exclusive and that there is some crossover (Zhang *et al.* 2008).

One of the more intriguing aspects of talin-2 compared to talin-1 is the considerable number of splice variants which leads to variations in the length and composition of the talin-2 proteins being produced (Figure 1.9). Both short and long versions of the isoform are produced, with the shorter isoforms losing the FERM domain integrin binding properties but retaining the C-terminal dimerisation domain. The functional significance

of these isoforms are yet to be explored but indicate an alternative function to the canonical integrin-actin linking and focal adhesion regulation.

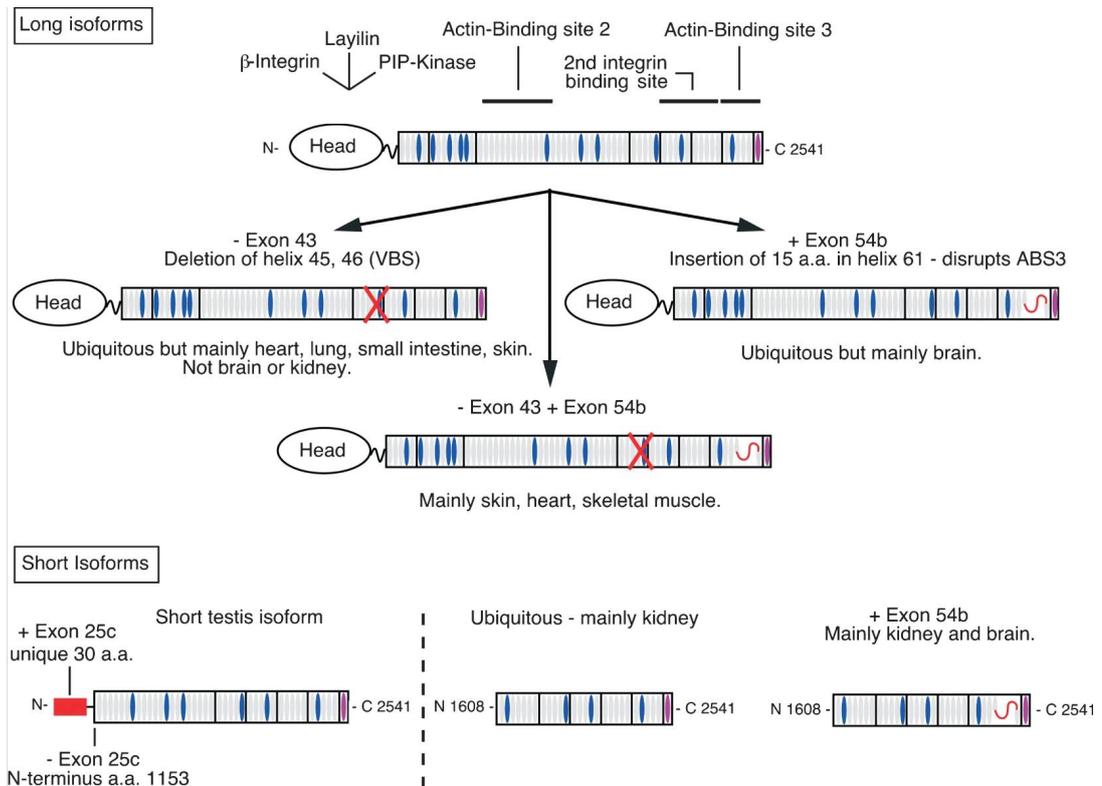


Figure 1.9. Schematic representation of talin-2 splice variants (Debrand *et al.* 2009)

Diagram showing both short and long splice variants of talin-2. The vinculin binding sites are indicated in blue, and red cross indicating loss of exon 43 leading to isoform truncation compared to full-length. Red curve indicates 15 residue insertion in ABS3.

Talin has an evolutionary history that stretches as far back as the protozoan single cell lineage amoebozoa, with talin-2 predicted to be the first ancestral gene to emerge (Senetar and McCann 2005). From this, it is likely that talin-1 emerged through a gene duplication event of talin-2 prior to the evolution of the chordate lineage, which includes vertebrates, marine invertebrates and small marine cephalochordate. Throughout evolution, the overall gene structure, protein length and protein domain composition has remained relatively well conserved, suggesting that the function of talin in its current form is essential for multicellular life and has remained the same throughout its evolution.

1.8. Talin-1 ligands

The talin molecule has a diverse set of interacting proteins which coordinate recruitment of adhesion complex components and regulate cellular responses. The binding sites for these ligands are located in numerous domains along the talin rod, with the binding of some, including RIAM and vinculin, being modulated mechanically by the folded state of the domains (Figure 1.10). Current knowledge of the roles that some key talin interactions have are explore in more detail in this section.

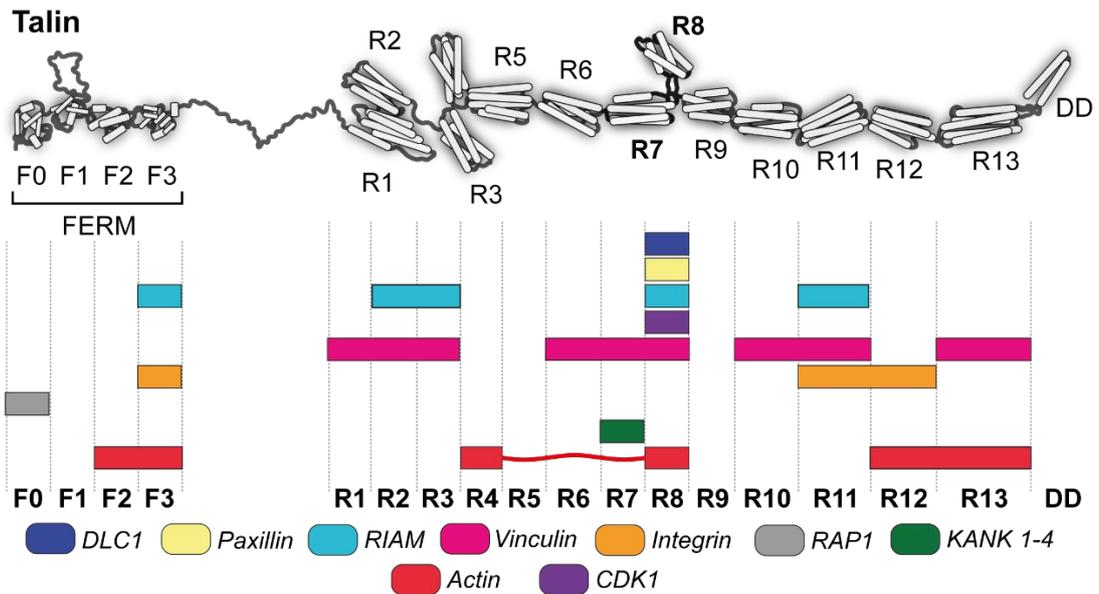


Figure 1.10. Schematic of talin interactors

Schematic of talin domain structure with ligand interaction sites indicated. Domains are shown with colour map indicating binding site locations and corresponding ligands along the talin molecule.

1.8.1. Integrins

Integrins are transmembrane heterodimeric adhesion receptors which directly interact with components of the extracellular matrix for environmental sensing and traction generation. They can be made up of up to 24 combinations of α and β subunits of which there are numerous types, with up to 12 integrin heterodimer combinations known to contain $\beta 1$ integrin (Figure 1.10). Different combinations of α and β heterodimers show different properties, with variations in tissue distribution, extracellular matrix ligand specificity and intracellular ligand interactions.

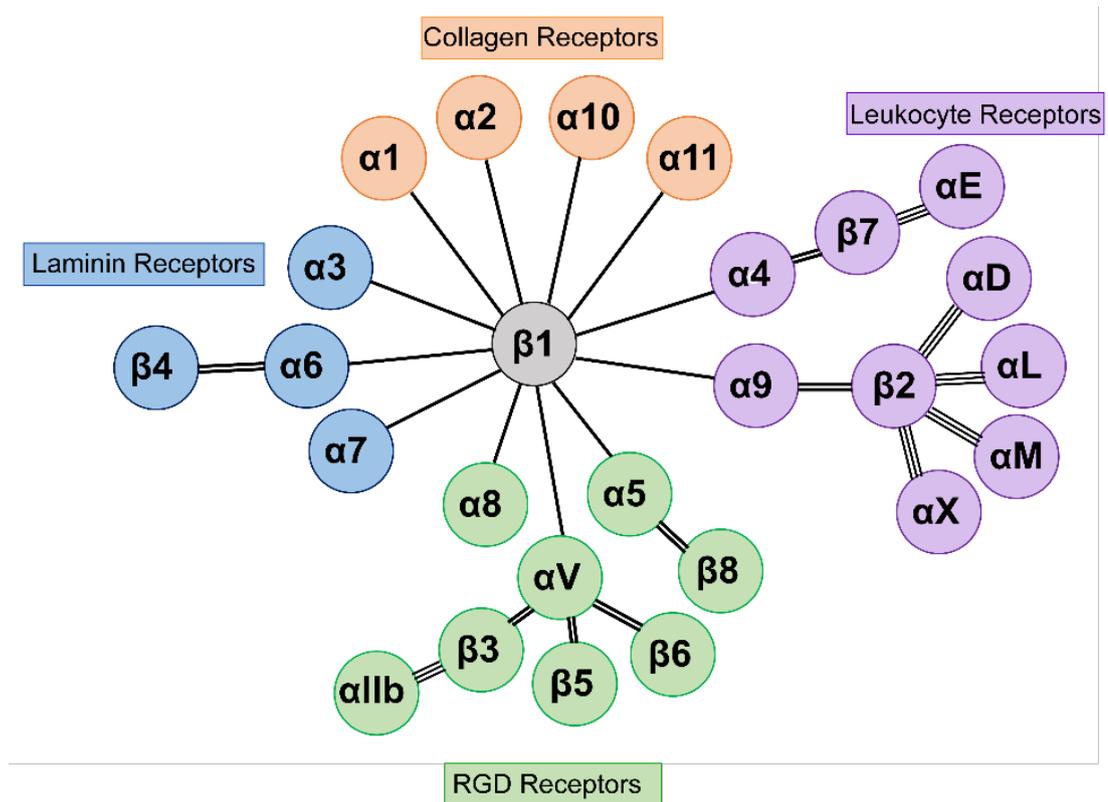


Figure 1.11. Diagram of all integrin α and β pairings

Schematic representing all possible combinations of α and β integrin heterodimers. Integrins are organised based on their target extracellular ligands including Laminin (blue), leukocyte receptors (purple), collagen (orange) and Arg-Gly-Asp (RGD) containing receptors (green).

Most integrins have a short cytoplasmic tail domain connected to a large extracellular ectodomain via a single transmembrane helix. The ectodomains of integrins contain a number of smaller domains connected via flexible linkers which enable the receptors to adopt multiple different conformations depending on its activation state. When inactive, integrin ectodomains adopt a 'bent' closed or 'extended' closed conformation which prevents or weakens interaction with ECM ligands (Li and Springer 2017). When activated either from ECM cues through outside-in signalling or from intracellular signalling events such as the binding of talin, integrin ectodomains adopt an 'extended' open conformation which enables high affinity stable interaction with ECM ligands (Figure 1.11).

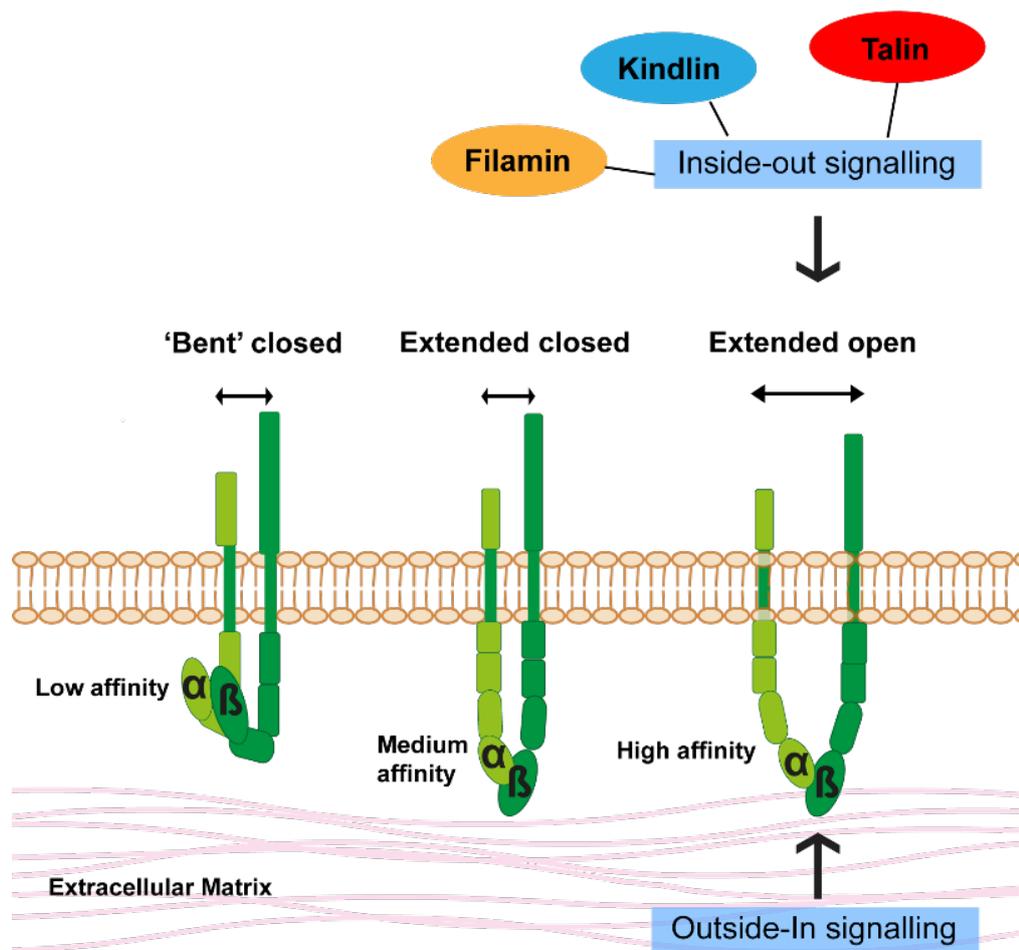


Figure 1.12. Integrin activation states

Illustration showing the different states of integrin activation. Integrins in a 'bent' closed conformation have low ligand affinity and is considered the inactive state, integrins in an extended closed conformation have slightly increased but unstable affinity for ligands and integrins in an extended open conformation are considered fully active with high affinity.

Inside-out signalling involves dynamic interaction of integrin activating proteins which bind to the integrin cytoplasmic tails. Talin is a key activator for integrin mediated adhesion which directly interacts with NPxY motifs in β -integrin tails. Talin activation of integrins involves disruption of a salt bridge between the α and β tails and additional talin interaction with PIP_2 phospholipids in the membrane, separating the membrane-proximal regions of α and β to activate the high affinity state of the heterodimer (Hughes *et al.* 1996; Elliott *et al.* 2010).

1.8.2. Vinculin

Vinculin is a cytoplasmic actin filament binding protein essential for mesoderm development and plays a key role in regulating adhesion complex stability and

maturation (Xu, Baribault and Adamson 1998; Case *et al.* 2015). Vinculin has been studied extensively for its mechanosensitive role in focal adhesion formation and its ability to stabilise actin filament association with adhesion complexes through direct interaction with the talin molecule (Burrige and Mangeat 1984; Atherton *et al.* 2015). Vinculin is thought to exist in an autoinhibited conformation with the head region, containing subdomains D1-D4, interacting with the D5 tail domain which masks a connecting proline rich region (Ziegler, Liddington and Critchley 2006). Interactions between the vinculin D1 domain (VD1) and adhesion related ligands such as talin relieves vinculin autoinhibition, allowing activation and subsequent reinforcement of adhesion complex association with the actin cytoskeleton (Case *et al.* 2015). The vinculin VD1 domain directly interacts with a number of binding sites along the talin rod when the domains unfold under mechanical force (Humphries *et al.* 2007; Yao *et al.* 2015).

1.8.3. Actin

Both talin-1 and talin-2 contain up to three identified actin binding sites, ABS1-3, with each having different affinities and properties (Hemmings *et al.* 1996). The first actin binding site, ABS1, resides in the N-terminal talin FERM F3 domain, close to the integrin binding site. ABS1 has been shown to inhibit barbed-end elongation of actin filaments in activated talin, with autoinhibited talin masking the ABS1 binding site and preventing association (Lee *et al.* 2004; Ciobanasu *et al.* 2018). ABS2 and ABS3 reside in the rod domains of talin and are required for anchoring with the actomyosin cytoskeleton, providing the tension on talin which enables it to regulate cellular behaviour in a mechanosensitive manner. The C-terminal ABS3 is in the R13 and dimerisation domains of talin with actin filament engagement only possible when talin is a dimer (Gingras *et al.* 2008). Recent research has shown the talin ABS3-actin interaction is required for the initial force generation which promotes focal adhesion assembly (Atherton *et al.* 2015). Finally, talin ABS2 is centrally located, encompassing both the R4 and R8 domains of talin and requires vinculin engagement with the R2 and R3 domains of talin (Atherton *et al.* 2015).

1.8.4. MRL family of proteins

The MRL family of adapter proteins include human members RIAM (Rap1-GTP-interacting adaptor molecule) and its paralog lamellipodin (Lpd), *C. elegans* ortholog MIG-10 and *Drosophila* Pico. MRL proteins are key regulators of actin cytoskeleton

assembly, with roles in lamellipodial protrusion, cell migration and proliferation (Lafuente *et al.* 2004; Lyulcheva *et al.* 2008; Hansen and Mullins 2015; Carmona *et al.* 2016). All MRL members contain a highly conserved centrally located Ras-association domain (RA) next to a coiled-coil region, a pleckstrin homology domain (PH), and a C-terminal proline rich region (Figure 1.12) (Lafuente *et al.* 2004).

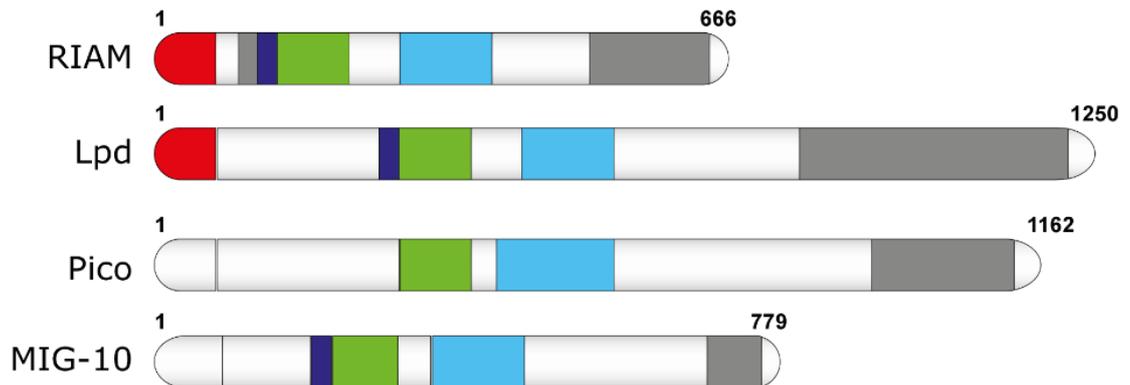


Figure 1.13. Schematic representation of MRL protein composition

Comparison of domain composition between different MRL family members. Included are proline rich regions (grey), Ras-association domain (RA; green), pleckstrin-homology domain (PH; light blue), coiled-coil regions (dark blue) and talin binding sites (TBS; red).

RIAM

RIAM, encoded by the *APBB1IP* gene, is an effector of the GTP-bound Ras GTPase RAP1 required for promotion of integrin activation. RIAM interacts with the actin cytoskeleton regulators profilin and Ena/VASP via six profilin binding motifs (XPPPPP) and six EVH1 motifs (D/E)(F/L/W/Y)PPPPX(D/E)(D/E) respectively (Lafuente *et al.* 2004). Between RIAM and its paralog Lpd, the RA and PH domains remain highly conserved, with the N-terminus and C-terminus of each protein showing increased variability. RIAM binds active Rap1 via its RA domain, promoting RIAM translocation to the plasma membrane (PM) where the PH domain stabilises the interaction with the PM through preferential binding of PI(4,5)P₂ (Wynne *et al.* 2012). RIAM contains two amphipathic α -helix talin binding sites, TBS1 and TBS2, distinguished by a similar motif to the canonical LD-motif distinctive of talin interactors, and has been shown to interact with both the R3 and R8 subdomains of talin (Goult *et al.* 2013a; Chang *et al.* 2014). When near the PM, Rap1 bound RIAM recruits talin to the leading edge of cells via its N-terminal talin binding site (TBS1) and promotes talin mediated activation of β -integrin (Lee *et al.* 2009; Goult *et al.* 2013a).

Lamellipodin

The RIAM paralog Lpd shares a highly similar structure to RIAM but the functional role of this protein is more closely associated with regulation of actin dynamics rather than integrin activation. One of the key roles of Lpd is in supporting lamellipodia and promoting migration. The exact mechanics for how Lpd does this is not yet fully understood, but it is known that direct interaction with active Rac1, actin filament binding and further interaction with actin polymerising proteins Ena/VASP and Scar/WAVE is important for this process (Law *et al.* 2013; Hansen and Mullins 2015; Carmona *et al.* 2016). More recently, it has been proposed that while Lpd is not essential for lamellipodia formation, it does play a key role in stabilising lamellipodia, regulating retraction for increased efficiency in nascent adhesion formation (Dimchev *et al.* 2020). Despite the differences in key roles for Lpd versus RIAM, this protein also contains a talin binding LD-motif at the N-terminus and like RIAM, it can promote integrin activation through this interaction with talin (Lee *et al.* 2009).

1.8.5. The KANK family of proteins

The first discovered KANK protein, originally called ANKRD15 (ankyrin repeat domain 15), was identified as a tumour suppressor in renal cell carcinoma (RCC), whereby reduced expression was found in a number of RCC cells (Sarkar *et al.* 2002). This deletion, or reduced expression, of KANK1 as a promotor of cancer progression has also been identified in other cancer types such as lung and gastric cancer (Chen, Wang and Tong 2017; Gu and Zhang 2018).

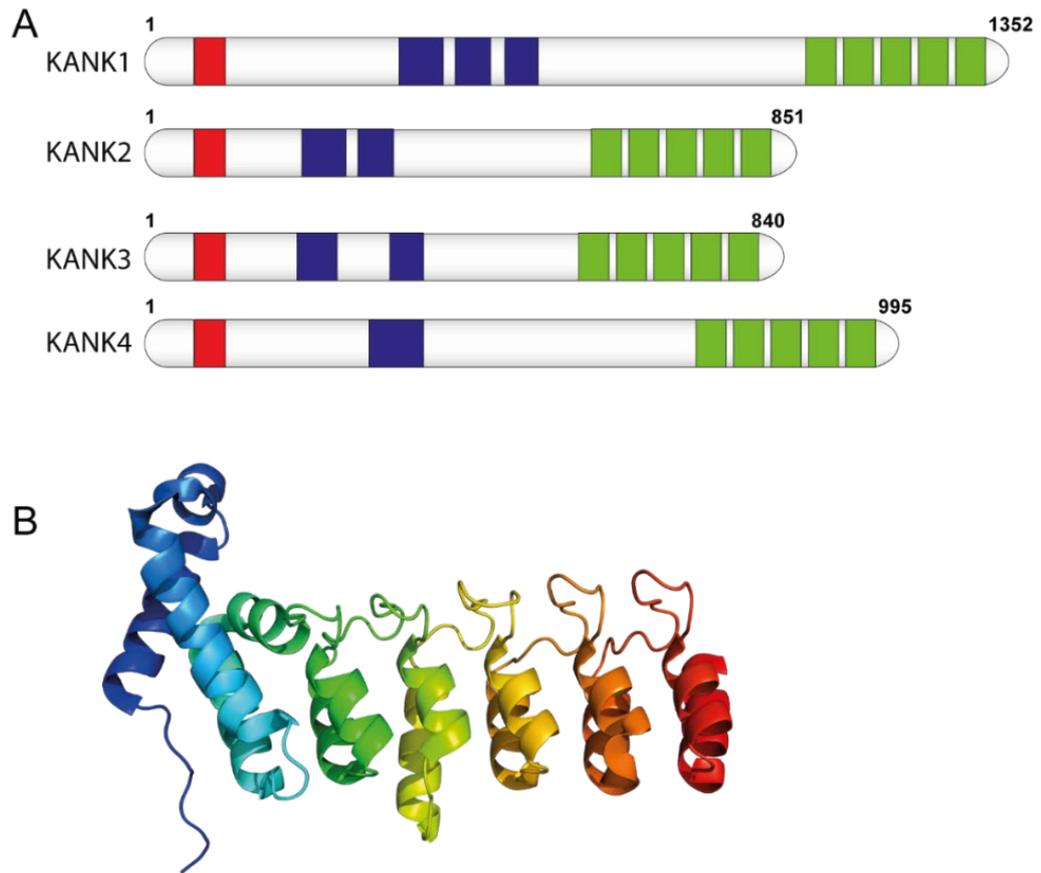


Figure 1.14. KANK Family of Proteins

(A) Comparison of domain positions between KANK isoforms 1-4. C-terminal ankyrin repeats are indicated in green, coiled-coil regions in blue and the N-terminal KN domains shown in red, (B) X-ray crystal structure (PDB: 5YBJ) of C-terminal KANK1 ankyrin repeats, a core structure of KANK proteins.

There are four members of the KANK family (KANK1-KANK4), named 'KN motif and ankyrin repeat domains' due to their distinct five ankyrin repeat structure and highly conserved N-terminal KN domain (Figure 1.13) (Zhu *et al.* 2008). Both the N-terminal and C-terminal regions of KANK proteins are highly conserved over evolution and between the different isoforms, with increased variability towards the central coiled coil domains.

KANK and the cell cytoskeleton

KANK proteins have been shown to regulate both the actin cytoskeleton and the microtubule cytoskeleton. Early work on the *C. elegans* KANK ortholog VAB-19 revealed a crucial role in epidermal actin organisation and distribution during development (Ding *et al.* 2003). KANK directly interacts with insulin receptor substrate p53 (IRSp53) to disrupt binding between IRSp53 and active Rac1, and in doing so KANK inhibits Rac1 dependent lamellipodia formation, cell spreading and membrane ruffling, processes dependent on reorganisation of the actin cytoskeleton (Chandra Roy, Kakinuma and

Kiyama 2009). Furthermore, KANK has been shown to negatively regulate stress fibre formation and actin polymerisation by binding 14-3-3 proteins, an interaction which subsequently inhibits Akt-mediated activation of the GTPase RhoA (Kakinuma *et al.* 2008).

More recent research into KANK's connection to the cell cytoskeleton has increasingly been focused on its role in microtubule targeting. Research on *Drosophila* Kank revealed a direct interaction with the microtubule plus end binding protein EB1 via an SxIP motif, which regulates Kank localisation to microtubule plus ends in *Drosophila* (Clohisey, Dzhindzhev and Ohkura 2014). KANK1 has also been shown to recruit KIF21A to cortical microtubule attachment sites at the cell periphery, a protein which inhibits microtubule plus end polymerisation at the cell edge in order to prevent overgrowth of microtubule filaments (van der Vaart *et al.* 2013).

Cortical microtubule stabilising complex

The KANK family of proteins play a core role in providing a connection between the microtubule cytoskeleton and FAs by initiating clustering of cortical microtubule stabilising complexes (CMSC's) to the periphery of FA complexes (Bouchet *et al.* 2016) (Figure 1.14). This multicomponent complex core includes LL5 β which interacts with PIP3, the liprins α 1 and β 1 known for their interactions with LAR receptor protein tyrosine phosphatase, and ELKS/Rab6-interacting/CAST family member 1 (ELKS) (Serra-Pagès *et al.* 1998; Parnavitane, Stephens and Hawkins 2007). Together these proteins have been shown to drive cellular migration by stabilising lamellipodia at the leading edge and promoting adhesion turnover at the cell rear (Astro *et al.* 2014). Through direct interaction with LL5 β , microtubule minus end tracking proteins called CLASPs, are also recruited to the CMSC complex, and KANK1 binding to liprin- β 1 promotes recruitment of the kinesin KIF21A (Lansbergen *et al.* 2006; Bouchet *et al.* 2016).

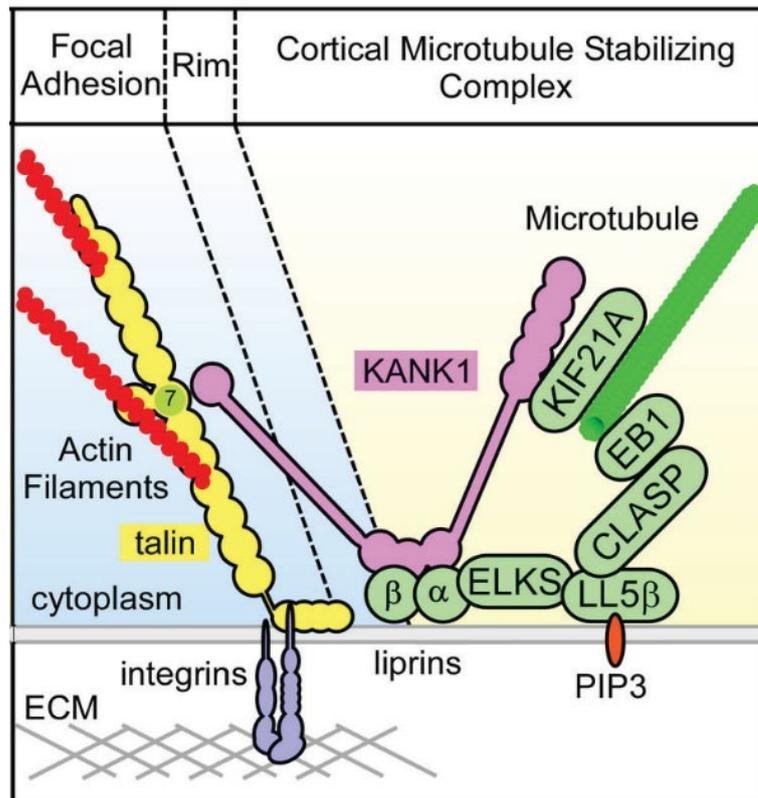


Figure 1.15. Schematic of a cortical microtubule stabilising complex (Bouchet *et al.* 2016)

Complete CMSC assembly showing KANK1 KN domain interaction with talin R7. ELKS, CLASPs, liprins, LL5 β and KIF21A are recruited to sites of complex formation surrounding the periphery of focal adhesions.

KANK interaction with talin

CMSC clustering typically occurs at the cell periphery and surrounds FA sites, but the exact mechanism for this recruitment to the vicinity of FA's has remained elusive until recently. It has now been shown that the KN domain of KANK1 and KANK2 can bind to the R7 domain of talin, leading to disruption of the talin-actin linkage via ABS2 (Figure 1.14) (Bouchet *et al.* 2016; Sun *et al.* 2016). The KANK1-talin interaction has been narrowed down to a modified LDxLLxxL binding motif in the KANK1 N-terminal KN domain and talin R7 domain. Unlike other LD motifs, KANK1 has a double LD motif, with the addition of a negatively charged aspartic acid residue replacing the canonical hydrophobic leucine. This region is predicted to be α -helical, supporting the proposed mode of binding via helix addition which is typical of known talin LD-ligand interactions. Comparison of all 4 KANK isoforms shows high conservation of the N-terminal KN LD containing region (Figure 1.15) and our own unpublished research has shown that talin R7R8 domain can interact with all four isoforms with high affinity. Each of these isoforms have unique differences in behaviour and localisation, suggesting that the function of the talin-KANK interaction may alter under different cellular contexts.

```

KANK1 MAHTTKVNGSASGKAGD I L - - SGD - - QDKEQKDPYFVETPYGYQLDLDLFLKYVDD I QKGN T I KRLNIQKRRKPSVPCPEPRTT -71
KANK2 MAQVLHVPAPFPPTPGPAS - PPAF - - PAKDPDPPYSVETPYGYRLDLDLFLKYVDD I EKGHTLRRVAVQRRPRLSS - - - - - LP -66
KANK3 MAKFA - LNQNLPDLGGPRLCPVPAAGGARSFSSPYSVETPYGFHLDLDLFLKY I EELERGPAARRAPGPPTS - - - - - RRRRAP -70
KANK4 MEKTD - - - - - AKDQS - SQGDEEKDPPKSHPYSVETPYGFHLDLDLFLKYVDD I EKGNT I KR IPIHRRAKQAKFSTLPRNF -75
*   :
      .   * * * * * . : * * * * * . : : : : *   : *

```

Figure 1.16. KANK isoforms KN domain sequence alignment

Amino acid sequence alignment of Human KANK isoforms 1 (Q14678), 2 (Q63ZY3), 3 (Q6NY19) and 4 (Q5T7N3). The double LD motif of each isoform is highlighted in red, with highly conserved region highlighted in blue.

1.8.6. Cyclin dependent kinase 1

The cell cycle and cyclin dependent kinases

For growth and division cells undergo a constant cycle through mitosis, transitioning through four phases starting with G₁, when cells expand, organelles are duplicated, and the cell physically prepares for mitosis. This is followed by the S phase, when chromosomes are replicated, and the G₂ cell checkpoint stage before proceeding through mitosis and cell division.

Cyclin dependent kinases (CDK's) are a family of serine/threonine kinases which regulate transitions through the cell cycle and transcriptional responses, becoming catalytically active upon association with their respective regulatory subunits, cyclins. Cyclin expression and degradation is coordinated in a cyclic fashion according to the stage of the cell cycle, controlling which cyclins are bound to their respective CDK's at different stages of the cell cycle (Figure 1.16). To date, there are ~26 identified CDK genes as determined by the HUGO Gene Nomenclature Committee (Malumbres *et al.* 2009).

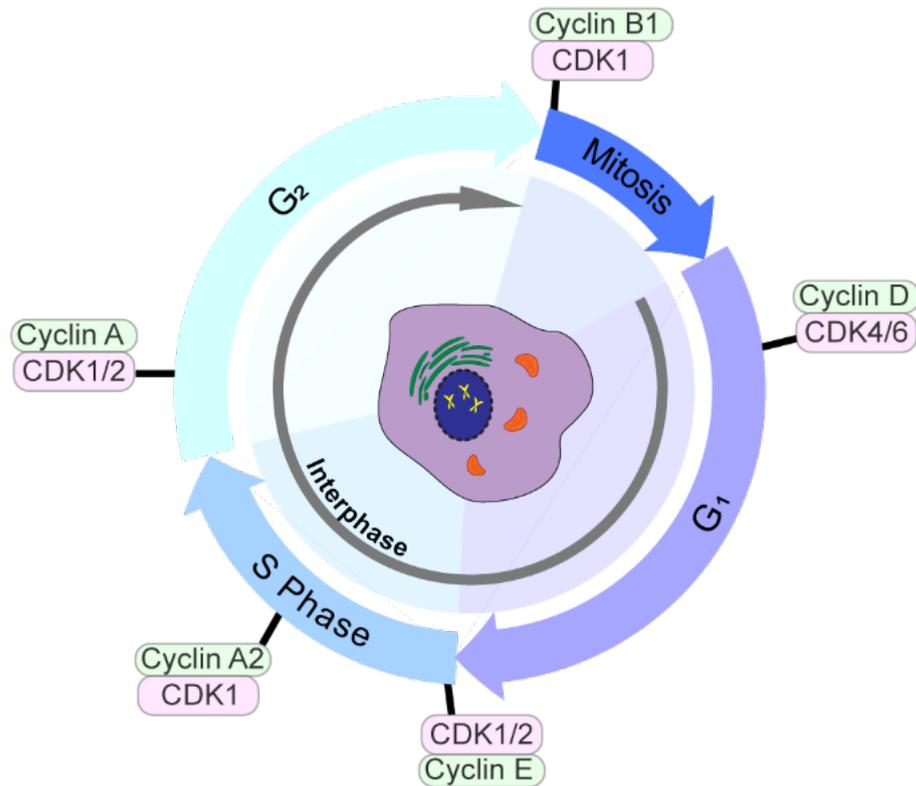


Figure 1.17. Schematic of the cell cycle with CDK and cyclin stages

Stages of the cell cycle and associated activity of the CDK/cyclin complexes. Cyclin D binds CDK4 and CDK6 during G₁, cyclin E binds CDK1 or CDK2 to start cell entry into S phase, cyclin A binds CDK1 and CDK2 during S phase and G₂, and cyclin B binds CDK1 to promote cell transition through mitosis.

Cyclin dependent kinase 1

Cyclin dependent kinase 1 (CDK1) is perhaps one of the most well studied members of the CDK family and was one of the first to be discovered. CDK1 is considered the only ‘essential’ CDK required for cell survival in mammalian cells and is vital for progression through mitosis (Santamaría *et al.* 2007). In the early stage of mitosis, cyclin B levels rise, increasing the amount of active CDK1-cyclin B complexes in the cytoplasm. These active complexes then translocate to the nucleus where it accumulates and triggers several key mitotic processes (Gavet and Pines 2010). One of these processes involves rounding of the cell for division, which requires disassembly of the adhesion complexes. Recently, it has been discovered that CDK1 is a key regulator of this adhesion assembly and disassembly in a cell cycle-dependent process. During interphase, CDK1 in complex with cyclin A2 maintains adhesion complex size and distribution, whereas in G₂, CDK1-cyclin B1 complexes are deactivated via tyrosine 15 phosphorylation, triggering adhesion disassembly (Jones *et al.* 2018). This connection between CDK1 and adhesion regulation is supported by previous research which showed that LAR

phosphatase regulation of CDK1 activity alters focal adhesion formation in mouse embryonic fibroblasts (Sarhan *et al.* 2016).

CDK1 and talin

Cell adhesion, cytoskeletal remodelling and cell cycle progression are intricately connected, with recent evidence demonstrating that CDK1 has a key role in orchestrating this connection (Sarhan *et al.* 2016; Jones *et al.* 2018). The exact mechanisms which underpin this regulation are yet to be discovered, however, recent unpublished research has identified CDK1 as a newly identified ligand for talin, an interaction which, like RIAM and KANK, is mediated via an LD-motif within the CDK1 structure. The interaction site was first narrowed down to the R7R8 domain of talin with an affinity of 15 μ M. It has since been identified that residues 206-223 of CDK1 interact with helices 32 and 33 of talin R8 domain with a K_d of 18 μ M, the same region which mediates interaction with other LD-motif containing ligands such as RIAM and DLC1 (Figure 1.17). Using mutational analysis, CDK1 LD-motif residue D211 has been identified as crucial for the interaction. Mutating residues I210 and D211 (CDK1_2A) disrupts the CDK1-talin R8 interaction.

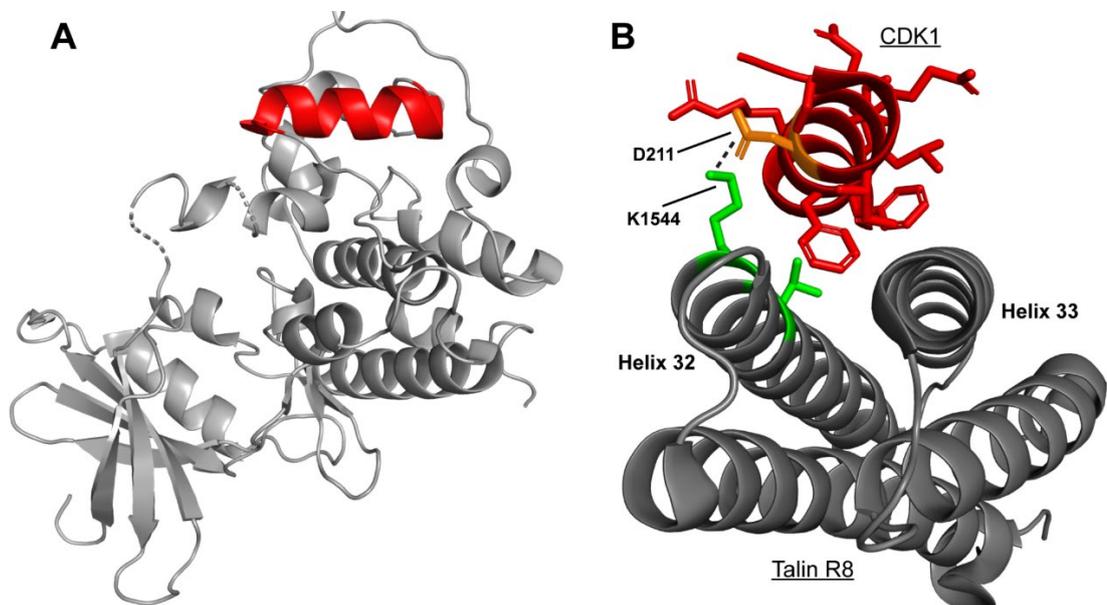


Figure 1.18. CDK1 talin binding region

(A) CDK1 crystal structure with LD containing region in red (PDB ID: 4YC6 (Brown *et al.* 2015) **(B)** Structural model of CDK1 (PDB ID: 4YC6) residues 207-222 (red) interacting with talin R8 helices 32 and 33 (grey). Predicted salt bridge formation between talin K1544 and CDK1 D211 is shown.

1.9. Talin Rod Domain Containing Protein 1 (TLNRD1)

Before 2018, the protein Talin Rod Domain Containing Protein 1 (TLNRD1) was known as Mesoderm Development Candidate 1 (MESDC1) and was identified as one of two genes in an *msd* genomic region on mouse chromosome 7 shown to be crucial for mesoderm development. Initial mouse knockout studies of this *msd* region led to embryos exhibiting failed primitive streak formation and mesoderm development, with apparently normal embryonic ectoderm suggesting a mesoderm specific function (Holdener *et al.* 1994). Later work characterising the genes within this *msd* region identified two genes, subsequently naming them *mesdc1* and *mesdc2* and identifying the location of the human equivalent genes on chromosome 15q25.1 (Wines *et al.* 2000; Wines *et al.* 2001). MESDC1 has a single exon gene encoding a 37 kDa protein which shares 22% similarity to the R7R8 region (residues 1359-1659) of talin. MESDC2 on the other hand encodes a 224 kDa protein, has 3 exons and doesn't show sequence similarities to any other proteins. Following the initial identification of these genes, it was revealed using transgene rescue that only *mesdc2* is able to rescue the *msd* deletion phenotype in mouse embryos, and it was concluded that this is the sole gene responsible for the phenotype therefore was subsequently renamed to MESD (Hsieh *et al.* 2003). MESDC1 shares no known similarities with MESD in sequence, structure or function. This lack of evidence linking MESDC1 to mesoderm development and no known relationship to MESD resulted in the protein being recently renamed to TLNRD1 due to its homology to talin.

Ever since the research identifying MESD as the main factor in the *mesd* deletion phenotype was published, very little research has been done on TLNRD1. Inspection of the proteins expression in the EBI expression atlas (Petryszak *et al.* 2016) shows that TLNRD1 is expressed across most tissues throughout embryonic stages and into adulthood, suggesting a widespread cellular function. In 2010, Gingras *et al.* characterised the protein further, showing that it is predicted to have the same domain topology as R7R8 with a 4-helix bundle inserted between two helices of a 5-helix bundle. They also revealed that unlike R7R8, TLNRD1 was unable to interact with the vinculin VD1 domain but did appear to interact really well with actin filaments showing a much higher affinity than talin R7R8.

1.9.1. TLNRD1 in Cancer

To date, very few papers have been published on TLNRD1 but those that have identify the protein as a target of tumour suppressing microRNAs. In bladder cancer cell lines

TLNRD1 mRNA is increased relative to normal cells and miR-574-3p was identified to reduce cancer cell migration and invasion through interaction with the TLNRD1 3' UTR (Tatarano *et al.* 2012). They also showed that siRNA knockdown of TLNRD1 exhibited the same effects on cell behaviour with a reduction in proliferation, invasion and wound healing indicating reduced migration speeds in 2D. Similarly, in hepatocellular carcinoma cells miR-508-5p also directly targeted the 3' UTR of TLNRD1 with increased expression of miR-508-5p leading to reduced migration, invasion and proliferation (Wu *et al.* 2017). They also showed that knockdown of TLNRD1 reduced proliferation and cell survival, with overexpression having the opposite effect of increasing proliferation, invasion and migration.

This connection between TLNRD1 and cancer is also evident from publicly available online databases showing expression and mutation analyses in cancerous tissues. In the COSMIC database for somatic mutations in cancer (Tate *et al.* 2019) TLNRD1 is identified as overexpressed in up to 468 cancer samples (as of of June 2020), with only 33 showing underexpression relative to normal healthy tissues (Figure 1.18A). This higher TLNRD1 expression is associated with reduced survival in patients with certain cancers such as lung cancer, as demonstrated by the generated Kaplan–Meier plot using data from Affymetrix microarrays (Győrffy *et al.* 2013) (Figure 1.18B). Additionally, there are up to 150 cancer-associated mutations identified in TLNRD1, over 51% of which result in missense mutation which could potentially impair protein function.

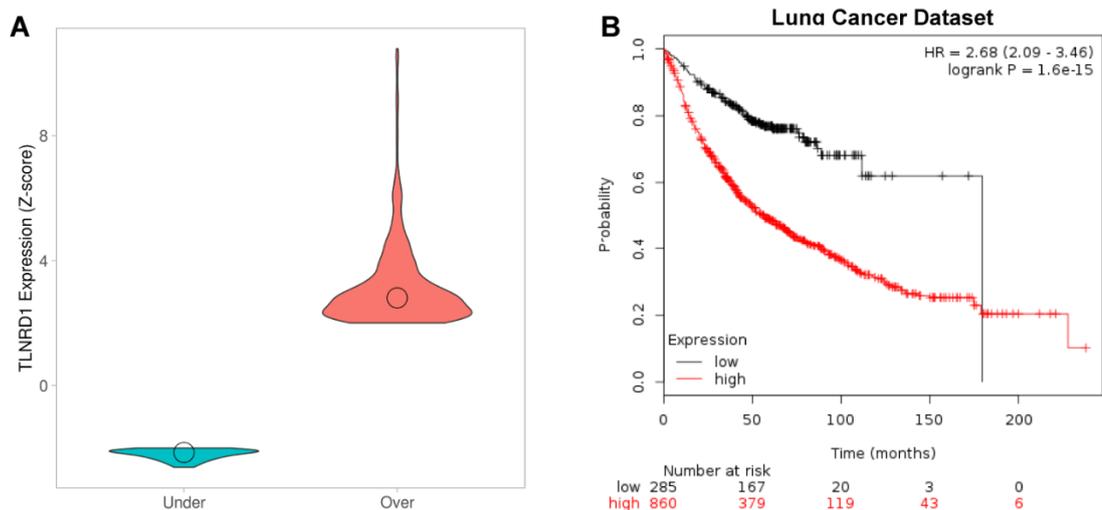


Figure 1.19. TLNRD1 overexpression in cancer and association with poor survival

(A) Violin plot comparing expression levels in cancer with data from the COSMIC database. **(B)** Kaplan-Meier plot showing reduced probability of survival in lung cancer patients with high TLNRD1 expression.

1.10. Project aims

Talin is a key mechanosensitive regulator of adhesion complexes with two known isoforms. Talin-1 and talin-2 are highly conserved in vertebrates and are essential for regulating adhesion complexes in response to extracellular and intracellular cues such as stages of the cell cycle or surrounding matrix stiffness. This leads to an incredible amount of complexity in terms of talins interaction network, with the R7R8 domain in particular behaving as an important signalling hub for mediating interaction with a number of different ligands. What is most intriguing, is the evolution of TLNRD1 which appears to be a replica of this unique R7R8 region of talin. Very little is known about this protein other than it is predicted to have the same domain structure as R7R8 and has been shown to interact with actin filaments. Therefore, the core aim of this study was to explore the intriguing structural and functional similarity between TLNRD1 and talin R7R8, with the hypothesis that TLNRD1 has evolved to interact with the same ligands as R7R8, enabling it to act as either a more mobile equivalent of this signalling hub or behave as a dominant negative regulator of R7R8 signalling. The future aim of this project is to use the structural and biochemical knowledge gained to guide exploration of the proteins cellular function.

1. TLNRD1 is predicted to have the same structure as talin R7R8, therefore the first aim is to explore the structural relationship between the two proteins. X-ray crystallography will be used to determine the structure of TLNRD1, with further characterisation using a range of biochemical techniques and solution structure analysis with SAXS. The information gained will be used to inform further experiments.
2. With TLNRD1's high similarity to R7R8, the second aim was to explore whether TLNRD1 can interact with the same LD-motif containing ligands. To achieve this, biochemical techniques including fluorescence polarisation and NMR spectroscopy will be utilised with the structural information gained from the first aim being used to guide investigations.
3. Previous published research has already demonstrated that full-length TLNRD1 can interact with F-actin *in vitro*. The third aim of this project is to investigate this further using the structural information gained from aim 1, to determine how

TLNRD1 is interacting with actin filaments and why this interaction is stronger in TLNRD1 compared to R7R8.

4. Finally, with talins heavy involvement in cancer progression, the final aim was to explore how cancer-associated mutations identified from publicly available databases alter talin behaviour using biochemical techniques such as fluorescence polarisation and circular dichroism. This will focus on mutations in the R7R8 region of talin, knowledge of which could be utilised to understand how mutations could also change TLNRD1 behaviour.

Chapter 2. Materials and Methods

2.1. Chemicals and Reagents

Name	Company
¹⁵ N Ammonium Sulphate ((NH ₄) ₂ SO ₄)	Cambridge Isotope Lab
InstantBlue™ Protein Stain	Expedeon
BME Vitamins 100X solution	Sigma Aldrich

Table 1. Chemicals and reagents

2.2. Buffers and Media

All buffers and solutions were made with ultrapure Milli-Q grade water (dH₂O).

2M9 Media for NMR

Solution A- 12.5 g Na₂HPO₄, 7.5 g KH₂PO₄, **Make up to 1 L and autoclave**

Solution B- 6 ml dH₂O, 10 ml BME vitamins, 4 g glucose, 1 g N¹⁵H₄Cl, 2 ml MgSO₄ (1M), 0.1 ml CaCl₂, 100 µg/ml ampicillin. **Filter sterilise with 0.2 µm pore filter and add to solution A.**

Antibiotic Stock and Working Solutions

Ampicillin was made up as a 1000x stock in dH₂O (1 g in 10 ml) and diluted to a working concentration of 100 µg/ml. Kanamycin 1000x stock was made up with dH₂O (0.5 g in 10 ml) and diluted to working concentrations of 50 µg/ml.

Buffer Name	Recipe
NiNTA Binding Buffer A	50 mM Imidazole, 500 mM NaCl, 20 mM Tris-HCL pH 8
NiNTA Elution Buffer B	500 mM Imidazole, 500 mM NaCl, 20 mM Tris-HCL pH 8
SP Affinity Buffer A and NMR buffer	20 mM phosphate buffer pH 6.5 (15 mM NaH ₂ PO ₄ , 6 mM NA ₂ HPO ₄), 20 mM NaCl, 2 mM Dithiothreitol (DTT)
SP Elution Buffer B	20 mM phosphate buffer pH 6.5, 1M NaCl, 2 mM DTT
Q Affinity Buffer A	20mM Tris-HCL pH 8, 50 mM NaCl, 2 mM DTT
Q Elution Buffer B	20mM Tris-HCL pH 8, 1M NaCl, 2 mM DTT
Phosphate Buffered Saline (PBS)	137 mM NaCl, 2.7mM KCl, 10 mM Na ₂ HPO ₄ , 1.8 mM KH ₂ PO ₄ pH 7.4
5x Sample Running Buffer	14.8 mM Tris-HCl pH 6.8, 2% SDS, 5% β-mercaptoethanol, 0.005% Bromophenol Blue, 12% glycerol
SDS-PAGE Running Buffer	50 mM MOPS, 50 mM Tris Base, 3.5 mM SDS, 1 mM EDTA, pH 7.7

MES Elution Buffer A	20 mM MES pH 6.5, 50 mM NaCl, 2 mM DTT
MES Elution Buffer B	20 mM MES pH 6.5, 1M NaCl, 2 mM DTT
Crystallisation Buffer	20 mM Tris-HCl pH 7.4, 100 mM NaCl, 2 mM DTT
Actin Co-sedimentation Buffer	10 mM Tris-HCl pH7, 50 mM NaCl, 2 mM MgCl ₂ , 1 mM NaN ₃ , 1 mM DTT

Table 2. Buffer composition

2.3. Plasmid Constructs- Biochemistry

Below is a list of all the constructs used in biochemical experiments with associated details included.

Name	Plasmid	Description	Selection	Source
hTLNRD1_FL	pET151 TOPO	Human TLNRD1 full-length protein	Amp	GeneArt, Invitrogen
hTLNRD1_4H	pET151 TOPO	Human TLNRD1 4-helix domain	Amp	Subcloned
hTLNRD1_4H_2E	pET151 TOPO	Human TLNRD1 4-helix domain with mutations K192E and R189E	Amp	GeneArt, Invitrogen
hTLNRD1_5H	pET151 TOPO	Human TLNRD1 5-helix domain	Amp	Subcloned
hTLNRD1_F250D	pET151 TOPO	Human TLNRD1 F250D dimerisation mutant	Amp	SDM of hTLNRD1_FL
mTalin1_R13DD_L2509P	pET151 TOPO	Mouse talin 1 domains R13 and DD with mutation L2509P	Amp	Dr Ben Goult
mTalin1_R13DD	pET151 TOPO	Mouse talin 1 domains R13 and DD	Amp	Dr Ben Goult
mTalin1_R7R8	pET151 TOPO	Mouse talin 1 domains R7 and R8	Amp	Dr Ben Goult
hTalin1_R7R8	pHisCavitag	Human talin 1 domains R7 and R8	Amp	Prof. Vesa Hytönen
hTalin1_R7R8_R1368W	pHisCavitag	Human talin 1 domains R7 and R8 with mutation R1368W	Amp	Prof. Vesa Hytönen
hTalin1_R7R8_Y1389C	pHisCavitag	Human talin 1 domains R7 and R8 with mutation Y1389C	Amp	Prof. Vesa Hytönen
hTalin1_R7R8_L1539P	pHisCavitag	Human talin 1 domains R7 and R8 with mutation L1539P	Amp	Prof. Vesa Hytönen
mTLNRD1_FL	pET151 TOPO	Mouse TLNRD1 full-length protein	Amp	Dr Ben Goult
mTLNRD1_4H	pET151 TOPO	Mouse TLNRD1 4-helix domain	Amp	Subcloned

Table 3. Biochemistry construct list

2.4. Plasmid Constructs- Cell Culture

Below is a list of all the constructs used in cell biology experiments with associated details included.

Name	Plasmid	Description	Selection	Source
pEGFP-TLNRD1	pGCT51	Mouse TLNRD1 full-length protein with N-terminal GFP	Kan	Dr Ben Goult
pEGFP_N3-TLNRD1	pEGFP-N3	Human TLNRD1 full-length protein with C-terminal GFP	Kan	Subcloned
pEGFP_N3-TLNRD1_4H	pEGFP-N3	Human TLNRD1 4-helix domain with C-terminal GFP	Kan	Subcloned
pEGFP_C1-TLNRD1_4H	pEGFP-C1	Human TLNRD1 4-helix domain with N-terminal GFP	Kan	Subcloned

Table 4. Cell biology construct list

2.5. Molecular Biology

2.5.1. DNA constructs

Full-length human TLNRD1 (TLNRD1_FL) spanning residues 1-362 was purchased as a codon optimised synthetic gene in the pET151 vector from GeneArt (Regensburg, Germany). Using structural prediction, TLNRD1 5-helix domain (TLNRD1_5H) was designed to span residues 1-362 with deletion of residues 142-273 (encompassing the predicted 4-helix domain) and purchased as a codon optimised gene in a pET151 vector from GeneArt. The TLNRD1 4-helix domain (TLNRD1_4H) spanning residues 143-273 was sub-cloned from TLNRD1_FL into a pET151 vector using PCR and restriction digest dependent ligation.

2.5.2. Calcium chloride treatment for making competent E. coli

Bacterial colonies of either DH10 β or BL21(DE3) were streaked and isolated on an antibiotic free Lysogeny broth (LB) agar plate. A single colony was used to inoculate 5 ml LB media and grown o/n at 37°C. The overnight starter culture was then used to inoculate 100 ml of LB media (1:100 dilution) which was grown at 37°C to an OD600 of 0.7. The cell culture was split into 2x 50 ml falcons and rapidly cooled on ice for 10 minutes before centrifugation at 3000 rpm for 10 minutes at 4°C. Pelleted cells in each tube were gently resuspended in 10 ml ice cold calcium buffer (0.1 M CaCl₂, 10% Glycerol) and incubated on ice for 15 minutes. The cells were re-pelleted by centrifugation at 3000 rpm for 10 minutes at 4°C. Final pelleted cells were resuspended in 2 ml ice cold calcium buffer, aliquoted into 0.5 ml Eppendorf tubes (50 μ l per tube) and rapidly frozen with liquid nitrogen before storage at -80°C.

2.5.3. DNA transformations

Plasmid DNA was transformed into DH10 β for DNA amplification and cloning, or BL21(DE3) for protein expression and purification. For each transformation, 2 μ l of DNA was incubated with 50 μ l of cells on ice for 20 minutes. The cells were then heat shocked at 42°C for 1.5 minutes before a further 2-minute incubation on ice. Cells were given a recovery period of 1 hour in 200 μ l LB media at 37°C before plating onto an agar plate containing appropriate antibiotic.

2.5.4. Primer design and Polymerase Chain Reaction

Primers were designed to include the appropriate restriction site for the target vector and meet the following criteria; a minimum GC content of 45%, length of 18-25 bases and melting temperatures (T_m) within the range of 50-60°C with a maximum of 5°C difference in T_m between forward and reverse primers. All PCR reactions for sub-cloning were performed using Hot Start *taq* polymerase (New England BioLabs) in a final reaction volume of 50 μ l using the recipe below.

PCR Mix

0.5 μ l Polymerase
1 μ l Template DNA
10 μ l Buffer (5x stock)
1 μ l Forward and reverse primers
1 μ l dNTPs (10 μ M)
3 μ l $MgCl_2$ (25 mM)
33.5 μ l dH_2O

Each PCR reaction started with pre-incubation at 95°C for 5 minutes, followed by 30 cycles of 92°C for 1-minute, annealing temperature for 30 seconds (5°C below T_m of primers) and 70°C extension at 1 minute per kb of target amplification region. The 30 cycles were completed with a final extension at 68°C for 10 minutes. 1 μ l of PCR fragments with 6x Gel Loading Dye were visualised on a 1% agarose gel containing 0.005% MIDORI^{Green} Advance (New England BioLabs). PCR reactions were then purified with QIAquick PCR purification kit using the supplied protocol (Qiagen).

2.5.5. Double restriction digest and ligation

For vector amplification, target plasmids were transformed into DH10 β cells and grown in 5 ml overnight cultures with the required antibiotic. Plasmid DNA was purified using a Qiagen spin miniprep kit according to the manufacturers protocol. Double digests of both vector and insert were prepared with 2 μ l of each enzyme, 10 μ l compatible digest buffer, 20 μ l of either PCR product/insert or vector and made up to a final volume of 100 μ l with dH_2O . Each digest reaction was incubated at 37°C for 2 hours and purified using agarose gel purification with the QIAquick gel extraction kit (Qiagen) according to the manufacturers protocol.

For ligation of the insert into the target vector, a range of molar ratios were prepared in a 10 μ l reaction with 1 μ l T4 DNA ligase (Promega) and 1 μ l T4 DNA ligation buffer. Ligations were incubated at 4°C overnight, transformed into DH10 β and grown on agar plates containing the required antibiotic for colony isolation, growth and miniprep purification.

2.5.6. Site-directed mutagenesis

All site-directed mutagenesis (SDM) was done using overlapping primers of around 30 bases with the mutation site centrally located and surrounded by the correct target sequence. Primers were designed for a minimum GC content of 45% with C or G bases capping each end and a T_m target of 78°C. For SDM PCR, a final SDM reaction of 50 μ l was used using the recipe below.

SDM Mix

1 μ l Pfu Polymerase
1 μ l Template DNA (50 ng)
5 μ l Pfu Polymerase Buffer (10x stock)
5 μ l Forward and reverse primers
5 μ l dNTPs (10 μ M)
33 μ l dH₂O

Each SDM reaction was initiated by pre-incubation at 95°C for 30 seconds, followed by 18 cycles of 95°C for 30 seconds, 66°C for 1 minute and 68°C at 1 minute per kb of DNA template. After the cycles were complete, a final extension of 68°C for 10 minutes was used. Once the SDM PCR reaction was complete, 1 μ l of the enzyme DpnI was added to each reaction and incubated at 37°C for 2 hours to remove template DNA. Final reactions were directly transformed into competent DH10 β cells and spread on an agar plate containing the appropriate antibiotic.

2.6. Protein Expression and Purification

2.6.1. Protein expression

Starter cultures for expression were prepared by transferring a single colony to 10 ml LB media with 100 μ g/ml ampicillin and grown overnight at 37°C. For each new construct, glycerol stocks were prepared by taking 750 μ l of the starter culture and mixing with 350 μ l filter sterilised 50% glycerol and storing at -80°C. Expression cultures in 750 ml LB

media with 100 µg/ml ampicillin were inoculated with 3 ml of the starter culture and grown to an OD₆₀₀ of 0.7. Protein expression was induced with 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells incubated at either, 18°C overnight, or at 37°C for 3 hours. The induced cells were harvested by centrifugation at 4000 rpm for 20 minutes, the supernatant discarded, and pellets resuspended in 20 ml NiNTA affinity buffer A. When not being used immediately, pellets were rapidly frozen in a 50 ml falcon and stored at -20°C until use.

2.6.2. Cell lysis

Stored pellets were defrosted, and cells were lysed by sonication on ice with an MSE Soniprep 150 sonicator at intervals of 20 seconds on and 40 seconds off x6. Sonicated samples were transferred to 30ml centrifuge tubes and spun at 20,000 rpm for 20 minutes at 4°C. The resulting supernatant was transferred to a clean 50 ml falcon and the pellets discarded. Before purification cell lysates were filtered through a 0.45 µm pore filter.

2.6.3. Protein purification by Ni-NTA column affinity

Most proteins used in this project were labelled with a His₆ polyhistidine tag which has high affinity for Ni²⁺ ions, allowing purification by nickel affinity chromatography. All Ni-NTA column purifications were performed using an ÄKTA start protein purification system (GE Healthcare) at room temperature. Sample loading was done at a flow rate of 3 ml/min and protein elution tracked by measuring absorbance at A₂₈₀. Sonicated cell lysates were loaded onto a 5 ml HisTrap HP column (GE Healthcare). Protein which bound to the column was eluted using an imidazole gradient with increasing amounts of Ni-NTA elution buffer B. After each Ni-NTA purification, a sample of eluted protein was taken for SDS-PAGE analysis. Samples were heated at 95°C for 5 minutes with 1x sample buffer.

Eluted protein was buffer exchanged using dialysis at 4°C overnight, into either SP affinity buffer A for cation exchange purification, Q affinity buffer A for anion exchange purification or PBS for gel filtration. His-tags were cleaved off during overnight dialysis with purified TEV protease (2 mg/ml) (Invitrogen) and protein lysates filtered before further purification using ion exchange chromatography. Columns used included either 5ml HiTrap SP HP cation exchange chromatography column or 5 ml HiTrap Q HP anion exchange chromatography column (GE Healthcare), depending on protein isoelectric

point (pI). The pI of each protein was determined using ExPASy ProtParam tool (Gasteiger *et al.* 2005).

2.6.4. Measuring protein concentration

All protein concentration measurements were performed using a microvolume NanoPhotometer N60/N50 (Implen). Purified protein was measured at a wavelength of 280nm and purity measured with a 260/280 nm ratio. Protein molecular weights and extinction coefficients were determined with the ExPASy ProtParam Tool. Calculated concentrations of the proteins were calculated using Beer-Lamberts Law and converted to molar concentrations.

2.6.5. SDS-PAGE

All protein samples taken for SDS-PAGE analysis were heated at 95°C for 5 minutes in 1x sample running buffer (diluted from 5x stock). All gels were prepared with 4% stacking gel and 12% separating gel in Novex empty 1mm gel cassettes (Invitrogen) according to the manufacturers protocol. Protein gels were stained with InstantBlue dye (Expedeon) for protein visualisation.

2.7. Biochemical Characterisation

2.7.1. Circular Dichroism (CD)

All CD analysis was performed on proteins at a concentration of 0.3 mg/ml in 20 mM phosphate buffer pH 6.5, 20 mM NaCl, 2 mM DTT. Samples were loaded into a cuvette with 1 mm path length in a JASCO J-715 spectropolarimeter equipped with a JASCO PTC-423S temperature control unit. Buffer blanks were taken for each new run with each new batch of buffer.

For secondary structure determination and to assess protein folding, spectra were measured at wavelengths from 190 to 260 nm with a total of four scans. The spectrum of the buffer sample alone was subtracted to remove buffer background from the final protein CD spectra. For protein thermostability analysis, CD measurements were taken at 222 nm at a temperature range starting at 20°C and ending at either 80 or 90°C, over a timescale of 1°C per minute.

2.7.2. Oligomerisation analysis with SEC-MALS

The combination of size-exclusion chromatography with multi-angle light scattering (SEC-MALS) was used to investigate the oligomeric state of TLNRD1 constructs. Size exclusion chromatography allows efficient separation of proteins by hydrodynamic volume, making it easier to distinguish between different oligomeric states of a protein. Multi-angle light scattering measures the scattering of light, known as the refractive index, through a solution using several different angles, which is dependent on molecule size and concentration.

SEC-MALS analysis of different TLNRD1 constructs were performed at room temperature in 20 mM Tris-HCl pH 7, 50 mM NaCl, 2 mM DTT. A total of 100 μ l of protein was loaded onto a Superdex 75 size-exclusion column (GE Healthcare Life Sciences) with a continuous flow rate of 0.75 ml/min and connected to a Viscotek RI detector VE3580 for refractive index detection, and a Viscotek Sec-Mals 9 for detection of laser light scattering (Malvern Panalytical). For each experiment, a BSA control sample was run through the system at 3 mg/ml to test system calibration. Data was collected and analysed using OmniSEC software (Malvern Panalytical) with the BSA sample used to calibrate the analysis.

2.7.3. Nano Differential Scanning Fluorimetry (nanoDSF)

In addition to protein thermostability analysis using CD, further investigations into protein stability was performed over a concentration range in different buffers to determine concentration dependent changes on stability and aggregation. NanoDSF uses measurements of protein intrinsic fluorescence signals from tyrosine's and tryptophan's over a temperature range to determine the folded state and T_m of a protein. Protein measurements using this technique were taken using the Prometheus NT.48 (NanoTemper Technologies). The high sensitivity and small sample volumes required allows easy measurement over a range of conditions.

Purified protein samples were loaded into Prometheus NT.48 Series nanoDSF Grade standard capillaries (NanoTemper Technologies) via capillary action and loaded into the sample tray in a concentration series order. Protein stability of TLNRD1_FL was measured over a concentration range in both crystallisation buffer and NMR buffer (Table 2). Emission wavelength measurements were taken at 350 and 330 nm from 20-95°C with a 3°C/min temperature slope. The ratio between 350 and 330 nm is used to generate

a curve which represents the folded state of the protein over a temperature range. T_m analysis, which is the point at which 50% of the protein has denatured, was performed in PR.ThermControl (NanoTemper Technologies) by determination of the first derivative maximum (Martin, Schwarz and Breitsprecher 2014).

2.7.4. Fluorescence Polarisation

Fluorescence polarisation (FP) allows determination of the binding constant, K_d for small fluorescently labelled ligand peptides against a larger target protein. In solution, the tumbling rate of a short peptide is rapid, but when binding to a larger globular protein occurs, this tumbling rate slows down. The addition of a fluorophore to the peptide allows measurements of the amount of polarised light emitted in one direction, which is dependent on this tumbling rate of the peptide in solution. When a labelled peptide binds to the target protein, the amount of polarised light emitted in one direction increases as the tumbling rate slows down (Figure 2.1). Using a titration of the target ligand against the target peptide allows polarisation measurements to be taken over a concentration range and thus determine the binding constant.

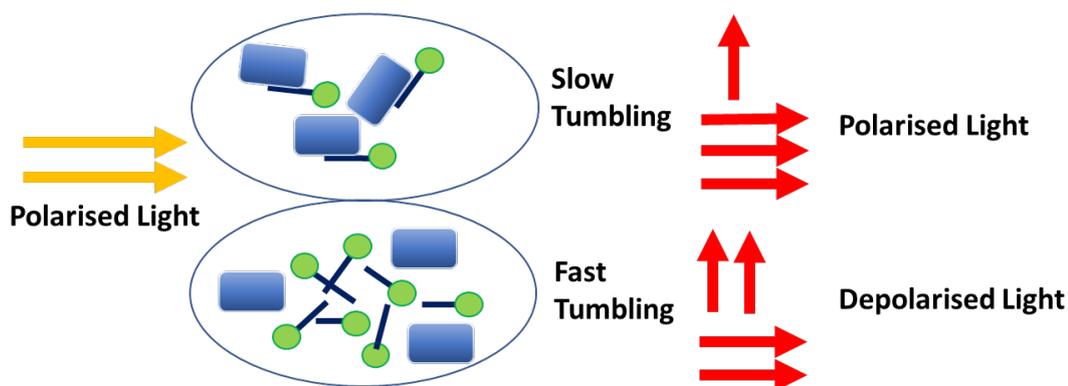


Figure 2.1. Schematic of fluorescence polarisation theory

In the absence of a ligand, the tumbling rate of a fluorophore labelled peptide in solution will be high, so when excited with polarised light the emitted light gets scattered in different directions. When binding occurs, the tumbling rate slows down, reducing the scattering of light and increasing the amount of polarised light emitted in one direction.

For peptide preparation, synthetic peptides were ordered from GLBiochem (China) with either an additional N-terminal or C-terminal cysteine residue for disulphide coupling to either Boron dipyrromethene (BODIPY) or fluorescein. Coupling reactions were prepared with 100 μ M peptide, 25 μ l fluorophore, 5 mM TCEP, 0.05% (v/v) Triton X-100 and made up to a volume of 1 ml with PBS. Reactions were incubated in the absence of light with rotation at room temperature for 2 hours. Final coupled reactions were purified

with a PD-10 column to remove any uncoupled dye from the solution and eluted into PBS. Collected coupled peptide was aliquoted into 50 μ l samples and rapidly frozen with liquid nitrogen before storage at -80°C in a light protected container.

All fluorescence polarisation assays were performed in black Nunc F96 Microwell polystyrene plates (Thermo-Scientific). On each 12-well row, the first well was filled with 100 μ l buffer, followed by 100 μ l of protein in the second well at a concentration between 50-200 μ M, depending on binding affinities. A serial dilution was created of the protein from wells 3-12. Fluorescently labelled peptide was then added to all wells to a final concentration of 100 nM. Fluorescence polarisation values were measured on a CLARIOstar plate reader (BMG Labtech) at 20°C. Excitation and emission wavelengths were adjusted according to the fluorophore used.

Fluorescence polarisation data was analysed with GraphPad Prism v7 using the non-linear regression one-site total binding equation:

$$Y = \frac{B_{max} * X}{K_D + X} + NS * X + Background$$

In this equation, Y is the total binding and X is the free ligand (labelled peptide) concentration, B_{max} is maximum specific binding in the same units as Y, NS is the nonspecific binding slope in Y units divided by X and background is the amount of nonspecific binding with no labelled ligand. This equation allows calculation of the binding constant K_d, which is the ligand concentration required to bind to half of receptor binding sites at equilibrium.

2.7.5. Microscale thermophoresis (MST)

Microscale thermophoresis (MST) is a highly sensitive and powerful technique which allows binding affinity measurements of protein-ligand interactions at nanomolar concentrations. Thermophoresis is defined as the movement of molecules in a solution along a temperature gradient. By heating samples with an infrared laser, MST measures the thermophoresis of a fluorescently labelled protein, a process highly dependent on molecule size or charge. Binding of a protein with a ligand induces changes in size and/or charge which can be measured through changes in thermophoresis.

All MST experiments were performed on a Monolith NT.115 (NanoTemper Technologies). Protein samples were prepared in 20 mM Tris-HCl pH 7, 50 mM NaCl and 2 mM DTT. For analysis of TLNRD1 dimerisation affinity, uncleaved His -tagged

TLNRD1_FL was diluted to a concentration of 100 nM and labelled with His-tag binding RED-tris-NTA dye (NanoTemper Technologies) with incubation at room temperature for 30 minutes. A serial dilution of unlabelled target protein, with the His-tag removed, was created with a starting final concentration of 5 μ M. A final working concentration of 50 nM of the labelled protein was then added to each dilution of the unlabelled protein. All samples were loaded into Monolith NT.115 Capillaries by capillary action and loaded in concentration order in the Monolith sample tray. All experiments were performed at 25°C with infrared laser power at 40%. Experiments were performed in triplicate and analysed using MO.Affinity Analysis v2.3 software (NanoTemper Technologies).

2.7.6. Nuclear magnetic resonance spectroscopy (NMR)

NMR Basic Theory

NMR uses specific atomic nuclei which have a magnetic moment, meaning they give off a magnetic signal which the NMR probe can detect. Most naturally occurring atomic nuclei have a nuclear spin of 0, which means that no magnetic moment is produced, rendering them undetectable to an NMR machine. The atomic nuclei ^1H , which is common in all proteins, and ^{15}N and ^{13}C which are introduced to proteins through media supplementation, all possess a nuclear spin number of $\frac{1}{2}$. This means they produce a magnetic signal which the NMR machine can detect.

Sample preparation

All constructs required for NMR experiments were transformed into BL21(DE3) cells with starter cultures grown overnight in 10 ml of 2M9 minimal media supplemented with filter sterilised solution B. Starter cultures were then used to inoculate (1:100) 500 ml of autoclaved 2M9 solution A, supplemented with the appropriate amount of filter sterilised solution B. Cells were grown to an OD600 of 0.7, induced with 0.4 mM IPTG and incubated at 18°C overnight. Expressed proteins were purified via Ni-NTA column affinity purification and ion exchange chromatography as described above. Final purified proteins for NMR analysis were buffer exchanged into NMR buffer through dialysis, concentrated to the required concentration and 5% (v/v) D_2O was added to the final protein samples before analysis.

1D NMR

All NMR measurements were performed at a temperature of 298K (24.85°C) on a Bruker Avance III 600 MHz spectrometer equipped with a QCI-P CryoProbe. Samples with 5%

D₂O were tested with a final volume of 450 µl in a 3mm Shigemi NMR microtube (Sigma-Aldrich).

Most typical NMR experiments start with a ¹H 1D spectrum measurement which produces signals for all protons in the protein. This allows us to first assess the quality of the protein sample, check that the protein is folded and check that the protein concentration is sufficient for further 2D experimentation. Running a 1D spectrum also allows optimisation of the solvent suppression, which is the background signal created by H₂O. 1D experiments can be used to detect the presence of a ligand in the sample, however, this generates a large number of signals within the 1D spectrum which can overlap, making it difficult to distinguish smaller details. To refine protein-ligand interaction data we use a second signal dimension based on the ¹⁵N nuclei introduced during protein expression.

2D Heteronuclear Single Quantum Coherence (HSQC)

To explore protein-ligand interactions and protein conformational changes, correlative 2D HSQC experiments were used which generates a spectrum of visible peaks for analysis. This approach produces two chemical shift signals from covalently bonded ¹H and ¹⁵N, allowing visualisation of all amino acids in a protein as well resolved peaks, except for proline which doesn't have ¹H-¹⁵N coupling. Almost all amino acids will produce a single peak on the spectrum, with the exception of tryptophan's and histidine's which produce a second peak for the aromatic H^N protons, and asparagine and glutamine which also produce extra signals from the NH₂ side chains.

All 2D experiments were performed using the same conditions as described for 1D NMR but with a ¹H-¹⁵N, HSQC experimental setup. The standard Bruker ¹H-¹⁵N HSQC pulse sequences were performed at 600 MHz Protein-ligand interaction studies were performed over a range of protein:ligand ratios to measure peak shifts. A shift in peak position upon addition of a ligand indicates a change in the chemical environment of the amide in the corresponding amino acid, allowing us to confirm protein-ligand interactions.

2D HSQC Transverse Relaxation-Optimised Spectroscopy (TROSY)

The resolution of a 2D HSQC spectra is highly dependent on the molecular weight of the protein being measured, with a loss of sensitivity and signal broadening with larger proteins (generally over ~25 kDa). To combat this, we can use a TROSY experiment which enhances sensitivity and resolution, allowing HSQC measurement of large proteins.

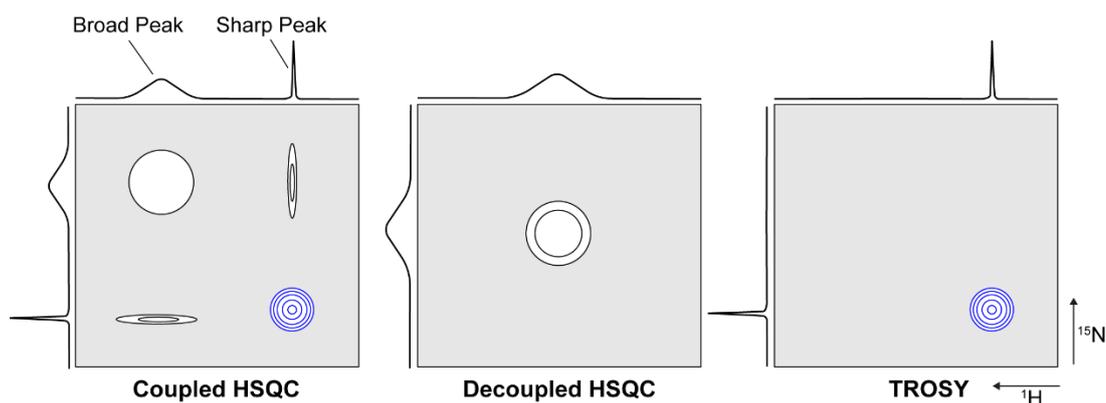


Figure 2.2. Representation of peak selection in a TROSY experiment

Multiplet peaks in a normal coupled HSQC experiment of a large protein. In the left hand panel, there is a broad peak called the 'anti-TROSY' high-field peak (top, left), the bottom right is a single low-field TROSY peak (blue), and the remaining two peaks are semi-TROSY peaks.

A normal HSQC experiment produces a multiplet with four peaks, which are averaged together to produce a single peak. With a smaller protein this multiplet of peaks all produce a sharp signal as relaxation rates remain the same, allowing easy visualisation of each amino acid on the spectrum, however, with a large protein, some of the peaks experience differential line broadening due to different relaxation times, which reduces resolution of the final single peak. The TROSY experiment is optimised to select only the single sharpest peak, or the peak not affected by line broadening, in the multiplet signal (blue, Figure 2.2). This removes influence from the different relaxation times of the broader peaks, therefore increasing resolution of the final spectrum.

Data collection and analysis

Data acquisition was performed using Bruker TopSpin software and analysed with CCPN CcpNmr suite version 2.5.2 (Vranken *et al.* 2005; Skinner *et al.* 2015).

2.7.7. Actin purification and polymerisation

All actin was isolated from rabbit muscle acetone powder kindly gifted by Professor Mike Geeves, and prepared using a cycle of polymerisation and depolymerisation as described previously (Spudich and Watt 1971). A total of 1.5 g of acetone powder was stirred on ice in pre-chilled buffer (10 mM Tris, 1 mM DTT, 0.5 mM Adenosine 5'-triphosphate [ATP], 0.2 mM CaCl₂, pH 8) for 30 minutes. Actin was filtered and the remaining solution spun at 30,000 rpm for 1 hour at 4°C. The actin was polymerised by adding KCl to a final concentration of 100 mM followed by MgCl₂ to a final concentration of 2 mM and stirred at room temperature for 1 hour. The polymerised actin was then pelleted by centrifugation at 30,000 rpm, 4°C for 3 hours. The pelleted actin was

resuspended in depolymerising buffer (5 mM Tris, 1 mM NaN₃, 0.2 mM CaCl₂, pH 7.5) by homogenisation and dialysed into the same buffer overnight at 4°C. The dialysed depolymerised actin was centrifuged the next day at 30,000 rpm at 4°C for 1 hour to remove sediments and resulting supernatant actin diluted to a concentration of 1 mg/ml. At this stage the actin was either flash frozen with 0.2 mM Na₂ATP, 3% sucrose and stored at -20°C or polymerised for further use. For final polymerisation of the diluted actin, ATP was added to a final concentration of 5 µM, followed by KCl to reach 100 mM then MgCl₂ to 2 mM. The actin was stirred at room temperature for 1 hour then pelleted by spinning at 30,000 rpm at 4°C for 3 hours. The polymerised now filamentous actin (F-actin) in the pellets was resuspended in a final actin co-sedimentation buffer (see Table 2). The polymerised actin was stored at 4°C for a maximum period of 1 month.

2.7.8. Actin co-sedimentation assay

For co-sedimentation assays, F-actin was diluted to a minimum of 10 µM and mixed with the appropriate ratio of protein, made up to a final volume of 100 µl with actin co-sedimentation buffer, and incubated for 1 hour at room temperature. Actin binding assays were performed using a high-speed spin method, whereby if the protein of interest binds to the F-actin, it will move to the pellet with the F-actin, whereas if there is no binding, the protein will remain in the supernatant. For the binding assays, samples were spun at 100,000 x g for 20 minutes at 4°C. A total of 50 µl of the supernatant was removed and added to 50 µl of 4x sample buffer. The remaining supernatant was removed, and the pellet resuspended in 50 µl 4x sample buffer and 50 µl actin co-sedimentation buffer. Both supernatant and pellet samples were heated at 95°C for 5 minutes before running equal volumes on an SDS-PAGE gel.

The F-actin bundling assays were performed using a similar method but instead this approach uses a low-speed spin. At low speed the F-actin will remain in the supernatant but if the actin filaments are bundled through addition of an actin bundling protein, then the F-actin will form a pellet with the bundling protein. Actin bundling samples are prepared using the same method as the binding assay and spun at 10,000 x g for 10 minutes at 4°C. Supernatant and pellets were prepared as before and equal volumes loaded onto SDS-PAGE. The density of the resulting bands on SDS-PAGE were analysed using Fiji ImageJ software (Schindelin *et al.* 2012).

2.7.9. Negative stain transmission electron microscopy

To visualise bundled actin filaments using transmission electron microscopy, F-actin was diluted to 25 μM in actin co-sedimentation buffer. Protein samples to be tested were mixed with F-actin at a 1:1 ratio and incubated at room temperature for 1 hour. After incubation and just before grid preparation, the samples were further diluted down to 2 μM with co-sedimentation buffer.

Samples were applied to carbon-coated copper grids for 30 seconds before removal of excess solution. The grids were negatively stained for 1 minute with 2% (w/v) uranyl acetate before excess removal and washing under a stream of ddH₂O. Grid samples were dried before imaging. All images were taken on a JEOL-1230 transmission electron microscope that is equipped with a Gatan One View 16 MP camera with an accelerating voltage of 80kV.

2.8. Structural Biology

2.8.1. Small angle x-ray scattering (SAXS)

Small angle X-ray scattering (SAXS) is a structural technique which has recently gained significant popularity in biomolecular investigations and has had recent rapid advancements in technology. The use of SAXS reveals information about a proteins overall shape in solution, allowing real-time investigations into large conformational changes and protein-protein interactions, as well as providing information on the scale of a protein's flexibility in solution. Despite the low-resolution of data which is obtained through this technique, it remains a hugely complementary structural technique to current crystallographic and biochemical methods.

Theory

SAXS delivers a high-power X-ray beam to a solution containing the target molecules and measures the resulting scattering pattern intensity (I) using a highly sensitive detector (Figure 2.3). SAXS only involves measurement of scattering at small angles up to 1°, as opposed to wide-angle scattering which measures from 5-60° which is generally used for determining the larger crystalline structures of polymers. All molecules in a solution will cause x-rays to scatter which could distort the scattering from the protein of interest, so a lot of care needs to be taken to ensure sample purity and monodispersity. For solutions where multiple states exist, for example a monomer and dimer, SAXS can be coupled with size-exclusion chromatography (SEC) to separate the different components and allow measurements of the different states individually. In this setup

scattering data is continuously recorded of the protein solution under flow as it elutes from the size-exclusion column (SEC) (Figure 2.3).

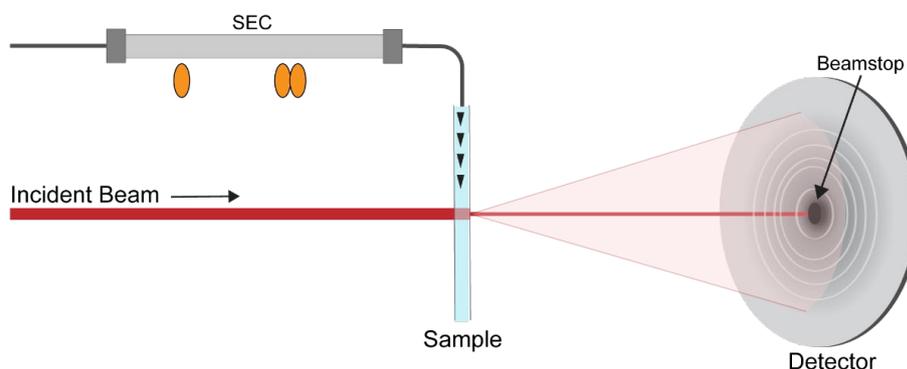


Figure 2.3. Representation of a typical setup for SEC-SAXS

Proteins are first separated by molecular mass on a size exclusion column and scattering of the eluted solution is continuously measured under flow in a quartz capillary. The solution is targeted with a high-power incident X-ray beam and the X-ray scattering pattern is measured at small angles. The beamstop prevents any remaining unscattered X-rays from hitting and damaging the detector, creating a shadow effect in the centre of the detector.

Data collection

Scattering data is collected in a continuous series of short exposures which are checked for radiation damage before being combined in a process of data reduction, which yields a typical 1D averaged SAXS curve. All SAXS data was collected at Diamond Light Source beamline B21 (Didcot, UK) with a Pilatus P3-2M silicon hybrid pixel detector (Dectris). SAXS experiments were combined with size-exclusion chromatography using a KW-403-4F 10-700 kDa column (Shodex) connected to an Agilent 1200 HPLC system. Proteins were tested at room temperature in different concentrations ranging between 2-10 mg/ml in 20 mM Tris-HCl pH 7, 50 mM NaCl, 2 mM DTT.

Data analysis

Raw data was automatically processed on site and data analysis was performed within open access software ScÅtter (version 3.0) available from <http://www.bioisis.net>. Buffer subtraction was performed using the buffer signal area that best matches the baseline of the protein peak signal on the size-exclusion column chromatogram. For further analysis averaged files were used.

Guinier analysis was used to generate an approximation of the SAXS curve and an estimation of the R_g and $I(0)$ values, with removal of any non-linear points with a q value at $q \times R_g$ at 1.3 or below. Using a Guinier plot also allows inspection for signs of aggregation in the sample and initial data quality. Generation of a normalised Kratky plot

based on the R_g value was used to inspect protein globularity or look for any protein unfolding. SAXS data can also be analysed to reveal protein flexibility by comparing earliest plateau point with increasing q max between a Porod-Debye plot (rigid), SIBYLS plot (partially flexible) and Kratky-Debye plot (flexible/unfolded). Finally, a volume-of-correlation (VC) plot was used to determine quality of buffer subtraction and $P(r)$ distribution optimised which gives an indication of the shape of the molecule.

With the data quality optimised and data parameters checked, DAMMIF available from ATSAS online web server, was used to generate an *Ab initio* single phase dummy atom bead model of the overall shape of TLNRD1, averaging 15 independent runs with P2 symmetry. In addition, GASBOR was used to create a dummy residue model to account for the number of amino acids in the protein.

2.8.2. X-ray crystallography

Theory

The use of X-ray crystallography offers the opportunity to solve protein atomic structures at incredibly high resolution. Knowing the detailed structure of a protein complements the use of solution based structural techniques such as SAXS and NMR and provides unique insights into proteins function and regulation. While all molecules will diffract X-rays, the ability for molecules to diffract X-rays in solution is weak which is why structural data obtained from solution-based techniques such as SAXS only offers limited structural information on a protein. To overcome this limitation, proteins can be encouraged to form crystals which produces lattices of protein molecules in a fixed orientation. This fixed orientation produces a distinct pattern of diffracted X-rays with greater intensity than SAX scattering. This pattern, described as a 'reflection', can be extrapolated to determine the three-dimensional structure of a protein.

Crystallisation screening

Protein crystallisation has three key stages; nucleation, growth and the cessation of growth. The nucleation process which initiates crystal formation requires a stable ordered clustering of molecules. This is promoted by pushing the balance of the molecule in a solution to a point between supersaturation and precipitation and achieving this balanced nucleation point varies greatly from protein to protein. In order to achieve the correct nucleation conditions for the protein of interest, a range of crystallisation conditions can be tested in a screening assay which involves varying the concentration of precipitants, additives and the protein itself, as well as pH and temperature.

All crystallisation trials for TLNRD1 were prepared using a hanging drop vapour diffusion method. Screening kits used included JCSG-plus, PACT Premier, Hampton crystal screen HT and Wizard classic I. A total 100 μ l of screen buffer solutions were plated in a 96-well plate.

Crystal optimisation

Optimisations of promising conditions were scaled up to 500 μ l in a 24 well-plate with 1 μ l drop ratios of protein to buffer in the range of 1:1, 0.5:1 and 1:0.5. Depending on the conditions, either the concentration of the salt or buffer pH was varied, and where possible precipitant percentage was also varied.

TLNRD1 4-helix domain crystal structure

TLNRD1 4-helix domain was purified using MES buffers as a replacement for phosphate, buffer exchanged into crystallisation buffer and concentrated to a final concentration of 5 mg/ml (350 μ M). Crystallisation screening was set up an equal volume of protein and well solution in 0.1 μ l drops with a Mosquito liquid handling robot (TTP Labtech). Crystals were achieved with the JCSG-plus screen after 4 days at 21°C in 200 mM ammonium citrate dibasic and 20% w/v PEG 3350. Optimisations were performed with a range between 50-300 mM ammonium citrate dibasic and 15-30% PEG 3350. Optimal crystals were grown after 4 days at 21°C in 300 mM ammonium citrate dibasic and 20% PEG 3350.

TLNRD1 full-length protein crystal structure

As with the TLNRD1 4-helix domain, TLNRD1 full-length protein was purified with MES buffers and dialysed into crystallisation buffer with a final concentration of 14.7 mg/ml (390 μ M). Crystal screens were also prepared in the same manner as the TLNRD1 4-helix above. Small crystals were visible after 5 days at 21°C in 200 mM sodium thiocyanate at pH 6.9 with PEG 3350. Optimisation screens were prepared with a sodium thiocyanate concentration ranging from 50-300 mM and PEG 3350 ranging 15-30%. Final optimised crystals were obtained in 300 mM sodium thiocyanate and 20% PEG 3350 after 1 week.

Mounting

Collection of X-ray scattering data requires diffraction images to be collected at multiple angles. Due to the high power of the X-ray beam, crystals can quickly become damaged and superheated. To combat this and prolong their lifespan, crystals are rapidly frozen and data collection performed under a continuous stream of liquid nitrogen. To protect the crystals from the freezing process, crystals are first prepared with a cryo-protectant and rapidly frozen. Both TLNRD1 full-length and 4-helix domain crystals were mounted

on CryoLoops (Hampton Research) or LithoLoops (Molecular Dimensions) with cryo-protectant containing the optimal crystal buffer condition and 25% ethylene glycol. Mounted crystals were rapidly frozen directly in pucks with liquid nitrogen in preparation for transportation.

Data collection

X-ray diffraction data of both TLNRD1 full-length protein and the 4-helix domain was obtained at Proxima-1 beamline at Soleil synchrotron (Paris, France) at 100K. Data was initially processed using the autoPROC pipeline (Vonrhein *et al.* 2011) which includes XDS for integration of spot intensities (Kabsch 2010), AIMLESS for reflection scaling (Evans and Murshudov 2013) and TRUNCATE for data merging which improves data quality (Evans 2011).

Molecular Replacement

Over recent years, the amount of publicly available structural data has grown exponentially. For some proteins, a homologous or similar related protein may have already had its structure solved and be available on online databases. These readily available structures can be used to estimate the phases of the new related structure in a process called molecular replacement.

For resolving the TLNRD1 4-helix domain structure, molecular replacement with the talin rod domain R7R8 structure (PDB ID: 2X0C; (Gingras *et al.* 2010)) as the search model was performed using PHASER followed by manual adjustment with COOT (McCoy *et al.* 2007; Emsley *et al.* 2010). Resolving the structure of the TLNRD1 full-length protein could not be fully resolved using this method, so BALBES molecular replacement timeline (Long *et al.* 2008) was used as an alternative using the same R7R8 structure. This approach was used to generate an initial model which had to be manually tweaked before manual adjustment in COOT and structure refinement with PHENIX.REFINE (Afonine *et al.* 2012). To check the interaction properties of the revealed dimerisation interface in both TLNRD1 structures, an assessment of dimerisation likelihood was performed using PISA (Krissinel and Henrick 2007). Final resolved structures and figure generation was performed using PyMOL version 2.3.3 (Schrödinger, LLC). All models were validated with MOLPROBITY (Chen *et al.* 2010) before deposition in the RCSB Protein Data Bank.

2.9. Cell Culture

2.9.1. Antibodies

Primary Antibodies	Dilution	Supplier/Source
MESDC1 Polyclonal Antibody	1:1000	ThermoFisher, Cat. No. PA5-70832
Raised Anti-TLNRD1 Polyclonal Antibody*	1:1000	Capra Science
Secondary Antibodies		Supplier/Source
IRDye 800CW Donkey Anti-Rabbit IgG	1:5000	Li-Cor, Cat. No. 926-32213

Table 5. Primary and Secondary Antibodies

* Polyclonal antibody raised by Capra Science (Sweden) in rabbit against recombinantly expressed full-length human TLNRD1 protein (residues 1-362). Full-length TLNRD1 was expressed as described above with nickel-affinity chromatography, cleavage of His⁶ tag via TEV, ion exchange chromatography and a final purification step using gel filtration.

2.9.2. Cell lines and culture maintenance

U2OS osteosarcoma cells were grown and maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% foetal bovine serum (FBS), 2 mM L-Glutamine, Penicillin and Streptomycin. All cells were cultured in 10 cm plastic dishes at 37°C with 5% CO₂.

For passaging and culture maintenance, cells were passaged when 70-90% confluency was reached. For each split, cells were washed 1X with PBS and incubated with the appropriate volume of pre-warmed Trypsin-EDTA for 5 minutes at 37°C. Once cells detached, 4 ml of warmed DMEM was added to cells to deactivate the trypsin. Appropriate number of cells were then added to a clean 10 cm dish (depending on requirements) and final volume increased to 10 ml with warmed DMEM.

2.9.3. Cell plasmid transfections

For each DNA construct to be transfected, 250 µl of optiMEM media was incubated with 2 µg of plasmid DNA and 5 µl of Lipofectamine P3000 reagent (mix 1) at room

temperature for 10 minutes. To a separate tube (mix 2), 5 μ l of Lipofectamine 3000 reagent was added to 250 μ l of optiMEM and incubated for 5-10 minutes. After incubation periods, 250 μ l of mix 2 was added to mix 1, mixed and incubated at room temperature for 5 minutes. The final 500 μ l was added either directly to adhered cells grown to 80% confluency in a 6-well plate or to cells in suspension which were transferred to a 6-well plate. Transfected cells were used up to 48 hours after transfection before being discarded.

2.9.4. siRNA silencing

All siRNA silencing of U2OS cells were performed using three TLNRD1 targeting siRNA's and nonsilencing Allstars negative control siRNA (QIAGEN). The siRNA's used for silencing of TLNRD1 were purchased from QIAGEN (siTLNRD1 #6 = Hs_MESDC1_6 FlexiTube siRNA Catalog No. SI04217605; siTLNRD1 #7 = Hs_MESDC1_7 FlexiTube siRNA Catalog No. SI04314569; siTLNRD1 #8 = Hs_MESDC1_8 FlexiTube siRNA Catalog No. SI04362820). 5 μ l of each siRNA was incubated with 250 μ l optiMEM media for 5 minutes at room temperature (mix 1). For siRNA transfection reactions, 5 μ l of Lipofectamine RNAiMAX transfection reagent was incubated with 250 μ l optiMEM per reaction for 5 minutes at room temperature (mix 2). After the initial incubation period, 250 μ l of mix 2 was added to each siRNA mix 1 and incubated at room temperature for 10 minutes. The total 500 μ l was each siRNA mix added to wild-type U2OS cells plated in a 6-well plate. Cells were used two days after transfection and discarded after four days.

2.9.5. Cell lysis

For all immunoprecipitation experiments cells were grown to 90% confluency in a 10cm dish. Cell culture media was aspirated, and cells washed 3x with ice cold PBS. To each 10 cm dish, 500 μ l of ice-cold cell lysis buffer (40 mM HEPES-NaOH pH7.4, 75 mM NaCl, 2 mM EDTA, 2% NP40) with protease inhibitor and phosphatase inhibitor cocktail, was added on top of the cells. Using a scraper, cells were removed from the dish and placed in a clean Eppendorf tube. Samples were centrifuged at 13,000 rpm for 5 minutes at 4°C and supernatant transferred to a clean tube. Final lysate samples were used immediately.

2.9.6. Immunoprecipitation

Protein immunoprecipitation experiments were performed using Dynabeads Protein G superparamagnetic beads (Invitrogen). 30 μ l of beads were added to a clean Eppendorf and washed with 800 μ l of PBS. Using a magnetic stand to hold the beads in place, the PBS was removed and fresh 800 μ l of PBS was added followed by 1 μ g/ml of anti-TLNRD1 rabbit antibody (raised by Capra Science) or mouse IgG control antibody. The beads were incubated with the antibody for 1 hour on rotation at room temperature. A total of 500 μ l of the cell lysates were added to the beads and incubated on rotation for 3 hours at 4°C. Following lysate incubation, a magnetic stand was used to hold the beads in the tube while the unbound lysate fraction was removed and 60 μ l samples taken for SDS-PAGE. The beads were washed with 800 μ l ice-cold PBS X3 before resuspension in 60 μ l of sample buffer.

2.9.7. Western blot

From the IP experiments, both cell lysates and bead bound fraction (25 μ l volume) were loaded onto an SDS-PAGE gel with a 4-20% gradient and run at 120 V for 1.5 hours. Transfers of the gels were done with a Bio-Rad Trans-Blot Turbo nitrocellulose transfer packs and Trans-Blot Turbo transfer system. The membrane was blocked at room temperature for 1 hour with SuperBlock (PBS) blocking buffer (ThermoScientific). The membrane was then incubated on rotation with 10 ml blocking buffer containing 1:1000 dilution of anti-TLNRD1 primary antibody at 4° overnight. Next day, the membrane was washed five times with PBS before 1 hour rotating incubation at room temperature with 10 ml blocking buffer and 1:5000 anti-rabbit 800 IgG secondary antibody.

2.9.8. Immunofluorescence- fixed confocal microscopy

Coverslip preparation

For each sample to be tested, 35 mm glass bottom dishes were used with a 14 mm micro-well and #1.5 cover glass. Each dish was coated with 1 ml poly-d-lysine diluted 1:1000 in PBS and incubated at 37°C for 2 hours. The poly-d-lysine was removed and 1 ml 10 μ g/ml fibronectin in PBS added with 1-hour incubation at 37°C. Following this, the fibronectin solution was removed and dishes washed 1x with PBS before addition of 2 ml of cells in supplemented DMEM media, and incubation for a further 2 hours at 37°C. Cells requiring transfection were transfected with either pEGFP_N3-TLNRD1, pEGFP_N3-TLNRD1_4H or pEGFP_C1-TLNRD1_4H the day before fixing. Finally, the cell culture media was removed, cells gently washed with PBS and incubated with 1.5 ml 4%

paraformaldehyde (PFA) at room temperature for 10 minutes. The PFA was removed and dishes washed 2x with PBS before final storage in PBS at 4°C.

Imaging

Fixed-cell imaging was performed on a 3i CSU-W1 spinning disk confocal microscope (intelligent Imaging Innovations) with a Hamamatsu scientific CMOS Orca Flash 4 v2 camera. Cells were imaged using either a 40x Zeiss LD C-Apochromat water immersion objective or 63x Zeiss Plan Apochromat oil immersion objective, with 100% 488nm GFP laser.

2.9.9. 2D Random Migration Assays

Cells were seeded at a density of 5×10^3 in 1 ml media on a plastic 24 well plate which was coated with 10 $\mu\text{g/ml}$ fibronectin and incubated for 2 hours at 37°C, 5% CO₂. The experiment was performed on a Nikon Eclipse Ti2-E widefield microscope with a heated CO₂ chamber, Hamamatsu scientific CMOS Orca Flash 4 v4 and 10x Nikon CFI Plan Fluor objective. Random migration of cells was measured over 24 hours in a time-lapse movie with images taken every 10 minutes. Tracking was performed in single cells with little or no cell-cell contact and dividing cells were avoided. ImageJ software available from Fiji was used for analyses with the MTrackJ plugin for manual tracking of cells. The tracked data was analysed using freely available Ibidi chemotaxis and migration tool to determine migration speeds, directionality and distance. Graphs of resulting data were produced using PlotsOfData (Postma and Goedhart 2019).

Chapter 3. Structural Characterisation of TLNRD1

In this chapter, the structure of TLNRD1 will be explored using a range of complementary techniques. From initial secondary structure analysis with circular dichroism, through to X-ray crystal structure analysis and solution structure analysis with SAXS and SEC-MALS to explore protein dynamics. To date, there is no published structural information on TLNRD1 apart from that which is predicted through sequence alignment with talin R7R8. This chapter reveals the striking features of TLNRD1 which share similarities with R7R8 as well as those which differentiate TLNRD1 from this distinctive region of talin.

3.1. Sequence analysis and secondary structure prediction

A crucial first step in this project is to understand the structure of TLNRD1 using evolutionary analysis and structural prediction, which will aid understanding of protein behaviour and guide initial construct design for biochemical analysis and later structural analysis. The high similarity with talin R7R8 was used to guide structural understanding of TLNRD1.

3.1.1. TLNRD1 shares sequence similarity with talin R7R8 and has a similar predicted secondary structure

Tlnrd1 encodes a 362 amino acid protein of unknown function with a molecular weight of 37.7 kDa. TLNRD1 was first characterised through its significant sequence identity with domains R7 and R8 of talin (residues 1359-1659) (Gingras *et al.* 2010). A BLAST search of the NIH GenBank genetic sequence database with the *Tlnrd1* gene identifies both talin-1 and talin-2 as the highest scoring hits aside from TLNRD1 itself, with specific alignment to the R7R8 regions of talin. Sequence alignment of human TLNRD1 amino acid sequence against talin-1 and talin-2 using Clustal Omega shows a high sequence identity of 22% (Figure 3.2).

The talin R7R8 region is a 9-helix module with an unusual domain topology, consisting of two domains, a 5-helix bundle (R7) with a 4-helix bundle (R8) inserted between two helices of the R7 domain (Figure 3.2A). Secondary structure prediction of TLNRD1 with Phyre2 protein fold recognition server (Kelley *et al.* 2015) suggests that TLNRD1 has the same 4 and 5 helix domain structure but has an additional unstructured region at the N-terminus spanning ~40 residues (Figure 3.2B). Alignment of predicted helix boundaries with those of talin R7R8 confirms that TLNRD1 may mimic the talin structure (Figure 3.2C). This significant similarity in predicted structure and sequence suggests TLNRD1 arose from a gene duplication event from the talin gene.

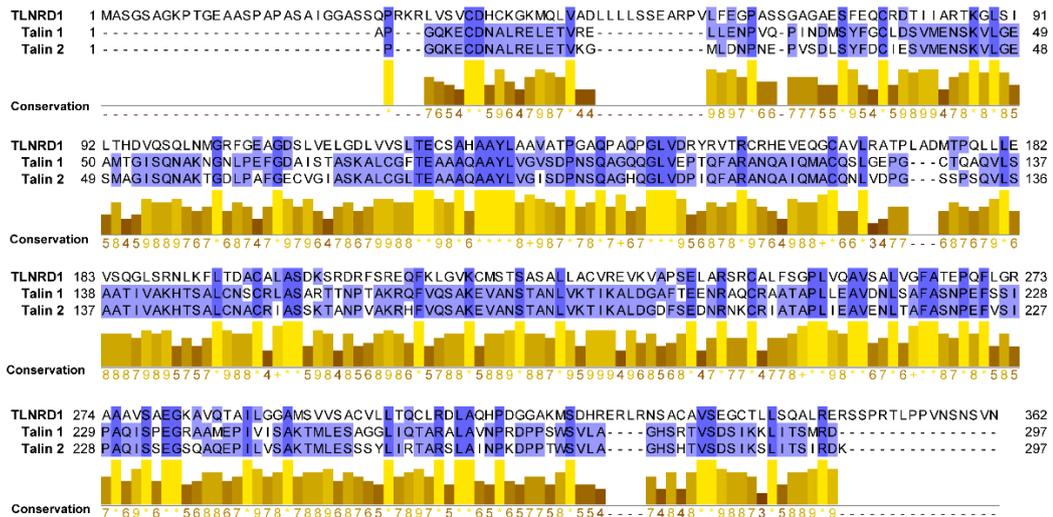


Figure 3.1. Sequence alignment of TLNRD1 with talin-1 and talin-2

Amino acid sequence alignment of human TLNRD1 (UniProt ID: Q9H1K6), talin-1 (Q9Y490) and talin-2 (Q9Y4G6) using Clustal Omega alignment in Jalview. Residues coloured according to % similarity, darker blue indicates high conservation. Conservation of each residue is also indicated below the alignment.

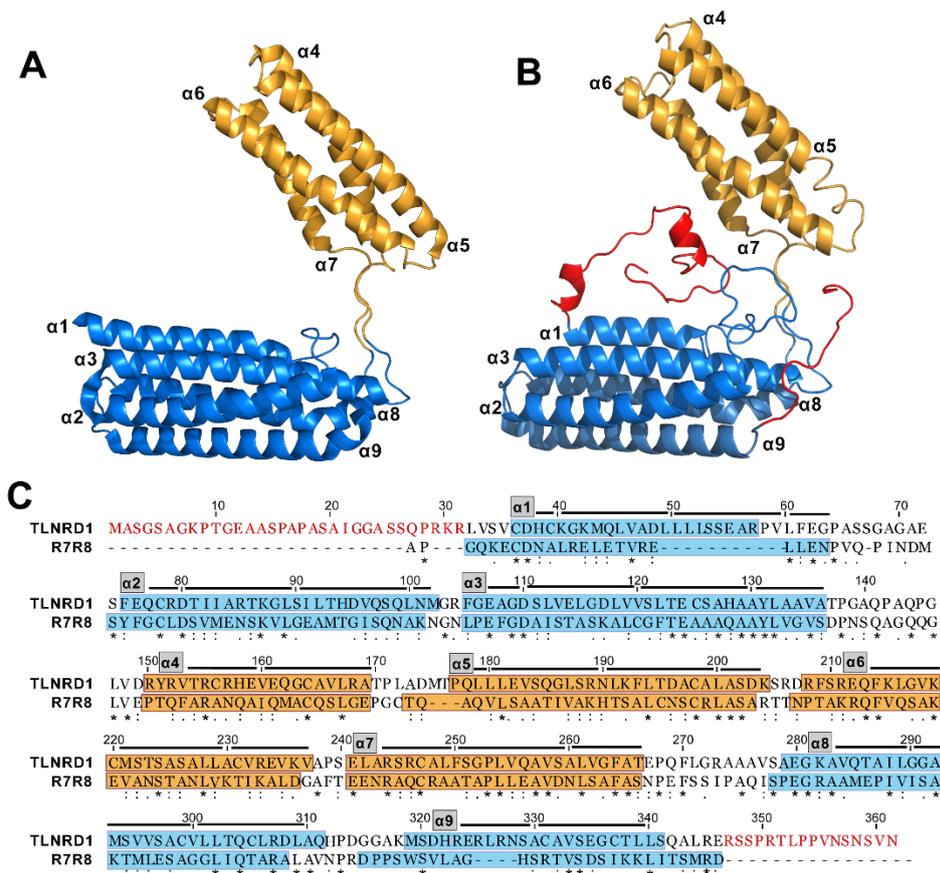


Figure 3.2. Talin R7R8 and TLNRD1 domain structure comparison

(A) Domain structure of talin R7R8 (PDB ID: 2XOC (Gingras *et al.* 2010)) showing the 4-helix R8 domain insertion (orange) between two helices of R7 5-helix domain (blue). (B) TLNRD1 Phyre2 structure prediction showing unstructured regions (red), 4-helix domain (orange) and 5-helix domain (blue). (C) TLNRD1 and R7R8 human sequence alignment from Clustal Omega with domain boundaries overlaid with corresponding colours from A and B. Helix numbers are indicated.

3.1.2. *TLNRD1* has an intriguing evolutionary history

The emergence of adhesion signalling occurred early in evolutionary history from unicellular nonbilaterian lineages (simple organisms lacking bilateral symmetry) but have diversified and shown independent losses throughout evolution. The first talin gene to be identified was *TLN1* (talin-1), with later discovery of *TLN2* (talin-2) which is thought to have arisen through gene duplication of talin-1 in the chordate lineage (animals distinguished by development of a notochord) (Senetar and McCann 2005). Both talin-1 and talin-2 have distinct functions despite sharing 88.4% similarity (76% identity), with talin-1 function more associated with focal adhesion regulation whereas talin-2 is not targeted to adhesion complexes. Altogether, *TLN1* and *TLN2* have up to 56 exons which are conserved throughout evolution, whereas the mammalian *Mesdc1* gene contains a vastly different gene structure having just a single large exon which encodes the whole protein. Only ~3% of the human genome encodes single exon genes, with a large proportion encoding key proteins for mammalian development and cell proliferation (Grzybowska 2012).

Exploration of *TLNRD1* homologues using an amino acid sequence BLAST search against the NIH GenBank database with full-length *TLNRD1* reveals that *TLNRD1* is highly conserved throughout vertebrate evolution. The *TLNRD1* gene first appears in *Salpingoeca rosetta* which is a species of choanoflagellate, organisms which are believed to be the closest living unicellular relatives to multicellular organisms (Brunet and King 2017). Choanoflagellates contain homologs of well characterised adhesion and cytoskeletal proteins such as cadherin and microtubules (King 2003; Karpov and Leadbeater 1998), with *S. rosetta* being adopted as a classic model organism for studying the early evolution of multicellularity (Hoffmeyer and Burkhardt 2016). A BLAST search of the talin-1 amino acid sequence in choanoflagellates reveals that they also have talin, however, it has been shown that choanoflagellates lack other key integrin adhesion machinery including integrins themselves and adhesion regulators FAK, ILK and paxillin (Sebé-Pedrós *et al.* 2010).

The *TLNRD1* gene is also present in the sponge *Amphimedon queenslandica*, organisms that are considered the oldest living metazoan lineage. Sponges express a number of proteins which regulate classic hallmarks of multicellular life, from cell-matrix adhesion regulators to proteins which regulate developmental signalling, making the species a good model for studying the earliest evolution of these complex processes (Srivastava

et al. 2010). What is most striking about TLNRD1 evolution is that while it is highly conserved in vertebrates and present in sponges and choanoflagellates, the gene is absent in nematodes, Cnidaria and arthropods (Figure 3.3). Independent losses in certain lineages have also been observed for other adhesion related genes as shown by the loss of integrins in choanoflagellates and fungi (Sebé-Pedrós *et al.* 2010). After these losses, the TLNRD1 gene reappears in echinoderms and remains highly conserved through to *Homo sapiens*. Functionally this suggests that while TLNRD1 is not essential for all multicellular life, nor is it essential for core adhesion related processes like talin is, it may play a more specific role which is unique to vertebrate development. Alignment of TLNRD1 amino acid sequences from mammals to choanoflagellates reveal high sequence conservation and preservation of hydrophobicity, showing little change over the course of its evolution (Figure 3.4).

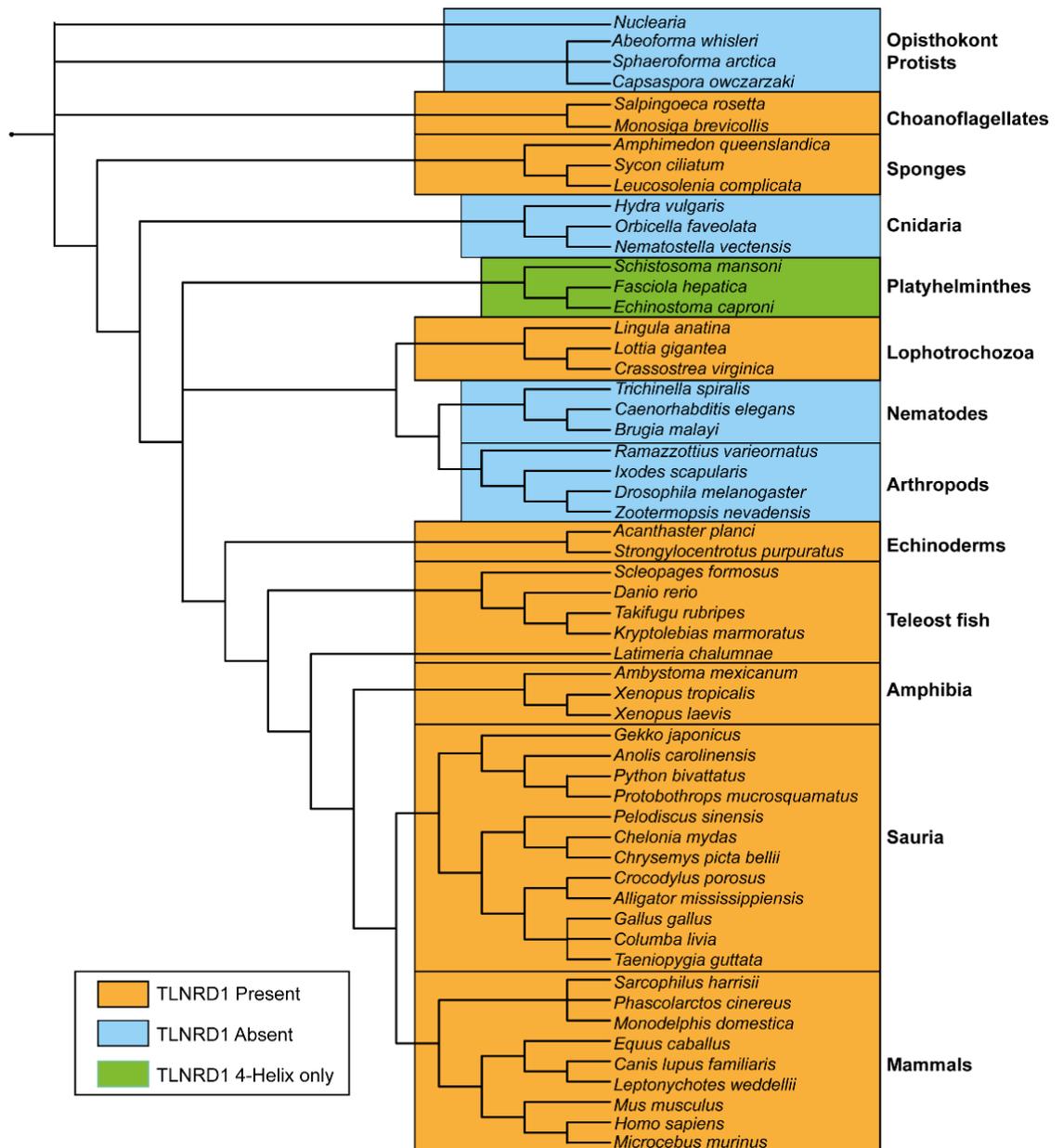


Figure 3.3. Evolutionary tree of TLNRD1

Tree diagram representing TLNRD1 presence and absence over evolution. TLNRD1 first appears in choanoflagellates and sponges and is conserved throughout vertebrate evolution. TLNRD1 has been lost from Nematodes, Arthropods and Cnidaria.

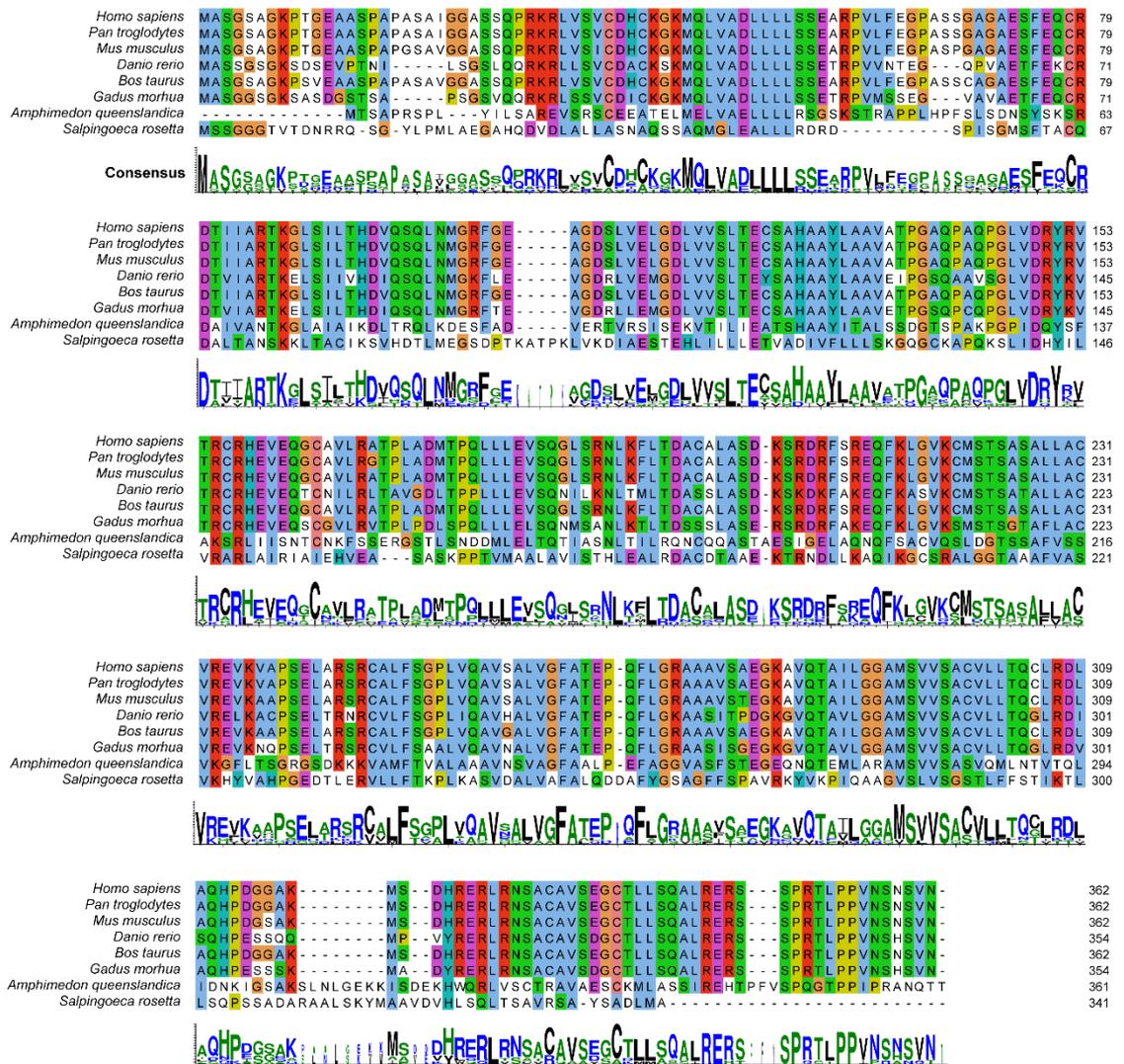


Figure 3.4. Amino acid sequence alignment of TLNRD1

Amino acid sequence alignment performed using Clustal Omega and coloured using Jalview (version 2.10.3) Clustal X colour scheme (Blue= hydrophobic, red= positive charge, magenta= negative charge, green= polar, pink= cysteines, orange= glycines, yellow= prolines, cyan= aromatic, uncoloured= unconserved). Consensus sequence generated using WebLogo 3 with residues coloured according to hydrophobicity (Black= hydrophobic, blue= hydrophilic, green= neutral). Sequences were downloaded from NCBI; Homo sapiens (NP_072088.1), Pan troglodytes (XP_003314880.1), Mus musculus (NP_109630.1), Danio rerio (XP_005166423.1), Bos Taurus (XP_001094728.1), Gadus morhua (XP_030233331.1), Amphimedon queenslandica (XP_011404074.1), Salpingoeca rosetta (XP_004997002.1).

3.1.3. TLNRD1 N-terminal unstructured region

The predicted N-terminal unstructured region of TLNRD1 from residues 1-40 is an intriguing addition which is absent in talin R7R8. Searching the GenBank database through BLAST using just residues 1-40 of TLNRD1 yields no significant hits showing that this region is unique to TLNRD1. Structural prediction from Phyre2 protein fold recognition server (Kelley *et al.* 2015) also suggests that short helical structures may be

able to form in this unstructured region (Figure 3.2B). To confirm disorder in this region, TLNRD1 protein disorder prediction was generated against the full sequence using PrDOS (Protein DisOrder System) (Ishida and Kinoshita 2007)). N-terminal residues 1-30 have a significantly high probability of disorder along with C-terminal residues 345-362 (Figure 3.5).

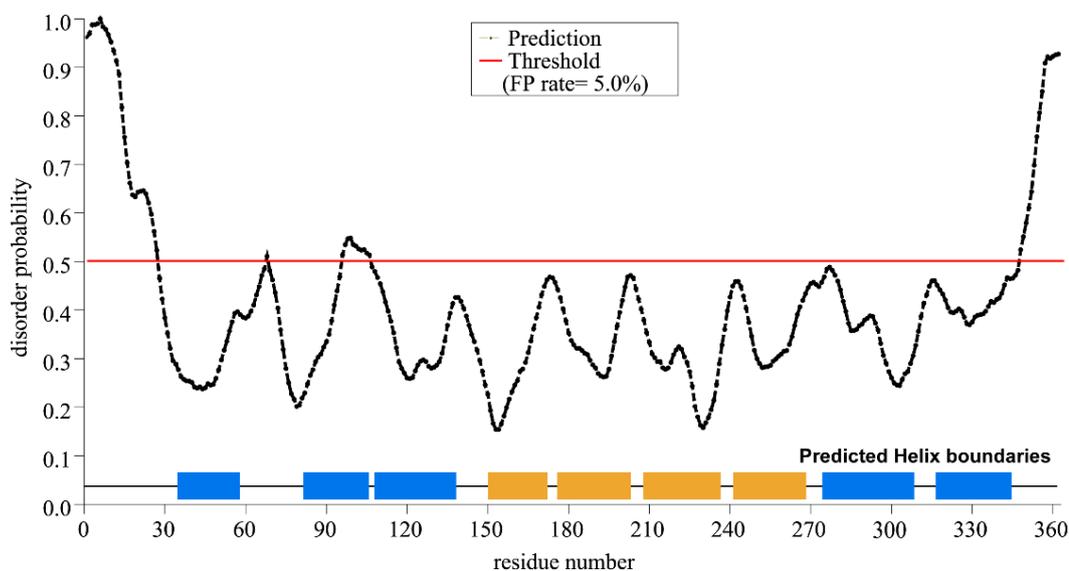


Figure 3.5. TLNRD1 PrDOS disorder prediction

Protein disorder prediction of full-length TLNRD1 showing high disorder from residues 1-30 and 345-362. TLNRD1 predicted helix boundaries are indicated with the 5-helix domain helices in blue and 4-helix domain helices in orange.

3.1.4. Predicted TLNRD1 phosphorylation sites

Protein phosphorylation is a common posttranslational modification which involves the covalent transfer of a phosphate group to the amino acids tyrosine, threonine or serine. Regulated by kinase and phosphatase enzymes, phosphorylation is important for regulating proteins through modification of protein conformational states and surface electrostatics, which modulate activation states, localisation, and ligand interactions. The prediction of phosphorylation sites through sequence information can be successful in identifying key regulatory kinases and provide clues to a protein's involvement in particular signalling responses.

Prediction of TLNRD1 phosphorylation sites alone using PhosPred-RF (Wei *et al.* 2017) predicts up to 22 phosphorylation sites in TLNRD1 with S15, S349 and T10 having a significant score above 0.8. Predicted sites with a score over 0.6 included S20, S26, S348, T137, T176 and T352. To support this, TLNRD1 sequence was also submitted to

iPhoPred (Li *et al.* 2019) which is adapted specifically for prediction of phosphorylation sites in human proteins and gives a percentage probability that identified residues may be phosphorylation sites. This predicted five sites of serine phosphorylation including S3 (69%), S15 (82%), S27 (81%), S35 (67%) and S55 (65%), one threonine site T10 (84%) and no tyrosine phosphorylation sites. Inspection of the protein phosphorylation database PhosphoSitePlus v.6.5.9.2 (Hornbeck *et al.* 2015) shows that phosphorylation of TLNRD1 has previously been detected in the N-terminus with S3 and S15 showing four references, S35 showing two references and S26 and S27 showing just one. Both the predictions and available data suggests that the N-terminal unstructured region is a target for regulation through phosphorylation.

While it is unknown exactly which kinases may phosphorylate the residues identified above, predictions can be made using both sequence information and functional features. This was done using two prediction models available online, the first being PhosphoPredict (Song *et al.* 2017) which was used with a low threshold to determine whether there are any predicted phosphorylation sites for different sets of kinases. Kinase specific prediction only yielded hits for residue S15 with the CDK family flagged with a low probability score of 0.423 and MAPK family with 0.35 probability, and amino acid S27 was flagged as a potential target for ATM family of kinases but with a low probability score of 0.311. To supplement this, the TLNRD1 sequence was also analysed for predicted sites using GSP 5.0 (Group based Prediction System) (Wang *et al.* 2020) using a high threshold and a search limited to *Homo sapiens*. For residues S3 and S15 the highest scoring predicted kinase was Vaccinia-related kinase VRK (score 465.9 and 439.7 respectively), for S27 the highest scoring hit was with tau-tubulin kinase (TTBK) (score 95.2) and for residue S35 protein kinase A PKA was the top hit with a score of 23.5. No kinases were predicted for S26. Collectively these results suggest that residues S3, S15 and S35 in the N-terminus of TLNRD1 may serve as targets for phosphorylation, however, predictions are still limited in accuracy so more research would be required to determine whether the predicted kinases can target TLNRD1 or whether other kinases may be involved in targeting these phosphosites.

3.1.5. Designing TLNRD1 constructs for structural and biochemical experiments

Design of TLNRD1 constructs for use in biochemical experiments were guided by known knowledge of the R7R8 structure domain boundaries, secondary structure predictions of TLNRD1 and protein disorder prediction. Constructs designed included full-length TLNRD1 protein residues 1-362 (TLNRD1-FL), the TLNRD1 4-helix domain from residues

144-273 (TLNRD1-4H) and the TLNRD1 5-helix domain from residues 1-362 with deletion of amino acids 144-273 encompassing the 4-helix module (TLNRD1-5H). All constructs were expressed in BL21(DE3) cells in a pET151 vector which has a N-terminal polyhistidine tag for nickel affinity purification followed by cation or anion exchange chromatography (Figure 3.6). Both TLNRD1-FL and TLNRD1-4H constructs expressed well and appeared relatively stable, however, expression of the TLNRD1-5H domain was unstable and variable, with low yield and multiple visible bands on SDS-PAGE making it unusable for most biochemical assays which require high concentrations or good quality protein.

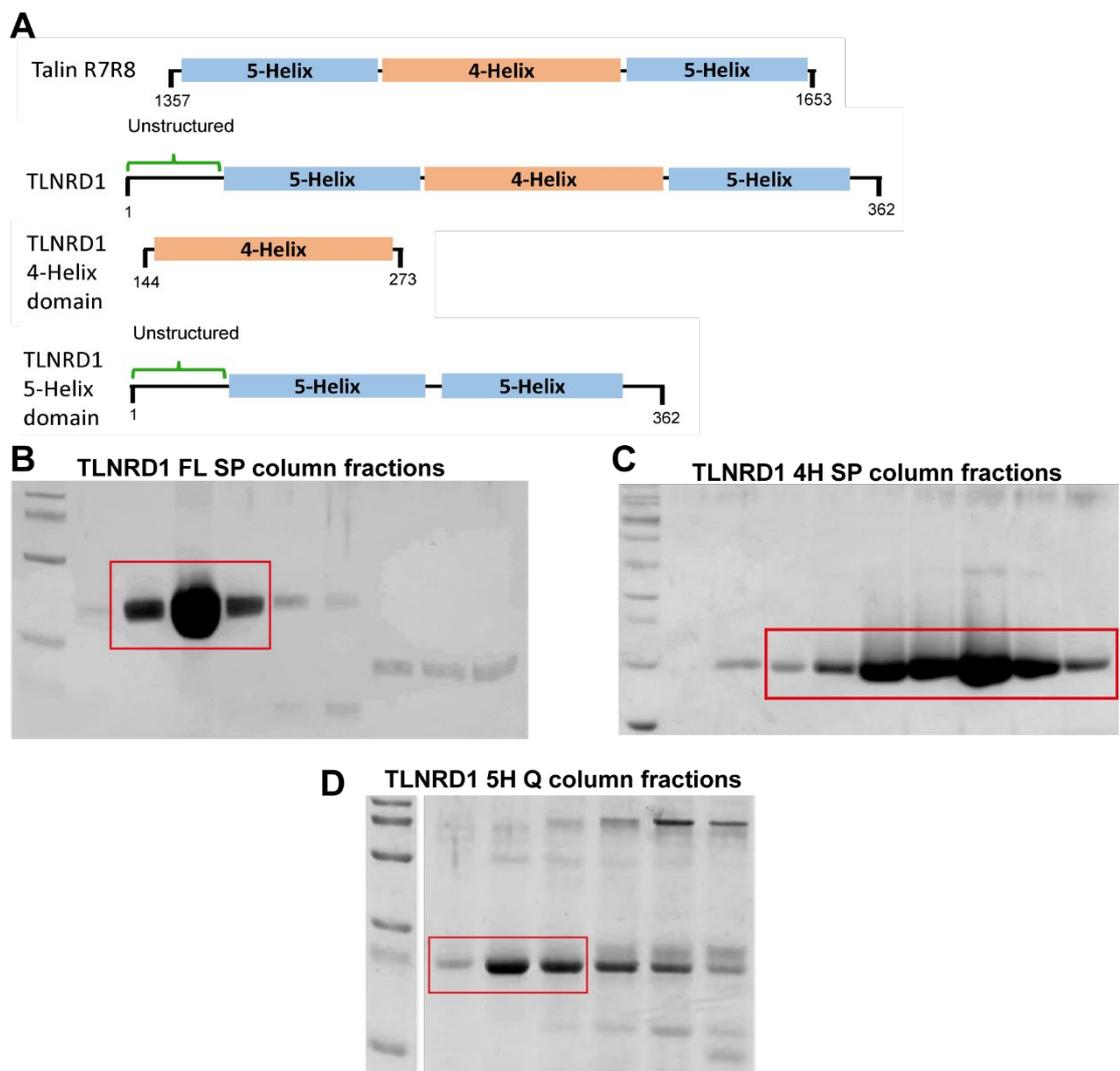


Figure 3.6. TLNRD1 construct design and purification

(A) Schematic representation of TLNRD1 constructs with 5-helix domain represented in blue and 4-helix domain represented in orange. Constructs were all expressed in an N-terminal His-tag expression vector. **(B-D)** SDS-PAGE gels of purified protein fractions of TLNRD1-FL, TLNRD1-4H and TLNRD1-5H after cation (FL and 4H) or anion exchange (5H). Fractions taken for biochemical assays indicated in red.

3.2. Stability and secondary structure analysis

3.2.1. Circular Dichroism theory

Circular dichroism (CD) is a highly sensitive technique which enables rapid evaluation of a protein's stability and secondary structure, allowing quick determination of whether a protein of interest is correctly folded. The basic principle of CD is the measurement of unequal absorption between left-handed and right-handed circularly polarised light. Light can be polarised using filters or prisms which causes sinusoidal oscillation of the light's electric field along a single plane. Viewed from the front, these oscillations can be seen as two circularly polarised waves in both a clockwise (left) and counter-clockwise (right) direction. Disruption by chiral molecules, such as a protein, of the two sets of polarised light can create differences in absorption between them. It is this difference in absorption which is measured in a CD experiment. Typical CD spectra for different protein compositions are represented below, showing the protein composition dependent changes in absorption over a range of Far-UV wavelengths. Each secondary structure element creates a distinctive pattern on the spectrum allowing easy visualisation of protein composition, for example proteins which are mostly α -helical will produce a positive curve at 190 nm and two negative minima at 208 and 222 nm in a classic 'double dip' formation (Figure 3.7) (Greenfield 2007).

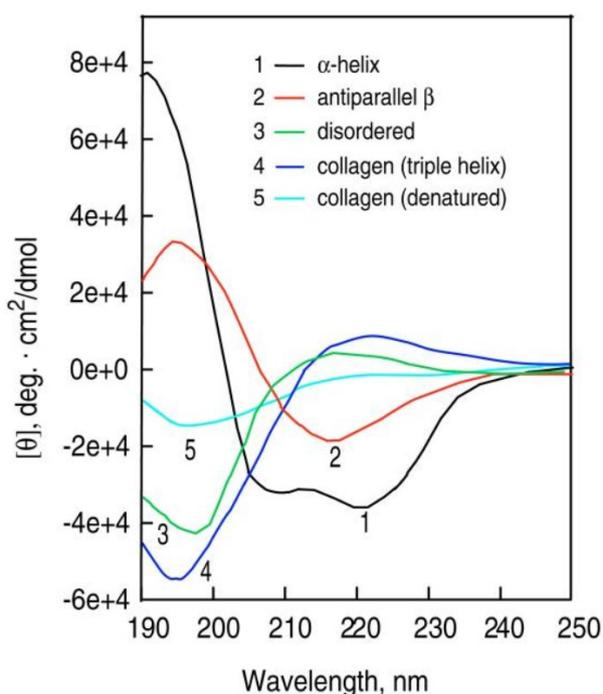


Figure 3.7. Circular Dichroism secondary structure spectra for proteins

A typical CD spectrum of proteins using wavelength measurements from 190-250 nm (Greenfield 2007). Degree of ellipticity or difference between left and right circularly polarised light exhibit protein composition dependent changes at different wavelengths.

3.2.2. Secondary structure and stability analysis of TLNRD1 constructs

CD spectroscopy was initially used to check protein stability and the folded state of the TLNRD1 constructs. CD measurements were taken in NMR phosphate buffer at 0.3 mg/ml. Secondary structure analysis was performed over the Far-UV wavelength range from 190-260 nm at 20°C. Protein thermostability measurements were measured at 222 nm wavelength with a temperature range from 20-80°C. Melting temperatures (T_m) are measured as the temperature at which 50% of the protein is unfolded.

The TLNRD1-FL construct showed a typical spectrum of an α -helix rich protein with melt curve analysis revealing that TLNRD1 is stable with a high T_m of 68°C (Figure 3.8 A/B). In comparison, talin R7R8 has a lower T_m of 53.5°C suggesting that there may be structural differences which increased TLNRD1 stability (Figure 3.8H). The TLNRD1-4H construct also had a typical secondary structure spectrum of a protein with high α -helical content and equally high stability with a T_m of 56°C (Figure 3.8 C/D). Finally, the TLNRD1-5H construct which was poorly expressed has a slightly altered secondary structure when compared to TLNRD1-FL and TLNRD1-4H, with reduced definition in the α -helical content (Figure 3.8E). This could be explained by the presence of the N-terminal unstructured region which will have a greater effect on the spectra of TLNRD1-5H when compared to TLNRD1-FL, as while TLNRD1-FL also has this unstructured region, it has a greater α -helical content through the presence of the 4-helix domain which masks any effect of protein disorder on the spectrum. TLNRD1-5H also showed a drastic reduction in thermostability compared to the other constructs with a T_m of 46.6°C suggesting that without the presence of the 4-helix domain the 5-helix module loses stability (Figure 3.8F).

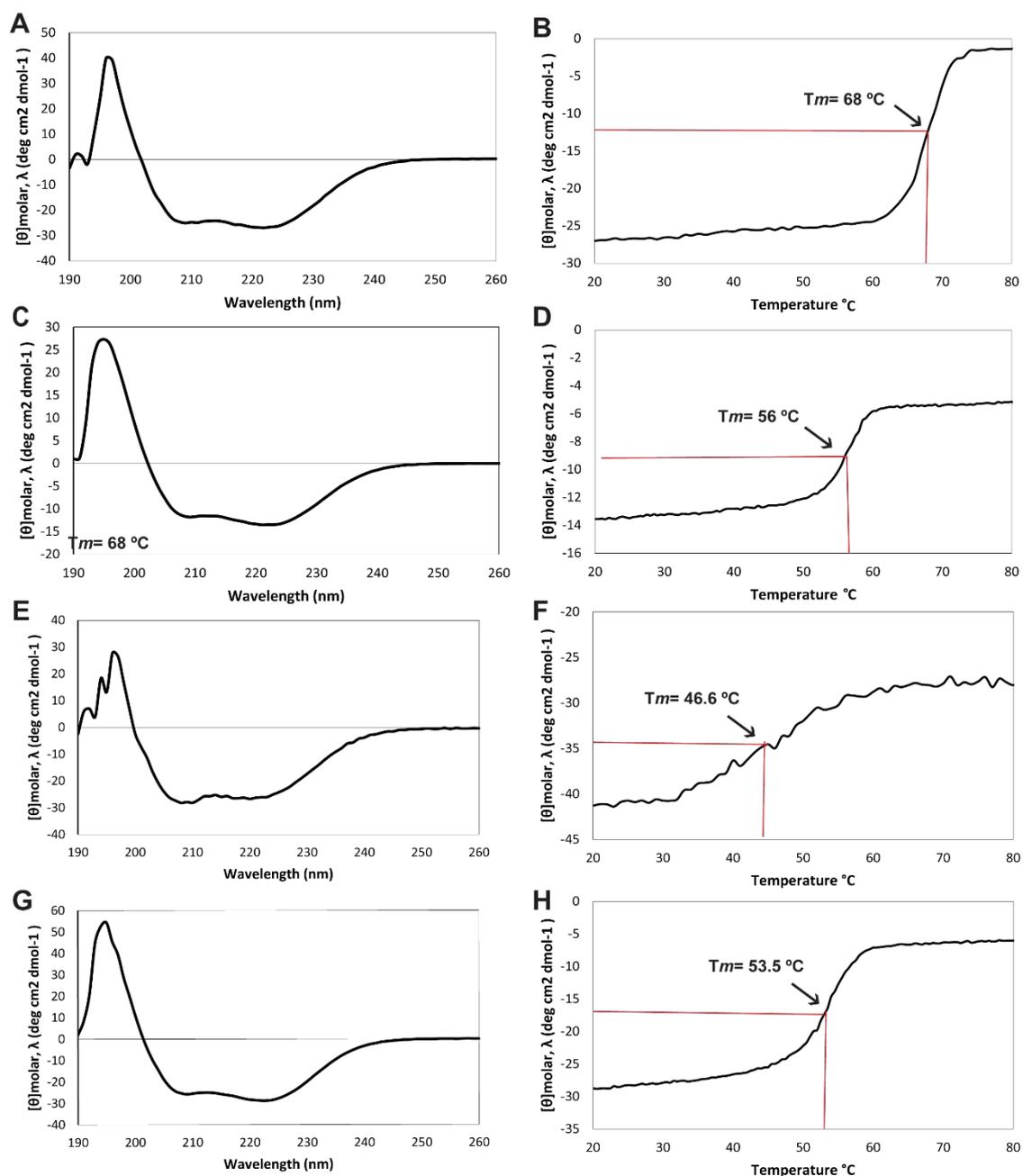


Figure 3.8. CD secondary structure and stability analysis

Secondary structure analysis of constructs with measurements over 190-260 nm wavelength. Protein stability analysis was performed at 222 nm wavelength over a temperature range of 20-80°C, measured T_m values are shown. **(A)** TLNRD1-FL secondary structure, **(B)** TLNRD1-FL melt curve, **(C)** TLNRD1-4H secondary structure, **(D)** TLNRD1-4H melt curve, **(E)** TLNRD1-5H secondary structure, **(F)** TLNRD1-5H melt curve, **(G)** Talin R7R8 secondary structure, **(H)** Talin R7R8 melt curve.

In addition to CD analysis, TLNRD1-FL thermal stability in two different buffers (NMR buffer and crystallisation buffer) was measured using nanoDSF with the Prometheus NT.48 (NanoTemper). This approach provides a quick method for checking a proteins stability over a concentration range with a minimal amount of protein. Both showed a

slight reduction in thermostability with melting temperatures around 59°C and minor concentration dependent changes in scattering showing that the protein remains stable at different concentrations in two commonly used buffers and isn't prone to detrimental aggregation.

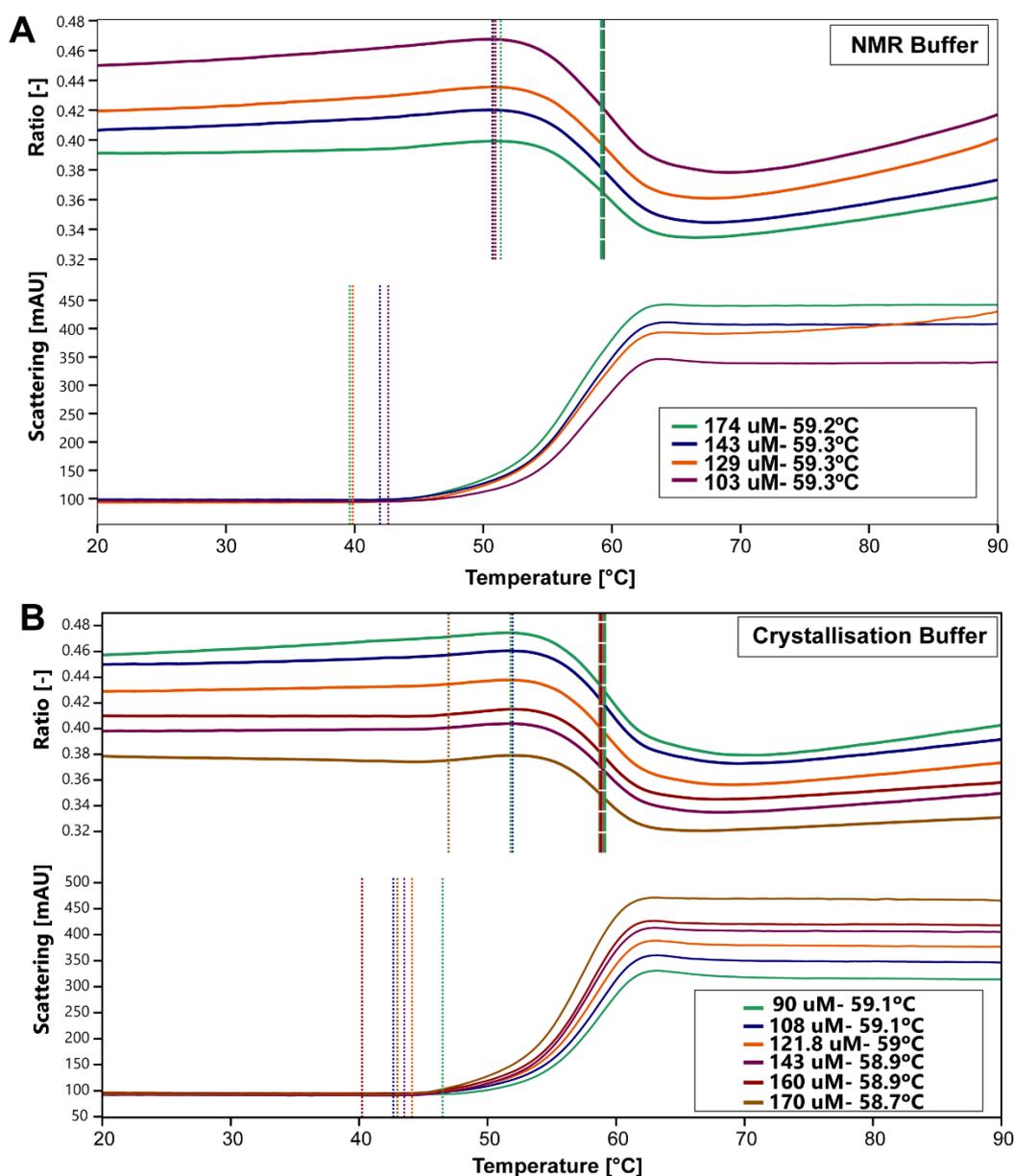


Figure 3.9. Thermal stability analysis of TLNRD1-FL

Stability measurements from 20-90°C of TLNRD1-FL. Ratio is the fluorescence ratio between 330 nm and 350 nm. Scattering measurements were taken to assess onset of protein aggregation. Dotted lines indicate the start of an unfolding event. Dashed lines indicate the inflection point of the curve.

(A) TLNRD1-FL in NMR buffer: 20 mM phosphate, 50 mM NaCl, 2 mM DTT, pH 6.5. **(B)** TLNRD1-FL in crystallisation buffer: 20 mM Tris, 100 mM NaCl, 2 mM DTT.

3.3. TLNRD1 crystallisation and structure determination

Both high stability and high expression of TLNRD1-FL and TLNRD1-4H makes them good candidates for crystallisation trials for X-ray diffraction data collection. Using this technique will provide detailed atomic information on the structure of TLNRD1, which is important to understanding its relationship with talin R7R8 and guiding future cell biology experiments.

3.3.1. TLNRD1-FL crystallisation

TLNRD1-FL construct was purified and prepared for crystallisation trials at a range of concentrations and screened as described in Materials and Methods section 2.8.2. Promising hits for optimisation were identified in two conditions, both from the JCSG-*plus* screen (Figure 3.10). No hits were identified in PACT Premier, Hampton crystal screen HT or Wizard classic I screens. Optimisation of crystallisation conditions for sodium acetate and 8% PEG 4000 did not yield crystals suitable for diffraction data collection. Optimisation of crystallisation conditions with sodium thiocyanate and PEG 3350 produced small square plate-like crystals which in most conditions were too small for data collection. Further optimisation produced larger thin crystals which could be grown to a suitable size for loop mounting within a week at 21°C. Final optimised crystallisation conditions used was 300 mM sodium thiocyanate pH 6.9 and 20% PEG 3350 with a protein concentration of 390 µM (Figure 3.10).

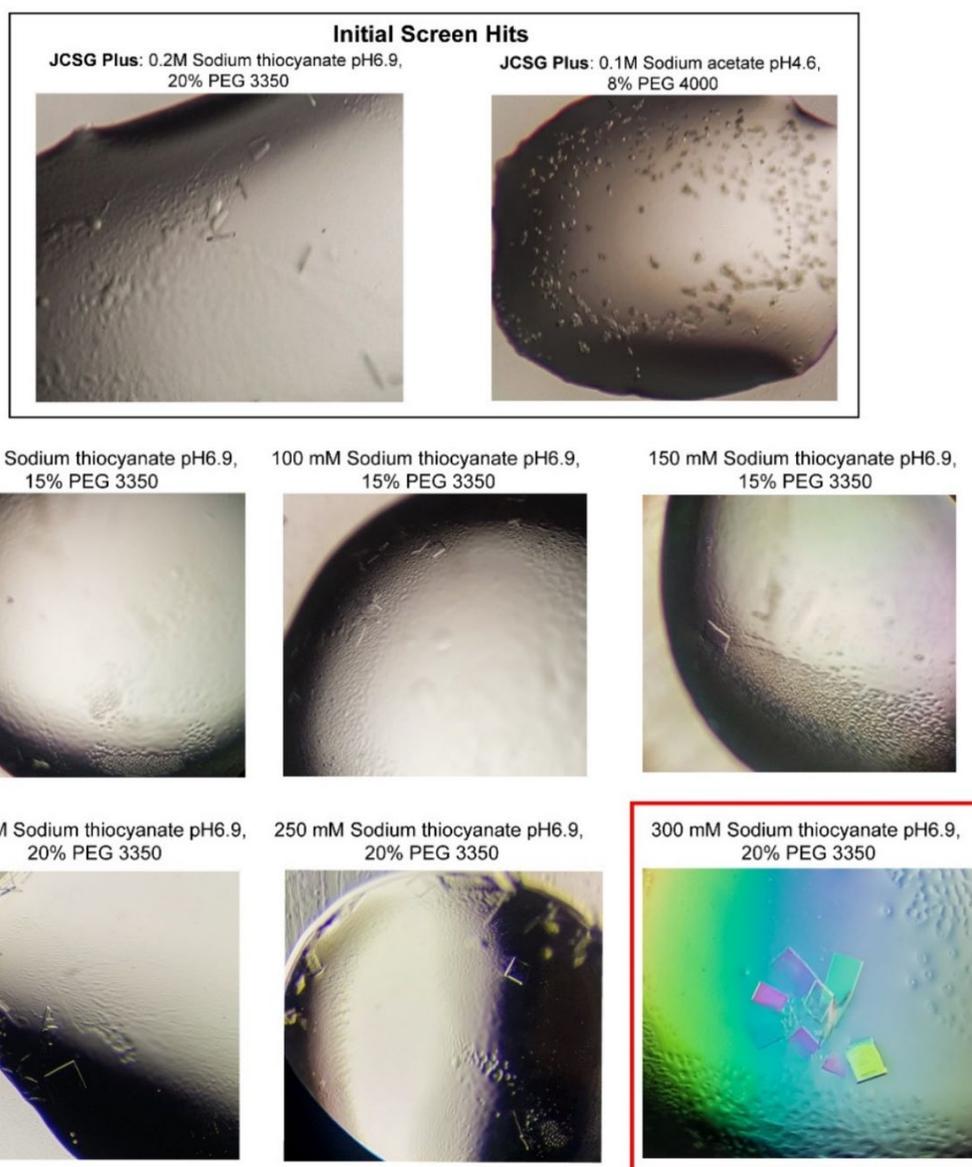


Figure 3.10. TLNRD1-FL screen hits and optimisation

Small crystals identified in two JCSG plus screen conditions: 0.2 M sodium thiocyanate pH 6.9, 20% PEG 3350 (top left) and 0.1M sodium acetate pH 4.6, 8% PEG 4000 (top right). Crystallisation conditions were optimised by varying sodium thiocyanate and PEG 3350. Diffracting crystals highlighted in red.

3.3.2. TLNRD1-4H crystallisation

TLNRD1-4H construct was prepared as described above. Promising hits for optimisation were identified in only one condition from the JCSG plus screen containing 0.2 M ammonium citrate dibasic pH5 and 20% PEG 3350 with a protein concentration of 350 μ M. No hits were identified in PACT Premier, Hampton crystal screen HT or Wizard classic I screens. Optimisation of crystallisation conditions produced a range of different crystal shapes including ‘snowflake’ needle like crystals, with the final optimised crystals being a medium size with a ‘boulder’ like square shape (Figure 3.11).



Figure 3.11. TLNRD1-4H screen hit and optimisation

A single crystal was initially identified in JCSG plus condition 0.2M ammonium citrate dibasic pH 5 and 20% PEG 3350 (top image). Crystallisation condition optimisation yielded different types of crystals with diffracting crystals highlighted in red.

3.3.3. Data collection and analysis for TLNRD1 constructs

X-ray diffraction data for both TLNRD1-FL and TLNRD1-4H crystals was obtained at Soleil synchrotron at the Proxima-1 beamline (Paris, France) and initially processed using the autoPROC pipeline (Material and Methods section 2.8.2). TLNRD1-FL crystallised in a $P2_1$ ($P\ 1\ 2_1\ 1$) monoclinic space group with two molecules per asymmetric unit and diffracted to a resolution of 2.3 Å. TLNRD1-4H crystallised in the $I4_122$ tetragonal space group with one molecule per asymmetric unit and diffracted to a resolution of 2.19 Å.

The TLNRD1-4H structure was solved using straightforward molecular replacement in PHASER with the ligand free talin R7R8 structure 2X0C (Gingras *et al.* 2010) as a search model. The TLNRD1-FL initial model was generated using an alternative molecular replacement pipeline to ensure the best possible solution for structure resolution. For this, the BALBES molecular replacement pipeline was utilised which uses a pre-processed database of structural information from the Protein Data Bank (PDB) to find the best solution. Both resulting models were then refined and validated with PHENIX refine with additional manual model refinement of side chains and water molecule

addition using Crystallographic Object-Oriented Toolkit (COOT). For both TLNRD1-4H and TLNRD1-FL models, the R_{free} and R_{work} values output from PHENIX.REFINE were used after each modification and round of refinement to monitor improvements to the model fit with the electron density map. MolProbity online web servers (Chen *et al.* 2010) were finally used to validate model plausibility before deposition in the Protein Data Bank. The TLNRD1 full-length structure was deposited with PDB ID 6XZ4 and TLNRD1 4-helix domain deposited with the PDB ID 6XZ3. Data collection and refinement statistics for both models are shown in Table 5 below.

Data collection	TLNRD1 Full-Length	TLNRD1 4-helix
Synchrotron and Beamline	Soleil Proxima-1	Soleil Proxima-1
Space group	$P2_1$	$I4_122$
Molecule/a.s.u	2	1
Cell dimensions		
a, b, c (Å)	69.17, 58.04, 84.32	114.59, 114.59, 59.40
α, β, γ (°)	90, 106.10, 90	90, 90, 90
Resolution (Å)	60.22 – 2.30 (2.38 – 2.30)*	57.30 – 2.19 (2.31 – 2.19)
R_{merge}	0.087 (0.682)	0.130 (1.034)
$I / \sigma I$	5.7 (1.3)	12.5 (2.6)
$CC(1/2)$	0.994 (0.831)	0.996 (0.938)
Completeness (%)	98.4 (98.7)	100 (99.9)
Redundancy	2.8 (2.8)	13.4 (13.6)
Refinement		
Resolution (Å)	2.30	2.19
No. reflections	27997 (2669)	10464 (2428)
$R_{\text{work}} / R_{\text{free}}$	0.26/0.31	0.22/0.26
No. atoms		
Protein	9047	1903
Water	70	38
B -factors (Å ²)		
Protein	74.25	72.05
Water	50.47	56.53
R.m.s. deviations		
Bond lengths (Å)	0.004	0.003
Bond angles (°)	0.792	0.620
Ramachandran plot		
Favoured/allowed/ outlier (%)	95.20/3.64/1.16	96.69/1.65/1.65
Rotamer		
Favoured/poor (%)	88.31/4.18	96.12/0.97
MolProbity scores		
Protein geometry	1.75 (97 th) [^]	1.01 (100 th) [^]
Clash score all atoms	1.88 (100 th) [^]	1.05 (100 th) [^]
PDB accession no.	6XZ4	6XZ3

Table 6. Data collection and refinement statistics for TLNRD1-FL and 4H

3.3.4. X-ray crystal structure of TLNRD1 full-length protein

For TLNRD1-FL, electron density could be traced for residues 40-344 of chain A and 39-342 of chain B, with no visible electron density for N-terminal residues 1-39/40 and C-terminal residues 344-362 in both chains. This is likely due to high flexibility in the N-terminal and C-terminal tails, supporting the secondary structure predictions and CD analysis which suggested that these residues are mostly disordered.

The crystal structure confirms that TLNRD1 is formed of 9 α -helices organised into two domains. Initial inspection of the TLNRD1-FL structure reveals that it shares the exact same domain structure and topology as the talin R7R8 domain with a 4-helix bundle inserted between two helices, α 3 and α 4, of a 5-helix bundle (Figure 3.12). Overlaying the talin R7R8 structure 2X0C with a single monomer of TLNRD1 shows highly similar helix positioning in both domains, but intriguingly, while the R7 and R8 domains crystallise in an open extended form, TLNRD1 appears to be a compact conformation with close association of both the 4-helix and 5-helix domain (Figure 3.12D). Interface analysis of the TLNRD1-FL structure with PDBePISA (Krissinel and Henrick 2007) fails to identify any interfacing residues which could potentially mediate an interaction of the two domains, suggesting it is either held in this position by the N-terminal disordered region not visible on the structure, or it is a result of crystal packing.

What is most striking about the TLNRD1 structure is that TLNRD1 appears to have adapted a dimerisation interface in the 4-helix domain of the protein. Previous biochemical characterisation of TLNRD1 in solution using gel filtration (Gingras *et al.* 2010) suggested that TLNRD1 is a dimer, however, this wasn't further investigated. Talin R7R8 is monomeric in solution with no evidence of being able to form a dimer and with the N-terminal unstructured region being the only region of TLNRD1 which distinguishes it from talin R7R8, we initially thought that this may be the region which mediates TLNRD1 dimerisation. However, this TLNRD1 structure reveals a surprising different mode of dimerisation from what we initially thought, forming an antiparallel symmetric dimer via an extensive interface with helices α 6 and α 7 of the 4-helix domain.

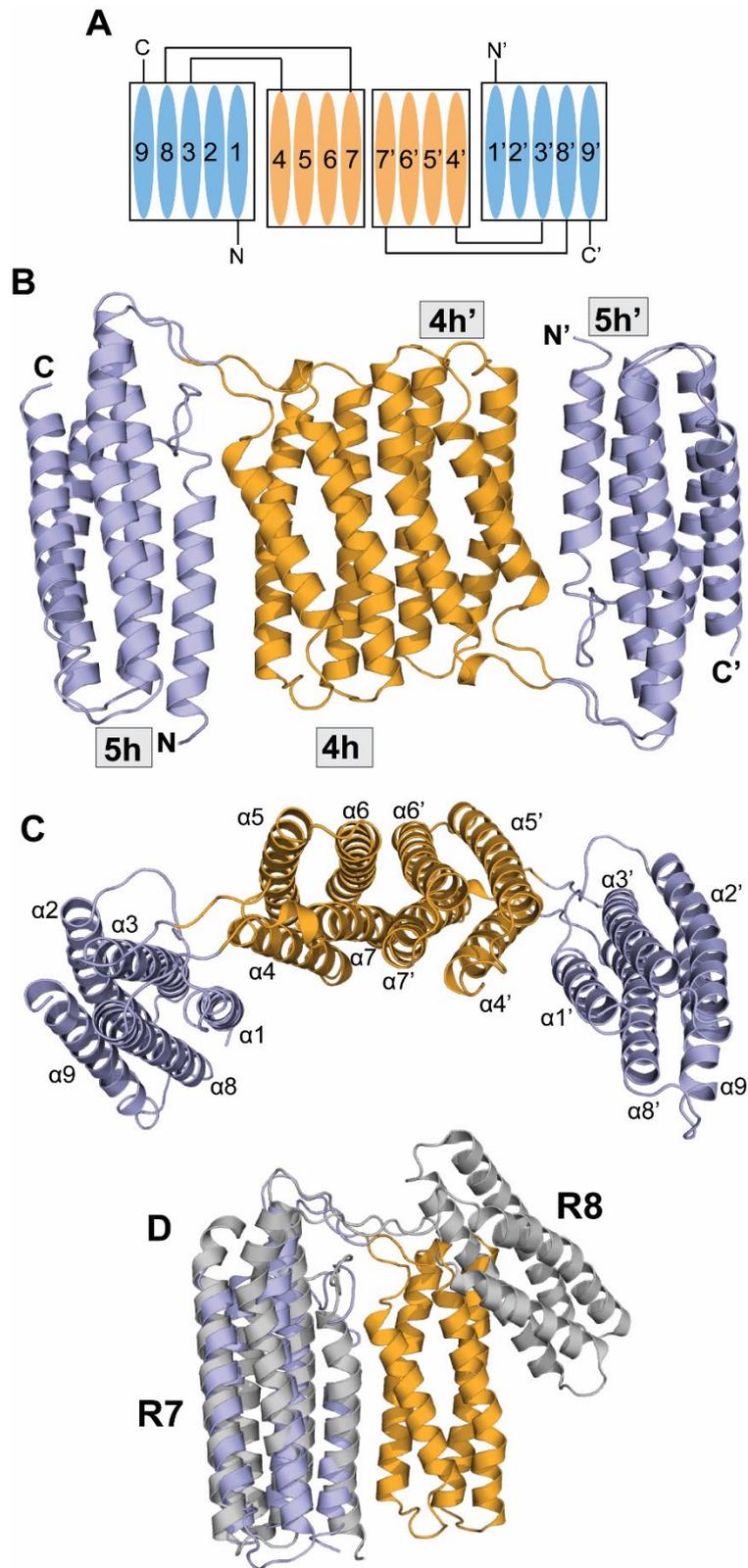


Figure 3.12. TLNRD1 crystal structure reveals an R7R8-like topology and antiparallel dimer
(A) Schematic representation of the TLNRD1 structure. Dimerisation is mediated via an extensive interface comprised of helices $\alpha 6$ and $\alpha 7$ in the 4-helix domain. **(B)** The crystal structure of TLNRD1 full-length protein with domains labelled and N and C-terminal regions indicated. **(C)** A top down view of (B) with the helices labelled, showing a curved structure **(D)** Overlay of one monomer of TLNRD1 (orange and blue) with talin R7R8 (grey; PDB ID: 2X0C).

3.3.5. X-ray crystal structure of TLNRD1 4-helix domain

With the TLNRD1-4H domain structure, electron density could be traced for residues 148 to 270, encompassing the entire 4-helix domain. In this structure, only one monomer molecule resides in the asymmetric unit, however, visualisation of nearby symmetry mates reveals that consistent with the TLNRD1-FL structure, the 4-helix domain forms an extensive dimerisation interface with an opposing monomer with helices α_6 and α_7 (Figure 3.13A). Alignment of the talin R8 4-helix structure (2X0C) with the TLNRD1-4H domain reveals only subtle structural differences. Helices α_6 and α_7 retain the same positioning as equivalent helices H33 and H34 of R8, but helices α_4 and α_5 show subtle shifts in positioning relative to H31 and H32 of R8, with α_5 shifting by 4.5Å, losing the marginal 9.8° bend seen in H32 of R8 (Figure 3.13B/C).

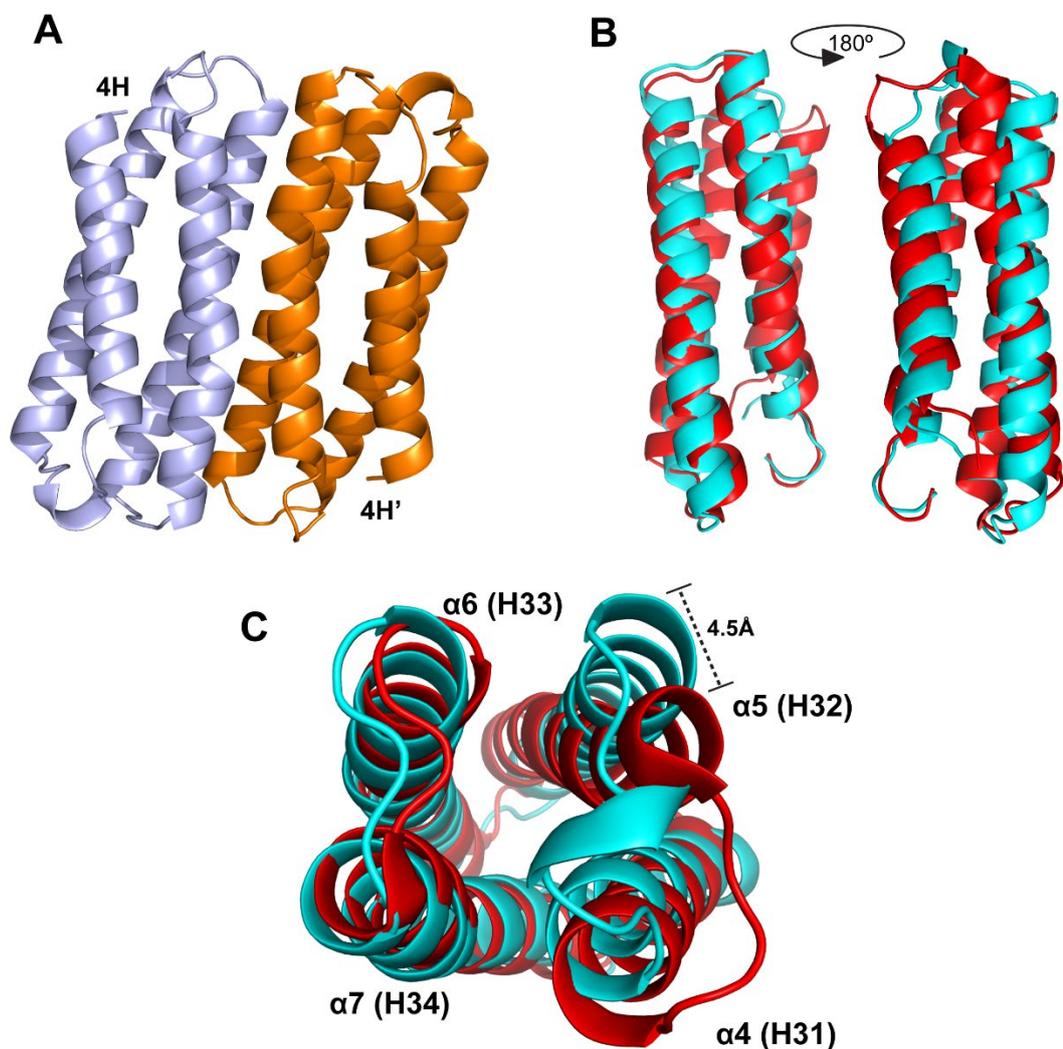


Figure 3.13. TLNRD1 4-helix domain crystal structure

(A) Crystal structure of TLNRD1-4H (blue) showing dimerisation with a second 4H monomer in an adjacent unit cell (orange). (B) PyMOL alignment of TLNRD1-4H structure (cyan) with talin R8 domain (red) (2X0C). (C) Top-down view of TLNRD1-4H and talin R8 aligned structures.

3.3.6. TLNRD1 is a symmetric antiparallel dimer

The dimerisation of TLNRD1 is created along the entire length of helices α_6 and α_7 via an extensive hydrophobic interface. Analysis of the TLNRD1-FL interaction properties with PDBePISA (Krissinel and Henrick 2007) produces a Complex Formation Significance Score of 1 (on a scale from 0-1) confirming that TLNRD1 is a *bone fide* dimer. The dimerisation of TLNRD1 appears to be mediated by the aromatic ring of phenylalanine F250 docking into a pocket on the opposing molecule created by Q213' and the small side chains of G217' and A260' (Figure 3.14A/B). The dimerisation interface is also stabilised by predicted salt bridge formation between R246 and E267 of the opposing monomer (Figure 3.14C). This unique configuration leads to an antiparallel configuration of the TLNRD1 dimer.

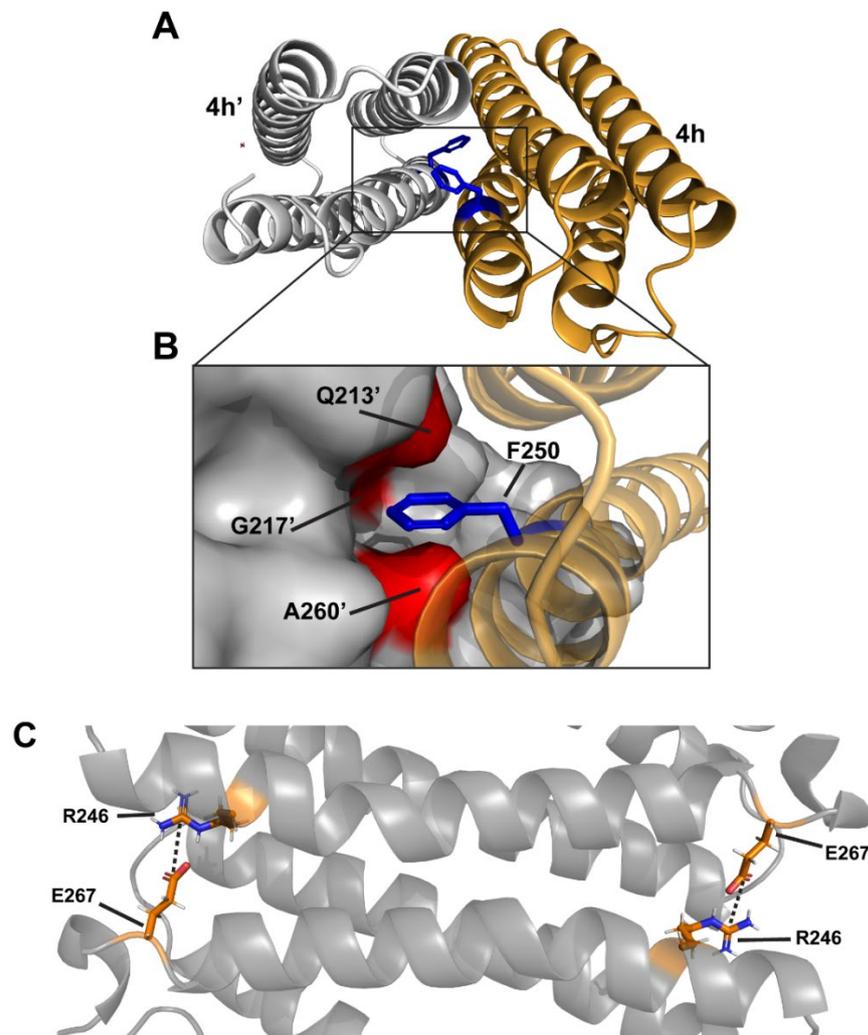


Figure 3.14. TLNRD1 dimerisation interface

(A) TLNRD1 antiparallel dimerisation is mediated by F250. **(B)** Zoomed-in image showing F250 docking into a pocket created by Q213', Q217' and A260' (red) of the opposing monomer (grey). **(C)** Dimer stabilisation by salt bridge formation between R246 and E267 (orange) of the opposing monomer.

To explore whether TLNRD1 has always been a dimer throughout its evolution from humans back to the choanoflagellate *Salpingoeca rossetta*, ConSurf analysis (Ashkenazy *et al.* 2016) was used to visualise the conservation of TLNRD1 amino acid residues. This analysis gave maximum conservation scores for F250 and the pocket creating residues Q213, and A260, showing that this mode of dimerisation has been highly conserved throughout its evolution and confirming that TLNRD1 is an obligate dimer (Figure 3.15). The two stabilising residues R246 and E267 showed moderate conservation suggesting that while these residues are not essential for dimer formation, they may have been a later addition to strengthen the dimerisation interface.

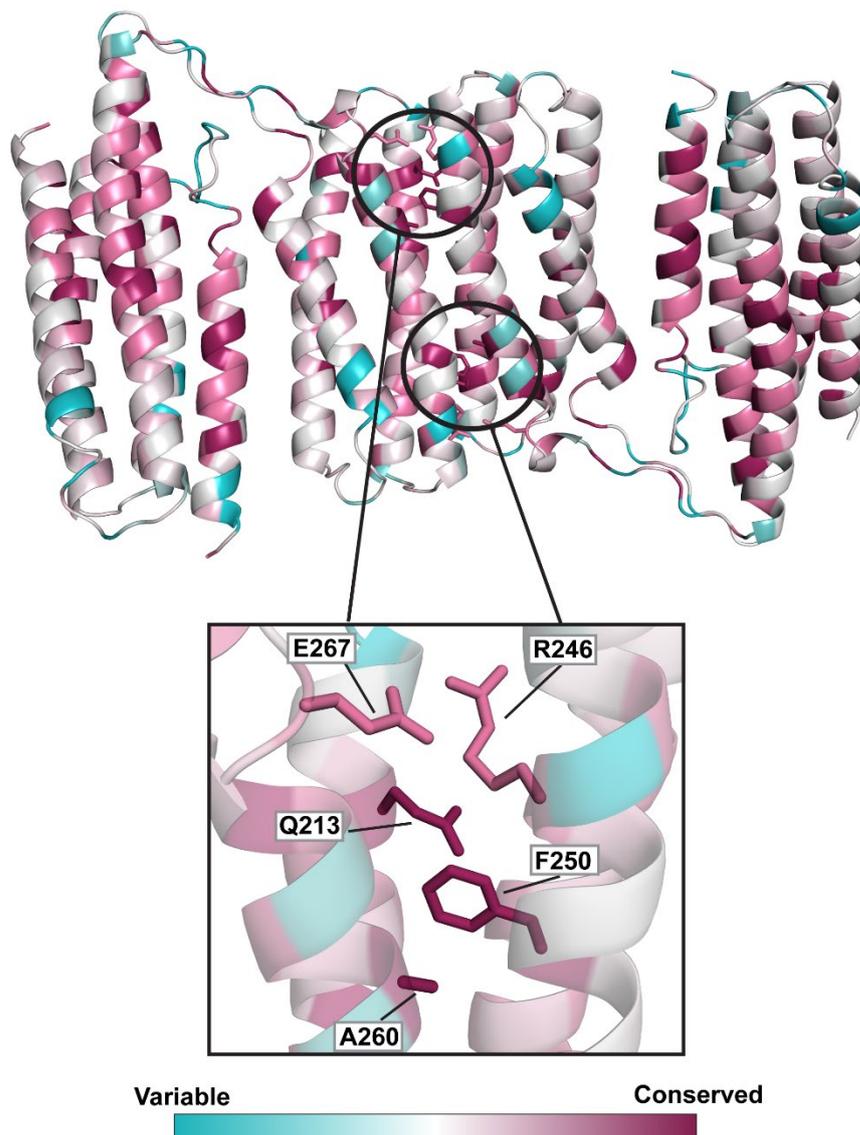


Figure 3.15. TLNRD1 dimerisation interface is highly conserved

TLNRD1-FL structure and sequence were analysed with ConSurf to explore residue conservation. TLNRD1 structure is coloured according to degree of conservation, with purple indicating highly conserved and blue showing variable residues. Zoomed-in image shows conservation of residues F250, Q213, G217, A260, R246 and E267.

3.3.7. Designing a TLNRD1 dimerisation knockout mutant

To understand the functional significance of TLNRD1 dimerisation, a mutant was designed which was predicted to prevent TLNRD1 from being able to form a dimer. With the residue F250 being so highly conserved, and its positioning into a pocket at each end of the dimerisation interface, residue F250 was a good candidate target for mutation. Using site-directed mutagenesis, the bulky F250 was swapped with a smaller negatively charged aspartic acid which should repel interaction with the opposing pocket (Figure 3.16). The mutation was introduced into TLNRD1-FL and the protein expressed well with CD analysis confirming that the protein was properly folded, however, CD thermostability analysis did show a drastic reduction in stability with a melting temperature of 48°C, which is an 11.7°C drop from wild-type stability (Figure 3.16). The impact of this mutation on TLNRD1 dimerisation is shown in the next section exploring TLNRD1 oligomerisation in solution.

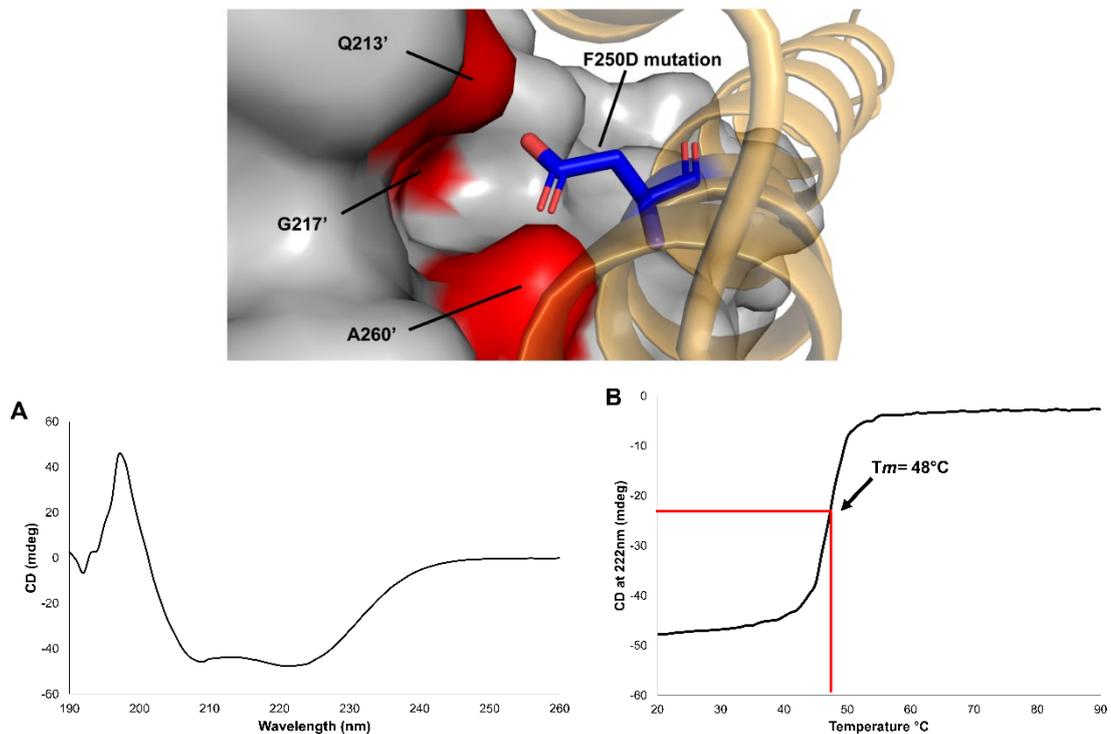


Figure 3.16. F250 mutation to aspartic acid

Dimer interfacing residue F250 was mutated to aspartic acid to destabilise TLNRD1 dimerisation. **(A)** TLNRD1 F250D CD spectra showing largely alpha-helical content. **(B)** TLNRD1 F250D melt curve analysis from 20-90°C at 222 nm wavelength. Melting temperature (T_m) shows a drastic reduction in thermostability.

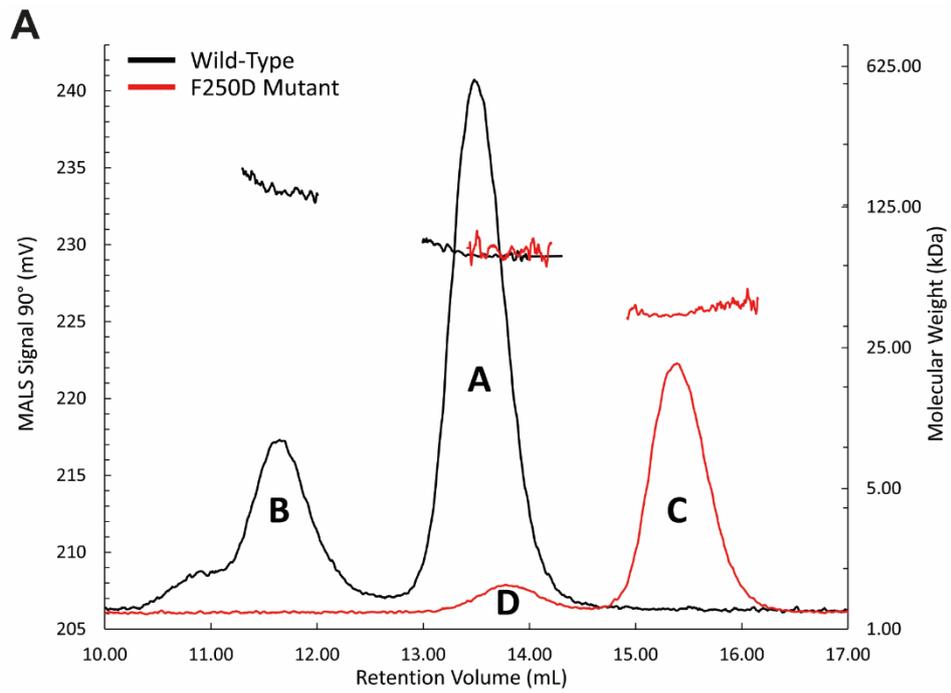
3.4. Solution analysis of TLNRD1 dimerisation

3.4.1. SEC-MALS analysis of TLNRD1 oligomeric state

To explore the oligomeric state of TLNRD1 in solution, size-exclusion chromatography was combined with multi-angle light scattering (SEC-MALS) to determine the molecular mass of any TLNRD1 molecules in solution. Size-exclusion chromatography allows separation of different oligomeric states such as monomer and dimer by separating them by molecular weight. Combining this with multi-angle light scattering allows accurate calculation of molecular mass for each separated peak/state of the protein.

For TLNRD1 full-length protein two peaks were present at 106 μ M showing TLNRD1 exists in two oligomeric states. MALS molecular weight calculation for both peaks revealed molecular weights of 73.4 kDa and 151.9 kDa, consistent with a dimer of TLNRD1 and a tetramer (Figure 3.17A). Approximately 70% of the TLNRD1 forms a dimer in solution with a smaller ~30% forming a dimer of dimer species. There was no evidence of any monomeric TLNRD1 which would have a calculated molecular weight of ~37.8 kDa, suggesting that the dimerisation is a high affinity interaction. Unlike TLNRD1-FL which forms two states, only one peak was observed for TLNRD1 4-helix domain alone, with an approximate calculated molecular weight of 28.6 kDa which is consistent with a dimer. The lack of a dimer of dimer species in TLNRD1-4H alone suggests that the TLNRD1 tetramer involves interaction of the 5-helix domain or either the C-terminal or N-terminal unstructured regions.

Finally, SEC-MALS analysis was used to establish whether F250 was a key residue in mediating TLNRD1 dimerisation. The F250D mutant produced two visible peaks on the elution spectra, with a drastic shift in peak position. MALS molecular weight calculation for both peaks revealed molecular weights of 75.7 and 37.8 kDa consistent with a dimer and monomer in solution. Nearly 90% of the protein appeared to be in a monomeric state confirming that mutation of F250 drastically reduces the ability of TLNRD1 to dimerise. The considerably smaller dimer peak which was observed could be due to either residual dimerisation interaction supported by the salt bridge between R246 and E267 or is a result of the second unidentified interacting interface that gives rise to TLNRD1 tetramer formation.



TLNRD1 Wild-Type		TLNRD1 F250D	
A	B	C	D
73.4 kDa	151.9 kDa	37.8 kDa	75.7 kDa

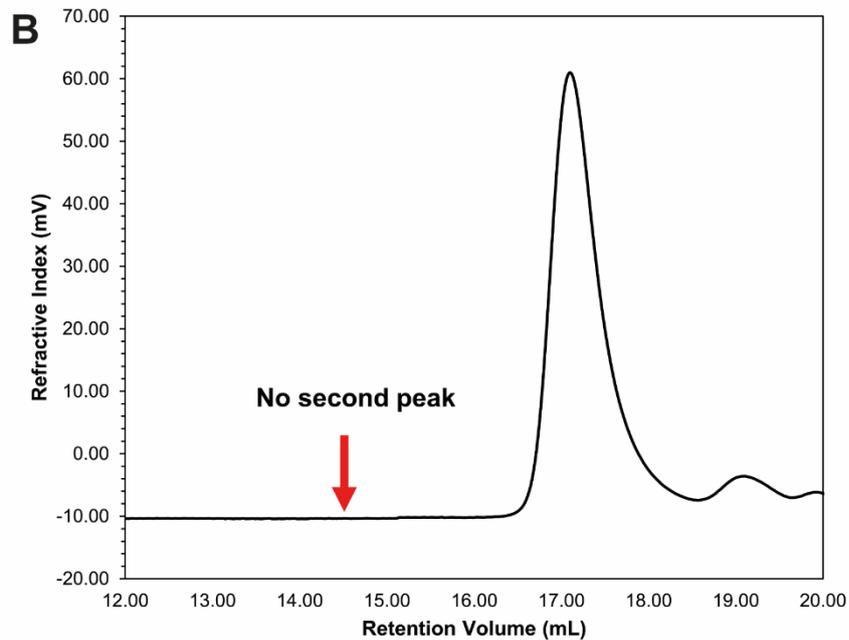


Figure 3.17. SEC-MALS analysis of TLNRD1 oligomer formation and F250D mutation

(A) SEC-MALS analysis of TLNRD1-FL and F250D at 106 μM . Calculated molar mass of TLNRD1-FL reveals dimer and tetramer species. F250D mutant shows drastic transition to monomeric state. **(B)** SEC analysis of TLNRD1-4H at 120 μM shows a single dimer peak.

3.4.2. MST analysis confirms TLNRD1 as a high affinity dimer

The complete absence of a monomeric peak in the SEC-MALS analysis and drastic impact of the F250D mutation suggests that the dimerisation of TLNRD1 is mediated by a high affinity interaction between F250 and the pocket on the opposing monomer. To explore the monomer-dimer equilibrium of TLNRD1 and confirm a high affinity interaction, microscale thermophoresis (MST) was used to calculate a dimerisation constant (K_d). The high sensitivity of MST allows binding experiments to be carried out on extremely low concentrations of protein, which is beneficial for determining a binding affinity for high affinity nanomolar interactions. The data output represents thermophoretic changes in normalised fluorescence over a ligand dilution series.

For these experiments, unlabelled TLNRD1 protein with the his-tag cleaved off was titrated against his-tagged TLNRD1 protein which was labelled with a red NTA dye which binds to the histidine tags of proteins. Wild-type TLNRD1-FL protein revealed a dimerisation K_d of 80 nm (± 0.6) confirming a high affinity interaction. MST analysis of the dimerisation mutant F250D showed that mutating F250 to an aspartic acid prevents TLNRD1 from being able to dimerise with no generatable K_d regardless of how much the target: ligand ratio was altered. Together, these results confirm the TLNRD1 exists as a high affinity obligate dimer which is mediated by F250 interaction with the Q213/A260 pocket on the opposing monomer.

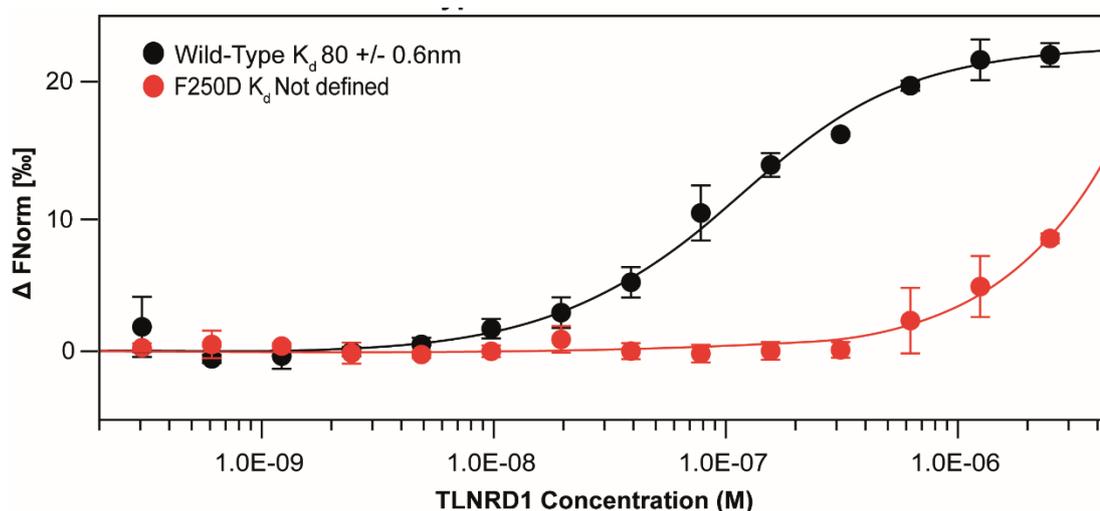


Figure 3.18. TLNRD1 is a high affinity dimer

Dose-dependent curve of the TLNRD1 monomer-dimer equilibrium reveals a high affinity interaction with a K_d of 80 nm (SE ± 0.6 nm). Thermophoresis curves represent change in the normalized fluorescence (ΔF_{norm}). N=3.

3.5. Small Angle X-ray Scattering (SAXS) of TLNRD1

Small angle X-ray scattering is a technique which enables exploration of a protein's dynamics in solution from concentration dependent aggregation effects to protein shape and flexibility. The TLNRD1-FL crystal structure shows TLNRD1 in a more compact conformation compared to talin R7R8. To explore whether this is just an artefact of crystal packing or whether TLNRD1 retains this 'autoinhibited' structure in solution, SAXS was used to explore shape and flexibility of the protein. Using SAXS will tell us whether TLNRD1 adopts an extended 'open' conformation in solution which would be presented as an elongated shape, or whether it retains the closed compact conformation which would be a more rounded globular shape. As TLNRD1 exists in two oligomeric states, SAXS was combined with size-exclusion chromatography (SEC-SAXS) to ensure scattering patterns are obtained only from single species and not a mixture of different states. Analysis was only performed on the TLNRD1 dimer as good quality data could not be obtained for accurate analysis of the tetramer.

Data quality inspection

SAXS is a highly sensitive technique which requires accurate buffer subtraction and a good quality monodisperse sample free from aggregation. 2D SAXS data is transformed into a 1D plot with a final buffer subtracted intensity plot. This is a logarithmic plot of intensity $I(q)$ against q which is defined as $q=4\pi\sin\theta/\lambda$ where 2θ is the scattering angle and λ is the radiation wavelength. The shape of this intensity plot allows inference of a proteins overall shape and weighting, including radius of gyration (R_g), maximum particle dimension (D_{max}) and hydrated particle volume (V_p), as well as assessment of a samples quality (Mertens and Svergun 2010). Initial inspection of the region of low q , termed the Guinier region, allows detection of any concentration-dependent effects, with an upward curve at low q indicating aggregation in the sample and a downwards slope indicating repulsion (Figure 3.19).

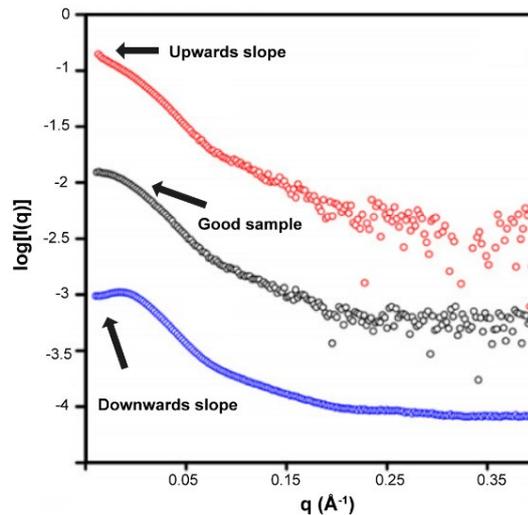


Figure 3.19. Typical SAXS scattering curves adapted from Mertens and Svergun (2010)

Scattering is represented in a logarithmic plot of intensity $I(q)$ against q . The top curve in red shows an upwards slope at low q indicative of sample aggregation. The bottom blue curve has a downwards facing slope at low q indicating inter-particle repulsion effects. The middle black curve represents a good quality sample.

Guinier analysis and R_g estimation

Guinier analysis, named after Andre Guinier, involves analysis of SAXS scattering data from the low q Guinier region of the intensity plot. Guinier analysis requires plotting of intensity $I(q)$ against q^2 to generate a straight Guinier plot (Figure 3.20). This information allows estimated calculation of the radius of gyration (R_g) which is the distribution of atoms around the protein's central axis, which serves as an indicator of how compact a molecule is in solution.

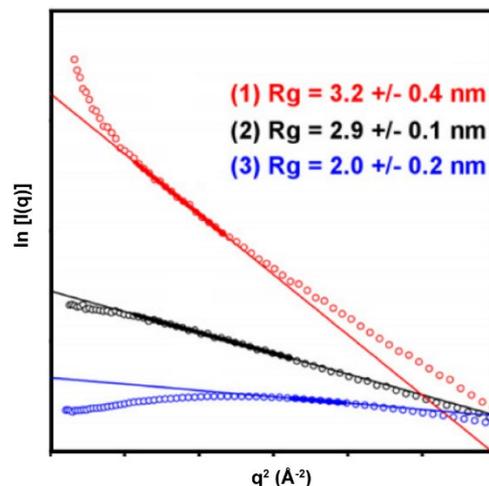


Figure 3.20. Typical Guinier analysis for R_g estimation adapted from Mertens and Svergun (2010)

Representative Guinier plot. The black plot represents a good quality sample with a linear downwards slope. Red shows typical curve of an aggregated sample, and blue represents inter-particle repulsion.

Kratky plot and flexibility analysis

Kratky plot analysis can be used to as a semi-quantitative analysis to check for signs of protein unfolding, and is created by plotting q^2 multiplied by intensity against q . A well folded globular protein will produce a defined bell-shaped curve whereas an unfolded sample would produce a plateau after an initial increase (Figure 3.21). A protein which has both folded and unfolded regions would produce a curve for both, with a classic bell curve which trails off with an increasing upwards trend.

One of the benefits of using ScÅtter software for data analysis is the ability to use an integration function which uses a power-law relationship between q and intensities $I(q)$ called the Porod-Debye law to estimate protein flexibility (Rambo and Tainer 2011). This uses data from the mid q region of the SAXS scatter curve called the Porod-region, the slope of which reflects protein flexibility and folded state. This technique produces 3 sub-graphs, a Porod-Debye plot with a q^4 axis, Kratky-Debye plot with a q^3 axis and SIBYLS plot with a q^2 axis. Each of these plots can be compared side by side to determine which produces a plateau, and this reflects the folded state of the protein with plateau in q^4 indicating a folded globular protein with no flexibility, q^3 indicates a globular protein with some unfolding for flexibility and q^2 indicating an unfolded highly flexible protein.

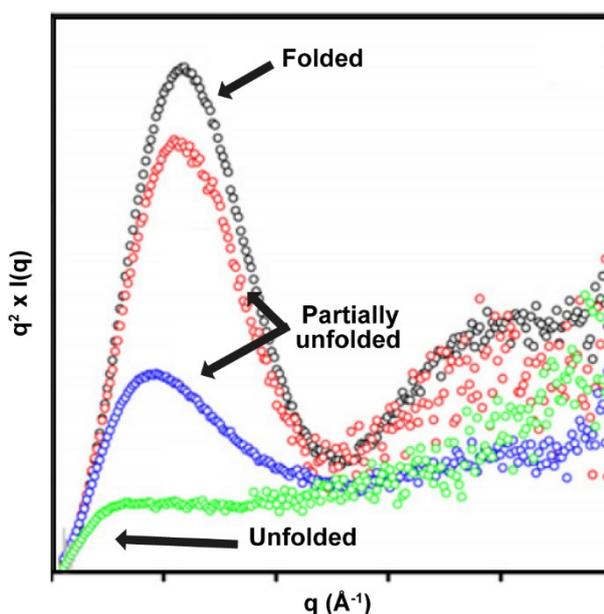


Figure 3.21. Typical Kratky plot adapted from Mertens and Svergun (2010)

Representative Kratky plot determination of protein flexibility. Folded and partially unfolded proteins produce a bell-shaped curve (black, red and blue), unfolded proteins produce a plateaued curve.

P(r) distribution

The pair distance distribution function, known as $P(r)$, gives information about a protein's overall shape by giving a real-space representation of the SAXS scattering curve. $P(r)$ is calculated from intensity $I(q)$ with an indirect Fourier transform. $P(r)$ can also be used to generate a more accurate R_g value as it accounts for all SAXS data, whereas R_g estimated from Guinier analysis only uses data from the Guinier region. The overall shape and right-hand tail of the curve near high r is a key indicator of protein shape. Most globular proteins will create a smooth bell-shaped curve, multi-domain proteins will produce areas of undulation in the curve and potentially form multiple peaks, and elongated proteins will have elongated tails at higher r values.

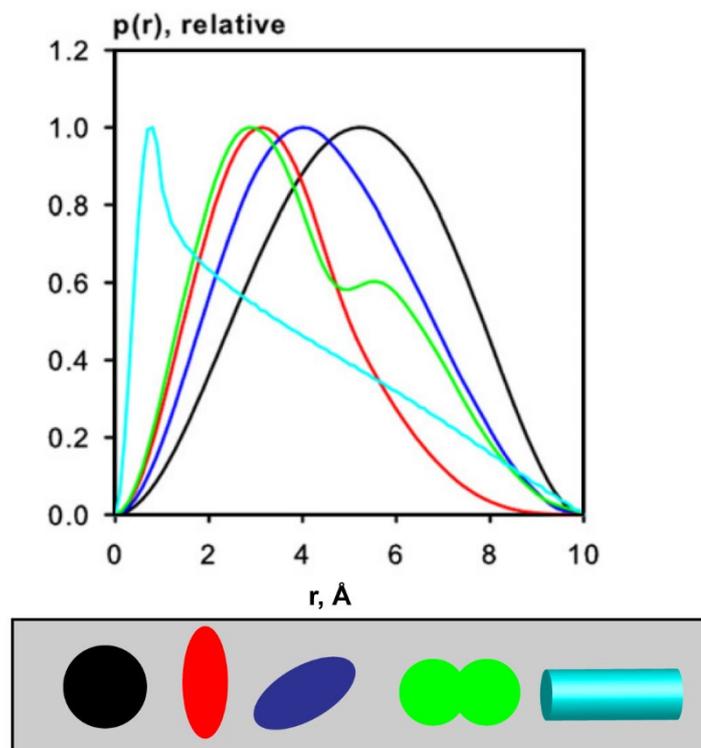


Figure 3.22. $P(r)$ distance distribution curves adapted from Mertens and Svergun (2010)

Representative $P(r)$ distance distribution curves and their relationship with protein shape. Rounded globular proteins will generate a bell-curve, multi-domain proteins may present as multiple peaks, and elongated proteins have a large trailing tail towards high r values.

3.5.1. SAXS analysis of the TLNRD1 dimer

SAXS analysis was performed on the TLNRD1-FL dimer peak with comparison at two concentrations, 3 mg/ml (79 μ M) and 7 mg/ml (185 μ M), to check for any concentration dependent effects on protein shape or dynamics. Analysed SAXS data and calculated parameters are summarised in Figure 3.23 below.

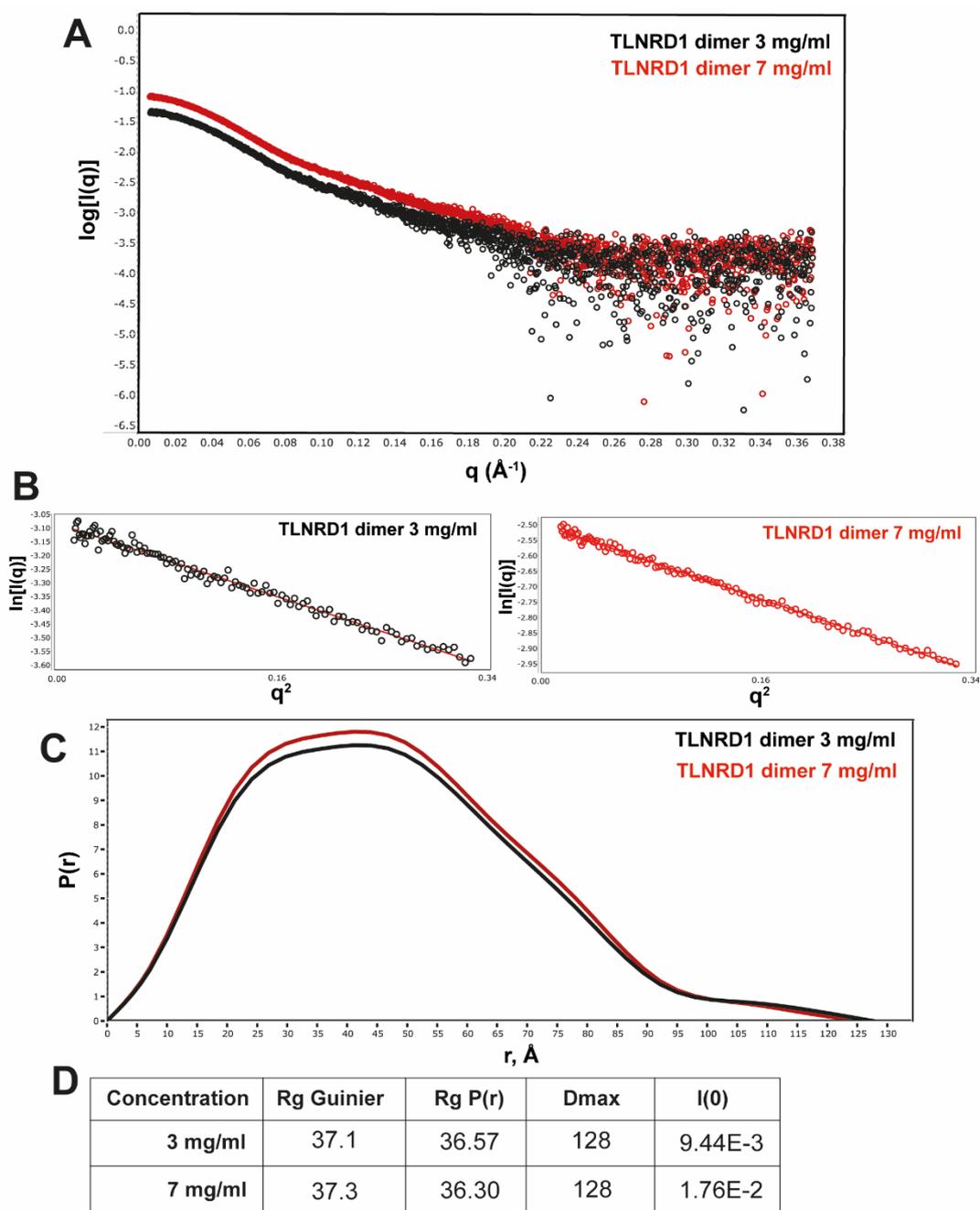


Figure 3.23. TLNRD1 SAXS data

(A) Comparison of buffer subtracted TLNRD1 scattering curves at 3 mg/ml and 7 mg/ml. **(B)** Guinier plot showing a linear trend with no deviations or signs of aggregation. Signal resolution improves at 7 mg/ml. **(C)** Comparison of TLNRD1 $P(r)$ distribution curves with D_{\max} of 128. **(D)** Table of calculated SAXS parameters for TLNRD1 at 3 mg/ml and 7 mg/ml.

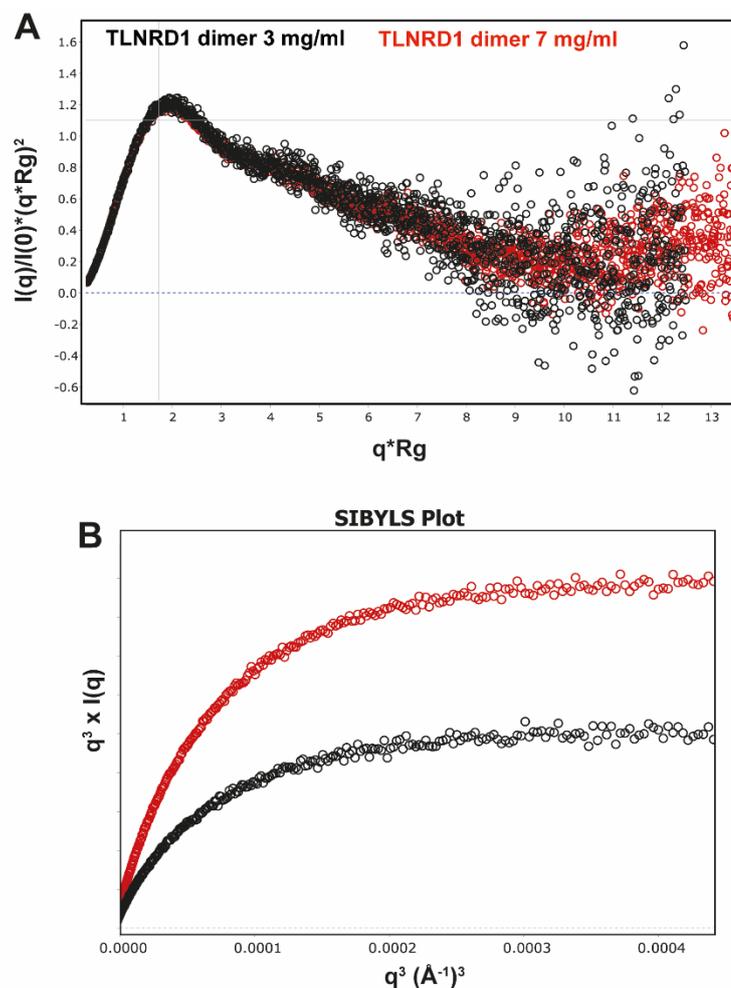


Figure 3.24. Normalised Kratky plot and SIBYLS flexibility plot

(A) Normalised Kratky plot for TLNRD1 dimer at 3 and 7 mg/ml. **(B)** ScÅtter flexibility plot analysis shows plateau in the SIBYLES plot with a q^3 axis scale indicating some protein flexibility.

Final TLNRD1 SAXS curves for 3 mg/ml and 7 mg/ml were compared after buffer subtraction and removal of noise from the beamstop parasitic scatter in low q . Estimated molecular weights were calculated for both concentrations using publicly available SAXS Molecular Weight package 2.0 (SAXSMoW) (Piiadov *et al.* 2018). At 7 mg/ml the calculated molecular weight was 88.8 kDa, and for 3 mg/ml it was 89 kDa, slightly over the expected molecular weight for the TLNRD1 dimer which would be ~ 75.5 kDa. This may be due to some elongation of a small region of the protein which would skew calculated molecular weight.

Initial inspection of the SAXS curves show consistency in distribution between the two curves with no indication of aggregation or inter-particle dispersion in the low q Guinier region (Figure 3.23A). A Guinier plot comparison shows a linear distribution with no deviations from the Guinier fit line again showing that there are no aggregation or

dispersion dependent effects in the sample (Figure 3.23B). Calculated R_g values from the Guinier plot were 37.1 for low and 37.2 for high concentration. Similarly, R_g values calculated from the more accurate $P(r)$ distribution were 36.57 at low and 36.30 at high concentration showing consistency between the two approaches and that there are no concentration dependent effects on the molecule radius of gyration. The distribution shown in the $P(r)$ curve shows a largely bell-curve shape with some minor undulations and a long trailing tail at high r values (Figure 3.23C). This reflects the multi-domain nature of TLNRD1 with the N-terminal unstructured region likely to be causing the extended tail at high r . The normalised Kratky plot showed a bell-shaped peak which corresponds to the folded globular domain of TLNRD1, and gradual slope towards higher $q \cdot R_g$ indicating the presence of an unstructured region which could correspond to the N-terminus. Finally, Porod law was used to establish flexibility of the TLNRD1 protein with a plateau at q^3 indicating partial flexibility in the protein.

3.5.2. *Ab initio* shape determination of the TLNRD1 dimer

SAXS data can be used to reconstruct an estimated shape for the protein of interest, allowing easy visualisation of a proteins overall conformation in solution. To explore whether TLNRD1 remains compact in solution or adopts a more extended shape, *Ab initio* modelling was used on the high concentration TLNRD1 data. Initial shape determination was performed using DAMMIF which produces single phase dummy atom model. DAMMIF was used to create 15 envelopes which were averaged to produce the final envelope bead model (Figure 3.25A). This yielded a model which matches the TLNRD1 dimer crystal structure relatively well suggesting that TLNRD1 is in a compact conformation in solution and shows an additional area which protrudes outwards in an antiparallel configuration, matching the expected N-terminal unstructured region positioning. In addition to DAMMIF, GASBOR was also used to create a dummy residue model which takes into account the number of expected amino acids in each asymmetric unit of the protein. This produced a similar shape to the DAMMIF model with an additional density at the top and bottom in an antiparallel configuration (Figure 3.25B). Viewing this model from a top-down view shows the same unique curvature in the structure as that seen in the crystal structure (Figure 3.25). Together these results suggest that TLNRD1 retains a closed compact conformation in solution which is consistent with the crystal structure.

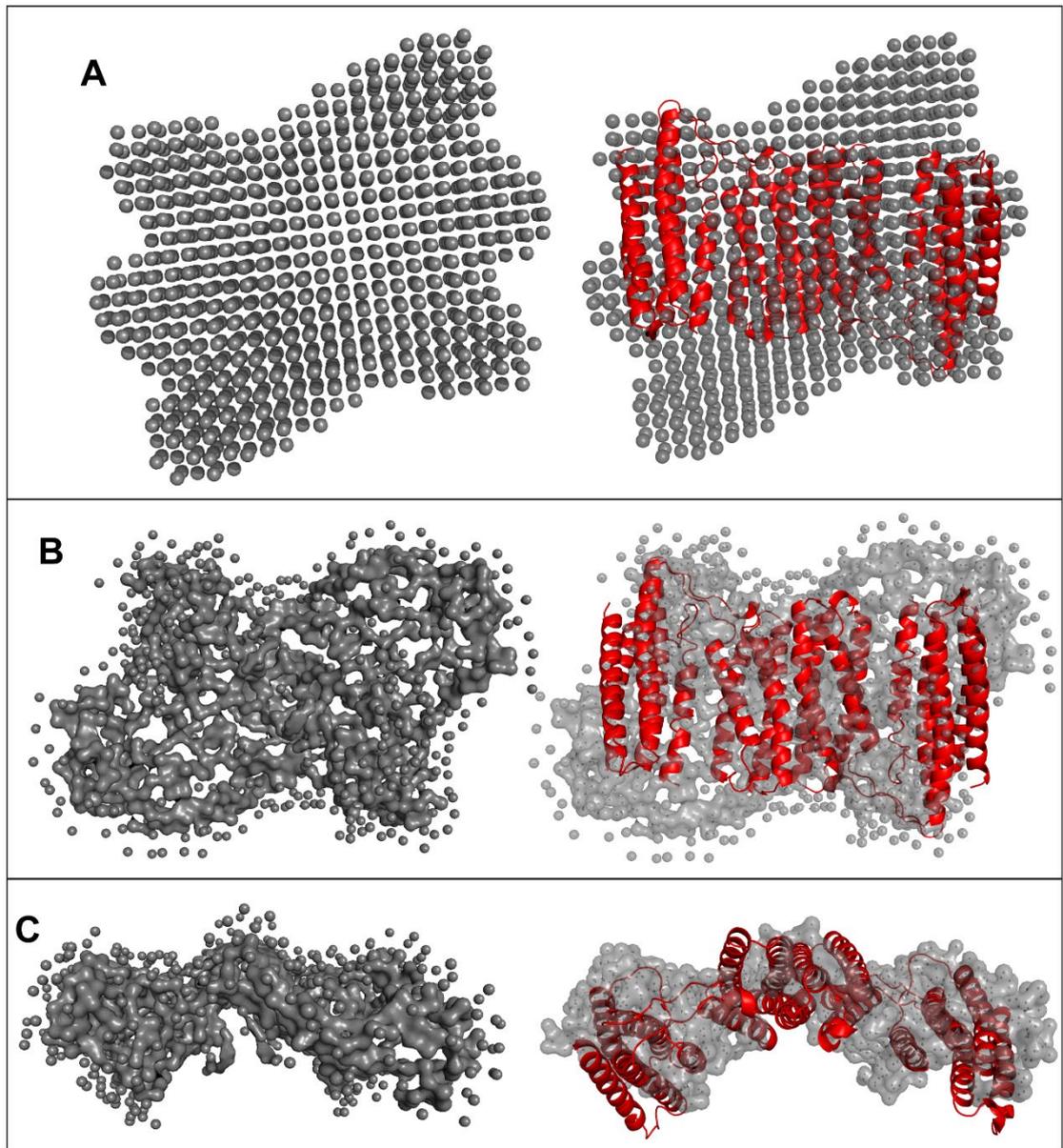


Figure 3.25. *Ab initio* shape determination of the TLNRD1 dimer

(A) DAMMIF averaged bead model of the TLNRD1 dimer with the TLNRD1-FL (red) structure modelled into the bead model on the right **(B)** GASBOR dummy atom bead model of the TLNRD1 dimer with TLNRD1 modelled on the right. **(C)** Top-down view of (B) showing curved nature of the envelope.

3.6. Discussion

The structural and behavioural similarities shown in this chapter raise interesting questions as to the functional origins of TLNRD1. Inspection of TLNRD1 evolution revealed independent losses in nematode, Cnidaria and arthropod lineages but high conservation in higher vertebrates, with the gene first appearing in choanoflagellates and sponges. These lineage specific losses have previously been observed for other components of integrin adhesions, for example integrins themselves are lost in choanoflagellates (Sebé-Pedrós *et al.* 2010). The evolutionary origins of cell adhesion components, including talin and integrins, precedes the emergence of Opisthokonts (eukaryotes including fungi and animals) whereas TLNRD1 emergence occurs after the divergence of choanoflagellates from this group. This suggests that either TLNRD1 function acts independently of adhesion complexes (as integrins, paxillin and FAK are absent in choanoflagellates) or could suggest that its emergence was dependent on the presence of talin alone and may act to modulate talin function. TLNRD1 and talin R7R8 share 22% sequence similarity and high structural homology, suggesting it arose through gene duplication but interestingly they each have drastically different gene structures with talin having multiple splice variants and TLNRD1 only having one transcript. Possible explanations for this striking difference between the talin and TLNRD1 gene structure include a splice variant of mature mRNA encoding just the R7R8 region of talin being reintroduced into an ancestral animal genome, an event which may have occurred by viral insertion to create a new independent gene, lateral gene transfer or by retroposition by retrotransposable elements in eukaryotic genomes (Baertsch *et al.* 2008).

This chapter provides the first unique structural insights into the TLNRD1 protein, showing its close structural relationship with the R7-R8 region of talin and striking modifications which also distinguish it. The structural data presented here reveals that TLNRD1 shares the exact same domain composition as the talin R7R8 domain but with some astonishing and unexpected differences. R7R8 has a unique domain structure in talin whereby the 4-helix R8 domain is inserted between two helices of the 5-helix R7 domain. This specific adaptation of the domain in talin enables it to sit outside the force transmission pathway, protected from mechanical unfolding in response to force by the R7 domain. This ensures that signalling through R8 interaction with certain ligands such as RIAM or DLC1 remains unaffected by the mechanical state of the talin molecule. Seeing this unique domain structure in TLNRD1 suggests that TLNRD1 may have retained that adaptation to withstand some kind of mechanical force, supported by the

high thermostability seen in CD analysis. Interestingly, this stability appears to be almost 15.5°C higher in TLNRD1 compared to R7R8 suggesting the structural differences observed allow TLNRD1 to be more stable and less likely to unfold in response to force. Another notable feature of TLNRD1 which distinguishes it from talin R7R8 is the 40 residues at the N-terminus which appear to be unstructured. Predictions show that this region contains multiple potential phosphorylation sites, and this is supported by published data showing that multiple serine residues are indeed phosphorylated. Three kinase families were predicted to phosphorylate these N-terminal residues, including Vaccinia-related kinase (VRK), Tau-tubulin kinase (TTBK) and Protein kinase A (PKA). Each of these have been implicated in different cellular processes with members of each family showing different localisation and expression profiles across different tissues. For example, VRK family consists of three members; VRK1, VRK2 and VRK3, each with different cellular localisation patterns and has been implicated in different processes including mitosis and transcription (Nichols and Traktman, 2004; Kang *et al.* 2008). It would be interesting for future work to explore how TLNRD1 is regulated and determine how phosphorylation of the N-terminus alters protein behaviour by creating phosphomimetics or preventing phosphorylation of these sites.

While it was originally predicted that TLNRD1 is a dimer from gel filtration, this comprehensive analysis of TLNRD1 reveals that it exists as an obligate dimer with no indication of forming a monomer except at extremely low nanomolar concentrations. It was originally thought that the N-terminal unstructured region of TLNRD1 may be responsible for this dimerisation as it is the only region which substantially differentiates it from talin R7R8, however, here it was discovered that TLNRD1 dimerises via an extensive interface on the 4-helix domain in an intriguing symmetric antiparallel configuration. This mode of dimerisation appears to be highly conserved throughout the proteins evolution even from the choanoflagellates and sponges, suggesting that TLNRD1 has always existed as a dimer and that this dimer configuration is crucial for its function. Interestingly, there was also consistent evidence of TLNRD1 forming a tetramer in solution and the ratio of dimer to tetramer always appeared the same when separating the states on gel filtration. Inspection of the structure shows no indication of how the tetramer may be forming, however, preliminary SAXS data (not shown) of the TLNRD1 tetramer indicates that it is the N-terminal unstructured region which mediates TLNRD1 tetramerisation. Further work should focus on repeating SAXS experiments to confirm N-terminal involvement and be supplemented by SEC-MALS analysis of N-terminal deletion mutants. If the N-terminal region of TLNRD1 is indeed responsible, it would also

be interesting to explore the importance of this in the context of the identified phosphorylation sites as this could be a way for TLNRD1 higher oligomer formation to be regulated in cells.

A final observation from this chapter is that the X-ray crystal structure of full-length TLNRD1 shows the protein adopting an apparent closed conformation, as opposed to talin R7R8 which appears extended and open in all published crystal structures to date. To explore this further and determine TLNRD1 flexibility in solution, SAXS was used to resolve protein globularity and check for elongation which would indicate an opening up of the domains. SAXS is a highly sensitive technique but the low-resolution data it produces must be taken with a pinch of salt and can sometimes lead to misleading results. That being said, the SAXS analysis presented here coincides well with the crystal structure of TLNRD1 showing that TLNRD1 is a globular dimer with an unstructured component and has the same consistent 'closed' structural conformation as seen in the crystal structure. This leads us to question whether the structure can open up, and if so, what are the factors that can induce this kind of conformational change in TLNRD1. No obvious predicted phosphorylation sites were identified on the interface between the 4-helix and 5-helix domains and there were no residues which may be creating a bond with the opposing domain. This could suggest that the N-terminus, not visible on the crystal structure, may promote this configuration of TLNRD1, or that an external factor such as force or interaction with specific ligands is required to promote an extension of the domains.

Chapter 4. Exploring TLNRD1 Interactions with Talin R7R8 Ligands

4.1. Overview

Talin is a large 250 kDa adapter protein which has a diverse range of ligands, binding of which is regulated by force-dependent changes in talin conformation and post-translation modification. The R7R8 region alone acts as a ‘signalling hub’ within talin, forming part of the second actin binding site with R4 and containing binding sites for multiple ligands such as vinculin, RIAM, KANK1 and CDK1, many of which contain conserved LD binding motifs. With the high similarity between TLNRD1 and talin R7R8, we set out to investigate whether TLNRD1 can interact with the same ligands as talin R7R8.

4.2. Leucine-Aspartic acid motifs

Leucine-Aspartic acid motifs (LD), first identified in paxillin, participate in a specific set of protein-protein interactions and are defined by their consensus sequence, LDxLLxL (x is any amino acid) and short α -helical structure (Brown, Perrotta and Turner 1996). LD motif interactions have been identified as important for the regulation of adhesion dynamics, cytoskeletal remodelling and mRNA transport to adhesion sites, coordinating intracellular responses to cues from the ECM (Alam *et al.* 2014). The LD motif binding domains (LDBD) of interacting proteins such as FAK, vinculin or talin are typically helical bundles containing four α -helices, which mediate interaction with LD motifs via helix addition (Figure 4.1). This LD motif recognition system is conserved throughout evolution and is so ingrained in the biology of cells that viruses have adapted mechanisms to hijack the system. By producing proteins which mimic LD motifs or LD binding domains, viruses can alter cell growth, adhesion and cytoskeletal organisation (Pol, Brown and Turner 1998; Zanier *et al.* 2013; Whitewood *et al.* 2018).

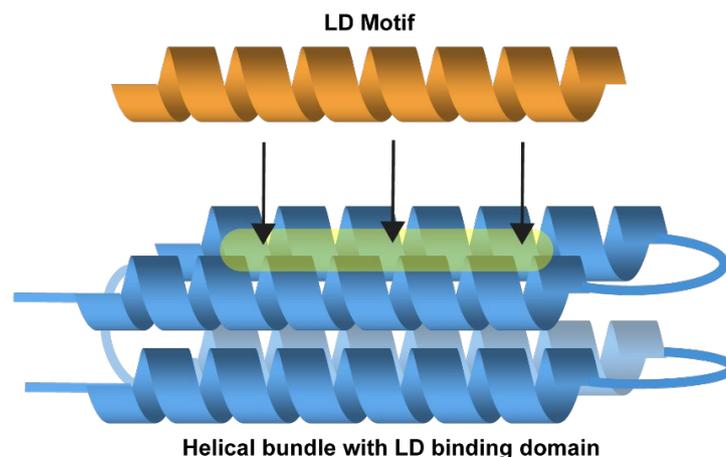


Figure 4.1. LD binding mechanism

Schematic representation of LD motif (orange) binding via helix addition to the helical bundle of a LD binding domain containing protein (blue). Binding region is highlighted in yellow.

4.3. Talin interactions with LD motifs

The rod domain of talin has a complex network of interactions with proteins which regulate adhesion and cytoskeletal dynamics. Most notably, the R7R8 domain of talin contains binding sites for multiple ligands including the tumour suppressor DLC1, scaffolding protein paxillin and small GTPase RIAM which interact with the R8 domain, and KANK proteins which interact with the R7 domain (Zacharchenko *et al.* 2016; Goult *et al.* 2013a; Bouchet *et al.* 2016; Sun *et al.* 2016). A feature consistent with all these interacting proteins is the presence of a canonical LD binding motif in the talin binding region (Figure 4.2).

Position	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
Consensus	X	X	E/D	L	D	X	L	L	X	X	L	X	X
RIAM (Zacharchenko, <i>et al.</i> 2016)	S	E	D	I	D	Q	M	F	S	T	L	L	G
Lpd (Lee <i>et al.</i> 2009)	L	G	E	L	D	R	L	T	Q	S	L	D	S
DLC1 (Zacharchenko, <i>et al.</i> 2016)	F	P	E	L	D	D	I	L	Y	H	V	K	G
KANK1 (Bouchet, <i>et al.</i> 2016)	Q	L	D	L	D	F	V	K	Y	V	D	D	I
KANK2 (Sun, <i>et al.</i> 2016)	R	L	D	L	D	F	V	K	Y	V	D	D	I
CDK1 (Unpublished)	D	S	E	I	D	Q	L	F	R	I	F	R	A
Paxillin LD1 (Zacharchenko, <i>et al.</i> 2016)	M	D	D	L	D	A	L	L	A	D	L	E	S

Figure 4.2. LD motif comparison of talin R7R8 ligands

Amino acid sequence alignment of talin interacting LD motifs. Associated published research is indicated. Acidic residues=red, basic residues= blue, hydrophobic residues= green. Sequences obtained from UniProt: Q7Z5R6 (RIAM), Q70E73 (Lpd), Q96QB1 (DLC1), Q14678 (KANK1), Q63ZY3 (KANK2), P06493 (CDK1) and P49023 (Paxillin LD1).

4.4. Results section 1: TLNRD1 interacts with RIAM and Lpd via the 4-helix domain

Previous research has identified talin binding LD motifs at the N-terminus of both RIAM and Lpd, with an N-terminal sequence homology of ~47% (Lee *et al.* 2009). Talin R8 contains one of 5 RIAM binding sites within talin and unpublished research has also shown that Lpd interacts with this R8 domain. With the high sequence and structural similarity between talin R8 and TLNRD1-4H, the binding potential of both RIAM and Lpd were tested against TLNRD1-4H using both FP assays and NMR spectroscopy. This revealed a high affinity interaction between TLNRD1 and RIAM, with a reduced but still substantial affinity for the RIAM paralog Lpd.

4.4.1. Structural comparison of TLNRD1 and the talin R8 RIAM binding site

The interaction between RIAM and talin has previously been mapped to N-terminal residues 6-30 of RIAM, with binding to domains R2, R3, R8 and R11 of talin (Goult *et al.* 2013a). Following this, an X-ray crystal structure of talin R7R8 in complex with a RIAM peptide spanning residues 5-25 revealed that RIAM binds to helices H32 and H33 in the R8 domain (Chang *et al.* 2014). This structure revealed a unique 55° kink in the RIAM helix when bound to talin R8 which does not appear in other domain structures complexed with RIAM. Comparison of the amino acid sequence between this RIAM binding site in the talin R8 domain and TLNRD1 using sequence alignment against full-length TLNRD1, reveals significant similarity with the equivalent helices $\alpha 5$ and $\alpha 6$, of the TLNRD1 4-helix domain (Figure 4.3A).

Mapping of identical (dark blue) and similar residues (light blue) on the talin R8 structure (PDB ID: 2X0C (Gingras *et al.* 2010)) in unbound form and the TLNRD1 4-helix domain show that the residues are also positioned in highly similar positions along the helices. To explore whether RIAM could adopt the same LD binding mechanism with TLNRD1, the RIAM peptide from the R8 complex structure (PDB ID: 4W8P (Chang *et al.* 2014)) was modelled with TLNRD1-4H with the assumption that TLNRD1 can interact with RIAM using the same binding surface. RIAM residue F12 docks into a pocket created by residues A1499 and V1540 on R8, whereas on TLNRD1 this pocket is created by similar aliphatic residues L191 and A225. RIAM residue L22 also binds in a pocket created by R1510, A1529 and K1530 on R8 while on TLNRD1 an equivalent pocket is created by residues A199, S202 and K215. These distinct similarities suggest that like talin R8, TLNRD1 may also interact with RIAM using the same binding surface.

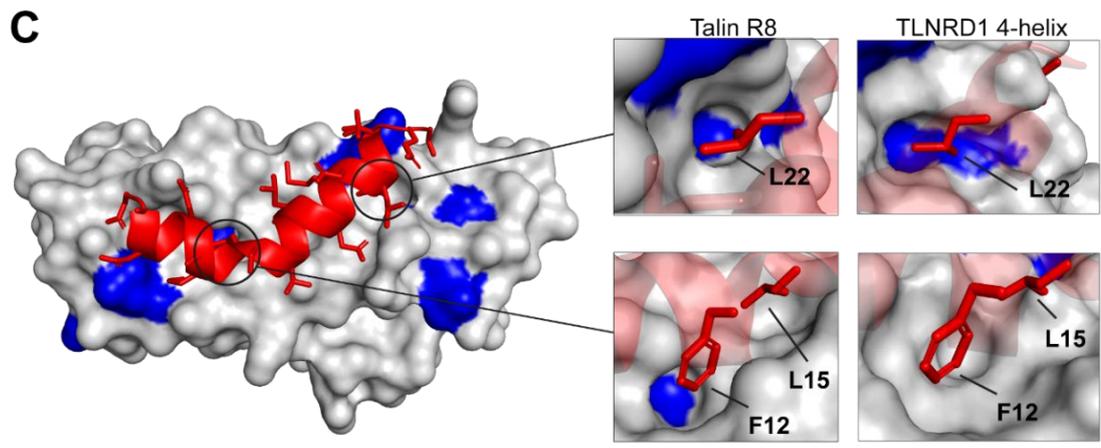
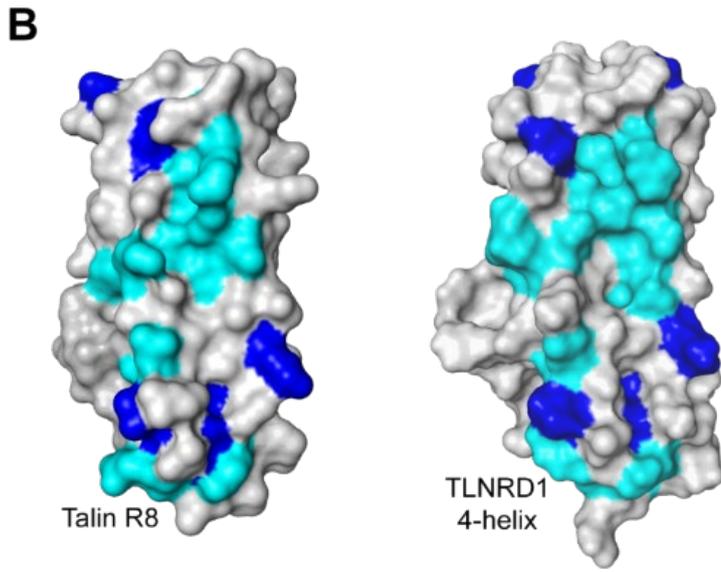
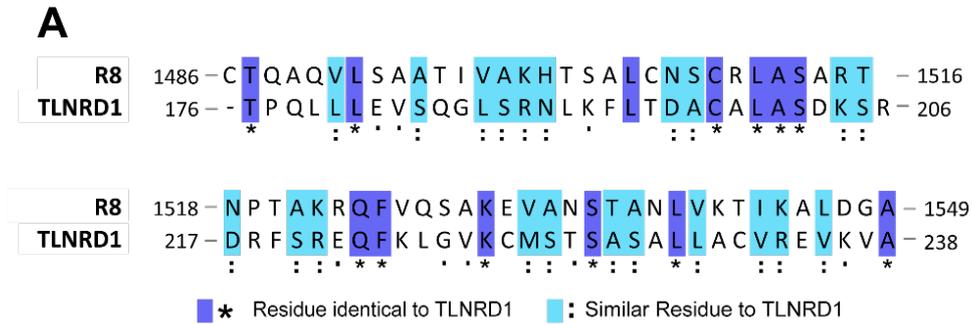


Figure 4.3. Comparison of RIAM binding sites between TLNRD1 and talin

(A) Human sequence alignment of residues 1486-1516 and 1518-1549 of talin R8 domain against the corresponding helices in TLNRD1. Residues in dark blue are identical between R8 and TLNRD1, with light blue indicating similarity. Alignment was created using Clustal Omega available from EMBL-EBI (Sievers et al. 2011), **(B)** Surface structure representation of talin R8 domain (PDB: 4W8P) and TLNRD1-4H domain with residue similarity highlighted according to the above sequence alignment, **(C)** Structure of talin R8 domain bound to RIAM residues 5-25 (PDB ID: 4W8P), dark blue indicates identical residues to TLNRD1. Close-ups shown of RIAM F12 and L22 binding sites on the R8 surface, with comparison to same regions in TLNRD1.

4.4.2. Fluorescence polarisation reveals TLNRD1 interaction with RIAM

To explore whether TLNRD1 can interact with RIAM we first used fluorescence polarisation assays to test both full-length TLNRD1 and the 4-helix domain against a RIAM peptide spanning residues 4-30, which contains the previously characterised LD motif that mediates binding to talin via helix addition. The RIAM peptide was coupled with fluorescein dye via an additional C-terminal cysteine added to the peptide sequence, and first tested against TLNRD1-FL at a starting concentration of 75 μM .

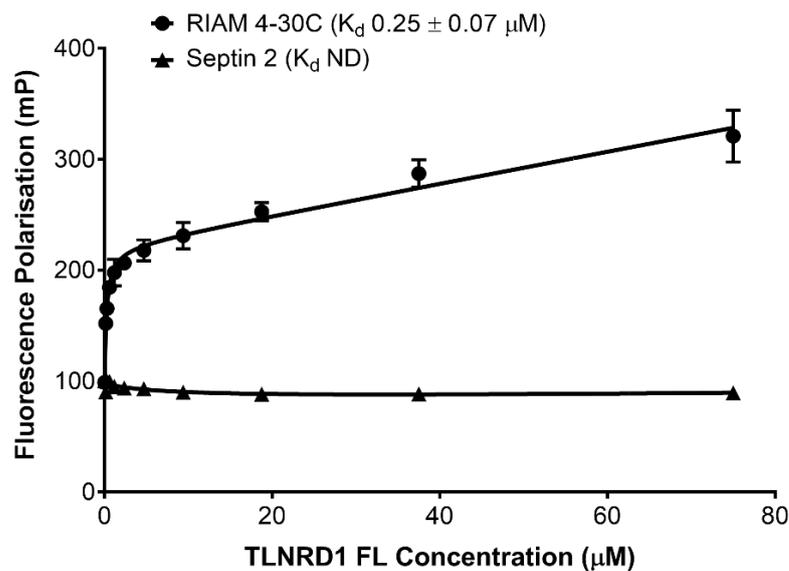


Figure 4.4. Full-length TLNRD1 binding with RIAM peptide spanning residues 4-30

FP binding assay of TLNRD1-FL against fluorescein labelled RIAM peptide spanning residues 4-30. The interaction dissociation constant (\pm SE) in μM is indicated in the graph legend. Binding experiments were repeated in triplicate. ND= Not defined.

The binding affinity of TLNRD1-FL and the RIAM peptide is shown in Figure 4.4. The FP assay reveals that TLNRD1 can interact with the RIAM LD motif peptide with high affinity with a K_d of 0.25 μM (\pm 0.07). To narrow down whether the 4-helix domain of TLNRD1 is mediating this interaction, the TLNRD1-4H construct was also tested against the RIAM peptide at a starting concentration of 78 μM . Again, this revealed a high affinity interaction with a K_d of 0.59 μM (\pm 0.02) (Figure 4.5), demonstrating that the TLNRD1 RIAM binding site resides within the 4-helix domain as predicted from the structural similarity with the talin R8 binding site.

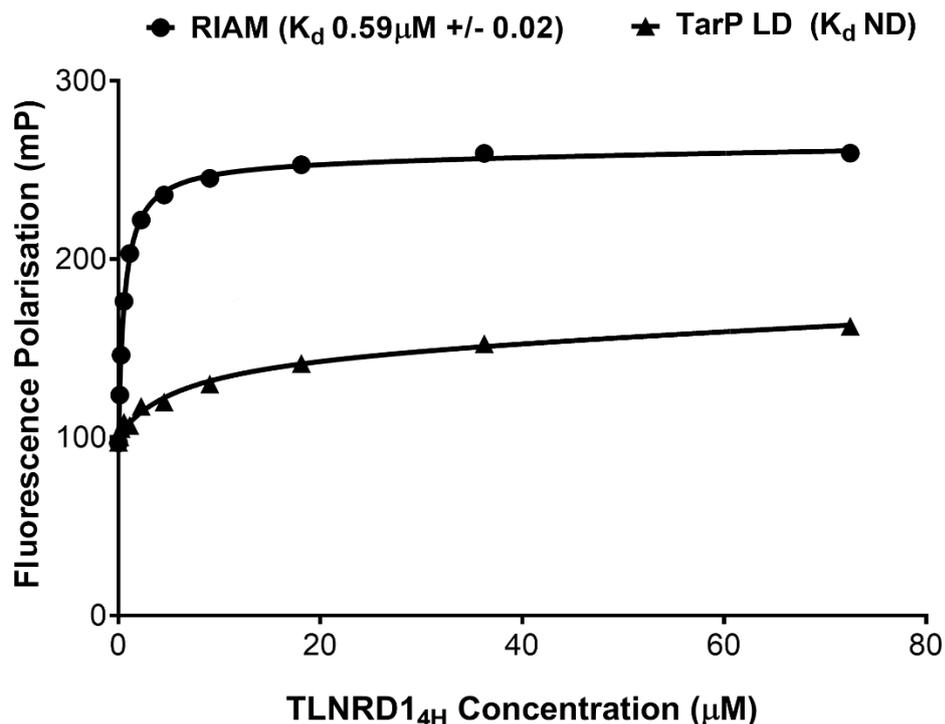


Figure 4.5. TLNRD1 4-helix domain interacts with RIAM peptide spanning residues 4-30

FP binding assay with TLNRD1-4H against fluorescein labelled RIAM 4-30 peptide with TarP LD motif as the negative control (Whitewood *et al.* 2018). The dissociation constant is indicated in the figure legend (\pm SE). Experiment was repeated in triplicate. ND= Not defined. N=3

4.4.3. HSQC titration of TLNRD1 4-helix domain with RIAM peptide

To validate and further explore the TLNRD1 interaction with RIAM, 2D NMR ^1H ^{15}N HSQC titrations were used. This allows us to visually explore changes in the chemical environment of each amino acid, allowing identification of specific binding sites, conformational changes and allosteric changes in response to ligand binding. This is a highly sensitive and dynamic technique which measures the chemical exchange between two nuclei, changes in which can be measured as a difference in chemical shift/relaxation rate. Depending on the rate of chemical exchange, a molecule can be in either slow, intermediate or fast exchange (Figure 4.6). When an interaction is in slow exchange, two signals will be produced on the NMR spectra as there is an equal mixture of proteins which are in a bound (δ_A) or unbound (δ_B) conformation, this is typical of molecules with strong interactions. Molecules in intermediate exchange are indicative of moderate interaction strength. In this state peak signals begin to broaden and can disappear when a ligand is added, but as the ligand concentration is increased the signal can be recovered if saturation is reached. Finally, weaker interactions produce a single peak in

a fast exchange regime, which is optimal for tracking amino acid interactions with a ligand as the chemical shift pattern on the NMR spectra is clear and easily tracked.

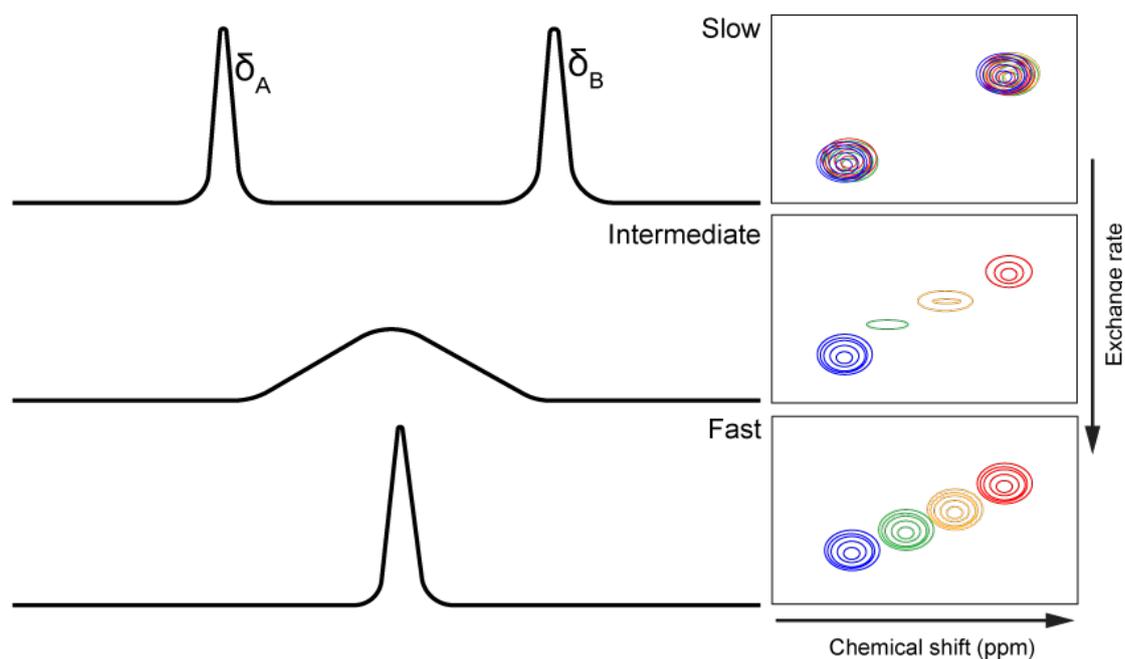


Figure 4.6. NMR exchange rates and chemical shift patterns

Schematic showing the relationship between chemical exchange rates and chemical shift spectra. Typical shift patterns for slow, intermediate and fast exchange are shown, with blue representing unbound, and increasing ligand concentration from green, orange and finally red when full saturation is reached.

For these experiments TLNRD1-4H was labelled with ^{15}N isotope as described in Materials and Methods section 2.7.6. Correlation was measured between the endogenous ^1H proton and the ^{15}N of each atom in the molecule. An initial baseline spectrum of TLNRD1-4H alone was obtained at $60\ \mu\text{M}$ (Figure 4.7) which showed defined well-dispersed peak resonances confirming that the protein was folded, with many of the peaks falling in the range (ppm) expected of α -helical protein (Wang and Jardetzky 2002).

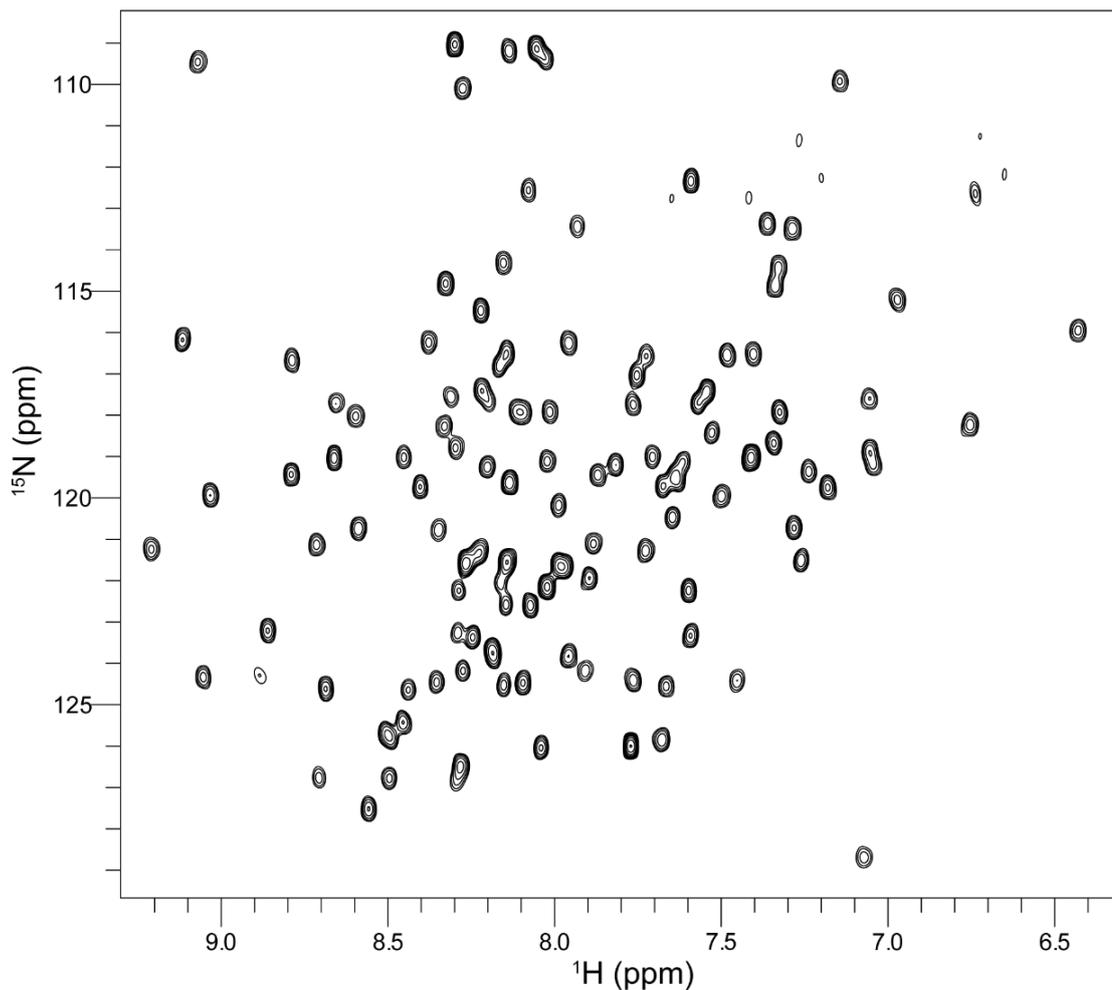


Figure 4.7. ^1H ^{15}N HSQC spectra of TLNRD1-4H

NMR spectra of ^{15}N labelled TLNRD1 4-helix module spanning residues 144-273. Spectra was obtained at $60\mu\text{M}$ in 20mM phosphate pH 6.5, 50mM NaCl, 2mM DTT at a temperature of 298K and field strength of 600MHz .

RIAM peptide 4-30 was titrated with TLNRD1-4H at a ratio of 0.5:1 initially (blue), followed by a 2:1 ratio (red). An overlay of the HSQC spectra of TLNRD1-4H alone vs the two titrations of RIAM peptide are shown in Figure 4.8. Comparison of spectra from the three experiments reveal evidence of both fast and intermediate exchange, with a large number of observable chemical shift changes which is consistent with an interaction. At 0.5:1 small chemical shifts were observed in several peaks in the spectra with all peaks remaining visible. When the ratio was increased to 2:1, the spectra showed drastic changes in distribution of some peaks, with an increase in the number of observable shifts. Some peaks were found to be in fast exchange, with chemical shift changes easily tracked with the increasing amount of ligand (Figure 4.8B). With a 2:1 peptide:protein ratio, some peaks began to disappear which is typical of a transition into intermediate

exchange, indicating a moderately strong interaction between TLNRD1-4H and RIAM LD motif (Figure 4.8C).

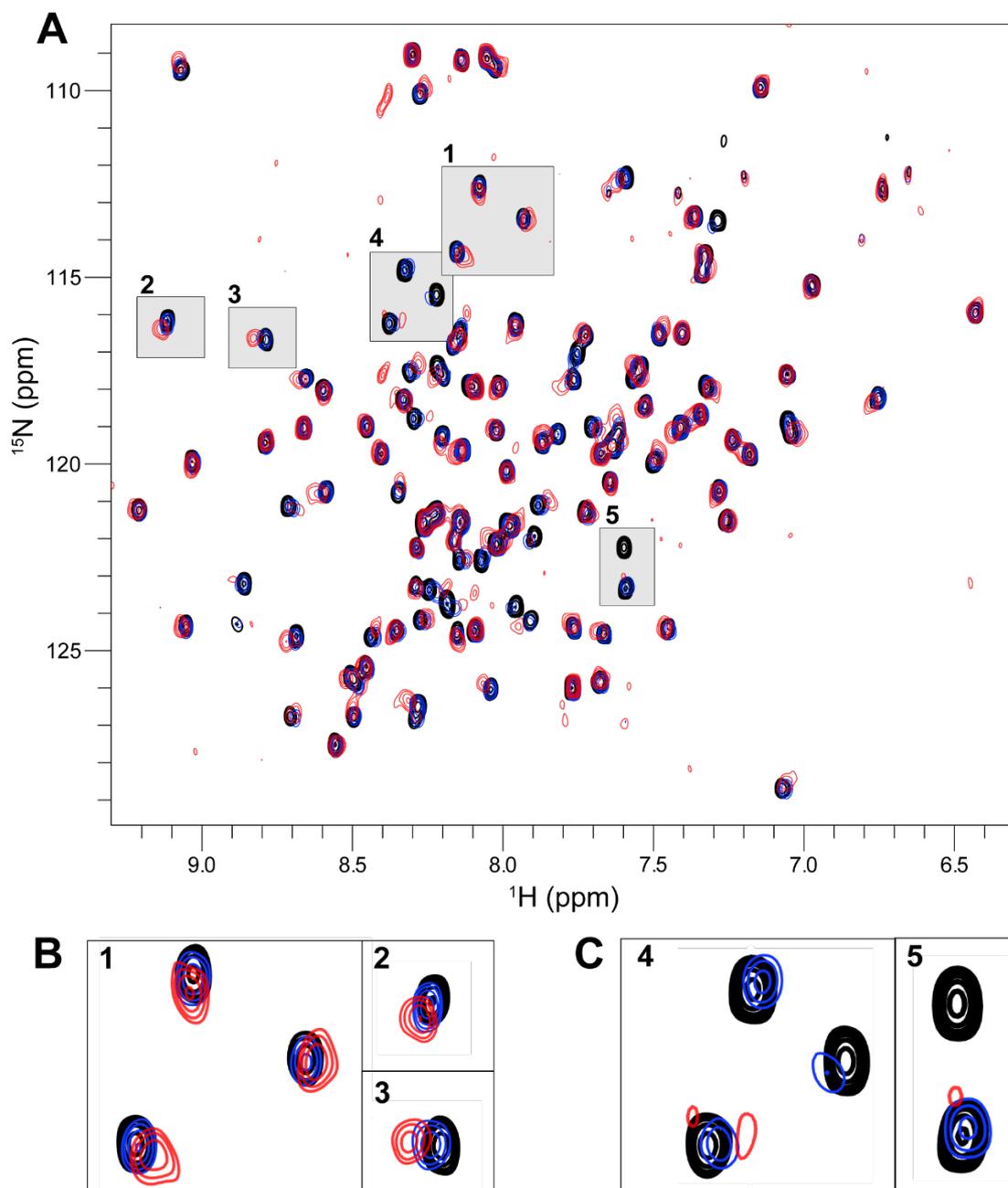


Figure 4.8. ^1H ^{15}N HSQC TLNRD1-4H titration against RIAM peptide 4-30

(A) ^1H ^{15}N HSQC spectra of 60 μM TLNRD1-4H in the absence of peptide (black) with titrations of RIAM peptide 4-30 at a peptide:protein ratio of 0.5:1 (blue) and 2:1 (red). **(B)** Zoomed in view showing amino acid chemical shifts in fast exchange with corresponding location highlighted in A. **(C)** Zoomed in view showing amino acid peak disappearance as they enter intermediate exchange with increasing concentration of peptide. Corresponding spectra location highlighted in A.

4.4.4. TLNRD1 4-helix domain binds the RIAM paralog Lamellipodin

As with RIAM, the paralogue lamellipodin contains an N-terminal talin binding site between residues 20-46. Sequence alignment of the RIAM talin binding LD motif and lamellipodin reveals high sequence similarity between the two regions and a highly similar LD motif in lamellipodin (Figure 4.9). Secondary structure prediction of both RIAM and lamellipodin LD motif regions reveal a predicted single amphipathic α -helix, consistent with the known LD binding mechanism via helix addition.

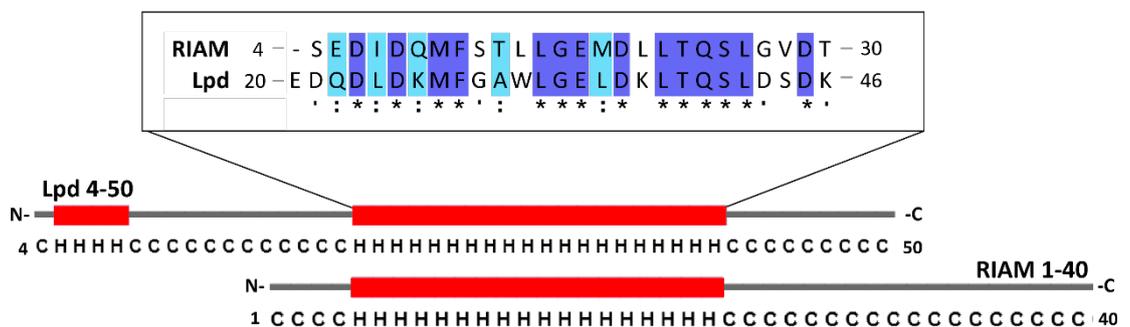


Figure 4.9. LD motifs of RIAM and Lamellipodin

Sequence alignment showing that lamellipodin contains a highly similar LD motif (residues 20-46) to RIAM LD motif (residues 4-30). Structural prediction of the LD region using PsiPred reveals that this region is also predicted to be alpha-helical.

Given that lamellipodin can interact with talin using the same LD motif recognition mechanism as RIAM, and we have now shown that RIAM can directly interact with TLNRD1, we next set out to determine whether lamellipodin can also bind to TLNRD1. Using fluorescence polarisation, TLNRD1-4H with a starting concentration of $110\mu\text{M}$ was tested against Lpd peptide spanning residues 20-46 with an additional C-terminal cysteine for fluorescein coupling. The assay revealed that TLNRD1 does interact with Lpd but with a lower affinity than RIAM with a K_d of $7.9\mu\text{M}$ ($\text{SE} \pm 1.69\mu\text{M}$) (Figure 4.10).

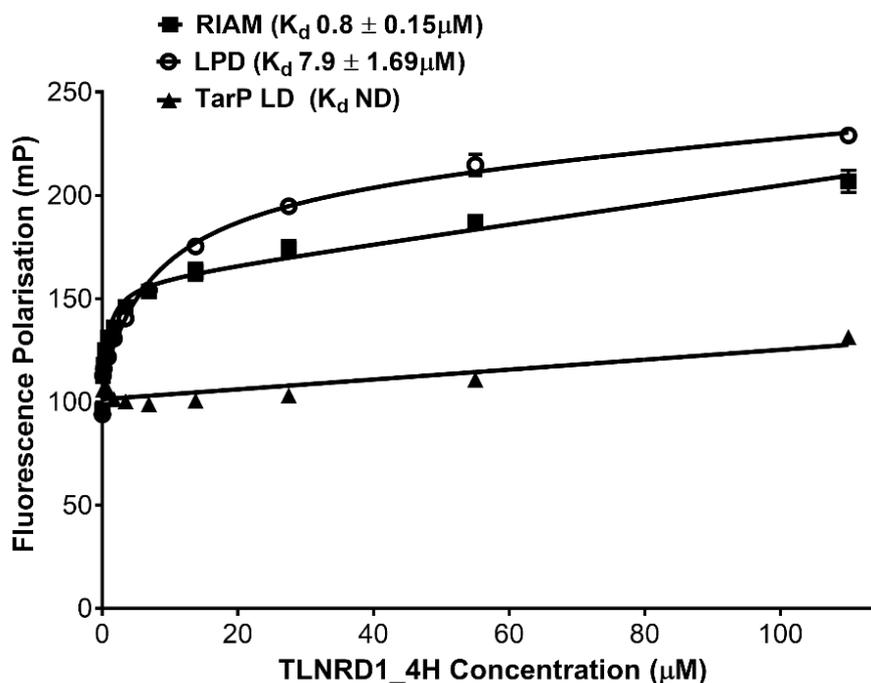


Figure 4.10. TLNRD1 4-helix domain interacts with Lpd peptide spanning residues 20-46

FP binding assay with TLNRD1-4H against fluorescein labelled Lpd 20-46 peptide with TarP LD motif as the negative control and RIAM as the positive control. The dissociation constant is indicated in the figure legend (\pm SE). Experiment was repeated in triplicate. ND= Not defined. N=3

4.4.5. Designing a TLNRD1 mutant to disrupt RIAM binding

With the knowledge that the TLNRD1 4-helix domain contains a binding site for the same LD motif of RIAM which mediates talin interaction, we set out to design a mutant within the 4-helix domain to perturb RIAM binding. This will not only validate the location of the RIAM binding site within TLNRD1 but provide a construct for future functional studies in a cellular context. For this, a double mutant was designed with the aim of using charge reversal to repel binding of the RIAM peptide whilst also maintaining structural integrity of the protein.

Close inspection of the TLNRD1-4H crystal structure reveals two positively charged basic residues, lysine 192 and arginine 233, sitting either side of the potential RIAM binding site, creating a basic surface area for protein-ligand interaction (Figure 4.11; Figure 4.12A). The positioning of these residues also makes them potential targets for ionic bond formation with the RIAM peptide. Modelling of the RIAM peptide with the predicted RIAM binding site on TLNRD1-4H positions RIAM E18 within close proximity to TLNRD1 K192, allowing the potential for ionic bond formation which could stabilise the TLNRD1-RIAM interaction (Figure 4.11B). Furthermore, TLNRD1 residue R233 sits close to I8 and D9 of

RIAM, key LD motif residues which are crucial for interaction with talin R8 LD binding region (Figure 4.11C). Previous studies have shown that by mutating either D8 or E13 of RIAM to alanine disrupts the talin R8-RIAM interaction, further supporting the theory that TLNRD1 binds using a similar mechanism (Chang *et al.* 2014). To explore whether these residues are important for mediating the TLNRD1-RIAM interaction, we mutated both K192 and R233 in TLNRD1-4H to glutamic acid (TLNRD1-2E), an amino acid similar in size but with a negative charge. Inspection of the surface electrostatic potential reveals these mutations abolish the positively charged basic area which may be important for the talin-RIAM interaction (Figure 4.12B).

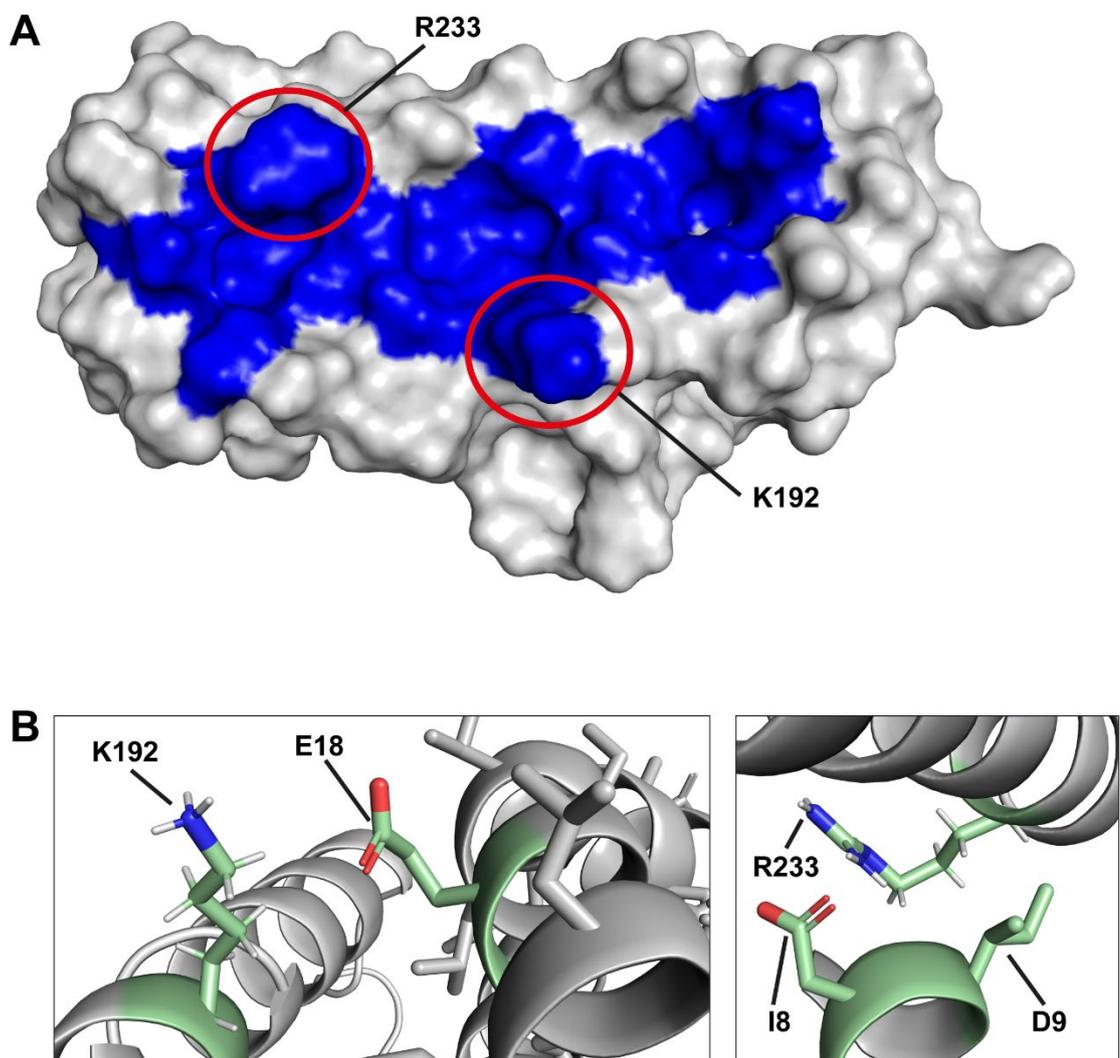


Figure 4.11. Identifying TLNRD1 residues which mediate TLNRD1-RIAM interaction

(A) Surface structure of TLNRD1 4-helix domain with predicted RIAM binding region highlighted in blue, positions of amino acids K192 and R233 are outlined in red. (B) Left= Residue K192 highlighted in green, sits within close proximity to E18 of the modelled RIAM peptide (green; PDB ID: 4W8P). Right= Residue R233, highlighted in green, sits within close proximity of I8 and D9 of RIAM.

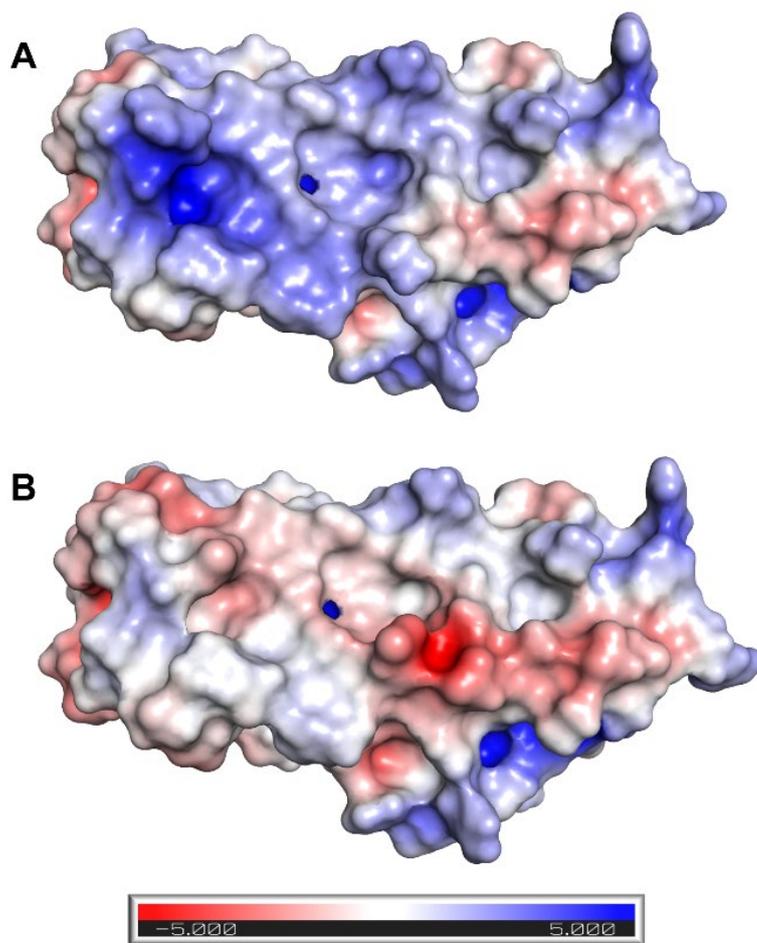


Figure 4.12. Comparison of surface electrostatics between wild-type and 2E TLNRD1

(A) Representation of wild-type TLNRD1-4H surface electrostatics in the predicted RIAM binding site.
 (B) Surface electrostatics after the introduction of K192E and R233E mutations.

Both wild-type TLNRD1 4-helix domain and 2E mutant were dialysed into the same batch of PBS with 2mM DTT and concentrated equally to a final concentration of 50 μM . Both were tested against fluorescein labelled RIAM 4-30 in a fluorescence polarisation assay to compare binding affinities. Wild-type TLNRD1-4H bound as expected with a high K_d of 0.46 μM (\pm 0.58), whereas introduction of the K192E and R233E mutations drastically reduced TLNRD1-4H binding to the RIAM peptide with no generated K_d (Figure 4.13). This loss of affinity for RIAM supports our theory that TLNRD1 interacts with RIAM LD motif using the same binding surface as talin R8 and provides a useful mutant for future work exploring the role of the interaction in cell biology.

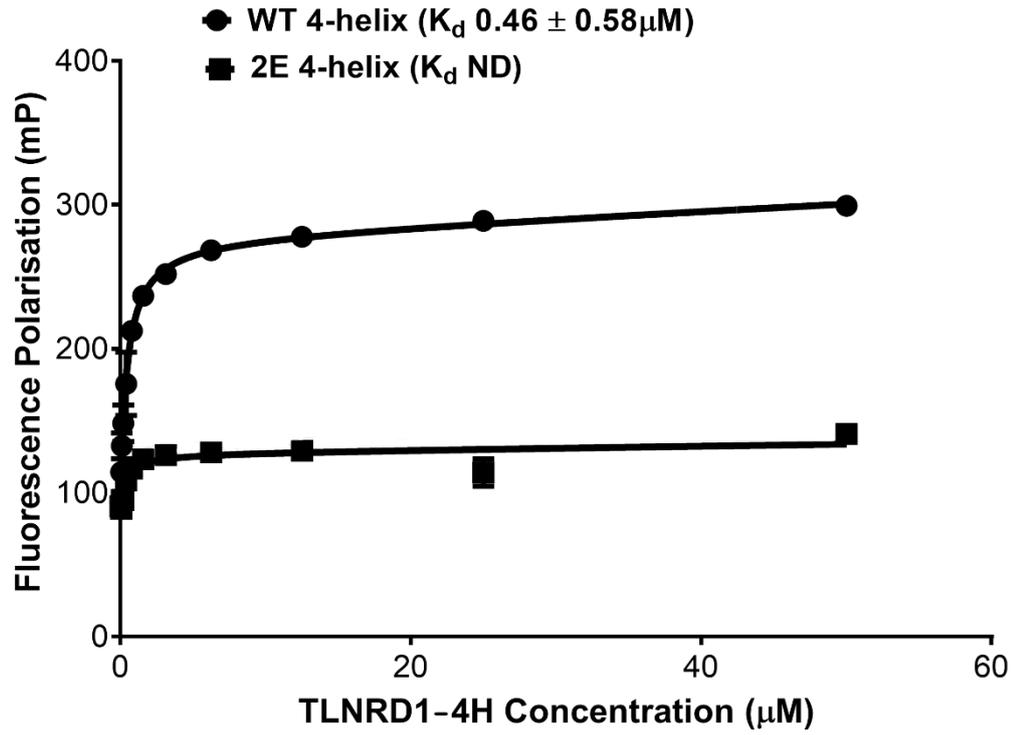


Figure 4.13. TLNRD1-2E no longer interacts with RIAM peptide

Fluorescence polarisation graph of wild-type vs 2E mutant binding to RIAM 4-30. K_d values \pm SE are indicated in the legend. Experiment was performed in triplicate. ND= Not defined.

4.5. Results section 2: TLNRD1 interaction with KANK proteins

Recently, KANK1 has been identified as a key talin R7 ligand for the formation of cortical microtubule stabilising complexes on the periphery of focal adhesion complexes (Bouchet *et al.* 2016). This interaction is mediated by an unusual double LD-motif in the KANK1 KN domain and is one of the first recognised LD mediated interactions with a 5-helix domain of talin. It has also been recently shown that the KANK1 isoform KANK2 can also interact with talin R7R8 domain via the N-terminal KN domain, promoting talin activation and disrupting ABS2 interaction with F-actin (Sun *et al.* 2016).

As with talin R7R8, TLNRD1 also contains a similar 5-helix bundle but with an additional 40 residue unstructured region at the N-terminus. The structural similarity between R7R8 and TLNRD1 led us to explore whether TLNRD1 has retained the KANK1 binding site as observed with the RIAM binding mechanism. Using previously characterised KANK peptides, we found the TLNRD1 can indeed interact with KANK1 and KANK2 but with a lower affinity than that observed with talin. TLNRD1 interaction with KANK3 could not be measured due to aggregation upon addition of KANK3 peptide.

4.5.1. Structural comparison of talin R7 KANK binding site with TLNRD1

The R7-KANK1 interaction has been narrowed down to the LD motif in KANK1 KN domain (30-68) and helices 29 (1389-1416) and 36 (1627-1657) of talin R7. Alignment of helices 29 and 36 to the same equivalent $\alpha 2$ and $\alpha 9$ helices of TLNRD1 (residues 75-101 and 316-342 respectively) reveals limited sequence similarity (Figure 4.14). While residues which are conserved between TLNRD1 helix $\alpha 2$ and talin helix 29 are retained in the same structural position, residues aligned between TLNRD1 helix $\alpha 9$ and talin helix 36 are not retained, suggesting modification to what is potentially the KANK binding site. Despite this low similarity in amino acid positioning in the KANK binding site, closer inspection of the structure shows that like talin R7, TLNRD1 has residues at each end of the helices, F75, Q97 and H322, which hold them apart at a similar distance to residues Y1389 and W1630 of talin (Figure 4.15). This increased separation between the helices compared to other 5-helix bundles has previously been shown to be important in creating the correct spacing distance of 8Å for KANK LD motif docking (Bouchet *et al.* 2016).

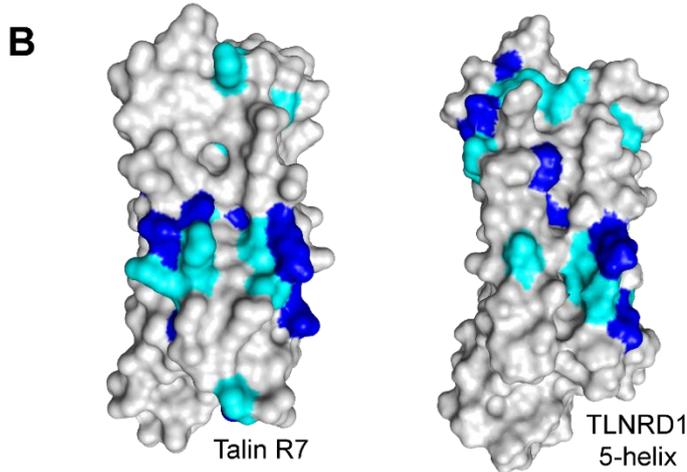
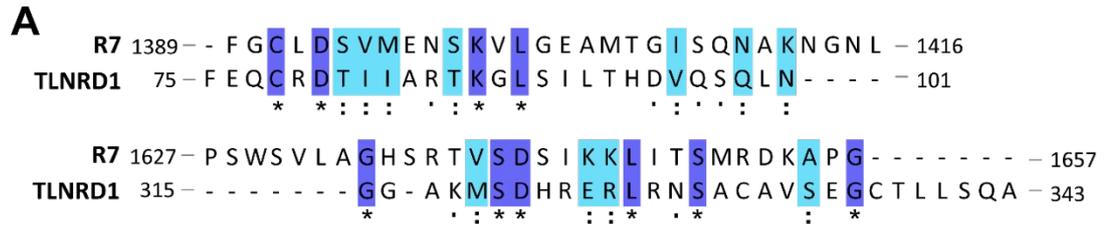


Figure 4.14. Comparison of talin R7 KANK binding domain with TLNRD1 5-helix domain

(A) Human sequence alignment of residues 1389-1416 and 1627-1657 of talin R7 domain against the corresponding helices in TLNRD1. Residues in dark blue are identical between R7 and TLNRD1, with light blue indicating similarity. Alignment was created using Clustal Omega. **(B)** Surface structure representation of talin R7 domain (PDB: 2X0C, (Gingras *et al.* 2010)) and TLNRD1_5H domain with residue similarity highlighted according to the above sequence alignment.

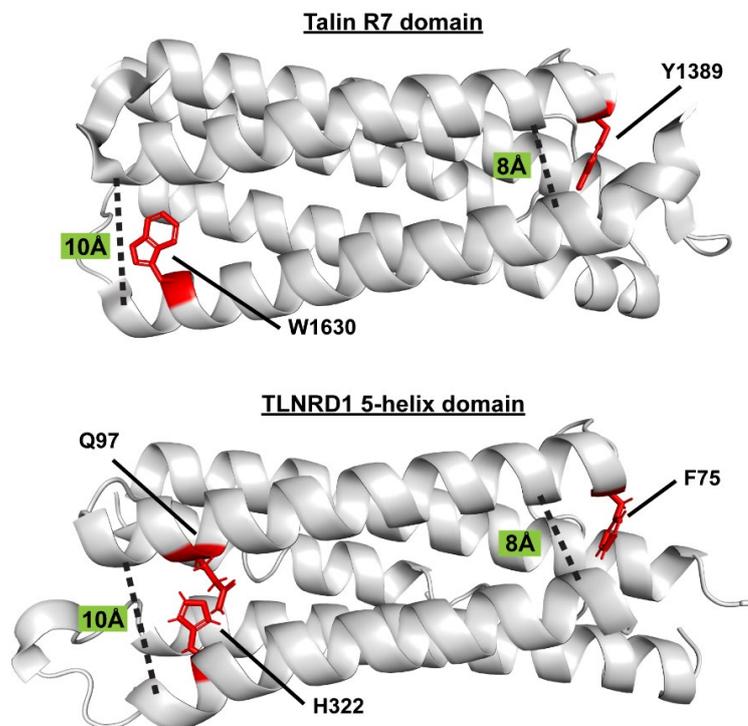


Figure 4.15. Helix arrangement for talin R7 vs TLNRD1

Comparison of helix spacing in the KANK1 binding site with TLNRD1. Residues W1630 and Y1389 are highlighted in R7, and residues F75, Q97 and H322 highlighted in TLNRD1.

To date, there are no published structures available of the KANK1 KN domain for modelling. With the knowledge that LD motifs typically bind via helix addition and can become helical upon interaction with a ligand, the already available DLC1 LD motif in a talin 1-DLC1 crystal structure (PDB ID: 5FZT) (Zacharchenko *et al.* 2016) was used to model the KANK1 LD motif. Using the DLC1 helical structure as a base model, mutagenesis was used in PyMOL to modify the DLC1 sequence to that of KANK1. Modelling of the KANK1 peptide with TLNRD1 was performed using previously published analysis of the talin-KANK1 interaction as a reference. In the talin R7 interaction with KANK1, amino acids K1401 and R1652 are predicted to form a salt bridge with nearby D44 and D42 of KANK1 respectively (Figure 4.16A). The TLNRD1 structure lacks a similar amino acid to R1652 in the equivalent position but does possess a lysine in the same position as K1402 (K87) which was predicted to form an important salt bridge with the KANK1 peptide (Figure 4.16B). This suggests that KANK1 may be able to interact with TLNRD1 using the same mechanism, but the lack of a second positively charged amino acid for salt bridge formation with D42 suggests that the affinity may be reduced compared to the talin-KANK interaction.

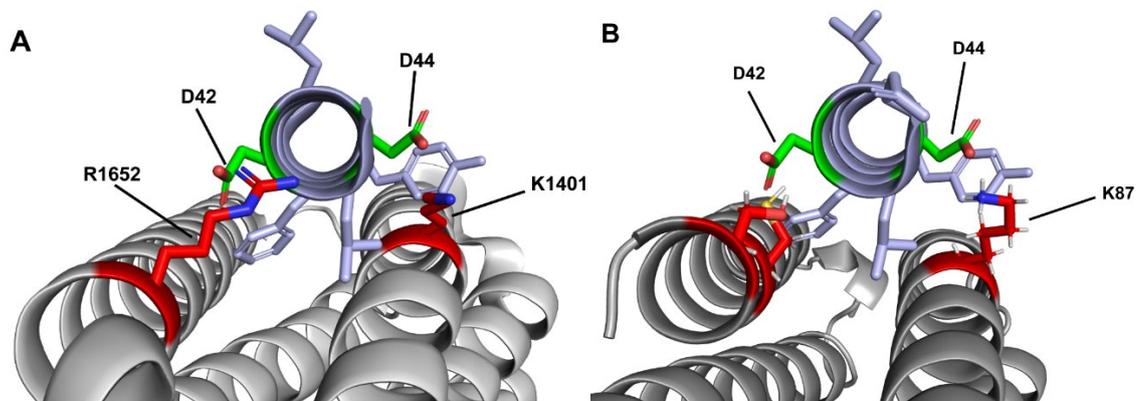


Figure 4.16. Structural modelling of KANK1 LD with talin R7 and TLNRD1

(A) Talin R7 domain (PDB ID: 5FZT) modelled with KANK1 residues 40-54 encompassing the double LD motif. Residues D42 and D44 are highlighted in green. Identified interacting residues K1401 and R1652 highlighted in red. **(B)** TLNRD1 5-helix domain modelled with KANK1 residues 40-54. Residues D42 and D44 are highlighted. Residue K87 of TLNRD1 sits in close proximity to D44 of KANK1.

4.5.2. TLNRD1 interaction with KANK1 LD motif

Using fluorescence polarisation, we sought to test whether TLNRD1 can interact with the same KANK1 LD motif containing peptides as shown through previous publication (Bouchet *et al.* 2016). For this experiment, full-length TLNRD1 protein at a starting concentration of 78 μM was titrated against BODIPY labelled KANK1 peptide spanning residues 30-68 of the KN domain. Previously, the 4 LDLD residues of the binding region in KANK1 was mutated to four alanine's which knocked out binding to talin R7. To test whether TLNRD1 is binding by a similar mode through LD motif recognition, the 4A peptide was also tested in this experiment. The septin 2 peptide was used as a negative control. The binding affinity of the KANK1 30-68 and KANK1 4A peptide is shown in Figure 4.17 below. With the shown similarity between the talin R7R8 KANK1 binding site and the equivalent region in TLNRD1 we expected the 30-68 peptide to bind and the 4A peptide to not bind. TLNRD1-FL protein bound to the KANK1 peptide with an affinity of 12 μM whereas the KANK1 4A peptide showed reduced binding with no K_d generated.

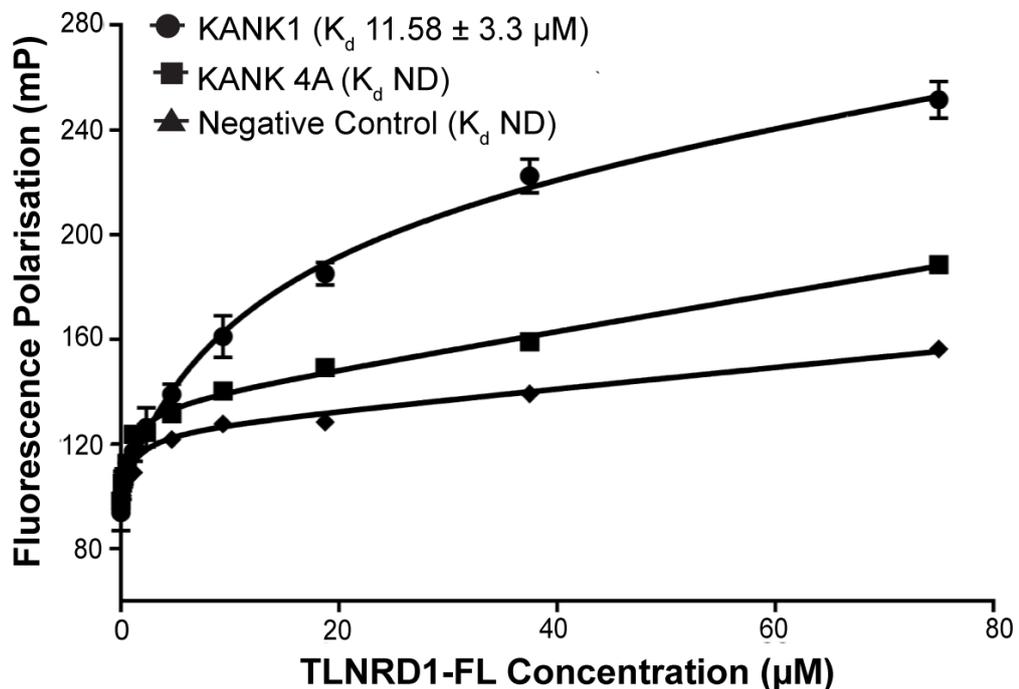


Figure 4.17. Full-length TLNRD1 interacts with KANK1 LD peptide spanning residues 30-68

Fluorescence polarisation binding assay of TLNRD1_FL against BODIPY labelled KANK1 peptide spanning residues 30-68, KANK1 4A mutant peptide and Septin 2 as a negative control. Interaction dissociation constants (\pm SE) in μM are indicated in the graph legend. All measurements were repeated in triplicate with ND indicating that a dissociation constant could not be defined. N=3

4.5.3. KANK1-TLNRD1 interaction is mediated by the 5-helix domain

KANK binding of talin occurs via the 5-helix R7 domain, however, when attempting to express the TLNRD1 5-helix domain in isolation (1-362 Δ 142-273) we found the construct was poorly expressed and highly insoluble. With the TLNRD1 5-helix domain proving too unstable to obtain protein levels required for FP or NMR analysis of binding, we instead tested KANK1 30-68 against the 4-helix domain of TLNRD1 to rule out the presence of the KANK1 binding site being in that domain. In this experiment, TLNRD1-4H with a starting concentration of 50 μ M was titrated against KANK1 30-68 peptide with RIAM peptide 4-30 as a positive control (Figure 4.18). As expected, the TLNRD1 4-helix domain did not bind to the KANK1 LD peptide, suggesting that the KANK1 binding site is on the 5-helix domain of the TLNRD1.

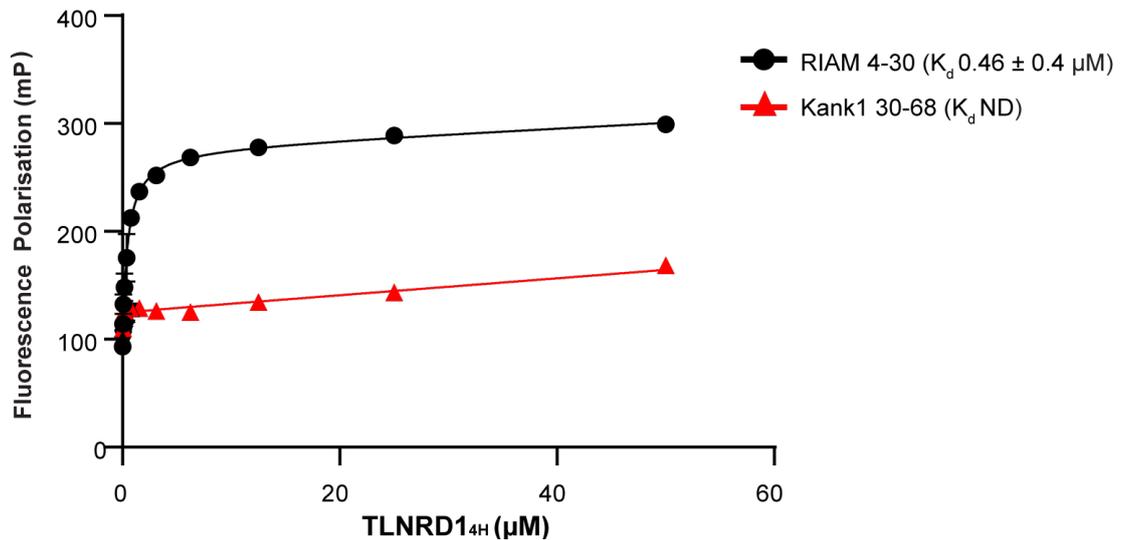


Figure 4.18. TLNRD1 4-helix domain does not interact with KANK1 peptide

Fluorescence polarisation binding assays of TLNRD1_{4H} against BODIPY labelled KANK1 30-68 peptide (red) with RIAM 4-30 as a positive control (black). All measurements were repeated in triplicate with dissociation constants \pm SE (μ M) displayed in the right-hand legend. ND= Not defined.

4.5.4. TLNRD1 interacts with KANK2

The high conservation of the LD region in KANK's enables binding of all isoforms to the R7R8 domain of talin. To explore whether this translates to TLNRD1-KANK interactions, a KANK2 peptide spanning residues 31-69 was tested against TLNRD1 full-length construct using fluorescence polarisation. KANK2 bound to TLNRD1 with a dissociation constant of 3.2 μ M (\pm 0.99), slightly tighter than the previously observed interaction with KANK1 (Figure 4.19). Fluorescence polarisation analysis of the TLNRD1 KANK3

interaction could not be obtained as addition of the equivalent KANK3 LD peptide caused sample aggregation.

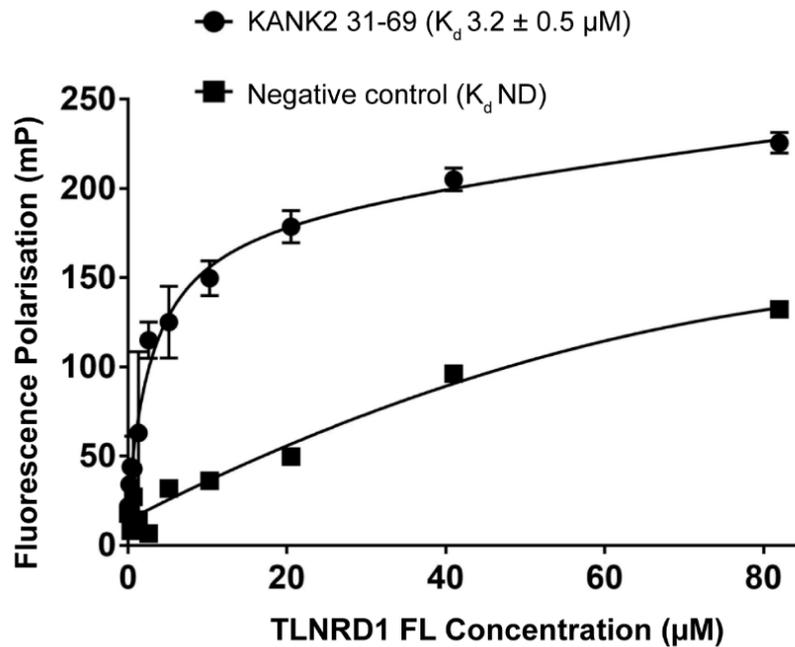


Figure 4.19. TLNRD1 interacts with KANK2 LD motif peptide spanning residues 31-69

Fluorescence polarisation binding assay of TLNRD1-FL against BODIPY labelled KANK2 peptide spanning residues 31-69, and tissue factor pS258 as negative control. Interaction dissociation constant (\pm SE) in μ M is indicated in the graph legend. Measurements were repeated in triplicate. ND= not defined.

4.6. Results section 3: TLNRD1 interacts with newly identified CDK1 LD motif

4.6.1. Structural comparison of talin R8 CDK1 binding site with TLNRD1

As shown previously in Figure 4.3A/B, TLNRD1 helices $\alpha 5$ and $\alpha 6$ share high sequence and structural similarity to talin helices H32 and H33 in the R8 domain. Manual structural modelling of TLNRD1-4H with CDK1 residues 206-223 using the predicted R7R8 binding site as a guide shows some distinct similarities between the two binding sites. In talin, D211 of CDK1 is predicted to form a salt bridge with K1544. In TLNRD1 this residue is replaced by the similar positively charged arginine (R233) positioned side-on to D211 suggesting the two residues may also be able to form a salt bridge interaction (Figure 4.20).

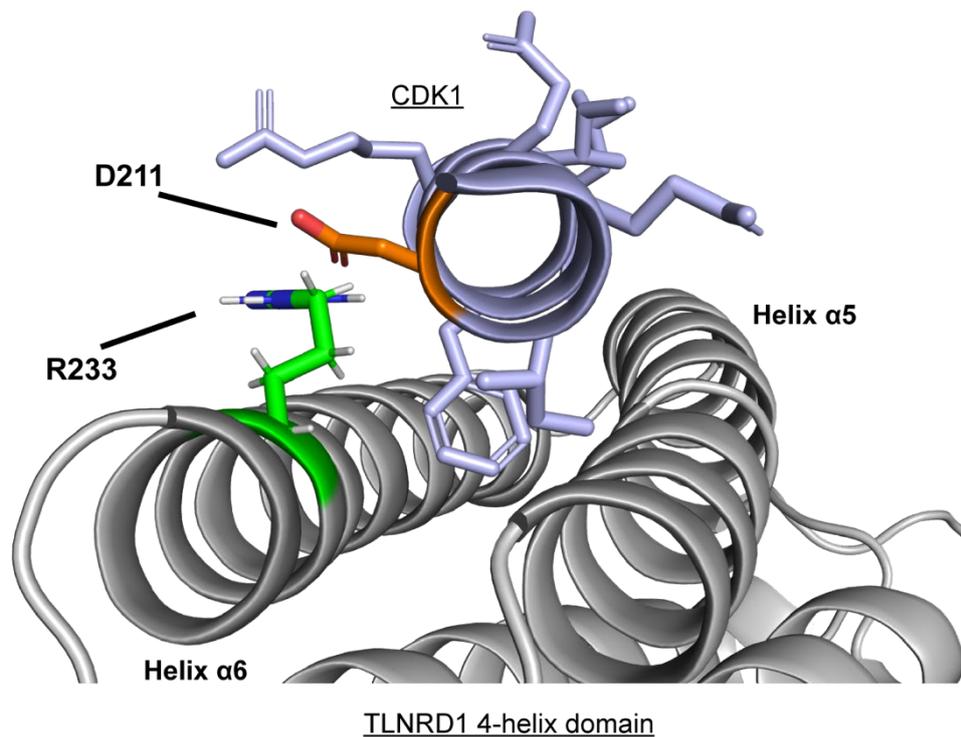


Figure 4.20. CDK1 207-222 modelling with TLNRD1 4-helix domain

Structural representation of CDK1 peptide binding based on known interaction site in talin R8. TLNRD1 R233 is highlighted in green, CDK1 residue D211 highlighted in orange.

4.6.2. TLNRD1 interacts with CDK1 via the 4-helix domain

Using BODIPY-labelled CDK1 206-223 and CDK1_2A peptides (Figure 4.21), fluorescence polarisation assays were used to determine whether TLNRD1 can also interact with the CDK1 LD motif. First, binding experiments with full-length TLNRD1 at 69 μM revealed an interaction with CDK1 206-223 at an affinity of 9.9 μM (± 1.12), with substantially reduced binding to the CDK1_2A mutant peptide with no K_d generated (Figure 4.21B). This reflects the binding profile of talin R7R8 which interacts with 206-223 but is unable to interact with the CDK1_2A peptide. Intriguingly, TLNRD1 was also found to interact with CDK1 with a higher affinity than that previously observed with talin.

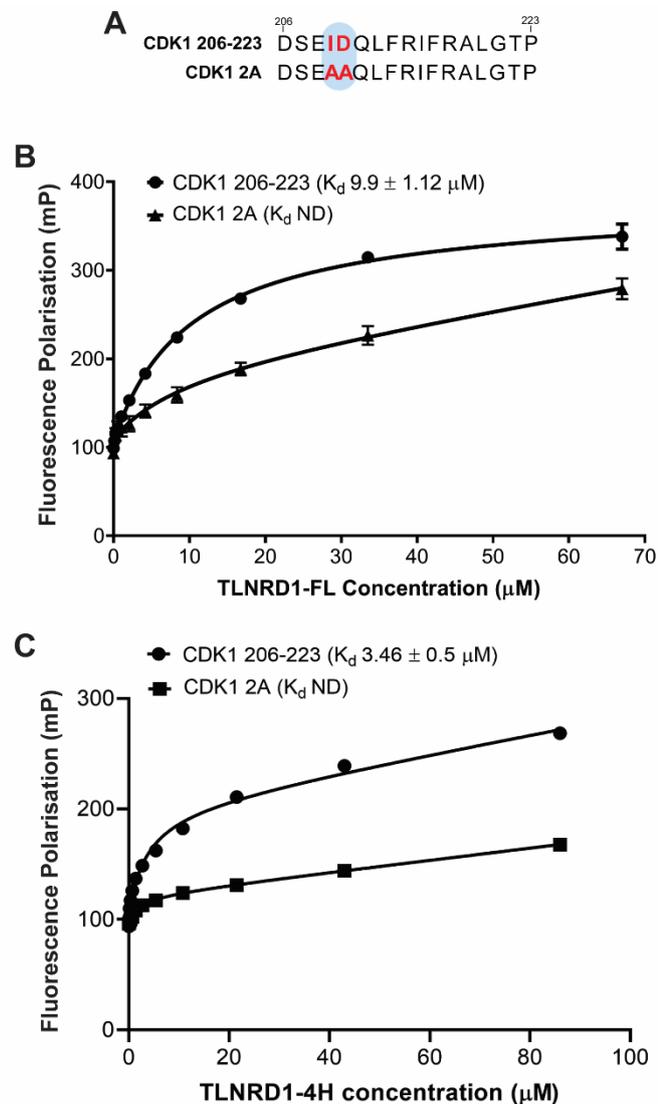


Figure 4.21. TLNRD1 interaction with CDK1 LD motif peptide spanning residues 206-223

(A) CDK1 LD motif wild-type and 2A mutant sequence spanning residues 206-223. Mutated residues are highlighted. (B) FP assay with full-length TLNRD1 and BODIPY labelled CDK1 206-223 and 2A mutant. (C) FP assay with TLNRD1 4-helix domain and CDK1 peptides. Experiments were performed in triplicate. ND= Not defined.

The same FP assay was performed using TLNRD1 4-helix domain to determine if the CDK1 binding site resides in the same domain as talin (Figure 4.21C). TLNRD1-4H with a starting concentration of 83 μM interacts with CDK1 206-223 with an even higher K_d of 3.46 μM ($\pm 0.5 \mu\text{M}$), and as with TLNRD1-FL, TLNRD1-4H was unable to interact with the CDK1_2A mutant peptide which suggests that TLNRD1 has indeed retained the CDK1 binding site through evolution from talin. The higher affinity interaction in the 4-helix domain compared to the full-length protein also suggests that the absence of the 5-helix module increases availability of the binding site.

4.6.3. NMR analysis of TLNRD1-CDK1 interaction

To confirm the results observed in the fluorescence polarisation assays, 2D NMR HSQC titrations were once again used to visualise amino acid interactions with the target peptide. ^{15}N labelled TLNRD1-4H at 75 μM was titrated against CDK1 206-223 at an initial 1:1 ratio and again at a 4:1 ratio. Multiple peak shifts could be observed on the overlaid spectra and could be easily tracked with an increasing concentration of ligand (Figure 4.22). The majority of peak shifts observed are in fast exchange which indicates a weak interaction between TLNRD1-4H and CDK1.

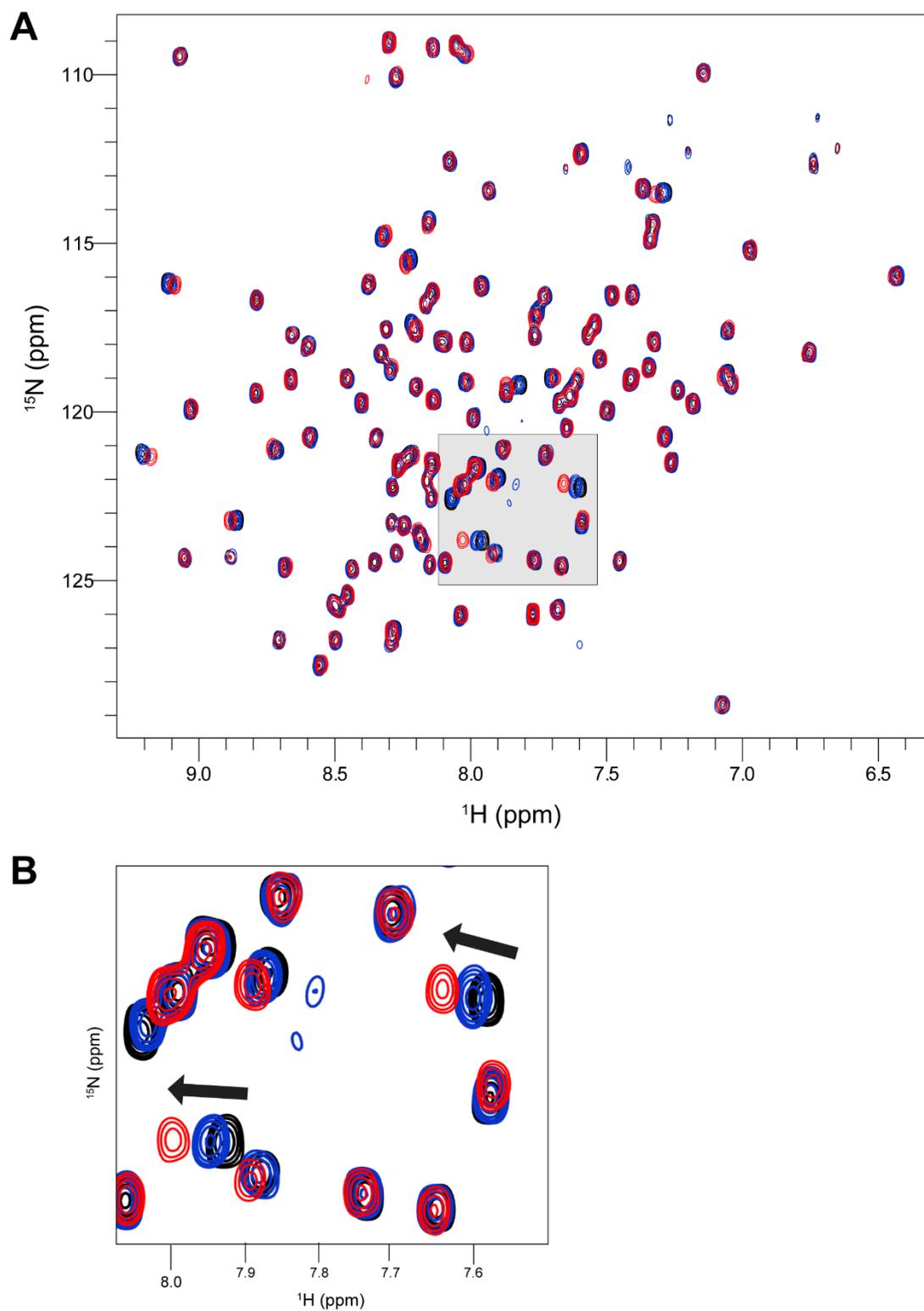


Figure 4.22. ^1H ^{15}N HSQC overlaid spectra of CDK1 206-223 titration against TLNRD1

(A) ^1H ^{15}N HSQC of 75 μM TLNRD1 4-helix domain alone (black) overlaid with spectra from an increasing titration of CDK1 206-223 at 1:1 (blue) and 4:1 (red) ratios. (B) Enlarged view of the spectra showing peak shifts which are indicative of fast exchange. Arrows show the direction of peak shift movement.

4.7. Results section 4: TLNRD1 is an actin bundling protein

The talin molecule contains a number of actin binding sites which provide coupling of the integrin machinery with the actomyosin cytoskeleton. The actin binding properties of talin are essential for its role in relaying mechanical forces into signals for the cell to sense and engage with its environment. The second actin binding site in talin, ABS2, requires the R7R8 domain module for actin filament binding. Previous research has already identified that like R7R8, TLNRD1 can engage actin filaments with much higher apparent affinity *in vitro* and has been shown to localise to actin stress fibres in NIH3T3 cells (Gingras *et al.* 2010). In light of the newly obtained structural information on TLNRD1 and identification of its antiparallel dimer configuration, the actin binding properties on TLNRD1 were investigated further. This section demonstrates TLNRD1 actin binding capability and reveals that unlike talin R7R8, TLNRD1 is an actin bundling protein which can engage more than one actin filament per dimer.

4.7.1. Structural comparison of R7R8 actin binding region and TLNRD1

Investigations into the talin R7R8 domain alone have shown that this region can interact weakly with F-actin, with both 4 and 5-helix domains required for engagement (Gingras *et al.* 2010). The surface of R8 contains an area of positively charged residues, giving the protein a high isoelectric point with residues R1500, R1510 and R1522 promoting electrostatic interaction with the negatively charged filamentous actin (Atherton *et al.* 2015) (Figure 4.23A,B). Interestingly, direct comparison of the surface electrostatic potential between R7R8 and TLNRD1 reveals that the R8 actin binding region has not been retained in the same equivalent position on TLNRD1 (Figure 4.23C). However, what is clear is that TLNRD1 has a strong positively charged region at the opposite end of the 4-helix module suggesting that not only is the binding region shifted, but it may have been modified to provide a higher affinity interaction with actin filaments.

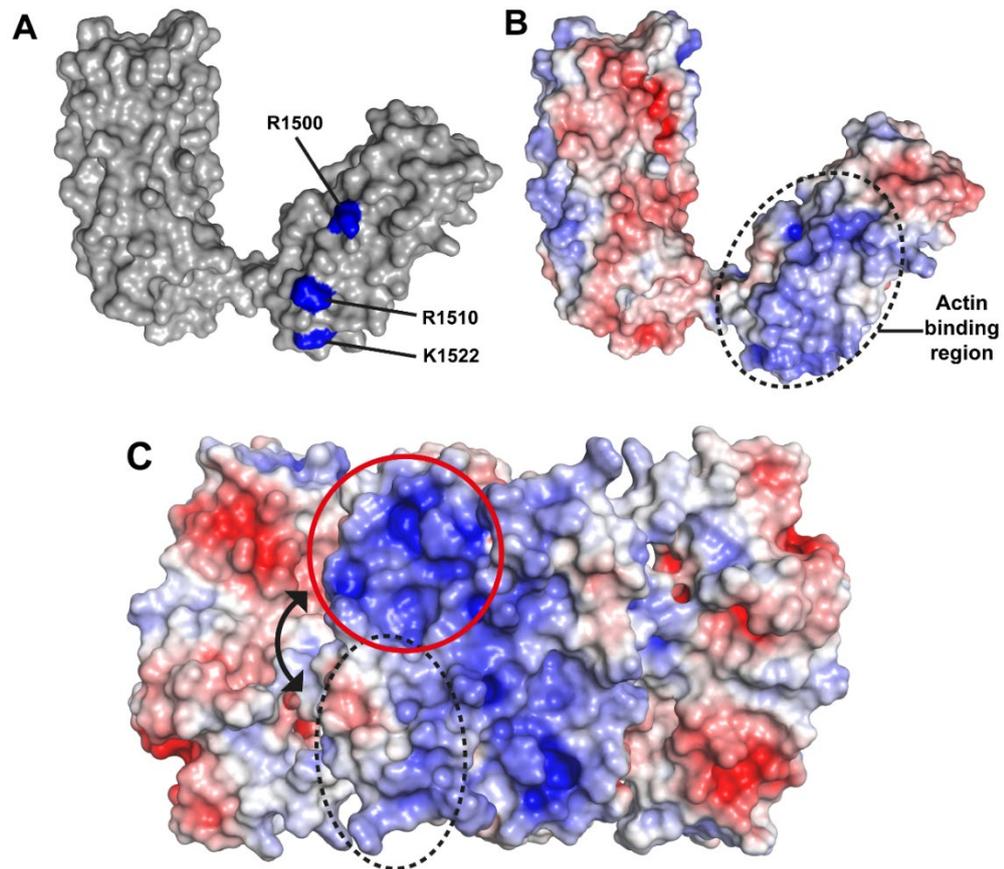


Figure 4.23. R8 actin binding site comparison with TLNRD1

(A) Talin R7R8 (PDB: 2X0C) showing previously characterised residues which mediate actin filament binding. **(B)** Surface electrostatics of (A) highlighting the positively charged actin binding region **(C)** TLNRD1-FL surface electrostatics showing potential change in location of the actin binding region.

4.7.2. Both TLNRD1 domains can interact with F-actin

The TLNRD1 overall structure has a relatively high pI at 8.54 which means the protein remains positively charged at physiological pH, a property indicative of an actin binding protein. Comparison of each 4 and 5-helix domain individually shows a slightly different pI with the 5-helix domain having a lower isoelectric charge at 6.51 suggesting that this domain is neutral at physiological pH, and the 4-helix domain having a much higher pI at 9.3 which suggests that this region contains the actin binding site of TLNRD1. To explore this further, actin co-sedimentation pull-down experiments were used to test the TLNRD1 interaction with filamentous actin. As these experiments require low concentrations of protein, the TLNRD1-5H domain was also tested. Actin binding experiments are performed using a high-speed spin which pellets filamentous actin, while proteins which are interacting with the F-actin will also go into the pellet, and if no binding occurs then the protein will remain in the supernatant. In support of previous observations, TLNRD1-

FL co-sediments well with F-actin with ~55% total protein going to the pellet indicating a direct interaction (Figure 4.24A). The TLNRD1 4-helix module alone can also interact with actin with ~57% going into the pellet, supporting the theory that TLNRD1 has retained the actin binding site within the same domain as talin R8. Furthermore, the 5-helix domain also appeared to show an interaction with F-actin, but this was weaker with just 29.6% of total protein going into the pellet. Talin ABS2 requires both talin R7 and R8 to interact with F-actin, whereas here it is clear that each domain on their own is sufficient for actin binding, with the 4-helix module containing the most dominant actin binding site.

The proposed actin binding region sits in opposing positions in the TLNRD1 dimer, with the dimeric structure offering the intriguing possibility of having more than one actin binding site in the TLNRD1 oligomer. To explore whether the F-actin interaction is dependent on the dimerisation of TLNRD1 the F250D mutant was also tested. F250 sits in the dimerisation interface of TLNRD1 and docks into a pocket on the opposing monomer to create a stable dimer. Mutating this residue to aspartic acid almost completely abolishes TLNRD1 dimerisation with only ~10% remaining as a dimer in solution. TLNRD1-F250D was tested alongside TLNRD1-FL in the high-speed spin assay and revealed that the actin binding property of TLNRD1 is largely retained in the monomeric protein, suggesting that as a dimer, TLNRD1 can engage more than one actin filament at a time (Figure 4.24B).

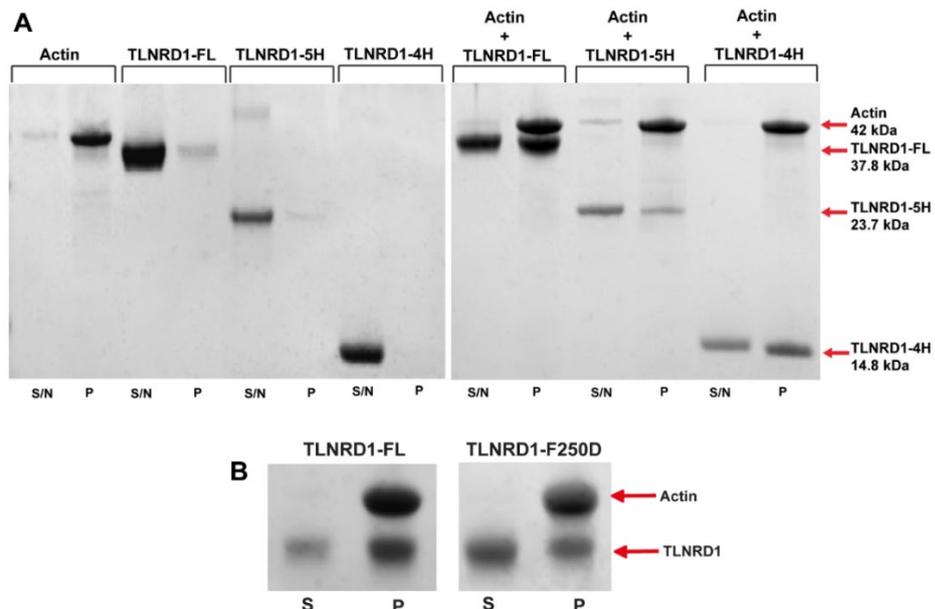


Figure 4.24. TLNRD1 actin co-sedimentation assay

Actin co-sedimentation assay with TLNRD1 constructs. **(A)** Left hand gel shows each protein on its own, with actin pelleting and the TLNRD1 constructs remaining in the supernatant. Right hand gel shows the TLNRD1 constructs incubated with F-actin and the transition to the pellet. **(B)** TLNRD1-F250D retains actin binding.

4.7.3. TLNRD1 2E mutant stops actin filament engagement

The basic region identified in the TLNRD1 4-helix domain resides in the same region which mediates interaction with the LD motif of RIAM and CDK1. Previously, it has been shown that mutating residues K192 and R233 on TLNRD1 to negatively charged glutamic acid drastically altered the surface electrostatics in this region, making it more neutral. It could therefore be predicted that if this region is important for mediating binding to actin filaments, we would see a difference in the binding from the co-sedimentation assay. TLNRD1 wild-type 4-helix domain and 4-helix 2E mutant was tested at 10 μ M against 10 μ M of F-actin. Consistently, 100% of the TLNRD1-4H wild-type construct bound to actin whereas with the 2E mutant actin binding was almost completely abolished with only limited pelleting observed, a result which was reproducible across multiple repeats. This loss of pelleted actin suggests that mutating residues K192 and R233 in the 4-helix domain disrupts its ability to interact with the filamentous actin, and that that either K192 and R233 are required for engagement with filamentous actin or the change in surface electrostatics weakens TLNRD1 affinity (Figure 4.25).

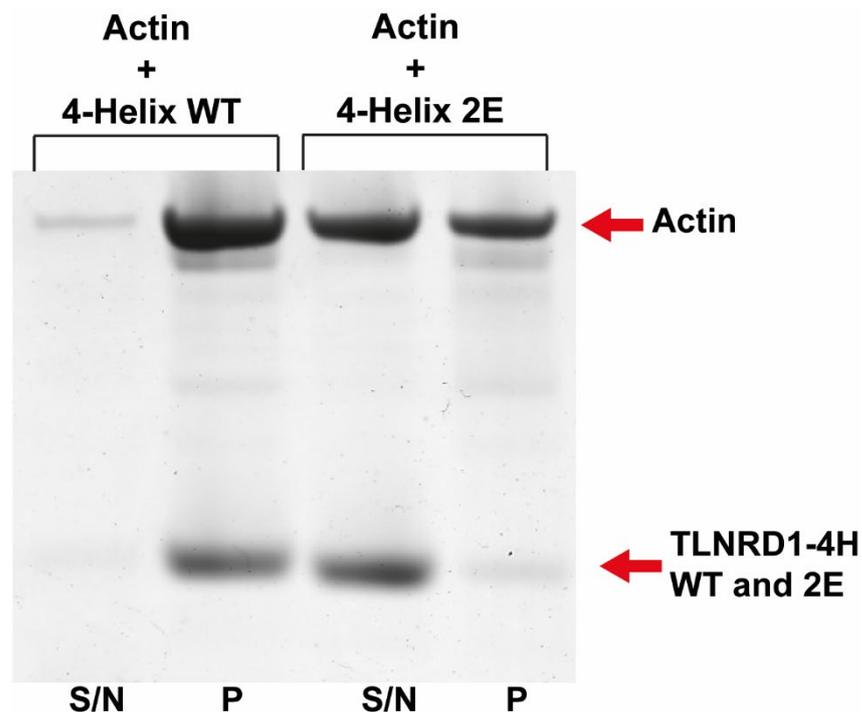


Figure 4.25. TLNRD1 4-helix wild-type vs 2E mutant

Mutating residues K192 and R233 to oppositely charged glutamic acid stops TLNRD1-4H interaction with actin filaments with no protein pelleting with actin at high speed.

4.7.4. TLNRD1 as an actin bundling protein

The structure of full-length TLNRD1 offers exciting insights into the similarities and differences in engagement of the same ligands as talin R7R8. In light of the newly identified antiparallel dimeric structure, we sought to explore whether TLNRD1 can bundle actin filaments via at least one binding site per monomer. Talin R7R8 is monomeric in solution and contains a single actin binding site so doesn't bundle actin filaments, whereas TLNRD1 is almost always either a dimer or tetramer making it potentially possible to engage more than one actin filament at a time.

To explore whether TLNRD1 is able to bundle actin filaments actin bundling assays were initially used, which in contrast to F-actin binding cosedimentation assays, use a low-speed spin which will pellet only actin filaments which are being bundled by the protein of interest. This revealed that both TLNRD1 full-length protein and the 4-helix domain were able to bundle actin filaments. The percentage of actin pelleting was calculated by subtracting measured band intensities for the pellet alone from the total intensity of actin measured from both supernatant and pellet for each concentration condition. TLNRD1-FL showed more efficient bundling with a high proportion of actin going into the pellet at all concentrations compared to the TLNRD1 free experiment which had a very low proportion of actin in the pellet. Actin percentage in the pellet for the 4H began to decrease after 0.937 μM (Figure 4.26A). Finally, to determine whether the bundling activity of TLNRD1 was dependent on dimerisation, the F250D mutant was also tested. This revealed a drastic reduction in pelleted actin showing that monomeric TLNRD1 is no longer able to bundle actin filaments (Figure 4.26C). This supports the theory that TLNRD1 has one binding site per monomer which enables it to bundle filaments as a dimer.

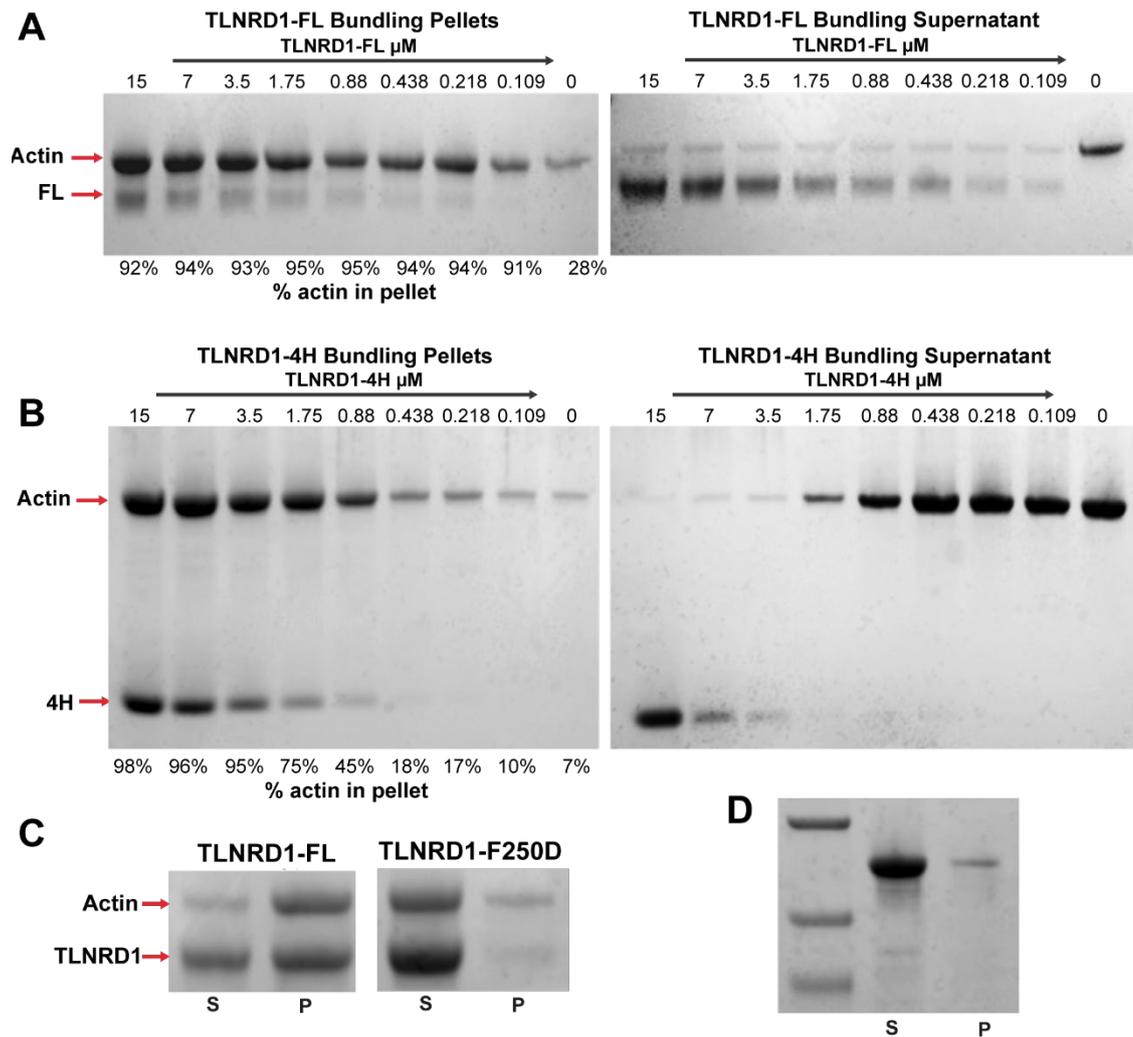


Figure 4.26. TLNRD1 bundles actin filaments in a dimerisation dependent manner

(A) Serial dilution of TLNRD1-FL starting at 15 μM and incubated with 25 μM F-actin with pellet and supernatant fractions shown. TLNRD1-FL appears as a high affinity actin bundling protein. **(B)** Serial dilution of TLNRD1-4H from 15 μM and incubated with 25 μM F-actin. TLNRD1 can bundle filaments without the 5-helix domain. **(C)** Comparison of TLNRD1-FL wild-type vs F250D mutant. TLNRD1 actin bundling is dependent on protein dimerisation. **(D)** Actin bundling fractions of F-actin alone.

4.7.5. Visual inspection of TLNRD1 bundling activity using electron microscopy

The spacing in between filaments is regulated by different sized actin bundling proteins with some creating larger distances between filaments such as α -actinin at 35 nm, and some creating tight bundles of filaments such as fascin at 6 nm ((Winkelman *et al.* 2016). This regulation determines the behaviour and properties of the bundled filaments, with the tight filament bundles formed by fascin being important for filopodia stability and stiffness (Vignjevic *et al.* 2006).

It is now apparent that TLNRD1 is an effective actin bundling protein. To explore the structure and spacing of bundled filaments, preliminary negative stain electron microscopy was used to visualise the actin filaments with and without TLNRD1 full-length protein, TLNRD1 4-helix domain and TLNRD1 F250D mutant. With full-length TLNRD1, we saw drastic changes in filament structure with the formation of large patches of bundled filaments with very tight packing which appeared to be slightly tighter than previously described bundles formed with fascin (Jansen *et al.* 2011) (Figure 4.27B). Similarly, addition of TLNRD1 4-helix domain alone was sufficient for creating bundles of actin filaments with the same tight spacing between filaments, however, the overall size of the bundles appeared smaller than those observed with the full-length protein (Figure 4.27C). Finally, the F250D dimerisation domain mutant showed some drastic changes in filament structure with the appearance of disorganised loose clusters of filaments with large irregular spacing in between each filament. Actin co-sedimentation assays with the F250D mutant shows that this protein retains the ability to interact with F-actin but no longer bundles. The images here suggest that without the antiparallel dimeric configuration, TLNRD1 binding creates distortions in the filament structure which leads to filament entanglement. This shows just how important the highly conserved dimerisation of TLNRD1 is to its bundling properties and thus potential cellular function. Further imaging is required to quantify filament distances, calculate overall size of bundles and provide comparison to other known bundlers but these first snapshots suggest that TLNRD1 is an actin bundling protein which promotes tight packing of filaments.

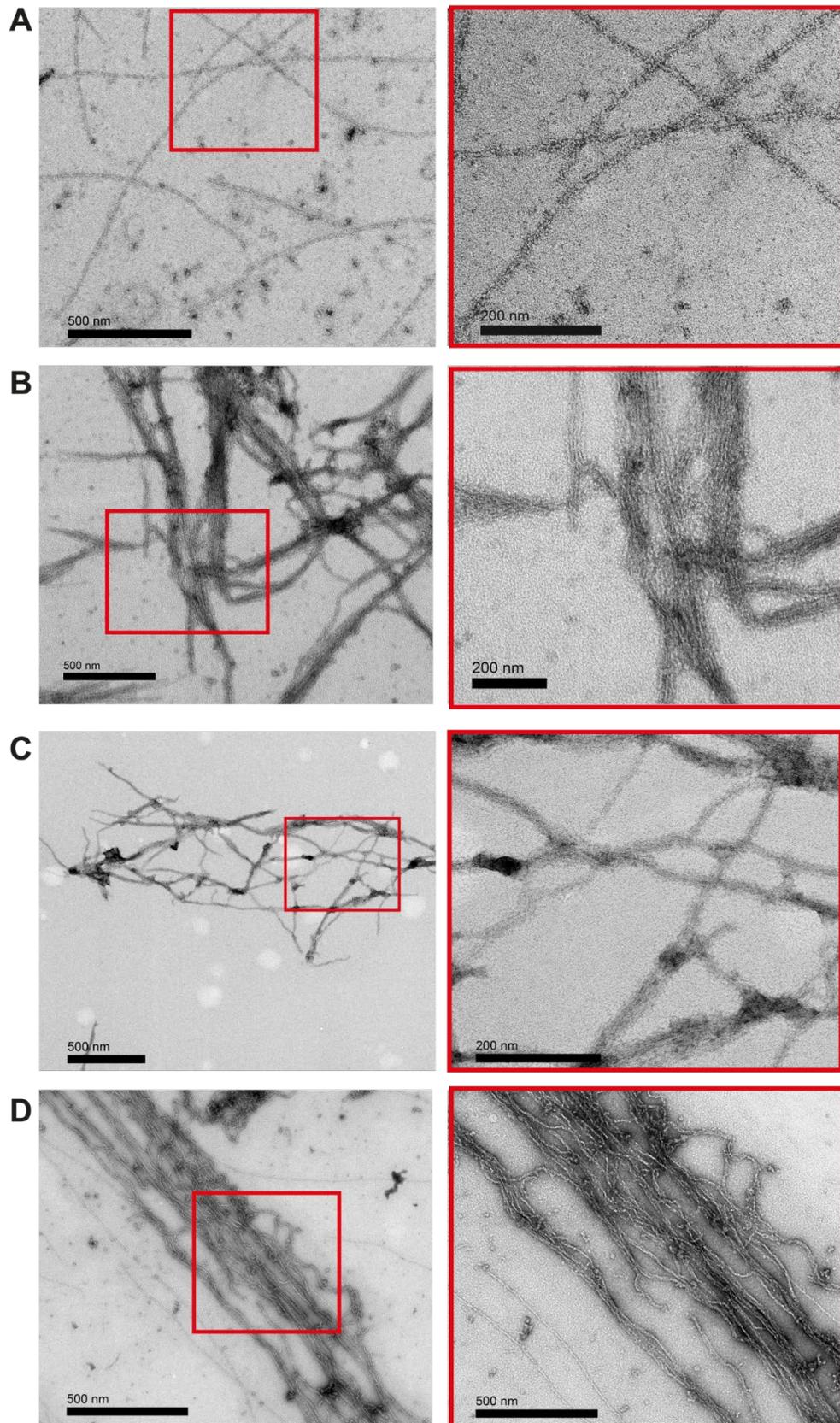


Figure 4.27. Transmission electron microscopy images of TLNRD1 actin bundling
(A) Actin filaments with no addition of protein. **(B)** TLNRD1 mediated actin bundle formation. **(C)** TLNRD1-4H actin bundling. **(D)** TLNRD1-F250D dimerisation mutant leads to disordered filament structures. Zoomed in image locations indicated in red.

4.8. Discussion

In this chapter, up to 5 new and novel interactions have been identified and characterised for the TLNRD1 protein, many of which are known to regulate different aspects of cellular behaviour. RIAM, Lpd, KANKs and CDK1 all share a similar LD motif amino acid sequences which both talin and TLNRD1 can recognise and directly interact with. Interestingly, many of these interactions occur within the same binding site on the TLNRD1-4H domain which is the equivalent binding site observed on talin R8. As an expansion to this work it would be good to explore the dynamics of these interactions, especially considering that they have variable affinities, suggesting that there may be competition or a hierarchy for binding to a particular LD recognition site. Binding experiments should also be repeated with full-length proteins rather than just peptides and supplemented with pull-down assays from cell culture to understand how they interact. Understanding how binding is regulated both biochemically and in cells will be important to deciphering the importance of these interactions in TLNRD1 function.

The results presented offer an intriguing insight into TLNRD1 behaviour and enhances our understanding of TLNRD1 functionality, showing that it has retained the ability to interact with numerous talin R7R8 ligands through LD motif recognition, behaviour which is commonly associated with adhesion related protein interactions. TLNRD1 itself, however, has not been identified as an adhesion localising protein and has lost the vinculin binding sites of R7R8, the binding of which to talin is important for anchoring talin to the actin cytoskeleton and promoting adhesion maturation in a force dependent manner (Atherton *et al.* 2016). This suggests that TLNRD1 may be acting as a regulator of different adhesion related proteins outside the mechanosensitive adhesion signalling pathways in the cell, perhaps acting as a mobile equivalent of the talin R7R8 signalling hub.

One of the most surprising and unexpected findings presented in this chapter is the identification of TLNRD1 as an actin bundling protein. It has been clear from previous research that TLNRD1 can directly interact with actin filaments *in vitro* and has been shown to localise to actin stress fibres in NIH3T3 cells (Gingras *et al.* 2010). Talin R7R8 interaction with actin filaments in the absence of the R4 domain is relatively weak, only exhibiting single filament interactions, with no binding observable from the individual domains R7 or R8 alone. TLNRD1 on the other hand, has adapted the actin binding site to drastically increase affinity, allowing binding and bundling of actin filaments by the 4-helix domain alone. Comparison with other actin bundling proteins reveals some

similarities and intriguing differences. Dimerisation is a feature observed in multiple actin bundling proteins, including α -actinin, filamin-A, formins and spectrin, however many of these have extended linear structures which allow them to increase spacing between filaments or wrap around actin filaments. TLNRD1 on the other hand appeared to have a more compact structure akin to actin bundling proteins which perform this function via two actin binding sites per monomer, one example being fascin which forms tightly packed bundles similar to those observed with TLNRD1. The question now remains on how exactly TLNRD1 is able to bundle filaments, whether it bundles parallel or antiparallel filaments and how does the structure change when in contact with actin filaments. Cryo-electron microscopy should be implemented in future work to directly observed the structure of TLNRD1 on actin filaments and compare filament organisation between monomeric, dimeric or even tetrameric TLNRD1. It could be speculated that TLNRD1 extends its structure when bundling filaments, allowing it to wrap around filaments while retaining tight packing. Additionally, the compact antiparallel configuration of TLNRD1 suggests that it may in fact bundle antiparallel filaments only. Notable cellular features which contain antiparallel actin filaments include stress fibres and transverse arcs, observing cellular localisation of TLNRD1 could provide clues as to whether there is a preference for these structures. While the discovery of TLNRD1 being an actin bundler is interesting, more work is needed to understand how TLNRD1 is interacting with actin, and how this compares with talin ABS2, particularly as the mode of binding appears drastically different with talin requiring both R7R8 and R4 engagement for actin filament binding. The loss of interaction with the 2E mutant suggests that the actin binding site overlaps with the LD motif recognition site suggesting that when bound to actin TLNRD1 can't engage these other ligands. Future work should focus on exploring whether this is the case as it alludes to the possibility that TLNRD1 can engage in different behaviours depending on cellular context.

The intriguing bundling behaviour observed with TLNRD1 is similar in some aspects to that of filopodia promoting fascin which also forms bundles of tightly packed actin filaments, the main difference being that each TLNRD1 monomer appears to have a single binding site rather than two, and that bundling is mediated by the incredibly tight dimerisation of the TLNRD1 4-helix domain. This exciting new discovery opens up a host of questions about TLNRD1 behaviour in the cell, particularly in the context of the research presented later on in Chapter 5 and leads us to question how interaction with all these different ligands are coordinated in the cell.

Chapter 5. Investigating TLNRD1 function

5.1. Introduction

In the current literature, research investigating the function of TLNRD1 in a cellular context has been extremely limited. Functional investigations using transgene recovery in mice suggested that TLNRD1 is not involved in mesoderm development (Hsieh *et al.* 2003), however, no published study has yet performed a targeted knockout or knockdown of TLNRD1 in an animal model. There is clear evidence that TLNRD1 can promote migration, proliferation and invasion of cancerous cells but the question of why or how increases in TLNRD1 expression causes this effect is yet to be explored.

Using the information gained from the structural and biochemical investigations into TLNRD1, we collaborated with Dr Guillaume Jacquemet, Dr Ilkka Paatero and the lab of Professor Johanna Ivaska (University of Turku, Finland) to explore the functional role of TLNRD1. This group has expertise in using zebrafish model organisms and advanced microscopy to explore protein function in the field of adhesion and migration. Exploration of TLNRD1 physiological function in zebrafish was performed by Dr Ilkka Paatero who is currently head of the zebrafish facility at Turku Bioscience.

Initial analysis of TLNRD1 cellular localisation and filopodia protrusion quantification was performed by Dr Guillaume Jacquemet at Turku Biosciences. Additional experiments including immunoprecipitation experiments for antibody testing, further visualisation of TLNRD1 localisation and migration assays exploring the impact of TLNRD1 expression on cell migration were performed by myself.

All cell biology experiments were performed using the osteosarcoma U2OS cancer cell line, which are highly adherent epithelial cells that form both lamellipodial and filopodial protrusions during migration. U2OS cells have been used extensively in studying talin function and adhesion complex regulation, and qPCR analysis revealed that TLNRD1 mRNA is expressed in this cell line making it a good starting point for exploring TLNRD1 function.

5.2. Results

5.2.1. *TLNRD1 zebrafish morpholino leads to developmental defects*

Zebrafish (*Danio rerio*) are incredibly useful animal models for the study of vertebrate development and disease, with biological systems which resemble those found in humans. Throughout embryogenesis and larval stages, zebrafish exhibit an optical transparency which allows easy visualisation of developing systems such as the cardiovascular system, and short lifespans coupled with large numbers of offspring allow

for easy and quick genetic manipulation for the study of protein function and expression (Dooley and Zon 2000). Inspection of the EBI expression atlas and Human Protein Atlas reveals that TLNRD1 RNA is expressed across all tissues in *Homo Sapiens* with expression highest in adipose tissue, lymphoid tissues and the brain (Uhlén *et al*, 2015; Papatheodorou *et al*, 2020). While it is unclear from the EBI database whether TLNRD1 has a similar expression pattern across tissues in zebrafish, it does show expression across all developmental stages with expression highest during the Pharyngula stages whereby the embryo lengthens and straightens starting with the tail, the pharyngeal arch's begin to form and the cardiovascular system begins to develop with the first heartbeat. To explore whether TLNRD1 plays a role in vertebrate development, morpholino antisense oligonucleotides (MOs) were designed by Dr Ilkka Paatero (Gene Tools LLC) to target TLNRD1 mRNA and knock-down TLNRD1 expression in the developing embryos using the sequenced region of the genome as a reference for design. As a negative control, embryos were injected with an MO which targets a mutated intron of human β -globin (Gene Tools LLC). Embryos were injected 2 days post-fertilisation (dpf).

As expected, the zebrafish injected with a 2.5 ng dose of control MO exhibited no phenotypic defects throughout the larval stages and into adulthood (Figure 5.1A). Injection with the TLNRD1 targeting MO, however, lead to drastic changes in phenotype (Figure 5.1B). The embryos developed observable defects consistent with both brain and pericardiac oedema (Figure 5.1B i and ii). They also exhibited highly abnormal tail morphologies with a loss of caudal fin-fold development, tail-curvature and defects in the tail-bud (Figure 5.1B iii).

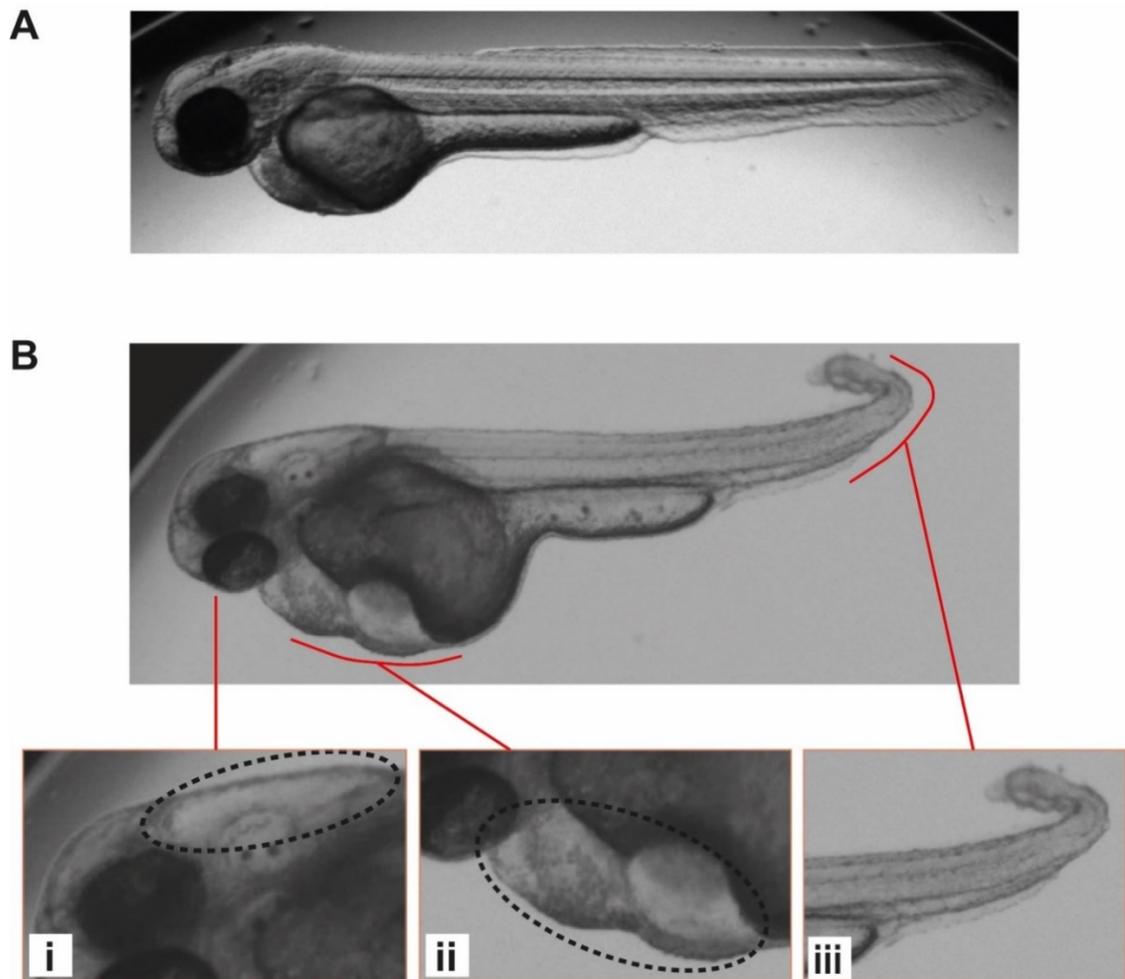


Figure 5.1. TLNRD1 zebrafish morpholino

(A) Brightfield image of zebrafish after injection with 2.5 ng of control MO. Image shows normal wild-type physiology with no developmental defects. **(B)** Brightfield image after injection with 2.5 ng of TLNRD1 morpholino. Zoomed in images highlighting brain oedema (i), pericardiac oedema (ii) and abnormal tail development (iii).

The oedema observed in the TLNRD1 MO zebrafish surrounding the heart and cranial region indicates that loss of TLNRD1 has an impact on development of these systems but whether this is direct or indirect could not be elucidated. To explore whether there were any vascular defects, transgenic *kdrl:mCherry-CAAX* zebrafish were used to directly visualise development of the cardiovascular system. Kdrl is a vascular endothelial growth factor receptor which is essential for blood vessel development, and as such serves as a useful marker for visualising the development of blood vessels in the transparent zebrafish. To further elucidate TLNRD1 impact on development and to discount any potential off-target binding of the MO causing the observed defects, CRISPR/Cas9 knock-outs were created using single guide RNA (sgRNA) specific for the *TLNRD1* gene locus.

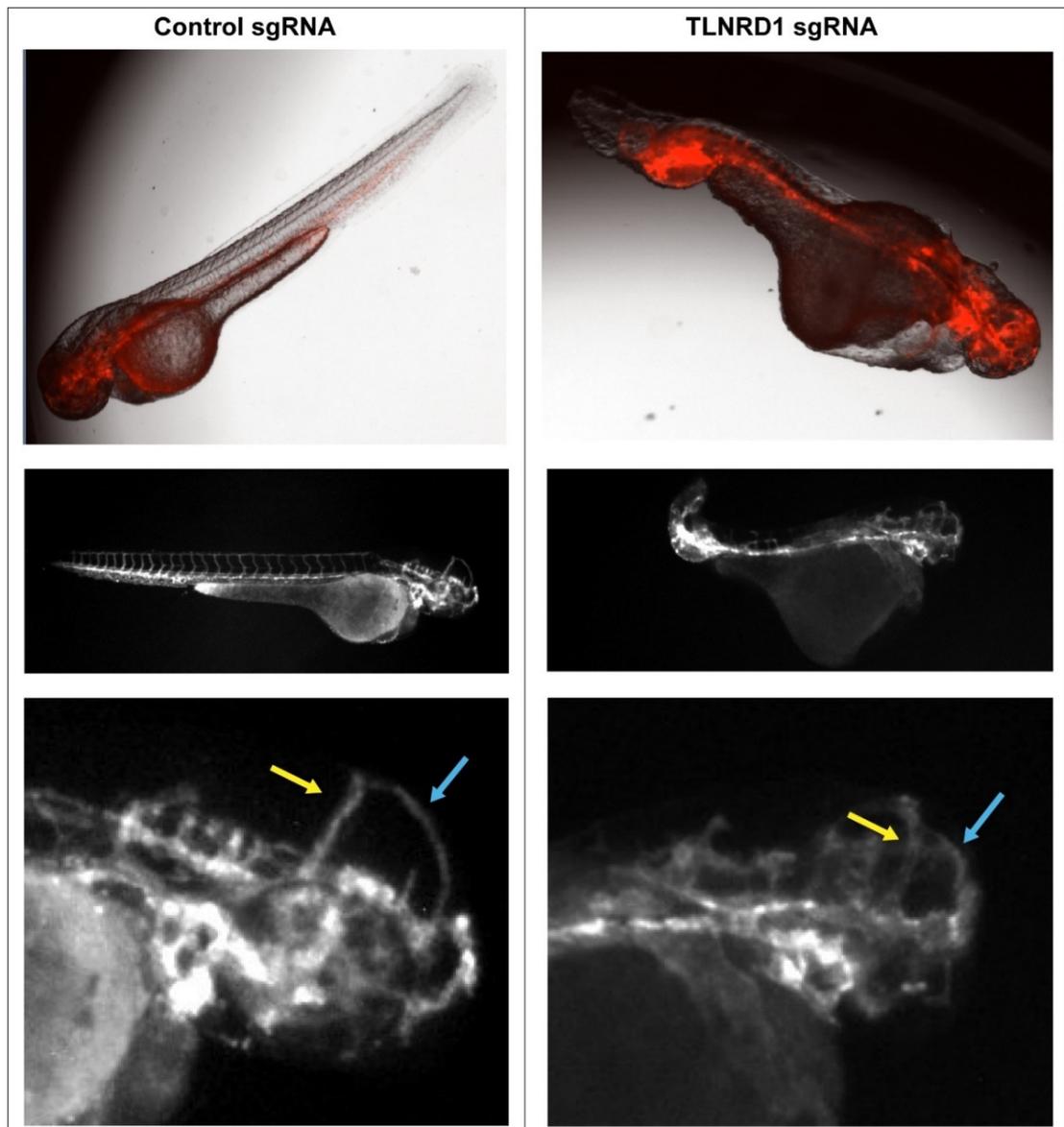


Figure 5.2. CRISPR/Cas9 mediated *tlrnd1* knock-out leads to severe developmental defects
 Comparison of control sgRNA vs TLNRD1 sgRNA in *kdrl:mCherry-CAAX* zebrafish. Knocking out TLNRD1 expression leads to severe vascular defects with large vessel dilations. Middle cerebral vein highlighted by yellow arrow; mesencephalic vein highlighted by blue arrow.

The loss of TLNRD1 resulted in severely affected zebrafish morphology, with large vascular dilation and developmental defects (Figure 5.2). Enlarged vessels could be observed both in the cranial region and the caudal hematopoietic tissue at the tail base. Close-up inspection of the cranial mid-cerebral vein and mesencephalic vein show severe deformity in vascular structure with vessel dilation.

5.2.2. *TLNRD1 localises to the tips of filopodia*

In order to explore the proteins function at a cellular level and understand why expression promotes increased migration, invasion and survival in cancer cells, our collaborator Dr Guillaume Jacquemet transfected U2OS cells with mouse TLNRD1 wild-type and F250D mutant carrying an N-terminal GFP tag to visualise localisation. For tracking potential involvement in filopodial protrusions, cell lines transiently expressing mScarlet/RFP labelled myosin-10 were used to promote filopodia formation. The cells were plated on fibronectin for 2 hours and imaged using spinning disk confocal microscopy. Protein localisation within filopodia was mapped as previously described (Jacquemet *et al.* 2019).

Talin largely localises to adhesion complexes which appear as distinct clusters under imaging. TLNRD1 on the other hand, despite being closely related to talin and interacting with adhesion related ligands RIAM and KANK, showed no localisation to sites of adhesion complex formation. Unexpectedly, the imaging revealed distinctive TLNRD1 localisation at the tips of filopodial protrusions, areas which are heavily enriched with receptors, adhesion related proteins including lamellipodin, integrins and talin, as well as cytoskeletal regulators such as the formin mDia2 (Jacquemet *et al.* 2019). This localisation to filopodial tips appeared to be dependent on TLNRD1 dimerisation, with the TLNRD1-F250D dimerisation knockout mutant completely losing filopodia localisation. What was most surprising from this imaging was the lack of co-localisation with actin filaments. Only a minimal amount of actin binding could be observed for TLNRD1 in the U2OS cells, largely occurring along the larger actin stress fibres which consist of bundles of ~10-30 filaments (Tojkander, Gateva and Lappalainen 2012). This unexpected result suggests that under normal conditions TLNRD1 actin binding is largely inhibited and requires activation by an as yet unknown signal.

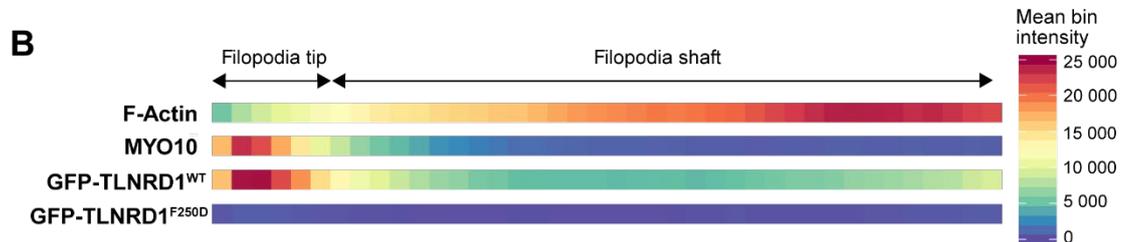
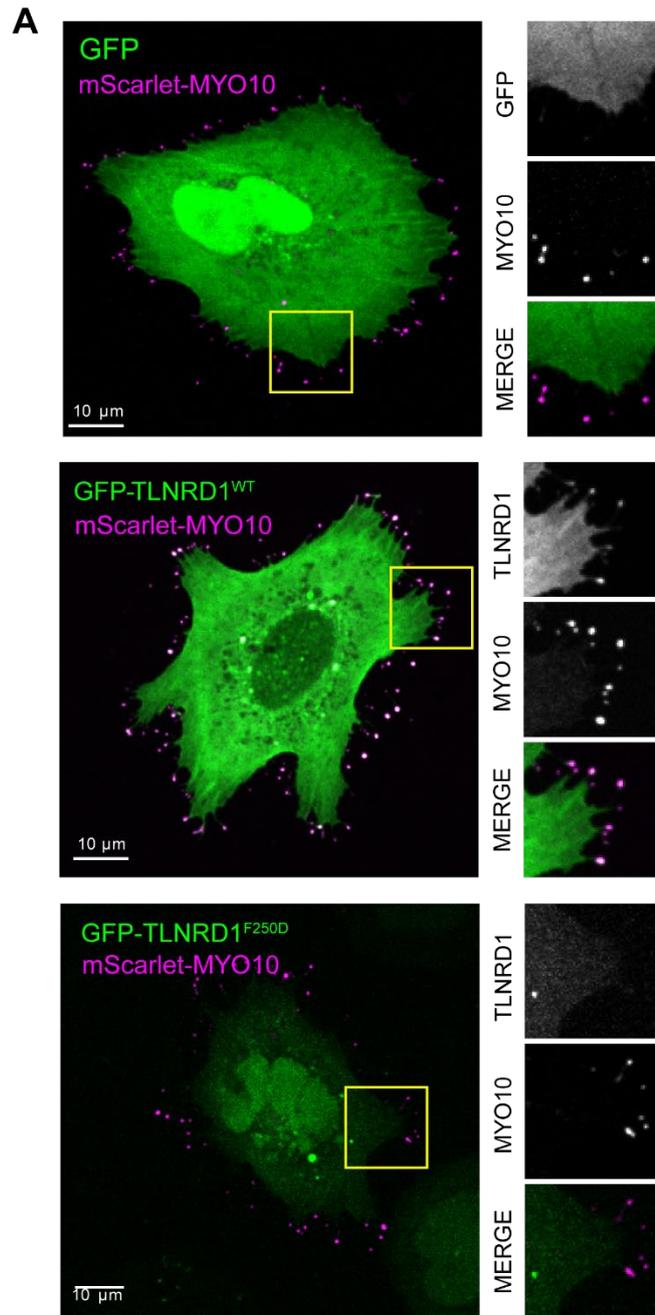


Figure 5.3. Spinning disk confocal microscopy of GFP-TLNRD1 imaged by Guillaume Jacquemet

(A) Spinning disk confocal images of U2OS cells transiently expressing Myosin 10 (MYO10) with GFP, GFP-TLNRD1 or GFP-TLNRD1 F250D. **(B)** Heat map representing TLNRD1 distribution along filopodia based on >360 profiles. Localisation determined as previously described (Jacquemet *et al.* 2019).

5.2.3. TLNRD1 promotes filopodia formation

With the identification of TLNRD1 localisation at the tips of filopodia, our collaborators also sought to explore whether the presence of TLNRD1 was affecting filopodia dynamics by regulating their formation. First, myosin-X positive filopodia numbers were quantified and compared between GFP only expressing cells, cells expressing GFP-TLNRD1 and cells expressing the GFP-TLNRD1 F250D dimerisation mutant (see section 3.3.7). This was followed by comparison of TLNRD1 expression silencing using two independent siRNA's (6 and 7) against cells silenced using AllStars negative control siRNA (Qiagen).

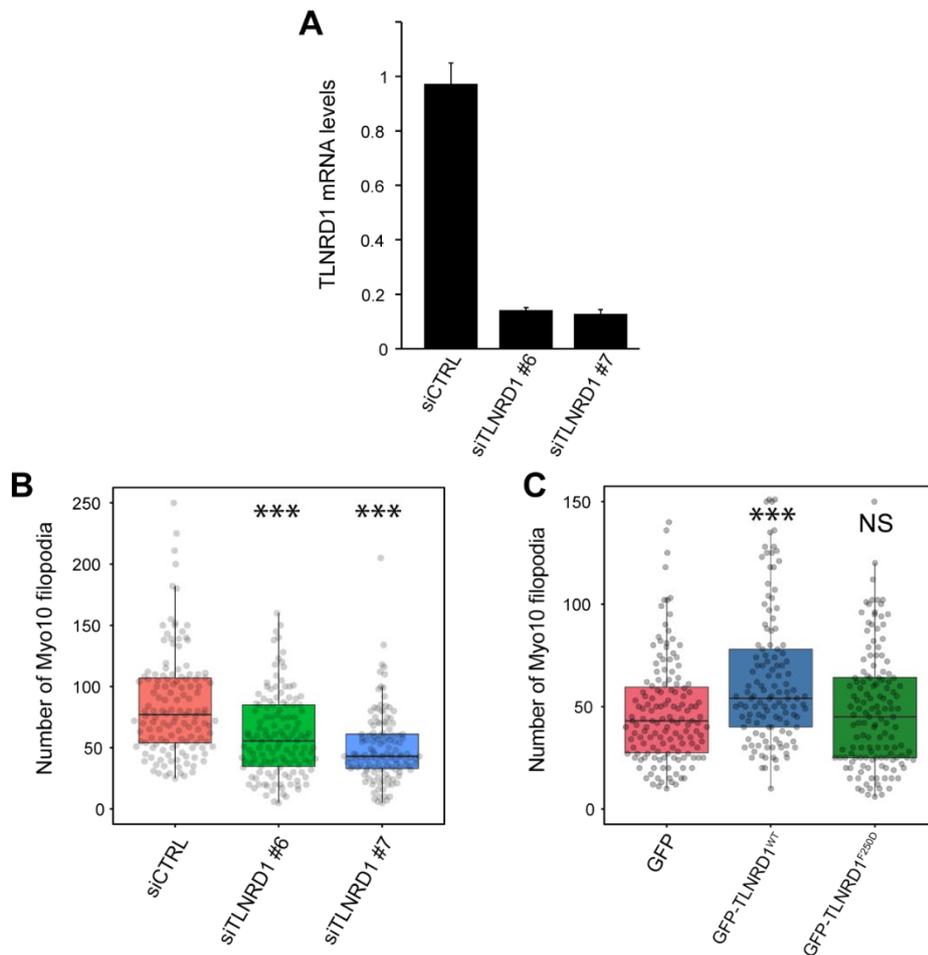


Figure 5.4. TLNRD1 expression and filopodia quantification from Guillaume Jacquemet

(A) TLNRD1 mRNA expression levels determined using qPCR confirming TLNRD1 knock-down. **(B)** Quantification of filopodia in cells treated with AllStar negative control siRNA and comparison with two TLNRD1 targeting siRNA's. **(C)** Quantification of filopodia in GFP (pink), GFP-TLNRD1 (blue) and GFP-TLNRD1 F250D (green) expressing cells. N = 125 cells, three independent repeats. ***P<0.001. NS = Not significant.

Comparison of GFP-TLNRD1 and GFP revealed an overall increase in the numbers of filopodial protrusions formed (Figure 5.4C), suggesting that TLNRD1 is somehow

facilitating their formation or increasing their stability. Interestingly, this effect appears to be dependent on TLNRD1 dimerisation, with the TLNRD1 F250D mutant showing no change in filopodial numbers compared to GFP alone. Furthermore, analysis with siRNA mediated knock-down of TLNRD1 showed a drastic reduction in filopodia formation, with the biggest effect seen with siRNA #7 (Figure 5.4B). Together, these results suggest that TLNRD1 facilitates filopodia elongation or plays a role in stabilising them once formed.

5.2.4. Identifying TLNRD1 endogenous protein

As of yet, no evidence of endogenous TLNRD1 protein could be identified in the U2OS cell line and experiments have so far been dependent on transfection with a GFP tagged TLNRD1 construct. Although TLNRD1 mRNA is consistently present in the cells, the protein itself could not be detected, either due to lack of antibody recognition or mRNA level suppression of TLNRD1 expression. To explore this further, work began to find an antibody which correctly recognises the TLNRD1 protein and works well for immunoprecipitation (IP) experiments. This was done using test IP experiments comparing a TLNRD1 antibody we generated which was raised against purified human full-length TLNRD1 protein in rabbit by CapraScience, and a commercially available rabbit TLNRD1 antibody which targets the TLNRD1 N-terminal region (ThermoFisher Scientific, PA5-70832). For this experiment, lysates were prepared from wild-type U2OS cells, cells expressing GFP only and cells expressing N-terminal tagged GFP full-length mouse TLNRD1. As a control GFP lysate was incubated with beads coated with an isotype matching IgG antibody. Both lysates (unbound) and bead bound samples were compared by western blot with corresponding primary antibody matching the test antibody, and IRDye 800CW conjugated donkey anti-rabbit IgG secondary antibody (LI-COR).

The initial IP test of both antibodies revealed that the human anti-TLNRD1 antibody which was raised by CapraScience correctly recognised the mouse GFP-TLNRD1 construct expressed in U2OS cells, successfully pulling down GFP-TLNRD1 with an expected approximate molecular weight of 65 kDa, as shown by the darker band in the bound fraction (Figure 5.5B). Additionally, there was evidence of an extra band at ~37 kDa which corresponds to the expected molecular weight of the TLNRD1 protein without the GFP-tag, however, this band is most prominent in the GFP-TLNRD1 bound fraction suggesting that some of the overexpressed protein is cleaved at the linker between TLNRD1 and the GFP tag. With the commercially available anti-MESDC1 antibody, despite binding to GFP-TLNRD1 in the unbound fraction, the antibody was unable to pull down the protein for

immunoprecipitation experiments, with the ~65 kDa band missing from the bound fraction (Figure 5.5A). Furthermore, there appeared to be a large amount of non-specific binding to a band with a molecular weight lower than expected for endogenous TLNRD1. For this reason, the commercial PA5-70832 antibody was not used for further experiments.

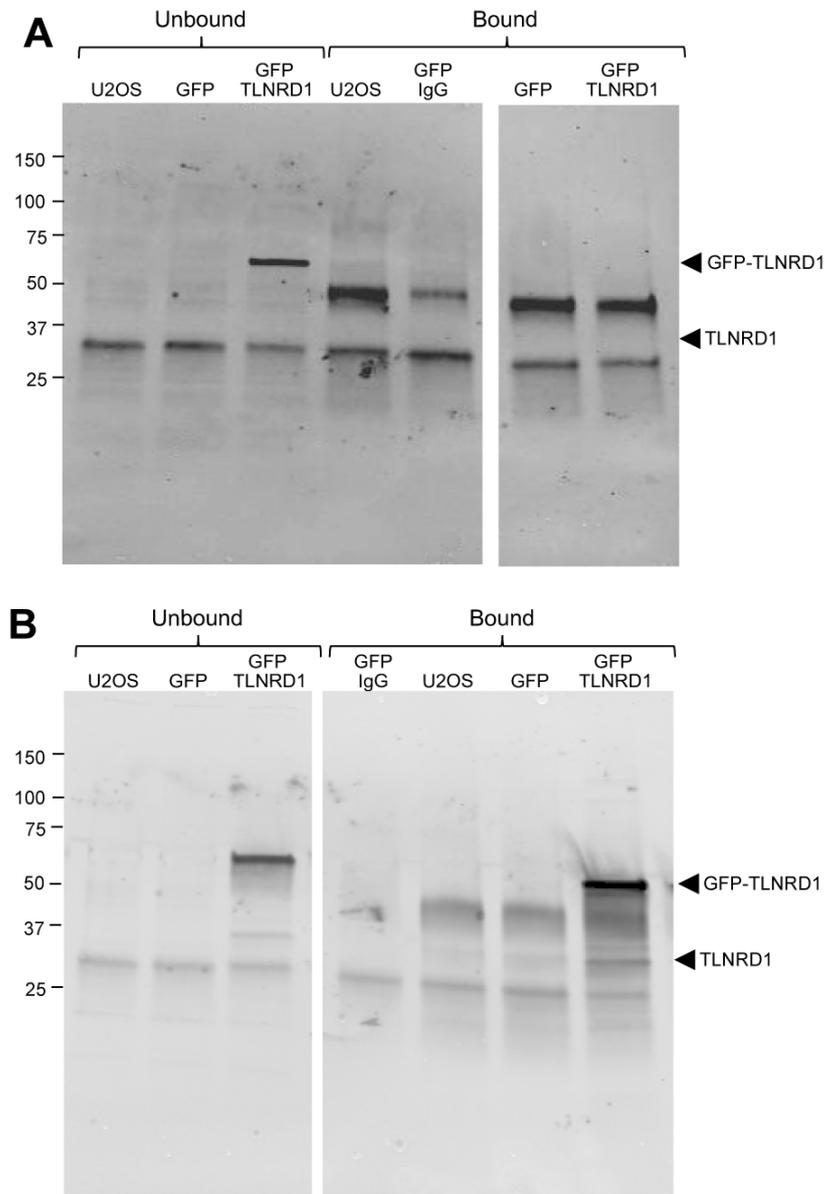


Figure 5.5. Immunoprecipitation test of TLNRD1 antibodies

(A) Western blot of IP with commercially available TLNRD1 antibody showing bound and unbound fractions **(B)** IP of newly raised TLNRD1 antibody showing bound and unbound fractions.

With evidence that the CapraScience raised antibody correctly recognises TLNRD1 protein, the IP experiments were scaled up to include a larger volume of cells in an attempt to find endogenous expression of TLNRD1. This was combined with 3

independent siRNA's (6, 7 and 8) which were designed to knock-down TLNRD1 endogenous expression. The siRNA knock-down will provide confirmation of any potential endogenous protein identification. AllStars negative control siRNA (Qiagen) was also used alongside TLNRD1 targeting siRNAs.

IP with wild-type U2OS cells, siRNA Allstar siRNA control cells and wild-type cells incubated with IgG isotype control antibody revealed a single well-defined band at ~37 kDa which corresponds with the expected molecular weight of TLNRD1. This band is not visible in any of the 3 TLNRD1 targeting siRNAs, suggesting that the band observed at 37 kDa is TLNRD1 and that the siRNAs efficiently knocked-down expression.

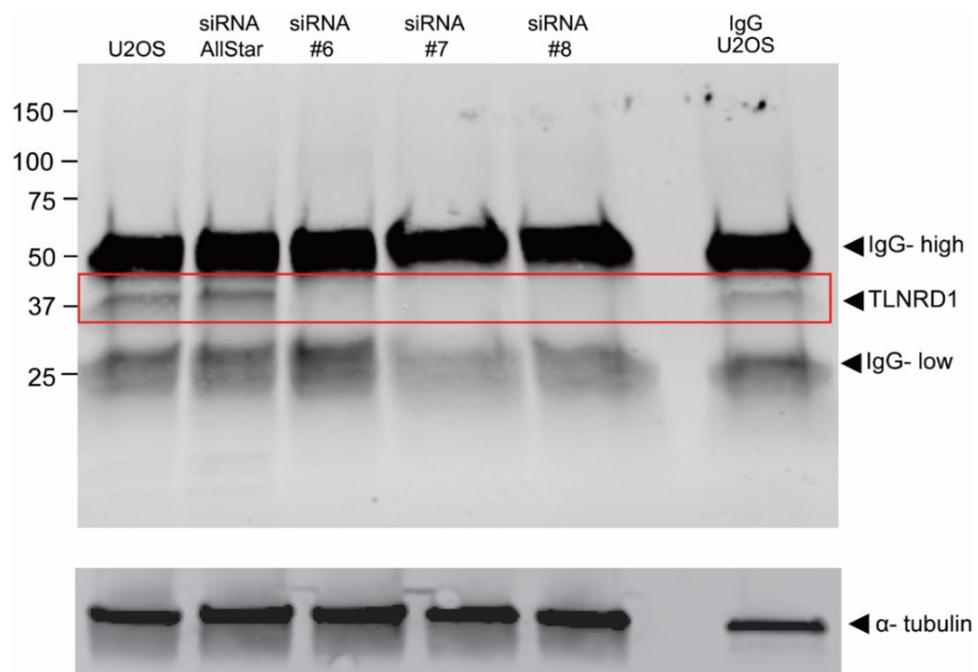


Figure 5.6. Identification of endogenous TLNRD1 expression

Western blot of IP experiment with CapraScience raised antibody, comparing bound fractions from wild-type, siRNA control and TLNRD1 siRNA's 6, 7 and 8. α -tubulin was used as a loading control. Region showing comparative levels of endogenous TLNRD1 protein is highlighted in red. N=2.

5.2.5. TLNRD1 4-helix domain shows increased F-actin localisation

One of the more surprising aspects of TLNRD1 localisation observed in the initial imaging was the limited amount of actin filament colocalization. TLNRD1 only appeared to localise to a few of the thicker actin stress fibres with no apparent preference in the type of stress fibre that TLNRD1 interacts with. This limited amount of actin binding observed for TLNRD1 suggests that the TLNRD1 actin binding properties are tightly regulated in a cellular context either through an autoinhibitory process or through post-translational

modification. To further assess the limited actin filament binding imaging was repeated using an Airyscan confocal microscope to obtain a video of U2OS cells expressing GFP-TLNRD1 with actin cytoskeleton labelling via SiR-actin. As before, a limited amount of clear filament binding was observed, however, there was some localisation with structures consistent with both dorsal stress fibres (Figure 5.7). These types of stress fibres are thick mechanosensing actin cables formed at the base of lamellipodial protrusions when cells migrate through 2D environments. These stress fibres grow outwards towards the cell periphery in a unipolar manner, whereas the proximal fibres situated closer to the centre of the cell have a mixture of polarities.

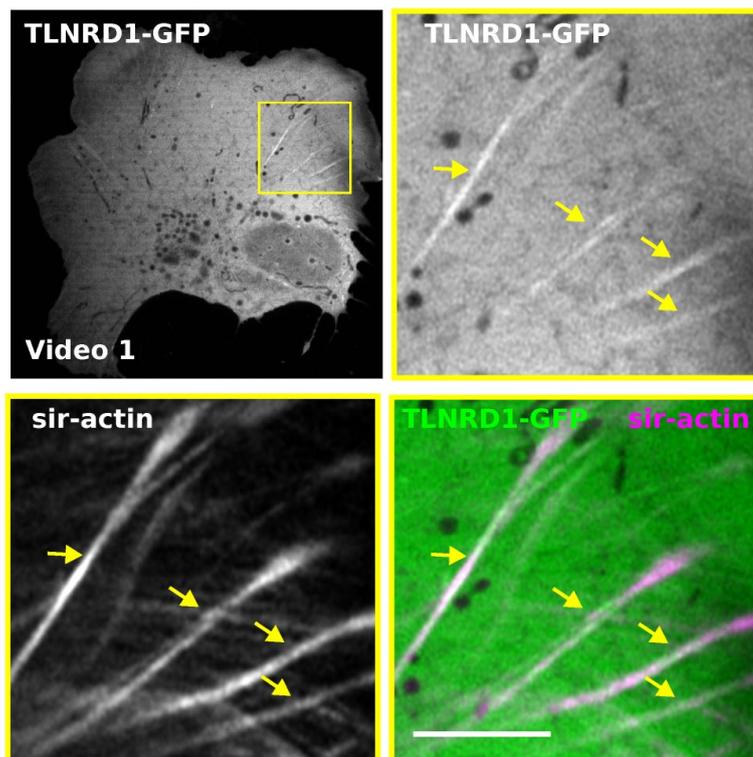


Figure 5.7. GFP-TLNRD1 actin stress fibre localisation as imaged by Guillaume Jacquemet

Snapshot taken from live imaging of GFP-TLNRD1 expressing U2OS cells plated in. Imaging was performed using an Airyscan confocal microscope. Cells were incubated with SiR-actin for 2 hours for labelling of the actin cytoskeleton.

In vitro actin cosedimentation assays revealed that the 4-helix domain alone can strongly interact with F-actin and induce tight bundling of actin filaments. To explore whether this would also be the case in a cellular context, the TLNRD1 4-helix domain (human) was cloned into a GFP vector for visualisation with confocal spinning disk microscopy. Interestingly, the 4-helix domain alone appeared to show an increase in colocalisation with actin filaments compared to full-length TLNRD1. Colocalisation was observed

predominantly along ventral stress fibres which extend across the cell and are connected at each end to focal adhesion complexes. Additionally, both N-terminal and C-terminal GFP tagged TLNRD1-4H was analysed with no differences in localisation between the two showing that the positioning of the tag did not interfere with filament binding. The increased colocalisation with actin filaments suggests that filament binding is somehow regulated by the 5-helix domain or the N-terminal unstructured region of TLNRD1.

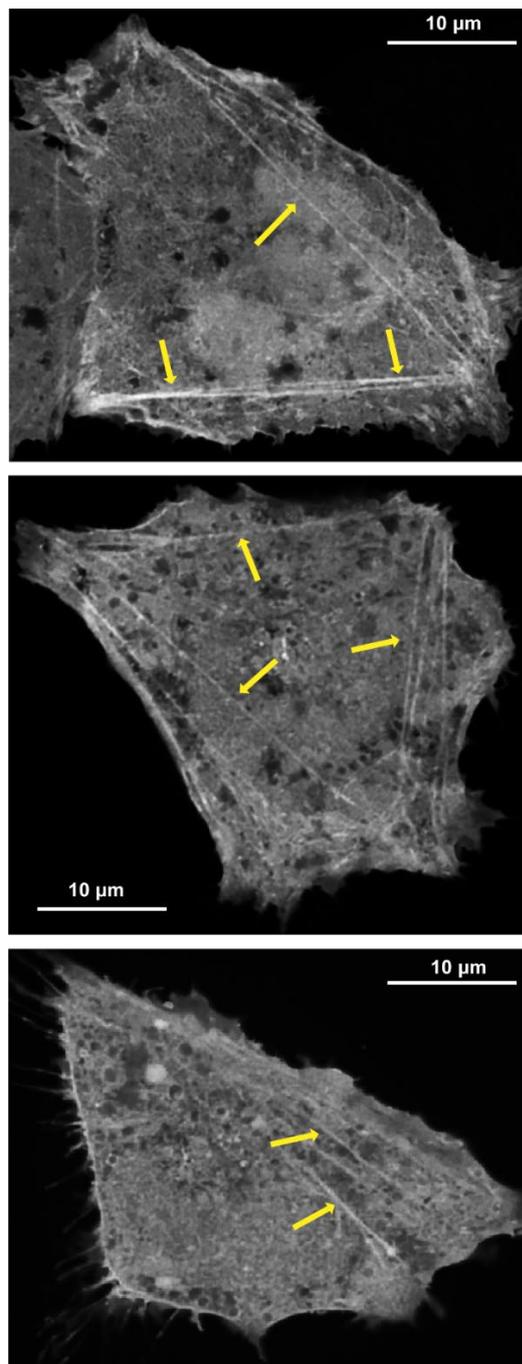


Figure 5.8. GFP-TLNRD1 4-helix domain localisation with ventral stress fibres

Spinning disk confocal microscopy of GFP-TLNRD1 4-helix domain. Colocalisation with ventral actin stress fibres indicated by yellow arrows.

5.2.6. TLNRD1 influences 2D random cell migration

Previous published research appeared to show that TLNRD1 expression or silencing influences migration in hepatocellular carcinoma and bladder cancer cells (Tatarano *et al.* 2012; Wu *et al.* 2017). To explore whether TLNRD1 changes the random migratory behaviour of cells, 2D migration experiments were performed over a 24-hour period with cells plated on 10 $\mu\text{g/ml}$ fibronectin. This was done using the same U2OS cells where TLNRD1 localisation to filopodial tips was observed, comparing untransfected wild-type cells with GFP-TLNRD1 expressing cells with GFP only as a control. Additional comparison was made with 3 siRNA's targeting TLNRD1 expression with AllStar siRNA used as a negative control.

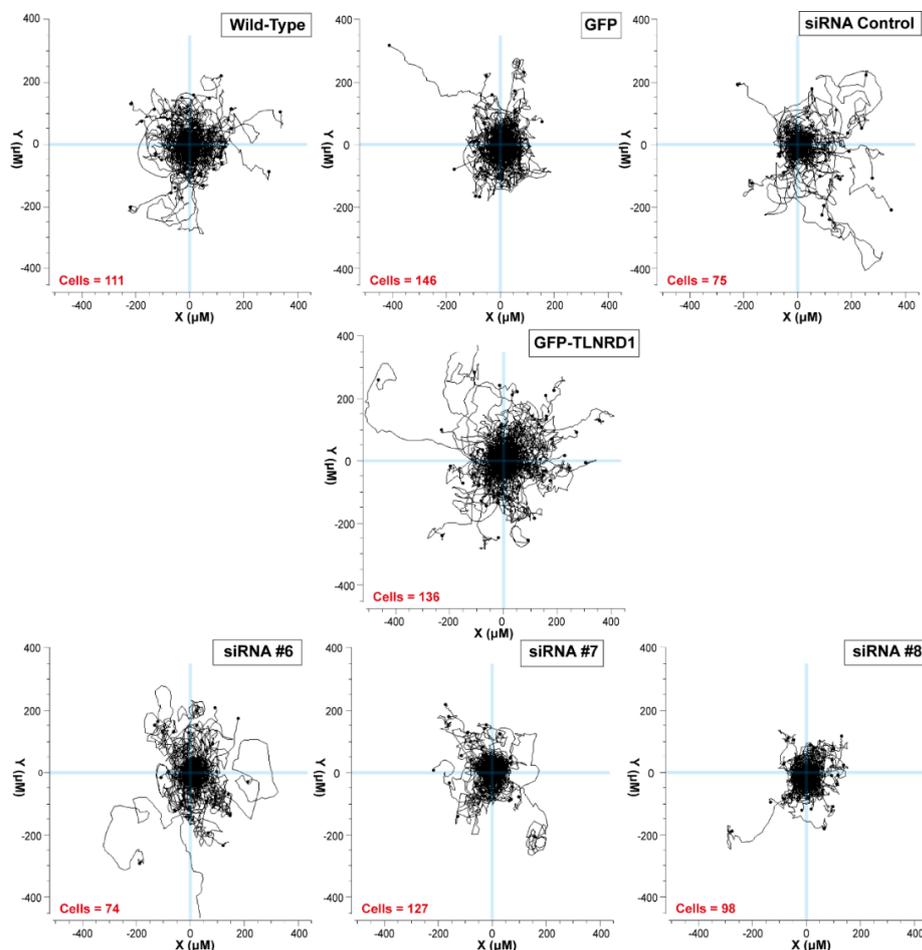


Figure 5.9. Spider plots of random migration in 2D

Trajectory spider plot of random migration with U2OS cells expressing GFP, GFP-TLNRD1 and cells transfected with TLNRD1 targeting siRNA's 6, 7 and 8. Migration measured over 24 hours. All tracks are displayed from the same origin with x/y intersection indicated. Total number of tracked cells indicated in red.

Initial inspection of cell migration through direct observation of the cell trajectories appear to show an increase in the distance in migration from the point of origin with GFP-

TLNRD1 expression. Conversely, no obvious differences can be observed for siRNA 6 however siRNA's 7 and 8 both appear to show a smaller area of migration from the point of origin suggesting that inhibition of TLNRD1 is reducing migration.

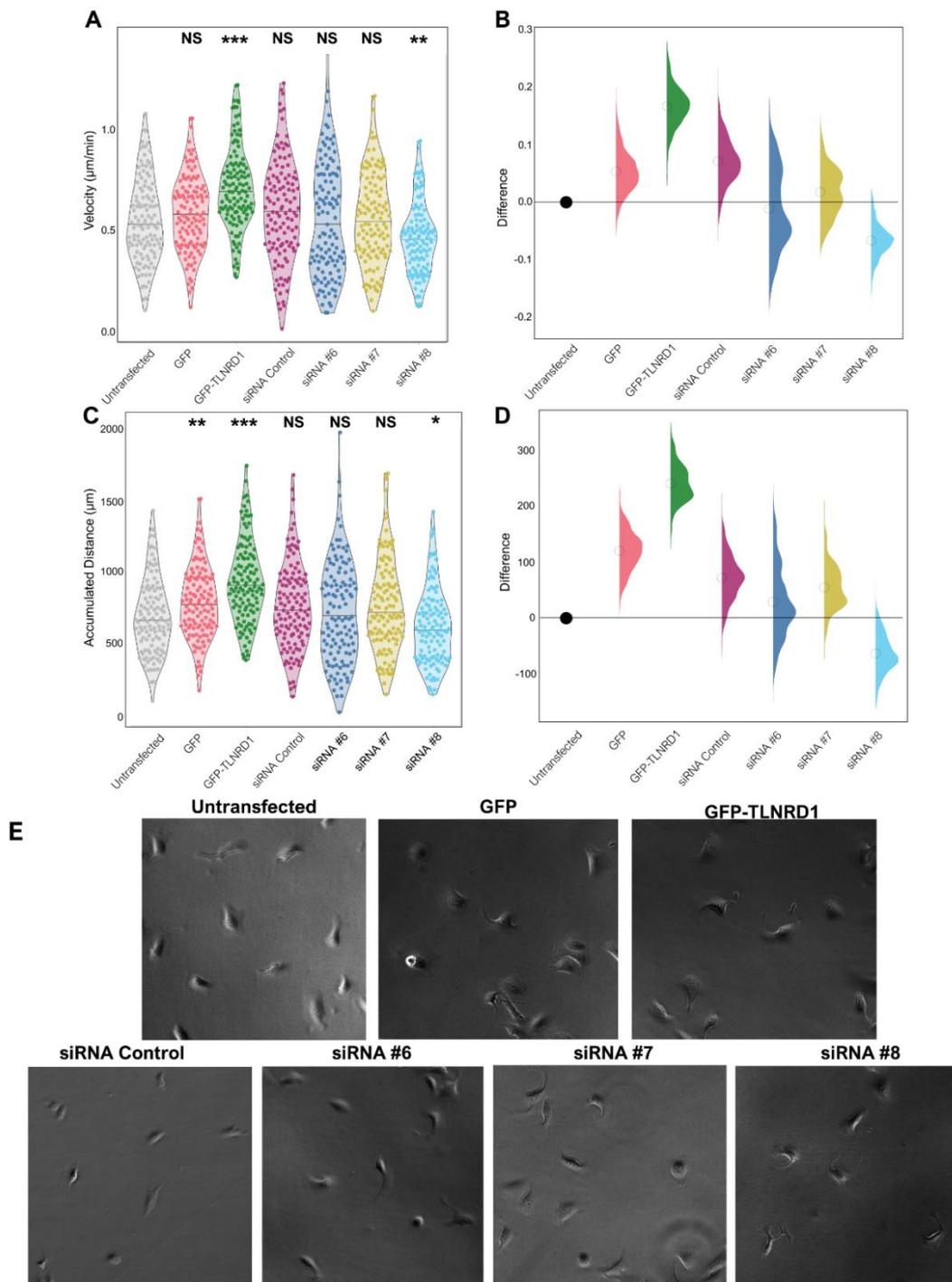


Figure 5.10. Migration analysis of wild-type vs. GFP-TLNRD1 and siRNA knockdowns

(A) Quantification of random cell migration velocities in $\mu\text{m}/\text{min}$. **(B)** Difference plot relative to untransfected control. **(C)** Accumulated distance that tracked cells have migrated (μm). **(D)** Difference plot of accumulated distance. $N=120$ cells from 3 independent repeats. Graphs and statistical analysis performed with PlotsOfDifferences (Goedhart, 2019). ***P value <0.001 , ** P value <0.01 , * P value <0.05 . **(E)** Still images taken from time-lapse imaging using brightfield microscopy.

The measured tracks were analysed to determine cell directionality, migration velocity and distance travelled over 24 hours using Ibidi chemotaxis and migration tool. Data was plotted as violin plots to show data distribution and as difference plots to visualise changes, then supplemented with randomised statistical analysis using the web application PlotsOfDifferences (Goedhart, 2019). This revealed an overall increase in both cell migration velocity and total accumulated distance with GFP-TLNRD1 (Figure 5.10A/C) expression when compared to untransfected cells, with $p < 0.001$ and a difference of $0.166 \mu\text{m}/\text{min}$ for velocity (Figure 5.10B) and $238.28 \mu\text{m}$ for accumulated distance (Figure 5.10D). With siRNA knock-down of endogenous TLNRD1 a slight reduction was observed in migration velocity and accumulated distance but the results varied slightly across the three siRNAs used. The strongest effect was observed with siRNA #8 with a P value of 0.003 ($P < 0.01$) and a difference of $-0.067 \mu\text{m}/\text{min}$ for migration velocity, and a P value of 0.034 ($P < 0.05$) and a difference of $-64.369 \mu\text{m}$ relative to control for accumulated distance. This supports previous research suggesting that there is a connection between increased TLNRD1 expression and increased cell migration. Knock-down of endogenous TLNRD1 showed no differences compared to wild-type.

5.3. Discussion

Both morpholino and CRISPR Cas9 knock-out studies to explore TLNRD1 function have revealed that despite previous research dismissing TLNRD1 involvement in mammalian development, the protein appears to be crucial in the zebrafish model. Zebrafish share highly similar developmental processes to complex mammalian organisms with a similar vascular architecture, central nervous system morphology and organised patterning of signalling pathways for limb formation. Targeting TLNRD1 expression resulted in drastic defects in zebrafish phenotype, with malformations in vascular structures, loss of fin-fold development and oedema around the brain and heart. Vascular development requires endothelial cells which stem from the mesoderm germ layer (Käßmeyer *et al.* 2009), suggesting that while TLNRD1 is not directly required for development of the mesoderm, it may be involved in vascular development either directly or indirectly which is dependent on cells from a mesodermal lineage. The expression data on the EBI database showed highest TLNRD1 expression during the pharyngula period where the embryo begins to straighten out, pectoral fins develop, heartbeat and circulation begins and the nervous system starts to expand anteriorly. Some of the defects observed through the targeting of TLNRD1 expression, such as abnormal tail curvature, loss of fin formation and vascular abnormalities suggests that TLNRD1 may be particularly required during

this stage of development, therefore future work should focus on exploring exactly where and when TLNRD1 is expressed. With zebrafish optical transparency and easy genetic manipulation, it would be possible to create transgenic zebrafish carrying fluorescently labelled TLNRD1 protein which would allow visual observation of TLNRD1 localisation and expression as the embryo develops. Work now continues to determine whether a similar phenotype is observed in a mouse model, which will give us greater understanding of TLNRD1 function during human development.

Analysis of cancer cell migration and TLNRD1 localisation reveals that GFP tagged TLNRD1 localises to filopodial tips and significantly increases cell migration speed in a 2D environment. Filopodia are important sensors for exploring the cells environment and are required for promoting cell migration through regulation of focal adhesion complexes (Schäfer *et al.* 2009). The increase in filopodia formed after TLNRD1 overexpression is likely associated with the increase in migration speeds observed and suggests that TLNRD1 behaves as an enhancer of cellular migration through the promotion, or stabilisation, of filopodial complexes. It must be noted, however, that while the cell migration results show an increase in migration with TLNRD1 overexpression, only a slight decrease was observed with one of the siRNA's with variable results across the three siRNA's tested. Additionally, while GFP expression had no significant affect on cell velocity, it did appear to alter cellular migration as shown by the statistically significant increase in accumulated distance. While overexpression of a protein of interest can lead to clues about a protein's behaviour, this does not reflect what may actually be happening in the cell with normal endogenous expression levels. Future work should build upon the research presented here and explore this connection between TLNRD1, cell migration and filopodia formation in a more endogenous context, removing artefacts introduced by the presence of a GFP tag and exploring whether this effect is more or less pronounced in a physiologically relevant 3D environment.

Filopodia have been shown to aid angiogenic branching during development by coordinating adhesion complex behaviour (Fischer *et al.* 2019). This, along with the zebrafish phenotypes observed, could lead to speculation that TLNRD1 promotes filopodia formation during early development in a process which aids cell migration necessary for the development of vascular structures. However, more work is now required to understand how this works in an endogenous context and determine whether the vascular defect observed is direct or indirect i.e. a consequence of impaired development in the development of a related system. TLNRD1 mRNA is expressed

across all tissues at high levels but the functional role of this continuous mRNA expression is yet to be discovered. In the experiments shown here, changes could only be observed with the overexpression of TLNRD1 protein from a transfected construct, with no clear changes to cell behaviour in siRNA knock-downs of endogenous protein. Despite the high mRNA levels in different tissues, evidence from the U2OS IP experiments reveals relatively low levels of endogenous protein being expressed that can be detected by antibody recognition. This could explain why endogenous TLNRD1 has little effect on cell behaviour under normal conditions, and suggests that activity may increase in response to factors such as the release of certain growth hormones during development. It may also indicate that antibody efficiency is low, while this is the first time being able to detect TLNRD1 protein via IP, further work is required to test different antibodies and determine how much protein is detected in whole cell lysates.

Overall, the work presented in this section reveals the importance for TLNRD1 during vertebrate development, particularly in development of the vascular system. It also reveals that TLNRD1 localises to filopodial tips, with increases in TLNRD1 expression promoting filopodia formation and subsequently increasing the speed of cellular migration. Work now continues to pick apart how TLNRD1 exerts this effect, why the increase in filopodia and migration may be required in a biological context and understand if TLNRD1 expression is regulated in a developmental dependent manner.

Chapter 6. Exploration of cancer associated point mutations in talin

6.1. Introduction

6.1.1. Talin and its association with cancer

The talin molecule is an essential scaffold for the formation of most integrin-mediated adhesion complexes, acting as a mechanosensitive signalling platform for the regulation of integrin engagement with the ECM. Not only does talin regulate direct ECM adhesion, but it's also intricately connected to the regulation of other cellular processes such as cell division and ion transport through interaction with numerous different ligands. This puts talin at the centre of an incredibly complex signalling network, transforming extracellular and intercellular mechanical forces into different signalling outputs for the regulation of cell behaviour (Goult, Yan and Schwartz 2018).

Extensive research over the years has revealed talin to be heavily involved in numerous different cancer types. High levels of talin-1 expression has been identified in several cancers, with overexpression of the protein being associated with advanced cancers such as oral squamous cell carcinoma and hepatocellular carcinoma (Lai *et al.* 2011; Kanamori *et al.* 2011). Increased talin-1 phosphorylation at serine 425 has also been linked with increased β 1 integrin activation and enhanced metastasis in prostate cancer (Jin *et al.* 2015). Furthermore, increasing evidence shows that both isoforms can promote cancer related processes with talin-2 overexpression in breast cancer increasing cancer invasiveness (Wen *et al.* 2019).

With talins integral role in orchestrating cellular adhesion and behaviour, the impact of any modification or mutation to the domains of talin could have drastic consequences on cell behaviour. Numerous studies have implemented single point mutations to explore and alter talin behaviour. Examples include the single point mutant of E1770 which when mutated to alanine stops talin autoinhibition, leading to increased integrin activation, persistent focal adhesion complex formation and impaired wound healing (Zhang *et al.* 2016; Haage *et al.* 2018). Single point mutations in the talin R3 subdomain have also been shown to be sufficient for increasing talin and vinculin recruitment to adhesion complexes, altering their dynamics (Haage *et al.* 2018; Elosegui-Artola *et al.* 2016).

6.1.2. Cancer-associated databases

Rapid advances in technology are making it easier to screen cancers for associated mutations and obtain expression data on a vast number of genes. This has led to the development of large public databases curating huge amounts of data describing

mutations in almost all genes in the human genome and analysing expression data to determine trends in a proteins involvement in cancer.

A substantial proportion of publicly available data comes from The Cancer Genome Atlas program (TCGA) which began in 2006 to take advantage of advancing genomic technologies. Through this program extraordinary amounts of genome sequencing and gene expression data has been generated, as well as RNA expression profiling and epigenetic modification profiling in tumours (Weinstein *et al.* 2013). A second large scale project, The Genotype-Tissue Expression project (GTEx) has also contributed substantial amounts of data on gene expression and genetic variation on healthy disease-free tissues. Together these two sets of data are used by databases to discover changes between at the genetic and protein level in healthy vs cancerous tissue.

Launched in 2004 by Sanger, the Catalogue of Somatic Mutations in Cancer (COSMIC) (Tate *et al.* 2019) has grown to be the largest publicly available database of curated somatic mutations, becoming a vast resource of genetic information for the exploration of cancer-associated mutations and aberrant expression. Other databases include the cBio Cancer Genomics Portal developed at Memorial Sloan-Kettering Cancer Centre with the aim of optimising interpretation and visualisation of the data (Cerami *et al.* 2012) and newly created Gene Expression Profiling Interactive Analysis (GEPIA) which provides differential expression and correlation analysis for the TCGA and GTEx data (Tang *et al.* 2017).

6.1.3. Chapter overview

In light of talins close connection with cancer progression and the evidence demonstrating how single point mutations can alter cell adhesion and migratory dynamics, we sought to explore whether mutations which may have a functional impact on a protein can be predicted using a bioinformatics pipeline. In collaboration with the lab of Associate Professor Vesa Hytönen at Tampere University, we combined biochemistry-based techniques from the Goult lab with cell biology from the Hytönen lab to explore the impacts of cancer-associated talin-1 mutations with the intention of developing a pipeline that will robustly screen cancer associated mutation data and identify potentially problematic mutations which could alter protein function. As previously described, the talin R7R8 domain acts as a signalling hub for a number of different ligands, many of which have now been shown to directly interact with TLNRD1. Therefore, most of the mutations biochemically explored here focus on single point

mutations in the R7R8 domain which could later aid understanding of TLNRD1 behaviour and its association with promoting migration and invasion in cancers.

6.2. Identification and evaluation of talin-1 single point mutations

To date there are nearly 700 identified mutations in the talin-1 gene on the COSMIC database, 55% of which are missense mutations resulting in an amino acid change in the expressed protein (Version 90; Accessed October 2019). These mutations have been identified in over 28 different types of cancerous tissue, showing the widespread implications for talin-1 mutation. The data also shows that in a vast number of cancers, talin-1 is overexpressed with up to 636 cases identified compared to 233 identified as underexpressing talin-1.

The Hytönen lab evaluated up to 260 talin-1 missense mutations identified in the COSMIC database to assess their potential impact on protein function. Analysis of predicted pathogenicity for each substitution was performed using the publicly available PON-P2-algorithm which uses conservation and biochemical analysis of residues to produce a probability score (Niroula, Urolagin and Vihinen 2015). This algorithm predicted up to 11 potentially pathogenic variants of talin-1 which were further investigated for predicted biochemical alterations such as hydrophobicity changes, structural impact based on location along the talin molecule as well as potential to disrupt binding to known ligands. The mutations were narrowed down to a final 10 (Table 6) which were to be explored further using cell biology by the Hytönen lab, four of which were also investigated using a range of biochemical techniques by myself as described in this chapter. Additionally, the E1770Q mutant which had a lower impact score was also included in cell biology analysis as this residue is important for regulating talin-1 behaviour through autoinhibition (Goult *et al.* 2009; Haage *et al.* 2018). The mutations were screened against the 1000 Genomes Project database (Clarke *et al.*, 2012) which revealed that all these mutations are only found on the COSMIC database and have not been identified in healthy individuals.

Mutant	Tissue and Histology	Talin-1 Domain	PON-P2 probability	Normalised BLOSUM 62-matrix score	Normalised ConSurf score	Final score
P229L	Skin - Carcinoma	F2	0.94	1	1	8.2
I392N	Pancreas - Carcinoma	F3	0.86	1	1	8.95
V577D	Liver - Carcinoma	R1	0.94	1	0.6	8.39
A893E	Central nervous system/pituitary - Glioma/Craniopharyngioma	R3	0.95	0.75	0.6	7.63
R1368W	Hematopoietic and lymphoid tissue/large intestine - lymphoid neoplasm/carcinoma	R7	0.98	1	1	7.75
Y1389C	Liver - Carcinoma	R7	0.9	0.87	1	8.58
L1539P	Liver - Carcinoma	R8	0.98	1	0.6	8.45
S1750F	Skin - Malignant melanoma	R9	0.89	0.87	0.6	7.55
E1770Q	Breast - Carcinoma	R9	0.81	0.37	1	6.01
D2086V	Breast - Carcinoma	R11	0.96	1	1	7.71
L2509P	Large intestine - Carcinoma	DD	0.96	1	0.64	7.46

Table 7. COSMIC talin-1 mutations identified for further investigation (Latifeh Azizi)

Summary of selected talin-1 point mutations from the COSMIC database and associated details, including domain location within talin-1. Values range from 0-1 with 1 having the highest probability of causing defects in protein function. Scoring coefficient and weight was used to generate a final score. Samples with further biochemical analysis are highlighted in green.

The selected mutations are dispersed across the entire talin-1 molecule which suggests they have the potential to alter different aspects of talin functionality (Figure 6.1). Mutations P229L and I392N both reside in the talin head FERM domain, in domains F2 and F3 respectively. The F2 domain is important for the initial talin interaction with the plasma membrane (Kalli *et al.* 2010; Arcario and Tajkhorshid 2014), whereas F3 houses the β -integrin binding site required for talin-mediated activation of integrin (Wegener *et al.* 2007).

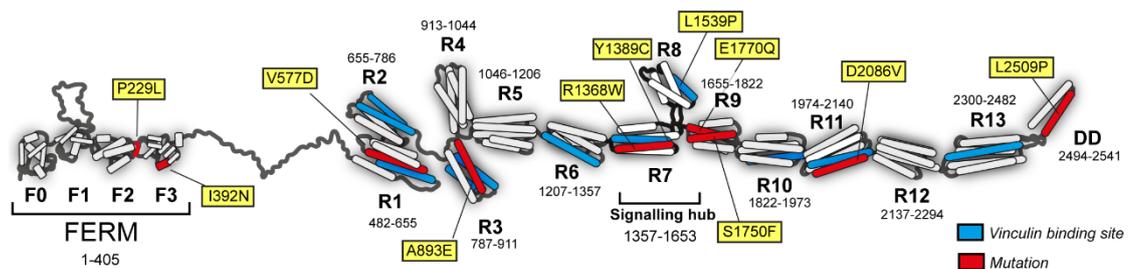


Figure 6.1. Location of selected mutations along the talin-1 molecule

Schematic representation of the talin-1 molecule and domain boundaries with helical location of each single point mutation selected for study highlighted in red. Vinculin binding sites are highlighted in blue.

The remaining mutations are found along the talin rod domain in locations which could potentially disrupt ligand binding or alter the mechanical properties of the associated helical bundle. Most notably, up to 3 potentially pathogenic mutations were identified in the R7R8 domain alone. These mutations in particular have the potential to disrupt interactions with key talin ligands. R1368W resides within the previously characterised

KANK1 binding site on the R7 domain with Y1389C also in close proximity and located in a region which has the potential to destabilise the helical bundle (Figure 6.2A, B). L1539P sits in the R8 domain (Figure 6.2C) which interacts with a number of different LD motif containing ligands and stabilises interaction with actin filaments as part of ABS2. A final notable mutation, L2509P, was identified in the dimerisation domain of talin and has the potential to disrupt the helix which mediates dimerisation (Figure 6.2D). Talin dimerisation is essential for assembly of focal adhesion complexes, cell spreading and actin filament engagement by ABS3, with mutations which prevent dimerisation leading to drastic phenotypic changes in migratory cells (Gingras *et al.* 2008; Goult *et al.* 2013b).

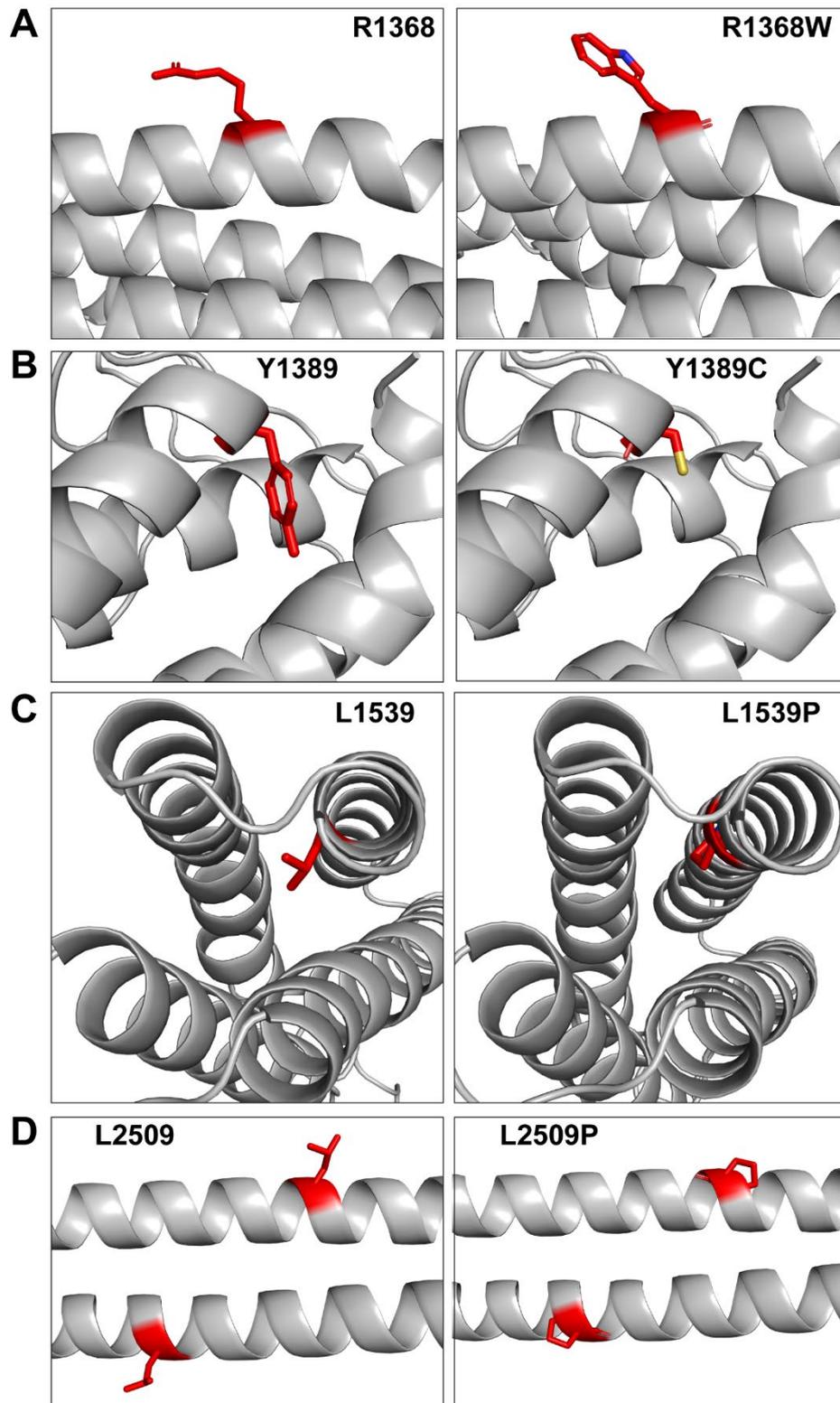


Figure 6.2. Structural comparison of wild-type and mutated residues

(A) Residue R1368 and mutation R1368W in R7 (PDB ID: 2X0C), **(B)** Residue Y1389 and mutation Y1389C in the core of the R7 helix bundle, **(C)** Residue L1539 and mutation L1539P in the core of the R8 domain, **(D)** Residue L2509 and mutation L2509P in the dimerisation domain (PDB ID: 2QDQ).

6.3. Phenotypic comparison of talin-1 point mutations

All cell biology experiments presented in this section were performed by the Hytönen lab using talin-null mouse kidney fibroblasts for introduction of the talin-1 gene carrying each point mutant to be investigated. With each mutation, a number of different parameters were measured including cell circularity and expression levels of focal adhesion markers as well as migration speed, cell division and invasiveness of the cells, increases in which are hallmarks of cancerous cells. For all transfected constructs, the total expression levels of the mutated talin-1 genes were compared against wild-type talin-1 expression levels to determine if there was any variance in translated protein. This revealed little variation in expression levels between each tested construct and wild-type, showing that they are directly comparable.

6.3.1. Most talin-1 point mutations have little impact on cell morphology

Comparison of each mutant against wild-type cells with non-mutated talin-1 revealed very little initial observable differences in cellular phenotype which includes cell area and circularity, factors which enable us to determine changes in the behaviour of the actomyosin cytoskeleton in response to the different talin mutations. Only one mutant produced a drastically altered phenotype and that was the L2509P mutation located in the dimerisation domain of talin. These cells had a considerably more rounded phenotype and much smaller cell area, showing that the cells are unable to spread or protrude as they normally would during migration. Previous research has demonstrated that the C-terminal dimerisation domain is important in maintaining interaction with actin filaments and thus targeting of talin to focal adhesions (Gingras *et al.*, 2008; Smith and McCann, 2007). This suggests that the L2509P disrupts either talin-1 ABS3 association with the actin cytoskeleton and/or talin dimerisation, both of which are thought to be required for talin activation, adhesion stabilisation and establishing cell polarity (Goult *et al.* 2013b; Rahikainen *et al.* 2019).

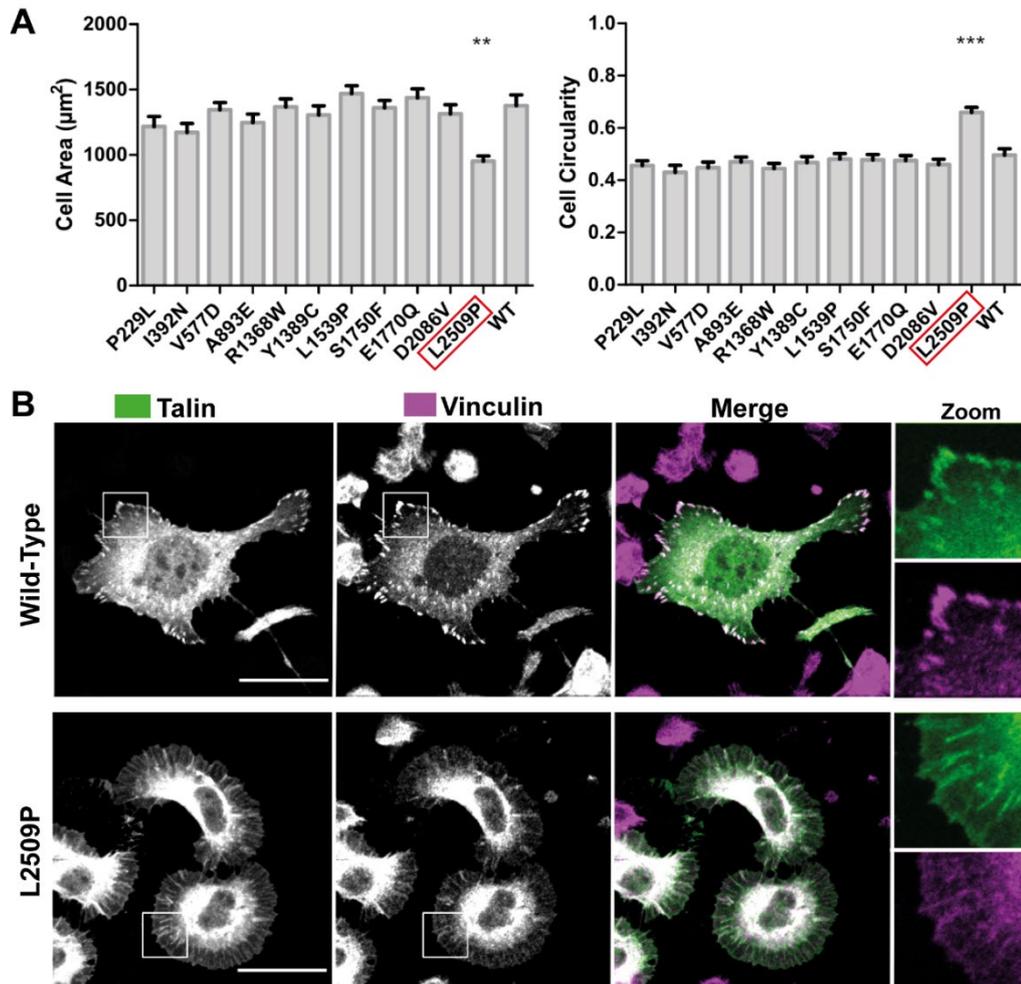


Figure 6.3. Comparison of wild-type vs. talin-1 mutant phenotypes (Latifeh Azizi)

(A) Comparison of cell area (μm^2) and cell circularity of each single-point mutant with wild-type cells. $N=40$ from four experiments. **(B)** Microscopy images comparing wild-type vs. L2509P mutant cells with GFP labelled talin-1 (green) and immunolabelled vinculin (purple).

6.3.2. Cancer-associated talin-1 point mutations alter cell behaviour

A large proportion of protein mutations identified in cancerous cells are introduced in a way which enhances rates of division, migration and overall invasiveness of the cell. To explore whether each mutant alters cellular behaviour, random 2D migration speed was measured with cells plated on $10 \mu\text{g/ml}$ fibronectin, cell division measured over 12 hours and cell invasion measured in a 3D Matrigel chamber after 24 hours (Figure 6.4). Cell division was measured with both full-serum (10% FBS) and reduced serum conditions with 0.2% FBS. Serum starvation puts cells under environmental stress and can reflect the nutrient depleted environments of tumours (Aghababazadeh, 2014).

With the mutants tested in this experiment, a lot of the cells exhibited subtle changes in migration speed and rates of division when compared to wild-type cells, with differing responses depending on whether cells were under serum starved conditions or fully

supplemented with foetal bovine serum (FBS). Only one point mutant, I392N which sits close to the integrin binding site of talin, increased cellular migration to a speed of 0.82 $\mu\text{m}/\text{min}$ compared to wild-type cells with a speed of 0.65 $\mu\text{m}/\text{min}$ (Figure 6.4A). Conversely, mutations R1368W and L2509P led to a drastic reduction in normalised migration speed with mutation L2509P having the greatest impact, reducing average speed to just 0.37 $\mu\text{m}/\text{min}$. This is unsurprising considering the drastic morphological changes seen with this mutant which suggest the cells are not able to form the mature adhesions required for migration in a 2D environment. Only subtle non-significant changes in speed were observed for the remaining mutants.

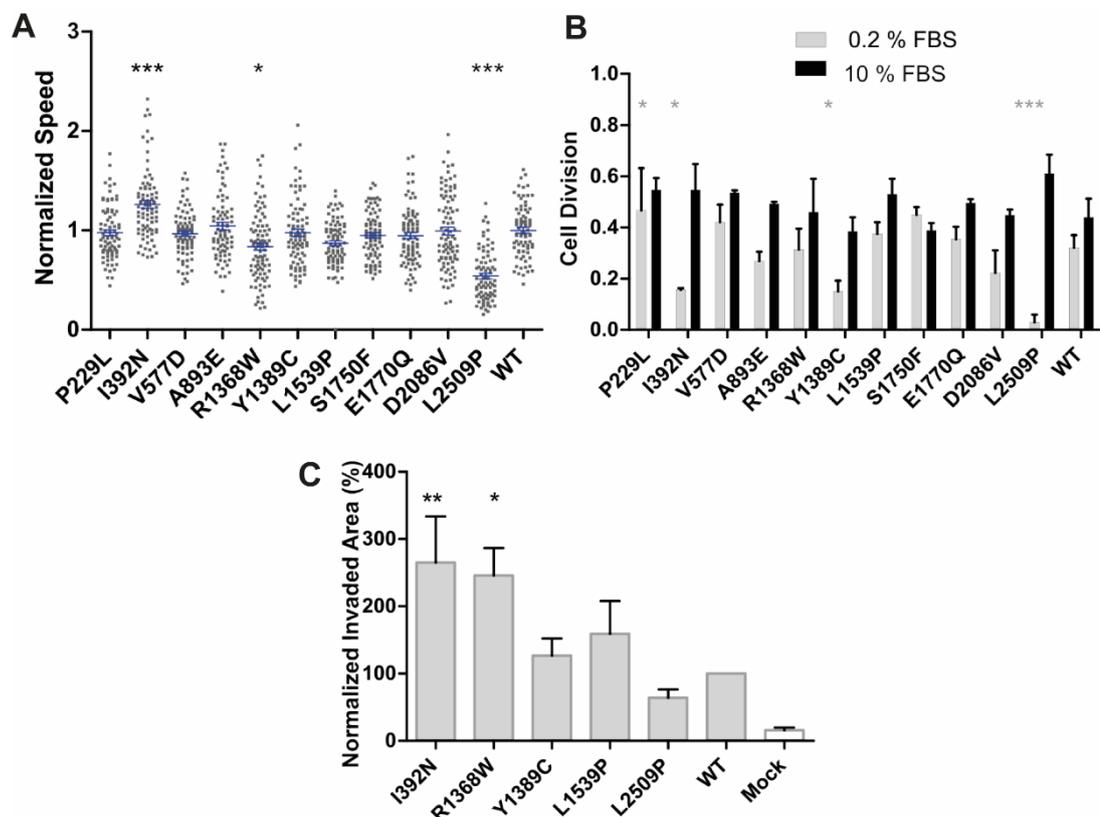


Figure 6.4. Comparison of cell migration, division and invasion in wild-type vs. mutants (Latifeh Azizi)

(A) Random cell migration speed ($\mu\text{m}/\text{min}$) across all tested constructs. **(B)** Number of dividing cells over a 12 hour period with comparison of 0.2% and 10% FBS conditions. **(C)** Cell invasion through 3D Matrigel towards media with 10% FBS. Mock control transfections performed using GFP expression vector alone. N=80 cells, t-test statistical analysis and non-parametric test of Mann-Whitney used: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Talin-dependent adhesion and integrin regulation is closely connected to cell division, with talin-1 depletion leading to defects in cellular proliferation and high talin-1 expression being associated with increased proliferation (Chen *et al.* 2017; Wang, Ballestrem and Streuli 2011).

To explore whether any of the identified mutations altered rates of cell division, the number of dividing cells were measured over 12 hours in 0.2% FBS serum starved and 10% FBS fully-supplemented conditions (Figure 6.4D). In full supplemented conditions, some mutants showed marginal increases in cell division compared to wild-type cells, including L2509P, L1539P and I392N. The most drastic differences were observed under serum starved conditions, with I392N, Y1389C and L2509P showing a considerably large drop in division rates compared to wild-type cells. Wild-type division dropped by 27%, whereas I392N and L2509P division dropped by 71% and 95%, with proliferation almost stalling entirely in the L2509P mutant expressing cells. While serum starvation can reflect nutrient and growth factor depleted environments, it is known that cells will stop dividing when starved for any period of time. The differences observed in the drop-off for replication suggests that some mutations may impact how cells respond to environmental stressors, however, this approach is not robust and can lead to spurious results (Pirkmajer and Chibalin, 2011).

The five mutants which showed the largest impact on migration speed and cell division were further investigated for differences in invasiveness through a 3D Matrigel using 10% FBS containing media as the chemoattractant (Figure 6.4C). Increases in invasiveness allows cancerous cells to migrate through the ECM and become metastatic as they migrate to other areas of the body through the blood circulation. It should also be noted that in a 3D environment, cells produce adhesion complexes with increased longevity and slightly different composition to the focal adhesions observed in cells migrating in a 2D environment (Doyle and Yamada 2016). Thus, examining cell migration in both 2D and 3D environments enables exploration of the different dynamics of adhesion complex formation and subsequent migration. Mutants Y1389C, L1539P and L2509P showed no significant difference in invasiveness compared to wild-type cells showing the migration in a 3D context was unaffected with these particular mutations. Mutations I392N and R1368W on the other hand showed a marked increase in invasiveness. Overall, I392N appears to increase cellular migration in both the random 2D context and 3D directional environments, whereas R1368W only showed an increase in the 3D Matrigel, suggesting that the location of this mutation is specifically altering the speed of directional cell movement towards a chemoattractant.

6.3.3. Cancer-associated talin-1 point mutations change adhesion composition

With talin having such a core role in integrin activation and adhesion complex formation, the Hytönen lab sought to investigate whether any of the cancer-associated point

mutations were also leading to changes in focal adhesion composition and overall expression of focal adhesion components vinculin, focal adhesion kinase (FAK) and paxillin. Vinculin is recruited to adhesions by paxillin and is activated through direct interaction with multiple sites along the talin rod domain when talin unfolds under mechanical load, and this promotes the maturation and stabilisation of adhesion complexes (Case *et al.* 2015). Focal adhesion kinase is one of the first proteins to be recruited to adhesion complexes and is required for recruitment of talin to sites of nascent adhesion formation (Lawson *et al.* 2012). FAK autophosphorylates at tyrosine residue 397 after recruitment to adhesion complexes which allows easy tracking of FAK activation and thus adhesion complex formation in cells (Schaller *et al.* 1994). Finally, paxillin a core component of adhesion complexes, is a scaffolding adaptor protein that is required for recruitment of multiple adhesion components such as FAK (López-Colomé *et al.* 2017).

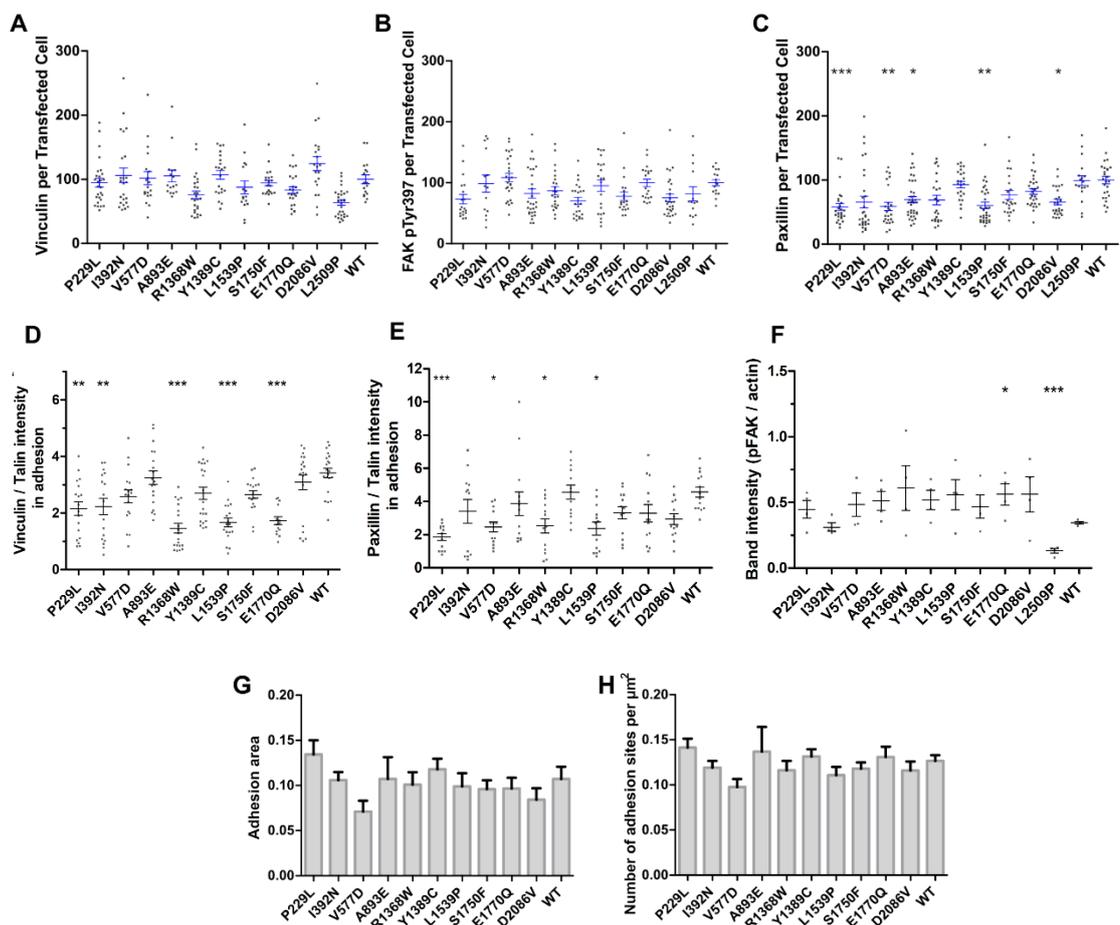


Figure 6.5. Comparison of adhesion component expression and localisation with talin (Latifeh Azizi)

(A, B, C) Comparison of expression levels of focal adhesion components vinculin, FAK pY397 and paxillin between mutants and wild-type cells using immunofluorescence. (D, E, F) Colocalisation comparison measuring intensity at sites of adhesion complex formation between components vinculin, FAK pY397

and paxillin with each talin construct. **(G)** Analysis of total adhesion area. **(H)** Quantification of total number of adhesion sites per μm^2 .

Analysis of protein expression levels in response to the introduction of each talin mutant showed only subtle differences in vinculin expression and phosphorylated FAK. Analysis of paxillin however showed significant reductions in expression with the introduction of mutants P229L, V577D, A893E, L1539P and D2086V (Figure 6.5C). Colocalisation analysis of the talin mutants with adhesion components vinculin, FAK pY397 and paxillin showed considerable variation in adhesion composition with the majority of mutants seeing a drop in recruitment of the adhesion components (Figure 6.5D, E, F). For paxillin recruitment, the biggest effect was seen with mutation P229L which resides in the F2 domain of talin which initiates talin interaction with the plasma membrane. With vinculin and FAK pY397 the most significant effects were seen with mutations R1368W and L1539P in talin R7 and R8 domain respectively, which saw the biggest reduction in FA component recruitment. This region of talin acts as a signalling hub and has a number of ligands including actin, vinculin and LD motif containing ligands such as KANK, RIAM and DLC1. It is likely that disruption of interactions in these particular regions are changing the signalling dynamics of adhesion complexes, leading to a reduction in protein recruitment. The mutation L2509P led to significant disruption of focal adhesion complex formation as observed in figure 6.3B and as such colocalisation analysis of focal adhesion components at the sites of adhesion complex formation could not be defined.

6.4. Biochemical analysis of talin-1 cancer associated mutations

Talin behaviour is incredibly complex with the potential of multiple interactions happening at any one time. In order to try and understand in a bit more detail about how talin-1 point mutations may be affecting its behaviour or structural integrity, a number of different biochemical techniques were used to explore the impact of selected mutations. The R7R8 mutations R1368W, Y1389C and L1539P were selected for further analysis along with the talin dimerisation domain mutant L2509P which appeared to show the greatest impact on cellular phenotype. Each R7 or R8 located mutation was introduced into an R7R8 construct (residues 1357-1653) as the wild-type R7R8 construct has been shown previously to express well and the L2509P mutant was introduced into an R13-DD construct which contains both the R13 actin binding domain and dimerisation domain. All mutants were first analysed by circular dichroism to determine protein thermostability and folded state. Talin is a highly mechanosensitive protein and minor changes to domain stability could lead to changes in how that particular domain unfolds under force, therefore this is a useful parameter to determine whether there may be any changes in talins dynamics.

6.4.1. Mutation L1539P disrupts the R8 helical structure

Initial purification of talin R7R8 containing the L1539P mutation showed a drastic reduction in overall yield with poor expression and an increased proportion of insoluble protein, as such only a low concentration of purified protein could be obtained, limiting the experiments which could be performed. With the soluble protein which was successfully purified, CD analysis revealed a drastically altered spectrum compared to wild-type R7R8 (WT) (Figure 6.6B). This spectrum shows a weaker high point at 215 nm suggesting the total alpha-helical content has been reduced compared to the wild-type and a lower signal at 190-200 nm suggests a degree of disorder in the protein. Overall, this CD spectra suggests that the L1539P is causing disorder and helical unfolding in the R8 domain, while the R7 domain retains the alpha-helical content. Analysis of the thermostability between wild-type R7R8 and L1539P only shows a minor reduction in overall thermostability by 2.5°C but has a more gradual slope of unfolding (Figure 6.6A).

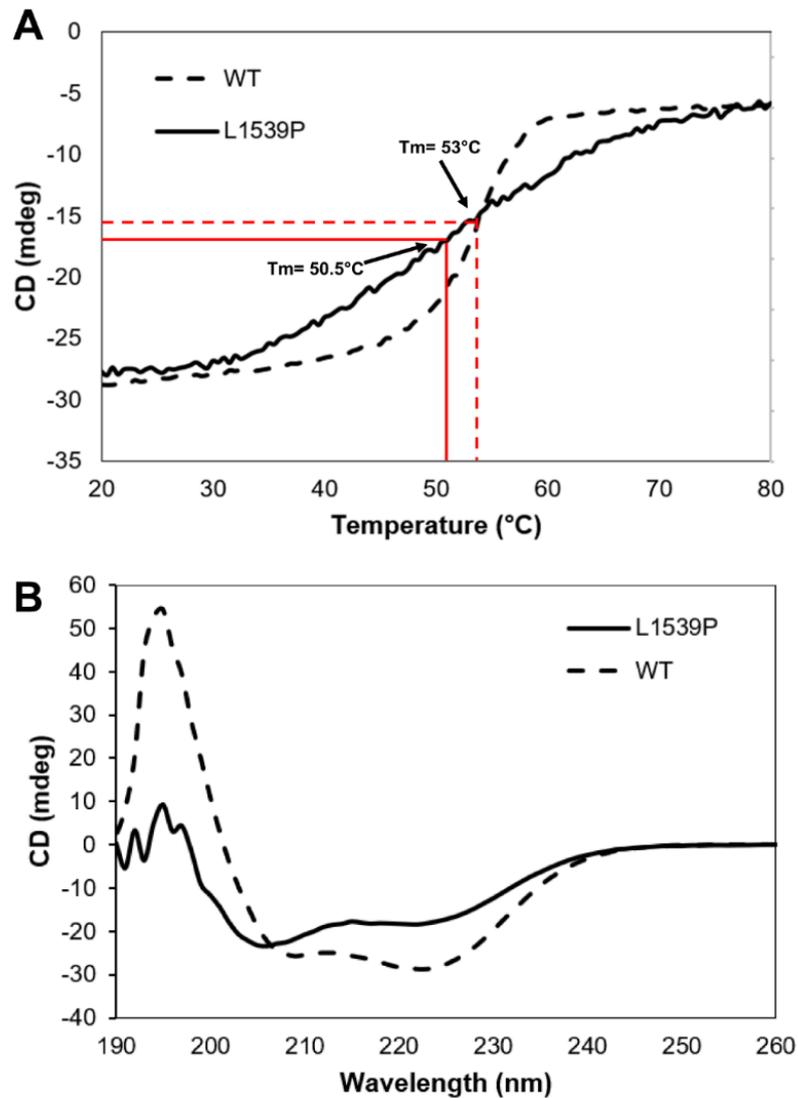


Figure 6.6. CD spectroscopy of wild-type R7R8 vs L1539P

(A) CD melt curve comparison of wild-type R7R8 and L1539P mutant from 20-80°C. Melting temperature is indicated for each construct. **(B)** CD spectrum comparison of wild-type R7R8 and L1539P.

6.4.2. R7 domain mutant Y1389C reduces protein thermostability

Both Y1389C and R1368W expressed and purified as well as wild-type R7R8 protein. Comparison of the CD spectra from wild-type and both mutants showed that there were no differences in secondary structure, confirming that with each mutation the protein was folded (Figure 6.7A). Thermostability analysis of the two mutants showed that while R1368W unfolds at the same temperature as wild-type R7R8 at ~53°C, Y1389C unfolded with a T_m of 47°C, showing a 6°C reduction in protein thermostability (Figure 6.7B). This decrease in thermostability suggests that introduction of the Y1389C mutation increases the unfolding susceptibility of the R7 domain.

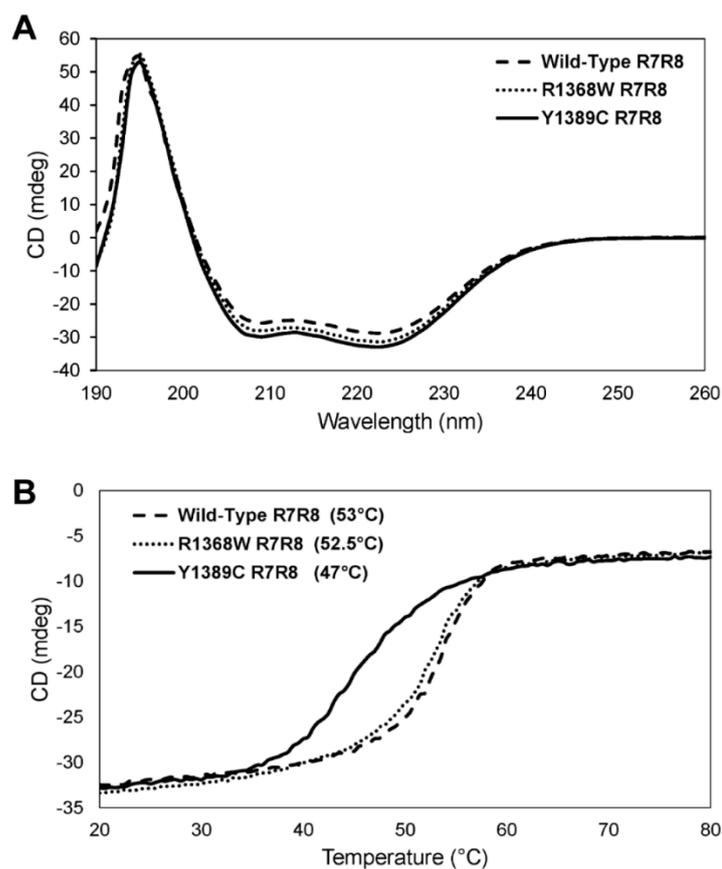


Figure 6.7. CD spectroscopy of wild-type R7R8 vs R1368W and Y1389C

(A) CD spectrum comparison between wild-type R7R8 and mutants R1368W and Y1389C. **(B)** CD melt curve comparison for each indicated construct. Melting temperatures are indicated in the figure legend.

6.4.3. NMR analysis of R1368W and Y1389C

To understand whether each mutant was causing any conformational or structural changes to the R7R8 domains, NMR HSQC spectra of each mutant was compared to wild-type. For this experiment, an NMR TROSY experiment was used which is optimised for larger proteins (see materials and methods section 2.7.6 for description). Each construct was analysed at 298K at a concentration of 160 μ M and peak distributions compared. Comparison of wild-type R7R8 and R1368W showed few changes in peak distribution showing that the protein largely remains the same conformationally and structurally (Figure 6.8A). The few residues which are altered are likely to correspond to the mutated residue itself and its surrounding residues. The Y1389C mutant on the other hand showed a surprisingly drastic change in peak distribution of the protein, with a large proportion of peaks disappearing from their original positions and clustering towards the centre of the spectrum (Figure 6.8B). This indicates that large scale changes have occurred across the molecule, drastically altering the chemical environment of the amino

acids, however some residues do remain in the same positions as the wild-type suggesting that only one of the domains is affected. To check whether the cysteine residue might be causing disulphide bond formation on an opposing R7R8 molecule, the spectra was compared both with and without DTT which showed no difference in the spectrum, suggesting that the peak distribution towards the centre is not likely due to increased protein size through disulphide bond formation.

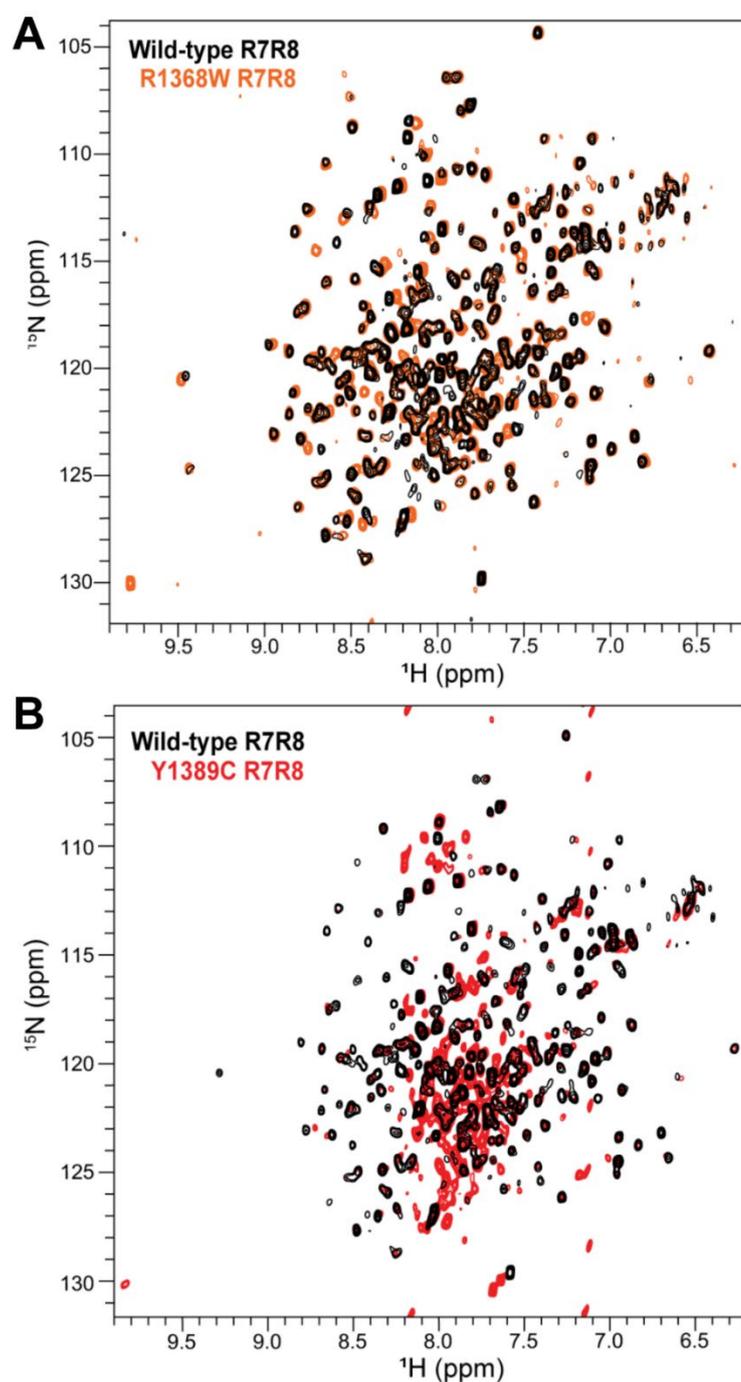


Figure 6.8. NMR ^1H ^{15}N HSQC TROSY spectra of R7R8 and mutants R1368W and Y1389C
(A) Overlaid HSQC TROSY spectra of 160 μM talin R7R8 wild-type (black) and R1368W (orange). (B) Overlaid HSQC TROSY spectra of 160 μM talin R7R8 wild-type (black) and Y1389C mutant (orange).

6.4.4. R7 domain mutant Y1389C influences vinculin binding

The reduced thermostability seen with the Y1389C mutant and its location in the core of the R7 domain offers the intriguing possibility that it may alter the unfolding properties of the domains. The R8 domain is largely protected from unfolding by the high stability of the R7 domain, with R8 domain unfolding being dependent on the initial unfolding of R7 at 15 pN of force (Yao *et al.* 2016). Both R7 and R8 contain a cryptic vinculin binding site which becomes exposed under force; therefore, vinculin binding is tightly regulated by the mechanical stability of the helical domains. Vinculin VD1 interacts with helix 36 in R7 and helix 33 in R8, with previous research using gel filtration showing that the R7 domain alone is unable to interact with vinculin at 20°C whereas R8 can interact with VD1 with high affinity (Gingras *et al.* 2010). Interestingly this high affinity of R8 for vinculin reduces with the presence of the R7 domain suggesting R7 is important for stability in both domains.

With the knowledge that vinculin binding to the R7R8 domain can be observed using gel filtration, we sought to use the same approach to determine whether R1368W or Y1389C have an impact on the R7R8 domains affinity for the vinculin VD1 domain. Vinculin VD1 was incubated at a 2:1 ratio with each construct and analysed using size-exclusion chromatography at room temperature (Figure 6.9). With wild-type R7R8, approximately half of the protein formed a complex with VD1, which is consistent with previous observations (Gingras *et al.* 2010), with the peak position showing that only a single vinculin VD1 bound to the protein. Similar binding was also observed with R1368W suggesting that neither R7 or R8 vinculin binding was affected by the presence of the mutation. In stark contrast to this, the Y1389C mutation which sits in the R7 domain core and reduces protein thermostability was able to bind two vinculin VD1 molecules as shown by the shift in peak position and increased band density on SDS-PAGE of the associated fraction. This indicates that the mutations destabilisation of R7 is enhancing the accessibility of the R7 vinculin binding site.

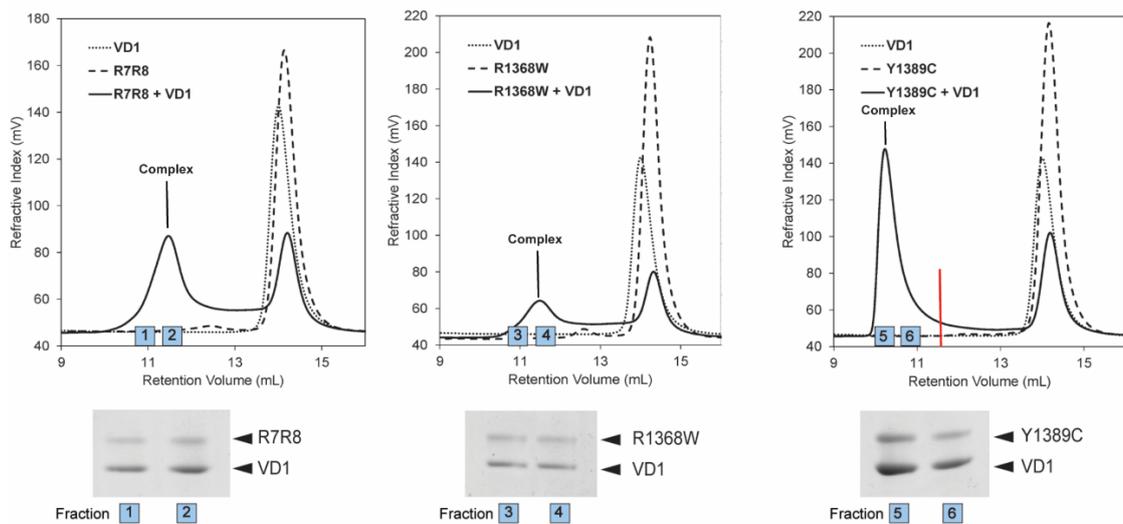


Figure 6.9. Analysis of vinculin binding to wild-type R7R8 and mutants R1368W and Y1389C
 Gel filtration analysis of vinculin VD1 domain binding to the R7R8 domain of talin and comparison with mutants R1368W and Y1389C at a 2:1 ratio. Peaks of each complex are indicated, with original wild-type complex peak position indicated in red on the Y1389C chromatogram (far right). SDS-PAGE analysis of the complex peak fractions is shown below and numbered accordingly.

6.4.5. Binding to KANK and RIAM appears unaffected by R1368W and Y1389C

The R7R8 domain is a ‘signalling hub’ within talin with a number of different ligands, many of which interact with talin through LD motif mediated helix addition. The R7 domain contains the binding site for interaction with KANK1, which promotes the formation of microtubule stabilising complexes at the periphery of adhesions. This interaction is capable of withstanding high forces and is mechanically stable (Yu *et al.* 2019), therefore any disruption to the R7 domain architecture has the potential to disrupt this stable interaction. Both R1368W and Y1389C sit close to the KANK1 binding site, with R1368W sitting on an exposed helix close to the binding site which has the potential to disrupt the talin-KANK interaction directly, whereas Y1389C is buried in the core of the domain but appears to affect the mechanical stability of the domain. To explore whether this interaction with KANK is affected, fluorescence polarisation assays were used with 50 μ M of each mutant construct against KANK1 LD peptide encompassing residues 30-68. Additionally, with the stability of the R8 domain being so dependent on R7 mechanical stability, we tested R8 binding to the RIAM LD peptide 4-30 to determine whether there were any knock-on effects on the domain’s interaction with LD containing ligands which all interact with the same binding surface.

Surprisingly, neither KANK1 nor RIAM showed any significant differences in binding affinity to the R7R8 domain carrying each of the mutations when compared to wild-type

protein (Figure 6.10A). Wild-type interacted with RIAM with a K_d of $3.6 \mu\text{M}$ (± 0.19) while R1368W had a K_d of $3.3 \mu\text{M}$ (± 0.22) and Y1389C also with $3.6 \mu\text{M}$ (± 0.17). This shows that the structural integrity of the R8 domain remains unaffected and the reduced stability of the R7 Y1389C mutation has no consequential effects on R8 interaction with LD motif containing ligands. With the KANK1 interaction, wild-type R7R8 interacted with a K_d of $0.35 \mu\text{M}$ (± 0.02) while R1368W had a K_d of $0.39 \mu\text{M}$ (± 0.02) and Y1389C also with a K_d of $0.35 \mu\text{M}$ (± 0.04) (Figure 6.10B). Again, this shows that even with the destabilising mutant Y1389C, the domain retains enough structural integrity to sustain interaction with KANK with no apparent differences in affinity. The interaction between talin and KANK1 promotes organisation and assembly of cortical microtubule stabilising complexes at the periphery of focal adhesions, and thus organises microtubules around the adhesion site (Bouchet *et al.* 2016). The Hytönen lab who showed that neither of the mutations had a drastic impact on the architecture of the microtubule cytoskeleton (Figure 6.10C) suggesting that the KANK1 interaction may be unperturbed in cells, however, better images are required to assess microtubule targeting to the periphery of focal adhesions to conclude that these mutations have no effect on the talin-KANK1 interaction.

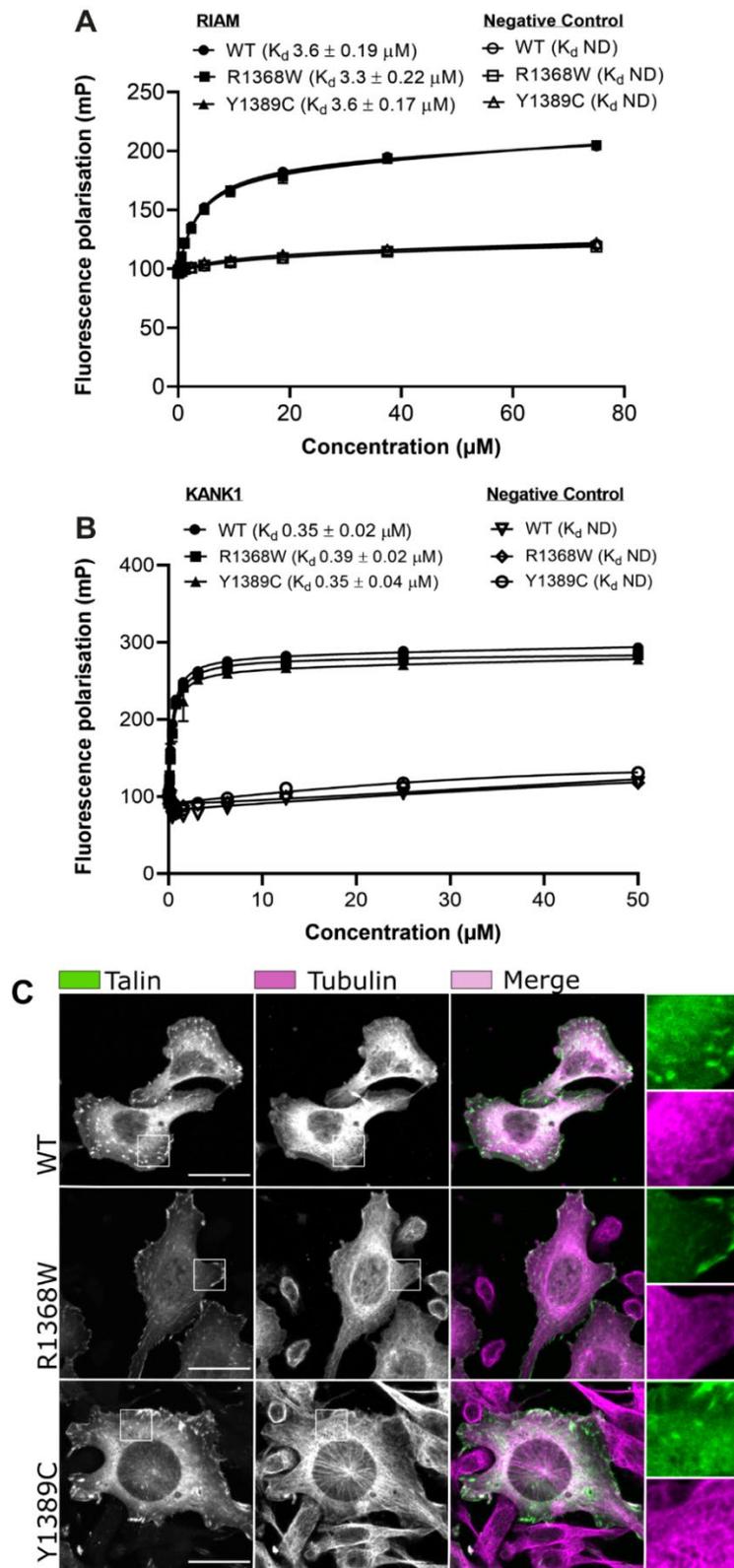


Figure 6.10. Mutations R1368W and Y1389C have no effect on ligand interaction. Cell imaging by Latifeh Azizi

(A) Fluorescence polarisation assay comparing RIAM binding with 78 μ M wild-type R7R8 (WT), Y1389C and R1368W. Calculated K_d and standard error are indicated in the legend. $N=3$ **(B)** Fluorescence polarisation comparing KANK1 binding with of 50 μ M R7R8, R1368W and Y1389C. $N=3$. **(C)** Confocal immunofluorescence showing tubulin organisation in WT vs R1368W and Y1389C.

6.4.6. C-terminal L2509P mutation may impair talin dimerisation

One of the most striking phenotypes observed in the studied cells was caused by the introduction of the L2509P mutation which resides at the centre of the talin dimerisation domain. The cells exhibited a loss of defined adhesion complex formation, disruption to the architecture of the actin cytoskeleton and had significantly reduced rates of migration in both 2D and 3D environments. It is highly likely that the introduction of a proline is causing either a bend in the helical structure of the dimerisation domain or completely breaking the helix, and thus preventing the two dimerisation domains interacting normally.

To test whether the mutation is disrupting talin dimerisation, size-exclusion chromatography (SEC) was used to visualise the oligomeric state of the protein and coupled with multi angle light scattering (MALS) for the calculation of molecular weight of the resulting peak. For this experiment, a construct containing talin R13 and the dimerisation domain (residues 2300-2541) was used (R13-DD) and L2509P introduced using mutagenesis. Wild-type R13-DD revealed a single peak showing a constitutive dimer which is consistent with previously published research (Gingras *et al.* 2008), and this was confirmed with the molecular weight calculated as 53.13 kDa. Comparison with the L2509P mutant revealed a drastic shift to the right in peak position with a calculated molecule weight of 26.57 kDa, half the weight of wild-type, confirming the transition to a monomer

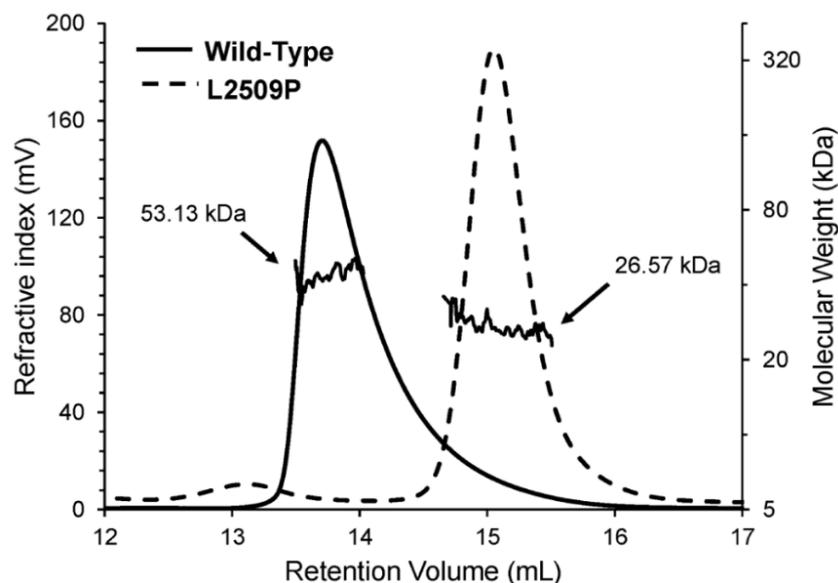


Figure 6.11. Talin L2509P mutation stops talin dimerisation

SEC-MALS analysis of wild-type R13-DD (53.13 kDa) and L2509P (26.57 kDa) confirming disruption of talin dimerisation.

6.4.7. L2509P leads to minor perturbation of ABS3 actin binding

Previous research has demonstrated how talin dimerisation is required for ABS3 mediated interaction with F-actin. With the knowledge that the L2509P mutation completely disrupts dimerisation, we also tested whether there were any differences in the actin binding properties of this domain *in vitro*. A range of different concentrations of wild-type R13-DD and L2509P from 5, 10 and 20 μM was incubated with 20 μM F-actin. Using a high-speed spin, the F-actin was pelleted with pellet and supernatant fractions ran on SDS-PAGE to determine to percentage of protein in the pellets with the F-actin. Consistent with expectations, the L2509P mutant appeared to reduce the amount of protein which pelleted with the F-actin, however, this was only a small reduction from 6-12% suggesting that binding may have been partially impaired compared to wild-type R13-DD but not so significantly that talin can no longer interact with actin (Figure 6.12).

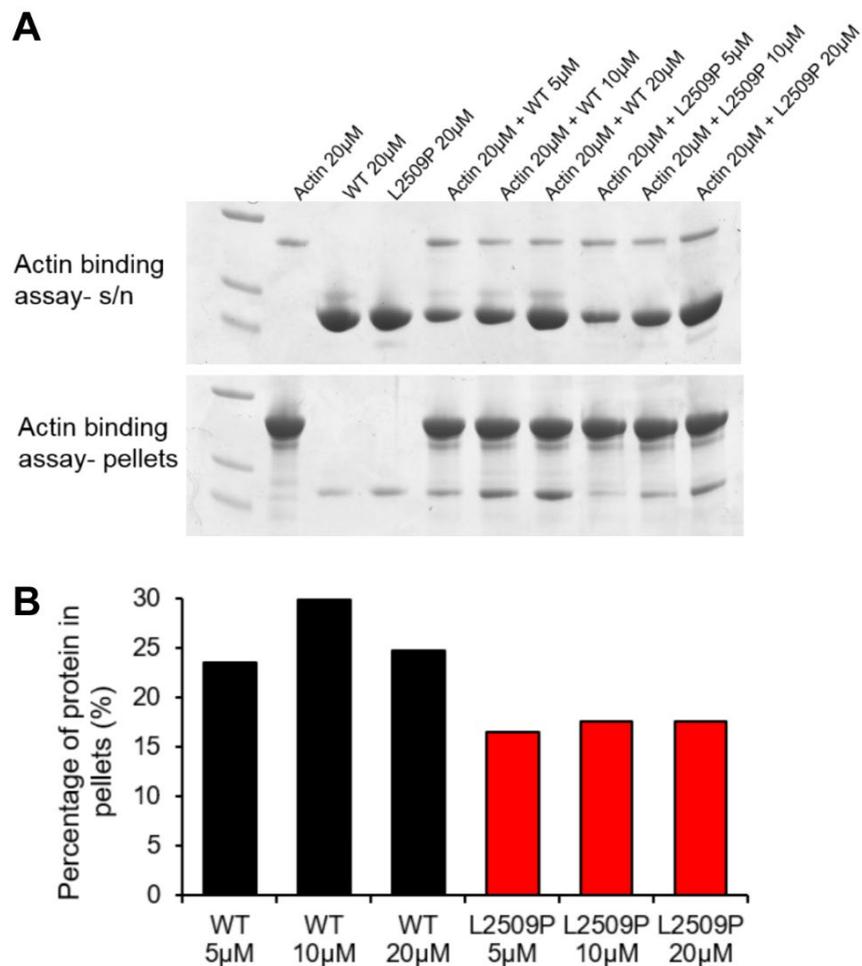


Figure 6.12. L2509P reduces actin binding

(A) Actin high-speed cosedimentation assay comparing interaction with 5, 10 and 20 μM of protein against 20 μM of F-actin. **(B)** Percentage of protein in pellets calculated from band density analysis in ImageJ. N=1

6.5. Discussion

Talin is a core mechanosensitive regulator of adhesion complex formation and is crucial for maintaining cell shape and migration. The series of 4 and 5-helix bundles along the talin rod unfold at differing rates in response to specific forces, with previous research demonstrating how a small change from a single residue mutation can lead to changes in force transmission and subsequently altering adhesion protein recruitment and sensing of the ECM (Rahikainen *et al.* 2017). The collaborative research presented here demonstrates how cancer associated single point mutations can alter talin properties and subsequently alter cellular behaviour in both 2D and 3D environments. Summary table of observed effects is included below (Table 8).

Mutant	Biochemistry	Cell morphology	Interactions (protein-protein* or within adhesions)	Division 0.2% FBS	Division 10% FBS	Migration speed ($\mu\text{m}/\text{min}$)	Invasion
P229L	-	•	vinculin↓ paxillin↓ pFAK•	↑	•	0.64 ±0.19	-
I392N	could not be produced	•	vinculin↓ paxillin↔ pFAK• β 1 active↓	↓	•	0.82 ±0.24	↑
V577D	-	•	vinculin↔ paxillin↓ pFAK•	•	•	0.63 ±0.16	-
A893E	-	•	vinculin• paxillin• pFAK•	•	•	0.68 ±0.22	-
R1368W	NMR • CD •	•	vinculin↓ paxillin↓ pFAK• Vd1• KANK• RIAM• tubulin•	•	•	0.54 ±0.23	↑
Y1389C	- CD↓	•	vinculin↓ paxillin↓ pFAK• Vd1↑ KANK• RIAM• tubulin•	↓	•	0.64 ±0.21	•
L1539P	could not be produced	•	vinculin• paxillin• pFAK• Vd1• DLC-1↓ tubulin•	•	•	0.57 ±0.15	•
S1750F	-	•	vinculin• paxillin• pFAK•	•	•	0.62 ±0.16	-
E1770Q	-	•	vinculin↓ paxillin• pFAK↑	•	•	0.61 ±0.19	-
D2086V	-	•	vinculin• paxillin• pFAK•	•	•	0.64 ±0.24	-
L2509P	SEC (monomer)	non-polarised	vinculin• paxillin• pFAK↓	↓	•	0.37 ±0.20	•

Table 8. Summary of observations for talin-1 mutations

Summary table showing effects observed across various experiments with each talin mutation. • = no change compared to wild-type talin, - = change not tested, increases or decreases relative to wild-type indicated by arrows. *Vinculin/VD1, KANK and RIAM assessed biochemically.

Cell based experiments from the Hytönen lab revealed that the I392N mutation exhibited the greatest enhancing effect on cellular migration in 2D and invasion through a 3D Matrigel. This mutation sits in the F3 domain close to the talin integrin binding site and resides in a region of talin which has the potential to interact with membrane lipids (Isenberg *et al.* 2002). The buried location of the residue could be changing domain stability, which is affecting association with integrins, or promoting dissociation from the membrane.

The R7R8 region of talin has a complex network of interactions and holds the R8 domain in a unique topology which allows protection from mechanical stretching of the talin rod. Mutations located in this region of talin have the potential to affect different aspects of the domains dynamic interactions from directly altering ligand binding sites or modifying domain stability which changes how and when vinculin would interact with the unfolded domain. Here we demonstrated how single cancer-associated point mutations in this region of talin can change these dynamics, translating into subtle changes in cellular behaviour which are likely to promote cancer cell migration and invasiveness. The mutation R1368W in the R7 domain is exposed on the surface and lies close to the KANK1 binding site. It was found that this mutation increased both random migration in a 2D environment and directional migration towards a chemoattractant in a 3D matrix. This mutation also appeared to drastically reduce recruitment of key focal adhesion components vinculin, active FAK and paxillin. What was most surprising is that this mutation had no impact on binding to the KANK1 LD region when compared to wild-type in biochemical assays. In mammalian cells, KANK1 interaction with talin promotes cortical microtubule stabilising complex formation at the periphery of focal adhesions, and shear force experiments have demonstrated that this interaction is capable of withstanding high mechanical forces (Bouchet *et al.* 2016; Yu *et al.* 2019). While talin is still capable of interacting with the KANK1 LD motif, more work is required to determine whether the interaction has the same mechanical stability as wild-type. It was shown that the microtubule architecture in cells appears undisrupted through introduction of these mutations, however, better imaging is required to determine microtubule organisation around the periphery of focal adhesions and check for potential changes in KANK1 localisation. Collectively these results suggest that the mutation is causing changes to an, as of yet, unidentified signalling function or ligand interaction which is drastically altering adhesion dynamics at the cellular level. The reduction in visible adhesion component recruitment could be a result of impaired signalling required for their recruitment or a result of increased focal adhesion turnover and disassembly. The

increased turnover of adhesion complexes is likely the most plausible explanation considering the increased migration speeds of cells containing this mutant. Future work would benefit from measuring the rates of adhesion turnover using live cell imaging as the cells migrate to explore whether this is the case.

One of the more interesting findings from this project was from the Y1389C mutant which lies buried in the core of the R7 domain. The impact of this mutation appeared subtle in the cells, slightly increasing cell invasiveness and decreasing association with FAK pY397. However, analysis of vinculin binding to R7R8 carrying the mutation revealed an enhancement in binding to the vinculin VD1 domain. Without force, the vinculin binding site (VBS) of talin R7 remains buried and unavailable for interaction, whereas the R8 domain can bind a single VD1 molecule in the absence of force with high affinity (Gingras *et al.* 2010). Size-exclusion chromatography analysis of vinculin VD1 binding with the R7R8 domain containing Y1389C revealed that the mutant was able to interact with two vinculin VD1 molecules whereas wild-type was only able to bind a single VD1. This result suggests that the location of the Y1389C mutation is leading to destabilisation of the R7 helical bundle, increasing the availability of the R7 VBS in the absence of force. This is further supported by CD thermostability analysis which revealed a reduction in the protein's overall stability. Domain destabilisation by the Y1389C mutation in R7 also has the potential to have a subsequent knock-on effect on the stability of the R8 domain as R8 stability is highly dependent on R7's folded state. While it was shown that R8 was still able to interact with the LD containing ligand RIAM the same as wild-type, the potential impact in actin binding was not explored. R8 and to a lesser extent R7, contribute to actin binding as part of the second actin binding site ABS2 in connection with the R4 domain (Atherton *et al.* 2015). Additionally, vinculin is required to promote actin association and stabilise the interaction. It would be interesting to explore whether the stability changes and altered vinculin recruitment observed here lead to changes in focal adhesion dynamics in association with actin recruitment and maturation into stable complexes.

Of the 11 mutations studied, the most striking phenotype was observed with the L2509P mutation which resides in the single helix dimerisation domain of talin. These cells lacked any form of polarisation, which is reflected in the almost complete loss in random migration in a 2D environment. The cells appeared unable to spread on the 2D surface, taking on a smaller more rounded phenotype and exhibited a loss in the ability to mature adhesion complexes for stable adhesion to the ECM. This same phenotype has previously been observed in talin deletion constructs, showing that deletion of the R13-

DD leads to drastic loss in cell polarisation and migration (Rahikainen *et al.* 2019). SEC-MALS analysis of wild-type R13-DD shows a constitutive dimer consistent with previous observations, whereas introduction of the L2509P mutation completely stopped talin dimerisation, with the domain becoming monomeric. This also appeared to partially impact the domain's ability to interact with actin filaments in an actin cosedimentation assay, suggesting that the loss of dimerisation may impair actin binding through ABS3, which prevents force transmission by the talin rod and inhibits adhesion maturation. More repeats are however required to confirm the reduction in actin binding presented and determine whether the observed phenotype is a result of impaired actin binding, through other knock on effects due to loss of dimerisation or some other unknown function for this region of talin. Additionally, the potential impact of this mutant needs to be assessed in full length talin in order to determine whether this dimerisation loss is relevant to the full protein. This is especially important in light of new research which raises questions about talin's dimeric behaviour in its autoinhibited form where it has been shown that autoinhibited full length talin is in fact monomeric (Dedden *et al.*, 2019).

Overall, this study selected 11 talin cancer-associated mutations for analysis from a screen of hundreds and showed a variety of effects in cellular behaviour and protein interaction dynamics. Two of the mutants, I392N and R1368W showed enhanced directional migration in a 3D context with I392N also showing increased migration speed in 2D. One mutant destabilised the R7 helical bundle, promoting vinculin binding in the absence of force, and another resulted in a loss of talin dimerisation, with a drastic impact on cell morphology. An increasing number of mutations are being added to publicly available online databases such as COSMIC, and with the potential impact each point mutation has on talin activity and subsequently cell migration, more attention should be paid to these subtle changes which can lead to big effects in cancer cell migration. The pipeline created here demonstrates that to an extent we can predict potential functional impacts of large numbers of cancer associated mutations by rapid screening and use this to determine which mutations may require further detailed analysis.

Chapter 7. Conclusions and future directions

7.1. TLNRD1 as a third talin gene

The aim of this thesis was to explore the relationship between the highly conserved TLNRD1 protein and the adhesion adapter protein talin, a protein essential for mammalian development and coordinating adhesion complexes. This was achieved using a range of structural and biochemical techniques to explore the structural relationship between TLNRD1 and the R7R8 region of talin and compare the proteins behaviour in terms of interaction with talin associated ligands. This intriguing project reveals the distinct similarities conserved between talin R7R8 and TLNRD1, yet also demonstrates how TLNRD1 has evolved a predominant actin bundling function through dimerisation of the 4-helix domain which is the equivalent of talin R8.

A number of novel ligands were identified for TLNRD1 which bind using the same LD-motif recognition system as talin. This included the KANK proteins which interact with talin R7 to connect adhesion complexes to the microtubule cytoskeleton, the essential cell cycle regulator CDK1 which is a newly identified ligand for R8, and finally RIAM, which promotes talin recruitment to the leading edge of the cell. Using binding affinity comparisons between TLNRD1 full-length protein, the 4-helix module alone, and a TLNRD1 4-helix domain mutant (2E), it was found that TLNRD1 interactions with the LD motif containing ligands have been conserved in the same positions after the protein's divergence from talin. These binding partners are known to regulate numerous different aspects of cell behaviour from the cell cycle to adhesion disassembly, and open up the possibility that TLNRD1 is behaving as a mobile equivalent of the R7R8 signalling hub coordinating similar pathways outside adhesion complexes or acting as a dominant negative regulator of R7R8 interactions (Figure 7.1).

There are many aspects with which future research can be directed, with many questions left unanswered. For example it would be good to assess competition between ligands which share the same binding site using full-length proteins rather than just peptides, and explore whether any of these LD motif containing ligands alter TLNRD1 association with actin filaments using cosedimentation assays and electron microscopy. To understand the biological relevance of these interactions work is required to understand how these are coordinated in a cellular context and determine whether TLNRD1 expression alters adhesion dynamics or associated signalling responses. One of the most intriguing interactions identified is with the RIAM and Lamellipodin LD motifs. Previous research has demonstrated that talin can form a complex with RIAM/Lamellipodin and integrin, called the MIT complex (MRL protein-integrin-talin) which localises at the tips of

filopodial protrusions (Lagarrigue *et al.* 2015). It would be interesting to explore whether the presence of TLNRD1 at the filopodial tips alters the behaviour of this complex or modulates the activation state of integrins at the filopodial tips. Further to this, the CDK1 LD motif interaction identified leads to questions of whether TLNRD1 behaviour is regulated in a cell cycle dependent manner. Does CDK1 actually phosphorylate TLNRD1 or is the interaction non-reactive? Phosphorylation detection assays with CDK1 and different cyclins would help answer this. It should also be noted that the region of amino acids containing the LD motif appears to be conserved across different members of the CDK family, raising the possibility of interaction with different CDK's with some perhaps interacting non-specifically or some interacting in a manner which leads to TLNRD1 phosphorylation and changes in protein behaviour. With this high sequence similarity in the LD region between different CDK's, it would be better to test potential interactions in the context of full-length CDK proteins rather than just the peptide.

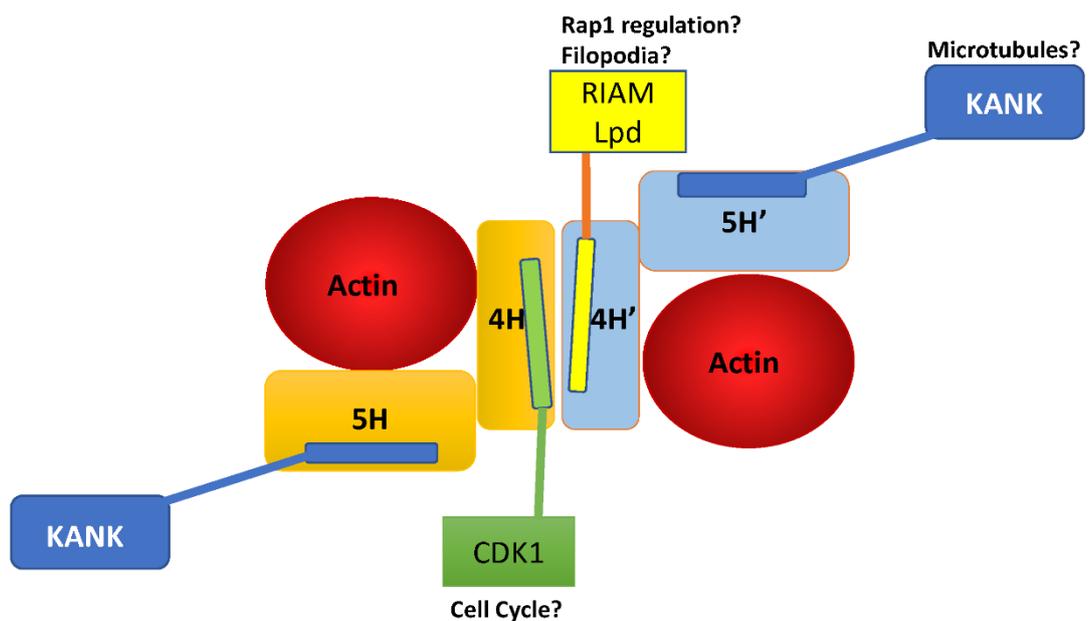


Figure 7.1. Model of TLNRD1 interactions

Schematic representation of TLNRD1 sites of interaction with identified ligands and their associated behaviours. KANK interaction may provide connection to microtubules, RIAM and lamellipodin are involved in Rap1 regulation and filopodia initiation, and CDK1 is connected to the cell cycle.

One of the more surprising findings of this study was the discovery that TLNRD1 is actually an actin bundling protein, capable of creating incredibly tight bundles of actin filaments *in vitro*. Talin R7R8 in connection with the talin R4 domain, forms the second actin binding site within talin which is required for talin mediated force transmission (Kumar *et al.* 2016). The R7R8 region on its own however shows relatively weak binding

to actin filaments without the support of the R4 domain, and neither R7 or R8 alone can interact with filaments. What is surprising about TLNRD1 when compared to R7R8 is the natural modifications introduced which enhance actin filament binding affinity, allowing the protein to hold filaments in close proximity via the 4-helix domain alone. Compared to R7R8, TLNRD1 has adapted a more defined area of basic residues which mediate the stronger interaction, and this basic region lies in a different position on TLNRD1 compared to the known actin binding regions of R7R8. This enhanced affinity for filament interaction likely provides the strength required to bring filaments so close together and suggests that the interaction may be mechanically stable under high cellular forces. A model of how TLNRD1 may be bundling actin filaments based on the clues given throughout this study is shown below (Figure 7.2). It is clear that the 4-helix domain of TLNRD1 is capable of forming actin bundles on its own, suggesting that this contains the core actin binding sites required for bundling, with the 5-helix domain stabilising the bundled structure. The antiparallel nature of TLNRD1 dimerisation and apparent decoration of actin stress fibres in cellular studies suggests that TLNRD1 has a preference for binding to or creating antiparallel actin bundles. There is also the potential that the N-termini of TLNRD1 can interact to form tetramers which could explain how such wide bundles were observed through electron microscopy. Again, improved imaging is required here to determine TLNRD1 distribution along the bundles, spacing between filaments and combining this with markers to confirm polarity of actin filaments. Filopodial protrusions require tightly bundled actin filaments to provide stability and rigidity, and filopodia formation is closely connected to increased migration in cancerous cells. TLNRD1 localisation to the tips of filopodia, ability to increase the number of filopodia being produced and its overall effect on increasing cell migration, suggests that these processes are closely connected and may explain why overexpression of TLNRD1 reduces patient survival with some cancers. The tight packing of filaments is characteristic of the actin bundling protein fascin, which is required for filopodia formation (Vignjevic *et al*, 2006; Zanet *et al*, 2012). It could be proposed that TLNRD1 bundling of actin filaments towards the filopodial tips helps to stabilise them as they form, evidenced by the increase in observed filopodial numbers and increased migration speeds. Alternatively, TLNRD1 role on filopodial tips may be entirely independent of bundling activity and this actin bundling activity may be required for other processes such as stabilisation of stress fibres.

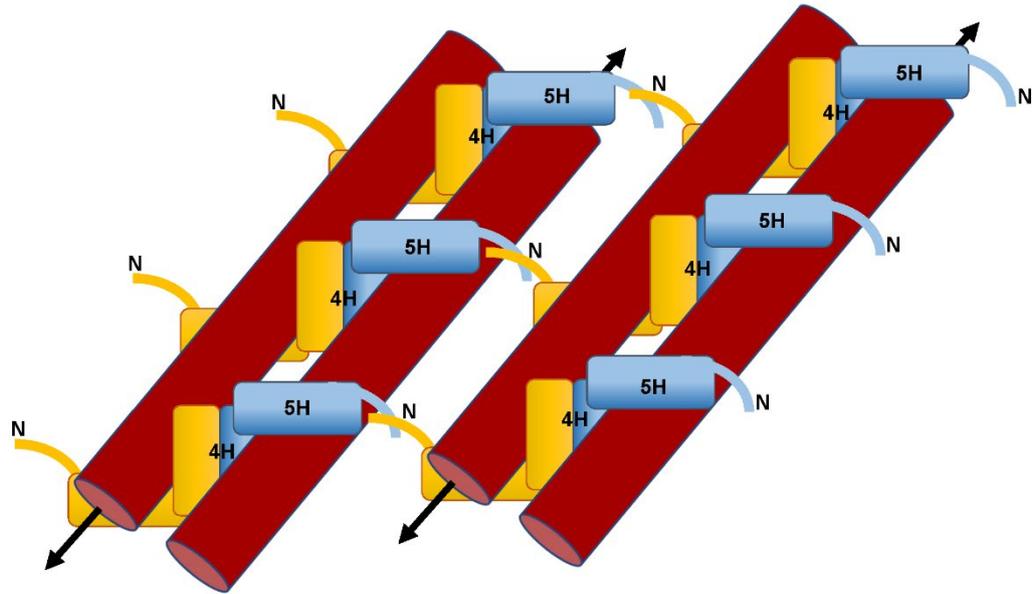


Figure 7.2. Schematic representation of TLNRD1 actin bundling

Diagram representing how TLNRD1 may be mediating actin bundle formation, taking into account the potential for TLNRD1 to extend into an open conformation and for the N-terminus to promote tetramer formation. Domains are indicated with N indicating N-terminal unstructured region. Arrows indicate actin filament polarity.

Finally, early investigations in zebrafish models using both morpholino and CRISPR knockout studies to directly target TLNRD1 expression reveals how TLNRD1 is essential for normal development, with TLNRD1 knockout leading to drastic defects in the development of vascular structures, tail morphology and fin-fold development. While previous research demonstrated that TLNRD1 is not required for mesoderm development, the exciting new evidence presented here reveals how it is required for mesoderm derived processes and likely plays a crucial role during development. This could have implications for the study of specific diseases related to vessel development or developmental defects but as of yet, TLNRD1 has not been identified as a disease-causing gene except in the case of malignant neoplasms (Tatarano *et al*, 2011; Wu *et al*, 2017). Work now needs to continue breaking down the functional importance of TLNRD1 in development, starting with understanding whether there are any specific cell types that require higher TLNRD1 expression, or whether TLNRD1 expression is regulated in a manner that is dependent on the different stages of development. Using transgenic labelling of TLNRD1 in zebrafish will help guide understanding of these key outstanding questions. For example, angiogenesis, the sprouting of new vessels from existing vessels, is a complex process dependent on endothelial cell migration and supported by

the formation of filopodial extensions in 'tip' cells can easily be visualised in developing embryos (Betz *et al*, 2016). With the increase in filopodial extensions observed with TLNRD1 expression in cells, it would be important to determine whether TLNRD1 is involved in promoting filopodial extension in this context by observing TLNRD1 localisation during vessel sprouting and formation of adhesive connections between vessels.

7.2. Limitations and future directions

From a structural perspective, there is still a lot to learn about TLNRD1. First, the crystal structure along with the solution SAXS envelope suggests that TLNRD1 exists in a compact closed conformation, which is in stark contrast to talin R7R8, the domains of which have always appeared extended. This leads us to question whether this compact structure is a process of autoinhibition in relation to the proteins function. It would be interesting to explore whether this is the case, and if so what process leads to protein extension and what would be the purpose of this regulatory feature. Further work from a cellular perspective to identify and characterise TLNRD1 ligands would help guide the exploration of this aspect of TLNRD1 function and coupling with SAXS to investigate which ligands (if any) lead to relief of autoinhibition. Secondly, the unstructured region at the N-terminus of TLNRD1 is a unique addition that is absent in talin R7R8. It is likely that this region plays a role in regulating TLNRD1 function or localisation, as it is predicted to contain multiple phosphorylation sites. The functional significance of this region was not explored in this project, but it would be important to study this further by generating N-terminal deletion constructs or selected mutagenesis of predicted phosphorylation sites to explore behaviour both in the cell and at a structural level. Additionally, it needs to be determined whether this region enables TLNRD1 to form tetramers, which may be of functional importance when it comes to understanding the proteins actin bundling activity.

The majority of the limitations in study relate to a lack of informative cell biology. First, while this project has identified numerous TLNRD1 ligands, the question of when and where these interactions occur in the cell has not been explored. This research would be especially important as thus far all LD motif interaction studies have just used peptides and not full-length proteins. Furthermore, of all the confocal microscopy imaging performed so far, there is little observed actin binding occurring in the cell which is unexpected considering how strongly TLNRD1 bundles filaments *in vitro*. One question which remains is how TLNRD1's interaction with LD motif containing ligands changes the

dynamics of actin filament binding, and whether TLNRD1 holds interaction with other ligands while bound to actin. Proteomics should be utilised here to determine exactly what proteins TLNRD1 is interacting with and identify new ligands which would aid functional understanding.

Multiple approaches can be taken to explore TLNRD1 function further. First, the evidence that the 4-helix domain of TLNRD1 alone appears to show increased binding to filaments *in cellulo* compared to the full-length protein suggest some sort of regulation from the 5-helix domain or N-terminal unstructured region. This may in part be due to some yet unknown autoinhibitory system within TLNRD1, potentially mediated by the N-terminal unstructured region. Here, N-terminal deletion constructs may be useful to determine whether TLNRD1 actin binding is regulated by this region in cells. Second, all the imaging done so far has relied on N-terminal tagged GFP constructs which would be attached to the unstructured region of TLNRD1. To exclude the possibility that this is affecting TLNRD1 behaviour in a cellular environment, C-terminal GFP tagged constructs should also be tested.

Finally, all the experiments so far have been conducted in U2OS cells to explore the connection between increased TLNRD1 expression and cancer cell migration, however, the phenotype presented by zebrafish knock-out studies suggest that TLNRD1 plays a key role in development with defects observed in the vasculature. Therefore, to understand whether TLNRD1 may be required for vascular development directly or indirectly, it will be important to explore how the protein functions in a vascular derived endothelial cell line. Additionally, research has unveiled how filopodial extensions are important for guiding cell migration during tissue patterning and vessel branching, with evidence emerging that there may be different types of filopodia required for these processes (De Smet *et al.* 2009; Sanders, Llagostera and Barna 2013). The key to understanding how TLNRD1 contributes to vascular development lies in taking a more developmental approach to the cell biology, taking into account the growth factors which regulate these processes such as VEGF and the role filopodia play in allowing cells to sense their environment and increase migration during mammalian development.

7.1. Pipeline development for assessment of cancer associated mutations

As an extension to the work presented here on TLNRD1, we tested a new bioinformatic pipeline developed by the Hytönen lab for the identification of somatic mutations which may alter the function of adhesion related proteins, using talin as a platform to explore the accuracy of the pipeline. A total of 11 mutations were identified which had the potential to alter talin function (Figure 7.3). Each individual mutation was assessed for their impact on cell morphology, migration, invasion and adhesion complex formation.

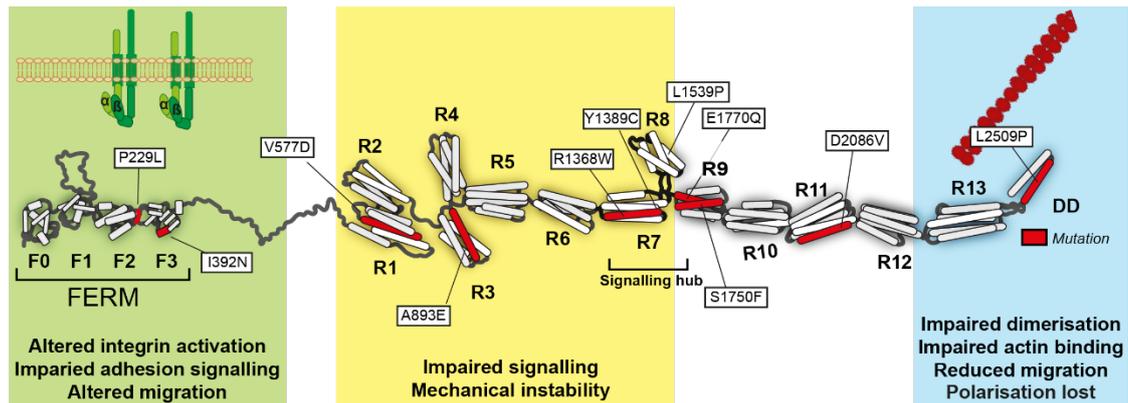


Figure 7.3. Schematic summary of mutation impact on talin functions

Schematic of talin structure with helices containing tested mutations highlighted in red. Distinct regions are highlighted in different colours and associated mutation impacts on talin behaviour are indicated.

The majority of mutations tested had moderate to minor impacts on all of these factors, suggesting that the mutations are causing subtle changes in signalling responses. Different behaviours were observed with each mutation and this appeared largely dependent on where along the talin rod structure each mutation was located (Figure 7.3). With talin behaviour being mechanically sensitive, it is possible that some of the subtle changes observed may be due to changes in mechanical unfolding of talin rather than directly impaired ligand interaction.

The pipeline used in this study was able to successfully identify mutants which subtly change signalling responses but did not aid understanding of exactly how some of these mutants may be contributing to disease. Some of the mutations identified using this pipeline actually appeared detrimental to the cells, for example L2509P transfected cells showed a complete loss in polarisation. On the other hand, some were able to promote cellular migration and invasion such as the I392N mutant which resides close to the integrin binding site of talin. In order to really understand how these mutations contribute to cancer, further research is required in the context of a 3D environment or solid

tumours. Cells exposed to different environmental factors will respond and behave differently than they would in a 2D cell culture setting and this is important in determining the true impact each of these mutations have. It is clear from this study that mutations in talin can alter cell behaviour and it is thus important to understand how they are doing this, but more research is required to really pick apart the interaction dynamics of talin and the altered responses observed.

Chapter 8. References

- Afonine, P. V. *et al.* (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D Biological Crystallography* [Online] **68**:352–367. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444912001308>.
- Aghababazadeh, Masoomah & Kerachian, Mohammad Amin. (2014). Cell fasting: Cellular response and application of serum starvation. *J. Fasting Health*. **2**:147-150. 10.22038/jfh.2014.3928.
- Alam, T. *et al.* (2014). How to find a leucine in a haystack? Structure, ligand recognition and regulation of leucine–aspartic acid (LD) motifs. *Biochemical Journal* [Online] **460**:317–329. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24870021>.
- Arcario, M.J. and Tajkhorshid, E. (2014). Membrane-Induced Structural Rearrangement and Identification of a Novel Membrane Anchor in Talin F2F3. *Biophysical Journal* [Online] **107**:2059–2069. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0006349514009965> [Accessed: 4 November 2019].
- Arjonen, A., Kaukonen, R. and Ivaska, J. (2011). Filopodia and adhesion in cancer cell motility. *Cell Adhesion & Migration* [Online] **5**:421–430. Available at: <http://www.tandfonline.com/doi/abs/10.4161/cam.5.5.17723> [Accessed: 18 October 2019].
- Ashkenazy, H. *et al.* (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* [Online] **44**:W344–W350. Available at: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw408>.
- Astro, V. *et al.* (2014). Liprin- α 1, ERC1 and LL5 define polarized and dynamic structures that are implicated in cell migration. *Journal of Cell Science* **127**:3862–3876.
- Atherton, P. *et al.* (2016). Mechanosensitive components of integrin adhesions: Role of vinculin. *Experimental Cell Research* **343**:21–27.
- Atherton, P. *et al.* (2015). Vinculin controls talin engagement with the actomyosin machinery. *Nature Communications* [Online] **6**:10038. Available at: www.nature.com/naturecommunications [Accessed: 24 October 2019].

- Van Audenhove, I. *et al.* (2016). Fascin Rigidity and L-plastin Flexibility Cooperate in Cancer Cell Invadopodia and Filopodia. *Journal of Biological Chemistry* [Online] **291**:9148–9160. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M115.706937> [Accessed: 18 October 2019].
- Baade, T. *et al.* (2019). Clustering of integrin β cytoplasmic domains triggers nascent adhesion formation and reveals a protozoan origin of the integrin-talin interaction. *Scientific Reports* [Online] **9**:5728. Available at: <http://www.nature.com/articles/s41598-019-42002-6> [Accessed: 5 October 2019].
- Baertsch, R. *et al.* (2008). Retrocopy contributions to the evolution of the human genome. *BMC Genomics* [Online] **9**:466. Available at: <http://bmcbgenomics.biomedcentral.com/articles/10.1186/1471-2164-9-466> [Accessed: 21 October 2019].
- Bear, J.E. and Gertler, F.B. (2009). Ena/VASP: towards resolving a pointed controversy at the barbed end. *Journal of Cell Science* [Online] **122**:1947–1953. Available at: <http://jcs.biologists.org/cgi/doi/10.1242/jcs.038125> [Accessed: 18 October 2019].
- Betz, C. *et al.* (2016). Cell behaviours and dynamics during angiogenesis. *Development JCS*, **143**:2249–2260. doi: 10.1242/dev.135616
- Bendas, G. and Borsig, L. (2012). Cancer Cell Adhesion and Metastasis: Selectins, Integrins, and the Inhibitory Potential of Heparins. *International Journal of Cell Biology* [Online] **2012**:1–10. Available at: <http://www.hindawi.com/journals/ijcb/2012/676731/> [Accessed: 13 October 2019].
- Blanchoin, L. *et al.* (2014). Actin Dynamics, Architecture, and Mechanics in Cell Motility. *Physiological Reviews* [Online] **94**:235–263. Available at: <http://www.physiology.org/doi/10.1152/physrev.00018.2013> [Accessed: 6 September 2018].
- Bouchet, B.P. *et al.* (2016). Talin-KANK1 interaction controls the recruitment of cortical microtubule stabilizing complexes to focal adhesions. *eLife* [Online] **5**. Available at: <https://elifesciences.org/articles/18124>.
- Brangwynne, C.P. *et al.* (2006). Microtubules can bear enhanced compressive loads in living cells because of lateral reinforcement. *The Journal of Cell Biology* [Online] **173**:733–741. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.200601060>

[Accessed: 13 October 2019].

Brasch, J. *et al.* (2012). Thinking outside the cell: how cadherins drive adhesion. *Trends in Cell Biology* [Online] **22**:299–310. Available at: <https://www-sciencedirect-com.chain.kent.ac.uk/science/article/pii/S0962892412000529> [Accessed: 13 September 2019].

Breitsprecher, D. and Goode, B.L. (2013). Formins at a glance. *Journal of Cell Science* [Online] **126**:1–7. Available at: <http://jcs.biologists.org/lookup/doi/10.1242/jcs.107250> [Accessed: 18 October 2019].

Brown, M.C., Perrotta, J. and Turner, C.E. (1996). Identification of LIM3 as the principal determinant of paxillin focal adhesion localization and characterization of a novel motif on paxillin directing vinculin and focal adhesion kinase binding. *The Journal of Cell Biology* [Online] **135**:1109–1123. Available at: <http://www.jcb.org/cgi/doi/10.1083/jcb.135.4.1109> [Accessed: 15 October 2019].

Brown, N.R. *et al.* (2015). CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nature Communications* [Online] **6**:6769. Available at: <http://www.nature.com/doi/10.1038/ncomms7769>.

Brunet, T. and King, N. (2017). The Origin of Animal Multicellularity and Cell Differentiation. *Developmental Cell* [Online] **43**:124–140. Available at: <https://doi.org/10.1016/j.devcel.2017.09.016>.

Burridge, K. and Mangeat, P. (1984). An interaction between vinculin and talin. *Nature* **308**:744–746.

Carmona, G. *et al.* (2016). Lamellipodin promotes invasive 3D cancer cell migration via regulated interactions with Ena/VASP and SCAR/WAVE. *Oncogene* [Online] **35**:5155–5169. Available at: <http://www.nature.com/articles/onc201647> [Accessed: 13 July 2019].

Case, L.B. *et al.* (2015). Molecular mechanism of vinculin activation and nanoscale spatial organization in focal adhesions. *Nature Cell Biology* [Online] **17**:880–892. Available at: <http://www.nature.com/articles/ncb3180> [Accessed: 24 October 2019].

Cerami, E. *et al.* (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discovery*

- [Online] **2**:401–404. Available at: www.aacrjournals.org [Accessed: 27 October 2019].
- Chandra Roy, B., Kakinuma, N. and Kiyama, R. (2009). Kank attenuates actin remodeling by preventing interaction between IRSp53 and Rac1. *The Journal of Cell Biology* [Online] **184**:253–267. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19171758> [Accessed: 27 August 2019].
- Chang, Y.-C. *et al.* (2014). Structural and Mechanistic Insights into the Recruitment of Talin by RIAM in Integrin Signaling. *Structure* [Online] **22**:1810–1820. Available at: <https://www.sciencedirect.com/science/article/pii/S0969212614003335> [Accessed: 5 August 2019].
- Changede, R. *et al.* (2015). Nascent Integrin Adhesions Form on All Matrix Rigidities after Integrin Activation. *Developmental Cell* **35**:614–621.
- Chen, P. *et al.* (2017). Talin-1 interaction network promotes hepatocellular carcinoma progression. *Oncotarget* [Online] **8**:13003–13014. Available at: <http://www.oncotarget.com/fulltext/14674> [Accessed: 26 October 2019].
- Chen, T., Wang, K. and Tong, X. (2017). In vivo and in vitro inhibition of human gastric cancer progress by upregulating Kank1 gene. *Oncology Reports* [Online] **38**:1663–1669. Available at: <https://www.spandidos-publications.com/10.3892/or.2017.5823> [Accessed: 27 August 2019].
- Chen, V.B. *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* [Online] **66**:12–21. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444909042073>.
- Choi, C.K. *et al.* (2008). Actin and α -actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nature Cell Biology* [Online] **10**:1039–1050. Available at: <http://www.nature.com/articles/ncb1763> [Accessed: 18 September 2019].
- Ciobanasu, C. *et al.* (2018). Integrin-bound talin head inhibits actin filament barbed-end elongation. *Journal of Biological Chemistry* [Online] **293**:2586–2596. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M117.808204> [Accessed: 28 October 2019].

- Clohisey, S.M.R., Dzhindzhev, N.S. and Ohkura, H. (2014). Kank Is an EB1 Interacting Protein that Localises to Muscle-Tendon Attachment Sites in Drosophila White-Cooper, H. ed. *PLoS ONE* [Online] **9**:e106112. Available at: <https://dx.plos.org/10.1371/journal.pone.0106112> [Accessed: 27 August 2019].
- Debrand, E. *et al.* (2009). Talin 2 is a large and complex gene encoding multiple transcripts and protein isoforms. *FEBS Journal* [Online] **276**:1610–1628. Available at: <http://doi.wiley.com/10.1111/j.1742-4658.2009.06893.x> [Accessed: 17 November 2019].
- Dechant, R. and Glotzer, M. (2003). Centrosome separation and central spindle assembly act in redundant pathways that regulate microtubule density and trigger cleavage furrow formation. *Developmental Cell* **4**:333–344.
- Dedden, *et al.* (2019) The Architecture of Talin1 Reveals an Autoinhibition Mechanism. *Cell*. **179**(1):P120-131. DOI:<https://doi.org/10.1016/j.cell.2019.08.034>
- Dimchev, G. *et al.* (2020). Lamellipodin tunes cell migration by stabilizing protrusions and promoting adhesion formation. *Journal of Cell Science* **7**: jcs239020.
- Ding, M. *et al.* (2003). C. elegans ankyrin repeat protein VAB-19 is a component of epidermal attachment structures and is essential for epidermal morphogenesis. *Development* [Online] **130**:5791–5801. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9570773> [Accessed: 27 August 2019].
- Dooley, K. and Zon, L. (2000). Zebrafish: a model system for the study of human disease. *Current Opinion in Genetics & Development* [Online] **10**:252–256. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0959437X00000745> [Accessed: 20 November 2019].
- Doyle, A.D. and Yamada, K.M. (2016). Mechanosensing via cell-matrix adhesions in 3D microenvironments. *Experimental Cell Research* [Online] **343**:60–66. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0014482715301403>.
- Elliott, P.R. *et al.* (2010). The Structure of the Talin Head Reveals a Novel Extended Conformation of the FERM Domain. *Structure* [Online] **18**:1289–1299. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S096921261000300X>.
- Elosegui-Artola, A. *et al.* (2016). Mechanical regulation of a molecular clutch defines force transmission and transduction in response to matrix rigidity. *Nature Cell Biology*

- [Online] **18**:540–548. Available at: <http://www.nature.com/articles/ncb3336>.
- Emsley, P. *et al.* (2010). Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography* [Online] **66**:486–501. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444910007493>.
- Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D Biological Crystallography* [Online] **67**:282–292. Available at: <http://scripts.iucr.org/cgi-bin/paper?S090744491003982X> [Accessed: 29 September 2019].
- Evans, P.R. and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography* [Online] **69**:1204–1214. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444913000061> [Accessed: 29 September 2019].
- Fischer, R.S. *et al.* (2019). Filopodia and focal adhesions: An integrated system driving branching morphogenesis in neuronal pathfinding and angiogenesis. *Developmental Biology* [Online] **451**:86–95. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S001216061830424X> [Accessed: 1 December 2019].
- Gan, Z. *et al.* (2016). Vimentin Intermediate Filaments Template Microtubule Networks to Enhance Persistence in Cell Polarity and Directed Migration. *Cell Systems* [Online] **3**:252–263.e8. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2405471216302629> [Accessed: 14 October 2019].
- Gasteiger, E. *et al.* (2005). Protein Identification and Analysis Tools on the ExpASY Server. In: *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press, pp. 571–607. Available at: <http://link.springer.com/10.1385/1-59259-890-0:571> [Accessed: 21 August 2019].
- Gavet, O. and Pines, J. (2010). Activation of cyclin B1–Cdk1 synchronizes events in the nucleus and the cytoplasm at mitosis. *The Journal of Cell Biology* [Online] **189**:247–259. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.200909144> [Accessed: 5 September 2019].
- Gingras, A.R. *et al.* (2010). Central region of talin has a unique fold that binds vinculin and actin. *Journal of Biological Chemistry* [Online] **285**:29577–29587. Available at:

<http://www.jbc.org/lookup/doi/10.1074/jbc.M109.095455>.

Gingras, A.R. *et al.* (2008). The structure of the C-terminal actin-binding domain of talin. *The EMBO Journal* [Online] **27**:458–469. Available at: <http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7601965>.

Goedhart, J. (2019). PlotsOfDifferences - a web app for the quantitative comparison of unpaired data. *bioRxiv* [Online]. Available at: <https://www.biorxiv.org/content/10.1101/578575v1>.

Goreczny, G.J., Forsythe, I.J. and Turner, C.E. (2018). Hic-5 regulates fibrillar adhesion formation to control tumor extracellular matrix remodeling through interaction with tensin1. *Oncogene* **37**:1699–1713.

Gough, R.E. and Goult, B.T. (2018). The tale of two talins - two isoforms to fine-tune integrin signalling. *FEBS Letters* [Online] **592**:2108–2125. Available at: <http://doi.wiley.com/10.1002/1873-3468.13081> [Accessed: 22 March 2019].

Goult, B.T. *et al.* (2013)a. RIAM and Vinculin Binding to Talin Are Mutually Exclusive and Regulate Adhesion Assembly and Turnover. *Journal of Biological Chemistry* [Online] **288**:8238–8249. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M112.438119> [Accessed: 5 August 2019].

Goult, B.T. *et al.* (2013)b. Structural studies on full-length talin1 reveal a compact auto-inhibited dimer: Implications for talin activation. *Journal of Structural Biology* [Online] **184**:21–32. Available at: <https://www.sciencedirect.com/science/article/pii/S1047847713001469> [Accessed: 15 March 2019].

Goult, B.T. *et al.* (2009). The Structure of an Interdomain Complex That Regulates Talin Activity. *Journal of Biological Chemistry* [Online] **284**:15097–15106. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M900078200> [Accessed: 6 September 2018].

Goult, B.T., Yan, J. and Schwartz, M.A. (2018). Talin as a mechanosensitive signaling hub. *The Journal of Cell Biology* [Online] **217**:3776–3784. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/30254032> [Accessed: 11 October 2018].

Greenfield, N.J. (2007). Using circular dichroism spectra to estimate protein secondary

structure. *Nature Protocols* **1**:2876–2890.

- Grzybowska, E.A. (2012). Human intronless genes: Functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochemical and Biophysical Research Communications* [Online] **424**:1–6. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0006291X12011874> [Accessed: 21 October 2019].
- Gu, Y. and Zhang, M. (2018). Upregulation of the Kank1 gene inhibits human lung cancer progression in vitro and in vivo. *Oncology Reports* [Online] **40**:1243–1250. Available at: <http://www.spandidos-publications.com/10.3892/or.2018.6526> [Accessed: 27 August 2019].
- Gyórfy, B. *et al.* (2013). Online Survival Analysis Software to Assess the Prognostic Value of Biomarkers Using Transcriptomic Data in Non-Small-Cell Lung Cancer. Chellappan, S. P. ed. *PLoS ONE* [Online] **8**:e82241. Available at: <http://dx.plos.org/10.1371/journal.pone.0082241> [Accessed: 14 November 2019].
- Haage, A. *et al.* (2018). Talin Autoinhibition Regulates Cell-ECM Adhesion Dynamics and Wound Healing In Vivo. *Cell Reports* [Online] **25**:2401-2416.e5. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2211124718317194> [Accessed: 26 October 2019].
- Hamidi, H. and Ivaska, J. (2018). Every step of the way: integrins in cancer progression and metastasis. *Nature Reviews Cancer* [Online] **18**:533–548. Available at: <http://www.nature.com/articles/s41568-018-0038-z> [Accessed: 13 October 2019].
- Hansen, S.D. and Mullins, R.D. (2015). Lamellipodin promotes actin assembly by clustering Ena/VASP proteins and tethering them to actin filaments. *eLife* [Online] **4**:1–29. Available at: <https://elifesciences.org/articles/06585> [Accessed: 13 July 2019].
- Harvey, A. *et al.* (2016). Vascular Fibrosis in Aging and Hypertension: Molecular Mechanisms and Clinical Implications. *Canadian Journal of Cardiology* [Online] **32**:659–668. Available at: <https://www.sciencedirect.com/science/article/pii/S0828282X16002154> [Accessed: 12 September 2019].
- Hemmings, L. *et al.* (1996). Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *Journal of cell science* **109** (Pt 1):2715–26.

- Hoffmeyer, T.T. and Burkhardt, P. (2016). Choanoflagellate models — *Monosiga brevicollis* and *Salpingoeca rosetta*. *Current Opinion in Genetics & Development* [Online] **39**:42–47. Available at: <http://dx.doi.org/10.1016/j.gde.2016.05.016>.
- Holdener, B.C. *et al.* (1994). *msd* is required for mesoderm induction in mice. *Development (Cambridge, England)* [Online] **120**:1335–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8026341> [Accessed: 29 May 2018].
- Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**:D512-20
- Hsieh, J.-C. *et al.* (2003). *Mesd* Encodes an LRP5/6 Chaperone Essential for Specification of Mouse Embryonic Polarity. *Cell* [Online] **112**:355–367. Available at: [https://www.cell.com/cell/references/S0092-8674\(03\)00045-X](https://www.cell.com/cell/references/S0092-8674(03)00045-X) [Accessed: 29 May 2018].
- Hughes, P.E. *et al.* (1996). Breaking the integrin hinge: A defined structural constraint regulates integrin signaling. *Journal of Biological Chemistry* **271**:6571–6574.
- Humphrey, J.D., Dufresne, E.R. and Schwartz, M.A. (2014). Mechanotransduction and extracellular matrix homeostasis. *Nature Reviews Molecular Cell Biology* [Online] **15**:802–812. Available at: <http://www.nature.com/articles/nrm3896> [Accessed: 12 September 2019].
- Humphries, J.D. *et al.* (2007). Vinculin controls focal adhesion formation by direct interactions with talin and actin. *The Journal of Cell Biology* [Online] **179**:1043–1057. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.200703036> [Accessed: 12 November 2019].
- Isenberg, G. *et al.* (2002). Membrane fusion induced by the major lipid-binding domain of the cytoskeletal protein talin. *Biochemical and Biophysical Research Communications* **295**:636–643.
- Ishida, T. and Kinoshita, K. (2007). PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Research* [Online] **35**:W460–W464. Available at: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkm363> [Accessed: 21 October 2019].
- Jacquemet, G. *et al.* (2019). Filopodome Mapping Identifies p130Cas as a

- Mechanosensitive Regulator of Filopodia Stability. *Current Biology* [Online] **29**:202-216.e7. Available at: <https://www.biorxiv.org/content/early/2018/07/26/377879> [Accessed: 27 July 2018].
- Jacquemet, G., Hamidi, H. and Ivaska, J. (2015). Filopodia in cell adhesion, 3D migration and cancer cell invasion. *Current Opinion in Cell Biology* [Online] **36**:23–31. Available at: <http://dx.doi.org/10.1016/j.ceb.2015.06.007> [Accessed: 12 March 2018].
- Jansen, S. *et al.* (2011). Mechanism of actin filament bundling by fascin. *The Journal of biological chemistry* [Online] **286**:30087–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21685497> [Accessed: 9 March 2018].
- Jevnikar, Z. *et al.* (2012). Three-dimensional invasion of macrophages is mediated by cysteine cathepsins in protrusive podosomes. *European Journal of Immunology* [Online] **42**:3429–3441. Available at: <http://doi.wiley.com/10.1002/eji.201242610> [Accessed: 1 October 2019].
- Jin, J.-K. *et al.* (2015). Talin1 phosphorylation activates β 1 integrins: a novel mechanism to promote prostate cancer bone metastasis. *Oncogene* [Online] **34**:1811–1821. Available at: <http://www.nature.com/articles/onc2014116> [Accessed: 26 October 2019].
- Jockusch, B.M. *et al.* (1995). The Molecular Architecture of Focal Adhesions. *Annual Review of Cell and Developmental Biology* [Online] **11**:379–416. Available at: <http://www.annualreviews.org/doi/10.1146/annurev.cb.11.110195.002115> [Accessed: 1 October 2019].
- Johnson, H.E. *et al.* (2015). F-actin bundles direct the initiation and orientation of lamellipodia through adhesion-based signaling. *The Journal of Cell Biology* [Online] **208**:443–455. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201406102> [Accessed: 18 October 2019].
- Jones, M.C. *et al.* (2018). Cell adhesion is regulated by CDK1 during the cell cycle. *The Journal of Cell Biology* [Online] **217**:3203–3218. Available at: <http://jcb.rupress.org/content/early/2018/06/20/jcb.201802088> [Accessed: 27 June 2018].
- Kabsch, W. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography* [Online] **66**:125–132. Available at: <http://scripts.iucr.org/cgi->

bin/paper?S0907444909047337 [Accessed: 28 September 2019].

Kakinuma, N. *et al.* (2008). Kank regulates RhoA-dependent formation of actin stress fibers and cell migration via 14-3-3 in PI3K–Akt signaling. *The Journal of Cell Biology* [Online] **181**:537–549. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18458160> [Accessed: 27 August 2019].

Kalli, A.C. *et al.* (2010). The Structure of the Talin/Integrin Complex at a Lipid Bilayer: An NMR and MD Simulation Study. *Structure* [Online] **18**:1280–1288. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0969212610003047> [Accessed: 4 November 2019].

Kanamori, H. *et al.* (2011). Identification by Differential Tissue Proteome Analysis of Talin-1 as a Novel Molecular Marker of Progression of Hepatocellular Carcinoma. *Oncology* [Online] **80**:406–415. Available at: <https://www.karger.com/Article/FullText/330734> [Accessed: 26 October 2019].

Kanchanawong, P. *et al.* (2010). Nanoscale architecture of integrin-based cell adhesions. *Nature* **468**:580–584.

Kang T. *et al.* (2008) VRK1 phosphorylates CREB and mediates *CCND1* expression. *Journal of Cell Science*. **121**: 3035-3041.

Karpov, S.A. and Leadbeater, B.S.C. (1998). Cytoskeleton Structure and Composition in Choanoflagellates. *The Journal of Eukaryotic Microbiology* [Online] **45**:361–367. Available at: <http://doi.wiley.com/10.1111/j.1550-7408.1998.tb04550.x> [Accessed: 20 October 2019].

Käßmeyer, S. *et al.* (2009). New Insights in Vascular Development: Vasculogenesis and Endothelial Progenitor Cells. *Anatomia, Histologia, Embryologia* [Online] **38**:1–11. Available at: <http://doi.wiley.com/10.1111/j.1439-0264.2008.00894.x> [Accessed: 1 December 2019].

Kechagia, J.Z., Ivaska, J. and Roca-Cusachs, P. (2019). Integrins as biomechanical sensors of the microenvironment. *Nature Reviews Molecular Cell Biology* [Online] **20**:457–473. Available at: <http://www.nature.com/articles/s41580-019-0134-2> [Accessed: 11 June 2019].

Kelley, L.A. *et al.* (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**:845–858.

- Khaitlina, S.Y. and Strzelecka-Gólaszewska, H. (2002). Role of the DNase-I-Binding Loop in Dynamic Properties of Actin Filament. *Biophysical Journal* [Online] **82**:321–334. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0006349502753976> [Accessed: 18 October 2019].
- King, N. (2003). Evolution of Key Cell Signaling and Adhesion Protein Families Predates Animal Origins. *Science* [Online] **301**:361–363. Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.1083853> [Accessed: 20 October 2019].
- Krause, M. and Gautreau, A. (2014). Steering cell migration: lamellipodium dynamics and the regulation of directional persistence. *Nature Reviews Molecular Cell Biology* [Online] **15**:577–590. Available at: <http://www.nature.com/articles/nrm3861> [Accessed: 18 October 2019].
- Krissinel, E. and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* [Online] **372**:774–797. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0022283607006420>.
- Kumar, A. *et al.* (2016). Talin tension sensor reveals novel features of focal adhesion force transmission and mechanosensitivity. *The Journal of Cell Biology* [Online] **213**:371–383. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201510012>.
- Lafuente, E.M. *et al.* (2004). RIAM, an Ena/VASP and Profilin Ligand, Interacts with Rap1-GTP and Mediates Rap1-Induced Adhesion. *Developmental Cell* [Online] **7**:585–595. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1534580704002850> [Accessed: 13 July 2019].
- Lai, M.-T. *et al.* (2011). Talin-1 overexpression defines high risk for aggressive oral squamous cell carcinoma and promotes cancer metastasis. *The Journal of Pathology* [Online] **224**:367–376. Available at: <http://doi.wiley.com/10.1002/path.2867> [Accessed: 26 October 2019].
- Lansbergen, G. *et al.* (2006). CLASPs Attach Microtubule Plus Ends to the Cell Cortex through a Complex with LL5 β . *Developmental Cell* **11**:21–32.
- Läubli, H. and Borsig, L. (2010). Selectins promote tumor metastasis. *Seminars in Cancer Biology* [Online] **20**:169–177. Available at: <https://www-sciencedirect-com.chain.kent.ac.uk/science/article/pii/S1044579X1000026X> [Accessed: 13 September 2019].

- Law, A.L. *et al.* (2013). Lamellipodin and the Scar/WAVE complex cooperate to promote cell migration in vivo. *Journal of Cell Biology* **203**:673–689.
- Lawson, C. *et al.* (2012). FAK promotes recruitment of talin to nascent adhesions to control cell motility. *The Journal of Cell Biology* [Online] **196**:223–232. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201108078>.
- Leduc, C. and Etienne-Manneville, S. (2015). Intermediate filaments in cell migration and invasion: the unusual suspects. *Current Opinion in Cell Biology* [Online] **32**:102–112. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0955067415000071> [Accessed: 14 October 2019].
- Lee, H.-S. *et al.* (2004). Characterization of an Actin-binding Site within the Talin FERM Domain. *Journal of Molecular Biology* [Online] **343**:771–784. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0022283604010319> [Accessed: 28 October 2019].
- Lee, H.-S.S. *et al.* (2009). RIAM Activates Integrins by Linking Talin to Ras GTPase Membrane-targeting Sequences. *Journal of Biological Chemistry* [Online] **284**:5119–5127. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19098287> [Accessed: 28 July 2019].
- Li, J. and Springer, T.A. (2017). Integrin extension enables ultrasensitive regulation by cytoskeletal force. *Proceedings of the National Academy of Sciences* [Online] **114**:4685–4690. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1704171114> [Accessed: 20 October 2019].
- Li *et al.* (2019) iPhoPred: A Predictor for Identifying Phosphorylation Sites in Human Protein, *IEEE Access*. **7**: 177517-177528. doi: 10.1109/ACCESS.2019.2953951
- Linder, S. and Wiesner, C. (2016). Feel the force: Podosomes in mechanosensing. *Experimental Cell Research* [Online] **343**:67–72. Available at: <https://www.sciencedirect-com.chain.kent.ac.uk/science/article/pii/S0014482715301671> [Accessed: 15 September 2019].
- Liu, J. *et al.* (2015). Talin determines the nanoscale architecture of focal adhesions. *Proceedings of the National Academy of Sciences* [Online] **112**:E4864–E4873. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1512025112>.

- Lock, J.G. *et al.* (2019). Clathrin-containing adhesion complexes. *The Journal of Cell Biology* [Online] **218**:2086–2095. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201811160> [Accessed: 10 October 2019].
- Lock, J.G. *et al.* (2018). Reticular adhesions are a distinct class of cell-matrix adhesions that mediate attachment during mitosis. *Nature Cell Biology* [Online] **20**:1290–1302. Available at: <http://www.nature.com/articles/s41556-018-0220-2> [Accessed: 23 October 2018].
- Long, F. *et al.* (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D Biological Crystallography* [Online] **64**:125–132. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444907050172>.
- López-Colomé, A.M. *et al.* (2017). Paxillin: a crossroad in pathological cell migration. *Journal of Hematology & Oncology* [Online] **10**:50. Available at: <http://jhonline.biomedcentral.com/articles/10.1186/s13045-017-0418-y> [Accessed: 8 November 2019].
- Lowery, J. *et al.* (2015). Intermediate Filaments Play a Pivotal Role in Regulating Cell Architecture and Function. *Journal of Biological Chemistry* [Online] **290**:17145–17153. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.R115.640359> [Accessed: 14 October 2019].
- Lyulcheva, E. *et al.* (2008). Drosophila Pico and Its Mammalian Ortholog Lamellipodin Activate Serum Response Factor and Promote Cell Proliferation. *Developmental Cell* [Online] **15**:680–690. Available at: <https://www.sciencedirect.com/science/article/pii/S1534580708004024?via%3Dihub> [Accessed: 5 August 2019].
- Malumbres, M. *et al.* (2009). Cyclin-dependent kinases: a family portrait. *Nature Cell Biology* [Online] **11**:1275–1276. Available at: <http://www.nature.com/articles/ncb1109-1275> [Accessed: 22 August 2019].
- Martin, L., Schwarz, S. and Breitsprecher, D. (2014). *Thermal Unfolding Analyzing Thermal Unfolding of Proteins: The Prometheus NT.48*.
- Mayor, R. and Etienne-Manneville, S. (2016). The front and rear of collective cell migration. *Nature Reviews Molecular Cell Biology* [Online] **17**:97–109. Available at: <http://www.nature.com/articles/nrm.2015.14> [Accessed: 13 October 2019].

- McCoy, A.J. *et al.* (2007). Phaser crystallographic software. *Journal of Applied Crystallography* [Online] **40**:658–674. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0021889807021206> [Accessed: 29 September 2019].
- Mehidi, A. *et al.* (2019). Transient Activations of Rac1 at the Lamellipodium Tip Trigger Membrane Protrusion. *Current Biology* [Online] **29**:2852-2866.e5. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0960982219308802> [Accessed: 21 August 2019].
- Mertens, H.D.T. and Svergun, D.I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* [Online] **172**:128–141. Available at: <https://www-sciencedirect-com.chain.kent.ac.uk/science/article/pii/S1047847710001905?via%3Dihub> [Accessed: 10 April 2019].
- Moudjou, M. *et al.* (1996). gamma-Tubulin in mammalian cells: the centrosomal and the cytosolic forms. *Journal of cell science* [Online] **109** (Pt 4):875–87. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8718679> [Accessed: 13 October 2019].
- Murphy, D.A. *et al.* (2011). A Src-Tks5 Pathway Is Required for Neural Crest Cell Migration during Embryonic Development Parsons, M. ed. *PLoS ONE* [Online] **6**:e22499. Available at: <https://dx.plos.org/10.1371/journal.pone.0022499> [Accessed: 15 September 2019].
- Murphy, D.A. and Courtneidge, S.A. (2011). The ‘ins’ and ‘outs’ of podosomes and invadopodia: characteristics, formation and function. *Nature Reviews Molecular Cell Biology* [Online] **12**:413–426. Available at: <http://www.nature.com/articles/nrm3141> [Accessed: 1 October 2019].
- Nelson, W.J. (2003). Adaptation of core mechanisms to generate cell polarity. *Nature* [Online] **422**:766–774. Available at: <http://www.nature.com/articles/nature01602> [Accessed: 12 September 2019].
- Nichols, J. R. and Traktman, P. (2004) Characterization of Three Paralogous Members of the Mammalian Vaccinia Related Kinase Family. *The Journal of Biological Chemistry*. **279**, **9**:7934-7946.
- Niroula, A., Urolagin, S. and Vihinen, M. (2015). PON-P2: Prediction Method for Fast and Reliable Identification of Harmful Variants Tosatto, S. C. E. ed. *PLOS ONE* [Online] **10**:e0117380. Available at: <http://dx.plos.org/10.1371/journal.pone.0117380>

[Accessed: 27 October 2019].

Pankov, R. *et al.* (2019). Characterization of stitch adhesions: Fibronectin-containing cell-cell contacts formed by fibroblasts. *Experimental Cell Research* [Online]:111616. Available at: [https://www.sciencedirect-com.chain.kent.ac.uk/science/article/pii/S0014482719304781](https://www.sciencedirect.com.chain.kent.ac.uk/science/article/pii/S0014482719304781) [Accessed: 1 October 2019].

Papatheodorou, I. *et al.* (2020) Expression Atlas update: from tissues to single cells, *Nucleic Acids Research*, **48**(D1):D77–D83. <https://doi.org/10.1093/nar/gkz947>.

Paranavitane, V., Stephens, L.R. and Hawkins, P.T. (2007). Structural determinants of LL5 β subcellular localisation and association with filamin C. *Cellular Signalling* **19**:817–824.

Parri, M. and Chiarugi, P. (2010). Rac and Rho GTPases in cancer cell motility control. *Cell Communication and Signaling* [Online] **8**:23. Available at: <https://biosignaling.biomedcentral.com/articles/10.1186/1478-811X-8-23> [Accessed: 18 October 2019].

Pasapera, A.M. *et al.* (2010). Myosin II activity regulates vinculin recruitment to focal adhesions through FAK-mediated paxillin phosphorylation. *Journal of Cell Biology* **188**:877–890.

Pirkmajer S, Chibalin AV. (2011) Serum starvation: caveat emptor. *Am J Physiol Cell Physiol*. **301**(2):C272-C279. doi:10.1152/ajpcell.00091.2011

De Pascalis, C. *et al.* (2018). Intermediate filaments control collective migration by restricting traction forces and sustaining cell–cell contacts. *The Journal of Cell Biology* [Online] **217**:3031–3044. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201801162> [Accessed: 14 October 2019].

Petryszak, R. *et al.* (2016). Expression Atlas update—an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Research* [Online] **44**:D746–D752. Available at: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1045> [Accessed: 14 November 2019].

Piiaodov, V. *et al.* (2018). SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative

- scale. *Protein Science* [Online] **28**:pro.3528. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/pro.3528> [Accessed: 25 October 2019].
- Pol, S.B. Vande, Brown, M.C. and Turner, C.E. (1998). Association of Bovine Papillomavirus Type 1 E6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif. *Oncogene* [Online] **16**:43–52. Available at: <http://www.nature.com/articles/1201504> [Accessed: 15 October 2019].
- Postma, M. and Goedhart, J. (2019). PlotsOfData—A web app for visualizing data together with their summaries. *PLOS Biology* [Online] **17**:e3000202. Available at: <http://dx.plos.org/10.1371/journal.pbio.3000202> [Accessed: 30 November 2019].
- Rahikainen, R. *et al.* (2017). Mechanical stability of talin rod controls cell migration and substrate sensing. *Scientific Reports* **7**.
- Rahikainen, R. *et al.* (2019). Talin-mediated force transmission and talin rod domain unfolding independently regulate adhesion signaling. *Journal of Cell Science* [Online] **132**:jcs226514. Available at: <http://jcs.biologists.org/lookup/doi/10.1242/jcs.226514> [Accessed: 6 March 2019].
- Rambo, R.P. and Tainer, J.A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* [Online] **95**:559–571. Available at: <http://doi.wiley.com/10.1002/bip.21638> [Accessed: 24 October 2019].
- Sabatier, L. *et al.* (2009). Fibrillin Assembly Requires Fibronectin. *Molecular Biology of the Cell* [Online] **20**:846–858. Available at: <http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E08> [Accessed: 10 October 2019].
- Sanders, T.A., Llagostera, E. and Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature* [Online] **497**:628–632. Available at: <http://www.nature.com/articles/nature12157> [Accessed: 1 December 2019].
- Santamaría, D. *et al.* (2007). Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* [Online] **448**:811–815. Available at: <http://www.nature.com/articles/nature06046> [Accessed: 27 August 2019].
- Sarhan, A.R. *et al.* (2016). LAR protein tyrosine phosphatase regulates focal adhesions

through CDK1. *Journal of Cell Science* **129**:2962–2971.

Sarkar, S. *et al.* (2002). A Novel Ankyrin Repeat-containing Gene (Kank) Located at 9p24 Is a Growth Suppressor of Renal Cell Carcinoma. *Journal of Biological Chemistry* [Online] **277**:36585–36591. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.M204244200> [Accessed: 15 August 2019].

Saunders, J.T. and Schwarzbauer, J.E. (2019). Fibronectin matrix as a scaffold for procollagen proteinase binding and collagen processing Ginsberg, M. H. ed. *Molecular Biology of the Cell* [Online] **30**:2218–2226. Available at: <https://www.molbiolcell.org/doi/10.1091/mbc.E19-03-0140> [Accessed: 10 October 2019].

Sayyad, W.A. *et al.* (2015). The role of myosin-II in force generation of DRG filopodia and lamellipodia. *Scientific Reports* [Online] **5**:7842. Available at: <http://www.nature.com/articles/srep07842> [Accessed: 20 October 2019].

Schäfer, C. *et al.* (2009). One step ahead: Role of filopodia in adhesion formation during cell migration of keratinocytes. *Experimental Cell Research* [Online] **315**:1212–1224. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0014482708004977> [Accessed: 1 December 2019].

Schaller, M.D. *et al.* (1994). Autophosphorylation of the focal adhesion kinase, pp125FAK, directs SH2-dependent binding of pp60src. *Molecular and Cellular Biology* **14**:1680–1688.

Schindelin, J. *et al.* (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods* **9**:676–682.

Sebé-Pedrós. *et al.* (2010) Ancient origin of the integrin-mediated adhesion and signaling machinery. *PNAS* **107**: 10142-10147.

Senetar, M.A. and McCann, R.O. (2005). Gene duplication and functional divergence during evolution of the cytoskeletal linker protein talin. *Gene* [Online] **362**:141–152. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0378111905005160> [Accessed: 21 October 2019].

Serra-Pagès, C. *et al.* (1998). Liprins, a Family of LAR Transmembrane Protein-tyrosine Phosphatase-interacting Proteins. *Journal of Biological Chemistry* [Online]

- 273**:15611–15620. Available at: <http://www.jbc.org/lookup/doi/10.1074/jbc.273.25.15611> [Accessed: 17 August 2019].
- Sievers, F. *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* [Online] **7**:539. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1038/msb.2011.75> [Accessed: 5 September 2019].
- Skinner, S.P. *et al.* (2015). Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallographica Section D Biological Crystallography* [Online] **71**:154–161. Available at: <http://scripts.iucr.org/cgi-bin/paper?S1399004714026662>.
- De Smet, F. *et al.* (2009). Mechanisms of Vessel Branching. *Arteriosclerosis, Thrombosis, and Vascular Biology* [Online] **29**:639–649. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19265031> [Accessed: 1 December 2019].
- Song, J., *et al.* (2017) PhosphoPredict: A bioinformatics tool for prediction of human kinase-specific phosphorylation substrates and sites by integrating heterogeneous feature selection. *Sci Rep* **7**:6862 <https://doi.org/10.1038/s41598-017-07199-4>
- Sottile, J. and Hocking, D.C. (2002). Fibronectin Polymerization Regulates the Composition and Stability of Extracellular Matrix Fibrils and Cell-Matrix Adhesions Schwartz, M. ed. *Molecular Biology of the Cell* [Online] **13**:3546–3559. Available at: <https://www.molbiolcell.org/doi/10.1091/mbc.e02-01-0048> [Accessed: 10 October 2019].
- Spudich, J.A. and Watt, S. (1971). The regulation of rabbit skeletal muscle contraction. *Journal of Biological Chemistry* **246**:4866.
- Spuul, P. *et al.* (2016). VEGF-A/Notch-Induced Podosomes Proteolyse Basement Membrane Collagen-IV during Retinal Sprouting Angiogenesis. *Cell Reports* [Online] **17**:484–500. Available at: <https://www.sciencedirect.com/science/article/pii/S2211124716312384> [Accessed: 15 September 2019].
- Srivastava, M. *et al.* (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature* **466**:720–726.

- Stevenson, R.P., Veltman, D. and Machesky, L.M. (2012). Actin-bundling proteins in cancer progression at a glance. *Journal of Cell Science* [Online] **125**:1073–1079. Available at: <http://jcs.biologists.org/cgi/doi/10.1242/jcs.093799> [Accessed: 18 October 2019].
- Sulimenko, V. *et al.* (2017). Regulation of microtubule nucleation mediated by γ -tubulin complexes. *Protoplasma* [Online] **254**:1187–1199. Available at: <http://link.springer.com/10.1007/s00709-016-1070-z> [Accessed: 13 October 2019].
- Sun, Z. *et al.* (2016). Kank2 activates talin, reduces force transduction across integrins and induces central adhesion formation. *Nature Cell Biology* [Online] **18**:941–953. Available at: <http://www.nature.com/articles/ncb3402>.
- Tang, Z. *et al.* (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research* [Online] **45**:W98–W102. Available at: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkx247> [Accessed: 27 October 2019].
- Tatarano, S. *et al.* (2012). Novel oncogenic function of mesoderm development candidate 1 and its regulation by MiR-574-3p in bladder cancer cell lines. *International Journal of Oncology* [Online] **40**:951–959. Available at: <https://www.spandidos-publications.com/10.3892/ijo.2011.1294> [Accessed: 3 April 2019].
- Tate, J.G. *et al.* (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* [Online] **47**:D941–D947. Available at: <https://academic.oup.com/nar/article/47/D1/D941/5146192> [Accessed: 26 October 2019].
- Tojkander, S., Gateva, G. and Lappalainen, P. (2012). Actin stress fibers - assembly, dynamics and biological roles. *Journal of Cell Science* [Online] **125**:1855–1864. Available at: <http://jcs.biologists.org/cgi/doi/10.1242/jcs.098087> [Accessed: 21 November 2019].
- Uhlén M, Fagerberg L, Hallström BM, *et al.* (2015) Proteomics. Tissue-based map of the human proteome. *Science*. **347**(6220):1260419. doi:10.1126/science.1260419
- Valiente, M. and Marín, O. (2010). Neuronal migration mechanisms in development and disease. *Current Opinion in Neurobiology* **20**:68–78.
- van der Vaart, B. *et al.* (2013). CFEOM1-Associated Kinesin KIF21A Is a Cortical

- Microtubule Growth Inhibitor. *Developmental Cell* [Online] **27**:145–160. Available at: <https://www.sciencedirect.com/science/article/pii/S1534580713005364> [Accessed: 27 August 2019].
- Venning, F.A., Wullkopf, L. and Eler, J.T. (2015). Targeting ECM Disrupts Cancer Progression. *Frontiers in Oncology* [Online] **5**:224. Available at: <http://journal.frontiersin.org/Article/10.3389/fonc.2015.00224/abstract> [Accessed: 12 September 2019].
- Vignjevic, D. *et al.* (2006). Role of fascin in filopodial protrusion. *The Journal of Cell Biology* [Online] **174**:863–875. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.200603013> [Accessed: 18 October 2019].
- Vonrhein, C. *et al.* (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D Biological Crystallography* [Online] **67**:293–302. Available at: <http://scripts.iucr.org/cgi-bin/paper?S0907444911007773> [Accessed: 28 September 2019].
- Vranken, W.F. *et al.* (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics* [Online] **59**:687–696. Available at: <http://doi.wiley.com/10.1002/prot.20449> [Accessed: 19 September 2019].
- Wai Wong, C., Dye, D.E. and Coombe, D.R. (2012). The Role of Immunoglobulin Superfamily Cell Adhesion Molecules in Cancer Metastasis. *International Journal of Cell Biology* [Online] **2012**:1–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22272201> [Accessed: 13 September 2019].
- Wakayama, Y. *et al.* (2015). Cdc42 mediates Bmp - Induced sprouting angiogenesis through Fmnl3-driven assembly of endothelial filopodia in zebrafish. *Developmental Cell* **32**:109–122.
- Wang, C. *et al.* (2020) GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins. *Genomics, Proteomics and Bioinformatics* [Online] S1672-0229(20)30027-9. Available at: <https://www.sciencedirect.com/science/article/pii/S1672022920300279#s0035> [Accessed 2 July 2020].
- Wang, H., Robinson, R.C. and Burtnick, L.D. (2010). The structure of native G-actin.

- Cytoskeleton* [Online] **67**:456–465. Available at: <http://doi.wiley.com/10.1002/cm.20458> [Accessed: 18 October 2019].
- Wang, P., Ballestrem, C. and Streuli, C.H. (2011). The C terminus of talin links integrins to cell cycle progression. *The Journal of Cell Biology* [Online] **195**:499–513. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201104128> [Accessed: 6 November 2019].
- Wang, Y. and Jardetzky, O. (2002). Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science* [Online] **11**:852–861. Available at: <http://www.proteinscience.org/cgi/doi/10.1110/ps.3180102> [Accessed: 4 October 2019].
- Ward, S.G. (2009). Millipede-like Lymphocyte Crawling: Feeling the Way with Filopodia. *Immunity* [Online] **30**:315–317. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1074761309001101> [Accessed: 20 October 2019].
- Warner, H., Wilson, B.J. and Caswell, P.T. (2019). Control of adhesion and protrusion in cell migration by Rho GTPases. *Current Opinion in Cell Biology* [Online] **56**:64–70. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S095506741830111X> [Accessed: 5 October 2019].
- Wawro, B. *et al.* (2005). Role of Actin DNase-I-Binding Loop in Myosin Subfragment 1-Induced Polymerization of G-actin: Implications for the Mechanism of Polymerization. *Biophysical Journal* [Online] **88**:2883–2896. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0006349505733415> [Accessed: 18 October 2019].
- Wegener, K.L. *et al.* (2007). Structural Basis of Integrin Activation by Talin. *Cell* [Online] **128**:171–182. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0092867406015960> [Accessed: 4 November 2019].
- Wehrle-Haller, B. (2012). Structure and function of focal adhesions. *Current Opinion in Cell Biology* [Online] **24**:116–124. Available at: <https://www.sciencedirect.com/science/article/pii/S0955067411001475?via%3Dihub> [Accessed: 1 October 2019].

- Wei, L. *et al.* (2017). PhosPred-RF: A Novel Sequence-Based Predictor for Phosphorylation Sites Using Sequential Information Only. *IEEE Transactions on NanoBioscience*, **16**: 240-247. doi: 10.1109/TNB.2017.2661756.
- Weinstein, J.N. *et al.* (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics* [Online] **45**:1113–1120. Available at: <http://www.cancergenome.nih.gov/>. [Accessed: 27 October 2019].
- Wen, Z. *et al.* (2019). Talin2 regulates invasion of human breast cancer MDA-MB-231 cells via alteration of the tumor microenvironment. *Oncology Letters* [Online] **17**:4835–4842. Available at: <http://www.spandidos-publications.com/10.3892/ol.2019.10175> [Accessed: 26 October 2019].
- Whitewood, A.J. *et al.* (2018). Chlamydial virulence factor TarP mimics talin to disrupt the talin-vinculin complex. *FEBS Letters* [Online] **592**:1751–1760. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/1873-3468.13074>.
- Wines, M.E. *et al.* (2001). Identification of Mesoderm Development (mesd) Candidate Genes by Comparative Mapping and Genome Sequence Analysis. *Genomics* [Online] **72**:88–98. Available at: <https://www.sciencedirect.com/science/article/pii/S0888754300964668?via%3Dihub> [Accessed: 9 April 2019].
- Wines, M.E. *et al.* (2000). Physical localization of the mesoderm development (mesd) functional region. *Genomics* [Online] **68**:322–329. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10995574> [Accessed: 29 May 2018].
- Winkelman, J.D. *et al.* (2016). Fascin- and α -Actinin-Bundled Networks Contain Intrinsic Structural Features that Drive Protein Sorting. *Current biology: CB* [Online] **26**:2697–2706. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27666967> [Accessed: 14 February 2018].
- Wong, S., Guo, W.-H. and Wang, Y.-L. (2014). Fibroblasts probe substrate rigidity with filopodia extensions before occupying an area. *Proceedings of the National Academy of Sciences* [Online] **111**:17176–17181. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1412285111> [Accessed: 18 October 2019].
- Wood, W. *et al.* (2002). Wound healing recapitulates morphogenesis in *Drosophila* embryos. *Nature Cell Biology* **4**:1–7.

- Worbs, T., Hammerschmidt, S.I. and Förster, R. (2017). Dendritic cell migration in health and disease. *Nature Reviews Immunology* **17**:30–48.
- Wu, S.G. *et al.* (2017). miR-508-5p acts as an anti-oncogene by targeting MESDC1 in hepatocellular carcinoma. *Neoplasma* [Online] **64**:40–47. Available at: http://www.elis.sk/index.php?page=shop.product_details&flypage=flypage.tpl&product_id=4976&category_id=85&option=com_virtuemart.
- Wynne, J.P. *et al.* (2012). Rap1-interacting adapter molecule (RIAM) associates with the plasma membrane via a proximity detector. *The Journal of Cell Biology* [Online] **199**:317–329. Available at: <http://www.jcb.org/lookup/doi/10.1083/jcb.201201157> [Accessed: 15 August 2019].
- Xiao, H. *et al.* (2013). MEK/ERK pathway mediates PKC activation-induced recruitment of PKC ζ and MMP-9 to podosomes. *Journal of Cellular Physiology* [Online] **228**:416–427. Available at: <http://doi.wiley.com/10.1002/jcp.24146> [Accessed: 1 October 2019].
- Xu, W., Baribault, H. and Adamson, E.D. (1998). Vinculin knockout results in heart and brain defects during embryonic development. *Development* **125**:327–337.
- Yao, M. *et al.* (2015). Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Scientific Reports* [Online] **4**:4610. Available at: <http://www.nature.com/articles/srep04610>.
- Yao, M. *et al.* (2016). The mechanical response of talin. *Nature Communications* [Online] **7**:11966. Available at: <http://www.nature.com/doi/10.1038/ncomms11966>.
- Yu, M. *et al.* (2019). Force-dependent regulation of KANK1-talin complex at focal adhesions. *Nano Letters* [Online]:acs.nanolett.9b01732. Available at: <http://pubs.acs.org/doi/10.1021/acs.nanolett.9b01732> [Accessed: 10 August 2019].
- Zacharchenko, T. *et al.* (2016). LD Motif Recognition by Talin: Structure of the Talin-DLC1 Complex. *Structure* [Online] **24**:1130–1141. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0969212616300776>.
- Zamir, E. *et al.* (2000). Dynamics and segregation of cell–matrix adhesions in cultured fibroblasts. *Nature Cell Biology* [Online] **2**:191–196. Available at: http://www.nature.com/articles/ncb0400_191 [Accessed: 10 October 2019].
- Zanet, J. *et al.* (2012) Fascin promotes filopodia formation independent of its role in actin

- bundling. *J Cell Biol* **197**(4): 477–486. doi: <https://doi.org/10.1083/jcb.201110135>
- Zanier, K. *et al.* (2013). Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* **339**:694–698.
- Zhang, H. *et al.* (2016). Structural and Functional Analysis of a Talin Triple-Domain Module Suggests an Alternative Talin Autoinhibitory Configuration. *Structure* [Online] **24**:721–729. Available at: <http://dx.doi.org/10.1016/j.str.2016.02.020>.
- Zhang, S. *et al.* (2018). Selective Filopodia Adhesion Ensures Robust Cell Matching in the Drosophila Heart. *Developmental Cell* [Online] **46**:189-203.e4. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1534580718305008> [Accessed: 20 October 2019].
- Zhang, X. *et al.* (2008). Talin depletion reveals independence of initial cell spreading from integrin activation and traction. *Nature Cell Biology* [Online] **10**:1062–1068. Available at: <http://www.nature.com/articles/ncb1765> [Accessed: 17 November 2019].
- Zhen, G. and Cao, X. (2014). Targeting TGF β signaling in subchondral bone and articular cartilage homeostasis. *Trends in Pharmacological Sciences* [Online] **35**:227–236. Available at: <https://www.sciencedirect.com/science/article/pii/S016561471400039X?via%3Dihub> [Accessed: 12 September 2019].
- Zhu, Y. *et al.* (2008). Kank proteins: A new family of ankyrin-repeat domain-containing proteins. *Biochimica et Biophysica Acta (BBA) - General Subjects* [Online] **1780**:128–133. Available at: <https://www.sciencedirect.com/science/article/pii/S0304416507002176?via%3Dihub> [Accessed: 15 August 2019].
- Ziegler, W.H., Liddington, R.C. and Critchley, D.R. (2006). The structure and regulation of vinculin. *Trends in Cell Biology* [Online] **16**:453–460. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0962892406001930> [Accessed: 12 November 2019].