



Kent Academic Repository

Khan, Rejina Begum (2021) *Investigating the structure, regulation and interactions of KANK proteins*. Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from

<https://kar.kent.ac.uk/86881/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.22024/UniKent/01.02.86881>

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, regulation and interactions of KANK proteins

Rejina Begum Khan

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

School of Biosciences

University of Kent

2021

DECLARATION

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

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Dr Ben Goult: I feel very lucky to have been assigned to work with you as part of my undergraduate final year project and I am so grateful for every opportunity you've given to me since. Your enthusiasm, guidance and constructive feedback throughout is unmatched.

I would like to thank members of the Goult lab, past and present, in addition to the Mulvihill and Geeves labs for the friendship and motivation and for creating an environment that was so enjoyable to be a part of. To Dr Karen Baker, Dr Lorena Varela Alvarez, Dr Jonathan Walklate, Dr Austin Whitewood, Dr Alana Cowell, Dr Chloe Johnson, Nyasha Allen, Gabriella Laker, Tom Paige and Dr Tara Eastwood in particular – our cake/coffee/snack breaks and socials were always such a highlight and I hope we can organise many more in the future (pandemic-permitting)! An extra shout out goes to Karen and Lorena for being the most wonderful and supportive post-doctoral scientists to work with – thank you for putting up with my endless questions over the years!

Thank you also to Dr Jose Ortega Roldan for his help with the NMR experiments in Chapter 5 – amongst the time constraints from lockdown, that chapter would not have been possible without your help. Thanks also to Dr Gary Thompson for NMR training and support throughout and to Dr Abhi Singh for assistance with X-ray crystallographic data analysis.

I would like to thank our collaborators from Prof Anna Akhmanova's group at Utrecht University, and also a very big thank you to Dr Jose M. de Pereda and his group at the Centro de Investigación del Cáncer in Salamanca for taking the time to teach me further X-ray crystallography skills (and attempting to teach me Spanish) and for making me feel so at home for 5 weeks.

Finally, a massive thank you to all my loved ones for supporting me throughout this time – especially those of you who've only ever known me as sleepy PhD student Rej and wanted to hang out anyway!

ABSTRACT

Many cellular activities rely on interactions between cells and the extracellular matrix (ECM) and these attachments are mediated by the integrin family of transmembrane receptors. Talin, a cytoplasmic adapter protein, activates integrin by binding to the cytoplasmic tails and ultimately couples integrin to the actin cytoskeleton. Once the adhesion has formed, talin can act as a mechanosensitive signalling hub and recruit additional proteins in a force-dependent manner.

KANK proteins, via direct interaction with talin and with the kinesin KIF21A, mediate the connection of these integrin-based adhesions with dynamic microtubules. A complex of proteins termed the cortical microtubule stabilising complex (CMSC) is recruited upon talin-KANK interaction which stabilises microtubules in the vicinity of adhesions. The talin-KANK connection results in mechanosensitive crosstalk between the actin and microtubule cytoskeletons, contributes to microtubule polarity, and provides a mechanism for turnover of adhesions during cell migration.

In this thesis, a combination of structural, biochemical and biophysical approaches were used to broaden our understanding of KANK proteins. With a recurring focus on identifying similarities and differences between the four mammalian KANK isoforms, particularly within specific domains, we discover some key isoform-specific differences in terms of the ankyrin repeat domains including dimeric propensity that is unique to KANK4, characterise a novel interaction between the KN and ankyrin repeat domains of KANKs and narrow down on the required binding surface on the KN domain, and identify a novel interaction between DYNLL2 and an intrinsically disordered region in KANK1. Additionally, as part of this work, two novel crystal structures of KANK ankyrin repeat domains were solved and solution NMR data providing structural information for DYNLL2 was collected. Ultimately, the work in this thesis explores some novel properties of KANKs and provides the foundations for an array of further study into this family of proteins, allowing the deduction of their structures and functions – either related to their role as linkers between actin filaments and microtubules or potentially in other unrelated and yet undiscovered cellular contexts.

CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 Cell adhesion.....	1
1.1.1 Cell-cell adhesion.....	2
1.2 Integrin-mediated cell-extracellular matrix adhesions.....	4
1.2.1 The extracellular matrix.....	4
1.2.2 Cell-extracellular matrix adhesions	4
1.3 Integrin	5
1.3.1 Integrins overview	5
1.3.2 Integrin structure	5
1.3.3 Integrin activation	6
1.3.4 The consensus integrin adhesome	8
1.4 Integrin adhesion complexes.....	9
1.4.1 Nascent adhesions.....	9
1.4.2 Focal complexes and focal adhesions	10
1.4.3 Fibrillar adhesions	11
1.4.4 Podosomes.....	11
1.4.5 Reticular adhesions	11
1.5 Talin	12
1.5.1 Talin structure	12
1.5.2 Talin autoinhibition and activation	14
1.5.3 Talin as a mechanosensitive signalling hub	15
1.6 Microtubules in cell adhesion and migration	16
1.6.1 Microtubules	16
1.6.2 Microtubule-mediated focal adhesion turnover	17
1.6.3 The cortical microtubule stabilising complex	19
1.7 KANK proteins	21
1.7.1 An introduction to KANK proteins.....	21
1.7.2 KANK proteins in disease	22
1.8 KANK links focal adhesions to cortical microtubules	23
1.8.1 KANK proteins and the cytoskeleton.....	23
1.8.2 The role of KANK proteins in the cortical microtubule stabilising complex..	25
1.8.3 KANK proteins interact with talin to allow FA-CMSC crosstalk	26

1.9 Project aims	29
CHAPTER 2: MATERIALS & METHODS.....	30
2.1 Materials.....	30
2.1.1 Reagents and equipment.....	30
2.1.2 Buffers and media	31
2.1.3 Plasmids	32
2.2 General molecular biology.....	33
2.2.1 Generating chemically competent <i>E. coli</i> using the calcium chloride method	33
2.2.2 DNA transformations.....	33
2.3 Molecular cloning.....	33
2.3.1 Restriction enzyme digest method	33
2.3.2 Site-directed mutagenesis	37
2.4 Protein expression.....	39
2.4.1 Unlabelled protein expression	39
2.4.2 Isotopically-labelled protein expression for NMR	40
2.5 Protein purification	41
2.5.1 Cell lysis.....	41
2.5.2 Protein purification by immobilised nickel ion affinity chromatography	41
2.5.3 Buffer exchange and TEV protease cleavage.....	41
2.5.4 Ion exchange chromatography	42
2.5.5 Measurement of protein concentration and storage	42
2.5.6 SDS-PAGE analysis	43
2.6 Peptides	43
2.6.1 Synthetic peptide design	43
2.6.2 Coupling peptides to BODIPY fluorescent dye.....	43
2.7 Biochemical assays	43
2.7.1 Circular dichroism.....	43
2.7.2 Fluorescence polarisation.....	45
2.7.3 Nuclear magnetic resonance titrations	46
2.7.4 Size-exclusion chromatography-multiangle light scattering	48
2.7.5 Fluorimetry.....	49
2.8 Structural studies	49
2.8.1 X-ray crystallography	49

2.8.2 Solution NMR	51
CHAPTER 3: CHARACTERISATION OF THE KANK ANKYRIN REPEAT DOMAIN	52
3.1 Overview.....	52
3.2 Ankyrin repeat domains are important mediators of protein-protein interactions throughout biology	52
3.3 Comparing the ankyrin repeat domains of all four mammalian KANKs	53
3.3.1 Determining the boundaries for the ankyrin repeat domain of each mammalian KANK isoform.....	53
3.3.2 The crystal structure of the KANK3 ankyrin repeat domain.....	56
3.3.3 KANK3 and KANK4 ankyrin repeat domains bind KIF21A with a weaker affinity compared to KANK1 and KANK2 ankyrin repeat domains.....	59
3.4 Characterising the Ala670Val disease-causing point mutation in KANK2.....	60
3.4.1 Ala670Val causes palmoplantar keratoderma and woolly hair	60
3.4.2 Circular dichroism indicates that KANK2 wildtype and Ala670Val mutant are structurally similar	60
3.4.3 The crystal structure of the Ala670Val mutant.....	61
3.4.4 KANK2 wildtype and Ala670Val mutant bind KIF21A equally	64
3.5 Characterising the Tyr801His disease-causing point mutation in KANK4.....	64
3.5.1 Tyr801His causes a steroid-resistant form of nephrotic syndrome.....	64
3.5.2 Circular dichroism indicates that KANK4 wildtype and Tyr801His mutant have some structural differences.....	65
3.5.3 SEC-MALS analysis indicates that KANK4 wildtype and Tyr801H mutant exist in different oligomeric states	67
3.5.4 KANK4 wildtype and Tyr801His mutant ankyrin repeat domains vary slightly in their binding to KIF21A.....	69
3.6 Interrogating the KANK1 proteomics dataset to search for new KANK ankyrin repeat ligands	70
3.6.1 The KANK1 proteomics dataset	70
3.6.2 Investigating a putative interaction with Scribble homologue in humans ...	70
3.6.3 Using the KANK1 proteomics dataset to continue identifying KANK1 ligands	71
3.7 Discussion	72
CHAPTER 4: INVESTIGATION OF AN INTERDOMAIN KANK COMPLEX.....	74
4.1 Overview.....	74
4.2 Autoinhibition as a major regulatory mechanism of integrin-mediated adhesion	74

4.3 KANK1 is a leading hit in KANK1 proteomics	75
4.4 Characterising the interaction between KANK1 ankyrin repeats and KANK1 KN domain.....	76
4.4.1 Using KANK1 KN domain peptides for fluorescence polarisation assays.....	76
4.4.2 The KANK1 ankyrin repeat domain and talin1 R7R8 bind to different regions of the KANK1 KN domain.....	76
4.4.3 Residues 61-68 of the KANK1 KN domain are required for the interaction between the KANK1 KN domain and the KANK1 ankyrin repeat domain	80
4.5 Investigating an interaction between the KN domain and ankyrin repeats domain in KANK2, KANK3 and KANK4	81
4.5.1 Designing KANK2, KANK3 and KANK4 KN domain peptides to investigate autoinhibition of the other KANK isoforms.....	82
4.5.2 Optimisation of KANK3 and KANK4 KN domain peptides	83
4.5.3 Using fluorescence polarisation to investigate self-interaction of KANK2, KANK3 and KANK4.....	84
4.6 Known disease-causing mutations in KANK2 and KANK4 ankyrin repeats may affect autoinhibition	88
4.6.1 Comparing autoinhibition in wildtype KANK2 ankyrin repeat domain and mutant KANK2 ankyrin repeat domain (Ala670Val)	88
4.6.2 Comparing autoinhibition in wildtype KANK4 ankyrin repeat domain and mutant KANK4 ankyrin repeat domain (Tyr801His)	89
4.7 Attempting FRET analysis to investigate KANK autoinhibition.....	90
4.7.1 FRET analysis can be used to investigate autoinhibition <i>in vitro</i>	90
4.7.2 Generating constructs to investigate KANK1 autoinhibition by FRET analysis <i>in vitro</i>	92
4.7.3 Investigating KANK1 autoinhibition by FRET analysis	92
4.8 Attempting X-ray crystallography to determine the structure of the interaction between the KN and ankyrin repeat domains of KANK.....	94
4.8.1 KN domain peptide presence increases frequency and quality of ankyrin repeat domain crystallisation despite no observed density for the peptide.....	96
4.8.2 Designing a “mini-KANK” for structural experiments	98
4.9 Discussion	101
CHAPTER 5: IDENTIFYING A NOVEL LIGAND FOR KANK1 IDR2	105
5.1 Overview.....	105
5.2 Large unstructured regions within KANK1	105
5.3 Dynein light chain proteins: the LC8 family.....	108

5.3.1 An introduction to dynein light chain proteins	108
5.3.2 LC8 proteins exist as pH-dependent dimers	109
5.3.3 LC8 as a hub protein	111
5.4 Selecting which LC8 family protein to use for investigating an interaction with KANK1.....	112
5.5 pH affects the oligomeric state of DYNLL2	112
5.6 Using triple resonance NMR to assign DYNLL2 residues at pH 7.5.....	114
5.6.1 Overview of backbone assignment theory using triple resonance NMR	114
5.6.2 NMR assignment of DYNLL2 backbone	115
5.6.3 Chemical shift indexing validates the structural information from the backbone assignment of DYNLL2	118
5.7 DYNLL2 binds KANK1 699-716	119
5.7.1 Designing a LC8b-binding KANK1 peptide	119
5.7.2 Using NMR titrations to test an interaction between DYNLL2 and KANK1 699-716.....	121
5.7.3 Mapping chemical shift perturbations of DYNLL2 residues upon KANK1 interaction	123
5.8 Discussion	128
CHAPTER 6: CONCLUSIONS	131
6.1 Summary	131
6.2 Potential limitations	134
6.3 Future work.....	135
CHAPTER 7: REFERENCES	137
APPENDIX	163

LIST OF FIGURES

FIGURE 1.1: THE FOUR MAJOR FAMILIES OF CELL ADHESION MOLECULES.	2
FIGURE 1.2: INTEGRIN A-B HETERODIMERS.	5
FIGURE 1.3: INTEGRIN ACTIVATION STATES.	ERROR! BOOKMARK NOT DEFINED.
FIGURE 1.4: INTEGRIN ADHESION COMPLEXES.	9
FIGURE 1.5: THE NANOSCALE ARCHITECTURE OF FOCAL ADHESIONS (KANCHANAWONG ET AL, 2010).	10
FIGURE 1.6: TALIN STRUCTURE AND LIGANDS.	13
FIGURE 1.7: TALIN AUTOINHIBITION.	15
FIGURE 1.8: THE DYNAMIC INSTABILITY OF MICROTUBULES.	17
FIGURE 1.9: THE CORTICAL MICROTUBULE STABILISING COMPLEX.	21
FIGURE 1.10: THE KANK FAMILY OF PROTEINS.	22
FIGURE 1.11: DISEASE-CAUSING MUTATIONS IN KANK.	23
FIGURE 1.12: KIF21A INTERACTS WITH THE ANKYRIN REPEAT DOMAIN OF KANK1.	26
FIGURE 1.13: CMSC PERTURBATION CAUSES DEFECTIVE MICROTUBULE ORGANISATION.	27
FIGURE 1.14: STRUCTURAL MODEL OF THE INTERACTION BETWEEN TALIN R7 AND KANK1 KN (BOUCHET ET AL, 2016).	27
FIGURE 1.15: KANK PROTEINS DIRECTLY FACILITATE CROSSTALK BETWEEN FOCAL ADHESIONS AND THE CMSC.	28
FIGURE 2.1: 1H 15N TROSY PEAK SELECTION FOR A LARGE PROTEIN.	47
FIGURE 3.1: DETERMINING THE KANK1 ANKYRIN REPEAT DOMAIN BOUNDARIES.	54
FIGURE 3.2: CIRCULAR DICHROISM REVEALS SOME UNIQUE PROPERTIES OF KANK ANKYRIN REPEAT DOMAIN IN DIFFERENT ISOFORMS.	55
FIGURE 3.3: KANK3 LOCALISES TO THE TIPS OF FILOPODIA.	56
FIGURE 3.4: THE CRYSTAL STRUCTURE OF THE KANK3 ANKYRIN REPEAT DOMAIN.	57
FIGURE 3.5: ALL HUMAN KANK ANKYRIN REPEAT DOMAINS BIND TO KIF21A 1144- 1173.	59
FIGURE 3.6: CIRCULAR DICHROISM INDICATES THAT KANK2 ANKYRIN REPEATS WILDTYPE AND ALA670VAL MUTANT HAVE SIMILAR STRUCTURAL PROPERTIES.	61
FIGURE 3.7: THE CRYSTAL STRUCTURE OF THE KANK2 ANKYRIN REPEAT DOMAIN CONTAINING THE DISEASE-CAUSING ALA670VAL MUTATION.	62
FIGURE 3.8: WILDTYPE AND ALA670VAL MUTATION VERSIONS OF THE KANK2 ANKYRIN REPEAT DOMAIN BIND SIMILARLY TO KIF21A 1144-1173.	64
FIGURE 3.9: CIRCULAR DICHROISM INDICATES THAT KANK4 ANKYRIN REPEATS WILDTYPE AND TYR801HIS MUTANT HAVE VARYING STRUCTURAL PROPERTIES.	66
FIGURE 3.10: SEC-MALS ANALYSIS OF WILDTYPE AND TYR801HIS MUTANT KANK4 ANKYRIN REPEAT DOMAIN REVEALS DIFFERENCES IN THE OLIGOMERIC STATES.	67
FIGURE 3.11: SECONDARY STRUCTURE PREDICTION INDICATES THAT TYR801 IS LOCATED CENTRALLY IN AN A-HELIX.	68

FIGURE 3.12: WILDTYPE AND TYR801HIS MUTANT VERSIONS OF THE KANK4 ANKYRIN REPEAT DOMAIN BIND SLIGHTLY DIFFERENTLY TO KIF21A 1144-1173.....	69
FIGURE 3.13: THE KANK ANKYRIN REPEAT DOMAINS DO NOT BIND TO SCRIBBLE 464-480.....	71
FIGURE 4.1: KANK1 KN DOMAIN PEPTIDES USED IN THIS WORK.	76
FIGURE 4.2: KANK1 ANKYRIN REPEATS AND TALIN R7R8 BIND DIFFERENTLY TO KANK1 KN DOMAIN.	77
FIGURE 4.3: TROSY NMR SPECTRUM OF THE KANK1 ANKYRIN REPEAT DOMAIN.....	78
FIGURE 4.4: TROSY NMR TITRATION SPECTRA OF KANK1 ANKYRIN REPEAT DOMAIN PROTEIN WITH KANK1 KN DOMAIN PEPTIDES.	79
FIGURE 4.5: KANK1 KN(30-73Δ60-68) PEPTIDE INTERACTS WITH TALIN1 R7R8 BUT DOES NOT INTERACT WITH KANK1 ANKYRIN REPEAT DOMAIN.	80
FIGURE 4.6: TROSY NMR TITRATION SPECTRUM OF KANK1 ANKYRIN REPEAT DOMAIN PROTEIN WITH KANK1 KN(30-73 Δ60-68) PEPTIDE.	81
FIGURE 4.7: MULTIPLE SEQUENCE ALIGNMENTS OF KANK KN DOMAINS.	82
FIGURE 4.8: OPTIMISATION OF KANK4 KN DOMAIN SYNTHETIC PEPTIDES.	84
FIGURE 4.9: AN INTERACTION BETWEEN THE ANKYRIN REPEAT DOMAIN AND THE KN DOMAIN CAN BE OBSERVED IN ALL FOUR KANK ISOFORMS.	85
FIGURE 4.10: ALL FOUR KANK FAMILY PROTEINS INTERACT WITH THE KN DOMAIN WITH SOME VARIATION.	86
FIGURE 4.11: WILDTYPE AND ALA670VAL MUTANT KANK2 ANKYRIN REPEAT DOMAIN PROTEINS INTERACT SIMILARLY WITH KANK1 AND KANK2 KN DOMAIN PEPTIDES.	88
FIGURE 4.12: WILDTYPE AND TRY801HIS MUTANT KANK4 ANKYRIN REPEAT DOMAIN PROTEINS VARY SLIGHTLY IN THEIR INTERACTIONS WITH KANK1 AND KANK4 KN DOMAIN PEPTIDES.	90
FIGURE 4.13: A SCHEMATIC OF A FRET EXPERIMENT USING KANK1.	91
FIGURE 4.14: ALTHOUGH A QUENCH IN EMITTED MCERULELAN-KN FLUORESCENCE IS OBSERVED, NO FRET IS DETECTED WHEN TESTED WITH MVENUS-ANKYRIN REPEATS.....	93
FIGURE 4.15: REFINED STRUCTURES OF KANK2 ANKYRIN REPEATS CRYSTALLISED ALONE AND IN PRESENCE OF KANK2 KN 31-69 PEPTIDE.....	97
FIGURE 4.16: PUBLISHED STRUCTURES OF KANK2 ANKYRIN REPEATS CRYSTALLISED ALONE AND CRYSTALLISED WITH KIF21A PEPTIDE.....	98
FIGURE 4.17: POTENTIAL MINI-KANK OLIGOMERIC STATES.....	100
FIGURE 4.18: A MINI-KANK PROTEIN BINDS TO THE KANK1 KN DOMAIN EQUALLY TO THE KANK1 ANKYRIN REPEATS.	101
FIGURE 5.1: SECONDARY STRUCTURE PREDICTION OF KANK1 480-1080 USING PSIPRED.....	106
FIGURE 5.2: EVOLUTIONARY CONSERVATION OF KANK1 IDR2 (501-1080) ESTIMATED USING CONSURF.	107

FIGURE 5.3: THE STRUCTURE OF THE DYNLL1 DIMER.	110
FIGURE 5.4: MULTIPLE SEQUENCE ALIGNMENT OF HUMAN AND D. MELANOGASTER LC8 PROTEINS.	112
FIGURE 5.5: SEC-MALS ANALYSIS OF DYNLL2 AT PH 3 AND PH 7.5.....	113
FIGURE 5.6: AN EXAMPLE STRIP WINDOW IN NMRVIEWJ.	115
FIGURE 5.7: B3 IS UNSTRUCTURED IN APO LC8.	116
FIGURE 5.8: THE BACKBONE ASSIGNMENT OF DYNLL2.....	117
FIGURE 5.9: TALOS-N OUTPUT FOR DYNLL2 IS CONSISTENT WITH EXISTING LC8 FAMILY STRUCTURES.	119
FIGURE 5.10: THE 699-716 REGION IS UNIQUE TO KANK1.....	121
FIGURE 5.11: HSQC NMR TITRATION SPECTRA OF DYNLL2 WITH KANK1 699-716 PEPTIDE.....	122
FIGURE 5.12: THE BACKBONE ASSIGNMENT OF DYNLL2 IN FREE AND BOUND STATES.	124
FIGURE 5.13: MAPPING THE DYNLL2 RESIDUES INVOLVED IN THE INTERACTION WITH KANK1 699-716.	125
FIGURE 5.14: THE STRUCTURE OF LC8 PROTEINS BOUND TO TQT-CONTAINING RECOGNITION MOTIFS.	127

LIST OF TABLES

TABLE 2.1: LIST OF REAGENTS, EQUIPMENT AND SOFTWARE USED.	30
TABLE 2.2: LIST OF BUFFERS AND MEDIA USED, EXCEPT WHERE OTHERWISE STATED.	31
TABLE 2.3: LIST OF PLASMID CONSTRUCTS USED.	32
TABLE 2.4: OLIGONUCLEOTIDE DESIGN FOR RESTRICTION ENZYME-BASED MOLECULAR CLONING.	34
TABLE 2.5: GENERAL PCR REACTION MIXTURE USING EXPAND HIGH FIDELITY DNA POLYMERASE (ROCHE).	35
TABLE 2.6: GENERAL PCR PROGRAMME FOR RESTRICTION ENZYME-BASED MOLECULAR CLONING USING EXPAND HIGH FIDELITY DNA POLYMERASE (ROCHE).	35
TABLE 2.7: GENERAL RESTRICTION ENZYME DOUBLE DIGEST REACTION MIXTURE.	36
TABLE 2.8: GENERAL LIGATION REACTION MIXTURE USING T4 DNA LIGASE (PROMEGA).....	36
TABLE 2.9: GENERAL SMALL-SCALE RESTRICTION ENZYME DOUBLE DIGEST REACTION MIXTURE.	37
TABLE 2.10: OLIGONUCLEOTIDE DESIGN FOR SITE-DIRECTED MUTAGENESIS.	38
TABLE 2.11: GENERAL PCR REACTION MIXTURE USING PFU DNA POLYMERASE (PROMEGA).....	38
TABLE 2.12: GENERAL PCR PROGRAMME FOR SITE-DIRECTED MUTAGENESIS USING PFU DNA POLYMERASE (PROMEGA).....	38
TABLE 2.13: LIST OF SYNTHETIC PEPTIDES USED.....	44
TABLE 3.1: X-RAY DATA COLLECTION AND REFINEMENT STATISTICS FOR KANK3 ANKYRIN REPEATS.	58
TABLE 3.2: X-RAY DATA COLLECTION AND REFINEMENT STATISTICS FOR KANK2 ANKYRIN REPEATS (ALA670VAL MUTANT).....	63
TABLE 3.3: GFP-KANK1 PROTEOMICS DATA.....	70
TABLE 4.1: GFP-KANK1 PROTEOMICS DATA.....	75
TABLE 4.2: ATTEMPTED CO-CRYSTALLISATION COMBINATIONS FOR KANK ANKYRIN REPEATS IN COMPLEX WITH KANK KN DOMAIN.	94
TABLE 5.1: GFP-KANK1 PROTEOMICS DATA.....	108
TABLE 5.2: POTENTIAL LC8 RECOGNITION MOTIFS PRESENT IN THE KANK1 IDR2.	120

ABBREVIATIONS

[]	Denotes concentration
Å	Ångström
°C	Degrees Celsius
2M9	Minimal media
ABS	Actin-binding site
ANK	Ankyrin repeat domain
BMRB	Biological magnetic resonance bank
BSA	Bovine serum albumin
CaCl ₂	Calcium chloride
CAM	Cell adhesion molecule
CD	Circular dichroism
C-terminal	Carboxy terminal
CMSC	Cortical microtubule stabilising complex
Da, kDa	Dalton, Kilodalton
DD	Dimerisation domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DYNLL1	Human dynein light chain 1
DYNLL2	Human dynein light chain 2
ECM	Extracellular matrix
FA	Focal adhesion
FP	Fluorescence polarisation
g, mg	Gram, milligram
GFP	Green fluorescent protein
HSQC	Heteronuclear single quantum spectroscopy
IDR	Intrinsically disordered region
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KANK	Kidney ankyrin repeat-containing protein

K _d	Dissociation constant
KN	KANK KN domain
L, mL, μ L	Litre, millilitre, microlitre
LB	Lysogeny broth
LC8 family	Dynein light chain protein family
LD motif	Leucine-aspartic acid motif
M, mM, μ M	Molar, millimolar, micromolar
MgCl ₂	Magnesium chloride
MT	Microtubule
N-terminal	Amino terminal
NaCl	Sodium chloride
nm	Nanometre
NMR	Nuclear magnetic resonance
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
rpm	Rotations per minute
RT	Room temperature
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC-MALS	Size exclusion chromatography-multiangle light scattering
TBS	Talin-binding site
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus
TROSY	Transverse relaxation optimised spectroscopy
VBS	Vinculin-binding site
WT	Wildtype

CHAPTER 1: INTRODUCTION

1.1 Cell adhesion

Cell adhesion is the tightly controlled process by which cells can form dynamic attachments with each other and with the surrounding extracellular matrix. This process allows cells of multicellular organisms to detect biological signals and translate them into appropriate biological outputs, dictating the cellular activity.

Specific transmembrane proteins at the cell surface, cell adhesion molecules (CAMs), are required for the formation of different adhesive structures. CAMs can be divided into four major families: selectins, cadherins and immunoglobulin superfamily, which are all primarily involved in cell-cell adhesion, and integrins which are primarily involved in cell-extracellular matrix adhesion (Figure 1.1). The range of structures formed by these CAMs contributes to a range of signal transduction pathways with links to an extensive list of cellular processes, including tissue development, cell polarity and cell migration (Gumbiner, 1996).

Disruption of cell adhesion can have a broad range of consequences, with loss of tight regulation resulting in a variety of disease phenotypes. This most notably includes cancer, where cancer progression and dysregulated epithelial-mesenchymal transition can be driven by the loss of adhesion regulation, causing abnormal cell migration (Witz, 2008; Desgrosellier & Cheresch, 2010; Hamidi & Ivaska, 2018). Other pathological states linked to aberrant cell adhesion include rheumatoid arthritis, associated with the migration of dendritic cells as part of the immune response (Worbs *et al*, 2017), and neuropathies including Alzheimer's disease (Nielsen & Wennström, 2012).

Additionally, a variety of pathogens have been identified which target adhesion sites by mimicking components of adhesive structures, enabling their invasion into the host cell. Examples include the *Shigella flexneri* invasion, IpaA (Izard *et al*, 2006), and *Chlamydia caviae* effector protein, TarP (Whitewood *et al*, 2018).

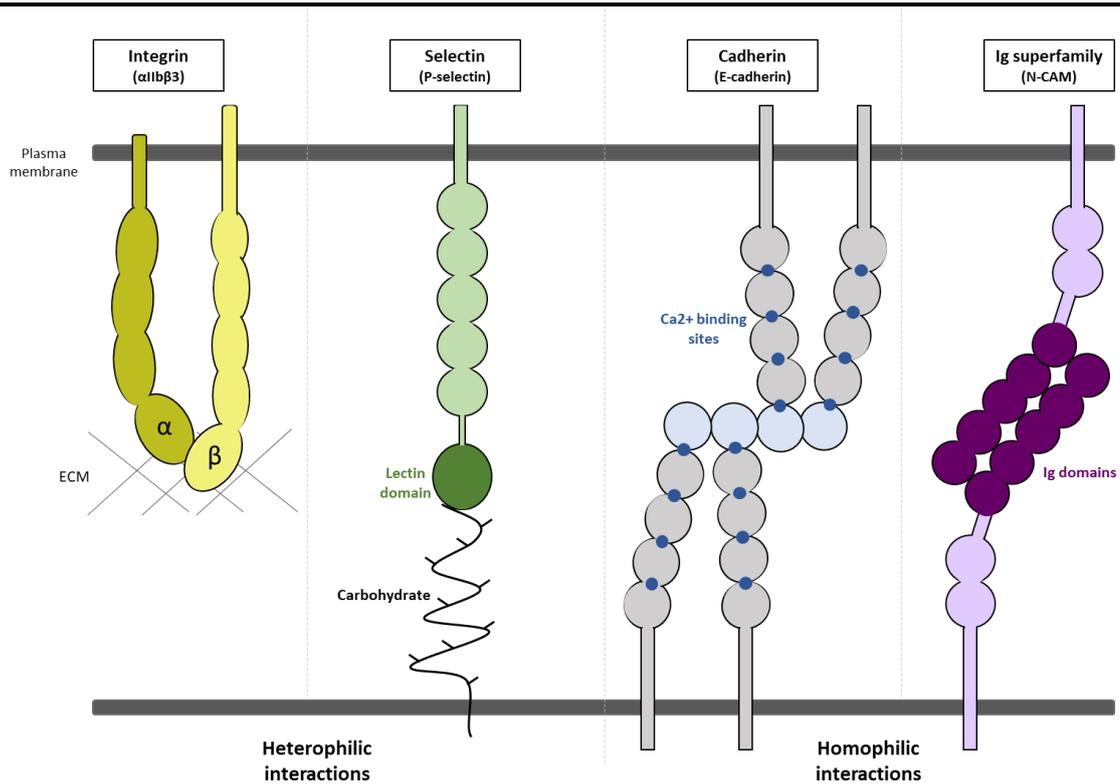


Figure 1.1: The four major families of cell adhesion molecules.

Integrins, selectins, cadherins and the immunoglobulin superfamily proteins are the major CAMs and allow for the formation of different adhesive structures. These can be further grouped by whether they form heterophilic or homophilic interactions.

1.1.1 Cell-cell adhesion

Selectins are heterophilic CAMs which mediate transient cell-cell adhesions between leukocytes and endothelial cells or platelets via binding to carbohydrates on the surface of another cell. The three members of the selectin family are L-selectin, expressed on leukocytes, E-selectin, expressed on endothelial cells, and P-selectin, expressed on leukocytes and platelets. Each of the selectins possesses a Ca^{2+} -dependent lectin domain at the N-terminus, highly conserved between species, which is responsible for carbohydrate binding. The most well-characterised function of selectins is their role in leukocyte extravasation: the process by which leukocytes migrate towards the site of infection (Muller, 2013). Selectins bind their ligands with very fast on and off rates (Alon *et al*, 1995), enabling the initial capture of a leukocyte and promoting its “rolling” along a blood vessel wall in the direction of the blood flow. E-selectin is involved in “slow rolling”, allowing the leukocyte to slowly arrest before it strengthens its adhesions and spreads, ultimately facilitating diapedesis (Ley *et al*, 2007).

Cadherins are homophilic CAMs involved in the formation of cell-cell adhesive structures including adherens junctions and desmosomes. Like selectins, cadherins also function in a calcium-dependent manner and all members of the cadherin superfamily share an extracellular Ca^{2+} -binding domain. Based on sequence similarity, several subfamilies of cadherins exist, including classical, desmosomal and proto-cadherins (Hulpiau & van Roy, 2009). Type I and type II classical cadherins were named according to the tissues in which they were initially identified and include E-cadherin (first identified in epithelial tissue), N-cadherin (first identified in neural tissue), and VE-cadherin (first identified in vascular epithelial tissue) though expression is not limited to these tissues (Saito *et al*, 2012). Desmosomal cadherins, comprising desmogleins and desmocollins, are primarily expressed on epithelial cells and cardiac muscle cells (Kowalczyk & Green, 2013). Adherens junctions, where classical cadherins mediate the adhesion, are formed via linking cytoskeletal actin. This is facilitated by the high affinity, 1:1 interaction of cytoplasmic cadherin with β -catenin which, in turn, forms a weaker-affinity bond with α -catenin (Hartsock & Nelson, 2008). This complex can interact with the actin cytoskeleton, directly through α -catenin itself or indirectly via actin-binding proteins such as vinculin (Seddiki *et al*, 2018). These adhesions are dynamic, mechanosensitive structures that respond to environmental changes, including force. Alternately, desmosomes, the junction formed by desmosomal cadherins, link adjoining cells via intermediate filaments. Here, intracellular desmoplakin directly links desmogleins/desmocollins to the intermediate filaments (Delva *et al*, 2009).

The immunoglobulin superfamily (IgSF) proteins are the most diverse group of CAMs, but all members possess at least one immunoglobulin-like extracellular domain: two antiparallel β -sheets stabilised by the presence of a disulphide bridge (Lesk & Chothia, 1982; Ioerger *et al*, 1999). Unlike selectins and cadherins, IgSF proteins function in a calcium-independent manner and have a more diverse range of ligands – they can be involved in both homophilic and heterophilic modes of binding, the latter of which occurs less frequently but typical ligands include carbohydrates and integrins (Barclay, 2003). The range of proteins belonging to the IgSF of course have quite diverse roles, and many review articles cover these – of note, these include the roles of IgSF proteins

in neuronal development, inflammation, and metastasis (Leshchyns'ka & Sytnyk, 2016; Golias *et al*, 2011; Wai Wong *et al*, 2012).

1.2 Integrin-mediated cell-extracellular matrix adhesions

1.2.1 The extracellular matrix

The extracellular matrix (ECM) is a complex, non-cellular scaffold which is present in all tissues to provide structural and biochemical support for the surrounding cells. ECM components are secreted by cells and generally consist of water, proteoglycans and fibrous proteins including elastins, fibronectins, laminins and collagens, the latter of which are typically the most abundant (Theocharis *et al*, 2016). While all these components are usually present, the dynamic surrounding microenvironment leads to unique, tissue-specific ECM compositions and topology, and this is determined during embryonic development (Rozario & DeSimone, 2010). In addition to the physical scaffold provided by the ECMs, anchoring cells in precise locations or providing a platform along which they can migrate, they also have vital roles in regulating a plethora of major cellular activity including tissue development, proliferation and wound healing in addition to migration. The reliance upon the ECM to regulate myriad essential cellular processes means that disruption can lead to a range of disease states – this is most commonly a result of the consequent disruption to adhesion itself, but conditions including angiogenesis and fibrosis can also occur due to changes in ECM composition (Vong & Kalluri, 2011; Liakouli *et al*, 2018).

1.2.2 Cell-extracellular matrix adhesions

In order to allow highly coordinated directional movement, and for adhesion itself, cells need to adhere to and detach from the ECM in perfect synchrony: as the cell crawls forward along the ECM via adhesion formation at the front, adhesions underneath the cell body and at the rear of the cell undergo tightly regulated disassembly. The directional migration of a cell thus requires coordination of cytoskeletal rearrangement and cellular polarity, a highly complex process allowing the force generation required to drive dynamic remodelling of adhesions connecting the cell to the ECM.

A range of integrin-based cell-extracellular matrix adhesion complexes are discussed in further detail in Section 1.4.

large ectodomain, a single transmembrane domain and a short cytoplasmic tail domain – the specific domains vary (Figure 1.3b). The ectodomain of the larger α -subunits contains either four or five subdomains depending on the isoform, whereas the β -subunits contains seven subdomains including a β -I domain which is homologous to the α -I domain of the α -subunit (Campbell & Humphries, 2011). Typically, the β -subunit is considered to be more flexible (Xie *et al*, 2010). Each subunit possesses a highly conserved transmembrane domain, typically consisting of a single helix, linking to a short cytoplasmic tail. The exception to this is β 4 wherein the cytoplasmic domain is much larger (Suzuki & Naitoh, 1990) and facilitates its preference for interacting with the intermediate filament network in contrast to other β -chains which preferentially interact with the actin network (Mercurio *et al*, 2001). The cytoplasmic tails appear to be extremely flexible and extended when there is no ligand bound, enabling their ability to form hub interactions (Wegener & Campbell, 2008) and, for both subunits, the cytoplasmic tails can bind a range of proteins. Of note, the β -cytoplasmic tail contains two NPxY-like motifs (Asn-Pro-x-Tyr), crucial for binding ligands containing phosphotyrosine-binding (PTB) domains which induce integrin activation (Calderwood *et al*, 2002; Uhlik *et al*, 2005). Ultimately, this results in huge protein complexes coalescing onto the cytoplasmic tails to give rise to large signalling complexes upon integrin activation.

This intricate structure facilitates the characteristic bidirectional signalling ability of integrins wherein signals that regulate countless crucial cellular activities can be transmitted across the plasma membrane.

1.3.3 Integrin activation

The concept of integrin activation has been discussed for decades, with the initial publication observing regulation of integrins via conformational changes being published in 1990 (Frelinger *et al*, 1990). Integrins can be activated via “inside-out” or “outside-in” signalling mechanisms. Here, “inside-out” activation, triggered by cytoplasmic signals, will be discussed.

Integrins have three major conformations: a bent, closed (autoinhibited) form, an extended form with the headpiece of the ectodomain still closed, and an extended form

with the headpiece open and the α - and β -cytoplasmic tails a greater distance apart (active) (Shimaoka *et al*, 2002; Li *et al*, 2017) (Figure 1.3).

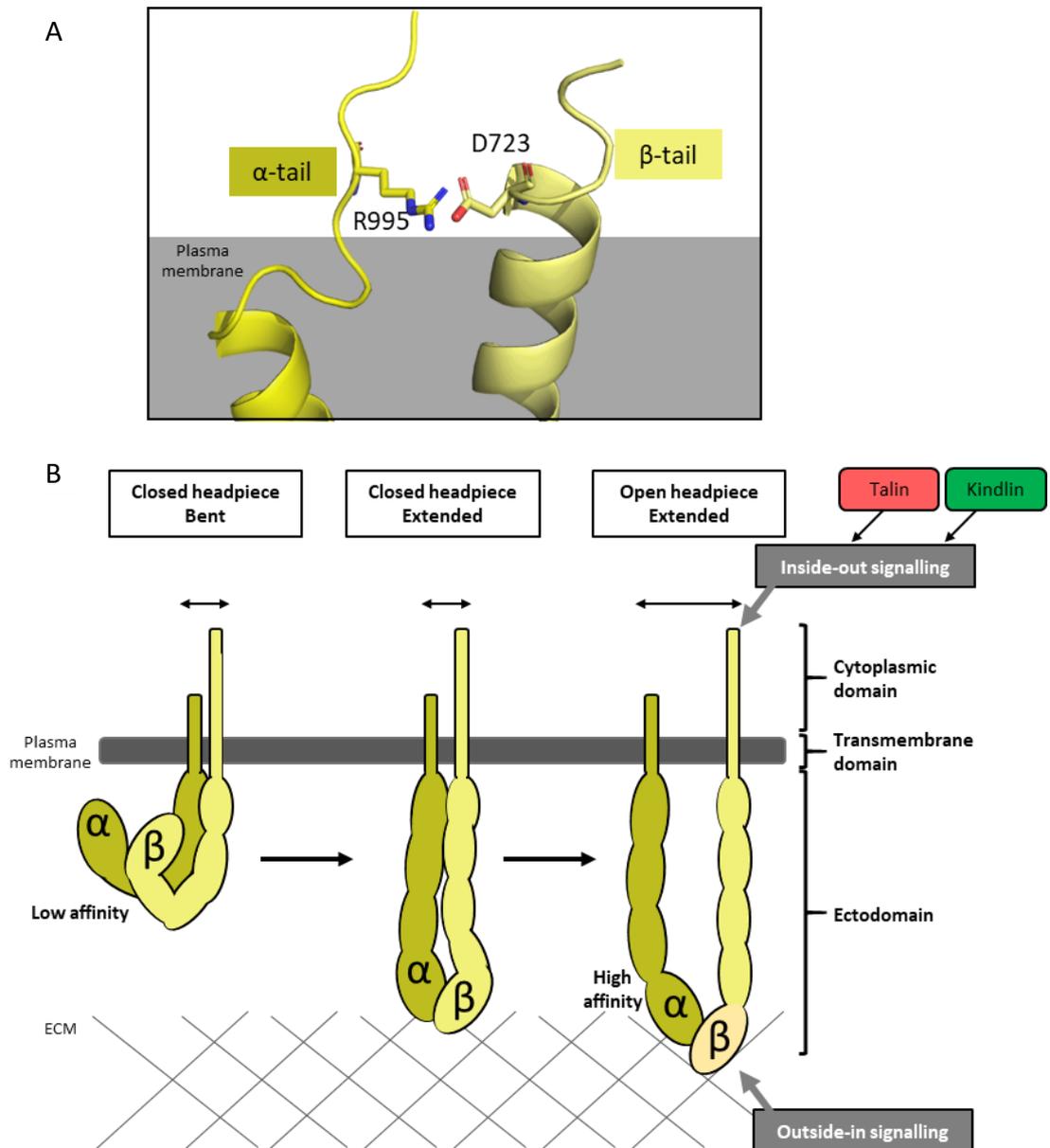


Figure 1.3: Integrin activation states.

(A) The structure of the complex between the transmembrane domains of integrin α IIb and β 3 [PDB ID: 2K9 (Lau *et al*, 2009)] showing the salt bridge formed which maintains the integrin in the autoinhibited state. (B) The three activation states of integrin: the inactive, autoinhibited form, the partially active form with unstable affinity for extracellular ligands, and the fully active form which has a high affinity for extracellular ligands.

Autoinhibition of integrins is maintained via an electrostatic interaction between the two cytoplasmic tails: for example, in α IIb β 3 integrin, Arg995 in the α IIb tail binds Asp723 in the β 3 tail (Figure 1.3a) (Anthis & Campbell, 2011). This crucial salt bridge maintains the low-affinity state, holding the legs of each subunit together and

preventing their separation (Hughes *et al*, 1996; Vinogradova *et al*, 2002; Kim *et al*, 2009; Lau *et al*, 2009).

The process of activation, relieving autoinhibition, causes extensive structural changes in integrin conformation resulting in the extended, open conformation wherein the ectodomains are subject to large-scale reorganisation (Ye *et al*, 2012) and both the transmembrane regions and cytoplasmic tail domains are further separated (Kim *et al*, 2003). This relief of integrin autoinhibition and driving of tail separation is induced by interactions between its β -cytoplasmic tail and the head domain of the adaptor protein talin (see section 1.5) (Calderwood *et al*, 1999). The crystal structure of the region of talin containing a PTB-like domain interacting with the cytoplasmic tail of integrin β 1d (PDB ID: 3G9W (Anthis *et al*, 2009)) indicates that, in addition to this talin-mediated activation breaking the salt bridge between the two cytoplasmic tails, an alternate salt bridge is formed via a conserved basic residue in talin. Thus, with some involvement of the FERM domain-containing protein kindlin as a co-activator (Moser *et al*, 2008; Calderwood *et al*, 2013), active talin is able to relieve integrin autoinhibition and drive conversion to the active conformation, exposing surfaces containing cryptic binding sites for further integrin ligands, including ECM components (Harburger & Calderwood, 2009).

1.3.4 The consensus integrin adhesome

Despite the structural and functional differences in integrin adhesion complexes (IACs) that can form, the composition of the majority of these structures overlap significantly. The integrin adhesome, the total network of structural and signalling components involved in the regulation of cell-extracellular matrix adhesions, contains greater than 200 proteins and, of these, a consensus integrin adhesome of 60 proteins have been identified to form the core of the integrin adhesion machinery (Thomson, 1992; Horton *et al*, 2015). With the crucial roles played by integrin adhesome components in tissue organisation and cellular activity, mutations in these key proteins are involved in various disease states, with 32 of the 60 core consensus integrin adhesome proteins involved in development and progression of cancer (Winograd-Katz *et al*, 2014).

1.4 Integrin adhesion complexes

Within cells, a variety of adhesion complexes have been identified. These complexes can vary in protein composition, morphology and cellular location and form in response to the specific cellular requirements – for example, whether the cell is forming cell-cell adhesions or forming cell-extracellular matrix adhesions for cell migration. Integrins are crucial for sensing the dynamic biophysical and biochemical properties of the extracellular environment which can differ significantly between tissues. To allow the highly coordinated, directional movement required for efficient cell migration, for example, the dynamic progression of adhesive structures formed is crucial. Here, some of the more widely studied types of IACs will be briefly introduced (Figure 1.4).

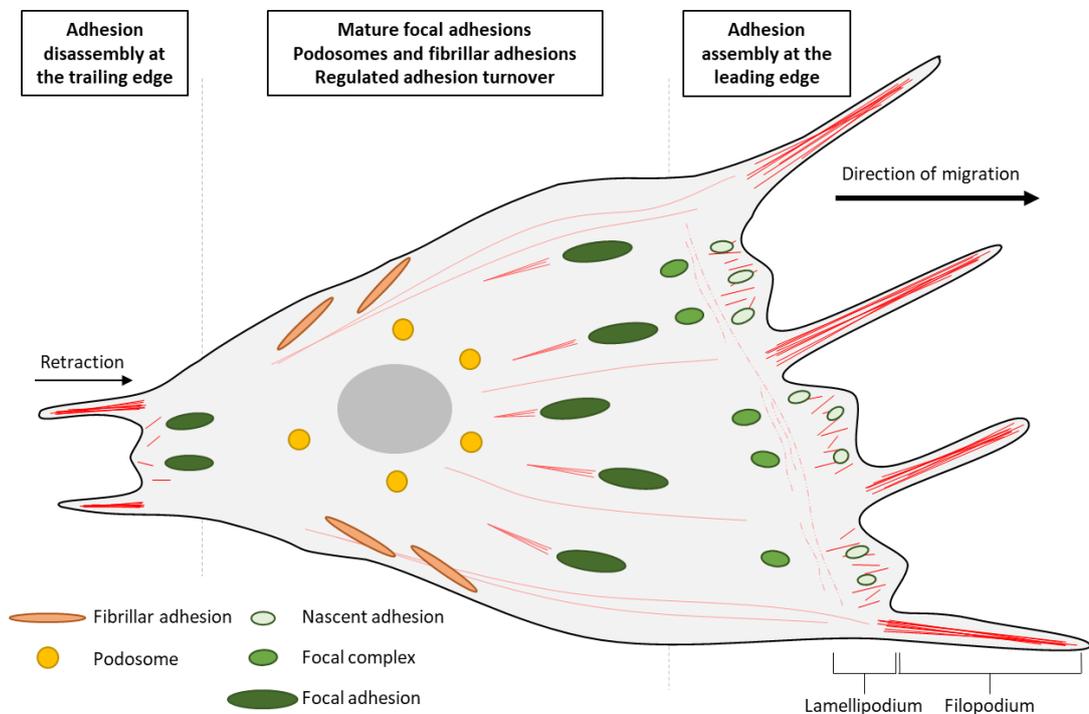


Figure 1.4: Integrin adhesion complexes.

Nascent adhesions form within the lamellipodium at the leading edge of the cell and either rapidly disassemble or mature into focal complexes. Focal complexes further mature into more stable focal adhesions or into fibrillar adhesions, and these structures are localised more centrally within the cell. Podosomes also form within this region. At the trailing edge, the disassembly of integrin-based adhesions allows for retraction at the rear, facilitating forward migration of the cell.

1.4.1 Nascent adhesions

Nascent adhesions are small, rounded, transient structures which form at the leading edge of a migrating cell (Zaidel-Bar *et al*, 2003), caused by rapid actin polymerisation that drives protrusion at the lamellipodium (Theriot & Mitchison, 1991; Webb *et al*, 2002; Ponti *et al*, 2004). These structures have been shown to consist of integrin and talin at the core, with involvement of other signalling proteins including focal adhesion

kinase (FAK) and p130Cas (Vicente-Manzanares & Horwitz, 2011; Changede *et al*, 2015). Although the majority of these rapidly disassemble (Parsons *et al*, 2010), nascent adhesions can mature into larger, more stable adhesive structures including focal complexes and focal adhesions. This maturation relies on mechanical forces (Choi *et al*, 2008; Han *et al*, 2019).

1.4.2 Focal complexes and focal adhesions

While nascent adhesions themselves form independently of mechanical cues, maturation and focal complex formation has been shown to be highly dependent on contractile force generated from crosslinking of actin by myosin II (Choi *et al*, 2008). Focal complexes share the dot-like morphology of nascent adhesions; however, focal complexes are larger with a diameter of $\sim 1 \mu\text{m}$ and are localised further towards the back of the lamellipodium. As increased mechanical force is experienced by the cell, the transient focal complex matures further into a focal adhesion (FA), a larger, more elongated structure with increased stability (Anderson *et al*, 2014). FAs are located close behind the leading edge, typically at the ends of actin bundles or stress fibres, and contain a broader range of proteins than NAs and focal complexes. These components have been proposed to organise into an integrin signalling layer including FAK and paxillin in addition to integrin itself, an intermediate force-transduction layer including talin and vinculin, and an actin-regulatory layer including zyxin, vasodilator-stimulated phosphoprotein (VASP) and α -actinin (Figure 1.5) (Kanchanawong *et al*, 2010). This

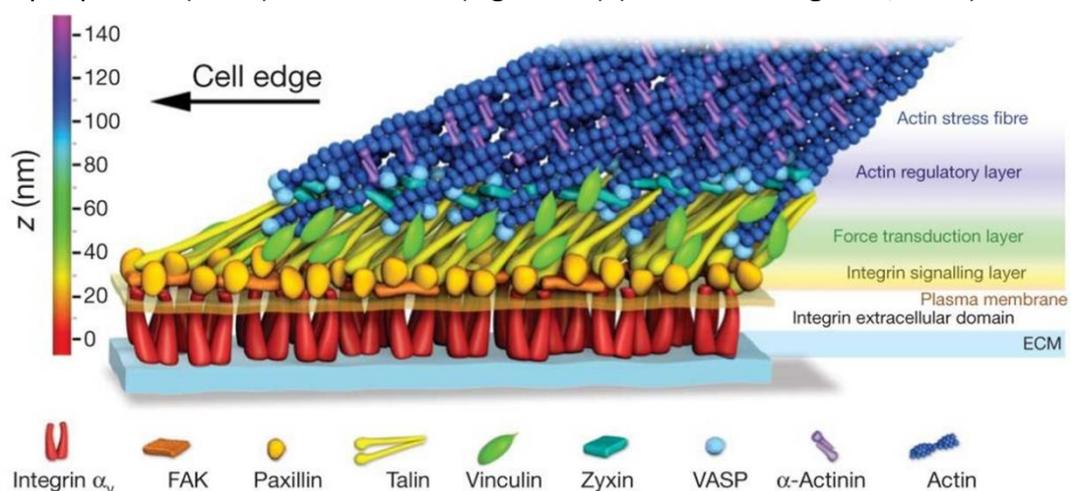


Figure 1.5: The nanoscale architecture of focal adhesions (Kanchanawong *et al*, 2010).

The signalling components of focal adhesions can be organised into distinct functional layers. The integrin signalling layer is located at the membrane and the force transduction layer, where talin is the key mechanosensitive component that helps to recruit further adhesion proteins, allows attachment to cytoskeletal actin.

nanoscale organisation and the mechanosensitive components of FAs allow response to force, translating and integrating mechanical forces into intracellular signals which can dictate cellular activity (Liu *et al*, 2015).

1.4.3 Fibrillar adhesions

Fibrillar adhesions are specialised structures in fibroblasts which have been reported to develop from mature FAs and are more actively involved in ECM organisation due to their role in promoting fibronectin remodelling (Zamir *et al*, 2000; Sottile & Hocking, 2002). This subsequently drives deposition of other components of the ECM including collagens and fibrillin (McDonald *et al*, 1982; Sabatier *et al*, 2009; Barber-Pérez *et al*, 2020). Located more centrally in the migrating cell, these adhesions are even larger, ranging in length from ~1-10 μm , and vary from FAs in their protein composition. Characteristically, fibrillar adhesions dominantly contain integrin isoform $\alpha 5\beta 1$, otherwise known as the fibronectin receptor, and contain tensin and zyxin (Katz *et al*, 2000; Pankov *et al*, 2000).

1.4.4 Podosomes

Podosomes are dot-like structures, often clustered behind the leading edge of the cell, which contain many of the typical proteins associated with cell-matrix adhesion but have distinct structural and functional differences compared to other integrin adhesion complexes (Linder & Kopp, 2005; Murphy & Courtneidge, 2011). Most notably, podosomes contain a unique actin-rich core comprised of filamentous actin and other actin-associated proteins, around which the adhesion proteins localise. Podosomes can form in a variety of cell types but are most prominently found in cells derived from monocytes including macrophages, dendritic cells and osteoclasts. In addition to adhesion, it has been suggested that the primary role of podosomes is to assist with ECM degradation via recruitment of specific proteases (Burgstaller & Gimona, 2005). In osteoclasts, specialised migratory cells found in the bone, podosomes play a key role in facilitating tightly regulated bone resorption, forming a sealing zone which defines the area of resorption (Luxenburg *et al*, 2007; Geblinger *et al*, 2010; Takito *et al*, 2018).

1.4.5 Reticular adhesions

Reticular adhesions are a newly characterised, distinct class of adhesion complex which form during interphase and persist throughout mitosis, allowing the attachment

between the cell and the extracellular matrix to be maintained during the crucial steps of cell rounding, spreading and cytokinesis. These structures differ to other IACs: mediated by $\alpha v\beta 5$ integrins, reticular adhesions lack many components of the consensus integrin adhesome (see Section 1.3.3), are able to form independently of talin and filamentous actin, and proteomic analysis highlights an enrichment in PIP₂-binding proteins (Lock *et al*, 2018). Many of the proteins identified overlap with the clathrin interactome, indicating that reticular adhesions possess a unique composition of both IAC and clathrin-mediated endocytosis components, with greater similarity to the latter which also forms independently of talin (Lock *et al*, 2019).

1.5 Talin

Talins are ~270 kDa mechanosensitive adaptor proteins involved in integrin-mediated adhesions and were first discovered in adhesion plaques in fibroblasts (BurrIDGE & Connell, 1983). Key to cell-ECM adhesions, talin directly links integrin with cytoskeletal actin, the basis of nascent adhesions, and facilitates the recruitment of additional proteins which allow the adhesion to stabilise (Klapholz & Brown, 2017). Additionally, talin has been shown to determine the geometry of focal adhesions (Liu *et al*, 2015) and undergoes tightly regulated, force-dependent conformational changes that dictate which ligands are able to bind at which time (Yan *et al*, 2015; Goult *et al*, 2018).

1.5.1 Talin structure

The two talin isoforms, talin1 and talin2, are encoded by separate genes (Senetar & McCann, 2005) but have 76% sequence identity and the same domain structure (Debrand *et al*, 2009; Gough & Goult, 2018). This consists of a non-canonical linear FERM (4.1 protein, ezrin, radixin, moesin) domain (F0-F3) (Elliott *et al*, 2010) connected, via a linker, to a rod domain comprised of 13 α -helical bundles (R1-R13) followed by a C-terminal dimerisation domain (DD) (Calderwood *et al*, 2013; Goult *et al*, 2013b) (Figure 1.6a). The FERM domain, otherwise known as the head, is considered non-canonical due to its composition of four subdomains arranged in a linear conformation as opposed to the classical FERM domain which contains three subdomains arranged in a compact cloverleaf-like conformation. This FERM domain of talin contains the integrin-activating PTB-like domain in the F3 region, termed integrin-binding site 1 (IBS1), in addition to binding sites for PIP₂, RAP1 and RIAM (Goult *et al*, 2010; Yang *et al*, 2014; Zhu *et al*, 2017)

in addition to various further membrane-associated ligands. The series of 4-helix (R2, R3, R4 and R8) and 5-helix (R1, R5, R6, R7, R9, R10, R11, R12, R13) bundles which comprise the 13 regions of the rod domain of talin can behave like individual mechanochemical switches (Goult *et al*, 2018), stretching and unfolding at specific levels of applied mechanical force. The rod domain contains binding sites for many talin ligands (Figure 1.6B) including two actin-binding sites (ABS), ABS2 spanning from R4-R8 and ABS3 at R13-DD, in addition to 11 vinculin-binding sites and three RIAM-binding sites. Within the rod domain, the R7-R8 fold has the greatest number of known ligands, with R8 in its folded state being particularly active as a scaffold. Folded R8 has binding sites for some of the most well-categorised talin ligands including actin, RIAM and DLC1 (Hemmings *et al*, 1996; Li *et al*, 2011; Goult *et al*, 2013b; Zacharchenko *et al*, 2016; Gough & Goult, 2018). This is interesting due to the very particular structure of this region: while the rest of talin adopts a linear arrangement when fully unfolded, R8 is held out of this path as a result of being inserted into a loop that emerges and terminates in R7 (Figure 1.6a) (Gingras *et al*, 2010). R7, in addition to being part of ABS2, contains a binding site for the microtubule-associated adapter protein KANK. The talin-KANK interaction allows microtubules to target adhesion sites and stimulate adhesion turnover (Bouchet *et al*, 2016; Sun *et al*, 2016a). This is discussed in further detail in Sections 1.7-1.8.

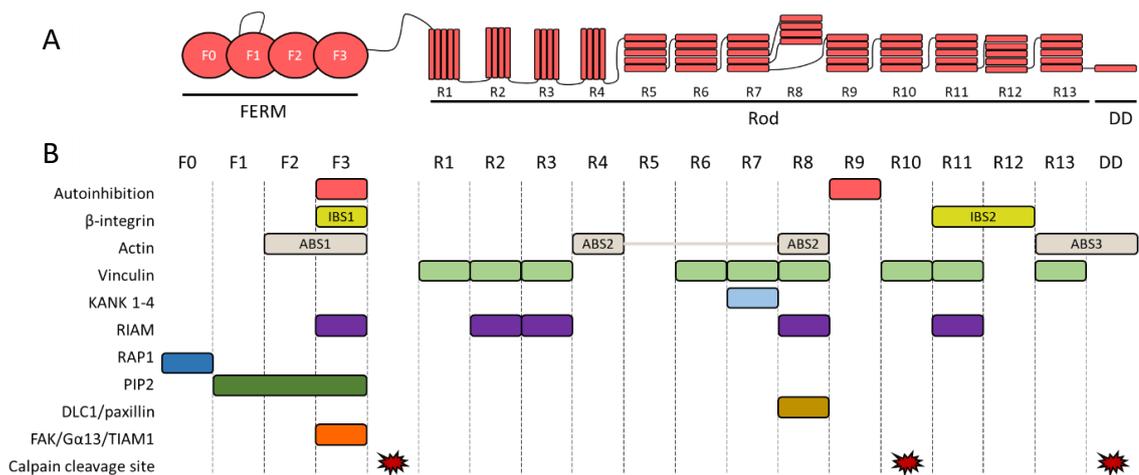


Figure 1.6: Talin structure and ligands.

(A) Talin consists of a non-canonical FERM domain linked to a rod domain comprised of a series of helical bundles (R1-R13). At the C-terminus is a single helix, the dimerisation domain (DD). (B) A summary of the binding sites of some of the most well-characterised talin ligands. Based on figure from Gough & Goult, 2018.

1.5.2 Talin autoinhibition and activation

Like integrin, talin exists in an autoinhibited conformation prior to activation. The autoinhibited, cytosolic talin has a globular conformation which primarily occurs via an interaction between F3 of the FERM domain and R9 of the rod (Goult *et al*, 2009b). This compact structure is facilitated by the formation of a talin homodimer (Molony *et al*, 1987; Goult *et al*, 2013a), formed via the dimerisation domain (DD) at the very C-terminal helix (Gingras *et al*, 2008). The crystal structure of the complex between F2F3 and R9 (PDB ID: 4F7G (Song *et al*, 2012)) and the recent cryo-EM structure of the autoinhibited form of talin (Dedden *et al*, 2019) have provided atomic detail of this primary autoinhibitory interface and confirmed the key role of a particular residue, Glu1770 (in murine talin1 R9), in mediating autoinhibition via a buried salt bridge with Lys318 of F3 (Figure 1.7). Mutations to Glu1770 have been shown to impair the autoinhibitory interaction (Goult *et al*, 2009b) and these mutants, particularly a Glu1770Ala mutant in talin1, have allowed detailed analysis of talin uncoupled from the upstream signalling pathways in a range of biological systems (Kopp *et al*, 2010). For example, the equivalent mutation in *Drosophila melanogaster* results in various defects due to adhesions of greater maturity and stability (Ellis *et al*, 2013) and impaired wound healing has been observed in mouse models (Haage *et al*, 2018).

R9 and integrin bind to the same site on F3, indicating that the autoinhibitory interaction of talin and its integrin binding are mutually exclusive (Goult *et al*, 2009b) and that integrin activation by talin cannot occur until its autoinhibition has been relieved. The mechanisms involved in activating talin and driving its localisation to the plasma membrane are not yet fully understood, though interactions with various ligands appear to be involved. The most well-categorised talin activator is Rap1-interacting adapter molecule (RIAM) (Han *et al*, 2006; Lee *et al*, 2009, 2013), though interactions with PIP2 (Goksoy *et al*, 2008a) and Gα13 switch region 2 (Schiemer *et al*, 2016) have also been shown to have roles in the relief of talin autoinhibition.

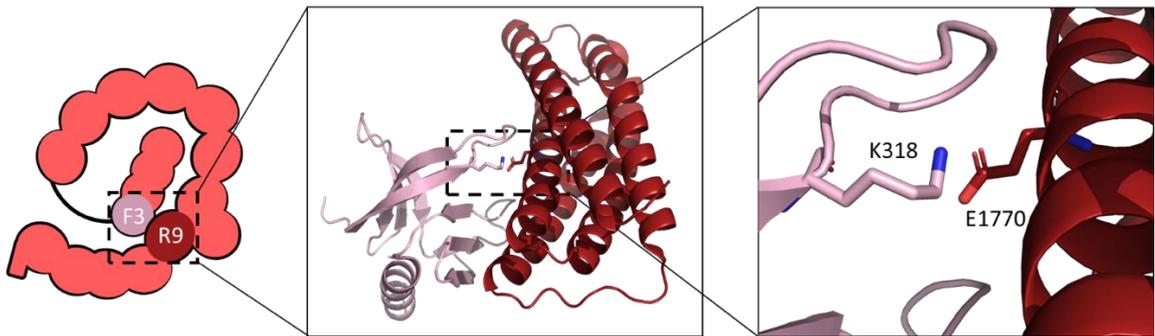


Figure 1.7: Talin autoinhibition.

F3 of the FERM domain and R9 of the rod are the primary interacting domains in talin autoinhibition. K318 and E1770 are the primary residues required for this interaction to take place via the formation of a key buried salt bridge. Crystal structure adapted from PDB ID: 4F7G (Song *et al*, 2012).

1.5.3 Talin as a mechanosensitive signalling hub

The central role of talin in cell adhesion, driving integrin activation and coupling the active integrin to the cytoskeleton, is largely facilitated by its mechanosensitive properties. Actomyosin force is applied and acts on the talin rod upon coupling to cytoskeletal actin, and the linear arrangement of the rod, with the exception of R8, creates a force transmission pathway that can withstand and respond to this tension (Yao *et al*, 2016). 11 vinculin-binding sites have been identified in nine of the subdomains of the rod (Gingras *et al*, 2005) and, although cryptic and buried within a helical core under a lack of tension, are exposed in response to sequential unfolding of the rod domain under mechanical force. Vinculin binding to talin promotes maturation into more stable adhesive structures. While vinculin is only able to bind to unfolded talin, other talin ligands are only able to bind to folded talin. A well-studied example of this is RIAM: where RIAM binding to talin is required for talin activation, this interaction is not required for adhesion maturation. For R3, the least stable of the rod subdomains due to a threonine belt in its hydrophobic core (Goult *et al*, 2013b), a relatively small amount of mechanical force (~ 5 pN) is sufficient to drive RIAM-bound, folded R3 to an unfolded conformation (Yan *et al*, 2015), causing RIAM to dissociate and exposing the cryptic vinculin binding site (Yao *et al*, 2014). Thus, talin behaves as a mechanosensitive signalling hub (Goult *et al*, 2018) where, as individual subdomains of the talin rod respond and unfold at specific levels of tension, this resembles a series of mechanochemical switches that dictate which ligands can bind at any particular time according to the requirements of the cell.

1.6 Microtubules in cell adhesion and migration

1.6.1 Microtubules

In mammals, the major tracks for vesicular transport are microtubules: these are highly dynamic structures with hollow, tube-like morphology and an outer diameter of ~25 nm. Microtubules have key roles in a range of cellular processes including mitosis, wherein the microtubule network is re-organised and facilitates the orientation of chromosomes for the separation of daughter cells in anaphase, in addition maintenance of cellular structure and integrity and transportation of cargo. The primary component of microtubules is tubulin, a heterodimer comprised of α - and β -tubulin subunits, which polymerises and arranges into long head-to-tail arrays to form protofilaments. 13 of these protofilaments assemble laterally to form the cylindrical structure of one microtubule.

Microtubules are intrinsically polar structures with two structurally and functionally distinct ends – a minus end at the free α -subunit which grows slowly, and a rapidly growing plus end at the exposed β -subunit – and this polarity facilitates directional movement. The polymerisation of tubulin dimers allows microtubules to grow (rescue), such as their depolymerisation causes microtubules to shrink (catastrophe). A tubulin dimer, GTP-bound at the β -subunit, is added to a growing microtubule and is followed by hydrolysis of the GTP to GDP. Microtubule growth will continue in the presence of a high concentration of GTP-bound tubulin, which can continue to be added and retain a GTP-cap at the growing end. If, however, GTP hydrolysis occurs more rapidly than the addition of new subunits, this will result in a GDP-bound tubulin dimer at the end of the microtubule. This will lead to disassembly of the more weakly bound terminal heterodimers, resulting in shrinkage (Erickson & O'Brien, 1992; Howard & Hyman, 2009) (Figure 1.8). Microtubules stochastically switch between rescue and catastrophe in rapid cycles as part of a phenomenon referred to as dynamic instability (Mitchison & Kirschner, 1984).

The minus ends of microtubules are anchored in and stabilised by microtubule-organising centres (MTOCs), typically the centrosome and the Golgi in mammalian cells. MTOCs contain γ -tubulin which nucleates new microtubules (Paoletti & Phong, 2007). In many mammalian cells, secretory trafficking thus occurs most widely towards the plus

ends. Preferentially, in migrating cells, the plus ends – and thus traffic – are directed towards the leading edge of the cell (Schmoranzer *et al*, 2003).

Microtubule structure and function is regulated by microtubule-associated proteins (MAPs), of which several groups exist. The two predominant types of motor are kinesins which transport cargo along microtubules in a plus end-directed manner and dyneins which travel in the opposite direction towards the minus end.

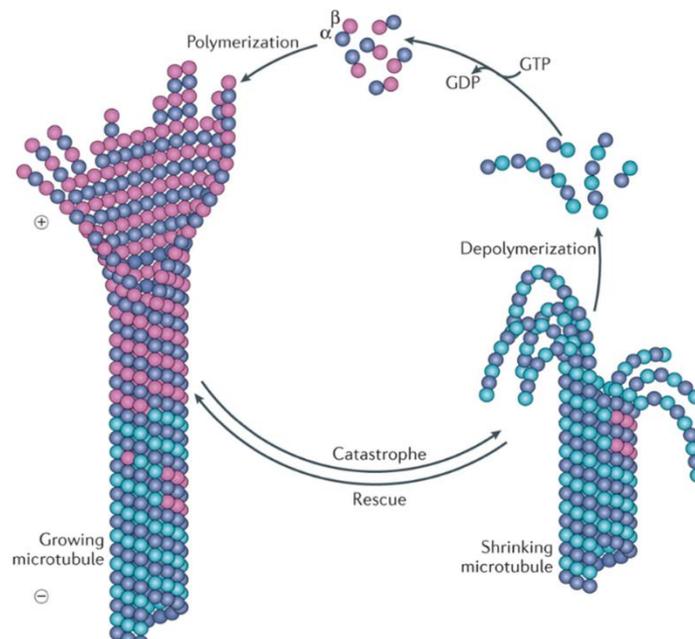


Figure 1.8: The dynamic instability of microtubules.

Microtubules exhibit dynamic instability, a process allowing them to grow via polymerisation of tubulin (rescue) and shrink via the depolymerisation of tubulin (catastrophe). This activity is driven by the hydrolysis of the GTP bound to the β -subunit of a tubulin heterodimer. In the presence of GTP-bound tubulin, microtubules will continue to grow as a result of maintaining a GTP cap. If GTP-bound tubulin is absent, GDP-tubulin will dissociate from the microtubule. Figure adapted from Akhmanova & Steinmetz, 2015.

1.6.2 Microtubule-mediated focal adhesion turnover

While actin is widely considered to be the primary cytoskeletal element associated with cellular adhesion, microtubules have been identified as associated with directional migration for many years (Vasiliev *et al*, 1970) and have more recently emerged as key regulators of adhesions. A proteomics-based study has shown enrichment of microtubule-associated components at active integrin complexes (Byron *et al*, 2015), supporting the observation that there is a dependence upon activation state, with a preference for active integrin and more mature adhesions, for cross-talk between dynamic microtubules and FAs (Hotta *et al*, 2010).

In contrast to the direct coupling of actin to FAs, microtubules instead associate with cell membrane sites in proximity to FAs. Microtubules have various FA-associated roles, including local regulation of signalling molecules such as Rho GTPase, with many functions serving to regulate FA dynamics via the role of microtubules in adhesion disassembly (Stehbens & Wittmann, 2012). For example, by their involvement in clathrin-dependent and clathrin-independent microtubule-mediated endocytosis, vital for integrin recycling (Caswell & Norman, 2006), in addition to serving as tracks for delivery of exocytic carriers (Stehbens *et al*, 2014), microtubules are crucial for adhesion turnover – an essential aspect of cellular migration.

For efficient forward migration of a cell across the ECM, spatial and temporal regulation of adhesion disassembly beneath the cell body and at the rear of the cell is required to facilitate appropriate responses to extracellular signals. While there are likely many components involved in the complex process of turnover, knowledge of the role played by dynamic microtubules in adhesion turnover has increased in prominence and research focus in recent years.

Individual microtubules have been shown to repeatedly target FAs, first observed using live cell microscopy to investigate migrating goldfish fibroblasts (Kaverina *et al*, 1998), and this repetitive targeting correlates to promotion of FA turnover (Kaverina *et al*, 1999). Compared to elsewhere in the cell, a seven times higher probability of microtubule catastrophes has been observed in the vicinity of FAs (Efimov *et al*, 2008).

Various proteins have been identified to be involved in the attachment and stabilisation of microtubules in the vicinity of FAs. Notably, this includes microtubule plus end-tracking proteins (+TIPs), which are highly conserved, specialised MAPs that specifically accumulate at the growing plus ends of microtubules. +TIPs have significant roles in many cellular processes, where they can perform both structural and signalling roles (Akhmanova & Steinmetz, 2008). Examples of +TIPs include adenomatous polyposis coli protein (APC) (Matsumoto *et al*, 2010), actin crosslinking family protein 7 (ACF7) (Kodama *et al*, 2003), and cytoplasmic linker-associated proteins (CLASP1/2) (Mimori-Kiyosue *et al*, 2005; Drabek *et al*, 2006; Stehbens *et al*, 2014) which are recruited to microtubule tips by proteins from the end binding protein (EB) family. EB1 and EB3 jointly function to regulate microtubule dynamics by suppressing catastrophe and

promoting growth, and potentially act to counteract the functions of factors that destabilise microtubules (Komarova *et al*, 2009). Conversely, EB2 does not seem to have a direct role in dictating dynamic instability of microtubules but, importantly, has been shown to recruit mitogen-activated protein kinase kinase kinase kinase (MAP4K4), a microtubule-dependent disassembly factor of FAs. The knockout of MAP4K4 causes the stabilisation of microtubules and causes defects in cell migration (Yue *et al*, 2014).

APC, ACF7 and the CLASPs function to guide growing microtubules to the cell periphery in the vicinity of FAs, where the cortical microtubule stabilising complex anchors the microtubule plus end in the cell cortex.

1.6.3 The cortical microtubule stabilising complex

Microtubules target to and interact with FAs via a cortical microtubule stabilising complex (CMSC), comprised of a network of scaffolding proteins that capture microtubules in the vicinity of adhesions and ultimately prevent their depolymerisation and their overgrowth (van der Vaart *et al*, 2013; Bouchet *et al*, 2016). Microtubules, CLASP-bound at the tips and anchored in the vicinity of FAs, can serve as tracks for transporting exocytic, Rab6-positive vesicles (Grigoriev *et al*, 2007). Membrane type 1-matrix metalloprotease (MT1-MMP) is delivered to the cell membrane as a result of secretory trafficking of these vesicles, allowing the degradation of ECM components (Wang & McNiven, 2012). This can lead to detachment of integrins from the ECM, no longer allowing contractile force to be withstood, ultimately leading to FA turnover (Stehbens *et al*, 2014; Akhmanova & Noordstra, 2017).

CLASPs at the tip of the growing microtubule can interact with the cell cortex in the presence of LL5 β , a phosphatidylinositol 3,4,5 trisphosphate (PIP₃)-interacting protein. This interaction between CLASPs and LL5 β is direct and is influenced by PI3 kinase activity (Lansbergen *et al*, 2006). Prominent components of the CSMC include the adapter proteins liprin- α 1 and liprin- β 1 which form heterodimers via their SAM domain repeats and are involved in synaptic vesicle exocytosis (Asperti *et al*, 2009; Astro *et al*, 2014), ELKS family proteins which are predominantly coiled coil structures that do not directly interact with microtubules but are required for vesicle trafficking in exocytosis (Lansbergen *et al*, 2006; Held & Kaeser, 2018), and KIF21A which is a kinesin-4 that inhibits microtubule polymerisation and depolymerisation at the cell cortex (van der

Vaart *et al*, 2013). Each of these proteins are known to have other roles in addition to their involvement in the CMSC. For example, liprin- α proteins interact directly with ELKS (Ko *et al*, 2003) and these are integral components of the cytomatrix at the active zone (CAZ), the site at which neurotransmitters undergo exocytosis at a neuronal junction in a Ca^{2+} -dependent manner (Gundelfinger & Fejtova, 2012). Further, various CMSC components have been observed to concentrate at podosomes (Proszynski & Sanes, 2013), with LL5 β - and CLASP-containing complexes shown to promote the delivery of acetylcholine receptor-containing vesicles at neuromuscular junctions (Kishi *et al*, 2005; Basu *et al*, 2015). Moreover, liprin- α 1 has been identified as a central component in driving endothelial cell polarity via its role in the internalisation of fibronectin and turnover of integrin α 5 β 1 (Mana *et al*, 2016).

The CMSC is observed as distinct patch-like clusters at the cell cortex which are often enriched at the leading edge of a migrating cell and are localised in very close proximity to, but do not overlap with, FAs (Lansbergen *et al*, 2006; van der Vaart *et al*, 2013; Bouchet & Akhmanova, 2017).

In various studies, it has been demonstrated that crosstalk between FAs and the CMSC is significant. For example, LL5 β and CLASPs have been shown in human keratinocytes to have important roles in FA turnover and ECM degradation, with targeted delivery of MMPs to the vicinity of FAs being promoted by the stabilisation of microtubule tracks by CLASPs (Stehbens *et al*, 2014), which we can now also attribute to the action of the CMSC as a whole.

More recently, KANK proteins have been identified as essential for FA-CMSC crosstalk, providing the link between the actin and microtubule cytoskeletal networks. In addition to directly binding both liprin- β 1 and KIF21A and participating in the clustering of CMSC components at the cell cortex (Figure 1.9) (van der Vaart *et al*, 2013), KANK proteins also directly bind to the rod domain of talin (Bouchet *et al*, 2016; Sun *et al*, 2016a). This is discussed in further detail in section 1.8.3.

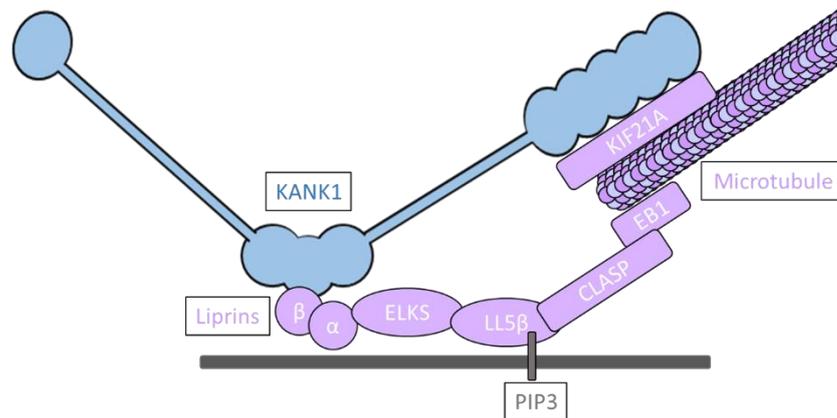


Figure 1.9: The cortical microtubule stabilising complex.

The cortical microtubule stabilising complex is comprised of a range of proteins which act to capture microtubules in proximity to adhesions. Upon binding of KANK with liprin- β 1, CMSC components including ELKS, LL5 β and CLASPs are recruited, ultimately resulting in microtubule capture via KIF21A.

1.7 KANK proteins

1.7.1 An introduction to KANK proteins

The human KANK1 gene (*KANK1*) was originally identified as a potential tumour suppressor gene for renal cell carcinoma, hence its given name (kidney ankyrin repeat-containing protein) (Sarkar *et al*, 2002), and is mapped to chromosome 9 at p24. Two types of KANK1 protein have been identified as a result of alternative first exons of the *KANK1* gene: KANK-S is the “short” protein, while KANK-L is the “long” protein that has an additional 158 residues at the N-terminus. While KANK-S was the initially identified form in renal cell carcinoma cell lines, it was also shown to express in a broad range of tissues and cell lines, seemingly with low-level, non-specific expression. KANK-L, alternatively, exhibits much more tissue-specific expression with notably increased levels in heart and kidney tissues. Interestingly, KANK-L expression was reduced in many renal cancer cell lines, leading to a hypothesis of its involvement in oncogenesis (Wang *et al*, 2005). Reduced expression of KANK-L has also more recently been linked to driving progression in other cancer types, including gastric cancer (Chen *et al*, 2017) and lung cancer (Gu & Zhang, 2018) wherein upregulation of the *KANK1* gene has been shown to inhibit progression. Henceforth, “KANK1” will be used in reference to KANK-L.

In mammals, there are 4 isoforms (KANK1-KANK4) that comprise the KANK family of proteins (Zhu *et al*, 2008). Each of these proteins contains a characteristic N-terminal motif (“KN domain”), coiled-coil domains towards the centre, and 5 ankyrin repeats

(“ankyrin repeat domain”) at the C-terminus, with all of these regions typically linked by unstructured regions (Kakinuma *et al*, 2009) (Figure 1.10). While the KN domain and the ankyrin repeat domain of each KANK is highly conserved between isoforms and species, the more central unstructured and coiled-coil regions show greater variability.

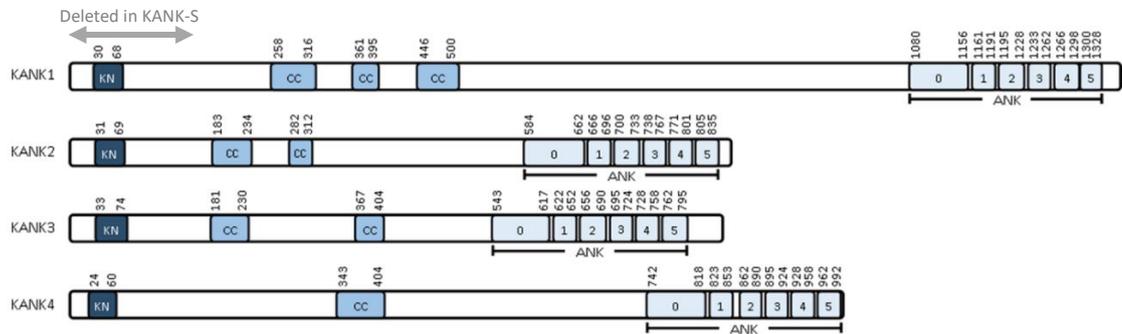


Figure 1.10: The KANK family of proteins.

Domain structure of the four human KANK proteins. KN = KN motif, CC = coiled coil region, ANK = ankyrin repeats.

1.7.2 KANK proteins in disease

In addition to their role in inhibiting cancer progression, KANKs have been linked to a variety of pathologies. The genomic deletion of the *KANK1* gene in humans, at chromosome position 9p24.3 and thus deleting both splice variants, is associated with neurodevelopmental disorders including a form of congenital cerebral palsy that is characterised by intellectual disability and quadriplegia (cerebral palsy with spastic quadriplegia, type 2). While this was originally hypothesised as a maternal imprinting inheritance pattern (Lerer *et al*, 2005), a further study indicates that there may be additional complexity and raises the possibility of random monoallelic expression or that aberrations in just one of the two different transcripts of the *KANK1* gene could be causing the neurodevelopmental disease phenotypes (Vanzo *et al*, 2013).

Further, various disease-causing point mutations in KANKs have been identified (Figure 1.11). A homozygous missense mutation in the *KANK2* gene in humans resulting in Ala670Val has been linked to palmoplantar keratoderma and woolly hair, characteristic symptoms of Naxos and Carvajal syndromes (Ramot *et al*, 2014). Moreover, recessive mutations in *KANK1*, *KANK2* and *KANK4* have been identified to cause a steroid-resistant form of nephrotic syndrome, the most common primary glomerular disease in children which results in end-stage kidney disease and requires dialysis or a kidney transplant (Gee *et al*, 2013). The mutations identified result in Glu454Lys in *KANK1*, Ser181Gly and

Ser684Phe in KANK2, and Tyr801His in KANK4 (Gee *et al*, 2015). This study highlights that KANK proteins are involved in podocyte function and nephrotic syndrome pathogenesis, roles that are evolutionarily conserved in humans, cultured human podocytes, *D. melanogaster* and zebrafish.

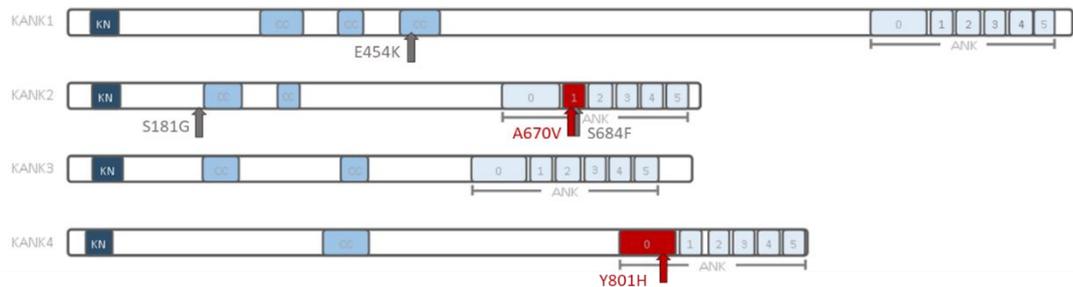


Figure 1.11: Disease-causing mutations in KANK.

Positions of disease-causing point mutations in KANKs. A670V in KANK2 and Y801H in KANK4, present within the ankyrin repeat domain of the respective isoform that they are associated with, are of particular interest. Data from Ramot *et al*, 2014 and Gee *et al*, 2015.

1.8 KANK links focal adhesions to cortical microtubules

1.8.1 KANK proteins and the cytoskeleton

For several years, the involvement of KANK proteins in the regulation of the actin and microtubule cytoskeletons has been observed. An orthologue of KANK1 in *C. elegans*, VAB19, was the subject of the earliest of some of this work. VAB19 functions, in conjunction with integrin INA1-/PAT-3, as a regulator of basement membrane gap formation (Ihara *et al*, 2011), an activity observed during development to facilitate tissue remodelling, cell migration and tumour invasion (Srivastava *et al*, 2007; Clay & Sherwood, 2015). It was found that, while VAB19 was not required for attachment structures to assemble, it was required for the localisation of attachment structures to muscle-adjacent epidermal regions during development and plays a crucial role in the regulation of epidermal actin (Ding *et al*, 2003). In the same study, mutants of VAB19 were shown to ultimately result in an embryonic lethal phenotype as a result of aberrant actin cytoskeletal organisation and defective muscle attachment to the epidermis. EPS-8, a PTB domain-containing and actin-binding adaptor protein in *C. elegans*, colocalises with and binds directly to the ankyrin repeat domain of VAB19, and it was suggested that this could be facilitating epidermal actin organisation (Ding *et al*, 2008).

The *D. rerio* KANK3 homologue, Numb-binding protein (NBP), was found to be essential in embryonic development, particularly during gastrulation and neurulation where the

protein is likely involved in the crucial role of establishing and maintaining cell polarity. NBP interacts with the PTB domain of Numb via a NGGY sequence that is required for Numb binding and is highly conserved amongst the majority of KANK proteins (Boggetti *et al*, 2012). The NGGY sequence is in accordance with the canonical NXXY motif that is found in many sequences known to bind PTB domains (Uhlik *et al*, 2005). NBP was shown to have non-redundant function, with its deficiency causing serious consequences for embryonic development that result in lethality within 48 hours (Boggetti *et al*, 2012).

In *D. melanogaster*, dKANK is the sole KANK orthologue. dKANK localises predominantly to sites of attachment between muscles and tendons. Similar to the aforementioned roles of *C. elegans* and *D. rerio* KANK proteins in embryonic development, dKANK has been suggested to have a role in *D. melanogaster* development, particularly in neurogenesis (Hong *et al*, 2013). A direct interaction between dKANK and end binding protein 1 (EB1) has been identified, an interaction which occurs via a SxIP motif in dKANK which, interestingly, is not evolutionarily well-conserved (Clohisey *et al*, 2014). This interaction is essential for the localisation of dKANK to microtubule plus ends. In contrast to the role of KANK in other the other species mentioned, knockdown of dKANK does not cause severe developmental defects and is compatible with life (Clohisey *et al*, 2014). However, dKANK knockdown has been shown to cause impaired function of cardiac nephrocytes (Gee *et al*, 2015) and disrupted neural phenotypes have been identified as a result of dKANK using RNAi (Sepp *et al*, 2008).

Studies in mammalian cells have highlighted the role of KANK proteins in the regulation of both actin and microtubule cytoskeletal architecture. Typically, the PI3 kinase-Akt signalling pathway, stimulated by insulin and growth factors such as epidermal growth factor, regulates various cellular functions including migration, growth and division (Engelman *et al*, 2006; Hopkins *et al*, 2020). KANK has been indicated to negatively regulate actin stress fibre formation by inhibiting the activation of the small GTPase RhoA, essential for stress fibre formation and for cell migration, resulting from the binding of KANK to 14-3-3 proteins during PI3 kinase-Akt signalling. This interaction between KANK and 14-3-3 occurs as a result of phosphorylation of one or both of two 14-3-3-binding motifs (RXRXX(S/T)) in KANK by Akt, with an indication that the site at Ser167 in KANK-S is dominant (Kakinuma *et al*, 2008). This is consistent with the

significantly reduced actin stress fibre formation observed in murine NIH3T3 cells expressing each KANK family protein (Zhu *et al*, 2008). Additionally, KANK1 directly binds to insulin receptor substrate p53 (IRSp53), inhibiting the interaction between IRSp53 and active Rac1 (Roy *et al*, 2009). This subsequently inhibits downstream Rac1 activity, including Rac1-dependent development of lamellipodia, but does not affect filopodia formation which relies on Cdc42 (Scita *et al*, 2008). Further, overexpression of KANK1 in NIH3T3 cells in this study inhibited both insulin-induced membrane ruffling and integrin-dependent spreading on fibronectin and the outgrowth of neurites induced by IRSp53 (Roy *et al*, 2009). It is important to note that many of the studies in mammalian cells mentioned here focused on the KANK-S form of KANK1 which does not include the first 158 residues, thus excluding the KN domain. It would be interesting to repeat the work and approaches used in these studies to investigate whether the roles of KANK-L (full-length KANK1) and KANK2, KANK3 and KANK4 in these systems are identical or display differences.

1.8.2 The role of KANK proteins in the cortical microtubule stabilising complex

As discussed briefly in section 1.6.3, KANK proteins have been increasingly highlighted in recent years to play integral roles in the cortical microtubule stabilising complex (CMSC). The first of the three coiled-coils towards the centre of KANK1 has been shown to bind directly to liprin- β 1 (van der Vaart *et al*, 2013), which then forms a heterodimer with liprin- α 1. Liprin- β 1 has been observed to associate strongly with both KANK1 and KANK2 and binds via its N-terminus (Luo *et al*, 2016). In addition to the involvement of the coiled-coil domain of KANK, the ankyrin repeat domain is also involved: the ankyrin repeat domain recruits KIF21A (Kakinuma & Kiyama, 2009), a kinesin-4, to the CMSC where it serves to inhibit microtubule catastrophe and suppress growth at the cell edge (Figure 1.12) (van der Vaart *et al*, 2013). The ankyrin repeat domain of KANK2 has also since been identified to interact with KIF21A via residues that are conserved between the two isoforms, while the corresponding residues in KANK3 and KANK4 include variations which appear to prevent their binding (Weng *et al*, 2018; Guo *et al*, 2018). Notably, heterozygous missense mutations in KIF21A which relieve its autoinhibition are linked to congenital fibrosis of the extraocular muscles types 1 (CFEOM1), a genetic disorder with an autosomal dominant inheritance pattern (Heidary *et al*, 2008).

CFEOM1-associated mutations in KIF21A were shown to enhance KANK1 translocation to the membrane and induce alterations in microtubule dynamics, potentially a key element that drives the pathogenesis of CFEOM1 (van der Vaart *et al*, 2013; Kakinuma & Kiyama, 2009).

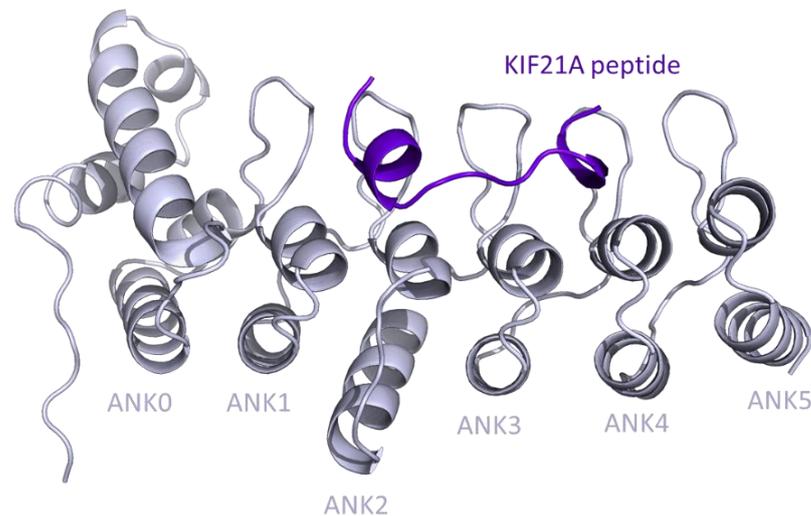


Figure 1.12: KIF21A interacts with the ankyrin repeat domain of KANK1.

A KIF21A peptide spanning 22 residues (purple) is sufficient for binding to the ankyrin repeats of KANK1 (grey) and to the ankyrin repeats of KANK2 (not pictured). A helical conformation is adopted by the peptide upon binding. Crystal structure adapted from PDB ID: 5YBU (Guo *et al*, 2018).

1.8.3 KANK proteins interact with talin to allow FA-CMSC crosstalk

Alongside the central role of KANKs within the CMSC, it has been discovered in recent years that KANK proteins are linked with FAs via a direct interaction with talin. For this interaction, an invariant leucine-aspartic acid motif (LD-motif) (Alam *et al*, 2014; Zacharchenko *et al*, 2016) within the KN domain of each mammalian KANK binds to the rod domain R7 of talin (Bouchet *et al*, 2016; Sun *et al*, 2016a; Yu *et al*, 2019), and it is likely that KANKs of other species will also exhibit this property (see section 1.8.1). This binding occurs via helix addition, a binding mode wherein the KANK KN domain packs against the 5-helix bundle that comprises R7 (Goult *et al*, 2018). Where R7 is included within actin-binding site 2 (ABS2, R4-R8) in talin, F-actin binding to the talin rod has been shown to be interrupted by KANK binding to talin, causing subsequent suppression of force transmission across the adhesion and ultimately resulting in a reduction in cell migration speed due to adhesion sliding (Sun *et al*, 2016a). Following the binding of KANK1 to talin1 within an integrin-based adhesion, further CMSC components (see sections 1.6.3 and 1.8.2) are recruited and cluster around the FA (Bouchet *et al*, 2016). In support of this model, KANK proteins and other CMSC proteins are continuously

observed to localise at the periphery of FAs and not the FA core (Bouchet *et al*, 2016; Rafiq *et al*, 2019; Stubb *et al*, 2019; Paradžik *et al*, 2020).

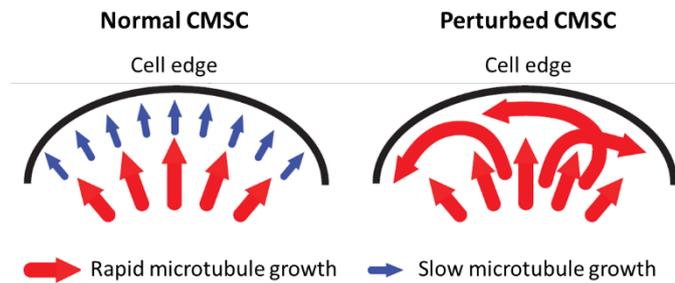


Figure 1.13: CMSC perturbation causes defective microtubule organisation.

A representation of the pattern of microtubule growth observed in HeLa cells and in cells where the CMSC is perturbed. Adapted from Bouchet *et al*, 2016.

Disruption of the KANK1-talin1 interaction in HeLa cells causes perturbed organisation of microtubule plus ends at the cell periphery – microtubule plus end growth occurs more rapidly and becomes less perpendicular to the cell edge in orientation compared to normal conditions (Figure 1.13). A single point mutation in murine talin1, Gly1404Leu (Figure 1.14), was sufficient to induce this perturbation of CMSC macromolecular assembly and thus microtubule dynamics and organisation at the cell periphery in proximity to FAs, without interfering with FA formation (Bouchet *et al*, 2016).

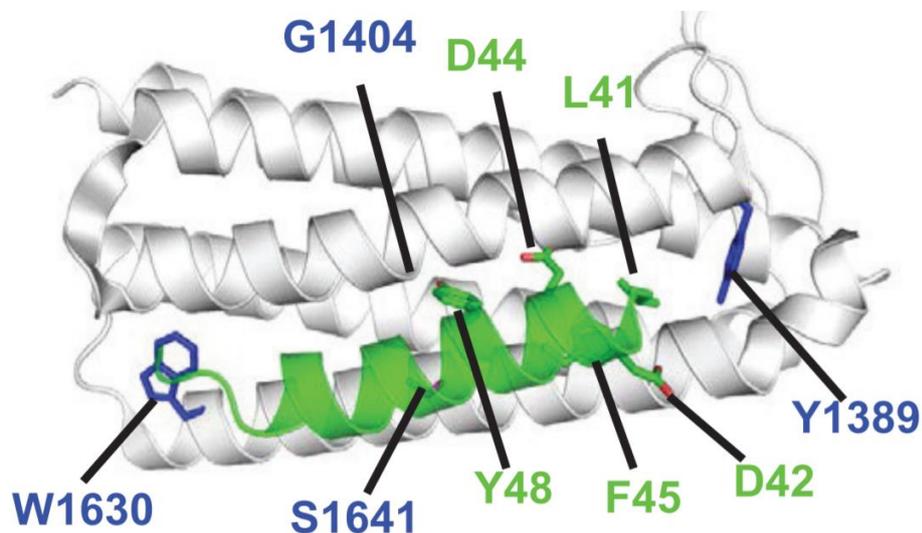


Figure 1.14: Structural model of the interaction between talin R7 and KANK1 KN (Bouchet *et al*, 2016).

Trp1630 and Tyr1389 (blue), two bulky residues, appear to hold two talin1 R7 helices apart to generate the binding surface for the KANK KN domain (green). Glu1404 creates a hydrophobic pocket which allows Tyr48 of the KN domain to dock.

More recently, it has been shown that KANK1 and KANK2 localise around podosomes in addition to FAs. This same study also suggests that the coupling by KANKs of

microtubules to integrin adhesions – both podosomes and FAs, in this case – limits the ability of GEF-H1, a microtubule-associated Rho-GEF known for many years to mediate cross-talk between the actin and microtubule cytoskeletons (Krendel *et al*, 2002), to release from microtubules (Rafiq *et al*, 2019). This suppression of GEF-H1 release and thus its activation causes low Rho-GTP levels and low activity of Rho-associated coiled-coil containing kinases (ROCKs), resulting in low levels of myosin IIA filaments and therefore limited actomyosin contractility in proximity to the adhesions – a condition that restricts FA growth (Seetharaman & Etienne-Manneville, 2019). This has been proposed as a negative feedback loop which prevents excessive actomyosin-induced forces at adhesion sites (Rafiq *et al*, 2019).

Collectively, these studies highlight that KANK proteins have various essential roles in linking focal adhesions to the cortical microtubule network by directly binding to the core adhesion protein, talin, and the microtubule-capturing kinesin-4, KIF21A, in addition to also causing downstream effects in signalling pathways that impact actomyosin contractility at adhesion sites. Ultimately, KANK proteins are key for mediating the attachment and cross-talk between two core cytoskeletal elements: actin filaments and microtubules (Figure 1.15).

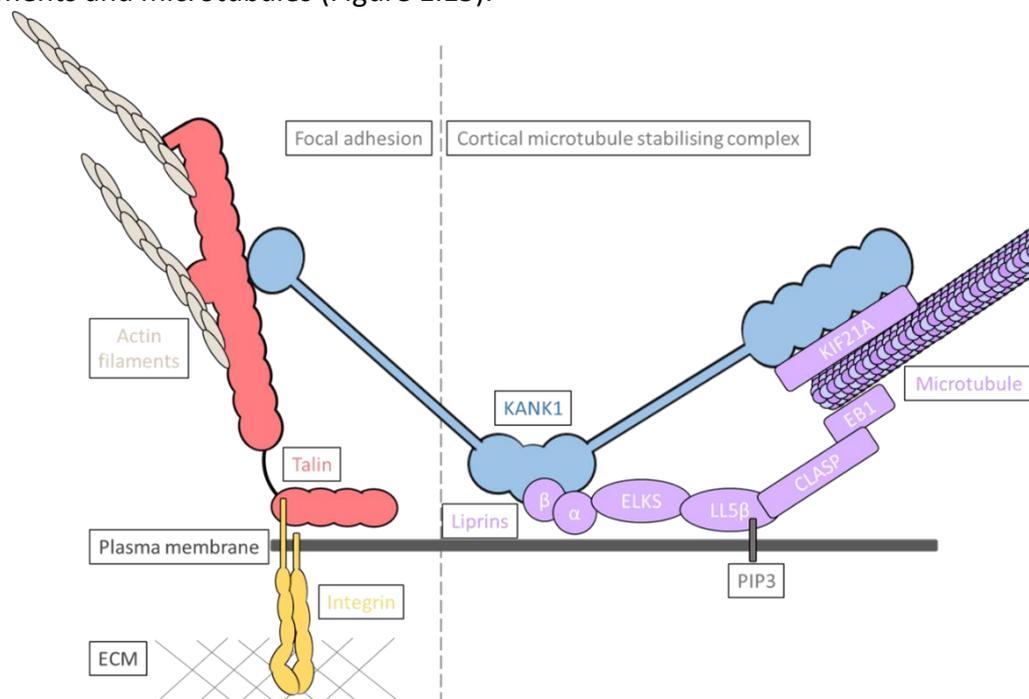


Figure 1.15: KANK proteins directly facilitate crosstalk between focal adhesions and the CMSC.

The cortical microtubule stabilising complex (CMSC, purple) assembles around the periphery of the integrin-based adhesion upon a direct interaction between the KANK KN domain and talin1 rod domain R7. KIF21A binds to the ankyrin repeats of KANK and stabilises microtubules in the vicinity of adhesions. Figure not to scale.

1.9 Project aims

Broadly, this project aims to further the understanding of KANK proteins, including their structure and function in addition to identifying and characterising some novel ligands.

This is discussed across three distinct chapters:

1. **Further characterisation of the ankyrin repeat domain of KANKs:** ankyrin repeat domains generally function as important mediators of protein-protein interactions. While the interaction with KIF21A is known, other ligands of the ankyrin repeat domain of KANK have yet to be identified. This chapter explores striking, previously uncharacterised differences between the ankyrin repeat domains of each mammalian KANK isoform and also aims to further understand some disease-causing point mutations which exist within this domain.
2. **Investigation of an interdomain KANK complex:** the ankyrin repeat domain and KN domain of KANK are both highly conserved structures which already have some key known functions, most notably including allowing their ability to provide a link between the actin and microtubule networks. This chapter identifies an interaction between these two KANK domains, likely a regulatory mechanism, and the binding surface on the KN domain is investigated. Further, the affinity of this interaction amongst all four isoforms of KANK and some disease-causing point mutations is compared.
3. **Identifying a novel ligand for the large intrinsically disordered region in KANK1:** LC8 family proteins, well-categorised hub proteins, are known to interact with disordered regions of proteins and promote their transition into secondary structure elements. In this chapter, the interaction between the LC8 protein DYNLL2 and a LC8-recognition motif discovered within the intrinsically ordered region of KANK1 is identified using NMR titrations and triple resonance experiments.

CHAPTER 2: MATERIALS & METHODS

Methods used throughout this work can also be found in Khan *et al*, 2021 – see Appendix 2.

2.1 Materials

2.1.1 Reagents and equipment

Table 2.1: List of reagents, equipment and software used.

All reagents were obtained from Fisherbrand, Melford, Promega or Sigma-Aldrich unless listed here.

Reagent/Equipment	Source
¹³ C (>99%) glucose	Cambridge Isotope Laboratories
¹⁵ N (>98%) NH ₄ Cl	Cambridge Isotope Laboratories
Autoclave	Prestige Medical
Centrifuge (benchtop)	Eppendorf
Centrifuge	Beckman
Dry block thermostat	Grant Instruments
Micropipettes	Gilson
N50 NanoPhotometer	IMPLEN
Soniprep 150 sonicator	MSE
Spectrophotometer	Varian
Weighing scales	VWR International
Oligonucleotides	Integrated DNA technologies
QIAprep miniprep kit	QIAGEN
QIAquick gel extraction kit	QIAGEN
³ Prime thermal cycler	Techne
ÄKTA Start	GE Healthcare
UNICORN start 1.1 software	GE Healthcare
HisTrap FF 5 ml column	GE Healthcare
HiTrap Q HP 5 ml column	GE Healthcare
HiTrap SP HP 5 ml column	GE Healthcare
Superdex 200 column	GE Healthcare
CLARIOstar plate reader	BMG LABTECH
Prism8 software	GraphPad
600 MHz Avance III spectrometer with CryoProbe	Bruker
J-715 spectropolarimeter	JASCO
JCSG Plus crystallisation screen	Molecular Dimensions
mosquito [®] LCP	TTP Labtech
Viscotek VE 3580 RI detector	Malvern Panalytical
Viscotek SEC-MALS 9	Malvern Panalytical
OmniSEC software	Malvern Panalytical

2.1.2 Buffers and media

Table 2.2: List of buffers and media used, except where otherwise stated.

All buffers were prepared at room temperature using Milli-Q ultrapure water and pH was adjusted using HCl and NaOH, where applicable.

Name	Components	pH
Nickel buffer A	20 mM tris, 500 mM NaCl, 20 mM imidazole	8.0
Nickel buffer B	20 mM tris, 500 mM NaCl, 500 mM imidazole	8.0
Q buffer A	20 mM tris, 50 mM NaCl	8.0
Q buffer B	20 mM tris, 1 M NaCl	8.0
S buffer A	20 mM phosphate (NaH ₂ PO ₄ , Na ₂ HPO ₄), 50 mM NaCl	6.5
S buffer B	20 mM phosphate (NaH ₂ PO ₄ , Na ₂ HPO ₄), 1 M NaCl	6.5
Phosphate-buffered saline (PBS)	100 mM Na ₂ HPO ₄ , 18 mM KH ₂ PO ₄ , 137 mM NaCl, 27 mM KCl	7.4
Phosphate/citrate buffer	100 mM sodium citrate/phosphate, 300 mM sodium sulphate, 5 mM DTT	3 or 7.5
NMR buffer	20 mM phosphate (NaH ₂ PO ₄ , Na ₂ HPO ₄), 50 mM NaCl, 2 mM DTT	6.5
2M9 minimal media solution A	Per L: 12.5 g Na ₂ HPO ₄ · ² H ₂ O, 7.5 g KH ₂ PO ₄	7.2
2M9 minimal media solution B	Per L solution A: 4.0 g (¹³ C (>99% ¹³ C)) glucose, 10.0 ml BME Vitamins (100x), 2.0 ml MgSO ₄ (1M), 0.1 ml CaCl ₂ (1M), 1.0 g (¹⁵ N (>98% ¹⁵ N)) NH ₄ Cl, diluted in 10.0 ml water	-
SDS-PAGE sample buffer	18.75 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, 0.1% SDS, 1 mM EDTA	7.7

2.1.3 Plasmids

Table 2.3: List of plasmid constructs used.

For further details for each plasmid, if applicable, see relevant chapter.

Protein	Gene	Uniprot AC	Species	Vector	Domain(s)	Residues	Source
KANK1 ANK	KANK1	Q14678	Human	pET-151 TOPO	Ankyrin repeats	1078-1328	Subcloned
KANK2 ANK	KANK2	Q63ZY3	Human	pET-151 TOPO	Ankyrin repeats	583-832	GeneArt, Invitrogen
KANK2 ANK A670V	KANK2	Q63ZY3	Human	pET-151 TOPO	Ankyrin repeats	583-832	SDM of KANK2 ANK
KANK3 ANK	KANK3	Q9Z1P7	Mouse	pET-151 TOPO	Ankyrin repeats	524-773	GeneArt, Invitrogen
KANK4 ANK	KANK4	Q6P9J5	Mouse	pET-151 TOPO	Ankyrin repeats	755-1002	Subcloned
KANK4 ANK Y801H	KANK4	Q6P9J5	Mouse	pET-151 TOPO	Ankyrin repeats	755-1002	SDM of KANK4 ANK
mCer-KN	KANK1	Q14678	Human	pET-151 TOPO	KN domain	1-78	Subcloned from Akhmanova group construct
mVen-ANK	KANK1	Q14678	Human	pET-151 TOPO	Ankyrin repeats	1078-1352	Subcloned from Akhmanova group construct
Mini-KANK	KANK1	Q14678	Human	pET-151 TOPO	KN domain and ankyrin repeats	30-73; 1078-1328	GeneArt, Invitrogen
Talin1 R7R8	Tln1	P26039	Mouse	pET-151 TOPO	R7R8	1357-1580	Dr Ben Goult
DYNLL2	DYNLL2	Q96FJ2	Human	pET-151 TOPO	Full-length	1-89	GeneArt, Invitrogen

2.2 General molecular biology

2.2.1 Generating chemically competent *E. coli* using the calcium chloride method

BL21(DE3) or DH10 β cells were streaked onto a LB agar plate and incubated at 37°C overnight. A starter culture was set up by inoculating 5 mL LB with an isolated colony and agitating at 37°C overnight. 1 mL of this starter culture was diluted into 100 mL LB and incubated at 37°C in an agitating incubator until OD₅₉₅ reached 0.6. Once this cell density was achieved, the *E. coli* culture was placed on ice for 10 minutes prior to centrifugation at 3000 rpm at 4°C for 10 minutes. The resulting pellet of *E. coli* cells was resuspended in 10 mL ice cold calcium-glycerol buffer (0.1 M CaCl₂, 10% glycerol) and placed on ice for 15 minutes. The centrifugation at 3000 rpm at 4°C for 10 minutes was repeated, and the resulting *E. coli* pellet was resuspended in 1 mL ice cold calcium-glycerol buffer. Keeping all materials at 4°C, the cells were separated into 40 μ L aliquots, flash frozen in liquid nitrogen, and stored at -80°C.

2.2.2 DNA transformations

Throughout this work, two strains of *E. coli* are used: DH10 β *E. coli* for molecular cloning and DNA amplification and BL21(DE3) *E. coli* for protein expression. For all transformations, the DNA used was from a ~90 ng/ μ L stock resulting from minipreparation (QIAGEN).

A 40 μ L aliquot of chemically competent *E. coli* was thawed on ice. For each transformation, 2 μ L of DNA was added to the cells and incubated on ice for 30 minutes. A heat-shock step was performed at 42°C for 90 seconds, followed by immediately placing the cells back on ice for a further 2 minutes. For the recovery step, the cells were incubated at 37°C for 1 hour in 200 μ L LB before being plated onto a LB agar plate containing appropriate antibiotic(s) and incubated at 37°C overnight.

2.3 Molecular cloning

2.3.1 Restriction enzyme digest method

Restriction enzyme digest followed by ligation of the generated sticky ends is a classic approach to molecular cloning. We use this method primarily for subcloning regions of interest, often from mammalian vectors or genomic DNA, into vectors for expression in BL21(DE3) *E. coli*.

For biochemical and structural analysis, we predominantly use the expression vector pET-151 TOPO. This vector encodes a 6xHis-tag followed by a TEV cleavage site and also encodes ampicillin resistance. The base vector contains a multiple cloning site with BamHI, NotI, EcoRI and XhoI restriction sites.

2.3.1.1 Oligonucleotide design

Forward and reverse oligonucleotides for the polymerase chain reaction (PCR) were designed using the DNA sequence of the region to be amplified and inserted (“insert”) into the pET-151 plasmid (“vector”). Both oligonucleotides were designed to consist of ~15-20 base pairs (bp) of the desired insert, introduce the suitable restriction site that is compatible with the restriction sites within the multiple cloning site of the vector, contain ~60-70% guanidine/cytosine content, and terminate with a guanidine or cytosine at the 5’ end. These parameters are vital for ensuring specificity and stability of the oligonucleotide. Table 2.4 shows the oligonucleotides designed for pET-151_KANK1_ankyrin-repeats and pET-151_KANK4_ankyrin-repeats. Oligonucleotides were purchased from Integrated DNA Technologies.

Table 2.4: Oligonucleotide design for restriction enzyme-based molecular cloning.

Construct	Restriction enzymes	Oligonucleotides	Notes
pET-151 KANK1ANK	BamHI	F: TT GGATCC GAAATCAGAGAGAGG	Subcloning from mammalian vector to expression vector
	XhoI	R: CTCGAG TCA TTTTGCAAAGTTGACATGGGC	
pET-151 KANK4ANK	BamHI	F: AA GGATCC CTGAGGGCTGAGAGATATAAACCC	Subcloning from mammalian vector to expression vector
	XhoI	R: CTCGAG TCA TGAGTGGGCCCCGC	

2.3.1.2 Polymerase chain reaction

All polymerase chain reactions (PCR) had a final reaction volume of 50 µL as shown in table 2.5 below and were performed using High Fidelity DNA Polymerase (Roche).

Table 2.5: General PCR reaction mixture using Expand High Fidelity DNA polymerase (Roche).

Component	Volume	Final concentration
10x reaction buffer	5 μ L	1X
Template DNA	X μ L	2 ng/ μ L
Forward oligonucleotide (IDT, 10 mM)	1 μ L	200 μ M
Reverse oligonucleotide (IDT, 10 mM)	1 μ L	200 μ M
dNTPs mix	1 μ L	100 mM each
Polymerase	0.5 μ L	2.5 U
Nuclease-free water	Make up to 50 μ L	-

The general PCR programme used for High Fidelity DNA Polymerase (Roche) is shown in table 2.6 below.

Table 2.6: General PCR programme for restriction enzyme-based molecular cloning using Expand High Fidelity DNA polymerase (Roche).

Stage	Temperature ($^{\circ}$ C)	Time	Cycles
Initial DNA denaturation	95	5 mins	1
DNA denaturation	95	30 secs	30
Oligonucleotide annealing	45-68*	1 min	
DNA extension	68	1 min/kb**	
Final DNA extension	68	10 mins	1
Final hold	4	-	1

* Annealing temperature should be \sim 5 $^{\circ}$ C below the melting point of the oligonucleotides

** Allow 1 minute per kilobase of the desired insert

Once the reaction was complete, agarose gel electrophoresis was performed to check the presence of a product at the expected size of the insert. The PCR product was then purified using the ethanol precipitation method, resuspended in 20 μ L nuclease-free water and stored at -20 $^{\circ}$ C until required for the next stage.

2.3.1.3 Restriction enzyme double digest and ligation

For the restriction enzyme double digests, separate reactions for the insert and vector were prepared. Table 2.7 shows the composition of a typical restriction enzyme double digest reaction mixture which was prepared for both the vector and the insert and incubated at 37 $^{\circ}$ C for 3 hours. Purification of the digested products was carried out using a Gel Extraction Kit (QIAGEN) according to manufacturer's instructions. The estimated

concentrations of the purified digested vector and the purified digested insert were determined by absorbance at 260 nm measured using a nanophotometer (IMPLEN).

Table 2.7: General restriction enzyme double digest reaction mixture.

In this example of a restriction enzyme digest, BamHI and XhoI were used.

Component	Volume	Final concentration
10x reaction buffer ("Buffer D", Promega)	10 μ L	1X
Insert/vector DNA	X μ L	~ 20 ng/ μ L
BamHI restriction enzyme (Promega)	2 μ L	20 u
XhoI restriction enzyme (Promega)	2 μ L	20 u
Nuclease-free water	Make up to 100 μ L	-

Ligation reactions were set up according to Table 2.8 at a 1:3 molar ratio of vector:insert and incubated at 4°C overnight. DNA transformation was then carried out into DH10 β *E. coli* (section 2.4.2). 5 mL starter cultures of isolated colonies were set up in LB media with 100 μ g/ μ L ampicillin and agitated at 37°C overnight. Isolation of plasmid DNA was performed using a QIAprep Spin Miniprep Kit (QIAGEN) according to manufacturer's instructions.

Table 2.8: General ligation reaction mixture using T4 DNA ligase (Promega).

In this example, a 1:3 ratio of vector to insert has been used.

Component	Volume	Final concentration
10x reaction buffer (Promega)	1 μ L	1X
Insert DNA	X μ L	3x vector concentration
Vector DNA	X μ L	1/3 insert concentration
T4 DNA ligase (Promega)	1 μ L	3 u
Nuclease-free water	Make up to 10 μ L	-

2.3.1.4 Validation

As a pre-validation step that the target sequence had been correctly inserted into the vector, a test restriction digest was performed for each of the isolated plasmids after minipreparation. This utilises the preserved restriction sites within the plasmid, cleaving the insert from the vector. This small-scale restriction digest was set up according to Table 2.9 and incubated at 37°C for 1 hour. Agarose gel electrophoresis was performed to visualise products at the expected size of the insert and the vector. Sanger sequencing was used to confirm correct insertion (Eurofins Genomics).

Table 2.9: General small-scale restriction enzyme double digest reaction mixture.

In this example of a restriction enzyme digest, BamHI and XhoI were used.

Component	Volume	Final concentration
10x reaction buffer ("Buffer D", Promega)	1 μ L	1X
Insert/vector DNA	X μ L	~ 20 ng/ μ L
BamHI restriction enzyme (Promega)	1 μ L	10 u
XhoI restriction enzyme (Promega)	1 μ L	10 u
Nuclease-free water	Make up to 10 μ L	-

2.3.2 Site-directed mutagenesis

Site-directed mutagenesis (SDM) allows introduction of targeted mutations into a plasmid. We used this method to introduce known disease-causing point mutations to allow investigation into the biochemical and structural implications.

2.3.2.1 Oligonucleotide design

Overlapping forward and reverse oligonucleotides for the polymerase chain reaction (PCR) were designed using the DNA sequence of the region to contain the desired mutation. Both oligonucleotides were designed to consist of ~25-30 base pairs (bp) containing the desired mutation in the centre, with the reverse oligonucleotide being the reverse complement of the forward oligonucleotide. Additionally, as above, oligonucleotides should contain ~60-70% guanine/cytosine content and terminate with a guanine or cytosine at the 5' end. These parameters are vital for ensuring specificity and stability of the oligonucleotide. Table 2.10 shows the oligonucleotides designed for pET-151_KANK2_ankyrin-repeats_A670V and pET-151_KANK4_ankyrin-repeats_Y801H. Oligonucleotides were purchased from Integrated DNA Technologies.

Table 2.10: Oligonucleotide design for site-directed mutagenesis.

Construct	Mutation	Oligonucleotides
pET-151_KANK2ANK	A670V	F: GCAATGGCAACACCCGTA CT GCATTATAGCG
		R: CGCTATAATGCAGTACGGTGTGGCCATTGC
pET-151_KANK4ANK	Y801H	F: GGCTGTGGCTGCCCAC CT CCTTGAGGTCC
		R: GGACCTCAAGGAGGTGGGCAGCCACAGCC

2.3.2.2 Polymerase chain reaction

All polymerase chain reactions (PCR) for site-directed mutagenesis had a final reaction volume of 50 μ L as shown in Table 2.11 below and were performed using Pfu polymerase (Promega).

Table 2.11: General PCR reaction mixture using Pfu DNA polymerase (Promega)

Component	Volume	Final concentration
10x reaction buffer	5 μ L	1X
Template DNA	X μ L	4 ng/ μ L
Forward oligonucleotide (IDT, 10 mM)	2 μ L	400 μ M
Reverse oligonucleotide (IDT, 10 mM)	2 μ L	400 μ M
dNTPs	1 μ L	100 mM each
Polymerase	1 μ L	1.25 U
Nuclease-free water	Make up to 50 μ L	-

The general PCR programme used for Pfu polymerase (Promega) is shown in Table 2.12 below.

Table 2.12: General PCR programme for site-directed mutagenesis using Pfu DNA polymerase (Promega)

Stage	Temperature ($^{\circ}$ C)	Time	Cycles
Initial DNA denaturation	92	5 mins	1
DNA denaturation	92	30 secs	30
Oligonucleotide annealing	42-65*	1 min	
DNA extension	68	1 min/kb**	
Final DNA extension	68	10 mins	1
Final hold	4	-	1

* Annealing temperature should be \sim 5 $^{\circ}$ C below the melting point of the oligonucleotides

** Allow 1 minute per kilobase of the desired insert

2.3.2.3 *DpnI* digest

The PCR product was incubated at 37°C for 1 hour with 1 µL (20 U) of *DpnI* restriction enzyme (New England Biolabs) to digest the template DNA. DNA transformation was then carried out into DH10β *E. coli* (section 2.4.2). 5 mL starter cultures of isolated colonies were set up in LB media with 100 µg/µL ampicillin and agitated at 37°C overnight. Isolation of plasmid DNA was performed using a QIAprep Spin Miniprep Kit (QIAGEN) according to manufacturer's instructions.

2.3.2.4 Validation

Sanger sequencing was used to confirm that the mutation was correctly introduced (Eurofins Genomics).

2.4 Protein expression

For protein expression, DNA transformation was carried out into BL21(DE3) *E. coli* with the plasmid encoding the protein of interest (section 2.4.2). Small-scale expression tests ahead of time elucidated which temperature, isopropyl β-D-1 thiogalactopyranoside (IPTG) concentration and expression time were most suitable for each individual construct before proceeding with larger scale expression as described below. A sample of the culture before addition of IPTG and a sample prior to harvesting of cells can be collected for SDS-PAGE analysis to validate expression of the protein of interest.

2.4.1 Unlabelled protein expression

A starter culture for each protein to be expressed was prepared by adding an isolated colony from transformation to 10 mL LB media with 100 µg/µL ampicillin. The starter culture was agitated at 37°C overnight. A glycerol stock for each construct was prepared by mixing 750 µL of a starter culture with 350 µL filter sterilised 50% glycerol. The glycerol stock was used in place of the isolated colony for starter cultures for subsequent rounds of protein expression.

3-5 mL of starter culture was used to inoculate a larger volume of LB, usually 750 mL in a 2 L baffled flask, including 100 µg/µL ampicillin. Cells were grown in an agitating incubator at 37°C to OD₅₉₅ 0.6-0.8, at which point they were induced with 200 µM IPTG at 18°C overnight or 1 mM IPTG at 37°C for 3 hours. Cells were harvested by

centrifugation at 6000 rpm for 20 minutes at 4°C and the pellet was resuspended in 15 mL Ni buffer A per L of culture. Resuspended pellets were stored at -20°C.

2.4.2 Isotopically-labelled protein expression for NMR

For ¹⁵N-labelled KANK ankyrin repeat constructs, a starter culture for each protein to be expressed was prepared by adding an isolated colony from transformation to 10 mL 2M9 minimal media (proportional solution A and B) with 100 µg/µL ampicillin. The starter culture was agitated at 37°C overnight. A glycerol stock for each construct was prepared by mixing 750 µL of a starter culture with 350 µL filter sterilised 50% glycerol. The glycerol stock was used in place of the isolated colony for starter cultures for subsequent rounds of protein expression.

3-5 mL of starter culture was used to inoculate a larger volume of 2M9 minimal media (proportional solution A and B), usually 750 mL in a 2 L baffled flask, including 100 µg/µL ampicillin. Cells were grown in an agitating incubator at 37°C to OD₅₉₅ 0.6-0.8, at which point they were induced with 200 µM IPTG at 18°C overnight. Cells were harvested by centrifugation at 6000 rpm for 20 minutes at 4°C and the pellet was resuspended in 15 mL Ni buffer A per L of culture. Resuspended pellets were stored at -20°C.

For ¹⁵N-labelled and ¹³C,¹⁵N-labelled full-length DYNLL2, which had very limited growth in 2M9 minimal media, the 4:1 condensation method of protein expression was used. For this, cells were grown as per the method for unlabelled protein expression (see Section 2.5.1) in 2 L of LB media. Upon reaching OD₅₉₅ 0.6-0.8, the cells were subject to centrifugation at 6000 rpm for 10 minutes at 4°C and the pellet was resuspended in 100 mL of 2M9 minimal media solution A before repeating the centrifugation. The washed cells were resuspended in 500 mL of 2M9 media (2M9 minimal media solution A, 2M9 minimal media solution B and 100 µg/µL ampicillin) – a quarter of the volume of LB used for the growth stage. The 500 mL of 2M9 media containing ampicillin and the *E. coli* was agitated at 18°C for 1 hour to allow the cells to equilibrate to the new media before being induced with 200 µM IPTG at 18°C overnight. Cells were harvested as described above.

2.5 Protein purification

2.5.1 Cell lysis

Harvested cell pellets were removed from -20°C and thawed. Cell lysis was performed on ice using six cycles of sonication using intervals of 25 seconds on followed by 35 seconds off at 50% amplitude. Centrifugation was used to separate the soluble fraction from the cell debris and was performed at 20,000 rpm at 4°C for 20 minutes. The supernatant was retained and a $0.45\ \mu\text{m}$ syringe filter was used to remove any excess debris. A sample of this supernatant was prepared for SDS-PAGE analysis to validate that the protein of interest is soluble.

2.5.2 Protein purification by immobilised nickel ion affinity chromatography

His-tagged proteins were purified using immobilised nickel ion affinity chromatography. This form of immobilised metal ion affinity chromatography (IMAC) works on the basis that the 6xHis tag at the N-terminus of the expressed protein binds with high affinity to Ni^{2+} ions. The 5 mL HisTrap HP column used (GE Healthcare) contains immobilised sepharose resin coupled to a nitrilotriacetic acid (NTA) chelating group which is charged with Ni^{2+} ions. This allows highly specific binding of the His-tagged protein of interest. All nickel ion affinity chromatography was performed using an ÄKTA Start (GE Healthcare) at room temperature.

The filtered cell lysate after sonication was loaded onto a column equilibrated in Ni buffer A at a flow rate of 2.5 mL/min. Gradient elution at 5 mL/min was used to elute the protein of interest – this increasingly applies Ni buffer B to the column, thus increasing the imidazole concentration applied to the column. Eventually, the imidazole concentration in Ni buffer B will outcompete the imidazole functional group of the His-tag which is bound to the Ni^{2+} , causing the protein to elute from the column. This can be monitored by measuring the absorbance at 280 nm during fractionation. Fractions corresponding to a peak were collected and samples were prepared for SDS-PAGE analysis to validate presence of a purified protein at the expected molecular weight.

2.5.3 Buffer exchange and TEV protease cleavage

Suitable fractions from nickel ion affinity chromatography were combined to be subjected to buffer exchange by dialysis. This was performed overnight at 4°C . The

appropriate buffer to exchange into was determined by the isoelectric point (pI) of the protein, calculated using the ExPASy ProtParam tool (Gasteiger *et al*, 2005). An isoelectric point of ~7 and below is generally suitable for anion exchange (Q column) so Q buffer A would be used; an isoelectric point above ~8 is suitable for cation exchange (S column) so S buffer A would be used. For any proteins with an isoelectric point within/close to this range, the buffer components could be slightly altered. For example, for talin R7R8, a Q column was used with a reduced concentration of tris – 10 mM in both Q buffer A and Q buffer B – allowing the protein to bind to the column.

For all proteins, recombinantly-expressed tobacco etch virus (TEV) protease (2.5 mg/ml stock) was used to cleave the His-tag from the purified protein. This was added to the protein during dialysis.

2.5.4 Ion exchange chromatography

Ion exchange chromatography was used to remove the TEV protease, the cleaved His-tag and other small impurities from the protein of interest. All ion exchange chromatography was performed using an ÄKTA Start (GE Healthcare) at room temperature and either a Q column (anion exchange) or S column (cation exchange) was used depending on the isoelectric point of the protein. After buffer exchange, the protein was filtered using a 0.45 µm syringe and was loaded onto a column equilibrated in Q/S buffer A at a flow rate of 2.5 mL/min. Gradient elution at 5 mL/min was used to increasingly apply Q/S buffer B to the column, causing the protein to elute. This can be monitored by measuring the absorbance at 280 nm during fractionation. Fractions corresponding to a peak were collected and samples were prepared for SDS-PAGE analysis to validate presence of a purified protein at the expected molecular weight.

2.5.5 Measurement of protein concentration and storage

Once purification of the protein of interest and complete TEV protease cleavage has been validated by SDS-PAGE analysis, each protein was buffer exchanged into a suitable buffer using a PD-10 desalting column (GE Healthcare). Protein concentration was measured with a nanophotometer (IMPLEN) using molecular weight and extinction coefficient values obtained from the ExPASy ProtParam tool and the protein was concentrated or diluted as required. Purified proteins were aliquoted and flash frozen using liquid nitrogen and stored at -20°C until use.

2.5.6 SDS-PAGE analysis

Samples for SDS-PAGE analysis were prepared using 1x SDS-PAGE sample buffer and were boiled at 90°C for 5 minutes. SDS-PAGE gels were cast in 1 mm Novex empty gel cassettes (ThermoFisher) and were comprised of 1% SDS, 4% acrylamide/BIS stacking gel (pH 8.8) and 10% acrylamide/BIS separating gel (pH 6.8) according to manufacturer's instructions. Coomassie blue stain and destain were used to visualise proteins.

2.6 Peptides

2.6.1 Synthetic peptide design

Synthetic peptides of small regions of interest were designed and ordered from GL Biochem. Various parameters had to be considered for peptide design, including solubility and preservation of secondary structure. All peptides were ordered with a terminal cysteine residue to allow coupling to fluorescent dyes. Peptides used in this work are shown in Table 2.13.

2.6.2 Coupling peptides to BODIPY fluorescent dye

For fluorescence polarisation experiments, synthetic peptides were coupled to the fluorescent dye BODIPY™ TMR C5-Maleimide (ThermoFisher) diluted according to manufacturer's instructions. The coupling reaction was comprised of 100 µM peptide (resuspended in water or PBS), 0.1 % Triton X-100, 5 mM TCEP, and 25 µL of fluorescent dye made up to 1 mL using PBS. This reaction was left stirring at room temperature for 2 hours in a dark environment. A PD-10 desalting column equilibrated with PBS was used to remove excess dye. The coupled peptide was elute using 2.5 mL PBS before being aliquoted, flash frozen using liquid nitrogen and stored at -20°C until use.

2.7 Biochemical assays

2.7.1 Circular dichroism

Circular dichroism (CD) as a tool was used for initial, rapid determination of protein properties including secondary structure, folding and thermal stability. CD measures the difference in absorption of left- and right-handed circularly polarised light within optically active (i.e. chiral) molecules. Far-UV measurements (180-250 nm region) are used to determine the secondary structure of a protein. Measurements in this region look at the backbone chain where the rotation of phi (ϕ) and psi (ψ) angles vary based

Table 2.13: List of synthetic peptides used.

All synthetic peptides were purchased from GL Biochem, Shanghai, China.

Name	Gene	Uniprot AC	Species	Domain(s)	Residues	Sequence
KANK1 KN (full)	KANK1	Q14678	Human	KN	30-68	PYFVETPYGFQLDLDFVKYVDDIQKGNTIKKLNIQKRRK-C
KANK1 KN (short)	KANK1	Q14678	Human	KN	30-60	PYFVETPYGFQLDLDFVKYVDDIQKGNTIKK-C
KANK1 KN 4A	KANK1	Q14678	Human	KN	30-68	PYFVETPYGFQAAAAFVKYVDDIQKGNTIKKLNIQKRRK-C
KANK1 KN Δ60-68	KANK1	Q14678	Human	KN	30-73 Δ60-68	PYFVETPYGYQLDLDFLKYVDDIQKGNTIKPSVP-C
KANK2 KN	KANK2	Q63ZY3	Human	KN	31-69	C-PYSVETPYGYRLDLDFLKYVDDIEKGHTLRRVAVQRRPR
KANK3 KN	KANK3	Q9Z1P7	Mouse	KN	32-73	PYSVETPYGFHLDLDFLKYVEEIERGPASRRTPGPPHARRPR-C
KANK4 KN	KANK4	Q6P9J5	Mouse	KN	24-65	PYSVETPYGFHLDLDFLKYVDDIEKGHTIKRIPIHRRAKQAK-C
KANK1 699-716	KANK1	Q14678	Human	TQT motif	699-716	CLSTLDKQTSTQTVETRT
KIF21A H1	KIF21A	Q7Z4S6	Human	Helix 1	1144-1173	CGEVKPKNKARRRTTQMELLYADSSSELAS
Scribble 464-480	SCRIB	Q14160	Human	CC	464-480	C-AAEKRGLQRRATPHPSE

on the conformation of the protein. α -helical, β -sheet, β -turn and random coils have varying characteristic spectra using this method. Spectral scan experiments are used for determining the thermal stability of a protein and this involves measuring the changes in the CD signal at a fixed wavelength over the course of a temperature gradient.

All CD experiments were performed using a J-715 spectropolarimeter (JASCO) and a quartz cuvette with 1 mm path length (Starna Scientific). 20 μ M samples in PBS were used for all experiments. For far-UV experiments, spectra were collected at 205-260 nm with 4 scans at 100 nm/min, 0.5 nm step resolution, 1 sec response and 0.5 nm bandwidth. For spectral scans, CD signal was set to 222 nm (maximum signal for an α -helix) and measurements were taken between 25-80°C with 20 seconds step resolution, 4 seconds of response and 1.0 nm bandwidth. Protein secondary structure was monitored during the spectra scan by collecting far-UV spectra at 10°C intervals using the aforementioned parameters.

2.7.2 Fluorescence polarisation

Fluorescence polarisation (FP) was used to investigate putative interactions between proteins and fluorescently-labelled peptides. When a fluorescent molecule is excited by polarised light, polarised light is emitted. This assay utilises the rapid tumbling that occurs when fluorescent peptides are in solution; as the molecule tumbles, it will lead to loss of polarisation. If an interaction occurs between the fluorescently-labelled peptide and the protein, the fluorescently-labelled peptide will tumble more slowly with increasing levels of protein, causing an increase in the polarisation of the light in one direction. This change in polarisation as a function of protein concentration can be quantified, allowing binding affinity to be determined.

All FP assays were set up in black Nunc F96 polystyrene plates (ThermoFisher) with one 12-well row used per assay. Using an appropriate starting stock concentration, usually 100-300 μ M, a serial dilution of protein was carried out, resulting in well 12 containing the highest concentration and well 1 containing no protein. BODIPY-labelled peptide was added to every well to a final concentration of 1 μ M. A CLARIOstar plate reader (BMG Labtech) was used to measure fluorescence polarisation at 25°C using correct excitation and emission wavelengths for the fluorophore. Analysis of data was

conducted with Prism software (GraphPad) using non-linear regression and the equation for one site total binding, shown below.

$$Y = \left(\frac{B_{max} * X}{K_d + X} \right) + (NS * X) + background$$

Here, Y = protein concentration, X = labelled peptide concentration, NS = the slope of nonspecific binding in Y units/X units, and background = the amount of nonspecific binding with no added ligand. B_{max} is the maximum specific binding, given in the same units as Y, and K_d is the binding constant, given in the same units as X. The K_d value defines the concentration of labelled peptide required to achieve half saturation of the protein at equilibrium, i.e. the amount of protein to achieve half of the B_{max} .

2.7.3 Nuclear magnetic resonance titrations

Nuclear magnetic resonance (NMR) is a highly sensitive, powerful technique which can provide assessment of protein sample quality, investigation of putative interactions, and determination of structural information. During an NMR experiment, the sample is placed in a strong magnetic field and irradiated with pulses of radio frequency electromagnetic radiation which cause the NMR-active nuclei to resonate at characteristic frequencies, and this can be measured.

All experiments were carried out at 298 K using a Bruker AVANCE III spectrometer equipped with a QCI-P CryoProbe. Proteins were prepared in 20 mM phosphate pH 6.5 (pH 7.5 for DYNLL2 experiments), 50 mM NaCl, 2 mM DTT and 5% v/v $^2\text{H}_2\text{O}$ to a final volume of 450 μM in a Shigemi NMR tube (Sigma-Aldrich). Data were processed using TopSpin and CcpNmr Analysis (Skinner *et al*, 2015).

2.7.3.1 1D experiments (^1H)

1D NMR experiments are an incredibly useful approach with a broad range of applications and were used in this work as part of all NMR investigations. Following spectrometer set up, which involves locking (H_2O , D_2O and salt), tuning, shimming and determining 90 degree pulse width (P1), 1D spectra were collected and subsequently Fourier transformed. The processed 1D spectrum allowed assessment of solvent suppression and protein quality: whether the protein is present, folded, and at the required concentration for further NMR experiments, where relevant. Additionally, as

samples for 1D NMR do not require isotopic labelling, this powerful tool was used to assess relative concentrations of synthetic peptides used in this work.

2.7.3.2 2D experiments (^1H , ^{15}N)

Performing titrations via ^1H , ^{15}N -2D NMR experiments works on the basis that the hydrogens and isotopically-labelled nitrogen atoms have active nuclei under the magnetic field. Labelling in this way means that each amide bond of each amino acid (with the exception of proline) in the protein backbone has a particular chemical shift, and these are represented as a peak in a specific location in the processed NMR spectrum. These chemical shifts are incredibly sensitive to the chemical environment. If a putative ligand is added to a protein and an interaction occurs, this causes a localised change to the chemical environment of the residues constituting and in proximity to the binding surface – this causes the location of those amino acid peaks to shift or alter intensity, a concept known as a chemical shift perturbation (CSP). These CSPs induced by ligand binding allow biochemical and structural characterisation of protein interactions. 2D experiments can be used to measure the chemical shifts and allow CSPs to be visualised.

Following spectrometer set up, which involves locking (H_2O , D_2O and salt), tuning, shimming and determining 90 degree pulse width (P1), 2D spectra were collected using pulse widths for specific 2D experiments according to the protein being investigated. Heteronuclear single-quantum correlation (HSQC) experiments were used for proteins up to ~ 20 kDa in size and, for proteins larger than ~ 20 kDa, transverse relaxation

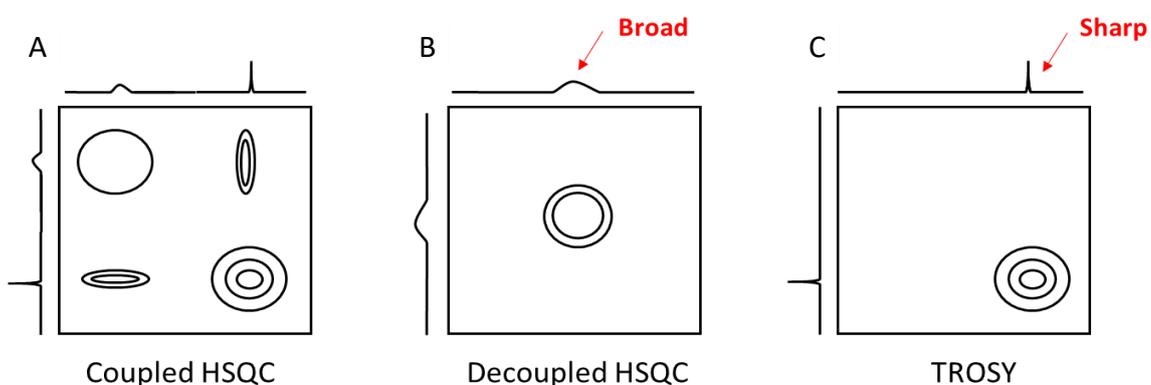


Figure 2.1: ^1H ^{15}N TROSY peak selection for a large protein.

(A) A typical coupled HSQC experiment for a larger protein will produce four peaks: a very broad “anti-TROSY” peak (top left), a “TROSY” peak (bottom right) and two “semi-TROSY” peaks. (B) The resulting peak in the decoupled HSQC spectrum generates a very broad signal. (C) A TROSY experiment is optimised to select the TROSY peak, resulting in a sharp signal and higher resolution.

optimized spectroscopy (TROSY) experiments were used. As proteins increase in size, the decay in the NMR signal occurs more rapidly – this causes the signal of the data collected to broaden as the sensitivity decreases. The TROSY experiment can increase sensitivity and ultimately the resolution of the data collected. Where a standard HSQC experiment generates a multiplet of four peaks, the average will be displayed as the final peak produced in the spectrum. For a small protein, each component of this multiplet generates a sharp peak due to uniform relaxation rates, resulting in a sharp peak and higher resolution in the representation of the corresponding amino acid in the spectrum. For a larger protein, however, the peaks in the multiplet vary as a result of different relaxation times and would result in a broader peak in a decoupled HSQC. In order to increase the resolution, TROSY experiments are optimised to only select the peak with the sharpest signal, the “TROSY peak”, resulting in an increased resolution of the spectrum.

For all titrations, the 90 degree pulse width and pH of each sample were monitored to ensure buffer uniformity between samples, eliminating false positives resulting from buffer changes induced upon addition of the potential ligand.

2.7.4 Size-exclusion chromatography-multiangle light scattering

Combining size-exclusion chromatography with multiangle light scattering (SEC-MALS) is a powerful tool for investigating oligomeric states and determining absolute molecular weight of individual components. While size-exclusion chromatography separates species within a sample according to their hydrodynamic volume, with larger components eluting first, multiangle light scattering measures the scattering of light through the sample at a range of angles and allows a Debye plot to be generated, allowing the determination of absolute molecular weight.

All SEC-MALS experiments were performed at room temperature using a Superdex 200 size-exclusion column (GE Healthcare) at a flow rate of 0.75 mL/min. A Viscotek VE 3580 RI detector was connected to a Viscotek SEC-MALS 9 to detect the light scattering (Malvern Panalytical). A 5 mg/mL BSA control was used to calibrate the system prior to each experiment in the appropriate buffer. For KANK4 ankyrin repeat domain oligomeric analysis, all samples were prepared and the system calibrated in PBS. For DYNLL2 oligomeric analysis, phosphate/citrate buffers at pH 3 and pH 7.5 were used. Data

collection and analysis, including absolute molecular weight determination, was performed with OmniSEC software (Malvern Panalytical) with BSA calibration.

2.7.5 Fluorimetry

In order to investigate *in vitro* FRET between KANK fusion proteins, fluorimetry was used. All fluorimetry experiments were performed using a Perkin Elmer Luminescence Spectrometer LS 50B with a slit width of 2.5 nm and measured at 200 nm/min. All proteins were dialysed into PBS prior to performing the experiments. An excitation wavelength of 425 nm was used to excite mCerulean-KN with emission measured at 450-600 nm. For the control wherein the mVenus-ankyrin repeat domain was excited, an excitation wavelength of 490 nm was set and emission measured at 500-600 nm.

2.8 Structural studies

2.8.1 X-ray crystallography

Crystallisation trials for KANK3 ankyrin repeat domain and Ala670Val mutant-containing KANK2 ankyrin repeat domain were performed using the hanging drop vapour diffusion method in 96-well plates with each well containing 100 μ L of a screening condition. Commercial screening kits used include JSCG+, Wizard Classic I & II (Molecular Dimensions) and Hampton HT (Hampton). Samples for crystallisation screening were prepared to >10 mg/mL in buffers comprised of 20 mM Tris pH 7.0-7.5, 50-100 mM NaCl and 1-3 mM DTT. For hanging drop vapour diffusion, a Mosquito LCP liquid handling robot (TTP Labtech) was used to create drops containing 100 nL each of protein sample and screening condition, and these were suspended over 100 μ L of the corresponding well condition. Plates were sealed and incubated at 4 °C or 21 °C. Plates were routinely observed to monitor crystal growth. Conditions containing small crystals or promising crystalline precipitate were optimised in a larger-scale fine screen wherein the pH or precipitant concentration were slightly varied around the hit condition. Fine screens were performed in 24-well plates with each well containing 500 μ L of an optimised screening condition. Larger drops of 3 μ L final volume were manually pipetted, maintaining the same 1:1 ratio of protein sample and screening condition.

Crystals were obtained at 21°C by hanging drop vapour diffusion in 0.2 M ammonium acetate, 0.1 M Bis-tris pH 5.7 and 31% w/v PEG 3350 for KANK2 ankyrin repeats A670V

mutant and 0.2 M ammonium formate and 20% w/v PEG 3350 for KANK3 ankyrin repeats. Crystals were cryoprotected in the same solution supplemented with 20% v/v glycerol prior to vitrification in liquid nitrogen, preserving the crystal. Diffraction datasets were collected at 100 K on beamline I04-1 at Diamond Light Source (Didcot, UK) using a Pilatus3 6M-F detector (Dectris, Baden, Switzerland). All crystallographic data was processed using autoPROC (Vonrhein *et al*, 2011), which incorporates XDS (Kabsch, 2010), AIMLESS (Evans & Murshudov, 2013) and TRUNCATE (Evans, 2011) for data integration, scaling and merging. The structures were determined using molecular replacement searches carried out with PHASER (McCoy *et al*, 2007) employing the published KANK2 ankyrin repeats structure (PDB ID: 4HBD) as search template. Manual model adjustment and refinement were performed with COOT (Emsley *et al*, 2010) and REFMAC (Murshudov *et al*, 2011) respectively. Figure preparation was carried out with PyMOL (Schrödinger LLC, Cambridge MA, USA). For data collection, phasing and refinement statistics, see Tables 3.1 and 3.2. The KANK2 ankyrin repeats A670V mutant and KANK3 ankyrin repeats structures have been deposited to RCSB Protein Data Bank (PDB) with accession codes 6TLH and 6TMD respectively.

The two KANK2 ankyrin repeat domain structures in Chapter 4 – apo and crystallised in the presence of a 3-fold molar excess of KANK2 KN 31-69 – were solved from crystals obtained at 21°C by sitting drop vapour diffusion. For sitting drop vapour diffusion, a Gryphon LCP liquid handling robot (Art Robbins Instruments) was used to create drops of the same volume in the sample well adjacent to the corresponding well condition. Crystals were obtained in 100 mM sodium phosphate dibasic/citric acid pH 4.2, 200 mM sodium chloride and 20% (w/v) PEG 8000 and were cryoprotected in NVH immersion oil prior to vitrification in liquid nitrogen. Diffraction datasets were collected at 100 K on beamline BL-13 at ALBA Synchrotron (Barcelona, Spain) using a Pilatus 6M detector (Dectris, Baden, Switzerland). Data processing, molecular replacement and refinement were performed as above. These structures were refined to allow preliminary visualisation of general structure and determine whether the KN domain peptide was present, and thus were not solved to completion.

2.8.2 Solution NMR

To assign the backbone resonances of a protein, double labelling (^{15}N and ^{13}C) was performed during the protein expression stage to be able to use three-dimensional, triple resonance NMR experiments. Triple resonance NMR experiments link three types of atomic nuclei: ^1H , ^{13}C and ^{15}N . Other than the ^{13}C labelling, the sample preparation and initial setup in the spectrometer were identical to that for 2D experiments (see Section 2.7.3.2) just with the appropriate three-dimensional experimental parameters used. A typical set of 3D NMR experiments to assign a protein are HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, HNCO and HN(CA)CO – all six of these were performed for the assignment of DYNLL2 (see Chapter 5). Triple resonance experiments were acquired with non-uniform sampling (NUS) to collect 15% of the data with a relaxation delay of 0.2 s and using fast pulsing sequences (BEST-type). Data processing was performed with NMRPipe (Delaglio *et al*, 1995). For the NUS 3D data, the iterative reweighted least squared (IRLS) algorithm was applied with 20 iterations using qMDD software (Orekhov & Jaravine, 2011). NMRView was used to analyse the triple resonance data (Johnson & Blevins, 1994).

CHAPTER 3: CHARACTERISATION OF THE KANK ANKYRIN REPEAT DOMAIN

3.1 Overview

This chapter focuses on developing a greater understanding of the ankyrin repeat domains of KANK proteins. Ankyrin repeat domains of proteins are widely recognised as mediators of protein-protein interactions throughout biology. Here, we compare the ankyrin repeat domains in the four mammalian KANK proteins in terms of their structure and function and identify a unique property of the KANK4 ankyrin repeat domain. Additionally, disease-causing single point mutations located within the ankyrin repeat domains of two KANK isoforms are characterised and we begin to explore some potential ligands based on KANK1 proteomics data.

3.2 Ankyrin repeat domains are important mediators of protein-protein interactions throughout biology

In recent years, tandemly repeating sequences have become increasingly recognised for their role as scaffolds for protein-protein interactions. Along with other types including leucine-rich repeats and armadillo repeats, ankyrin repeats are one of the most abundant tandem repeat motifs (Mosavi *et al*, 2004). Unlike other known motifs or domains known to be significantly involved in protein-protein interactions, including SH2 and SH3 domains which are prominent amongst adhesion-associated proteins (Malabarba *et al*, 2001; Panni & Dente, 2002), tandem repeat domains form a surface which can bind a wide range of ligands and are not limited to the recognition of a specific structures or sequences of amino acids (Li *et al*, 2006).

An ankyrin repeat is defined as a 33-residue sequence motif consisting of two α -helices connected by loops. As a repeating motif, they were first described in *S. cerevisiae* SWI6, *S. pombe* Cdc10 and *D. melanogaster* Notch (Breedon & Nasmyth, 1987). Shortly after, they were named after the protein ankyrin which contains 24 ankyrin repeats (Lux *et al*, 1990). In the years following, ankyrin repeats have been identified across eukarya, bacteria and archaea in addition to some viral genomes, though the majority of ankyrin repeat-containing proteins exist within eukaryotes. The ankyrin repeat-containing proteins identified thus far range from containing one to a potential 34 repeats, with this largest ankyrin repeat domain predicted to be a protein in *Giardia lamblia*

(Elmendorf *et al*, 2005). A region of two to six sequential ankyrin repeats occurs most frequently (Mosavi *et al*, 2004).

The KANK ankyrin repeat domain consists of five sequential canonical ankyrin repeats (ANK1-ANK5) immediately preceded by a seemingly unique ankyrin repeat-like structure (ANK0) which has been described as a regulatory or capping region (Weng *et al*, 2018) and is comprised of five α -helices which form a compact helical bundle. ANK0 appears to be an essential structural element for the KANK ankyrin repeat domain activity (Guo *et al*, 2018). Thus, the entire KANK ankyrin repeat domain refers to ANK0-ANK5.

3.3 Comparing the ankyrin repeat domains of all four mammalian KANKs

3.3.1 Determining the boundaries for the ankyrin repeat domain of each mammalian KANK isoform

At the start of this project, the only existing KANK structure in the Protein Data Bank (PDB) was that of the KANK2 ankyrin repeats (PDB ID: 4HBD, solved by a Structural Genomics Consortium). Using the amino acid sequence of the residues involved in this structure (583-832 of human KANK2), the corresponding residues in human KANK1, KANK3 and KANK4 were determined using multiple sequence alignment. Residues 1078-1328 of KANK1, 541-789 of KANK3 and 740-987 of KANK4 were the resulting boundaries (Figure 3.1B).

The cDNA of murine full-length KANK3 and KANK4 (a generous gift from Reinhard Fässler), were each used to generate the respective ankyrin repeat domain constructs. The corresponding boundaries in the murine isoform were therefore used to create these two ankyrin repeat domain constructs: 524-773 in murine KANK3 and 755-1002 of murine KANK4. The KANK1 ankyrin repeat domain was generated using human KANK1 cDNA (a generous gift from Anna Akhmanova). PSIPRED (Jones, 1999; Buchan & Jones, 2019), a secondary structure propensity calculating tool, predicts that these regions contain a mixture of predominantly unstructured and α -helical regions, with a few short regions of low-confidence β -strand propensity – this prediction is mostly consistent with the existing KANK2 ankyrin repeat domain crystal structure. The multiple sequence alignment and PSIPRED data for the KANK1 ankyrin repeat domain is shown in Figure 3.1 as an example of this process.

secondary structure elements of the folded protein, and again at 80 °C (red) with the expectation of observing a loss in the secondary structure. This was the case for all four isoforms with an α -helical spectrum containing the characteristic “double dip” observed for each protein, one minima at ~ 206 nm and another at ~ 222 nm. The CD melt experiments revealed varying melting temperatures across isoforms and the KANK3 ankyrin repeat domain appeared to be have the lowest thermal stability with a melting temperature of 38.6 °C. KANK1 and KANK2 ankyrin repeat domains were more similar in

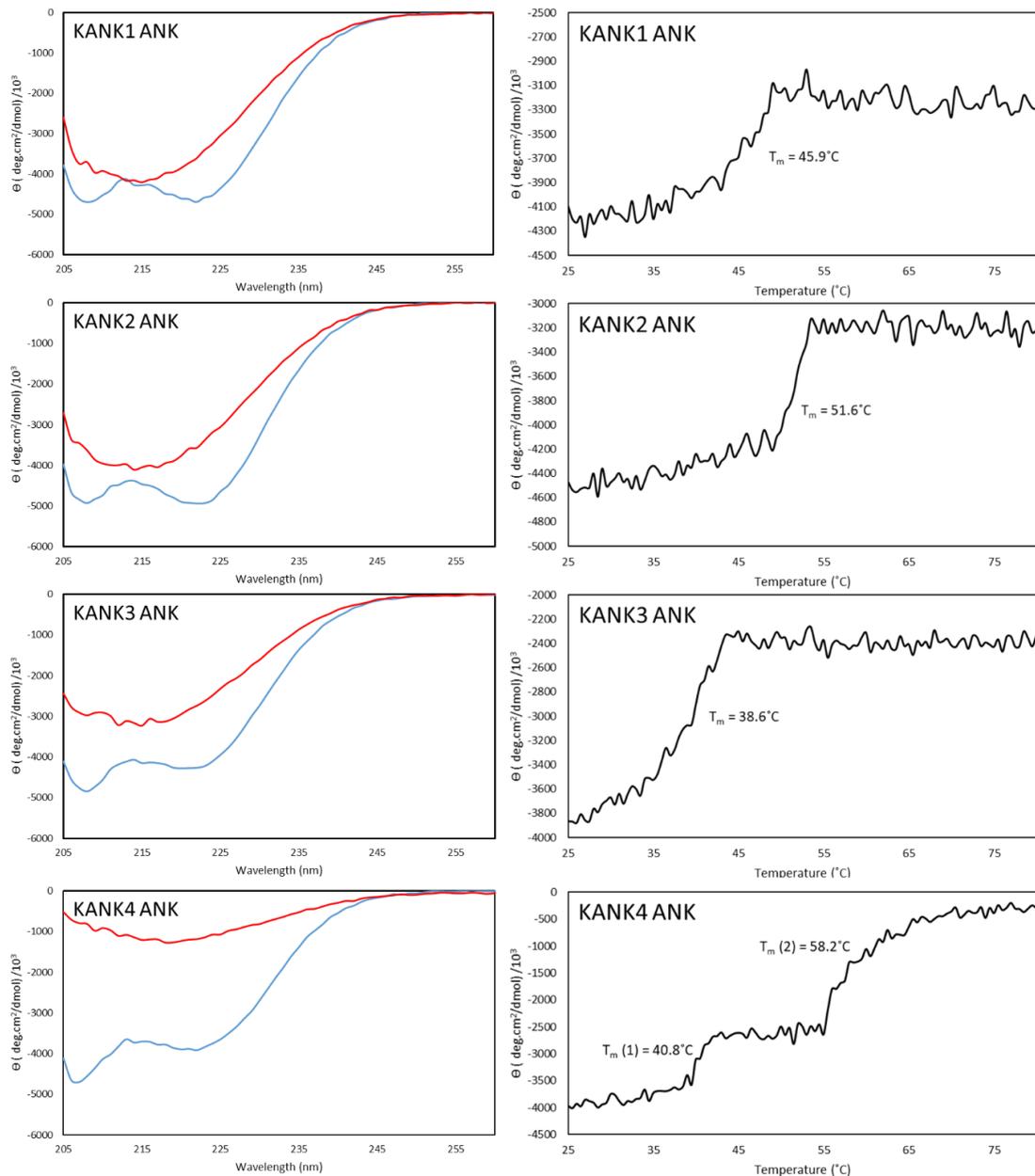


Figure 3.2: Circular dichroism reveals some unique properties of KANK ankyrin repeat domain in different isoforms.

Left: Far-UV spectra for each ankyrin repeat domain protein measured at 25 °C (blue) and 80 °C (red). Right: CD melt curves measured at 222 nm for each ankyrin repeat domain indicates varying melting temperatures between isoforms and a double transition unique to the KANK4 ankyrin repeats. Θ denotes molar ellipticity.

their melting temperatures – 45.9°C and 51.6°C, respectively – with the two most similar looking far-UV spectra with defined dips at the characteristic α -helix positions. Interestingly, the dataset for the KANK4 ankyrin repeat domain displays some unique properties: although a characteristic α -helical double dip can be observed at the typical wavelengths, the minima at 208 nm appears more defined. Most strikingly, however, the CD melt curve reveals a double transition state with two melting temperatures: one at a relatively low temperature of 40.8°C and another at 58.2°C. This property of the KANK4 ankyrin repeat domain will be discussed in more detail in Section 3.5.

3.3.2 The crystal structure of the KANK3 ankyrin repeat domain

With such little known about KANK3 and KANK4 compared to KANK1 and KANK2, including no structures deposited in the PDB, we sought to crystallise and solve the structure of the KANK3 and KANK4 ankyrin repeat domains. Investigation of KANK3 was particularly of interest due to its unique localisation as observed in U2OS cells: in addition to its localisation at the periphery of adhesion sites, KANK3 is the only KANK isoform which also can be observed at the tips of filopodia (Figure 3.3) when tagged with GFP for visualisation. Although the main domains and structural elements of KANKs appear mostly similar in terms of their amino acid sequence and predicted secondary structure, we aimed to investigate this further.

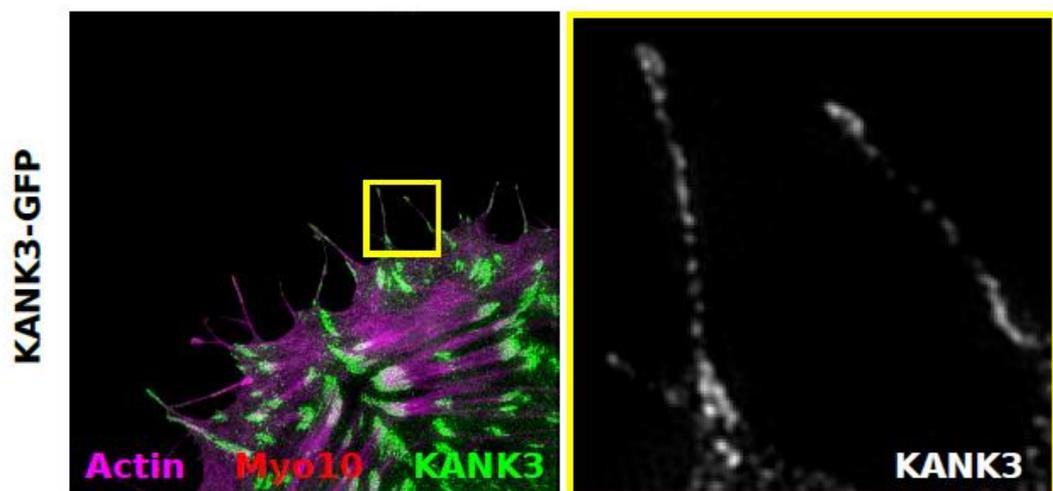


Figure 3.3: KANK3 localises to the tips of filopodia.

In addition to localisation at the periphery of adhesions, observed here at the ends of actin filaments, KANK3 is the only KANK isoform also localises at the tips of filopodia. SIM images of U2OS cells courtesy of Guillaume Jacquemet (unpublished) visualising endogenous KANK3.

The novel crystal structure of the KANK3 ankyrin repeat domain at 1.80 Å was solved as part of this work and is shown in Figure 3.4A. The data collection and refinement statistics are listed in Table 3.1.

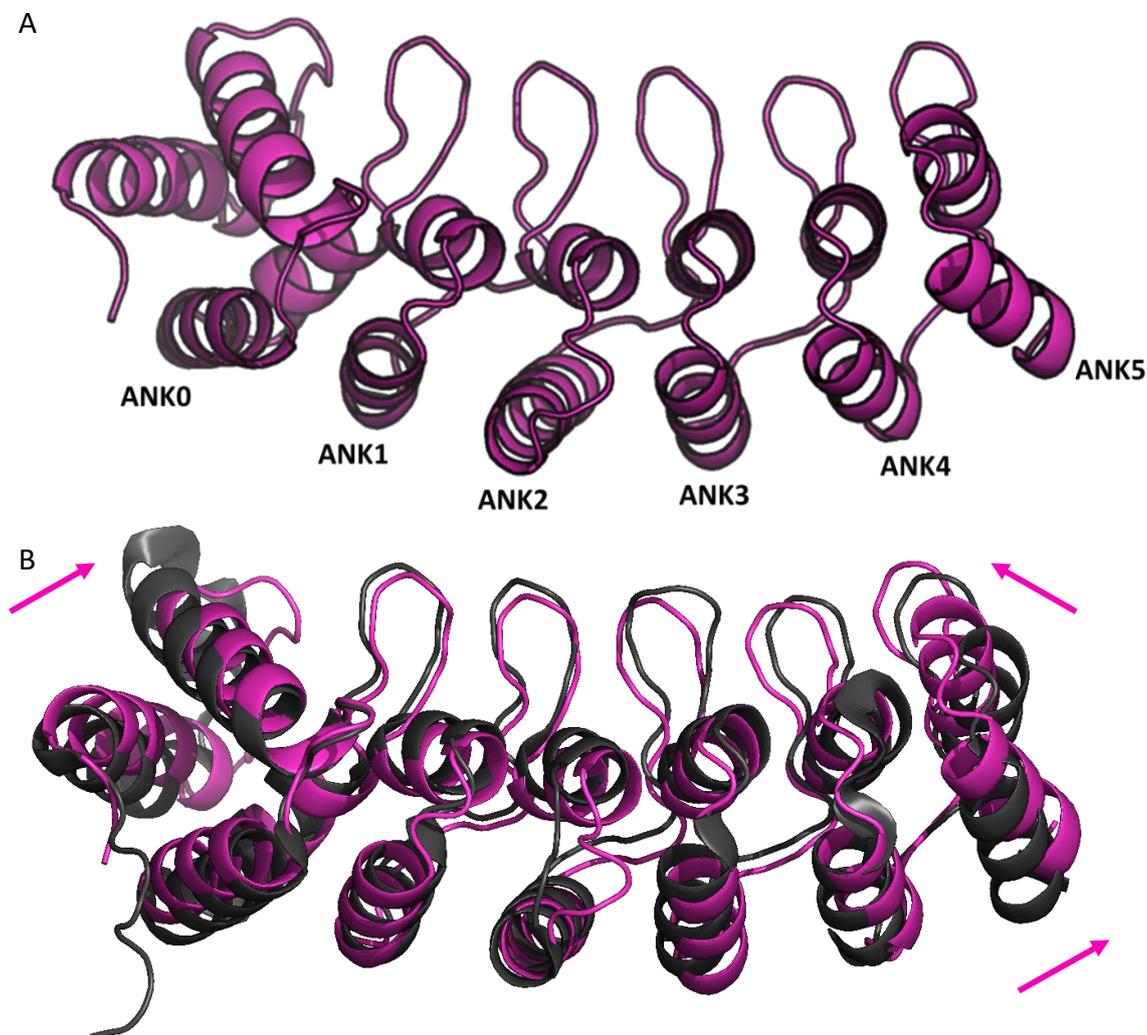


Figure 3.4: The crystal structure of the KANK3 ankyrin repeat domain.

(A) The crystal structure of the KANK3 ankyrin repeat domain at 1.8 Å resolution shows the expected ANK0-ANK5 domain structure and has been deposited in the PDB with the accession number 6TLH. (B) The KANK3 ankyrin repeat domain structure (pink) overlaid with the KANK1 ankyrin repeat domain (PDB ID: 5YBJ, grey) reveals a slightly more compact structure of the KANK3 ankyrin repeat domain. Pink arrows indicate the direction of the structural shifts in KANK3 ankyrin repeat domain in relation to the KANK1 ankyrin repeat domain.

Although the overall structure appears very similar to the other KANK ankyrin repeat domain structures, alignment with the KANK1 ankyrin repeat domain indicates that the KANK3 ankyrin repeat domain adopts a slightly more compact structure. This is highlighted by observing the loops, which all are oriented more towards the centre of the whole domain. Consistent with the loops oriented more inwards, the N-terminus of

the final helix in ANK5 tilts outwards and seemingly enables the more compacted overall structure.

Table 3.1: X-ray data collection and refinement statistics for KANK3 ankyrin repeats.

Data collected from a single KANK3 ankyrin repeat domain crystal. *Values in parentheses are for highest-resolution shell. ^Values in parentheses indicate percentile scores as determined by MolProbity.

Data collection	KANK3 ankyrin repeats
Synchrotron and Beamline	Diamond I04-1
Space group	C2
Molecule/a.s.u	2
Cell dimensions	
<i>a, b, c</i> (Å)	149.43, 45.77, 69.59
<i>a, b, g</i> (°)	90, 95.06, 90
Resolution (Å)	69.32 – 1.80 (1.84 – 1.80)*
<i>R</i> _{merge}	0.062 (0.776)
<i>I</i> / <i>σI</i>	10.5 (1.5)
<i>CC</i> (1/2)	0.994 (0.811)
Completeness (%)	97.4 (96.7)
Redundancy	3.5 (3.7)
Refinement	
Resolution (Å)	1.8
No. reflections	40455 (2956)
<i>R</i> _{work} / <i>R</i> _{free}	0.19/0.23
No. atoms	
Protein	3547
Water	370
<i>B</i> -factors (Å ²)	
Protein	33.53
Water	38.36
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.529
Ramachandran plot	
Favoured/outlier (%)	98.74/0
Rotamer	
Favoured/poor (%)	93.53/2.43
MolProbity scores	
Protein geometry	1.36 (97 th)^
Clash score all atoms	2.81 (99 th)^
PDB accession no.	6TLH

3.3.3 KANK3 and KANK4 ankyrin repeat domains bind KIF21A with a weaker affinity compared to KANK1 and KANK2 ankyrin repeat domains

The interaction between KANK ankyrin repeat domains and KIF21A, the kinesin-4 which is part of the CMSC that captures microtubules in the vicinity of adhesions, has been recognised for several years (Kakinuma *et al*, 2009; van der Vaart *et al*, 2013). More recently, structures of this interaction with the KANK1 and KANK2 ankyrin repeat domains were published during the early stages of this work (Weng *et al*, 2018; Guo *et al*, 2018; Pan *et al*, 2018) though these studies agree that neither KANK3 nor KANK4 ankyrin repeat domains interact with the KIF21A peptides used in these works. While the peptides used in these papers were of residues 1142-1169 (murine), 1142-1167 (human), or 1138-1156 (murine) of KIF21A, our KIF21A peptide used throughout in the following work spanned residues 1144-1173 of human KIF21A.

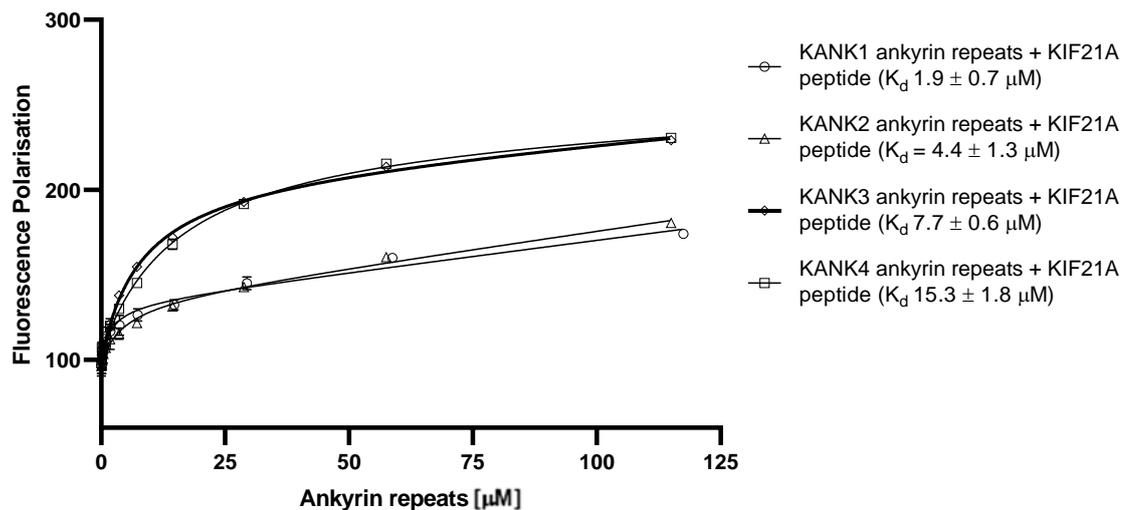


Figure 3.5: All human KANK ankyrin repeat domains bind to KIF21A 1144-1173 with varying affinity.

Binding of KANK1-4 ankyrin repeat domains to BODIPY-labelled KIF21A 1144-1173 was measured using a fluorescence polarisation assay. Dissociation constants \pm SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

Using this KIF21A peptide, an FP assay was performed that shows that KANK 1 and KANK2 ankyrin repeat domains bind KIF21A with the highest affinities – 1.9 μM and 4.4 μM , respectively – but, in contrast to published data, KANK3 and KANK4 ankyrin repeat domains also bound to the KIF21A peptide with K_d values of 7.7 μM and 15.3 μM , respectively. It would be interesting to deduce whether the more inward-facing loops within the KANK3 ankyrin repeat domain are contributing to the weakened affinity and whether the KANK4 ankyrin repeat domain adopts a similar structure.

3.4 Characterising the Ala670Val disease-causing point mutation in KANK2

3.4.1 Ala670Val causes palmoplantar keratoderma and woolly hair

KANK2, also known as steroid receptor coactivator (SRC)-interacting protein (SIP), has been identified to sequester SRCs in the cytoplasm (Zhang *et al*, 2007), a function facilitated by its ankyrin repeat domain. As mentioned in Chapter 1, one particular mutation – Ala670Val, located in ANK1 – is associated with phenotypes including palmoplantar keratoderma and woolly hair. Although this combination of symptoms is usually accompanied by cardiomyopathy and resulting from the known Naxos and Carvajal syndromes which are caused by mutations in desmosomal proteins, that is not the case for this KANK2 mutation (Ramot *et al*, 2014). Because this single point mutation can cause such significant phenotypes, we decided to investigate the structure and some functions of this mutant form of the KANK2 ankyrin repeat domain compared to the wildtype KANK2 ankyrin repeat domain protein.

3.4.2 Circular dichroism indicates that KANK2 wildtype and Ala670Val mutant are structurally similar

As in Section 3.1.1 for each wildtype mammalian KANK ankyrin repeat domain, CD was used to compare the major secondary structure elements and determine the thermal stability of the wildtype and Ala670Val mutant KANK2 ankyrin repeat domains (Figure 3.6).

As previously shown in Figure 3.2, the wildtype protein displays a clear double dip in a far-UV spectrum at 25°C (blue): characteristic dips at specific wavelengths which are caused by α -helical elements. The Ala670Val mutant displays this similar profile, consistent with the structures of the wildtype and this mutant being broadly similar. The far-UV spectrum measured at 80°C (red) displays a loss of α -helical secondary structure for both proteins. The CD melt curves measured over 25-80°C at 222 nm indicate that the melting temperatures for these two proteins are quite similar, with the Ala670Val mutant having a slightly reduced stability ($T_m = 47.6^\circ\text{C}$) compared to the wildtype ankyrin repeat domain ($T_m = 51.6^\circ\text{C}$). Combined, these data indicate that the wildtype and mutant likely have very similar structures – thus, the Ala670Val mutant does not have a hugely significant effect on the structure of the KANK2 ankyrin repeats. Alanine and valine are both amino acids containing hydrophobic side chains and so a significant

perturbation in structure or inter- and intramolecular interactions are unlikely based on this. Valine is a bulkier residue, however, so this could cause shifts in sub-structures within the domain.

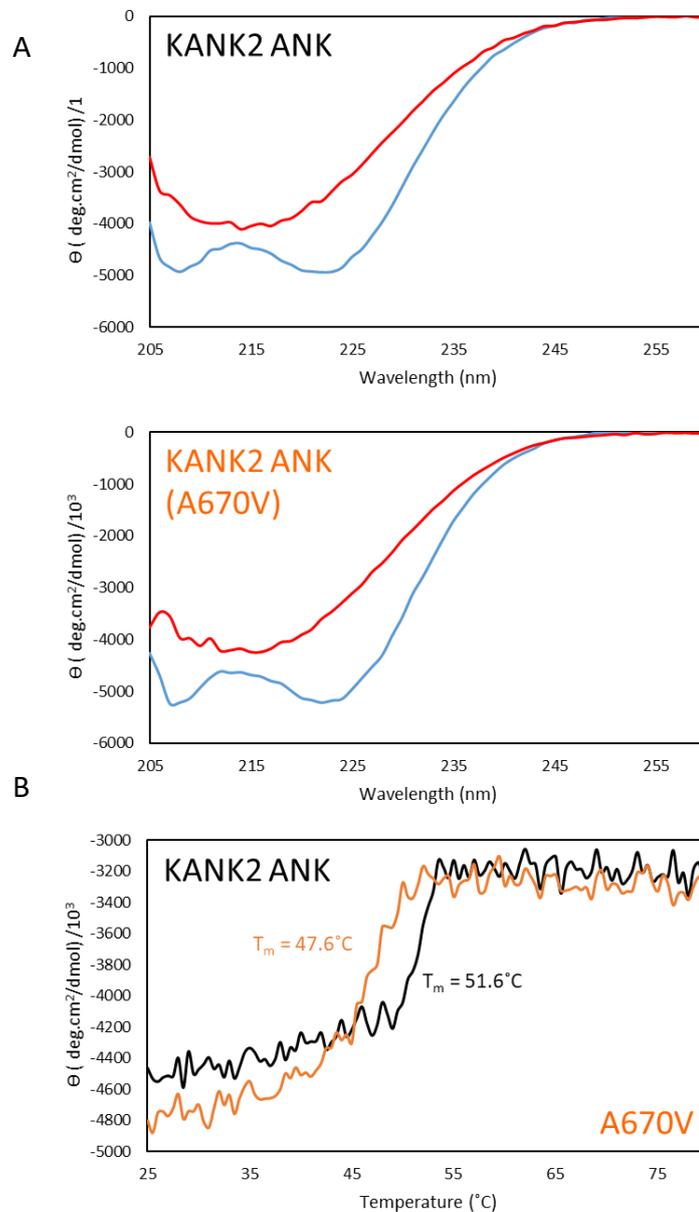


Figure 3.6: Circular dichroism indicates that KANK2 ankyrin repeats wildtype and Ala670Val mutant have similar structural properties.

(A) Far-UV spectra for wildtype and Ala670Val mutant versions of KANK2 ankyrin repeat domain proteins measured at 25 °C (blue) and 80 °C (red). (B) CD melt curves measured at 222 nm indicate similar melting temperatures for the wildtype and mutant versions of the protein, with a slight decrease in thermal stability for the Ala670Val mutant. Θ denotes molar ellipticity.

3.4.3 The crystal structure of the Ala670Val mutant

In order to develop an even greater understanding of the Ala670Val mutant and validate the CD findings, protein crystals of the Ala670Val mutant version of the KANK2 ankyrin repeats were grown and the crystal structure solved to 1.50 Å.

This novel crystal structure of KANK2 ankyrin repeat domain containing Ala670Val is aligned with the wildtype KANK2 ankyrin repeat domain protein (PDB ID: 4HBD) is shown in Figure 3.7 and the data collection and refinement statistics are listed in Table 3.2.

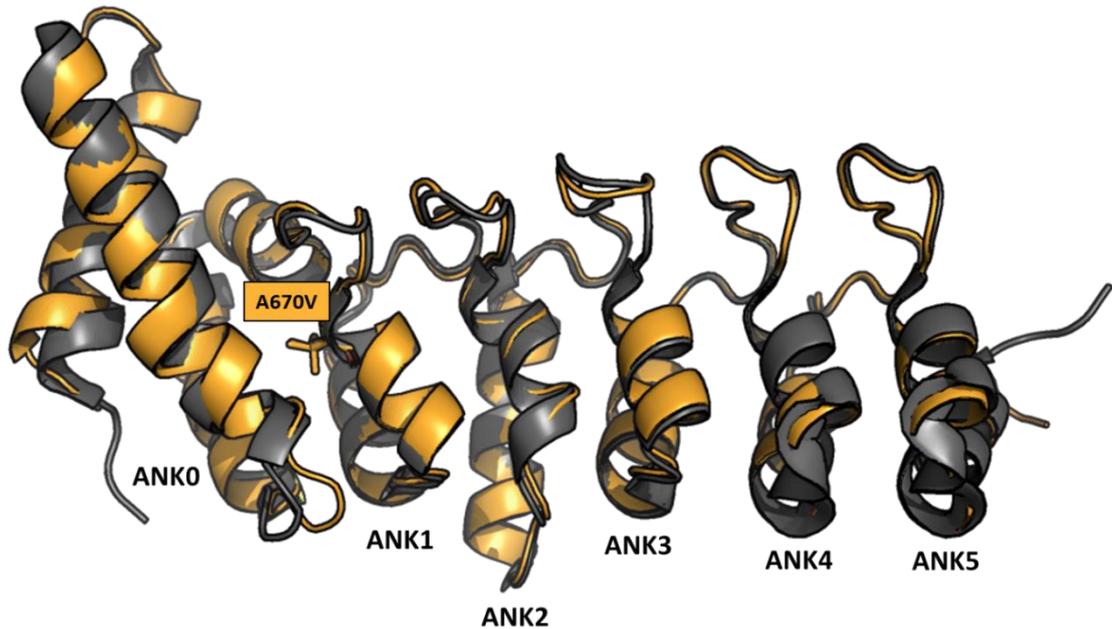


Figure 3.7: The crystal structure of the KANK2 ankyrin repeat domain containing the disease-causing Ala670Val mutation.

The crystal structure of the KANK2 ankyrin repeat domain containing the Ala670Val mutation at 1.5 Å resolution (orange) shows the expected ANK0-ANK5 domain structure and has been deposited in the PDB with the accession number 6TMD. This structure is overlaid with the wildtype KANK2 ankyrin repeat domain (PDB ID: 4HBD, grey) and this indicates very few differences between the wildtype and Ala670Val mutant.

Ala670Val (orange) is overall structurally similar to the wildtype KANK2 ankyrin repeat domain protein (grey), as expected with the relatively minor mutation. There are some small differences which can be observed: notably, the loops connecting ANK1-ANK2 and ANK2-ANK3 are slightly different in their positioning and, interestingly – despite ANK1, the region containing the mutation, being overall unaffected – the α -helices of ANK0, ANK4 and ANK5 also vary slightly in their positioning. This is particularly striking as it seems that the majority of the changes seemingly induced by the Ala670Val mutation are occurring further away from the mutated residue. Overall, although the differences are subtle, the Ala670Val mutation in the KANK2 ankyrin repeat domain appears to have a more significant effect on the KANK2 ankyrin repeat domain structure than the binding of a KIF21A peptide to the domain (see Section 4.8.1).

Table 3.2: X-ray data collection and refinement statistics for KANK2 ankyrin repeats (Ala670Val mutant).

Data collected from a single crystal of the KANK2 ankyrin repeat domain containing the Ala670Val mutation. *Values in parentheses are for highest-resolution shell. ^Values in parentheses indicate percentile scores as determined by MolProbity.

Data collection	KANK2 ankyrin repeats (A670V mutant)
Synchrotron and Beamline	Diamond I04-1
Space group	C2
Molecule/a.s.u	1
Cell dimensions	
<i>a, b, c</i> (Å)	95.77, 45.65, 51.07
<i>a, b, g</i> (°)	90, 100.38, 90
Resolution (Å)	47.10 – 1.50 (1.52 – 1.50)*
<i>R</i> _{merge}	0.056 (2.239)
<i>I</i> / <i>σI</i>	10.3 (0.5)
<i>CC</i> (1/2)	0.999 (0.329)
Completeness (%)	99.2 (99.9)
Redundancy	3.4 (3.2)
Refinement	
Resolution (Å)	1.5
No. reflections	32838 (2070)
<i>R</i> _{work} / <i>R</i> _{free}	0.17/0.22
No. atoms	
Protein	1884
Water	186
<i>B</i> -factors (Å ²)	
Protein	28.21
Water	41.55
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.565
Ramachandran plot	
Favoured/outlier (%)	98.33/0
Rotamer	
Favoured/poor (%)	95.71/0.95
MolProbity scores	
Protein geometry	1.47 (87 th)^
Clash score all atoms	8.8 (68 th)^
PDB accession no.	6TMD

3.4.4 KANK2 wildtype and Ala670Val mutant bind KIF21A equally

Because a peptide of KIF21A, a known ankyrin repeat domain binder, was available, the wildtype and Ala670Val mutation versions of the KANK2 ankyrin repeat domain were subject to a FP assay to compare their binding. From this assay, very similar affinities were calculated for each protein with the KIF21A peptide: 4.4 μM for the wildtype protein and 3.8 μM for the Ala670Val mutant. There is a very small difference here which may just be the overlap in the standard error but, if a valid difference, might be due to the small differences in the connecting loops between ANK1-ANK2 and ANK2-ANK3 which form part of the binding surface for the KIF21A interaction (see section 4.8.1).

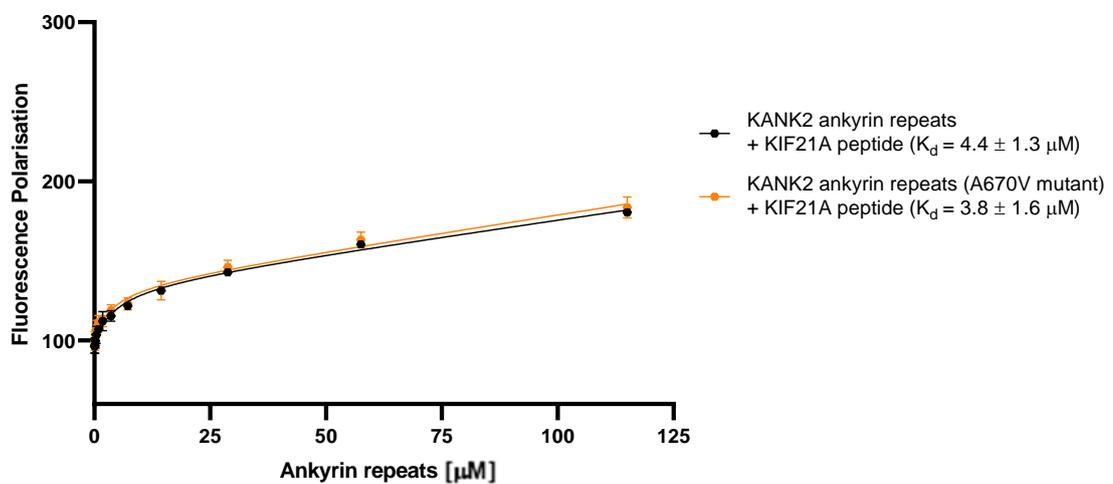


Figure 3.8: Wildtype and Ala670Val mutation versions of the KANK2 ankyrin repeat domain bind similarly to KIF21A 1144-1173.

Binding of wildtype and Ala670Val mutant KANK2 ankyrin repeat domains to BODIPY-labelled KIF21A 1144-1173 was measured using a fluorescence polarisation assay. Dissociation constants +/- SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

3.5 Characterising the Tyr801His disease-causing point mutation in KANK4

3.5.1 Tyr801His causes a steroid-resistant form of nephrotic syndrome

As mentioned in section 1.7.2, one of the diseases associated with single point mutations in KANKs includes a steroid-resistant form of nephrotic syndrome (Gee *et al*, 2013). In addition to Glu454Lys in KANK1 and Ser181Gly and Ser684Phe in KANK2, one of these single point mutations is Tyr801His in KANK4 (Gee *et al*, 2015). This particular mutation occurs within ANK0 of the KANK4 ankyrin repeat domain and so was of interest for this work. Where the murine KANK4 ankyrin repeat domain is used here, the equivalent of the Tyr801His mutation (Tyr816His) was created using site-directed mutagenesis but will be referred to as Tyr801His throughout this work.

3.5.2 Circular dichroism indicates that KANK4 wildtype and Tyr801His mutant have some structural differences

CD was used here to compare the major secondary structural elements and determine the thermal stability of the wildtype and Tyr801His mutant KANK4 ankyrin repeat domain (Figure 3.9).

Although both the wildtype and mutant proteins display the characteristic α -helical double dip at the typical wavelengths at 25°C, these traces in the far-UV are much less defined than the previous spectra shown in this chapter. This may be due to interference from other secondary structure elements or other properties that are unique to the KANK4 ankyrin repeat domain. The far-UV spectrum measured at 80°C (red) displays a loss in this α -helical secondary structure for both proteins.

The CD melt curves measured over 25-80°C at 222 nm for wildtype and Tyr801His mutant KANK4 ankyrin repeat domains reveal a significant difference between the two proteins. Strikingly, the wildtype KANK4 ankyrin repeat domain undergoes a double transition in its unfolding, with one transition state at 40.8°C and a second transition state at 58.2°C. Perhaps even more interestingly, Tyr801His from a single point mutation only undergoes a single transition, similarly to all other ankyrin repeat domain proteins investigated so far, with a melting temperature of 58.2°C. Preliminarily, this remarkable observation was hypothesised to be the result of dimerisation: if the KANK4 ankyrin repeat domain exists in a dimeric or other higher order oligomeric state, the first transition may be the breaking of this quaternary structure and the second transition, which has an equal melting temperature as the Tyr801His mutant, is the loss of the α -helical secondary structure elements.

Unfortunately, crystallographic data was not able to be collected for either the wildtype or Tyr801His mutant KANK4 ankyrin repeat domain proteins, eliminating the ability at this time to understand why the mutation of a single tyrosine residue is able to induce such a dramatic change. Further, a corresponding residue to Tyr801 is difficult to confidently identify within the other three mammalian KANK isoforms which renders

this property difficult to model and further implies that this residue at position 801 has a vital structural/functional property that is unique to KANK4.

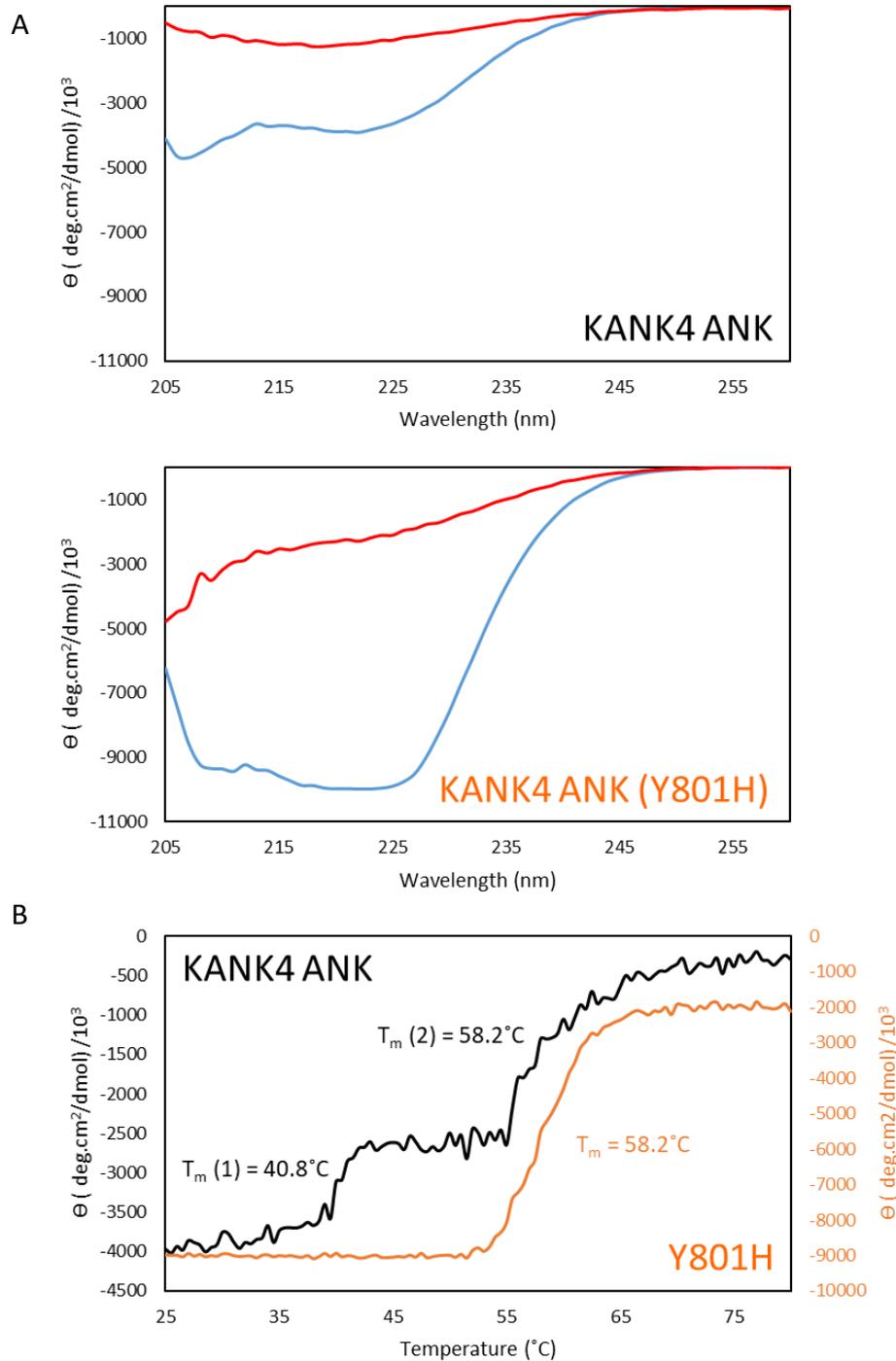


Figure 3.9: Circular dichroism indicates that KANK4 ankyrin repeats wildtype and Tyr801His mutant have varying structural properties.

(A) Far-UV spectra for wildtype and Tyr801His mutant versions of KANK4 ankyrin repeat domain proteins measured at 25 °C (blue) and 80 °C (red). (B) CD melt curve measured at 222 nm indicates a double transition in the wildtype KANK4 ankyrin repeat domain, a property seemingly unique amongst KANK ankyrin repeat domains. The Tyr801His mutant causes a switch to a single transition. Θ denotes molar ellipticity.

3.5.3 SEC-MALS analysis indicates that KANK4 wildtype and Tyr801H mutant exist in different oligomeric states

In lieu of a crystal structure of the wildtype or Tyr801His KANK4 ankyrin repeat domain, SEC-MALS was employed to further investigate the oligomeric states associated with KANK4.

It is immediately noticeable that these proteins, which theoretically are of the same molecular weight, elute differently from the SEC column: while the wildtype ankyrin repeat domain elutes with a peak at a retention volume of ~10 mL, the Tyr801His mutant elutes at ~11 mL (Figure 3.10A). Further analysis and absolute molecular weight determination reveals a molecular weight of 54142 Da for the wildtype KANK4 ankyrin repeat domain and a molecular weight of 27118 Da for the Tyr801His mutant KANK4 ankyrin repeat domain (Figure 3.10B). Thus, the wildtype ankyrin repeat domain appears to exist exclusively as a dimer whereas the single point mutation resulting in Tyr801His causes a switch to an exclusively monomeric form of the ankyrin repeat domain.

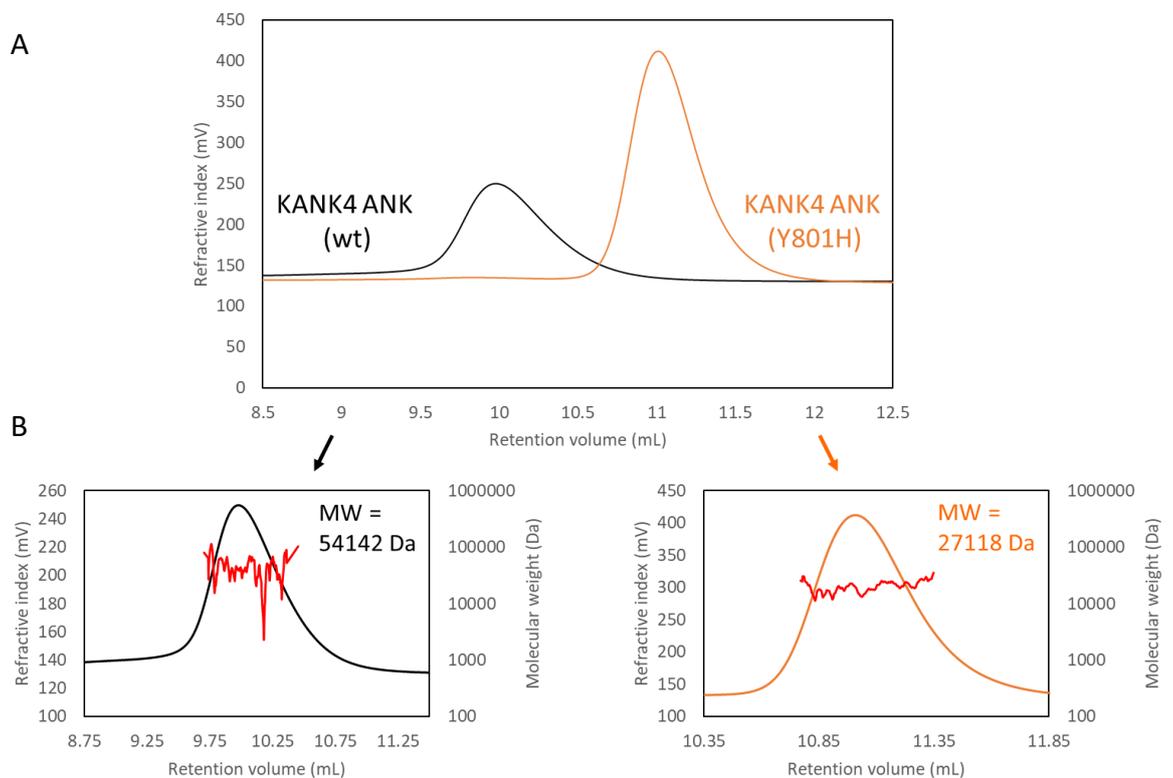


Figure 3.10: SEC-MALS analysis of wildtype and Tyr801His mutant KANK4 ankyrin repeat domain reveals differences in the oligomeric states.

(A) SEC-MALS analysis of 100 μ M wildtype or Tyr801His mutant KANK4 ankyrin repeat domain shows that each protein elutes at a different retention volume. (B) Absolute molecular weight calculations are 54142 Da for wildtype KANK4 ankyrin repeat domain and 27118 Da for the Tyr801His mutant KANK4 ankyrin repeat domain.

disulphide bond formation, and key to the unique dimeric property of KANK4 ankyrin repeat domain, this cysteine would likely be unique to KANK4 and not conserved between isoforms.

3.5.4 KANK4 wildtype and Tyr801His mutant ankyrin repeat domains vary slightly in their binding to KIF21A

While the literature indicates that the KANK4 ankyrin repeat domain does not bind to KIF21A, the wildtype and Tyr801His KANK4 ankyrin repeat domains were tested by a FP assay to compare their binding. As shown in Figure 3.5, the wildtype KANK4 ankyrin repeat domain protein binds with a K_d of 15.3 μM and this is the weakest of all wildtype ankyrin repeat domains. The Tyr801His mutant binds to the KIF21A peptide even more weakly, but only slightly so, with a K_d of 19.6 μM (Figure 3.12). Although this small difference may be insignificant, the oligomeric state of the KANK4 ankyrin repeat domain could be affecting the interaction with KIF21A. The Tyr801His mutant could perhaps be inducing a change which causes the loops constituting the KIF21A binding surface to be further compacted (as in KANK3 ankyrin repeat domain, Figure 3.4) or potentially even more dispersed, ultimately affecting the accessibility to the ANK-connecting loops. In order to predict this with more confidence, a structure of the KANK4 ankyrin repeat domain would be integral.

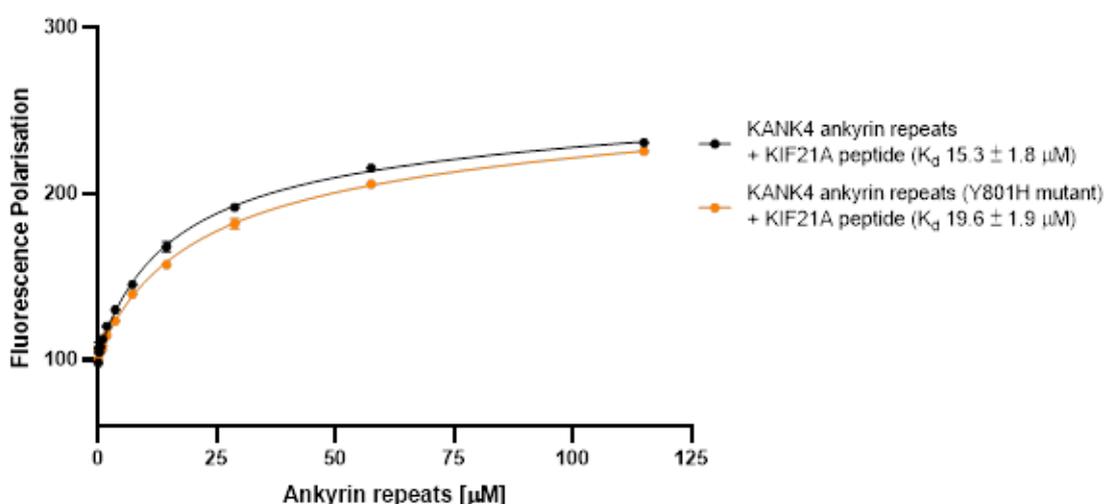


Figure 3.12: Wildtype and Tyr801His mutant versions of the KANK4 ankyrin repeat domain bind slightly differently to KIF21A 1144-1173.

Binding of wildtype and Tyr801His mutant KANK4 ankyrin repeat domains to BODIPY-labelled KIF21A 1144-1173 was measured using a fluorescence polarisation assay. Dissociation constants +/- SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

3.6 Interrogating the KANK1 proteomics dataset to search for new KANK ankyrin repeat ligands

3.6.1 The KANK1 proteomics dataset

A proteomics study was carried out by the Akhmanova group (Utrecht University) using full-length bioGFP-KANK1 in HEK239T cells (experimental procedures are detailed in Bouchet *et al*, 2016). Over 1200 hits with unique peptides were identified (unpublished data). Of these hits, it is very likely that a significant proportion are found in proximity to KANK1 or are involved in the same pathways, and some may have been identified by the bioGFP bait. Others, however, will be direct ligands, and many previously known ligands are some of the primary hits in the proteomics dataset– including KIF21A, talin1, talin2, liprin β 1 and liprin α 1.

Where ankyrin repeat domains are involved primarily in mediating protein-protein interactions, we aimed to start interrogating this proteomics dataset in the search for KANK ankyrin repeat domain ligands. One particularly interesting hit was the human Scribble homologue.

Table 3.3: GFP-KANK1 proteomics data.

A small selection of the data from a full-length GFP-KANK1 proteomics study carried out by the Akhmanova group (Utrecht University) highlighting a particularly interesting hit: the human Scribble homologue.

Protein	Accession number	Σ Unique Peptides
KIF21A	Q7Z4S6	75
Talin1	Q9Y490	70
KANK1	Q14678	63
Liprin β 1	Q86W92	49
Talin2	Q9Y4G6	47
Scribble	Q14160	37
Liprin α 1	Q13136	11
LC8	Q96FJ2	5
LC8b	P63167	5
Tensin3	Q68CZ2	1

3.6.2 Investigating a putative interaction with Scribble homologue in humans

Scribble, the human homologue of the *D. melanogaster* Scribble protein, has roles in many cellular processes including cell migration but also, of note, in cell polarity. In particular, this protein is involved in the planar cell polarity pathway which is involved

in patterning the nervous system during embryonic development (Montcouquiol *et al*, 2003). Where cell-extracellular matrix adhesions are intrinsically polar structures, this link to the establishment of polarity was of particular interest.

The scribble protein contains 16 leucine-rich repeats and four PDZ domains. In selecting a region to order as a synthetic peptide for investigation, a peptide of a coiled coil region at 464-480 was chosen due to possessing some sequence homology with the N-terminal region of the KIF21A peptide used in this work. FP was used to test if an interaction occurred between this scribble peptide and any of the KANK ankyrin repeat domains (Figure 3.13) though, despite what is likely a non-specific element rendering the curves unable to plateau, these were all negative and a K_d value could not be calculated. While only a very small region of the scribble protein was used here for the binding experiment, it is very possible that a different region of KANK binds to this coiled coil of scribble, or a different region of scribble binds to the ankyrin repeat domain or elsewhere in KANK.

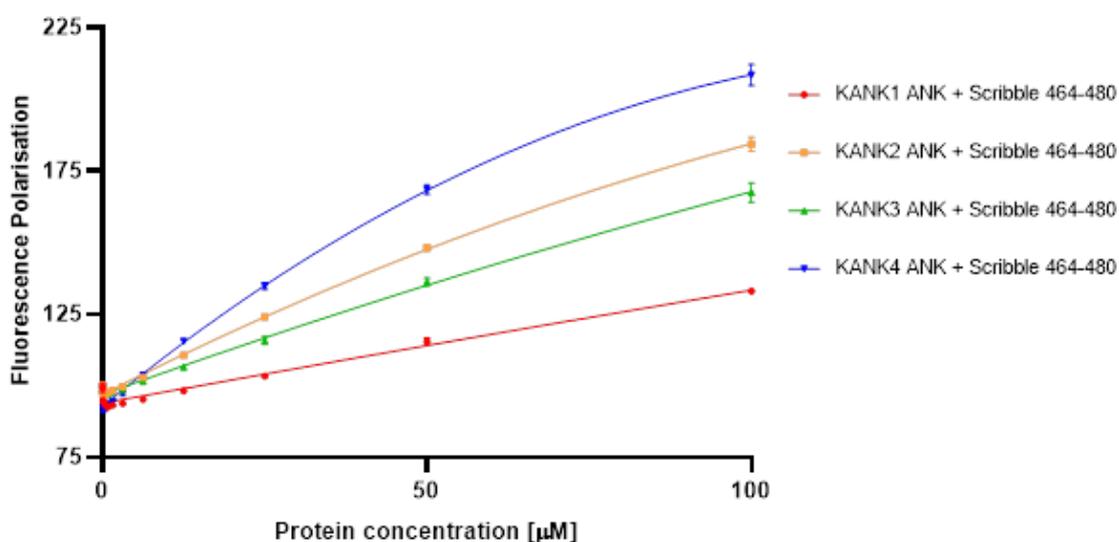


Figure 3.13: The KANK ankyrin repeat domains do not bind to scribble 464-480.

Binding of wildtype KANK ankyrin repeat domain proteins to BODIPY-labelled scribble 464-480 was measured using a fluorescence polarisation assay. All measurements were repeated in triplicate. No K_d values were able to be calculated.

3.6.3 Using the KANK1 proteomics dataset to continue identifying KANK1 ligands

The KANK1 proteomics dataset is a powerful resource which will be useful to continue to identify interesting potential ligands for KANK proteins. Throughout this thesis, the

KANK1 proteomics dataset will be revisited and mentioned as a reference from which to investigate putative novel ligands for KANK1 and other KANK isoforms.

3.7 Discussion

In this chapter, the ankyrin repeat domains of different KANKs – including different wildtype mammalian isoforms and disease-causing point mutations – have been further characterised.

Initial boundaries for the ankyrin repeat domain in the four human KANK isoforms were deduced, and CD analysis confirmed the predominantly α -helical secondary structures amongst all isoforms and the disease-causing mutants investigated, though differences in thermal stability were observed. The most striking difference was in the KANK4 ankyrin repeat domain wherein a double transition is observed in the CD melt curve, suggesting the loss of structure occurring in two stages. This is likely attributed to a dimer that is formed via hydrogen bonds involving the phenol side chain hydroxyl group from a tyrosine residue, as a disease-causing Tyr801His mutation results in a monomeric protein with a single transition. This was supported by SEC-MALS analysis with calculated molecular weights for the wildtype and Tyr801His mutant versions of the KANK4 ankyrin repeat domains indicating an exclusively dimeric or monomeric oligomeric state, respectively. However, it is difficult to accurately predict the structural basis for this difference without a structure of the KANK4 ankyrin repeats, either in the wildtype or Tyr801His form. Further, no obvious corresponding residue to Tyr801 can be identified within other ankyrin repeat domains, likely due to this being an isoform-specific property, which renders this structural change difficult to model using other KANK ankyrin repeat domains. Tyr801His in KANK4 is a known mutation which causes a steroid-resistant form of nephrotic syndrome (Gee *et al*, 2015) and it would be interesting to deduce whether the other known KANK mutations which cause these diseases, located in different regions of their respective isoforms, are attributed to similar structural changes causing similar successive effects in protein function.

Another disease-causing mutation in KANKs which was investigated here was the Ala670Val mutation in KANK2. Overall, this mutant is largely very similar to the wildtype protein and behaves similarly in its stability and its interaction with known ligand,

KIF21A. This was validated by a crystal structure of the KANK2 ankyrin repeat domain containing the Ala670Val mutation which was solved as part of this work and has been deposited into the PDB. For each of these disease-causing single point mutations in KANK, it would be very interesting to see if the observations are consistent when investigated with full-length KANK *in vitro* to see if this property is maintained and *in vivo* to give further insights into how these single point mutations can cause the specific symptoms of the associated disease. Further advancement of our knowledge of KANK proteins and fine details of their various roles will also facilitate further understanding of the development of these disease states when their normal activity is compromised.

A second novel crystal structure solved and deposited into the PDB as part of this work was that of the wildtype murine KANK3 ankyrin repeat domain. Although an overall very similar structure to the previously solved KANK1 and KANK2 ankyrin repeat domains can be observed, one interesting observation is that sub-structures of the KANK3 ankyrin repeat domain are arranged in a slightly more compacted manner, forcing the ANK-connecting loops to be oriented further towards the centre of the ankyrin repeat domain. This difference in the arrangement of the loops could be causing the weaker affinity of KANK3 to KIF21A, or perhaps causing the varied localisation of KANK3 in U2OS cells where the protein has been observed at the tips of the filopodia.

Finally, in this chapter, a proteomics dataset obtained by the Akhmanova group at Utrecht University is introduced and this has the potential to be a very powerful resource in identifying and characterising novel ligands for KANK1. This dataset will be referred to throughout this thesis.

CHAPTER 4: INVESTIGATION OF AN INTERDOMAIN KANK COMPLEX

4.1 Overview

In this chapter, the self-interaction activity of KANK proteins is explored. Many adhesion-associated proteins have been shown to undergo a regulatory intramolecular interaction, usually in a head-tail manner. Here, an interaction between the ankyrin repeat domain and KN domain of KANK1 is identified, explored primarily using a combination of fluorescence polarisation and NMR titration methods to determine the binding interface on the KN domain. Additionally, we tested whether the other KANK isoforms and KANK ankyrin repeat domains containing disease-causing mutations can autoinhibit similarly to KANK1.

4.2 Autoinhibition as a major regulatory mechanism of integrin-mediated adhesion

The concept of autoinhibitory domains – regions of proteins which can form intramolecular interactions which regulate behaviour and activity – can be found throughout biology (Pufall & Graves, 2002). Integrin-mediated adhesion, a highly complex process, relies in part on the ability to regulate by controlling the proteins that assemble to form adhesive structures and we have discussed this at length in a review article (Khan & Goult, 2019). One way that proteins can be dynamically regulated is via formation of an intramolecular interaction that maintains the protein in an inactive state, often in proximity to the adhesion, until adhesion assembly is required. The regulation of autoinhibition, and the factors that enable regulation of the activity of the protein, provides regulatable checkpoints in the system. Most notably, integrin (Anthis & Campbell, 2011) in addition to talin (Goksoy *et al*, 2008b; Goult *et al*, 2009a; Song *et al*, 2012), focal adhesion kinase (Lietha *et al*, 2007) and vinculin (Borgon *et al*, 2004) each adopt an autoinhibited conformation. A more comprehensive list of adhesion-associated proteins that undergo autoinhibition as well as further information about the roles of these conformations can be found in our review focused on autoinhibition as a major regulatory mechanism in cell-extracellular matrix adhesion (Khan & Goult, 2019) (see Appendix 1).

4.3 KANK1 is a leading hit in KANK1 proteomics

A proteomics study carried out by the Akhmanova group (Utrecht University) using full-length bioGFP-KANK1 in HEK239T cells identifies KANK1 as one of the leading hits, with 63 total unique peptides (Table 4.1).

Table 4.1: GFP-KANK1 proteomics data.

Data from full-length GFP-KANK1 proteomics study carried out by the Akhmanova group (Utrecht University) indicating that KANK1 itself is one of the top hits (red).

Protein	Accession number	Σ Unique Peptides
KIF21A	Q7Z4S6	75
Talin1	Q9Y490	70
KANK1	Q14678	63
Liprin β 1	Q86W92	49
Talin2	Q9Y4G6	47
Scribble	Q14160	37
Liprin α 1	Q13136	11
LC8	Q96FJ2	5
LC8b	P63167	5
Tensin3	Q68CZ2	1

Based on this data, the prevalence of autoinhibition throughout biology and amongst adhesion-associated proteins in particular (Pufall & Graves, 2002; Khan & Goult, 2019) and the highly conserved domains at each terminus of the protein, the hypothesis of KANK self-association in the form of head-tail autoinhibition began to develop. The N-terminal KN domains of all four KANK proteins have been shown to bind to the R7 domain of talin (Bouchet *et al*, 2016; Sun *et al*, 2016b) while the C-terminal ankyrin repeats of KANK1 and KANK2 have been shown to engage the kinesin KIF21A (van der Vaart *et al*, 2013). As these interactions are critical for the localisation and function of KANKs to adhesions, we hypothesised that the regulation of KANK activity might be controlled by limiting the exposure of the KN domain and the ankyrin repeats via an autoinhibitory interaction.

4.4 Characterising the interaction between KANK1 ankyrin repeats and KANK1 KN domain

In order to investigate the interaction between the terminal domains of KANK1, recombinantly expressed KANK1 ankyrin repeat domain and synthetic peptides of the KANK1 KN domain were used throughout.

4.4.1 Using KANK1 KN domain peptides for fluorescence polarisation assays

An *in vitro* fluorescence polarisation (FP) assay was the first method used. For each FP assay, BODIPY TMR C₅-Maleimide was used to fluorescently label a KANK1 KN domain peptide which was titrated with an increasing amount of the KANK1 ankyrin repeat domain protein.

As shown in previous work by our group, the interaction between talin1 R7R8 and KANK1 KN domain requires the leucine-aspartic acid motif at residues 41-44 (Leu-Asp-Leu-Asp) of KANK. Thus, when investigated via FP, talin1 R7R8 binds both a KANK1 KN domain peptide consisting of residues 30-68, the full KN domain, and a peptide consisting of residues 30-60, a truncated version of the KN domain. In order to abolish the interaction, a “4A mutant” peptide was designed: the four residues required in the KN domain for talin1 R7R8 binding were instead replaced with four alanine residues (Bouchet *et al*, 2016). These peptides are depicted in Figure 4.1.



Figure 4.1: KANK1 KN domain peptides used in this work.

Three of the KANK1 KN domain peptides used in this work are shown here (from Bouchet *et al*, 2016). The sequence of residues 30-68, which constitute the KN domain of KANK1, are shown above.

4.4.2 The KANK1 ankyrin repeat domain and talin1 R7R8 bind to different regions of the KANK1 KN domain

Using FP, the experiments from Bouchet *et al*. (2016) were repeated and the findings validated here as a control dataset for further experiments: talin1 R7R8 binds with high affinity to wildtype KANK1 KN domain peptides of residues 30-60 (green) and 30-68 (red), whereas the interaction is abolished with the 4A mutant-containing peptide (purple) (Figure 4.2A).

Next, the KANK1 ankyrin repeat domain protein was tested against the same set of KANK1 KN domain peptides. As shown in Figure 4.2B, the KANK1 ankyrin repeat domain protein binds to KANK1 KN domain residues 30-68 (red), the full KN domain, with a K_d of $\sim 30 \mu\text{M}$. This indicates that the KN domain and ankyrin repeat domain of KANK1 interact and supports the hypothesis of an autoinhibitory head-tail interaction. Notably, this interaction is much weaker than that between talin1 R7R8 and the KN domain and this was expected due to several factors. Firstly, the two domains of KANK are tethered together and thus may not require such high affinity to assist in bringing the two components into proximity. Additionally, a lower-affinity autoinhibitory interaction, typically associated with inactivity of the protein, may facilitate activation.

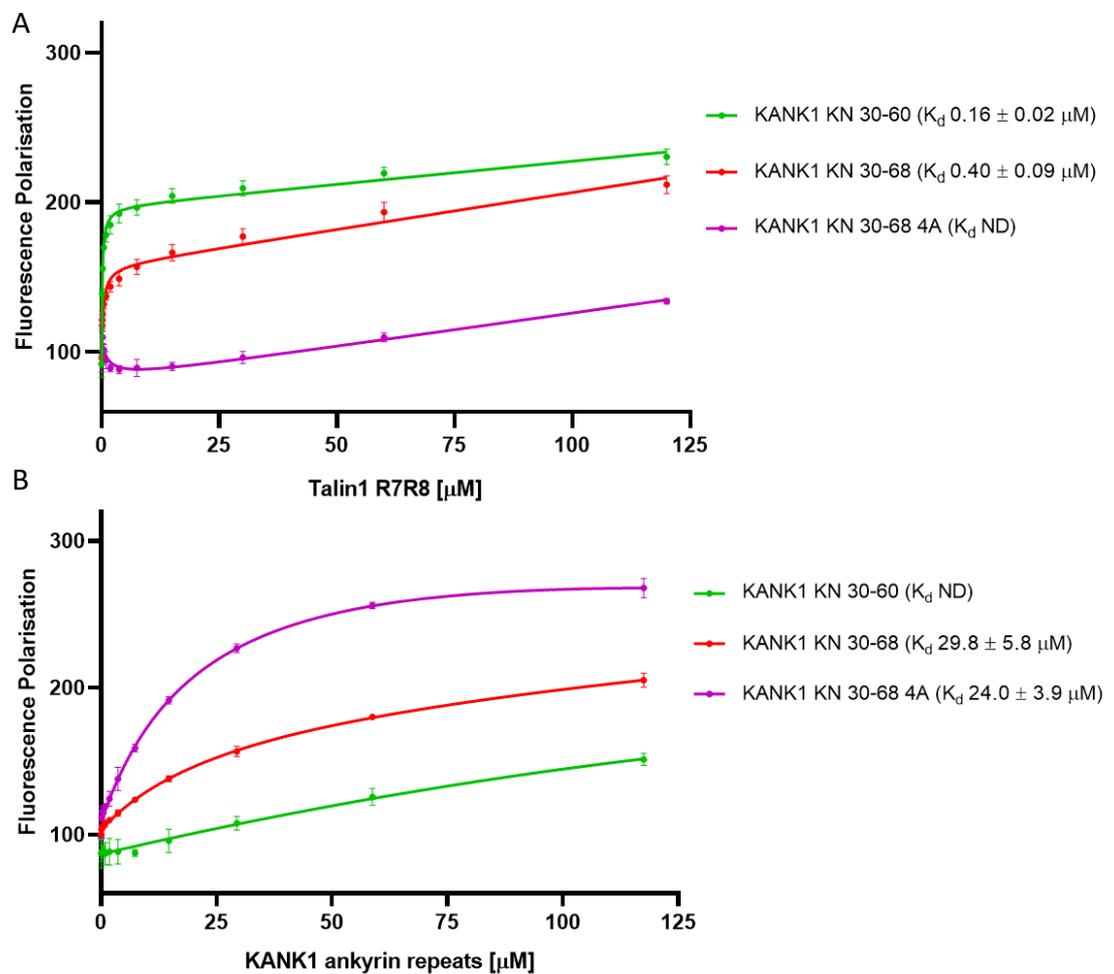


Figure 4.2: KANK1 ankyrin repeats and talin R7R8 bind differently to KANK1 KN domain.

Binding of (A) talin1 R7R8 and (B) KANK1 ankyrin repeats and to BODIPY-labelled KANK1 KN domain peptides was measured using a fluorescence polarisation assay. Dissociation constants \pm SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

Strikingly, the FP data for the KANK1 ankyrin repeat domain protein indicates binding to a different region of the KANK1 KN domain compared to talin1 R7R8. While the 4A

mutant-containing KN domain peptide (purple) abolishes talin1 R7R8 binding, this peptide still interacts with the KANK1 ankyrin repeat domain protein with a similar affinity as the wildtype KN domain peptide. This suggests that the LD motif at residues 41-44 of the KN domain are not required for the KANK head-tail interaction. Secondly, the truncated KN domain peptide of residues 30-60 (green), which binds with high affinity to talin1 R7R8, does not interact with the KANK1 ankyrin repeat domain. This suggests that the KANK1 head-tail interaction requires residues 61-68.

To support these findings from the FP assays, we decided to further investigate using NMR. Due to the ability to recombinantly express the KANK1 ankyrin repeat domain from an expression construct, this was used as the “labelled” component in the NMR experiments. In order to do this, the KANK1 ankyrin repeat domain protein was expressed in 2M9 medium with $^{15}\text{N-NH}_4\text{Cl}$ as the sole nitrogen source as described in Section 2.4.2.

A $^1\text{H},^{15}\text{N}$ TROSY NMR experiment was performed using 100 μM purified ^{15}N -KANK1 ankyrin repeat domain. This provides a reference spectrum of the protein alone for comparison upon addition of putative ligands. A TROSY experiment, as opposed to a standard HSQC experiment, was required to increase sensitivity and resolution – the KANK1 ankyrin repeat domain construct, at ~ 27 kDa, is considered too large for highly resolved HSQC experiments. Figure 4.3 shows the resulting spectrum from this TROSY

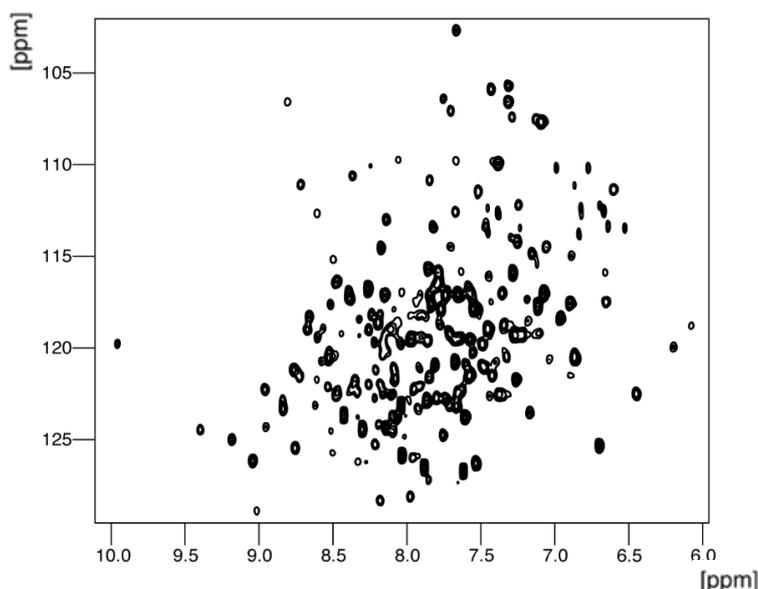


Figure 4.3: TROSY NMR spectrum of the KANK1 ankyrin repeat domain.

$^1\text{H},^{15}\text{N}$ TROSY NMR spectrum of 100 μM ^{15}N -labelled KANK1 ankyrin repeat domain.

experiment, with good peak dispersion indicating that the KANK1 ankyrin repeat protein was folded. Each peak in the spectrum results from the N-H of the amide bond of each amino acid, and thus each amino acid of the protein backbone is displayed as a peak with the exception of any proline residues.

In order to support the findings from FP analysis of the KANK head-tail interaction, each peptide was added to the ^{15}N -KANK1 ankyrin repeat domain and the $^1\text{H},^{15}\text{N}$ TROSY experiment was repeated. As shown in Figure 4.4A, when the full KN domain peptide (residues 30-68, red) is added to the KANK1 ankyrin repeats and the resulting spectrum is overlaid with the reference spectrum (black), many chemical shift perturbations can be observed. The right-hand panel highlights a particularly notable region of this spectrum where many peaks have shifted significantly while other peaks have

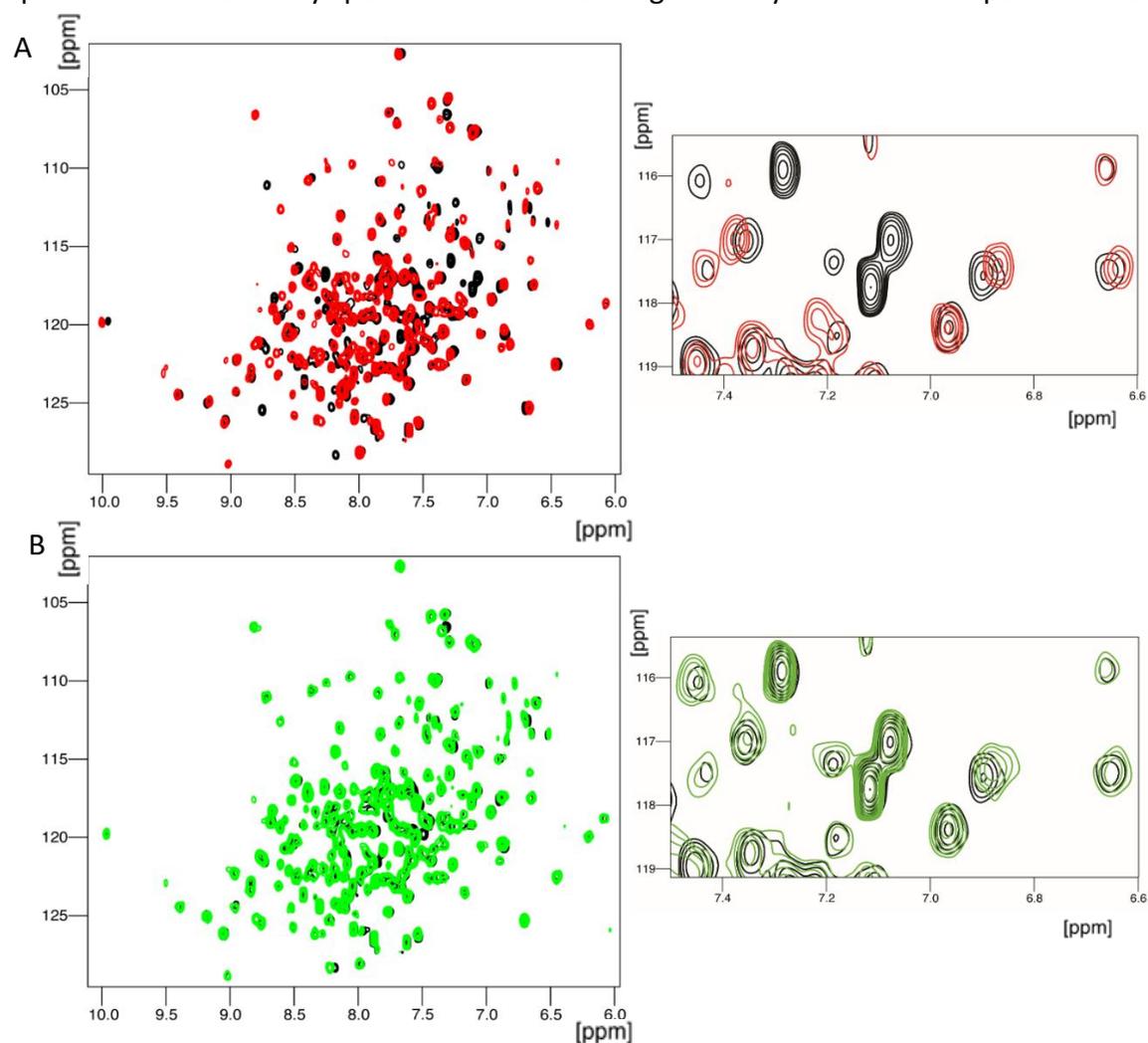


Figure 4.4: TROSY NMR titration spectra of KANK1 ankyrin repeat domain protein with KANK1 KN domain peptides.

(A) $^1\text{H},^{15}\text{N}$ TROSY NMR spectra of 100 μM ^{15}N -labelled KANK1 ankyrin repeats in absence (black) or presence (red) of 100 μM KANK1 KN (30-68). (B) $^1\text{H},^{15}\text{N}$ TROSY NMR spectra of 100 μM ^{15}N -labelled KANK1 ankyrin repeats in absence (black) or presence (green) of 100 μM KANK1 KN (30-60).

disappeared entirely. These changes in peak position indicate that the corresponding residue has had its chemical environment altered due to the presence of the peptide. This suggests that an interaction is taking place and supports the findings from the FP experiments.

In contrast, when the truncated KN domain peptide (residues 30-60, green) is added to the KANK1 ankyrin repeats and the resulting spectrum is overlaid with the reference spectrum (black), significantly fewer chemical shift perturbations can be observed (Figure 4.4B). The right-hand panel highlights the identical region as in panel Figure 4.4A and, where there were significant changes when the 30-68 peptide was added, there are very few changes when the 30-60 peptide is added to the KANK1 ankyrin repeat domain protein. Interestingly, the interaction does not seem to be entirely abolished as there are some very small changes throughout the spectrum, but these are miniscule in comparison. Taken together, these data indicate that residues 61-68 are required for an effective head-tail interaction to take place in KANK1.

4.4.3 Residues 61-68 of the KANK1 KN domain are required for the interaction between the KANK1 KN domain and the KANK1 ankyrin repeat domain

As part of the validation pipeline, a KANK1 KN domain peptide ranging from residues 30-73 excluding residues 60-68 was synthesised. This KN(30-73 Δ 60-68) peptide was used in

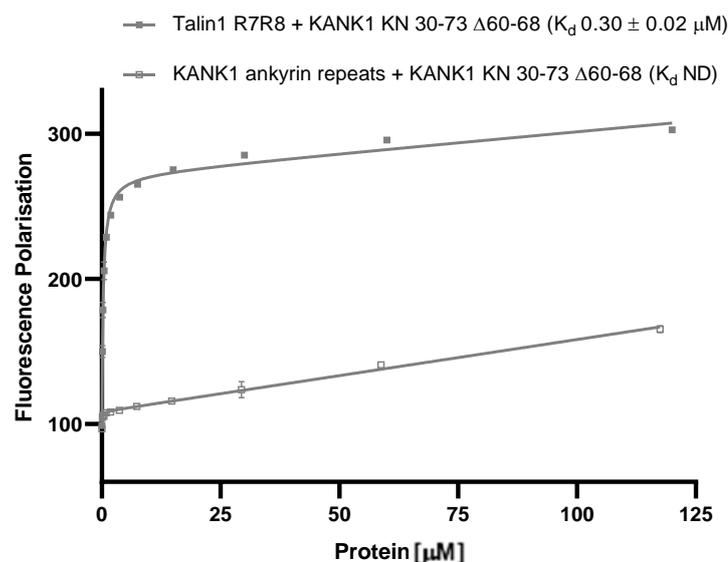


Figure 4.5: KANK1 KN(30-73 Δ 60-68) peptide interacts with talin1 R7R8 but does not interact with KANK1 ankyrin repeat domain.

Binding of talin1 R7R8 and KANK1 ankyrin repeats to BODIPY-labelled KANK1 KN domain peptides measured using a fluorescence polarisation assay. Dissociation constants +/- SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

FP assays against both talin1 R7R8 and the KANK1 ankyrin repeat domain (Figure 4.55). Talin1 R7R8 bound with high affinity as expected due to residues 30-60 being sufficient for maximal binding. However, with the KANK1 ankyrin repeat domain, deletion of this region abolished binding and a K_d value could not be calculated.

In line with this reduced binding, the NMR spectrum shows very few changes in peak positions when the KN(30-73 Δ 60-68) peptide is added to ^{15}N -labelled KANK1 ankyrin repeats (Figure 4.6), indicating a significantly weaker interaction than with 30-68. The very minor changes which can be observed are in line with the NMR spectrum of the KANK1 ankyrin repeat domain with the KANK1 30-60 peptide added, further supporting the suggestion that there may be other KN domain residues involved which assist in stabilising the head-tail interaction, similar to the interaction between KANK1/KANK2 ankyrin repeat domain and the KANK-binding region of KIF21A (Weng *et al*, 2018). Overall, it can be concluded from these data that residues 61-68 of the KN domain of KANK1 are required for autoinhibition of KANK1.

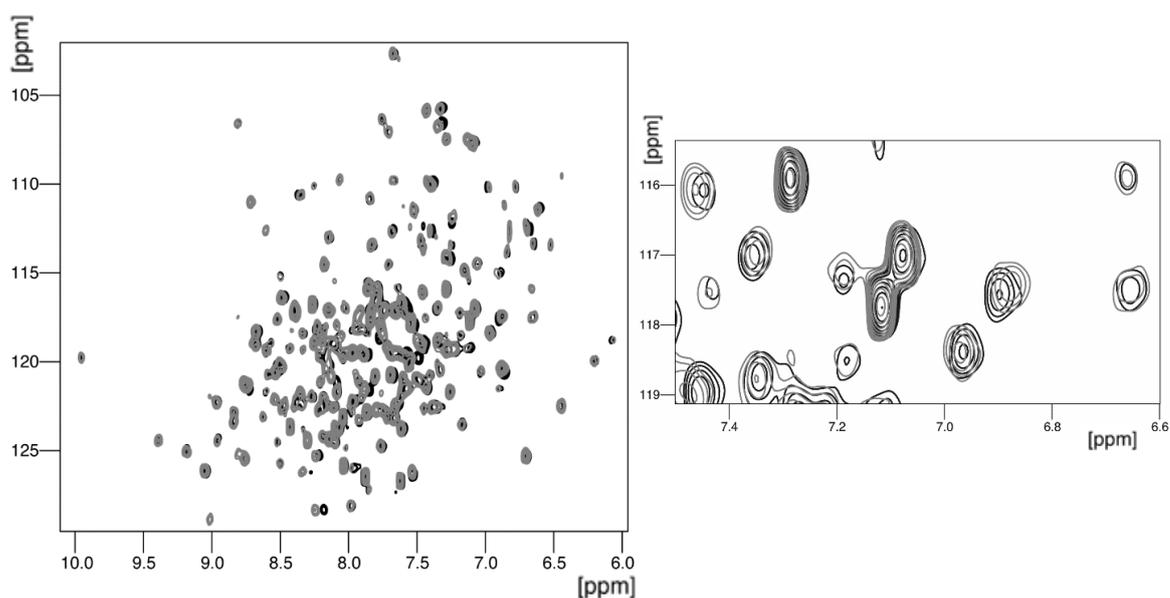


Figure 4.6: TROSY NMR titration spectrum of KANK1 ankyrin repeat domain protein with KANK1 KN(30-73 Δ 60-68) peptide.

^1H , ^{15}N TROSY NMR spectra of 100 μM ^{15}N -labelled KANK1 ankyrin repeats in absence (black) or presence (grey) of 100 μM KANK1 KN(30-73 Δ 60-68).

4.5 Investigating an interaction between the KN domain and ankyrin repeats domain in KANK2, KANK3 and KANK4

As mentioned in Chapter 1 and Chapter 3, each of the four human KANK isoforms share clear similarities but there are also significant differences. While it is clear by

biochemical and biophysical analysis that the KN domain and ankyrin repeat domain of KANK1 interact, we next sought to investigate whether this property is common amongst all KANK isoforms.

4.5.1 Designing KANK2, KANK3 and KANK4 KN domain peptides to investigate autoinhibition of the other KANK isoforms

While recombinantly expressed KANK2, KANK3 and KANK4 ankyrin repeat domain proteins was already optimised, synthetic peptides of the KN domains of each of the isoforms were required. In order to deduce the boundaries for the KN domain of each protein, multiple sequence alignment was performed (Figure 4.7A). As a result, the boundaries for the KN domain of each human isoform were residues 31-69 for KANK2, residues 33-74 for KANK3, and residues 24-62 for KANK4. Interestingly, it is clear from this analysis that KANK3 contains a three-residue insertion towards the centre of the 61-68 region. As the KANK3 and KANK4 ankyrin repeat domain proteins were designed from the murine protein sequence of each respective isoform, it was decided that the KN domain peptides designed to be tested against these should also be based upon the murine protein sequences. Multiple sequence alignment between the human and murine KANK3 and KANK4 proteins sequences was used to deduce appropriate boundaries, resulting in residues 32-73 for murine KANK3 KN domain (Figure 4.7B) and residues 24-62 for murine KANK4 KN domain (Figure 4.7C).



Figure 4.7: Multiple sequence alignments of KANK KN domains.

(A) Multiple sequence alignment of human KANK KN domains to identify the homologous region of KANK2, KANK3 and KANK4 to residues 30-68 in KANK1. (B-C) Multiple sequence alignments of the determined regions between human and murine KANK3 and KANK4. All multiple sequence alignments were performed using EMBL-EBI T-coffee (Madeira et al, 2019).

As a result of this analysis, it was decided that a good initial set of peptides to order would be those of an equivalent length to the KANK1 30-68 peptide. Thus, residues 31-69 of human KANK2, residues 32-70 of murine KANK3 and residues 24-62 of murine

KANK4 were initially produced as synthetic peptides with a cysteine residue at one of the termini to allow for coupling to maleimides for FP.

4.5.2 Optimisation of KANK3 and KANK4 KN domain peptides

While the KANK2 KN domain and KANK3 KN domain peptides were highly soluble in both water and assay buffer, it was immediately clear that the KANK4 KN domain peptide was not suitably soluble despite no predicted solubility issues using ThermoFisher Scientific Peptide Analyzing Tool. 1D NMR analysis, performed routinely as a validation step to ensure that the peptides were of similar concentration relative to each other, highlighted that an undetectable amount of KANK4 KN domain peptide was present in comparison to the three other isoforms (Figure 4.8C). Additionally, while the KANK3 KN domain peptide was appropriately soluble, the peptide would not adequately couple to fluorescent dyes for FP assays. After attempting various assay buffers and conditions to improve the outcomes, the next step was optimisation of each peptide itself.

Using the Peptide Analyzing Tool in combination with sequence analysis, three additional residues were added to the KANK3 KN domain and KANK4 KN domain peptides, now residues 32-73 of murine KANK3 and residues 24-65 of murine KANK4. For KANK3, this resulted in a KN domain peptide which contained the entirety of the region homologous to 61-68 in KANK1, including the three-residue insertion. This optimised KANK3 KN domain peptide coupled well to BODIPY. For KANK4, this resulted in a KN domain peptide where the added amino acids were Glu-Ala-Lys, the positive charge of the lysine increasing both the charge and the isoelectric point (Figure 4.8A, 4.8B). For the initial KANK4 KN peptide (24-62), the net charge at pH 7 was 1.2; the net charge at pH 7 of the optimised KANK4 KN peptide (24-65) was 2. This increased charge resulted in increased hydrophilicity and the peptide had increased solubility in both water and buffer. This was confirmed by ¹H NMR (Figure 4.8C, 4.8D).

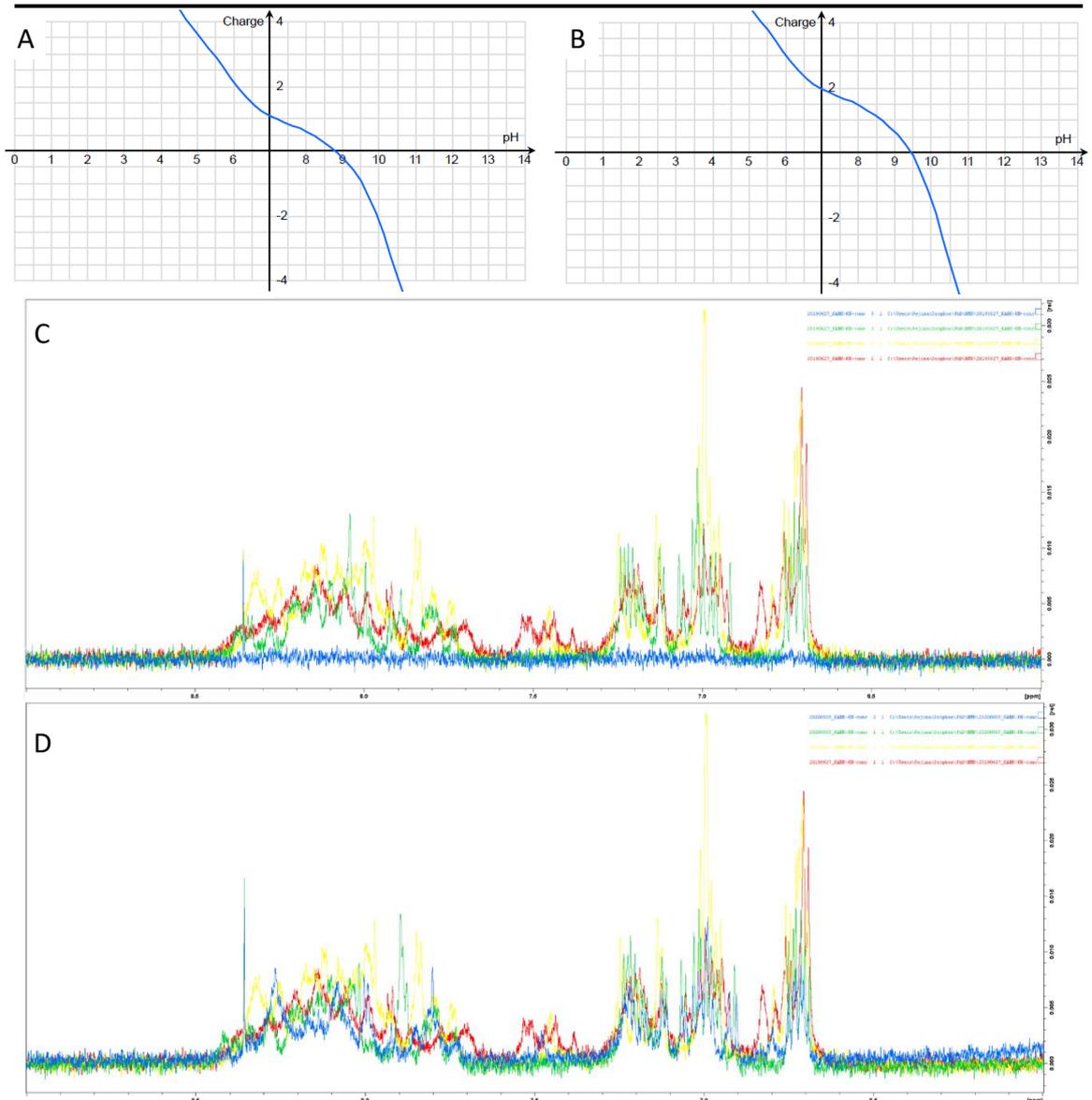


Figure 4.8: Optimisation of KANK4 KN domain synthetic peptides.

(A-B) Predicted charge and isoelectric point of synthetic peptides based on peptide sequences input into ThermoFisher Scientific Peptide Analyzing Tool for (A) initial KANK4 KN domain peptide (residues 24-62) and (B) optimised KANK4 KN domain peptide (residues 24-65). (C) Amide region from ^1H NMR analysis of KANK KN domain peptides: KANK1 KN 30-68 (red), KANK2 KN 31-69 (yellow), KANK3 KN 32-70 (green) and KANK4 KN 24-62 and (blue). (D) Amide region from ^1H NMR analysis of KANK KN domain peptides: KANK1 KN 30-68 (red), KANK2 KN 31-69 (yellow), optimised KANK3 KN 32-73 (green) and optimised KANK4 KN 24-65 (blue).

4.5.3 Using fluorescence polarisation to investigate self-interaction of KANK2, KANK3 and KANK4

With optimised KN domain peptides, FP was used to test whether an interaction between the KN and ankyrin repeat domain was a common regulatory feature amongst all four KANK isoforms. Using this technique, it was shown that the ankyrin repeat domain of each of the four KANK isoforms interacts with its corresponding KN domain (Figure 4.9). Interestingly, KANK1, KANK2 and KANK4 have similar affinities between

their own ankyrin repeat domains and KN domains, with calculated K_d values of 29.8 μM , 28.2 μM and 19.6 μM respectively. KANK3 appears to be somewhat of an anomaly here, with a calculated K_d of 92.6 μM between its ankyrin repeat domain and KN domain. This ~ 4 -fold weaker affinity could result from the 3-residue insertion (Gly-Pro-Pro) within the 61-68-equivalent region of the KANK3 KN domain (Figure 4.7A).

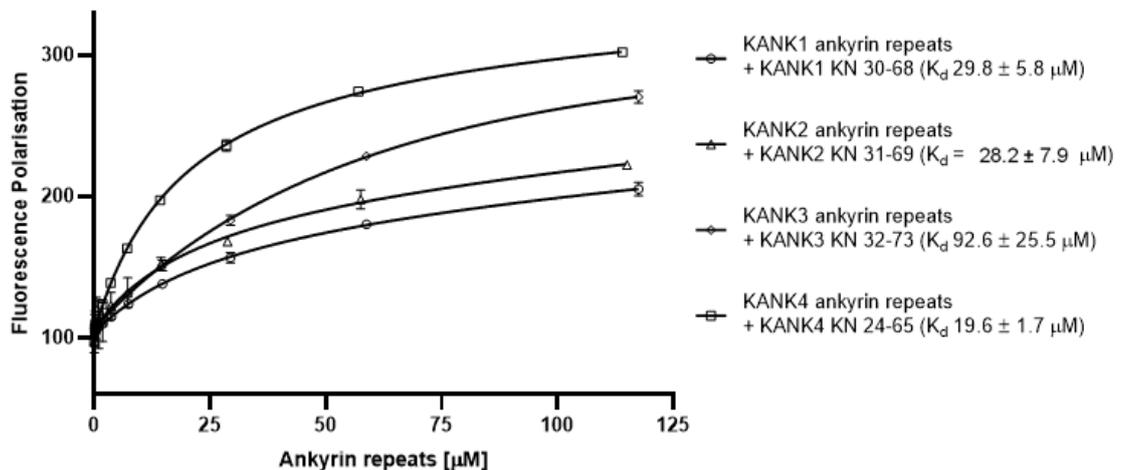


Figure 4.9: An interaction between the ankyrin repeat domain and the KN domain can be observed in all four KANK isoforms.

Binding of KANK ankyrin repeats to BODIPY-labelled peptides of corresponding KANK KN domain using a fluorescence polarisation assay. Dissociation constants \pm SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

Figure 4.10 shows some further analysis as to how the ankyrin repeat domain of each of the isoforms compares to the KANK1 ankyrin repeats. While the KANK2 ankyrin repeat domain appears to interact comparably with the KANK1 KN domain peptides as the KANK1 ankyrin repeat domain, the KANK3 and KANK4 ankyrin repeats domain datasets appear to have some differences. As mentioned, the KANK3 ankyrin repeat domain binds, although more weakly compared to the other isoforms, to the KANK3 KN domain. Despite this, it binds with a K_d of 20.3 μM to the KANK1 KN domain (30-68) peptide. Unlike the KANK1 ankyrin repeats, the KANK3 ankyrin repeat domain still supposedly binds to the truncated KANK1 KN domain (30-60) peptide, although 2-fold weaker interaction compared to the full-length KN domain with a K_d of 41.5 μM and the shape of the curve is not entirely convincing. Although this could suggest that there could be a region of the KN domain upstream of residues 61-68 that the ankyrin repeats domain can interact with, the KANK1 KN domain (30-73 Δ 60-68) peptide still results in a non-determinable K_d value.

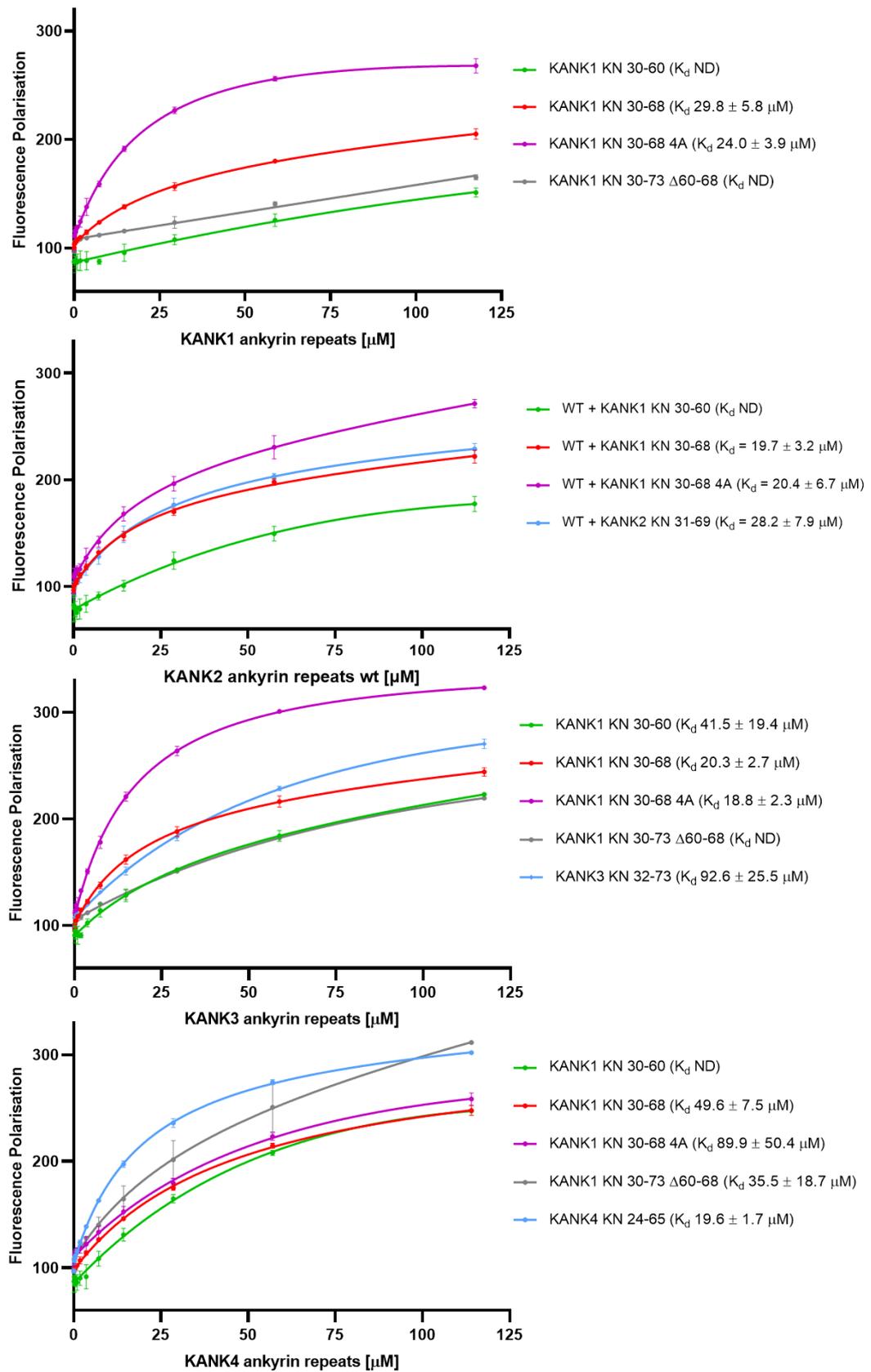


Figure 4.10: All four KANK family proteins interact with the KN domain with some variation. Binding of KANK1-4 ankyrin repeat domains to BODIPY-labelled KANK KN domain peptides measured using a fluorescence polarisation assay. Dissociation constants +/- SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

The dataset for the KANK4 ankyrin repeat domain appears to vary significantly to the KANK1 ankyrin repeat domain dataset. While no dissociation constant could be calculated between the KANK4 ankyrin repeat domain and the truncated KANK1 KN domain peptide (30-60), a K_d of 35.5 μM was calculated for the KANK1 KN domain (30-73 Δ 60-68) peptide. It must be noted here that the latter experiment clearly did not reach saturation, which may have been caused by the measurement of a non-specific element or perhaps aggregation caused by the peptide, resulting in an unreliable K_d and a larger standard error. Of all KANK ankyrin repeat domain isoforms, the KANK4 ankyrin repeat domain bound weakest to the wildtype KANK1 KN domain (30-68) peptide with a K_d value of 49.6 μM . The interaction with the mutant KANK1 KN domain (30-68) 4A peptide is almost 2-fold weaker with a K_d value of 89.9 μM , though the standard error for this calculated value is relatively high. These varied data for the KANK4 ankyrin repeat domain when investigated for binding to KANK1 KN domain peptides – especially when taking into account that the ankyrin repeat domain-KN domain interaction within KANK4 is the tightest of all isoforms – may be caused by the oligomeric differences exhibited by the KANK4 ankyrin repeat domain in comparison to the other isoforms (as discussed in Chapter 3).

In summary thus far, the data in Figure 4.10 indicates that an interaction between the KN and ankyrin repeat domains can occur in all four KANK proteins. Where residues 61-68 are required for this interaction in KANK1, the similarities in binding suggest that the homologous region is required for the interaction in KANK2: residues 62-69. In KANK3 and KANK4, for which slightly longer KN domain regions were used for investigation, there were some notable differences. An interaction between the KN and ankyrin repeat domains of KANK3 occurs, albeit more weakly than the other isoforms. The KANK3 ankyrin repeat binds KANK1 KN domain peptides with similar affinity to the KANK1 ankyrin repeats, suggesting that the differences in the KANK3 KN domain – likely the three-residue insertion – cause the weaker interaction. In KANK4, almost the opposite occurs: compared to the other isoforms, the ankyrin repeat domain does not bind the KANK1 KN domain peptides as tightly but, despite this, exhibits the strongest interaction with its own KN domain.

4.6 Known disease-causing mutations in KANK2 and KANK4 ankyrin repeats may affect autoinhibition

Further to the characterisation of two disease-causing mutations in KANK ankyrin repeats – Ala670Val in KANK2 ANK1 and Tyr801His in KANK4 ANK0 – we decided to investigate whether the interaction between the KN and ankyrin repeat domains observed in the wildtype proteins is affected.

4.6.1 Comparing autoinhibition in wildtype KANK2 ankyrin repeat domain and mutant KANK2 ankyrin repeat domain (Ala670Val)

Using FP, the wildtype and Ala670Val mutant KANK2 ankyrin repeat domain proteins were tested and compared for binding to KANK1 and KANK2 KN domain peptides (Figure 4.11). Although minor differences can be observed in the measured K_d values, which generate overlapping standard error values, the two proteins were very similar overall:

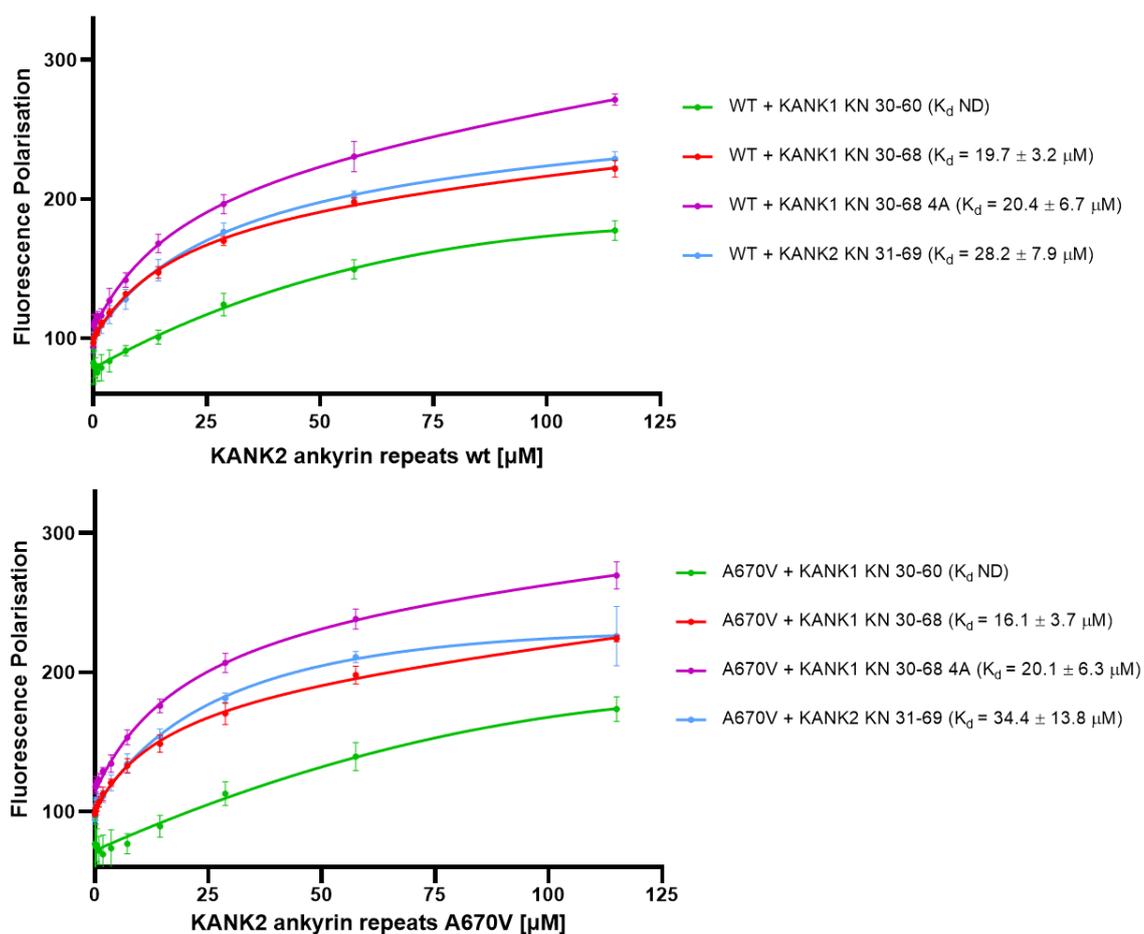


Figure 4.11: Wildtype and Ala670Val mutant KANK2 ankyrin repeat domain proteins interact similarly with KANK1 and KANK2 KN domain peptides.

Binding of KANK2 ankyrin repeat domain proteins to BODIPY-labelled KANK KN domain peptides measured using a fluorescence polarisation assay. Dissociation constants +/- SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

both bind KANK1 and KANK2 KN domain peptides, and interaction with the latter was slightly weaker. This lack of significant difference between the wildtype and mutant forms was expected due to the positioning of the Ala670Val mutation further away from the ankyrin repeat-connecting loops that collectively form the hypothesised binding surface for the KN domain (see section 4.8.1). Additionally, this indicates that the slightly increased protrusion of the valine residue does not disrupt the arrangement of the structural elements of the ankyrin repeats, namely the α -helices, in a manner that affects the interaction between the KN and ankyrin repeat domains.

4.6.2 Comparing autoinhibition in wildtype KANK4 ankyrin repeat domain and mutant KANK4 ankyrin repeat domain (Tyr801His)

Using FP, the wildtype and Tyr801His mutant KANK4 ankyrin repeat domain proteins were tested and compared for binding to KANK1 and KANK4 KN domain peptides (Figure 4.12). Unlike the KANK2 wildtype and mutant ankyrin repeat domains, a slightly more pronounced difference is observed here when comparing binding to KANK1 KN (30-68). Where the wildtype KANK4 ankyrin repeat domain binds to this peptide with a K_d of 45.3 μ M, the Tyr801His mutant has a slightly increased affinity with a K_d of 34.1 μ M. Interestingly, although this difference may appear insignificant, this data for the mutant KANK4 ankyrin repeat domain is more similar to the interactions of the other KANK isoforms with the KANK1 KN domain. This indication could be linked with the findings in Chapter 3: unlike the other isoforms, the ankyrin repeat domain of KANK4 likely favours a dimeric form. CD and SEC-MALS analysis suggested that the Tyr801His mutant instead favours a monomeric form. Consistent with these findings, this mutant appears to interact more similarly to the wildtype ankyrin repeat domains of the other KANK isoforms and this could be because it, too, adopts a monomeric form, potentially increasing the availability for a required binding surface. For both the wildtype and Tyr801His mutant proteins, the interaction with the KANK4 KN (24-65) peptide has the highest affinity. Taking all of these data into account, it seems possible that the wildtype KANK4 exhibits an interaction between its KN and ankyrin repeat domains, though the precise mode of interaction may differ from other the ankyrin repeat domains of other isoforms either in terms of specific residues involved in the binding surface or a different location of the binding surface. Alternatively, if there is occlusion of the typical binding site due to dimerisation, this may be causing the interaction to be weaker due to

reduced contact points. The Tyr801His mutant version of the ankyrin repeat domain may then be able to bind the KN domain either via these otherwise occluded sites in wildtype KANK4 ankyrin repeats which are similar to that in other KANK isoforms or via the site which is used by wildtype KANK4 ankyrin repeats, should these two surfaces be distinct, and therefore displays a higher affinity.

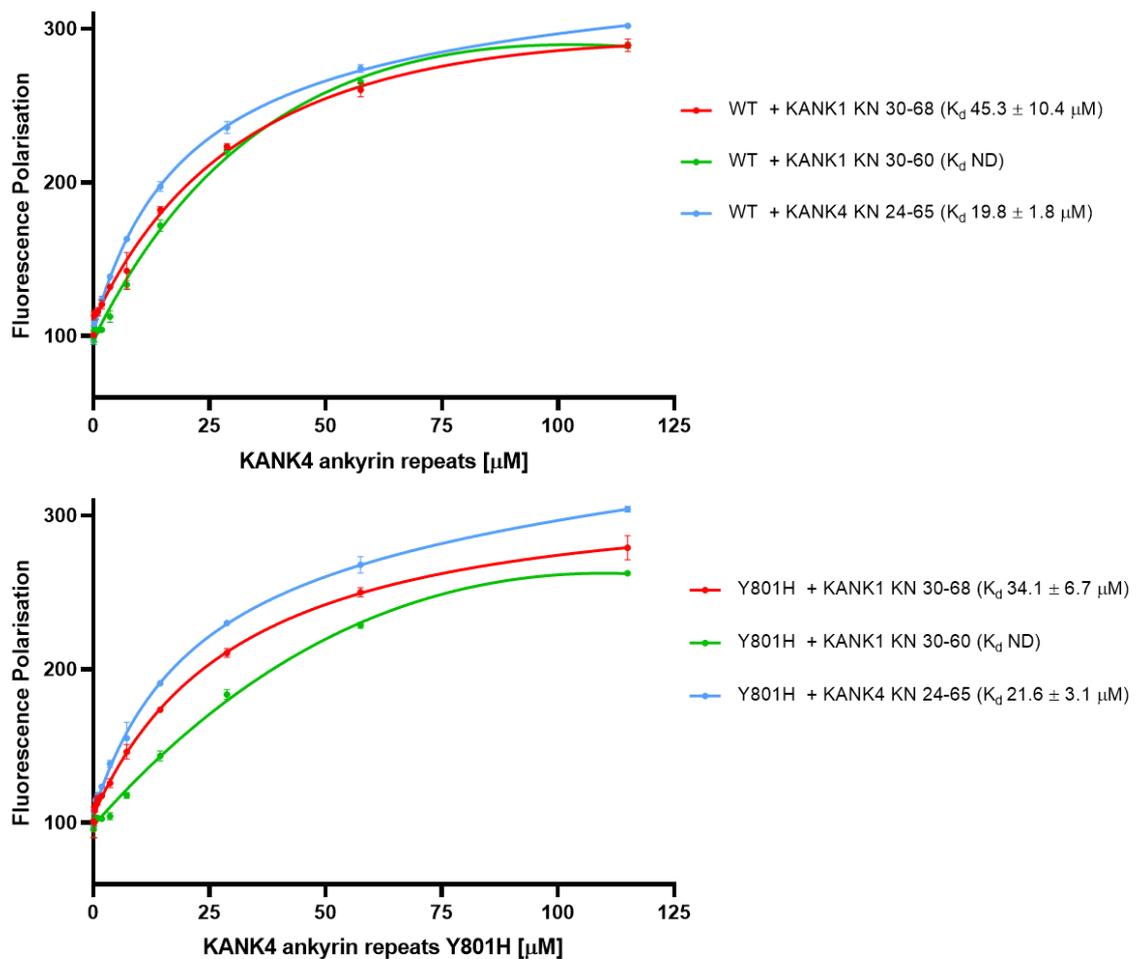


Figure 4.12: Wildtype and Tyr801His mutant KANK4 ankyrin repeat domain proteins vary slightly in their interactions with KANK1 and KANK4 KN domain peptides.

Binding of KANK4 ankyrin repeat domain proteins to BODIPY-labelled KANK KN domain peptides measured using a fluorescence polarisation assay. Dissociation constants \pm SE are indicated in the legend. All measurements were repeated in triplicate. ND, not determined.

4.7 Attempting FRET analysis to investigate KANK autoinhibition

4.7.1 FRET analysis can be used to investigate autoinhibition *in vitro*

Förster resonance energy transfer (FRET) has been used extensively to investigate interactions within a variety of scientific fields. This highly sensitive method works by measuring energy transfer between two fluorophores: when a donor fluorophore is

excited at the appropriate wavelength, the excitation energy is transferred to an acceptor fluorophore within a small distance and causes the acceptor to emit fluorescence at its corresponding wavelength (Forster, 1946). For this method, fusion proteins are created by attaching the protein of interest, via a linker, to a monomeric fluorescent protein – it is crucial that these are monomeric in order to avoid the formation of homodimers or other higher order oligomers of the fluorescent protein which would interfere with the data (Chudakov *et al*, 2010). Additionally, the linker between the two components is typically kept shorter in order to reduce the susceptibility to proteolytic cleavage.

A FRET experiment was planned in order to support the findings thus far that the KN and ankyrin repeat domains interact. mCerulean (blue) was attached at the N-terminus of KANK, generating a mCerulean-KN region, while mVenus (yellow) was attached at the C-terminus, generating a mVenus-ankyrin repeats region (Bajar *et al*, 2016). The experiment entails exciting the mCerulean: if the KN and ankyrin repeat domains do not interact in this system, low or no FRET occurs and mCerulean emission is observed; if an interaction between the KN and ankyrin repeat domains occurs, high FRET occurs wherein the excess excitation energy transfers to the mVenus attached to the ankyrin repeats and thus mVenus emission is observed. A schematic of this is shown in Figure 4.13.

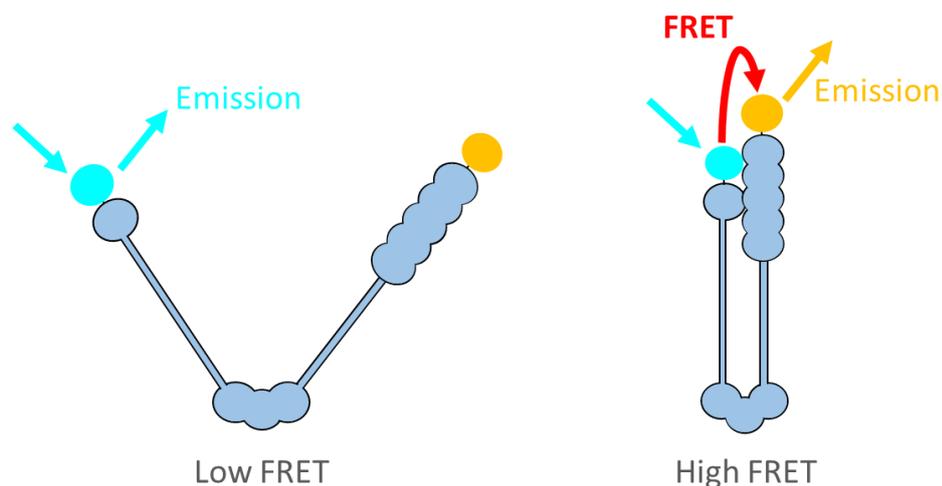


Figure 4.13: A schematic of a FRET experiment using KANK1.

In this model, a FRET experiment begins by exciting mCerulean (blue) at the N-terminus of KANK. If the KN domain at the N-terminus does not interact with the ankyrin repeat domain, no/low FRET occurs and emission for mCerulean is measured (left). If the KN domain does interact with the ankyrin repeat domain at the C-terminus of KANK, fused to mVenus (yellow), high FRET occurs and emission for mVenus is measured (right).

4.7.2 Generating constructs to investigate KANK1 autoinhibition by FRET analysis *in vitro*

A mCerulean-hKANK1-mVenus construct within a mammalian vector was designed by the Akhmanova group at the University of Utrecht for FRET analysis. This construct contained a Ser-Gly-Leu (SGL) linker between mCerulean and the N-terminus of human KANK1, while a Pro-Val-Ala-Thr (PVAT) linker was located between the C-terminus of human KANK1 and mVenus. In order to enable *in vitro* FRET analysis, two constructs were generated by subcloning from this construct into a pET-151 expression vector: 1) mCerulean-SGL-KANK1 KN, containing residues 1-78 of human KANK1, and 2) KANK1 ankyrin repeats-PVAT-mVenus, containing residues 1078-1352 of human KANK1. Once successfully cloned, each fusion protein was expressed and purified as described in Chapter 2.

4.7.3 Investigating KANK1 autoinhibition by FRET analysis

To investigate FRET between mCerulean-tagged KANK1 KN domain (mCer-KN) and mVenus-tagged KANK1 ankyrin repeat domain (mVen-ANK), three control experiments were initially performed. Firstly, 2 mM mCer-KN alone was excited at 425 nm and its emission measured, generating a broad peak between ~450-570 nm, with the majority of the signal at 450-510 nm (Figure 4.14, blue). Secondly, 2 mM mVen-ANK was excited at 490 nm, the ideal wavelength to visualise mVenus signal, generating a peak from ~500-590 nm (Figure 4.14, yellow). Finally, 2 mM mVen-ANK was excited at 425 nm and only a small signal was observed, though this corresponded to the peak observed for mVen-ANK excited at 425 nm (Figure 4.14, brown). Collectively, these controls provide reference spectra for each of the fluorescent proteins alone and validate that they are functioning correctly by generating emission data at appropriate wavelengths.

For the FRET experiment, a sample containing 2 mM mCer-KN and 2 mM mVen-ANK was excited at 425 nm (Figure 4.14, black). The measured fluorescence emission generated peaks corresponding to mCer-KN with no indication of fluorescence emission for mVen-ANK. Had the two domains interacted and caused the fluorescent proteins to be in close enough proximity to generate a FRET signal, fluorescence emission would have been measured with a peak at ~530 nm, corresponding to the mVen-ANK emission. Thus, in these data, a FRET signal has not been measured and this would suggest that the two proteins have not interacted. However, it must be noted that a significant quench in the

fluorescence occurs compared to the mCer-KN alone, an observation that would have been expected in conjunction with an increase at ~ 530 nm had a FRET signal been measured. There are various possibilities as to why this may have occurred, one being that the linkers between the fluorescent protein and the KANK domain of interest is suboptimal. For example, the linker may be too short or too rigid to allow the KANK domain of interest to have the binding surface for the other domain sufficiently exposed, or steric hindrance from the fluorescent protein itself could have contributed. Further, strain placed upon the fluorescent protein may have affected the experiment. Each fluorescent protein consists of a scaffold formed by a β -barrel secondary structure which contains the unique chromophore (Tsien, 1998). The integrity of this β -barrel is crucial for protecting the chromophore and multiple noncovalent interactions contribute to ensure its chemical, thermal and anti-proteolytic stability (Stepanenko *et al*, 2013). Despite this high stability, any strain placed upon the β -barrel and/or its components that could be disrupting its structure would likely immediately affect the chromophore, and this could potentially have contributed to the quenched fluorescence. If this was the case, no FRET would occur due to this complication with the donor mCerulean, preventing transfer of energy from the donor to the acceptor. Further, the N-terminal region of KANK which was part of this fusion protein is relatively short – 78 residues –

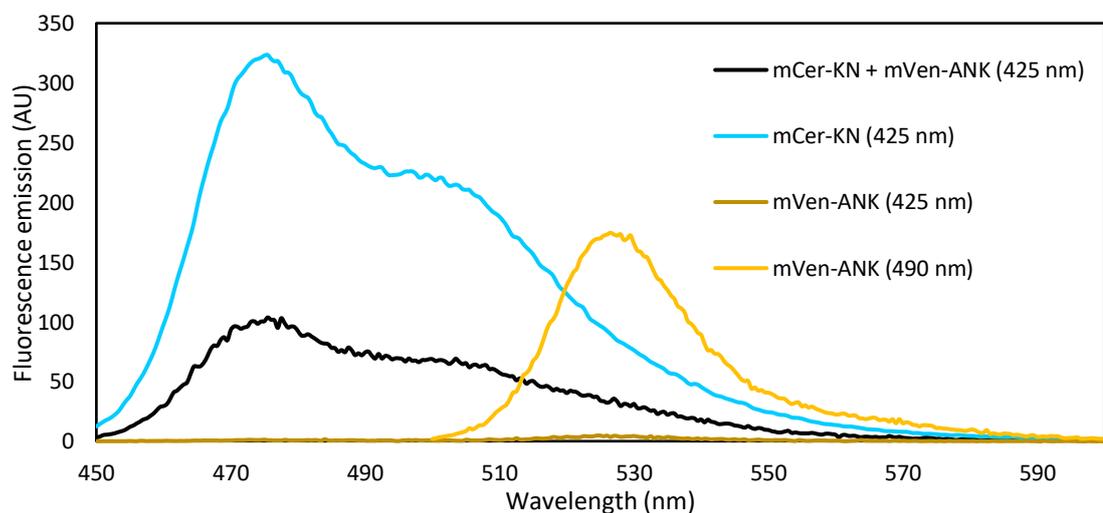


Figure 4.14: Although a quench in emitted mCerulean-KN fluorescence is observed, no FRET is detected when tested with mVenus-ankyrin repeats.

An *in vitro* fluorimeter-based FRET experiment was performed with mCerulean-KANK1 KN domain (mCer-KN) and mVenus-KANK1 ankyrin repeats (mVen-ANK). When mixed, a significant quench in the mCer-KN emission is observed though no emission occurs for mVen-ANK, indicating low/no FRET

with, according to our findings, potentially only the very C-terminal region of this involved in the interaction. This in itself could be a limitation for this experiment.

4.8 Attempting X-ray crystallography to determine the structure of the interaction between the KN and ankyrin repeat domains of KANK

To determine the structure of the complex formed between KANK KN and KANK ankyrin repeat domains and identify the residues involved in the binding surfaces, X-ray crystallography was chosen as a suitable approach. Typically, ankyrin repeat domains throughout biology are widely considered as somewhat prone to crystallisation due to their highly ordered and organised structures consisting of sequential predominantly α -helical regions.

Over the course of this project, a wide range of crystallisation trials were performed. For these, co-crystallisation of various combinations of ankyrin repeat domain proteins and KN domain peptides were attempted, and trials performed using a range of commercial crystallisation screens (Table 4.2). Although co-crystallisation of components from the same isoform of the protein would ideally be performed, many mixed combinations were attempted – for example, KANK2 ankyrin repeats and KANK1 KN domain was considered a useful combination due to these having one of the highest affinity interactions according to FP experiments (Figure 4.10).

In addition to the range of domain combinations and crystallisation screens used, a variety of crystallisation conditions were also attempted: from initial buffer components for both peptide and protein being altered to cover a range of concentrations and buffering systems, including pH and salt conditions, to varying crystallisation solution preparation approaches and incubation temperatures of the completed screens. In terms of preparing the crystallisation solution, the seemingly most successful approach was diluting the protein and the peptide separately by $\sim 10x$ in the crystallisation buffer before mixing, followed by incubation and concentration by ultrafiltration to prepare the final sample. This method appeared to reduce visible precipitation of the peptide when mixing with the protein, a recurring issue throughout the process.

Table 4.2: Attempted co-crystallisation combinations for KANK ankyrin repeats in complex with KANK KN domain.

A selection of the attempted combinations of KANK ankyrin repeat protein and KANK KN peptide subject to co-crystallisation and some of the commercial crystallisation screens used.

Protein	Peptide	Commercial screen
KANK1 ankyrin repeat domain	KANK1 KN 30-68	3D structure Hampton JCSG+ Pi minimal Stura Wizard
	KANK1 KN 30-68 (4A mutant)	Hampton JCSG+ Wizard
	KANK2 KN 31-69	Hampton JCSG+ Wizard
KANK2 ankyrin repeat domain	KANK1 KN 30-68	3D structure Hampton JCSG+ Pi minimal Wizard
	KANK2 KN 31-69	Hampton JCSG+ Wizard
KANK2 ankyrin repeat domain (A670V mutant)	KANK2 KN 31-69	Hampton JCSG+
KANK3 ankyrin repeat domain	KANK1 KN 30-68	Hampton JCSG+
KANK4 ankyrin repeat domain	KANK1 KN 30-68	Hampton JCSG+ Wizard
	KANK1 KN 30-68 (4A mutant)	Hampton JCSG+
	KANK4 KN 24-65	Hampton JCSG+ Wizard

Despite these many approaches used to co-crystallise the KANK ankyrin repeat domain with the KANK KN domain and a large number of stable, single crystals formed, there was no success in generating crystals of the protein with a peptide bound. However, one interesting observation throughout this process was that crystallisation attempts

including a peptide generated a greater number of apo crystals of a higher quality compared with a crystallisation attempt of an intentionally apo ankyrin repeat domain protein.

4.8.1 KN domain peptide presence increases frequency and quality of ankyrin repeat domain crystallisation despite no observed density for the peptide

While working at the Centro de Investigación del Cáncer with the Pereda lab, various simultaneous crystal trials were set up. One of these included plating the KANK2 ankyrin repeat domain on its own using a variety of commercial screens, and a co-crystallisation solution of KANK2 ankyrin repeats with KANK2 KN (31-69) was performed using identically prepared samples at the same time in the same crystallisation screens. Using the Wizard screen which contains 96 unique conditions as an example, more than 20 conditions for the co-crystallisation solution contained single crystals within 24 hours while only eight conditions for the apo protein solution exhibited the same observation. During the process of mounting the crystals onto loops for vitrification in liquid nitrogen to allow X-ray diffraction and data collection to take place, it was observed that the crystals from the co-crystallisation solution were very stable, whereas the crystals from the apo crystallisation solution were very fragile and prone to cracking or being completely destroyed. Finally, the quality of the crystallographic data acquired varied quite significantly: the highest resolution for the co-crystallisation solution which included the KANK2 KN domain peptide was 1.9 Å while the highest resolution for the apo crystallisation solution of KANK2 ankyrin repeats only was 3.2 Å. These two datasets were subject to processing, molecular replacement and several rounds of refinement and manual model adjustment and the resulting structures from these steps are shown in Figure 4.15. Despite these stark variations observed throughout the process – with the only difference in the set up being either the presence or absence of the KANK2 KN peptide – no density was observed for the peptide.

When comparing the two structures, there appear to be some clear differences: ANK0 of the apo KANK2 is unstructured in places where the protein that was crystallised with the KN domain peptide possesses the predicted α -helical elements. Additionally, some small positioning differences can be observed, particularly in ANK1 and in the loop of ANK3. It must of course be noted that the lower resolution data for the apo protein is

likely a contributing factor for these changes, particularly in ANK0 where the data quality was particularly poor and lacking density in places.

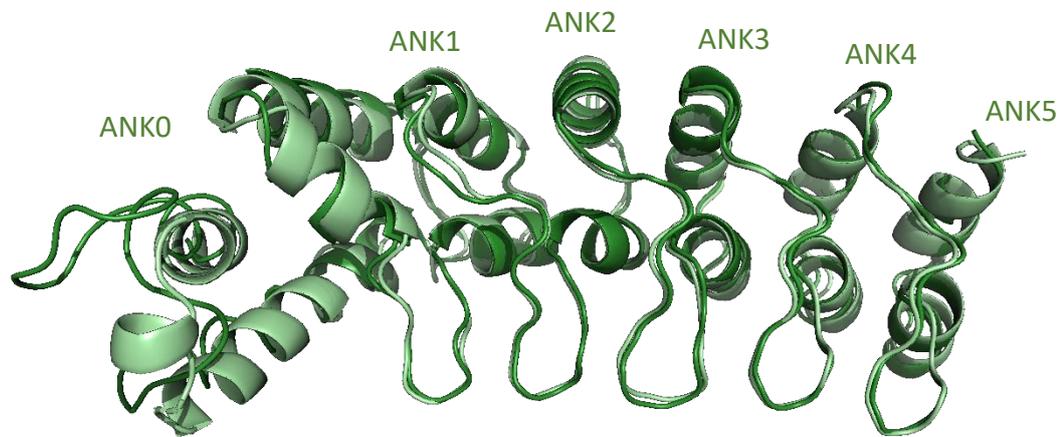


Figure 4.15: Refined structures of KANK2 ankyrin repeats crystallised alone and in presence of KANK2 KN 31-69 peptide.

The structures of KANK2 ankyrin repeats crystallised alone (light green) and in the presence of the KANK2 KN 31-69 peptide (dark green) are shown here overlaid. Images produced with PyMOL.

In terms of existing KANK structures in the PDB, few exist. They include human KANK1 ankyrin repeats and KANK2 ankyrin repeats, each in their apo forms (PDB IDs: 5YBJ and 4HBD, respectively) and their KIF21A-bound forms (PDB IDs: 5YBU and 5YBV, respectively), in addition to structures of these same domains from the murine KANK proteins. Prior to the structures of the ankyrin repeat domains in complex with the KIF21A peptide, the primary hypothesis was that the KN domain forms an additional helix which aligns itself or packs in adjacently to the α -helices of the ankyrin repeat domains, akin to how ligands including the KANK KN domain may bind to the talin rod domain. Thus, significant changes in the protein conformation, and likely therefore a change in the space group for generated crystals, was expected. However, the published structures for the KANK1 and KANK2 ankyrin repeat domains each in complex with a KIF21A peptide indicate that this peptide actually binds somewhat perpendicularly along the loops which connect each ANK region, with intermolecular interactions formed to that between ANK1-ANK2, ANK2-ANK3 and ANK3-ANK4. Most interestingly, this complex does not appear to dramatically alter the conformation of the protein and thus the space group is also identical for each KIF21A-bound protein and its apo form – for example, both apo KANK2 ankyrin repeats (PDB ID: 4HBD) and KANK2 ankyrin repeats in complex with KIF21A (PDB ID: 5YBV) are in the space group $P2_12_12_1$ (Figure 4.16). Based on this information, the KN domain of KANK may bind to the ankyrin repeats in a

very similar way and, if mostly unstructured, it would be difficult to differentiate between an apo or ligand-bound version of an ankyrin repeat protein based on space group determination and perhaps also without very high resolution data to confidently observe density for the KN domain peptide. Binding onto a very similar or equal binding surface to the KIF21A peptide may allow a form of competition *in vivo*, where the KANK

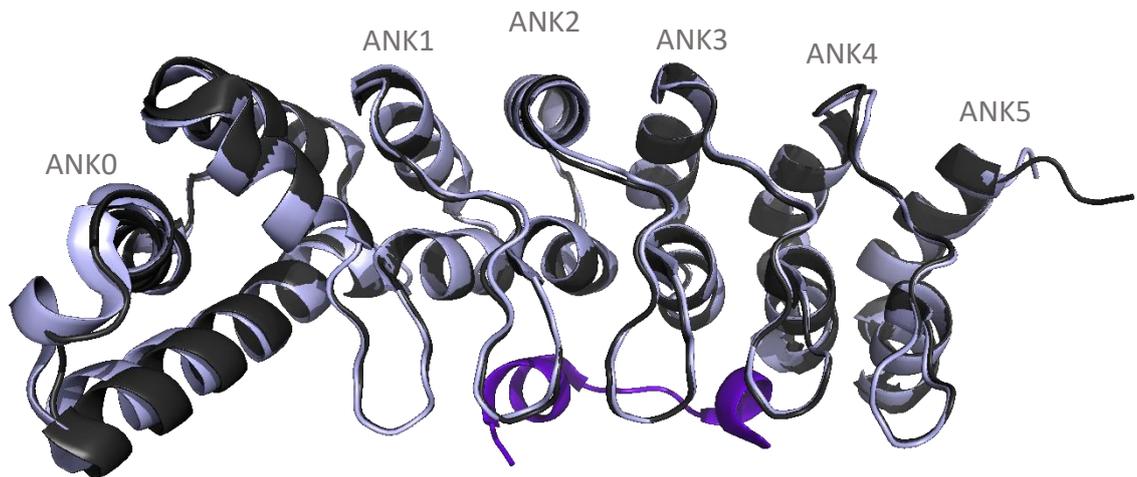


Figure 4.16: Published structures of KANK2 ankyrin repeats crystallised alone and crystallised with KIF21A peptide.

The structures of KANK2 ankyrin repeats crystallised alone (light grey) from PDB ID 4HBD and in the presence of the KIF21A peptide (dark grey) from PDB ID 5YBV (Guo et al, 2018) are shown here overlaid. The KIF21A peptide is shown in purple. Images produced with PyMOL.

protein, via the ankyrin repeats, is either autoinhibited (i.e. KN domain bound) or bound to KIF21A and providing the link to the microtubule cytoskeletal network, with these two interactions being mutually exclusive.

4.8.2 Designing a “mini-KANK” for structural experiments

Another factor which is likely preventing a crystal structure for the KANK KN-ankyrin repeat domain complex is the affinity of the interaction. Typically, higher affinity interactions will have a higher propensity to crystallise as a complex, as would a catch bond interaction wherein increasing force can increase the lifetime of a bond – an example of the latter is the interaction between the vinculin tail and actin (Huang *et al*, 2017) and, more recently, the interaction between the KANK KN domain and talin R7 has also been identified to exhibit catch bond kinetics at forces below 6 pN (Yu *et al*, 2019). The interaction between the KN and ankyrin repeat domains of KANK varies according to the isoform but is ~20-30 μM for the majority of KANK isoforms (Figure 4.10) – this affinity is considered fairly weak for co-crystallisation. In order to counter

this and increase the likelihood of crystallisation of the complex, various strategies were attempted. One example of this was including a large molar excess of the peptide in the crystallisation solution, but this also was not successful in the trials carried out.

Another strategy employed was the creation of a chimeric protein which contains both the KN and ankyrin repeat domains in one construct, tethering them together to increase the affinity and therefore increase the probability of the observing the ankyrin repeat domain in the bound state with the KN domain present. In doing so, this protein is actually more reminiscent of the full-length KANK protein although a significant portion is missing. The construct was designed using human KANK1 domains: the KN domain (residues 30-73) linked to the ankyrin repeat domain (residues 1078-1328) via a GGGGSGGGGSGGGGS linker. This is a $(GGGGS)_n$ linker, widely known as an option which provides flexibility and allows interaction between domains (Chen *et al*, 2013). Using this mini-KANK chimeric protein, the plan was to use SEC-MALS to identify and collect fractions corresponding to an eluted complex. Theoretically, a variety of species could elute when subjected to size-based separation and some of these are depicted schematically in Figure 4.17. One option would be the two domains not interacting with each other, or alternatively a bound “monomeric” form wherein the KN domain of one mini-KANK interacts with the ankyrin repeat domain of the same molecule. Other potential species include a “dimer” wherein each domain interacts with the other domain of another mini-KANK, or potentially higher order oligomeric structures which essentially form “strings” of mini-KANK proteins. If the “dimer” or higher order oligomer can be identified from the SEC-MALS analysis – i.e. the oligomeric forms most likely to contain an interaction between the two domains – each of these would be collected. The SEC-MALS experiment would then be repeated to confirm that the sample is not in a form of equilibrium between multiple oligomeric states and rather does, in fact, remain in the same oligomeric state. If confirmed, this eluted species that is collected would be subject to crystallisation trials.

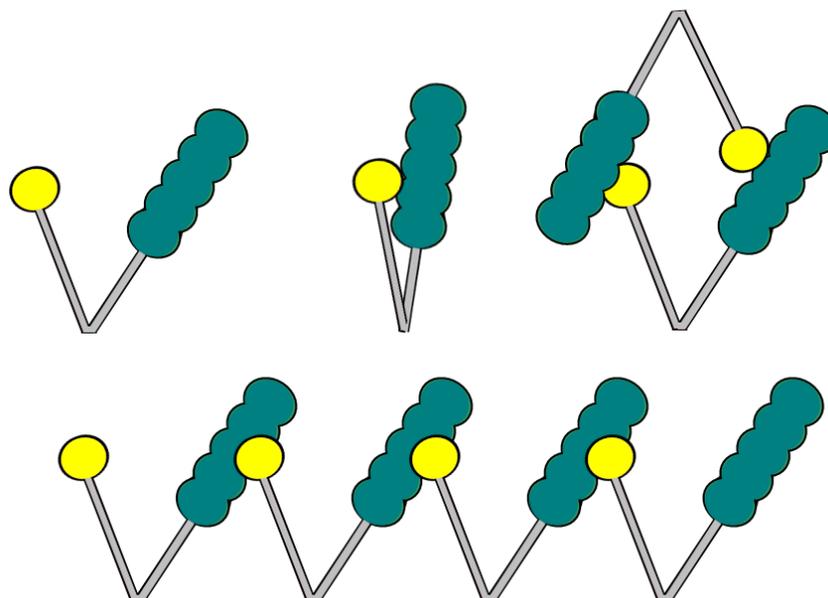


Figure 4.17: Potential mini-KANK oligomeric states.

Some of the potential oligomeric states of the mini-KANK protein are depicted here. The KANK1 KN domain is shown in yellow and the KANK1 ankyrin repeat domain is shown in turquoise. The linker between the two domains is shown in grey.

As a preliminary assessment step, FP was performed. For this experiment, the KANK1 ankyrin repeat domain protein and the mini-KANK protein were both tested with the KANK1 KN domain (30-68) peptide. The logic behind this experiment was that, if the ankyrin repeat domain within mini-KANK is occupied by the KN domain of the mini-KANK, the fluorescent KANK1 KN domain 30-68 which is tested against the protein will be competing for the binding surface on the ankyrin repeat domain. Thus, if the mini-KANK protein is functioning as intended, the calculated affinity for binding the fluorescent KANK1 KN domain 30-68 peptide should be weaker. When performing this experiment, the two K_d values calculated are very similar, with mini-KANK actually exhibiting a slightly tighter affinity (Figure 4.18). This data suggests that the mini-KANK protein was not a state where the intended SEC-MALS experiments would be effective in generating a bound protein suitable for crystallisation. Despite this, preliminary analysis via SEC-MALS suggested that the species was mainly monomeric. Fractions corresponding to a small potential dimer peak were collected and the SEC-MALS experiment repeated, and an identical spectrum was observed with a large monomer peak – indicating that the system is in an equilibrium wherein the monomeric form is favoured (data not shown). This approach to generate a pre-bound KN-ankyrin repeat domain complex for crystallisation is a promising one but clearly needs some

optimisation. Some possibilities for this may be altering the position of the KN domain relative to the ankyrin repeat domain or adapting the linker. On one hand, the linker may not be long enough to provide the required flexibility for the KN domain to be able to reach the correct binding surface on the ankyrin repeats but, equally, increasing the length of this linker may increase the susceptibility to proteolytic cleavage. Proteolytic cleavage in this mini-KANK protein could potentially also already be a contributing factor as to why this protein behaves so identically to the KANK1 ankyrin repeats.

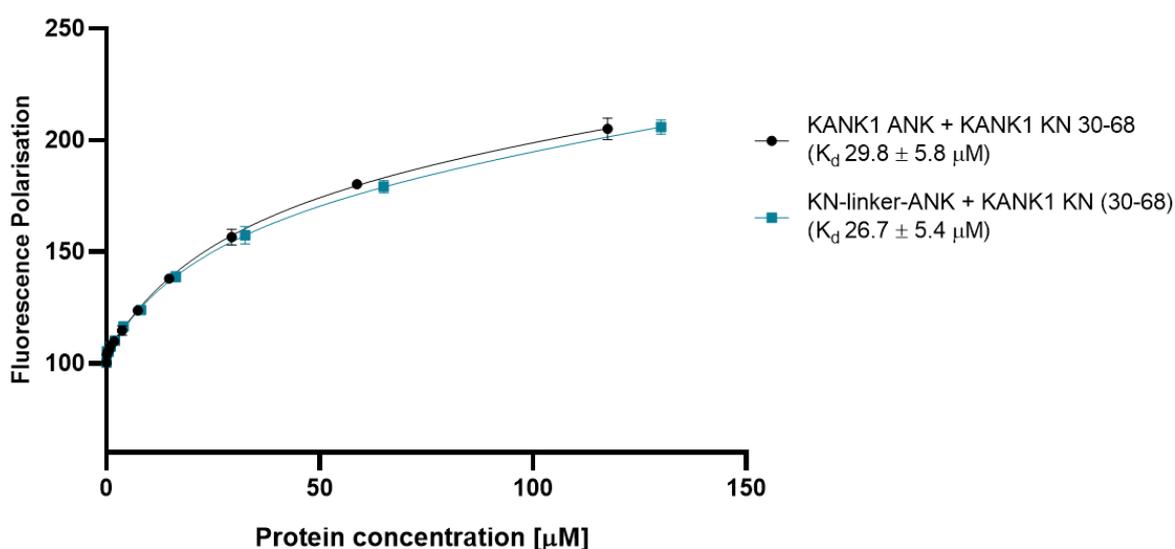


Figure 4.18: A mini-KANK protein binds to the KANK1 KN domain equally to the KANK1 ankyrin repeats.

A mini-KANK protein (KANK1 KN-linker-ANK) was designed to increase the affinity of the KN-ankyrin repeat domain interaction by tethering these two domains together. Binding of mini-KANK and KANK1 ankyrin repeat domain proteins to BODIPY-labelled KANK1 KN 30-68 domain was measured using a fluorescence polarisation assay. Dissociation constants \pm SE are indicated in the legend. All measurements were repeated in triplicate.

4.9 Discussion

In this chapter, an interaction between the KN and ankyrin repeat domains of various KANK proteins was identified – this includes all four mammalian isoforms in addition to two disease-causing mutation-containing ankyrin repeat domain proteins. Although all of these versions of the protein demonstrate the interaction, they do so in various ways – most notably, KANK3 has the weakest affinity between its KN and ankyrin repeat domains of all combinations tested, and this may be as a result of a 3-residue insertion located within an essential region of the KN domain. This region of the KANK KN domain, residues 61-68 in KANK1 and the corresponding residue numbers in the other isoforms, were found to be essential for this interaction to take place. These findings were shown using a combination of approaches, with particular focus on FP and NMR titration

methods. NMR titrations revealed that some very small changes can still be observed with the KN domain peptide spanning residues 30-60 and also the 30-73 Δ 60-68 peptide, suggesting that there may also be a small region upstream of residues 61-68 which could be binding weakly and contributing to the stabilisation of the overall complex between the KN and ankyrin repeat domains.

Identifying the specific binding surface on both domains and the specific interactions involved would be a significant step forward for this project. While x-ray crystallographic methods were attempted extensively and were thus far unsuccessful in generating a structure of the complex, further optimisation of the discussed approaches, including the mini-KANK construct, may be useful. The relatively weak interaction between the KN and ankyrin repeat domains is likely one of the most significant contributing factors to the lack of success. While the relatively low affinity may be viewed by some as questionable to persist investigating, this weaker affinity is logical – where the two domains are tethered together *in vivo*, and likely as part of an autoinhibition mechanism adopted when inactive, the lower affinity compared with the KN-talin R7 and ankyrin repeat domain-KIF21A interactions would actually allow these two important KANK ligands to preferentially bind, activating KANK as and when required within the cell. Further investigation into the potential competition between talin R7 and the ankyrin repeat domain for the KN domain and the competition between the KN domain and KIF21A for the ankyrin repeat domain would therefore be very interesting to explore. Another important factor which may be preventing the crystallisation of the complex is the unstructured nature of the KN domain. While we have hypothesised that a helix addition mechanism of the KN domain occurs upon binding, it is important to consider that this region may remain partially or fully unstructured, which could potentially affect the observation of density for this component if only few specific contacts are made – especially if no significant structural changes are induced in the ankyrin repeats, such as in KIF21A binding.

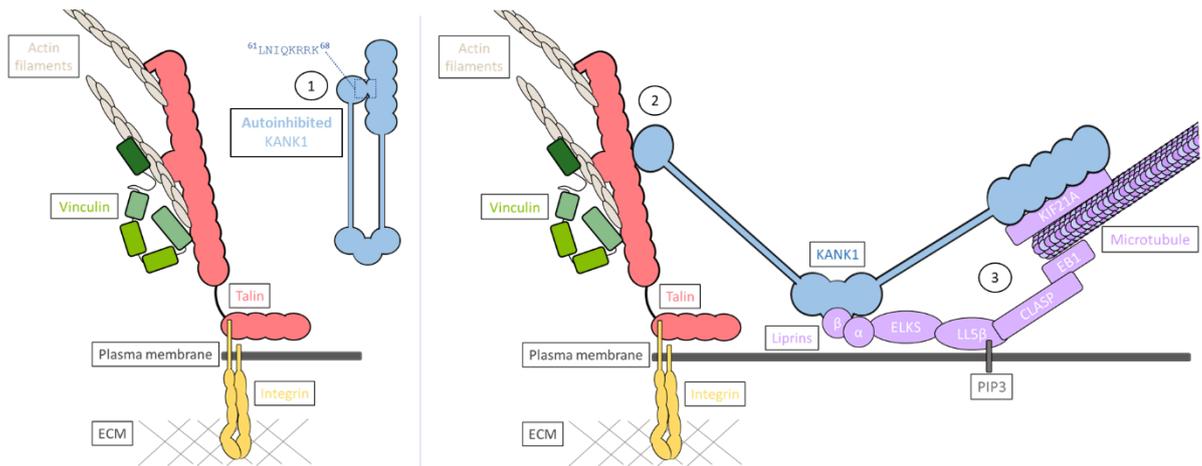


Figure 4.19: Schematic model of the role of KANK autoinhibition in proximity to integrin-based adhesions. 1) KANK1 may exist in an autoinhibited conformation in proximity to the developing or mature integrin-based adhesion, and this autoinhibition requires residues 61-69 which are located at the C-terminus of the KN domain. 2) KANK1, now activated, binds talin R7 via the KN domain. 3) CMSC components are recruited to the proximity of the adhesion following the KANK1-talin interaction.

The greatest challenge with this body of work was supporting the findings using more *in vivo* methods. While, biochemically and biophysically, we can establish consistently that an interaction between the KN and ankyrin repeat domain occurs, validation in mammalian cells has been challenging. Some factors to consider are that this interaction between the KN and ankyrin repeat domains could be working in a tissue-specific manner or prevalent during a different stage of development than the mammalian cells so far used for investigation. It is also possible that particular signalling components which are lacking in the somewhat artificial mammalian cell culture system are required to influence cellular activity related to the KN-ankyrin repeat interaction. Collaborations with various academics to further characterise this interaction in a more *in vivo* environment are ongoing, particularly with the Akhmanova group at Utrecht University, with the aim of identifying the consequences on structure, function and activity within mammalian cells as a result of a constitutively active KANK in the form of a $\Delta 60-68$ version of the protein.

Autoinhibition, typically in the form of a head-tail interaction, has emerged as a key regulatory mechanism amongst integrin adhesion-associated proteins. One increasingly developing hypothesis is the formation of pre-complexes: the coalescence of proteins in the vicinity of the adhesion ahead of assembly (Khan & Goult, 2019). In terms of what the role of autoinhibited KANK may be, one hypothesis is that it exists in this

conformation prior to its association with talin (Figure 4.19) – and, as adhesions assemble and talin is extended, perhaps R7 of the talin rod domain outcompetes the KANK ankyrin repeat domain for the KANK KN domain. KANK may exist in a pre-complex with the liprin proteins, an interaction which could be occurring simultaneously to KANK autoinhibition, or in association with the other CMSC proteins which have previously been hypothesised to be recruited after the interaction with talin has taken place (Bouchet *et al*, 2016). Evidence for pre-complex formation for core adhesion proteins has been observed via fluorescent fluctuation analysis experiments which show association between talin and vinculin prior to that between integrin and talin in a nascent adhesion (Bachir *et al*, 2014), knock-sideways experiments which show pre-complex formation between paxillin and autoinhibited talin and vinculin (Atherton *et al*, 2019), and a pre-complex between integrin and kindlin has also been suggested (Rossier & Giannone, 2016). Together, these studies support the hypothesis that the presence of the key proteins localised together – typically in their autoinhibited forms – could be a requirement for rapid and successful assembly of adhesions. With this finding that KANK proteins also autoinhibit, perhaps advancing techniques within this field could be applied to the investigation of pre-complex formation between KANK proteins and CMSC components in the vicinity of integrin-based adhesions and the effect this may have on adhesion and microtubule dynamics.

CHAPTER 5: IDENTIFYING A NOVEL LIGAND FOR KANK1 IDR2

5.1 Overview

In this chapter, we sought to identify a ligand for the unstructured region of KANK. Based on a literature search, analysis of proteomics data and use of online databases, the hub protein LC8 was identified as a promising candidate. Here, an interaction between KANK and LC8 family proteins is explored. SEC-MALS analysis is used to confirm that a human LC8 equivalent, DYNLL2, also forms a pH-dependent dimer as observed in the literature for *Drosophila* LC8, and triple resonance NMR is used to assign DYNLL2 and map the interaction with a novel ligand, KANK1.

5.2 Large unstructured regions within KANK1

All four mammalian KANKs contain regions predicted to be unstructured between the major structured elements that are the KN domain, the coiled coil region and the ankyrin repeat domain. These intrinsically disordered regions (IDRs) vary between each isoform: for example, KANK1 contains larger unstructured regions, most notably a 580-residue region located between the coiled coil region and ankyrin repeat domain (IDR2) (see Figure 1.10). Interestingly, despite these regions in KANKs being implied as unstructured throughout literature (Zhu *et al*, 2008; Kakinuma *et al*, 2009), structure prediction programmes such as PSIPRED (Jones, 1999; Buchan & Jones, 2019) indicate that various areas throughout have secondary structure propensity, predominantly for β -strand but also some α -helical regions (Figure 5.1).

An increasingly prominent pattern within scientific literature is that, upon binding to a ligand, some IDRs can transition to have increased order and/or fold into secondary or tertiary structural elements. A well-cited example of this is the phosphorylated kinase-inducible domain (pKID) of cyclic-AMP response element-binding protein (CREB) which, although unstructured when free in solution (Richards *et al*, 1996; Radhakrishnan *et al*, 1997), folds upon interaction with KID-binding (KIX) domain of CREB-binding protein (CBP) (Radhakrishnan *et al*, 1997). In this example, the region of pKID which is involved in this interaction shows helical propensity when subject to amino acid sequence analysis (Dyson & Wright, 2005). Another example of a IDR transitioning to an

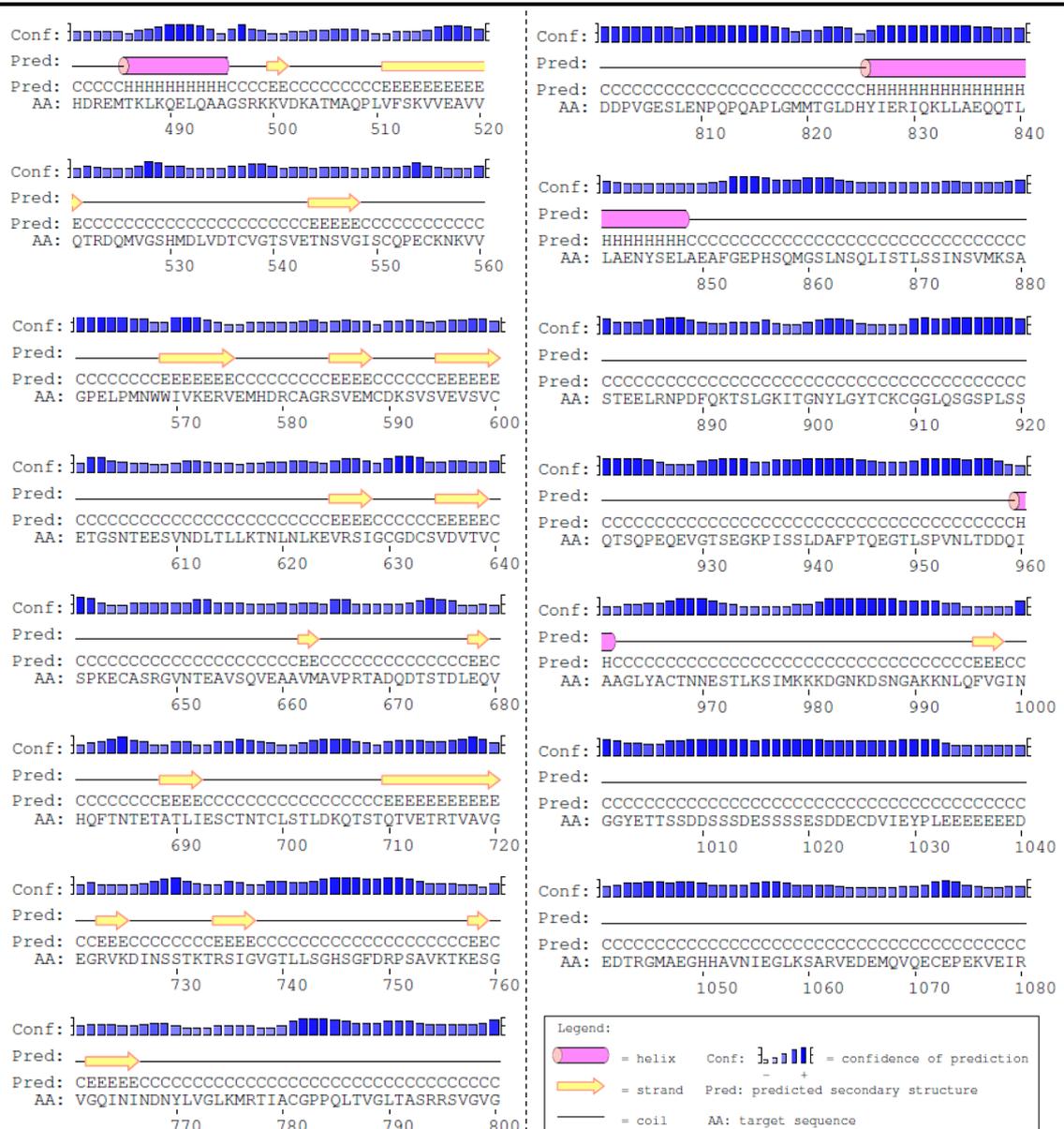


Figure 5.1: Secondary structure prediction of KANK1 480-1080 using PSIPRED.

Secondary structure prediction by PSIPRED (Jones, 1999; Buchan & Jones, 2019) of what is referred to in literature as a large unstructured region within KANK1 (501-1080, IDR2) reveals several smaller regions with secondary structure propensity.

ordered/folded state is the helix addition mechanism which occurs for some LD motifs binding to their ligands: for example, while paxillin LD2 exists as a stable helix in solution, LD4 is mostly random coil as a free peptide in solution – yet both of these form a helix when bound to the focal adhesion-targeting domain of focal adhesion kinase (Bertolucci *et al*, 2005; Alam *et al*, 2014).

KANK1 501-1080 (IDR2), the largest region suggested to be intrinsically disordered, exhibits large areas of high evolutionary conservation throughout (Figure 5.2) (Ashkenazy *et al*, 2016), suggesting that these residues might have essential functions.

Many of these regions with high degrees of conservation overlap with those that have secondary structure propensity.

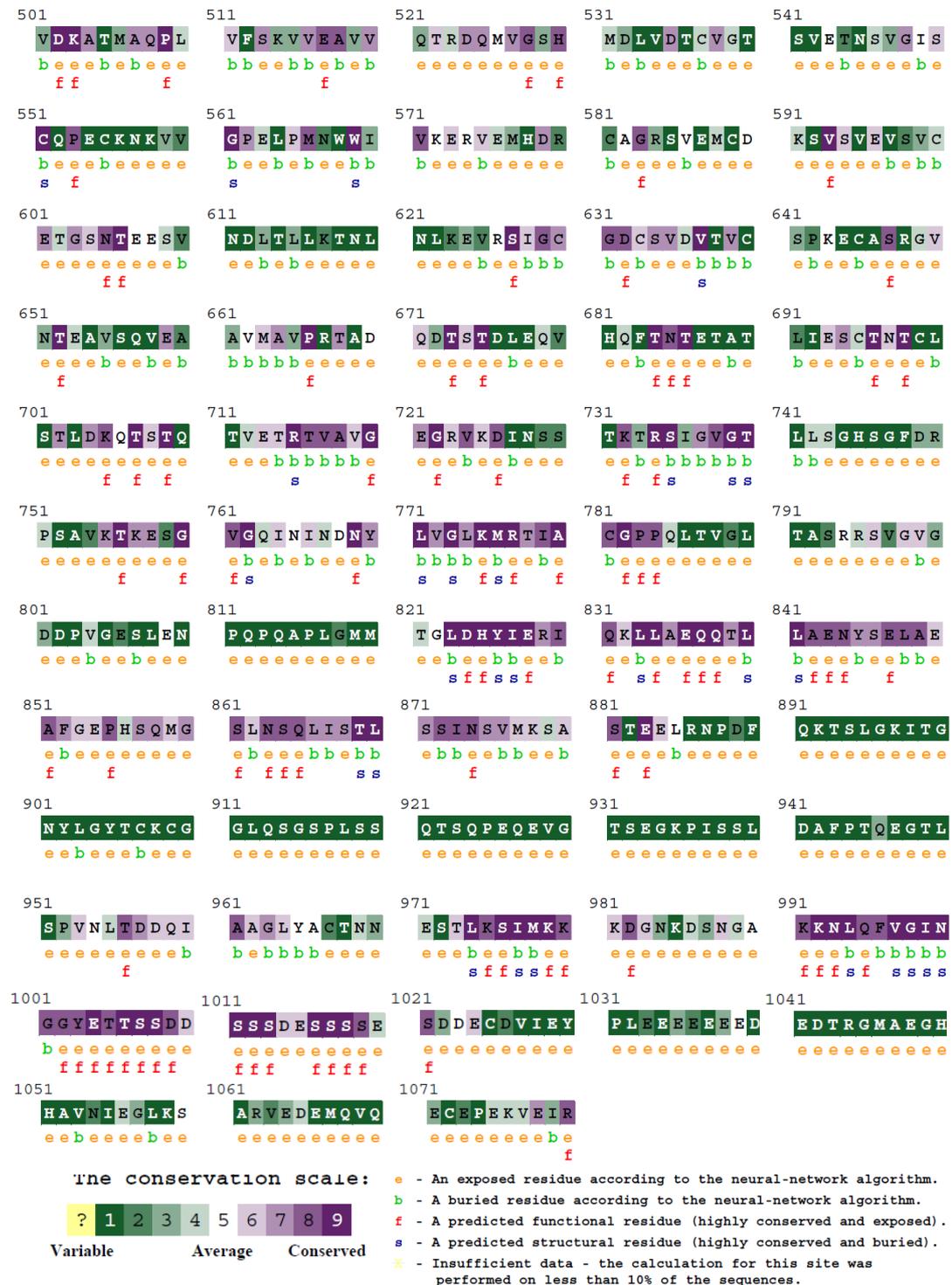


Figure 5.2: Evolutionary conservation of KANK1 IDR2 (501-1080) determined using ConSurf. Evolutionary conservation of what is referred to in literature as a large unstructured region within KANK1 (501-1080, IDR2) reveals several smaller regions with high conservation, likely implicated in vital structural and functional roles.

In an attempt to identify potential ligands for this region of KANK, the KANK1 proteomics dataset obtained by the Akhmanova group at Utrecht University was interrogated. Two interesting hits were dynein light chain 1 (DYNLL1) or LC8, and dynein light chain 2 (DYNLL2) or LC8b (Table 5.1). These proteins are discussed further in Section 5.3.

Table 5.1: GFP-KANK1 proteomics data.

Data from full-length GFP-KANK1 proteomics study carried out by the Akhmanova group (Utrecht University) indicating that DYNLL1 and DYNLL2 are two significant hits (red).

Protein	Accession number	Σ Unique Peptides
KIF21A	Q7Z4S6	75
Talin1	Q9Y490	70
KANK1	Q14678	63
Liprin β 1	Q86W92	49
Talin2	Q9Y4G6	47
Scribble	Q14160	37
Liprin α 1	Q13136	11
DYNLL2	Q96FJ2	5
DYNLL1	P63167	5
Tensin3	Q68CZ2	1

LC8pred, a prediction software based on a proteomic phage display library screen that can predict LC8-binding motifs within a protein sequence with reported 78% accuracy (Jespersen *et al*, 2019), was used to determine potential LC8-binding sites within KANK1. 69 potential sites were identified, of which 31 are located within IDR2, including the top four hits. These are discussed further in Section 5.7.1.

5.3 Dynein light chain proteins: the LC8 family

5.3.1 An introduction to dynein light chain proteins

Cytoplasmic dynein is the major minus-end directed microtubule motor protein and is involved in a variety of cellular activities including, most notably, its roles in the formation and orientation of the mitotic spindle during mitosis (Busson *et al*, 1998; Sharp *et al*, 2000; Vallee *et al*, 2004) and the transportation of various cargos including protein complexes and membrane-bound organelles from the cell periphery towards the centre of the cell during interphase (Reck-Peterson *et al*, 2018). The dynein motor

complex is comprised of heavy, intermediate, light intermediate and light chain polypeptides. Dynein light chain proteins are further split into three groups: the t-complex-associated family, the Roadblock family, and the LC8 family (Pfister *et al*, 2006). LC8 proteins are highly conserved, 10 kDa elements of the dynein complex and bind directly to the dynein intermediate chain. Humans have two isoforms of LC8: dynein light chain 1 (DYNLL1/LC8) and dynein light chain 2 (DYNLL2/LC8b). These two proteins share >93% sequence identity, differing only in six of 89 residues (Wilson *et al*, 2001). Although each protein is encoded by genes located on different chromosomes (12 and 17, respectively), knockout or knockdown of DYNLL1 causes embryonic lethality in various animal models (Dick *et al*, 1996; Lightcap *et al*, 2009; Li *et al*, 2015), suggesting non-redundancy between the two mammalian paralogues, despite no clear functional difference having been observed between the individual DYNLL1 and DYNLL2 proteins (Radnai *et al*, 2010). Some examples of target specificity have been suggested in the literature (Puthalakath *et al*, 1999, 2001; Day *et al*, 2004), though many ligands show no significant difference in binding affinity between the two proteins. It is worth noting, however, that DYNLL2 expression levels are generally reduced (Jurado *et al*, 2012b), though overall expression levels in different tissues can vary significantly, and the transcriptional activator ASCIZ only regulates DYNLL1 expression (Jurado *et al*, 2012a; Zaytseva *et al*, 2014). The notation of DYNLL1/DYNLL2 will be used throughout this work to refer to the mammalian forms of the protein, with LC8 being used as a general term for any LC8 family proteins including the *D. melanogaster* homologue.

5.3.2 LC8 proteins exist as pH-dependent dimers

Both DYNLL1 and DYNLL2, in addition to other well-characterised LC8 family proteins, exist as dimers and both the apo and bound forms of each protein are symmetrical structures. Each monomer consists of a β -strand at the N-terminus followed by two α -helices and a further four β -strands (Liang *et al*, 1999). The dimer interface occurs along the second and third of the four consecutive β -strands at the C-terminal end (Figure 5.3A, Figure 5.3B). At physiological pH, this protein exists as a dimer though it has been widely observed that $\text{pH} < 4.8$ causes dissociation, resulting in a shift to a folded monomeric form (Barbar *et al*, 2001). Interestingly, this pH-dependent dimerisation is unique to the LC8 family of dynein light chain proteins. This has since been identified to

be caused by the protonation of a pH-sensitive histidine residue at position 55 in the protein chain which protrudes into the dimer interface from each subunit (Figure 5.3C) (Nyarko *et al*, 2005). Additionally, phosphorylation has been identified to play a significant role: a study involving *in vitro* phosphomimetic mutants at position 88 of the protein chain (Ser88Glu) promotes dissociation of the dimer, shifting the equilibrium to favour the folded monomeric form. This dimer dissociation eliminates a binding groove essential for many LC8 family ligands, including intermediate chain proteins of the dynein complex, and may provide a mechanism for discriminating between different LC8 binding partners (Song *et al*, 2007, 2008).

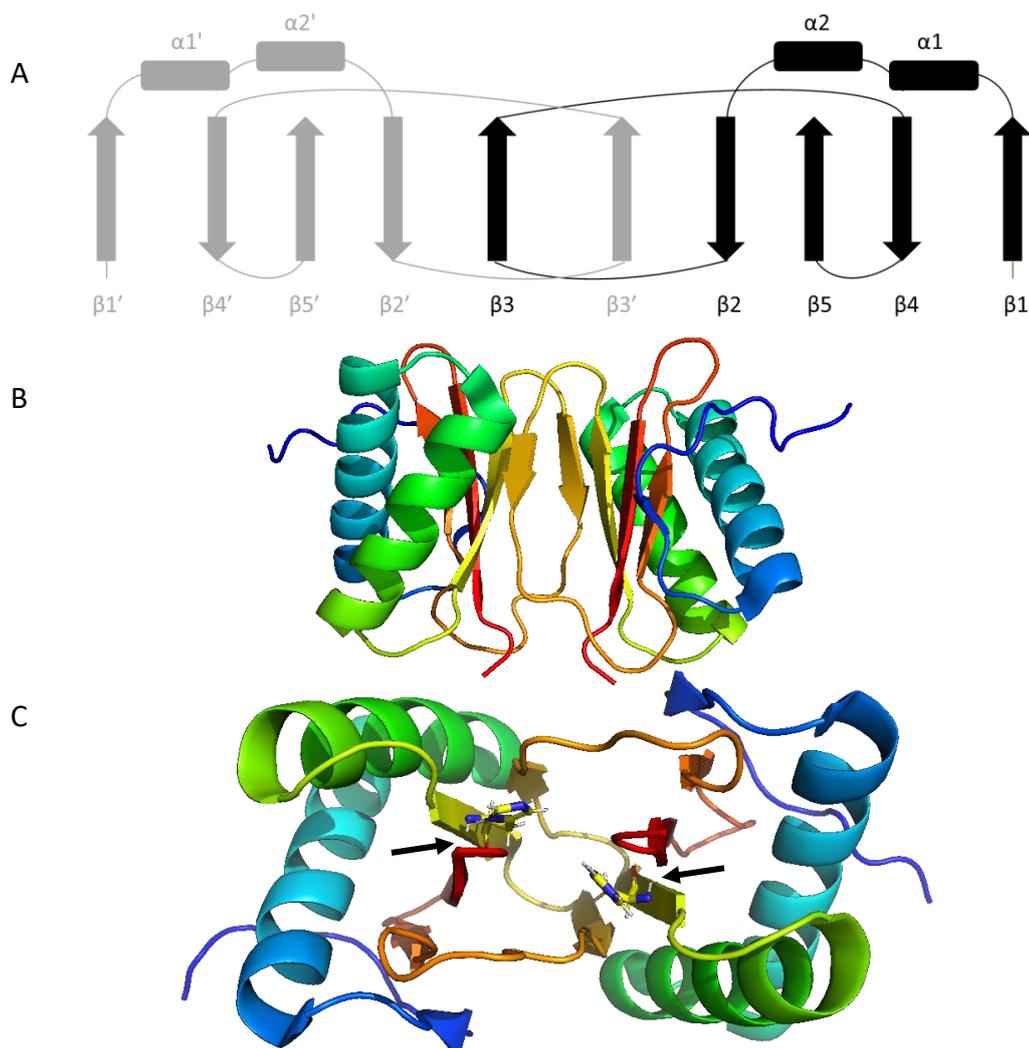


Figure 5.3: The structure of the DYNLL1 dimer.

(A) Simplified schematic of secondary structure organisation of LC8 proteins, highlighting the strand-switch of $\beta 2$ and $\beta 3$ from each monomeric component to create the dimeric interface. Based on figure from Makokha *et al*, 2004. (B) The solution structure of the *Rattus norvegicus* DYNLL1 symmetrical dimer. Rainbow colouring is used to highlight the symmetry of the molecule: the N-terminus of each subunit is shown in blue, resulting in a C-terminus shown in red. (C) The dimer interface of the molecule includes a histidine residue at position 55, shown as sticks, which protrudes into the cavity. Images produced with PyMOL using NMR structure adapted from PDB ID: 1F3C.

5.3.3 LC8 as a hub protein

In addition to its interaction with the N-terminal domain of the dynein intermediate chain as part of its function within the dynein motor complex (Makokha *et al*, 2002), LC8 family proteins also interact with a wide range of proteins. Of note, this includes nitric oxide synthase (Jaffrey & Snyder, 1996) and myosin V (Espindola *et al*, 2000), in addition to transcription factors including Swallow (Schnorrer *et al*, 2000) and the pro-apoptotic factor Bim (Puthalakath *et al*, 1999). For many years, many LC8 family ligands were assumed to be dynein cargo wherein DYNLL1 was thought to link cargos to the dynein motor complex. However, a significant proportion of cytosolic DYNLL1 and DYNLL2 is not dynein-bound (King *et al*, 1996), suggesting that these proteins have alternate functions in addition to their roles in the dynein motor.

Many proteins form complexes with only a few proteins. Others, often known as hubs, have a broad range of ligands and are integral in a variety of protein-protein networks, typically allowing for regulation of various cellular processes (Han *et al*, 2004). Generally, one or both of the hub protein and its binding partners consist intrinsically disordered regions, enabling interaction with a range of partners (Haynes *et al*, 2006) which all compete for the same sites (Patil *et al*, 2010). It is now widely agreed that LC8 proteins function as ordered hub proteins which bind to unstructured regions of their diverse partners, promoting increased order or dimerisation in the ligand.

In both the apo and bound forms, LC8 proteins are symmetrical dimers with two binding sites present within a groove at the dimer interface, and it has been suggested that binding at one of these identical sites can promote binding at the other (Benison *et al*, 2008). This dimer interface consists of a highly flexible (Fan *et al*, 2002), dynamic pocket wherein short recognition motifs of 7-10 amino acids within intrinsically disordered regions of their ligands can bind, forming what is essentially a sixth β -strand within each monomer that integrates into the antiparallel β -sheet (Barbar, 2008). Thus, this results in an LC8 dimer wherein a 12-stranded β -sandwich, comprised of 5 β -strands from each monomer and 1 β -strand from each ligand, is flanked by 2 helices on either side (Clark *et al*, 2016). Occasionally, ligands have several LC8-binding motifs in tandem. The recognition motif varies amongst ligands but most frequently is anchored by a Thr-Gln-Thr (TQT) motif towards the C-terminus wherein the Gln is denoted as position zero, or

the anchor residue, as a result of being the most highly conserved residue (Benison *et al*, 2007; Jespersen *et al*, 2019).

5.4 Selecting which LC8 family protein to use for investigating an interaction with KANK1

While the majority of published LC8 data uses the *D. melanogaster* form, with decades of comprehensive structural and biophysical research, we sought to use a human LC8 protein in this work. As mentioned, there are two LC8 proteins in humans – DYNLL1 (LC8) and DYNLL2 (LC8b) – and these proteins are almost identical in their structure and, seemingly, function. Upon performing multiple sequence alignments for each of these with the *D. melanogaster* homologue (dLC8), five differences were observed between dLC8 and DYNLL1 but only three differences between dLC8 and DYNLL2 (Figure 5.4), and thus DYNLL2 was identified as having the highest homology. For each human protein, the motif-binding interface is preserved between species, though a consistent change from Arg to Lys can be observed at position 71, likely a small, species-specific change. Ultimately, a full-length DYNLL2 construct was designed and used for all subsequent work.

```

dLC8      MSDRKAVIKNADMSEEMQQDAVDCATQALEKYNIEKDIAAYIKKE
DYNLL1    MCDRKAVIKNADMSEEMQQDSVECATQALEKYNIEKDIAAHIKKE
* .*****:*****:*****:*****:*****:*****:*****
          :
dLC8      FDKKYNPTWHCIVGRNFGSYVTHETRHFIYFYLQVAILLFKSG
DYNLL1    FDKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLQVAILLFKSG
*****:*****:*****:*****:*****:*****:*****
          :
-----
dLC8      MSDRKAVIKNADMSEEMQQDAVDCATQALEKYNIEKDIAAYIKKE
DYNLL2    MSDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKE
*****:*****:*****:*****:*****:*****:*****
          :
dLC8      FDKKYNPTWHCIVGRNFGSYVTHETRHFIYFYLQVAILLFKSG
DYNLL2    FDKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLQVAILLFKSG
*****:*****:*****:*****:*****:*****:*****
          :

```

Figure 5.4: Multiple sequence alignment of human and *D. melanogaster* LC8 proteins.

Each human LC8 family protein was compared against *D. melanogaster* LC8 using multiple sequence alignment to identify the protein with highest homology. dLC8 refers to *D. melanogaster* LC8.

5.5 pH affects the oligomeric state of DYNLL2

As an initial validation step to show that DYNLL2 functions identically to dLC8 and DYNLL1, SEC-MALS was performed at low and physiological pH to calculate the absolute molecular weight and ascertain the oligomeric state (Figure 5.5). At pH 3, DYNLL2

appears to run as a monomer and eluted at ~13.5 mL, with a calculated absolute molecular weight of 10140 Da – consistent with a monomeric oligomeric state. As expected, DYNLL2 at pH 7.5 runs as a dimer and eluted at ~13 mL – eluting sooner and thus as a larger product – with a calculated molecular weight of 20810 Da. For each experiment, one single peak was present, indicating that the protein exists exclusively in one oligomeric state, either monomer or dimer, depending on the pH. This initial validation step allowed confirmation that the protein is behaving as expected: forming secondary and quaternary structure, exclusively in a dimeric form at physiological pH. This indicates that the crucial dimer interface and thus the two pockets that allow the hub functionality of DYNLL2 are intact.

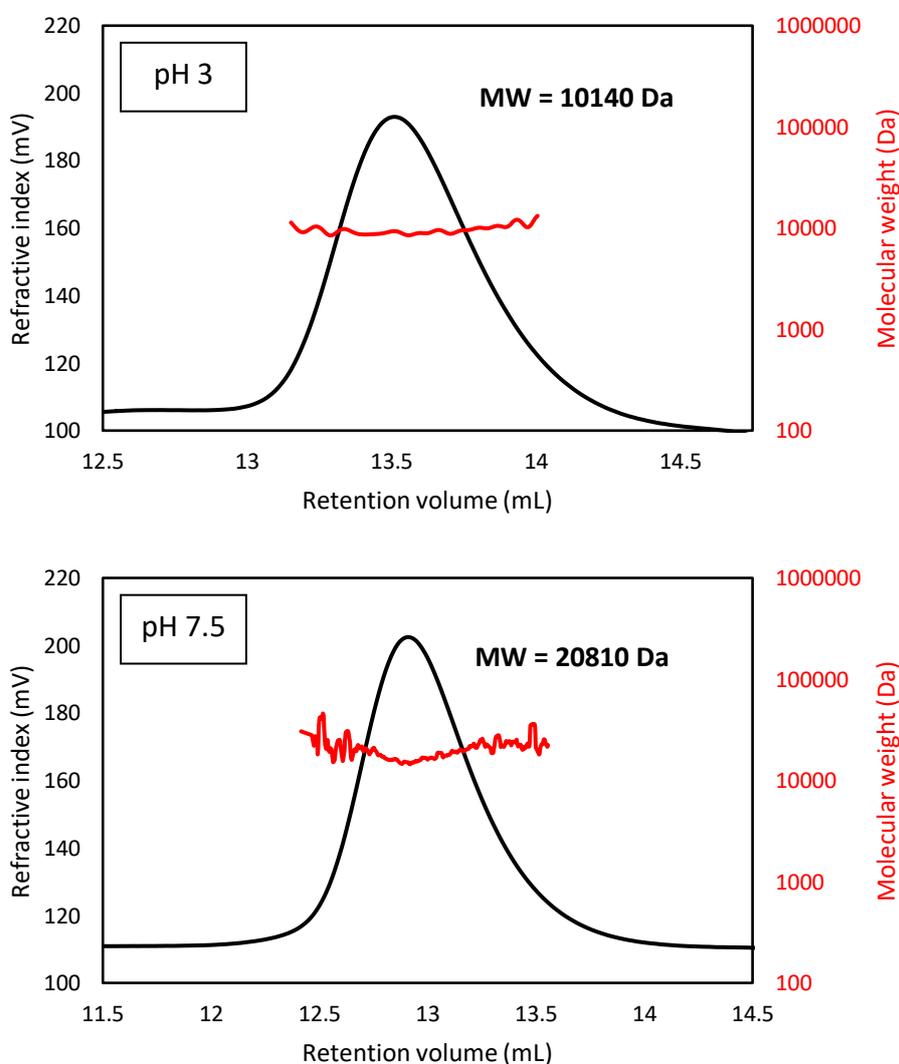


Figure 5.5: SEC-MALS analysis of DYNLL2 at pH 3 and pH 7.5.

SEC-MALS analysis of 75 μ M DYNLL2 at varying pH indicates a monomeric form at pH 3 and a dimeric form at pH 7.5, with calculated absolute molecular weight of 10140 Da and 20810 Da, respectively. A phosphate/citrate buffering system was used to allow the broad pH range without alteration of any other buffer components.

5.6 Using triple resonance NMR to assign DYNLL2 residues at pH 7.5

Triple resonance NMR has been widely used in the research of LC8 family proteins and has played a crucial role in characterising the interaction with many binding partners. The majority of data and structures deposited in both the PDB and BMRB were determined using *D. melanogaster* LC8 with some using human DYNLL1 and that of other species. As the structure of human DYNLL2 has yet to be solved and no NMR assignments are available and deposited for this, initial characterisation and triple resonance assignment of DYNLL2 was performed. Typically, the use of triple resonance NMR to solve the structure of a protein, as opposed to X-ray crystallography, has the advantage of being in solution for the duration of the experiment as opposed to a solid protein crystal and so can be considered more physiologically relevant in terms of being investigated while in a more native state and, further, NMR allows the study of protein dynamics (Kleckner & Foster, 2011). To extend this physiological relevance, all NMR experiments were performed at pH 7.5 to ensure that the DYNLL2 is in its dimeric state.

5.6.1 Overview of backbone assignment theory using triple resonance NMR

To assign the backbone, a series of triple resonance experiments were performed using ^{13}C , ^{15}N -DYNLL2: HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, HNCO, and HN(CA)CO where, for each residue, NH refers to the amide, $\text{C}\alpha$ refers to the α -carbon, $\text{C}\beta$ refers to the β -carbon, i.e. the first carbon of the amino acid R group, and CO refers to the carbonyl. Each of these experiment names corresponds to which atoms are observed in that experiment, with the components in brackets referring to atoms through which magnetisation flows but are not observed. HNCA, CBCAHN and HN(CA)CO provide information on both the current residue (“i”) and the previous residue (“i-1”), typically with higher intensity peaks for i. HN(CO)CA, CBCA(CO)NH and HNCO provide information on only i-1. Using the spectra from these six experiments provided the foundation for assigning the DYNLL2 residue responsible for each NMR peak by enabling the linkage of neighbouring spin systems. An example of how this is performed with CBCA(CO)NH and CBCANH spectra is shown in Figure 5.6.

The $\text{C}\alpha$ and $\text{C}\beta$ of each residue can be assigned using the CBCAHN and CBCA(CO)HN spectra in addition to the HNCA and HN(CO)CA spectra, whereas HNCO (i-1) and HN(CA)CO (i and i-1) can be used to assign the CO of each residue.

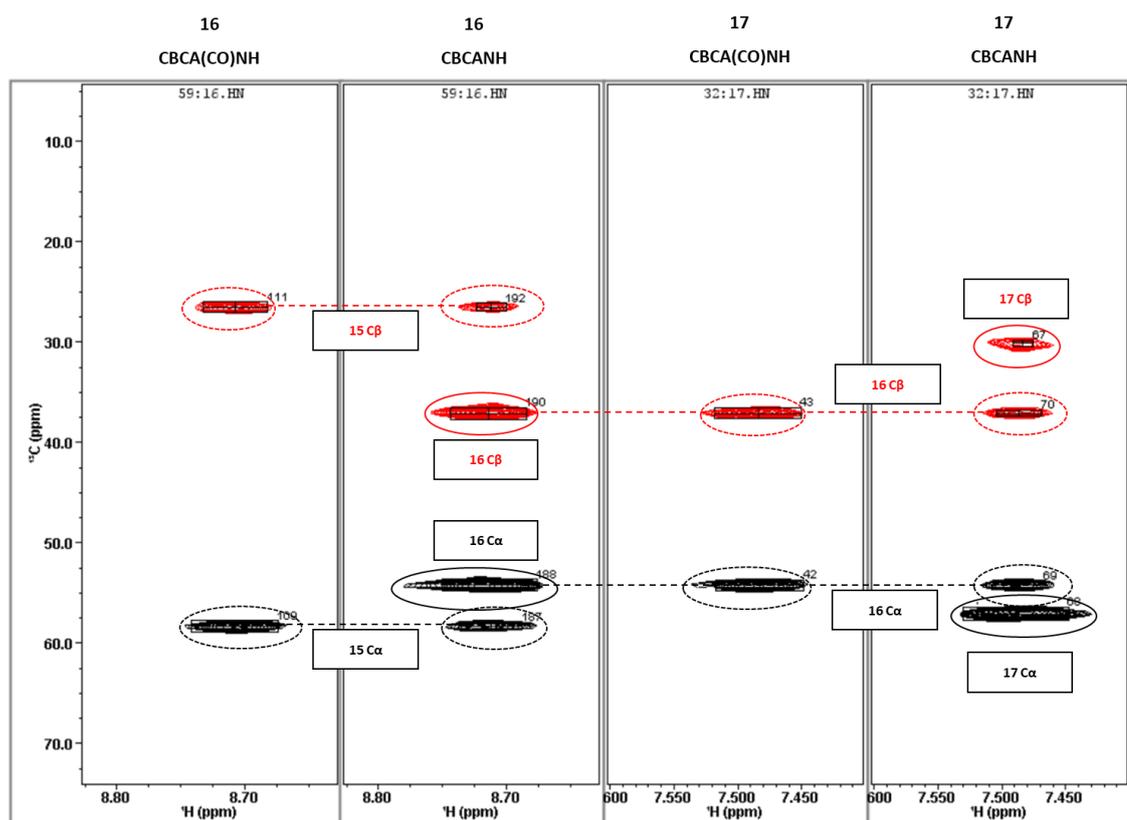


Figure 5.6: An example strip window in NMRViewJ.

Here, a strip window has been set up to show CBCA(CO)NH and CBCANH spectra for two consecutive amino acids in the DYNLL2 chain: residues Glu15 and Asp16. α -carbon ($C\alpha$) peaks are shown in black and β -carbon ($C\beta$) peaks are shown in red. For simplicity, the matches in just the CBCA(CO)NH and CBCANH spectra are highlighted, wherein the peaks for each residue have identical ^{13}C chemical shift values across the different spectra. Thus, where the CBCANH contains peaks from residues i and $i-1$, the $i-1$ peaks match those in the CBCA(CO)NH spectrum and the i peaks can therefore be deduced. These i peaks will be observed in the CBCA(CO)NH spectrum of the next residue in the protein chain where it is now residue $i-1$. This same principle can be used to assign the CA peaks in the HN(CO)CA and HNCA spectra.

5.6.2 NMR assignment of DYNLL2 backbone

Of the 89 amino acids in the DYNLL2 protein chain, the amide of all but 12 of these were assigned to a corresponding peak. Of those that were unassigned, one string of six residues is particularly notable: residues 66-71 which are part of the highly flexible region of free LC8. As shown in the crystal structure for apo LC8 (of *D. melanogaster*), these residues form the C-terminal end of the β 3 strand (see Figure 5.7) upon ligand binding, and this strand only becomes fully ordered in the bound state (Benison *et al*, 2008). These differences between the free and ligand-bound forms of LC8 proteins shown in the crystal structure are supported by previous NMR chemical shift perturbation studies of these protein complexes in solution: the β 3 strand, in addition to the α 2 helix, are subject to the most significant perturbations as a result of forming the most integral region of the binding pocket (Benison *et al*, 2007). The other six

residues which were unassigned are Met1 – which does not have a N-terminal amide bond – and Ser2 which comprise the unstructured N-terminal end of the protein in addition to residues Ala11, Glu15, Glu35 and Gln80. With further data collection and further optimisation of parameters, for example the sample temperature, it may be possible to acquire sufficient data to assign these residues.

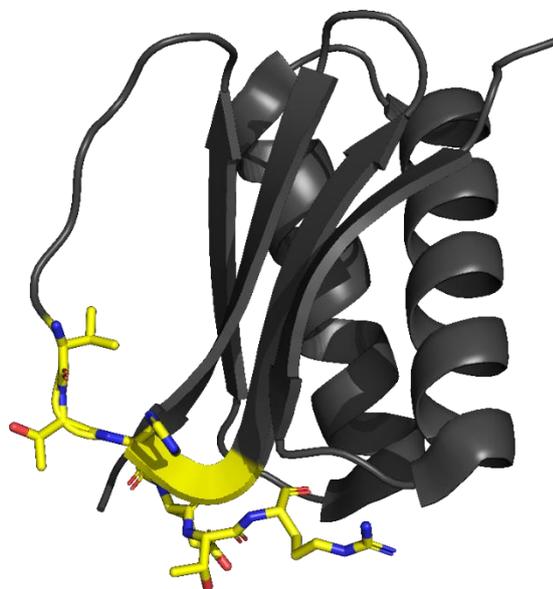


Figure 5.7: β 3 is unstructured in apo LC8.

The crystal structure of *D. melanogaster* LC8 shows that the region typically referred to as the β 3 strand, which forms an integral part of the dimer interface, is unstructured in apo LC8. A string of 6 unassigned residues, shown here as yellow sticks, are part of a flexible region at the C-terminal region of what would be the β 3 string. Image produced with PyMOL using X-ray crystallography structure adapted from PDB ID 3BRI.

Figure 5.8 shows the HSQC spectrum of DYNLL2 with the assigned peaks labelled. Although the protein is a dimer, its symmetry (Figure 5.3B) results in each amino acid amide bond of the monomer protein chain (with the exception of that of the one proline residue) being represented by one peak, i.e. ~ 88 backbone peaks as opposed to ~ 176 as might be expected. As previously shown in literature for other LC8 family proteins and exemplified in the DYNLL2 spectra here, these proteins are well-suited for NMR – the peaks observed are well-resolved and dispersed, with minimal overlap between peaks. This makes the protein ideal for backbone assignment. In addition to backbone peaks, some sidechain peaks are also clearly visible due to containing protons attached to nitrogen atoms. These include a peak generated by the indole ring of the single tryptophan, observed at the bottom left of the HSQC spectrum, and the cluster of peaks generated by the glutamine and asparagine side chain amide groups observed in the top right of the HSQC spectrum.

Phe86, Lys87 and Ser 88; the missing C β peaks were Tyr65, Phe73, Ile 74, Leu85, Phe86, Lys87 and Ser 88; and the missing CO peaks were Thr53, Tyr65, Ser 88 and Gly89.

5.6.3 Chemical shift indexing validates the structural information from the backbone assignment of DYNLL2

C α , C β and CO assignment for each residue allows the use of various programmes for the identification of secondary structure, one of which is TALOS-N. TALOS-N is an artificial neural network that can predict backbone torsion angles (ϕ/ψ) and sidechain torsion angles (χ_1) and thus, as shown here, protein secondary structure. In order to carry out this prediction, HN, N, C α , C β and CO frequencies and the protein sequence are used as the input. The secondary structure prediction uses information from existing high-resolution structural databases: looking at a residue of interest and one residue either side, a consensus from the top 10 matches for the secondary chemical shifts are required to predict the backbone angles of the residue of interest (Shen & Bax, 2013). TALOS-N is thus also able to predict secondary structure information quite reliably for any small areas where assignment information is lacking – this is useful here due to the few areas where the backbone assignment for DYNLL2 is incomplete.

The secondary structure prediction of DYNLL2 from the NMR assignment data using TALOS-N is consistent with the existing structure of DYNLL1 and other homologues of LC8 family proteins in other species (Figure 5.9A, 5.9B): two distinct regions of high α -helical propensity with one region of high β -strand propensity at the N-terminus and four distinct consecutive regions of high β -strand propensity on the C-terminal side of the helices.

In addition to secondary structure prediction, TALOS-N also predicts the backbone flexibility of each peptide. This is calculated as S^2 , a generalised order parameter. For all of the assigned DYNLL2 peaks, the prediction for order is relatively high (Figure 5.9C). It would be interesting to view this prediction for the string of residues 66-71, which were unassigned due to lack of sufficient data in the triple resonance spectra, as it is likely that flexibility and dynamic disorder of this region is implicated in the transition between the free and bound states of DYNLL2.

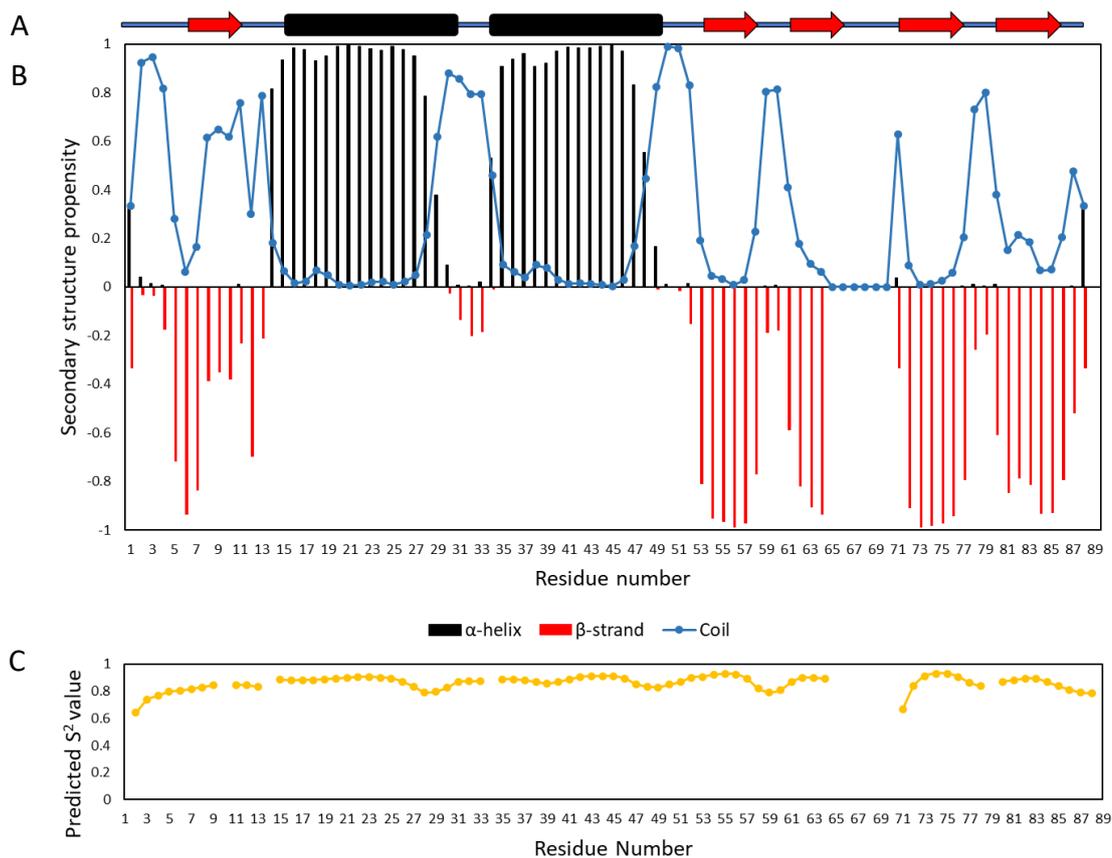


Figure 5.9: TALOS-N output for DYNLL2 is consistent with existing LC8 family structures.

DYNLL2 assignment data was input to TALOS-N for secondary structure and dynamic disorder prediction (Shen & Bax, 2013). (A) A depiction of the secondary structure of DYNLL2 as predicted by PSIPRED (Jones, 1999; Buchan & Jones, 2019) which is consistent with existing solution NMR and X-ray crystallography structures for DYNLL1 and various homologues. (B) Secondary structure prediction for DYNLL2 based on assigned HN, N, C α , C β and CO information: the magnitude of the black bars above and the red bars below 0 of the y-axis indicate the propensity for α -helical and β -strand secondary structure for each residue of the protein chain. The propensity for random coil is shown in blue. (C) The predicted S² value for each residue. For this prediction, residues with unassigned HN and N were excluded.

5.7 DYNLL2 binds KANK1 699-716

5.7.1 Designing a LC8b-binding KANK1 peptide

For this chapter of work, the aim was to identify any DYNLL2-binding sites within IDR2 of KANK1. As previously mentioned, LC8Pred – a programme which can predict LC8-binding motifs with reported 78% accuracy (Jespersen *et al*, 2019) – was used. The 31 hits located within IDR2, of the 69 total, are shown in Table 5.2 and presented in order of their anchor position. Upon interrogation of these hits, it is clear that many of these do not display the very typical, canonical series of residues expected for a LC8 recognition motif. Although not all TQT anchor-containing peptides in the LC8Pred-generating study actually bound to LC8 proteins (Jespersen *et al*, 2019) due to flanking residues also being vital for facilitating LC8 protein interactions, a TQT-containing hit

within KANK1 was searched for in an attempt to select a region reminiscent of the most widely characterised LC8 recognition motif as the most promising candidate. As shown in Figure 5.10, this is found at anchor position 710 in the KANK1 protein chain – and this was the hit ranked #2 from the total DYNLL2 prediction.

Table 5.2: Potential LC8 recognition motifs present in the KANK1 IDR2.

LC8Pred was used to predict any potential LC8 protein-binding motifs in KANK1. Here, 31 hits located in IDR2 are shown in order of their anchor position. A particularly promising hit containing the well-characterised TQT motif is highlighted in the red box.

Hit #	Motif	Amino Acid Score	Volume and Polarity Score	Anchor position
3	VDKAT TMA Q	6.359157018	1.088843942	506
17	VEAV VQ TR	0.449317595	-3.774815276	521
13	TSVE TNS V	-0.9507047	-0.32117154	545
4	CETG SN TE	7.072428439	-1.872204234	605
54	VNDL TLL K	-20.70803136	-2.705071305	615
49	TLLK TNL N	-18.11708479	-3.637689678	619
66	LKTN LNL K	-26.04542655	-8.392642525	621
1	ASRG VNT E	10.89995912	1.840588335	651
16	TEAV SQ VE	-1.369373521	-1.563496231	657
24	VEAA VMA V	-6.709134363	-1.242826817	663
7	VHQF TNT E	5.557840585	-2.582313223	685
11	TETA TL IE	0.938538933	-1.251856193	691
6	IESC TNT C	5.502571158	-1.662899811	697
32	CTNT CL ST	-8.043422504	-2.293576639	700
2	KQTST TQT V	4.572145154	3.950033948	710
15	RVKD INS S	-1.78803646	-0.162251995	728
30	IGVG TLL S	-7.713587186	-2.269500993	741
60	GVGT L LSG	-25.47895476	-4.257773437	742
45	GVGQ I NIN	-14.26839264	-4.371649797	765
41	RIQK LLA E	-15.39453878	0.254267858	834
26	AEQQ TLL A	-6.395725223	-1.742793014	840
57	EQQT LLA E	-20.52849598	-3.699101619	841
43	QMGS LNS Q	-12.74255433	-5.041759997	863
29	GSLN SQ LI	-7.370255428	-2.211281491	865
22	QLIS TL SS	-3.639151423	-1.558607008	870
52	TLSS INS V	-17.14706248	-5.872404528	874
14	KCGG LQS G	-3.91543141	1.984000515	913
8	SPLS SQT S	2.19372608	0.409586511	921
40	TQEG TL SP	-13.9761318	-0.81535065	950
61	TLSP VNL T	-23.06485997	-7.033722359	954
23	GHHA VNI E	-4.982611918	-1.572310274	1054

A peptide spanning residues 699-714 of KANK1 was designed and ordered to investigate binding with DYNLL2. These boundaries allow for a few residues either side of the LC8 recognition motif. Interestingly, this particular region of KANK1 does not appear to have

a homologous region in any other KANK isoforms – this could be a potential indicator of some specific differences between the KANKs, and perhaps some or all LC8-binding regions are unique to KANK1.

hKANK1	⁶⁹⁹ CLSTLD <u>QNSTQT</u> ETRT ⁷¹⁶
hKANK2	-----
hKANK3	-----
hKANK4	-----

Figure 5.10: The 699-716 region is unique to KANK1.

Residues 699-716 of KANK1, containing a top LC8Pred hit comprised of a TQT-containing recognition motif, is unique to KANK1 – multiple sequence alignments performed with IDR2 of each KANK isoform do not have any aligned regions with KANK1 688-716. The 8-residue LC8 recognition motif is highlighted in yellow and the TQT motif, where the Gln residue serves as the “anchor”, is underlined.

Notably, the ConSurf output indicates that the first Thr residue of the TQT motif, in addition to several upstream residues, is very highly conserved, although the subsequent Gln and Thr are not (Figure 5.2). This would suggest that the first part of this LC8 recognition site plays a functional role in KANK1 activity. However, it should be noted that ConSurf often does not differentiate between isoforms when selecting its candidates for analysis, perhaps resulting in some KANK2, KANK3 or KANK4-related proteins in its input. This may have caused the Gln and second Thr of the recognition motif to score poorly in terms of conservation because the data has not been drawn exclusively from KANK1 and equivalent proteins. Considering all of this information, in addition to the high β -strand propensity for the anchor Gln onwards (Figure 5.1) which usually dictates high affinity binding with LC8 proteins, this KANK1 699-716 peptide was selected to test for DYNLL2 binding.

5.7.2 Using NMR titrations to test an interaction between DYNLL2 and KANK1 699-716

To investigate an interaction between full-length DYNLL2 and the KANK1 699-716 peptide, NMR was used. For this, HSQC NMR experiments were performed on full-length DYNLL2 on its own and with various concentrations of KANK1 699-716 peptides added. This data is shown in Figure 5.11A and indicates that an interaction occurs. As previously shown in literature for other LC8 proteins binding to their ligands, this interaction appears to be in slow exchange – where the rate of exchange between different conformations is smaller than the duration of the NMR experiment being recorded – and thus, in places where the population is neither fully free nor fully bound, a particular amino acid can be observed as peaks in both the free and bound positions.

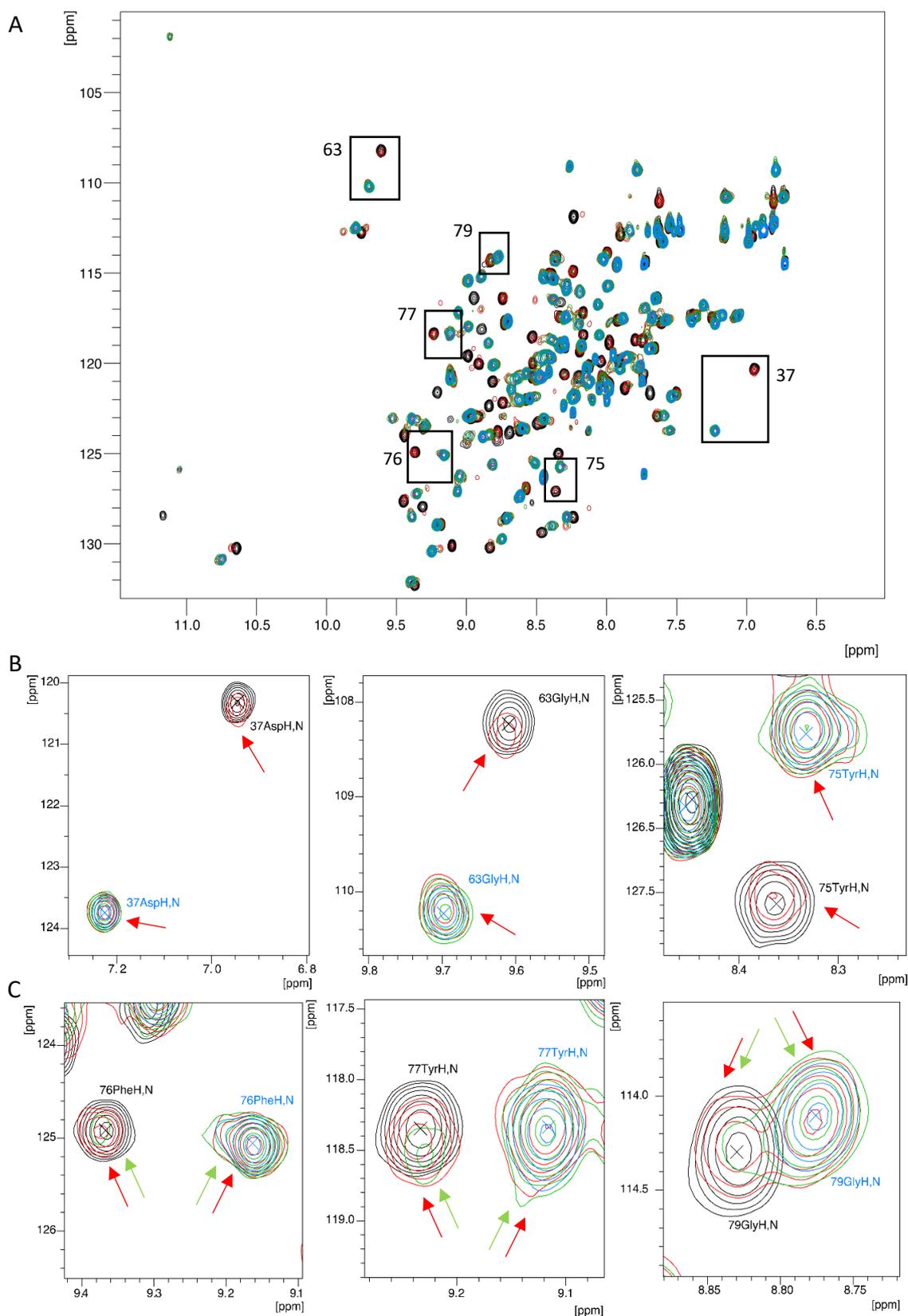


Figure 5.11: HSQC NMR titration spectra of DYNLL2 with KANK1 699-716 peptide.

(A) ^1H , ^{15}N HSQC NMR spectra of 250 μM ^{15}N -labelled DYNLL2 in absence (black) or presence of 125 μM (red), 250 μM (green) or 500 μM (blue) KANK1 699-716 peptide. (B) Three residues which indicate that the reaction is in slow exchange: here, at 1:0.5 ratio of DYNLL2:KANK1 699-716 (red), a peak can be observed in both the free and bound states. By 1:1 ratio, DYNLL2 is exclusively in the bound state. (C) Three residues which still have a proportion of DYNLL2 in the free state at 1:1 ratio of DYNLL2:KANK1 699-716 (green). By 1:2 ratio, DYNLL2 is exclusively in the bound state.

Figures 5.11B and 5.11C show examples of some residues in the HSQC which have clear chemical shift perturbations and are in a slow exchange state, displaying peaks in a free and/or a bound position. Interestingly, some of these appear to have a saturated interaction wherein DYNLL2 is exclusively in the bound state at 1:1 ratio of the protein to the KANK1 699-716 peptide (Figure 5.11B), whereas some others still have some protein in the free state at 1:1 ratio of protein to peptide and the protein is then exclusively in the bound state at 1:2 ratio of DYNLL2 to KANK1 699-716 (Figure 5.11C). Although this may have occurred due to slight inaccuracies in the sample in terms of achieving intended ratios, perhaps due to a small level of protein or peptide precipitation, it is an interesting observation. In order to deduce the significance, the backbone of DYNLL2 in its bound state in the presence of KANK1 699-716 was also assigned and the chemical shift perturbations were mapped, allowing identification of the residues represented by the peaks highlighted in Figures 5.11B and 5.11C.

5.7.3 Mapping chemical shift perturbations of DYNLL2 residues upon KANK1 interaction
 In order to have a greater understanding of what is occurring in the interaction between DYNLL2 and KANK1 699-716, the backbone of the bound state of the protein was also assigned. This was accomplished by collecting HNCA, HNCO and CBCAHN spectra of ^{13}C , ^{15}N -DYNLL2 with KANK1 699-716 at a 1:2 ratio. These data were compared to the data for DYNLL2 alone and the corresponding peaks in the bound protein form identified. All previously assigned residues, with the exception of Thr53, Phe62, Ser64, Leu85, Phe86, Lys87 and Ser88, were assigned in the bound form of the protein (Figure 5.12).

From this data, combined chemical shift difference ($\Delta\delta$) for each residue can be calculated using the following equation:

$$\Delta\delta = \sqrt{\left(\frac{\Delta\delta(\text{N})}{R_{\text{scale}}}\right)^2 + (\Delta\delta(\text{HN}))^2}$$

where $\Delta\delta(\text{HN})$ = amide proton chemical shift perturbation in the ^{15}N dimension

$\Delta\delta(\text{N})$ = nitrogen chemical shift perturbation in the ^{15}N dimension

R_{scale} = a scaling factor used to standardise shifts across the ^{15}N and ^1H dimensions – here, a R_{scale} value of 6 was used as the ^{15}N scale of the data is approximately 6x greater than the ^1H scale.

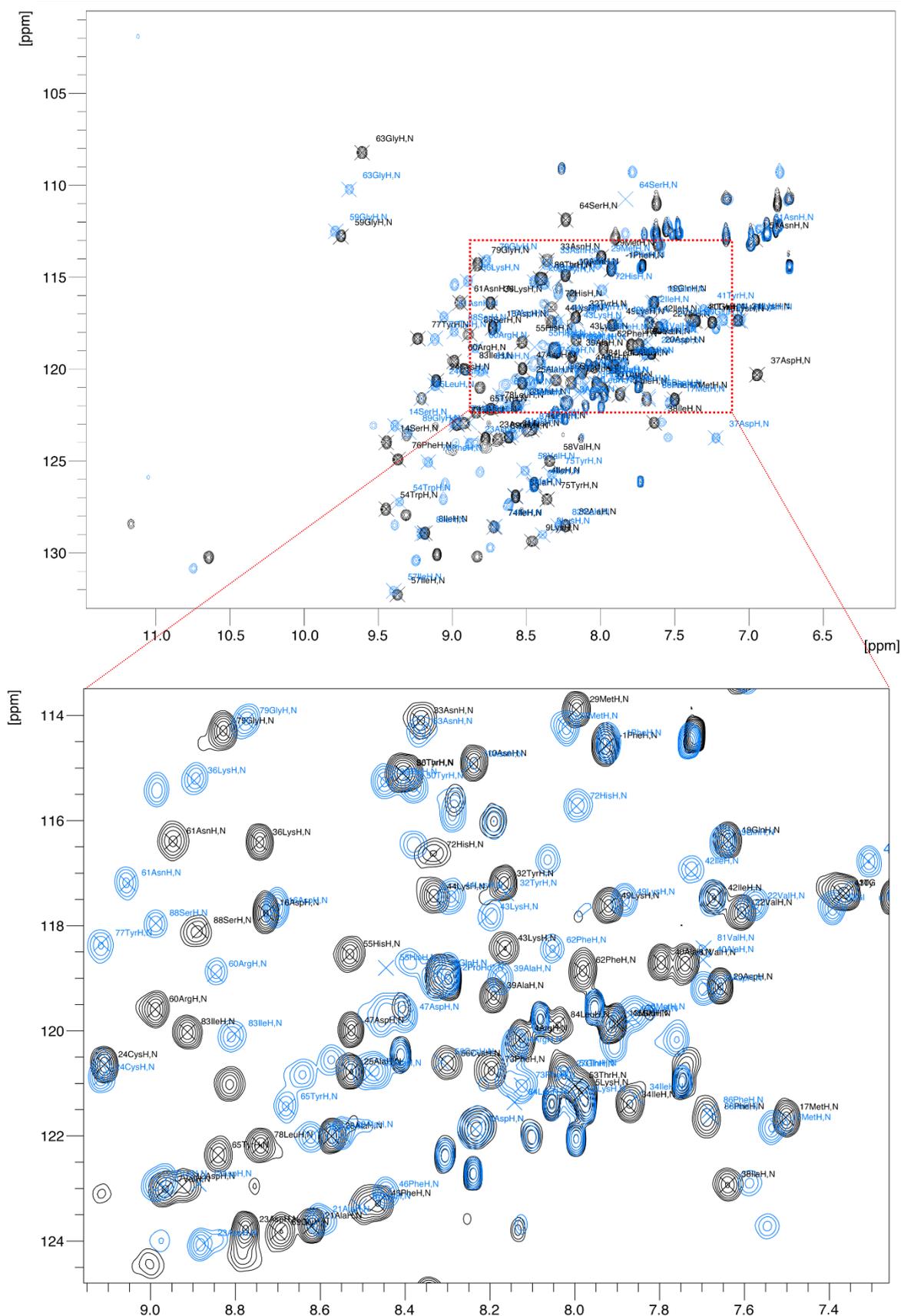


Figure 5.12: The backbone assignment of DYNLL2 in free and bound states.

$^1\text{H}, ^{15}\text{N}$ HSQC spectra of $250\ \mu\text{M}$ ^{15}N -labelled DYNLL2 alone (black) or with $500\ \mu\text{M}$ KANK1 699-716 (blue) with assigned peaks labelled. The central region of the spectrum is shown in more detail in the lower panel.

The calculated chemical shift distance ($\Delta\delta$) of each assigned residue in the DYNLL2 chain is shown in Figure 5.13A. To allow identification of those residues which shifted a significant distance as a result of the interaction with KANK1 699-716, the mean shift distance across data for all residues, calculated as 0.096, was applied as a threshold. Those residues which have a higher calculated shift distance than this mean are coloured in orange, while those which a calculated shift distance greater than two times this mean are coloured in pink. This colour scheme has been used to identify these residues with significant shift distance values on the apo LC8 crystal structure, 3BRI (Figure 5.13B).

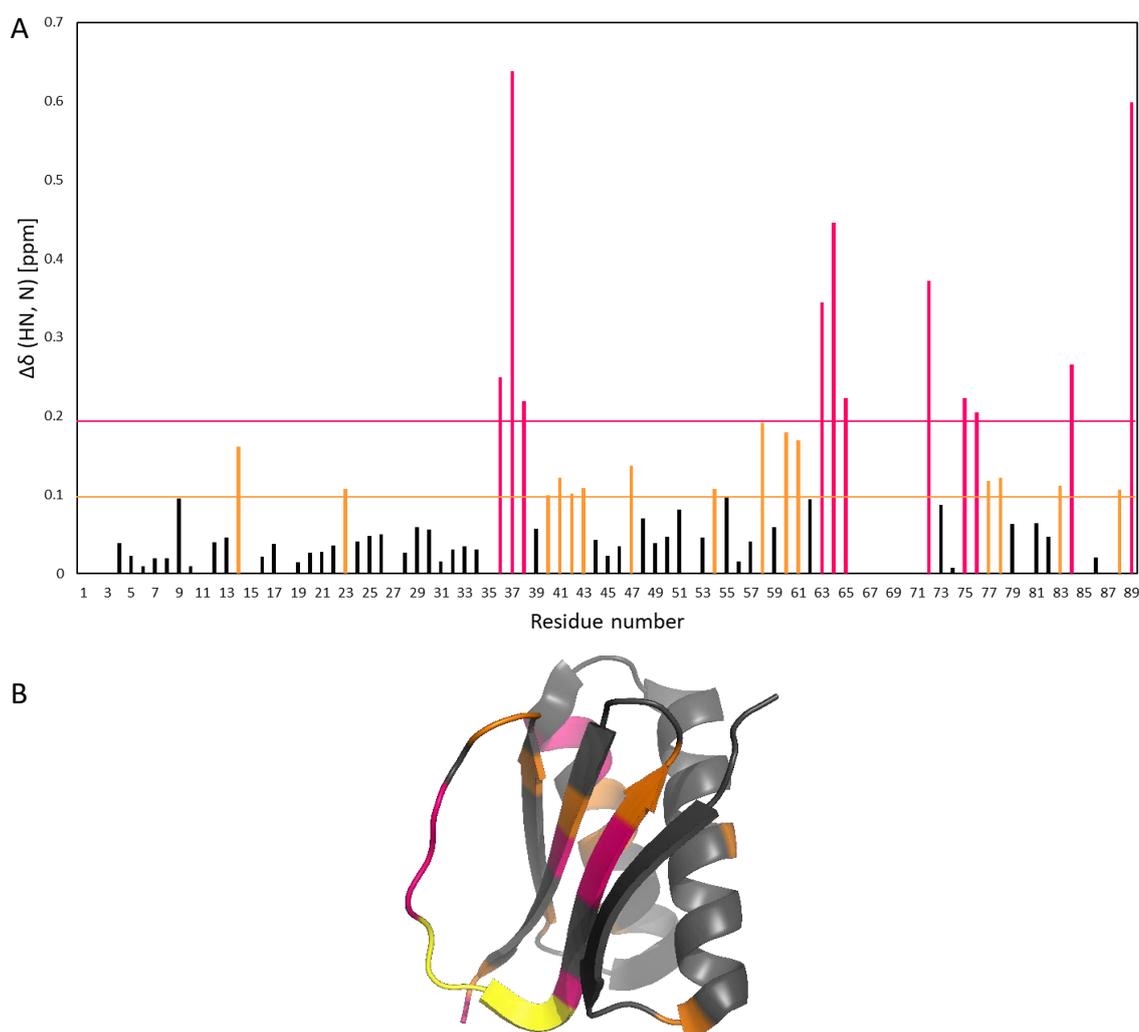


Figure 5.13: Mapping the DYNLL2 residues involved in the interaction with KANK1 699-716.

(A) The calculated chemical shift distance ($\Delta\delta$) for each DYNLL2 residue is shown, calculated from the peak positions for the free protein and the bound protein after addition of KANK1 699-716. Residues with $\Delta\delta$ greater than the mean shift distance are shown in orange, and those greater than double the mean shift distance are shown in magenta. (B) Using the same colour scheme as (A), the residues with a greater $\Delta\delta$ than the mean shift distance have been mapped onto the apo LC8 crystal structure (monomeric for clarity). In yellow, the string of 6 unassigned residues from Figure 5.7 is shown. Image produced with PyMOL using X-ray crystallography structure adapted from PDB ID: 3BRI.

Upon alignment of free and bound LC8 proteins, some structural differences are immediately clear. The most significant change, as already mentioned, is the transition from a highly dynamic, unstructured β 3 in the free form to a considerably more ordered β -strand in the ligand-bound form. In addition to this, structural differences can also be observed in β 1, β 4 and β 5 with altered positioning of the strands (Figure 5.14A). Using the crystal structure of human DYNLL1 bound to a TQT-recognition motif-containing Nek9 peptide (PDB ID: 3ZKE (Gallego *et al*, 2013), the TQT-recognition motif-containing KANK1 699-716 peptide used in this work was modelled. As with other TQT-recognition motif-containing ligands, in addition to other LC8-recognition motifs in general, the KANK1 peptide likely forms a sixth β -strand within the binding pocket, resulting in a 12-stranded β -sandwich across the entire dimer (Figure 5.14B). In order for this additional β -strand to be incorporated into the existing β -sandwich, and also factoring in where β 3 becomes ordered and forms the β -strand, strain would be generated and would disrupt the hydrogen bonds maintaining the antiparallel β -sheet structure. Thus, in order to allow for the transition into the bound state, several of the other β -strands appear to shift – possibly in a somewhat sequential knock-on effect – facilitating the maintenance of the secondary and quaternary structure. Interestingly, the largest chemical shift differences of the assigned DYNLL2 residues outside of those corresponding to β 3 appear to largely occur within the central regions of the other β -strands (Figure 5.14C, orange and magenta). Although these residues seem very distant to be implicated in binding itself, this suggests that the knock-on effect for maintaining the hydrogen bonding pattern of the β -sheets is a significant event that occurs as part of DYNLL2 binding ligands such as KANK1.

Of those DYNLL2 residues which have the largest chemical shift differences between the free and KANK-bound forms, the residues within β 3 and the adjacent structural elements provide information that is consistent with that in literature for other TQT-based ligands (Lo *et al*, 2001; Benison *et al*, 2007). Based on these data, it can be deduced that the TQT motif in KANK1, consisting of Thr709, Gln710 and Thr711, forms a network of hydrogen bonds with DYNLL2: the two threonine residues form interactions within the binding pocket to residues of β 3 of one monomer, including Gly63 and Ser64 in particular, whereas Gln710 additionally interacts with the α 2 of the

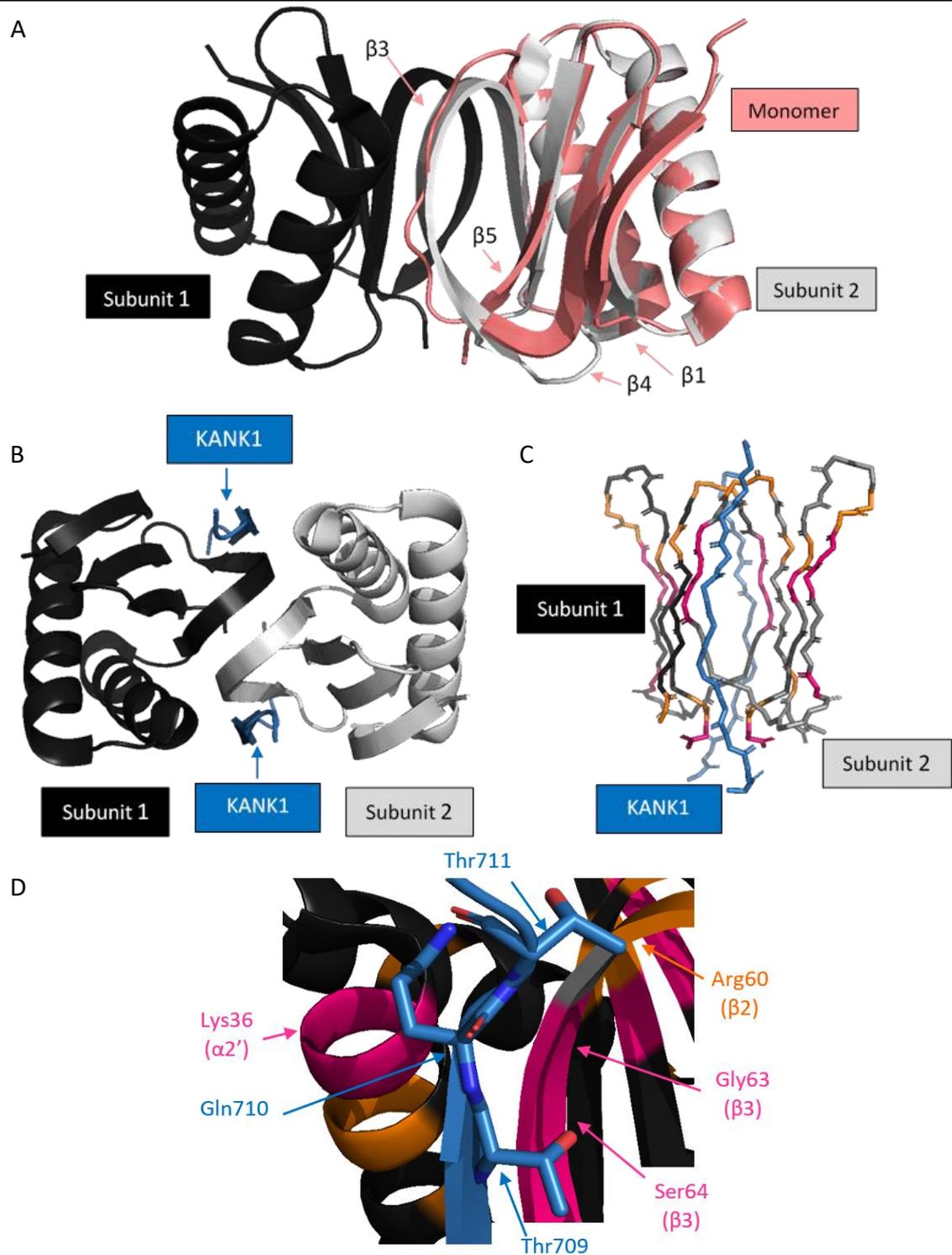


Figure 5.14: The structure of LC8 proteins bound to TQT-containing recognition motifs.

(A) Free LC8 (from PDB ID: 3BRI, pink) aligned with the bound form wherein two monomeric subunits are observed (from PDB ID: 3ZKE, black and grey). Significant structural differences, particularly in $\beta 3$, are highlighted. The TQT-containing peptide from 3ZKE is not shown. (B) The ligand-bound β -sandwich: when modelling KANK1 702-712 in complex with the protein, an additional β -strand is formed on each side which contributes to the overall β -sandwich. (C) Using the colour scheme from Figure 5.13, the significantly shifted regions of the protein are likely due to knock-on effects to individual β -strands as a result of the insertion of the ligand β -strands, allowing tight order of the β -sandwich to be maintained. (D) The modelled TQT recognition motif in KANK1 forms specific hydrogen bonds, particularly with Ser64 and Phe62 of the Subunit 2 $\beta 3$ and Lys36 of the Subunit 1 $\alpha 2'$. Images produced with PyMOL.

other DYNLL2 subunit, particularly Lys36 (Figure 5.14D). Hence, the glutamine of the TQT motif – and the equivalent residue in other LC8-recognition motifs – truly functions as an anchor and allows the entire dimeric, ligand-bound DYNLL2 structure to be reinforced. The remaining residues which comprise the KANK1 β -strand encompassed within the complex also form hydrogen bonds with their respective opposing β 3 residues.

The DYNLL2 residues mentioned to be directly involved with the TQT motif have some of the largest calculated chemical shift distances from the NMR titration experiments with KANK1 699-716 and, in conjunction with the large shift distances observed for other β -strands as a result of accommodation for the ligand, the data presented in this chapter indicates that KANK1 699-716 interacts with DYNLL2 in a canonical TQT-recognition motif-based manner.

5.8 Discussion

In this chapter, it has been shown by NMR titration that an interaction occurs between DYNLL2 and the 699-716 region of KANK1. Using triple resonance NMR to assign the backbone structure of DYNLL2, the residues implicated in and/or affected by this interaction were able to be identified. As mentioned, the backbone assignment was incomplete: a string of 6 residues at the C-terminus of the β 3-strand, known to be a highly dynamic and flexible region, as well as 6 other residues throughout the protein, did not have assigned amides as part of this work. Based on the existing literature for LC8 proteins and their ligands, it can be assumed that these β 3 residues would be involved in the interaction with the residues upstream of the KANK1 TQT-motif. However, because of this missing data, this cannot be conclusively stated. If the triple resonance NMR experiments were to be repeated, performing at a higher temperature may alter the dynamics of the unassigned string, allowing acquisition of sufficient data to then assign these residues. If the free form is assigned, it may then be possible to measure any chemical shift distance that occurs and likely confirm that these residues are involved in KANK binding. Alternatively, solving the structure using X-ray crystallography could also provide this information.

Throughout this chapter, various isoforms and homologues of LC8 proteins have been mentioned. All experimental work was performed with human DYNLL2, a less well characterised LC8 family protein. Working with this specific form of the protein allowed some characterisation of it to be performed, including validation that it exists as a dimer at physiological pH and that low pH causes a switch to the monomeric form, NMR backbone assignment, and preliminary work indicating interaction with TQT-recognition motif-containing ligands very similarly to other LC8 family proteins. Where the binding surface is very highly conserved amongst LC8 family proteins, it is assumed that this interaction with KANK will occur throughout – however, this would be important to validate, particularly with the more widely studied LC8 family proteins including human DYNLL1 and the *D. melanogaster* homologue. In addition to this comparison for other LC8 proteins, it would also be very fascinating to investigate whether other KANK family proteins bind LC8 proteins. Although already deduced that there is not a corresponding region to 699-716 of KANK1, the region used in this work, using LC8Pred to predict any LC8-binding sites in the other isoforms and investigating any very promising hits could be very informative. Perhaps KANK1 is the only KANK able to bind LC8 proteins, optimised for doing so by having the largest IDR2 amongst the mammalian KANKs, or perhaps all KANKs exhibit this property but in different capacities.

Unpublished work by our collaborators, the Akhmanova group at Utrecht University, suggests that, in mammalian cells, the coiled coil region of KANK is required in addition to an intact IDR2 for the recruitment of LC8 proteins to focal adhesion sites. This could imply that an interaction between LC8 proteins and the coiled coil region of KANK occurs – of the 69 LC8Pred hits for KANK1, 10 of these are located within the coiled coil region – or it could be that the function/role of the interaction between IDR2 and LC8 proteins is linked to that of the coiled coil region, and simultaneous or sequential activity of these regions is a requirement for normal recruitment and subsequent cellular activity. Further work will be required to increase our understanding of this very interesting observation.

One very curious aspect of this interaction *in vivo* would be the stoichiometry involved. Where a short peptide was used in these investigations and thus it is likely that two peptides bound two DYNLL2 molecules due to the interaction being almost completely

saturated at a 1:1 molar ratio (Figure 5.11), this of course, despite allowing a greater understanding of the specific region involved in the binding surface, is not entirely physiologically relevant. As mentioned, there are 69 predicted LC8 protein-binding sites in KANK1 and a significant proportion of these are located within IDR2. Potentially, therefore, it could be that one DYNLL2 dimer recruits a single KANK1 but the two binding pockets are occupied by different binding sites within the single KANK. Alternatively, one DYNLL2 dimer may instead be binding two KANK1 proteins. With such a high number of predicted LC8 protein-binding sites, it is important to consider what the purpose would be of having several binding sites for the same/similar protein. Further analysis of the LC8Pred results and systematic interrogation of the more promising hits will be useful to shed light on this. It must be noted that a significant proportion of these hits are unlikely to bind, with the region investigated in this chapter being the only hit comprising the most canonical LC8 recognition motif, but these results are a useful reference from which to plan future work.

Continuing this investigation into the interaction between KANK and LC8 proteins would be vital to ultimately allow of a greater understanding of what the purpose of driving unstructured regions of KANK1 into an ordered secondary structure would be, and what implications this has for the function of the KANK1 protein overall. On a broader level, it would be particularly of interest to deduce the context of this interaction – what is the purpose of the interaction between an LC8 family protein and a KANK family protein? Where LC8 family proteins have such diverse contexts within which they are involved, understanding whether they are functioning as part of the dynein motor complex with KANK1 as cargo or as hub proteins for a different cellular context of KANK1 would be important. One possibility is that LC8 family proteins may be key in driving the assembly of the large KANK complexes comprising the KANK-rich belts at the periphery of adhesions (Bouchet *et al*, 2016) or perhaps involved in maintaining this localisation of KANK proteins. Further, given the prominent association of both dynein and KANK proteins with microtubules, it would be illogical to overlook this link – one possibility for investigation could be the observation of microtubule dynamics when the DYNLL2-KANK1 interaction is abolished, though prior investigation into whether LC8 proteins bind to other KANK1 sites or to the other KANK isoforms would be useful.

CHAPTER 6: CONCLUSIONS

6.1 Summary

The overall aim of this thesis was to further develop our understanding of KANK proteins – the linker protein which, via direct interaction with both the core adhesion protein talin and the kinesin KIF21A, allows crosstalk between the actin and microtubule cytoskeletons, ultimately contributing to microtubule polarity and facilitating adhesion turnover during cell migration. Where KANK proteins consist of such defined characteristic regions which vary in their similarity between isoforms, each chapter of this thesis sought to magnify specific regions to further characterise their structural and functional properties.

Firstly, the C-terminal ankyrin repeat domain was subject to investigation. While the notable binding partner KIF21A had been identified and crystal structures of KANK1 and KANK2 ankyrin repeat domains both alone and in complex with KIF21A were deposited by other groups early into this project, further understanding of this domain was lacking. Comparing all four mammalian KANK ankyrin repeat domains using circular dichroism, it was identified that the KANK4 ankyrin repeat domain has a unique double transition in the melt curve, indicating two distinct stages of structural changes. Using SEC-MALS, it was confirmed that the KANK4 ankyrin repeat domain exists exclusively as a dimer throughout this pipeline. When attempting to characterise the disease-causing single point mutants in KANK ankyrin repeat domains, one particular mutation – Tyr801His within ANK0 of the KANK4 ankyrin repeat domain – caused a shift to an exclusively monomeric form. Additionally, although it is stated in the existing literature that neither KANK3 nor KANK4 ankyrin repeat domains interact with KIF21A, it is shown in this work that an interaction does in fact occur, though weaker than those with KANK1 and KANK2 ankyrin repeat domains. As part of this project, the crystal structures of the KANK3 ankyrin repeat domain and the Ala670Val mutant-containing KANK2 ankyrin repeat domain were solved and deposited into the PDB (PDB IDs: 6LTH and 6TMD, respectively). These two novel structures provide detail on subtle changes in these two regions compared with the existing KANK ankyrin repeat domain structures available, notably a more compact arrangement within KANK3 which alters the KIF21A – and potentially the KN domain – binding loops connecting each ankyrin repeat, possibly the cause for the

weakened affinity. Determining the structure of the KANK4 ankyrin repeat domain would give valuable insight into the structural variations within that isoform and more definitively determine the dimer interface, in addition to perhaps indicating why the KANK4 interaction with KIF21A is weaker.

Following this, it was identified by a combination of NMR binding experiments and FP assays that the KN and ankyrin repeat domains of all four mammalian KANKs interact in what is likely an autoinhibitory interaction. By designing peptides of varying regions of the KN domain and some peptides containing mutations, it was deduced that residues 61-68 of KANK1 – and the corresponding residues in the other three isoforms – are required for this interaction. Interestingly, the KANK3 KN domain contains a three-residue insertion within this region which appears to affect the affinity of the autoinhibitory interaction: where interactions between the KN and ankyrin repeat domains of KANK1, KANK2 and KANK4 have similar affinities, that of KANK3 is ~4-fold weaker. Determining the structure of this autoinhibitory interaction in KANKs will be key in acquiring more information and may provide further insight into subtle structural changes which could be resulting in a weaker affinity for KANK3. In terms of the broader purpose of an interdomain KANK complex, it is easy to speculate that having a “closed” conformation of the protein correlates to inactivity and perhaps pre-complex formation, similarly to other adhesion-associated proteins including talin and vinculin (Khan & Goult, 2019). In terms of inactivity, it is also likely that competition could occur with the autoinhibitory complex: perhaps, in response to signals for the control of adhesion disassembly, for example, R7 of talin competes with the ankyrin repeat domain to bind to the KN domain or, alternatively, KIF21A competes with the KN domain to bind to the ankyrin repeat domain. The structure of the autoinhibitory interaction between the KN and ankyrin repeat domains of KANK will be key in determining whether the KN domain binding to talin R7 or the ankyrin repeat domain, or the ankyrin repeat domain binding to the KN domain or the KIF21A KANK-binding helix, are mutually exclusive. The required residues in the KN domain for ankyrin repeat domain binding, residues 61-68, clearly do not overlap with the LD-motif required for talin binding, although there has been some indication that other residues upstream of residues 61-68 may be involved in stabilising the interdomain KANK complex (Figure 4.4). Taking into account their significantly

higher binding affinities, both talin R7 and the KANK-binding helix of KIF21A could easily outcompete the autoinhibitory conformation, though the overlap in binding surfaces must also be considered and identified in order to investigate this fully. Preliminary efforts to investigate competition using FP and NMR-based methods were attempted as part of this work but further optimisation will be required to provide conclusive data. Where difficulties have occurred in attempting to observe autoinhibited KANK within mammalian cells, it is very possible that this particular conformation occurs only in certain regions or at certain points in time in response to specific signals and, although the work in this thesis provides an excellent basis from which to work, investigating this further will be key in developing our understanding of the purpose and context of KANK autoinhibition.

Finally, we sought to develop a greater understanding of the largest regions of KANK proteins, the intrinsically disordered regions (IDRs). We focused particularly on KANK1 where IDR2 is significantly larger than all other KANKs. Despite being an IDR, this large region contains microdomains with high secondary structure propensity, so we aimed to investigate the possibility of KANK1 IDR2 in its unstructured state binding ligand proteins and transitioning into the predicted secondary structure. Using a combination of literature searches and interrogation of the KANK1 proteomics dataset, LC8 family proteins were identified as potential ligands. One protein within this family, DYNLL2, was shown using NMR titration methods to bind to an LC8 recognition motif within KANK1 IDR2 containing a characteristic Thr-Gln-Thr sequence. Triple resonance NMR was used to assign the DYNLL2 backbone resonances and the structural changes induced by KANK1 binding were identified, shown to be overall consistent with existing LC8 protein interactions that have been identified. This novel KANK1-DYNLL2 interaction is the first IDR2 interaction characterised and provides a foundation for further studies – for example, investigating whether there are further LC8 protein-binding regions within KANK1 and other KANKs, and what the implications of this would be within an *in vivo* environment. One potential hypothesis to explore would be that the KANK interaction with LC8 proteins could generate a platform for recruiting other proteins – where KANK proteins already initiate cortical platform assembly in the context of the CMSC which provides the cortical microtubule link (Bouchet *et al*, 2016), a similar activity may be

occurring here for the assembly of a platform for an entirely different or very similar process.

Overall, this work has allowed an increased understanding of the KANK family of proteins, a key component of regulated adhesion turnover during cell migration, by characterising some key differences between isoforms and identifying novel interactions and binding partners. The isoform-specific intricacies are becoming increasingly elucidated within this area of research: very recent work identifies cell type-specific expression of KANK isoforms which rarely overlaps (Guo *et al*, 2021) – supporting the hypothesis that the subtle differences amongst KANK proteins determine their distinct cellular functions despite the similarities in their major structural elements.

The findings within this thesis have set the foundations for even more studies of KANK proteins, ultimately allowing a greater understanding of the *in vivo* implications of missing or defective KANK.

6.2 Potential limitations

Although the structural, biochemical and biophysical data contained within this body of work is comprehensive and efforts were made throughout to validate the findings with various techniques, a limitation of this work is the lack of *in vivo* approaches. Structural, biochemical and biophysical methods provide excellent insight and fine detail of components within a biological system and are an extremely useful approach for characterising interactions and determining binding affinities. However, with the overall aim of understanding what is occurring within a biological system, it is important to also take the findings of the *in vitro* work and validate the biological relevance within a true *in vivo* system – a combination of these approaches will result in a more confident determination of what is actually occurring in these systems that we study.

The investigation of KANK proteins in mammalian cells and in animal models has posed some difficulties as a result of gene redundancy and compensation, resulting in isoform-specific conclusions being difficult to make. Further, a significant limitation observed as part of the work in collaboration with the Akhmanova group at Utrecht University is the lack of a functional readout for both KANK3 and KANK4 in many mammalian cell systems. The data discussed in Chapter 3, for example identifying KANK3 localising to

the filopodia tips (Figure 3.3) which may provide a new direction for identifying the function of this isoform, will facilitate more targeted future investigation into the autoinhibition of KANKs in a cellular context.

6.3 Future work

The work presented in this thesis provides excellent insight into the significant role of KANK proteins and sets the foundation for many more avenues for future investigation.

As already mentioned throughout, there are several structures which would be invaluable to determine. These include both the wildtype and Tyr801His mutant versions of the KANK4 ankyrin repeat domain – key in understanding the oligomeric switch and to allow an understanding of the difference in KIF21A binding across isoforms – and the structure of the autoinhibitory complex formed between the KN and ankyrin repeat domain of any KANK. In terms of this autoinhibition interaction, various approaches were attempted to crystallise this complex, ranging from varying screens and mixing techniques to designing a chimeric protein containing the two domains tethered together. Optimisation of this chimeric protein construct could be a particularly promising approach, and perhaps re-designing with the KN domain instead at the C-terminal end could be attempted. This design strategy would be particularly useful in the case of steric hindrance/geometry causing restricted access between the binding surfaces in the original construct. Alternatively, triple resonance NMR experiments could be used to attempt assigning the protein structure and mapping the binding surface to determine the solution structure of the interaction – however, the size of the ankyrin repeat domain proteins would make this quite challenging, with the inevitable broadened signal and reduced resolution affecting final results.

Another area where further study would be useful would be to deduce with increased certainty whether the KANK3 and KANK4 ankyrin repeat domains bind KIF21A. Although published work states that these two isoforms do not interact with KIF21A (Pan *et al*, 2018; Guo *et al*, 2018; Weng *et al*, 2018), FP experiments here indicate otherwise, with relatively tight K_d values calculated. Use of a technique with increased sensitivity, including perhaps NMR titrations, could be used to validate the findings. It could also be insightful to characterise the other known disease-causing point mutations in KANKs,

with many of those not discussed in depth here located within other KANK regions, notably the coiled coil region and IDR2.

The interaction with the LC8 family protein DYNLL2 is the first characterised ligand for IDR2 in KANK1. With several other potential sites within IDR2 and elsewhere within KANK1 which could also interact, it would be interesting to investigate some of these and determine what the stoichiometry of this interaction would be. *In vivo* studies will be ideal to assist in determining what the ultimate role of this interaction is – perhaps KANK, in addition to its interaction with the kinesin KIF21A, has direct involvement with the dynein motor complex, or DYNLL2 is binding in a dynein-independent context – and any other significant events/components involved.

Finally, the KANK1 proteomics dataset from the Akhmanova group at Utrecht University is an invaluable resource which could be key in identifying novel ligands and roles of KANK proteins. Interrogating this data further will contribute to developing our understanding of KANK proteins within a biological system.

Overall, the data presented in this thesis drives the characterisation of KANK family proteins as individual proteins with similar overall functions yet with attributes unique to each isoform. These isoform-specific differences will be key in furthering our understanding of the exact roles that KANK proteins possess, defining the contexts in which they can compensate for each other and those wherein a loss or mutation in an isoform could have significant repercussions. This thesis has provided further characterisation in the context of the protein-protein interaction-mediating ankyrin repeat domain, which will facilitate the future identification of new ligands and functions, and two novel interactions of KANK proteins have been identified: an autoinhibitory complex which could be key in KANK regulation and the complex formed with the LC8 family of hub proteins. Continuing to develop our understanding of this protein family that is key in linking two major cytoskeletal elements, the actin filaments and cortical microtubules, will be essential in unveiling further contexts in which KANK proteins are involved and ultimately increasing our insight into these vital cellular processes.

CHAPTER 7: REFERENCES

- Abercrombie M (1961) The bases of the locomotory behaviour of fibroblasts. *Exp. Cell Res.* **8**: 188–198
- Akhmanova A & Noordstra I (2017) Linking cortical microtubule attachment and exocytosis. *F1000Research* **6**: 469
- Akhmanova A & Steinmetz MO (2008) Tracking the ends: A dynamic protein network controls the fate of microtubule tips. *Nat. Rev. Mol. Cell Biol.* **9**: 309–322
- Akhmanova A & Steinmetz MO (2015) Control of microtubule organization and dynamics: Two ends in the limelight. *Nat. Rev. Mol. Cell Biol.* **16**: 711–726
- Alam T, Alazmi M, Gao X & Arold ST (2014) How to find a leucine in a haystack? Structure, ligand recognition and regulation of leucine-aspartic acid (LD) motifs. *Biochem. J.* **460**: 317–329
- Alon R, Hammer DA & Springer TA (1995) Lifetime of the P-selectin-carbohydrate bond and its response to tensile force in hydrodynamic flow. *Nature* **377**: 86
- Anderson LR, Owens TW & Naylor MJ (2014) Structural and mechanical functions of integrins. *Biophys. Rev.* **6**: 203–213
- Anthis NJ & Campbell ID (2011) The tail of integrin activation. *Trends Biochem. Sci.* **36**: 191–198
- Anthis NJ, Wegener KL, Ye F, Kim C, Goult BT, Lowe ED, Vakonakis I, Bate N, Critchley DR, Ginsberg MH & Campbell ID (2009) The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO J.* **28**: 3623–3632
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T & Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**: W344–W350
- Asperti C, Astro V, Totaro A, Paris S & de Curtis I (2009) Liprin- α 1 promotes cell spreading on the extracellular matrix by affecting the distribution of activated integrins. *J. Cell Sci.* **122**: 3225–3232
- Astro V, Chiaretti S, Magistrati E, Fivaz M & De Curtis I (2014) Liprin- α 1, ERC1 and LL5 define

- polarized and dynamic structures that are implicated in cell migration. *J. Cell Sci.* **127**: 3862–3876
- Atherton P, Lausecker F, Carisey A, Gilmore A, Critchley D, Barsukov I & Ballestrem C (2019) Force-independent interactions of talin and vinculin govern integrin-mediated mechanotransduction. *bioRxiv*: 629683 Available at: <https://www.biorxiv.org/content/10.1101/629683v1>
- Bachir AI, Zareno J, Moissoglu K, Plow EF, Gratton E & Horwitz AR (2014) Integrin-associated complexes form hierarchically with variable stoichiometry in nascent adhesions. *Curr. Biol.* **24**: 1845–1853
- Bajar BT, Wang ES, Zhang S, Lin MZ & Chu J (2016) A guide to fluorescent protein FRET pairs. *Sensors (Switzerland)* **16**:
- Barbar E (2008) Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. *Biochemistry* **47**: 503–508 Available at: <https://pubs.acs.org/sharingguidelines> [Accessed March 30, 2019]
- Barbar E, Kleinman B, Imhoff D, Li M, Hays TS & Hare M (2001) Dimerization and folding of LC8, a highly conserved light chain of cytoplasmic dynein. *Biochemistry* **40**: 1596–1605 Available at: <https://pubs.acs.org/sharingguidelines> [Accessed January 22, 2019]
- Barber-Pérez N, Georgiadou M, Guzmán C, Isomursu A, Hamidi H & Ivaska J (2020) Mechano-responsiveness of fibrillar adhesions on stiffness-gradient gels. *J. Cell Sci.*
- Barclay AN (2003) Membrane proteins with immunoglobulin-like domains - A master superfamily of interaction molecules. *Semin. Immunol.* **15**: 215–223
- Basu S, Sladeczek S, De La Peña Y Valenzuela IM, Akaaboune M, Smal I, Martin K, Galjart N & Brenner HR (2015) CLASP2-dependent microtubule capture at the neuromuscular junction membrane requires LL5β and actin for focal delivery of acetylcholine receptor vesicles. *Mol. Biol. Cell* **26**: 938–951
- Benison G, Karplus PA & Barbar E (2007) Structure and Dynamics of LC8 Complexes with KXTQT-Motif Peptides: Swallow and Dynein Intermediate Chain Compete for a Common Site. *J. Mol. Biol.* **371**: 457–468 Available at: <https://www.sciencedirect.com/science/article/pii/S0022283607006821> [Accessed March

30, 2019]

- Benison G, Karplus PA & Barbar E (2008) The Interplay of Ligand Binding and Quaternary Structure in the Diverse Interactions of Dynein Light Chain LC8. *J. Mol. Biol.* **384**: 954–966
- Bertolucci CM, Guibao CD & Zheng J (2005) Structural features of the focal adhesion kinase-paxillin complex give insight into the dynamics of focal adhesion assembly. *Protein Sci.* **14**: 644–652
- Boggetti B, Jasik J, Takamiya M, Strähle U, Reugels AM & Campos-Ortega JA (2012) NBP, a zebrafish homolog of human Kank3, is a novel Numb interactor essential for epidermal integrity and neurulation. *Dev. Biol.* **365**: 164–174 Available at: <http://dx.doi.org/10.1016/j.ydbio.2012.02.021>
- Borgon RA, Vonrhein C, Bricogne G, Bois PRJ & Izard T (2004) Crystal structure of human vinculin. *Structure* **12**: 1189–1197
- Bouchet BP & Akhmanova A (2017) Microtubules in 3D cell motility. *J. Cell Sci.* **130**: 39–50
- Bouchet BP, Gough RE, Ammon YC, van de Willige D, Post H, Jacquemet G, Maarten Altelaar AF, Heck AJR, Goult BT & Akhmanova A (2016) Talin-KANK1 interaction controls the recruitment of cortical microtubule stabilizing complexes to focal adhesions. *Elife* **5**: 1–23
- Breedon L & Nasmyth K (1987) Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*. *Nature* **329**: 651–654
- Buchan DWA & Jones DT (2019) The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **47**: W402–W407
- Burgstaller G & Gimona M (2005) Podosome-mediated matrix resorption and cell motility in vascular smooth muscle cells. *Am. J. Physiol. - Hear. Circ. Physiol.* **288**:
- Burridge K & Connell L (1983) A new protein of adhesion plaques and ruffling membranes. *J. Cell Biol.* **97**: 359–367
- Busson S, Dujardin D, Moreau A, Dompierre J & De Mey JR (1998) Dynein and dynactin are localized to astral microtubules and at cortical sites in mitotic epithelial cells. *Curr. Biol.* **8**: 541–544

- Byron A, Askari JA, Humphries JD, Jacquemet G, Koper EJ, Warwood S, Choi CK, Stroud MJ, Chen CS, Knight D & Humphries MJ (2015) A proteomic approach reveals integrin activation state-dependent control of microtubule cortical targeting. *Nat. Commun.* **6**:
- Calderwood DA, Campbell ID & Critchley DR (2013) Talins and kindlins: Partners in integrin-mediated adhesion. *Nat. Rev. Mol. Cell Biol.* **14**: 503–517
- Calderwood DA, Yan B, De Pereda JM, Alvarez BG, Fujioka Y, Liddington RC & Ginsberg MH (2002) The phosphotyrosine binding-like domain of talin activates integrins. *J. Biol. Chem.* **277**: 21749–21758
- Calderwood DA, Zent R, Grant R, Rees DJG, Hynes RO & Ginsberg MH (1999) The talin head domain binds to integrin β subunit cytoplasmic tails and regulates integrin activation. *J. Biol. Chem.* **274**: 28071–28074
- Campbell ID & Humphries MJ (2011) Integrin structure, activation, and interactions. *Cold Spring Harb. Perspect. Biol.* **3**: 1–14 Available at: www.cshperspectives.org [Accessed January 23, 2019]
- Caswell PT & Norman JC (2006) Integrin trafficking and the control of cell migration. *Traffic* **7**: 14–21
- Changede R, Xu X, Margadant F & Sheetz MP (2015) Nascent Integrin Adhesions Form on All Matrix Rigidities after Integrin Activation. *Dev. Cell* **35**: 614–621
- Chen T, Kun W & Xianzhou T (2017) In vivo and in vitro inhibition of human gastric cancer progress by upregulating Kank1 gene. *Oncol. Rep.* **38**: 1663–1669
- Chen X, Zaro JL & Shen WC (2013) Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* **65**: 1357–1369
- Choi CK, Vicente-Manzanares M, Zareno J, Whitmore LA, Mogilner A & Horwitz AR (2008) Actin and α -actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nat. Cell Biol.* **10**: 1039–1050
- Chudakov DM, Matz M V., Lukyanov S & Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.* **90**: 1103–1163
- Clark S, Nyarko A, Löhner F, Karplus PA & Barbar E (2016) The Anchored Flexibility Model in LC8

- Motif Recognition: Insights from the Chica Complex. *Biochemistry* **55**: 199–209
- Clay MR & Sherwood DR (2015) Basement Membranes in the Worm: A Dynamic Scaffolding that Instructs Cellular Behaviors and Shapes Tissues. *Curr. Top. Membr.* **76**: 337–371
- Clohisey SMR, Dzhindzhev NS & Ohkura H (2014) Kank is an EB1 interacting protein that localises to muscle-Tendon attachment sites in drosophila. *PLoS One* **9**:
- Day CL, Puthalakath H, Skea G, Strasser A, Barsukov I, Lian LY, Huang DCS & Hinds MG (2004) Localization of dynein light chains 1 and 2 and their pro-apoptotic ligands. *Biochem. J.* **377**: 597–605
- Debrand E, El Jai Y, Spence L, Bate N, Praekelt U, Pritchard CA, Monkley SJ & Critchley DR (2009) Talin 2 is a large and complex gene encoding multiple transcripts and protein isoforms. *FEBS J.* **276**: 1610–1628
- Dedden D, Schumacher S, Kelley CF, Zacharias M, Biertümpfel C, Fässler R & Mizuno N (2019) The Architecture of Talin1 Reveals an Autoinhibition Mechanism. *Cell* **179**: 120-131.e13
Available at: <https://doi.org/10.1016/j.cell.2019.08.034> [Accessed May 11, 2020]
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277–293
- Delva E, Tucker DK & Kowalczyk AP (2009) The desmosome. *Cold Spring Harb. Perspect. Biol.* **1**:
- Desgrosellier JS & Cheresch DA (2010) Integrins in cancer: Biological implications and therapeutic opportunities. *Nat. Rev. Cancer* **10**: 9–22
- Dick T, Ray K, Salz HK & Chia W (1996) Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in *Drosophila melanogaster*. *Mol. Cell. Biol.* **16**: 1966–1977
- Ding M, Goncharov A, Jin Y & Chisholm AD (2003) *C. elegans* ankyrin repeat protein VAB-19 is a component of epidermal attachment structures and is essential for epidermal morphogenesis. *Development* **130**: 5791–5801
- Ding M, King RS, Berry EC, Wang Y, Hardin J & Chisholm AD (2008) The cell signaling adaptor protein EPS-8 is essential for *C. elegans* epidermal elongation and interacts with the ankyrin repeat protein VAB-19. *PLoS One* **3**:

- Drabek K, van Ham M, Stepanova T, Draegestein K, van Horssen R, Sayas CL, Akhmanova A, ten Hagen T, Smits R, Fodde R, Grosveld F & Galjart N (2006) Role of CLASP2 in Microtubule Stabilization and the Regulation of Persistent Motility. *Curr. Biol.* **16**: 2259–2264
- Du X, Plow EF, Frelinger AL, O'Toole TE, Loftus JC & Ginsberg MH (1991) Ligands 'activate' integrin α IIb β 3 (platelet GPIIb-IIIa). *Cell* **65**: 409–416
- Dyson HJ & Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**: 197–208
- Efimov A, Schiefermeier N, Grigoriev I, Brown MC, Turner CE, Small JV & Kaverina I (2008) Paxillin-dependent stimulation of microtubule catastrophes at focal adhesion sites. *J. Cell Sci.* **121**: 196–204
- Elliott PR, Goult BT, Kopp PM, Bate N, Grossmann JG, Roberts GCK, Critchley DR & Barsukov IL (2010) The Structure of the Talin Head Reveals a Novel Extended Conformation of the FERM Domain. *Structure* **18**: 1289–1299
- Ellis SJ, Goult BT, Fairchild MJ, Harris NJ, Long J, Lobo P, Czerniecki S, Van Petegem F, Schöck F, Peifer M & Tanentzapf G (2013) Talin autoinhibition is required for morphogenesis. *Curr. Biol.* **23**: 1825–1833
- Elmendorf HG, Rohrer SC, Khoury RS, Bouttenot RE & Nash TE (2005) Examination of a novel head-stalk protein family in *Giardia lamblia* characterised by the pairing of ankyrin repeats and coiled-coil domains. *Int. J. Parasitol.* **35**: 1001–1011
- Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**: 486–501
- Engelman JA, Luo J & Cantley LC (2006) The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat. Rev. Genet.* **7**: 606–619
- Erickson HP & O'Brien ET (1992) Microtubule dynamic instability and GTP hydrolysis. *Annu. Rev. Biophys. Biomol. Struct.* **21**: 145–166
- Espindola FS, Suter DM, Partata LBE, Cao T, Wolenski JS, Cheney RE, King SM & Mooseker MS (2000) The light chain composition of chicken brain myosin-Va: Calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN. *Cell Motil. Cytoskeleton* **47**: 269–281

- Evans PR (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**: 282–292
- Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**: 1204–1214
- Fan JS, Zhang Q, Tochio H & Zhang M (2002) Backbone dynamics of the 8 kDa dynein light chain dimer reveals molecular basis of the protein's functional diversity. *J. Biomol. NMR* **23**: 103–114
- Forster T (1946) Energiewanderung und Fluoreszenz. *Naturwissenschaften* **33**: 166–175
- Frelinger AL, Cohen I, Plow EF, Smith MA, Roberts J, Lam SCT & Ginsberg MH (1990) Selective inhibition of integrin function by antibodies specific for ligand-occupied receptor conformers. *J. Biol. Chem.* **265**: 6346–6352
- Gallego P, Velazquez-Campoy A, Regue L, Roig J & Reverter D (2013) Structural analysis of the regulation of the DYNLL/LC8 binding to Nek9 by phosphorylation. *J. Biol. Chem.* **288**: 12283–12294
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* pp 571–607.
- Geblinger D, Addadi L & Geiger B (2010) Nano-topography sensing by osteoclasts: (Journal of Cell Science 123, (1503-1510)). *J. Cell Sci.* **123**: 1814
- Gee HY, Saisawat P, Ashraf S, Hurd TW, Vega-Warner V, Fang H, Beck BB, Gribouval O, Zhou W, Diaz KA, Natarajan S, Wiggins RC, Lovric S, Chernin G, Schoeb DS, Ovunc B, Frishberg Y, Soliman NA, Fathy HM, Goebel H, et al (2013) ARHGDI1 mutations cause nephrotic syndrome via defective RHO GTPase signaling. *J. Clin. Invest.* **123**: 3243–3253
- Gee HY, Zhang F, Ashraf S, Kohl S, Sadowski CE, Vega-Warner V, Zhou W, Lovric S, Fang H, Nettleton M, Zhu JY, Hoefele J, Weber LT, Podracka L, Boor A, Fehrenbach H, Innis JW, Washburn J, Levy S, Lifton RP, et al (2015) KANK deficiency leads to podocyte dysfunction and nephrotic syndrome. *J. Clin. Invest.* **125**: 2375–2384
- Gingras AR, Bate N, Goult BT, Hazelwood L, Canestrelli I, Grossmann JG, Liu HJ, Putz NSM, Roberts GCK, Volkmann N, Hanein D, Barsukov IL & Critchley DR (2008) The structure of the C-

- terminal actin-binding domain of talin. *EMBO J.* **27**: 458–469
- Gingras AR, Bate N, Goult BT, Patel B, Kopp PM, Emsley J, Barsukov IL, Roberts GCK & Critchley DR (2010) Central region of talin has a unique fold that binds vinculin and actin. *J. Biol. Chem.* **285**: 29577–29587
- Gingras AR, Ziegler WH, Frank R, Barsukov IL, Roberts GCK, Critchley DR & Emsley J (2005) Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. *J. Biol. Chem.* **280**: 37217–37224
- Goksoy E, Ma YQ, Wang X, Kong X, Perera D, Plow EF & Qin J (2008a) Structural Basis for the Autoinhibition of Talin in Regulating Integrin Activation. *Mol. Cell* **31**: 124–133 Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2522368/pdf/nihms-59842.pdf> [Accessed January 23, 2019]
- Goksoy E, Ma YQ, Wang X, Kong X, Perera D, Plow EF & Qin J (2008b) Structural Basis for the Autoinhibition of Talin in Regulating Integrin Activation. *Mol. Cell* **31**: 124–133
- Golias C, Batistatou A, Bablekos G, Charalabopoulos A, Peschos D, Mitsopoulos P & Charalabopoulos K (2011) Physiology and pathophysiology of selectins, integrins, and IgSf cell adhesion molecules focusing on inflammation. A paradigm model on infectious endocarditis. *Cell Commun. Adhes.* **18**: 19–32
- Gough RE & Goult BT (2018) The tale of two talins – two isoforms to fine-tune integrin signalling. *FEBS Lett.* **592**: 2108–2125
- Goult BT, Bate N, Anthis NJ, Wegener KL, Gingras AR, Patel B, Barsukov IL, Campbell ID, Roberts GCK & Critchley DR (2009a) The structure of an interdomain complex that regulates Talin activity. *J. Biol. Chem.* **284**: 15097–15106 Available at: <http://www.rcsb.org/> [Accessed March 15, 2019]
- Goult BT, Bate N, Anthis NJ, Wegener KL, Gingras AR, Patel B, Barsukov IL, Campbell ID, Roberts GCK & Critchley DR (2009b) The structure of an interdomain complex that regulates Talin activity. *J. Biol. Chem.* **284**: 15097–15106
- Goult BT, Bouaouina M, Elliott PR, Bate N, Patel B, Gingras AR, Grossmann JG, Roberts GCK, Calderwood DA, Critchley DR & Barsukov IL (2010) Structure of a double ubiquitin-like domain in the talin head: A role in integrin activation. *EMBO J.* **29**: 1069–1080

- Goult BT, Xu XP, Gingras AR, Swift M, Patel B, Bate N, Kopp PM, Barsukov IL, Critchley DR, Volkmann N & Hanein D (2013a) Structural studies on full-length talin1 reveal a compact auto-inhibited dimer: Implications for talin activation. *J. Struct. Biol.* **184**: 21–32
- Goult BT, Yan J & Schwartz MA (2018) Talin as a mechanosensitive signaling hub. *J. Cell Biol.* **217**: 3776–3784
- Goult BT, Zacharchenko T, Bate N, Tsang R, Hey F, Gingras AR, Elliott PR, Roberts GCK, Ballestrem C, Critchley DR & Barsukov IL (2013b) RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J. Biol. Chem.* **288**: 8238–8249
- Grigoriev I, Splinter D, Keijzer N, Wulf PS, Demmers J, Ohtsuka T, Modesti M, Maly I V., Grosveld F, Hoogenraad CC & Akhmanova A (2007) Rab6 Regulates Transport and Targeting of Exocytotic Carriers. *Dev. Cell* **13**: 305–314
- Gu Y & Zhang M (2018) Upregulation of the Kank1 gene inhibits human lung cancer progression in vitro and in vivo. *Oncol. Rep.* **40**: 1243–1250
- Gumbiner BM (1996) Cell adhesion: The molecular basis of tissue architecture and morphogenesis. *Cell* **84**: 345–357
- Gundelfinger ED & Fejtova A (2012) Molecular organization and plasticity of the cytomatrix at the active zone. *Curr. Opin. Neurobiol.* **22**: 423–430
- Guo Q, Liao S, Zhu Z, Li Y, Li F & Xu C (2018) Structural basis for the recognition of kinesin family member 21A (KIF21A) by the ankyrin domains of KANK1 and KANK2 proteins. *J. Biol. Chem.* **293**: 557–566
- Guo SS, Seiwert A, Szeto IYY & Fässler R (2021) Tissue distribution and subcellular localization of the family of Kidney Ankyrin Repeat Domain (KANK) proteins. *Exp. Cell Res.*
- Haage A, Goodwin K, Whitewood A, Camp D, Bogutz A, Turner CT, Granville DJ, Lefebvre L, Plotnikov S, Goult BT & Tanentzapf G (2018) Talin Autoinhibition Regulates Cell-ECM Adhesion Dynamics and Wound Healing In Vivo. *Cell Rep.* **25**: 2401-2416.e5
- Hamidi H & Ivaska J (2018) Every step of the way: integrins in cancer progression and metastasis. *Nat. Rev. Cancer*: 1–16
- Han J, Lim CJ, Watanabe N, Soriani A, Ratnikov B, Calderwood DA, Puzon-McLaughlin W, Lafuente

- EM, Boussiotis VA, Shattil SJ & Ginsberg MHH (2006) Reconstructing and Deconstructing Agonist-Induced Activation of Integrin α IIb β 3. *Curr. Biol.* **16**: 1796–1806
- Han JDJ, Berlin N, Hao T, Goldberg DS, Berriz GF, Zhang L V., Dupuy D, Walhout AJM, Cusick ME, Roth FP & Vidal M (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**: 88–93
- Han S, Dean K, Whitewood A, Bachir A, Guttierrez E, Groisman A, Horwitz A, Goult B & Danuser G (2019) Formation of talin-vinculin pre-complexes dictates maturation of nascent adhesions by accelerated force transmission and vinculin recruitment. *bioRxiv*: 735183 Available at: <https://www.biorxiv.org/content/10.1101/735183v1>
- Harburger DS & Calderwood DA (2009) Erratum: Integrin signalling at a glance (Journal of Cell Science vol. 122 (159-163)). *J. Cell Sci.* **122**: 1472
- Hartsock A & Nelson WJ (2008) Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton. *Biochim. Biophys. Acta - Biomembr.* **1778**: 660–669
- Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M & Iakoucheva LM (2006) Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* **2**: 0890–0901
- Heidary G, Engle EC & Hunter DG (2008) Congenital fibrosis of the extraocular muscles. *Semin. Ophthalmol.* **23**: 3–8
- Held RG & Kaeser PS (2018) ELKS active zone proteins as multitasking scaffolds for secretion. *Open Biol.* **8**:
- Hemmings L, Rees DJG, Ohanian V, Bolton SJ, Gilmore AP, Patel B, Priddle H, Trevithick JE, Hynes RO & Critchley DR (1996) Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *J. Cell Sci.* **109**: 2715–2726
- Hong JW, Park KW & Levine MS (2013) Temporal regulation of single-minded target genes in the ventral midline of the *Drosophila* central nervous system. *Dev. Biol.* **380**: 335–343
- Hopkins BD, Goncalves MD & Cantley LC (2020) Insulin–PI3K signalling: an evolutionarily insulated metabolic driver of cancer. *Nat. Rev. Endocrinol.* **16**: 276–283
- Horton ER, Byron A, Askari JA, Ng DHJ, Millon-Frémillon A, Robertson J, Koper EJ, Paul NR,

- Warwood S, Knight D, Humphries JD & Humphries MJ (2015) Definition of a consensus integrin adhesome and its dynamics during adhesion complex assembly and disassembly. *Nat. Cell Biol.* **17**: 1577–1587
- Hotta A, Kawakatsu T, Nakatani T, Sato T, Matsui C, Sukezane T, Akagi T, Hamaji T, Grigoriev I, Akhmanova A, Takai Y & Mimori-Kiyosue Y (2010) Laminin-based cell adhesion anchors microtubule plus ends to the epithelial cell basal cortex through LL5 α / β . *J. Cell Biol.* **189**: 901–907
- Howard J & Hyman AA (2009) Growth, fluctuation and switching at microtubule plus ends. *Nat. Rev. Mol. Cell Biol.* **10**: 569–574
- Huang DL, Bax NA, Buckley CD, Weis WI & Dunn AR (2017) Vinculin forms a directionally asymmetric catch bond with F-actin. *Science (80-.).* **357**: 703–706
- Hughes PE, Diaz-Gonzalez F, Leong L, Wu C, McDonald JA, Shattil SJ & Ginsberg MH (1996) Breaking the integrin hinge: A defined structural constraint regulates integrin signaling. *J. Biol. Chem.* **271**: 6571–6574
- Hulpiau P & van Roy F (2009) Molecular evolution of the cadherin superfamily. *Int. J. Biochem. Cell Biol.* **41**: 349–369
- Hynes RO (1992) Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell* **69**: 11–25
- Hynes RO (2002) Integrins: Bidirectional, allosteric signaling machines. *Cell* **110**: 673–687
- Ihara S, Hagedorn EJ, Morrissey MA, Chi Q, Motegi F, Kramer JM & Sherwood DR (2011) Basement membrane sliding and targeted adhesion remodels tissue boundaries during uterine-vulval attachment in *Caenorhabditis elegans*. *Nat. Cell Biol.* **13**: 641–652
- Ioerger TR, Du C & Linthicum DS (1999) Conservation of cys-cys trp structural triads and their geometry in the protein domains of immunoglobulin superfamily members. *Mol. Immunol.* **36**: 373–386
- Izard T, Tran Van Nhieu G & Bois PRJ (2006) *Shigella* applies molecular mimicry to subvert vinculin and invade host cells. *J. Cell Biol.* **175**: 465–475
- Jaffrey SR & Snyder SH (1996) PIN: An associated protein inhibitor of neuronal nitric oxide synthase. *Science (80-.).* **274**: 774–777

- Jespersen N, Estelle A, Waugh N, Davey NE, Blikstad C, Ammon YC, Akhmanova A, Ivarsson Y, Hendrix DA & Barbar E (2019) Systematic identification of recognition motifs for the hub protein LC8. *Life Sci. Alliance* **2**:
- Johnson BA & Blevins RA (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**: 603–614
- Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**: 195–202
- Jurado S, Conlan LA, Baker EK, Ng JL, Tennis N, Hoch NC, Gleeson K, Smeets M, Izon D & Heierhorst J (2012a) ATM substrate Chk2-interacting Zn²⁺ finger (ASCIZ) is a bi-functional transcriptional activator and feedback sensor in the regulation of dynein light chain (DYNLL1) expression. *J. Biol. Chem.* **287**: 3156–3164
- Jurado S, Gleeson K, O'Donnell K, Izon DJ, Walkley CR, Strasser A, Tarlinton DM & Heierhorst J (2012b) The Zinc-finger protein ASCIZ regulates B cell development via DYNLL1 and Bim. *J. Exp. Med.* **209**: 1629–1639
- Kabsch W (2010) Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**: 125–32 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20124692> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2815665>
- Kakinuma N & Kiyama R (2009) A major mutation of KIF21A associated with congenital fibrosis of the extraocular muscles type 1 (CFEOM1) enhances translocation of Kank1 to the membrane. *Biochem. Biophys. Res. Commun.* **386**: 639–644
- Kakinuma N, Roy BC, Zhu Y, Wang Y & Kiyama R (2008) Kank regulates RhoA-dependent formation of actin stress fibers and cell migration via 14-3-3 in PI3K-Akt signaling. *J. Cell Biol.* **181**: 537–549
- Kakinuma N, Zhu Y, Wang Y, Roy BC & Kiyama R (2009) Kank proteins: Structure, functions and diseases. *Cell. Mol. Life Sci.* **66**: 2651–2659 Available at: <http://www.hugo-international.org> [Accessed January 22, 2019]
- Kanchanawong P, Shtengel G, Pasapera AM, Ramko EB, Davidson MW, Hess HF & Waterman CM (2010) Nanoscale architecture of integrin-based cell adhesions. *Nature* **468**: 580–584
- Katz BZ, Zamir E, Bershadsky A, Kam Z, Yamada KM & Geiger B (2000) Physical state of the

- extracellular matrix regulates the structure and molecular composition of cell-matrix adhesions. *Mol. Biol. Cell* **11**: 1047–1060
- Kaverina I, Krylyshkina O & Small JV (1999) Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *J. Cell Biol.* **146**: 1033–1043
- Kaverina I, Rottner K & Small JV (1998) Targeting, capture, and stabilization of microtubules at early focal adhesions. *J. Cell Biol.* **142**: 181–190
- Khan RB & Goult BT (2019) Adhesions Assemble!—Autoinhibition as a Major Regulatory Mechanism of Integrin-Mediated Adhesion. *Front. Mol. Biosci.* **6**:
- Khan RB, Varela L, Cowell AR & Goult BT (2021) Biochemical Characterization of the Integrin Interactome. In *Methods in Molecular Biology*
- Kim C, Lau TL, Ulmer TS & Ginsberg MH (2009) Interactions of platelet integrin α IIb and β 3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood* **113**: 4747–4753
- Kim M, Carman C V. & Springer TA (2003) Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science (80-.)*. **301**: 1720–1725
- King SM, Barbarese E, Dillman JF, Patel-King RS, Carson JH & Kevin Pfister K (1996) Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved M(r) 8,000 light chain. *J. Biol. Chem.* **271**: 19358–19366
- Kishi M, Kummer TT, Eglen SJ & Sanes JR (2005) LL5 β : A regulator of postsynaptic differentiation identified in a screen for synaptically enriched transcripts at the neuromuscular junction. *J. Cell Biol.* **169**: 355–366
- Klapholz B & Brown NH (2017) Talin - The master of integrin adhesions. *J. Cell Sci.* **130**: 2435–2446
- Kleckner IR & Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta - Proteins Proteomics* **1814**: 942–968
- Ko J, Na M, Kim S, Lee JR & Kim E (2003) Interaction of the ERC Family of RIM-binding Proteins with the Liprin- α Family of Multidomain Proteins. *J. Biol. Chem.* **278**: 42377–42385

- Kodama A, Karakesisoglou I, Wong E, Vaezi A & Fuchs E (2003) ACF7: An essential integrator of microtubule dynamics. *Cell* **115**: 343–354
- Komarova Y, De Groot CO, Grigoriev I, Gouveia SM, Munteanu EL, Schober JM, Honnappa S, Buey RM, Hoogenraad CC, Dogterom M, Borisy GG, Steinmetz MO & Akhmanova A (2009) Mammalian end binding proteins control persistent microtubule growth. *J. Cell Biol.* **184**: 691–706
- Kopp PM, Bate N, Hansen TM, Brindle NPJ, Praekelt U, Debrand E, Coleman S, Mazzeo D, Goult BT, Gingras AR, Pritchard CA, Critchley DR & Monkley SJ (2010) Studies on the morphology and spreading of human endothelial cells define key inter- and intramolecular interactions for talin1. *Eur. J. Cell Biol.* **89**: 661–673
- Kowalczyk AP & Green KJ (2013) Structure, function, and regulation of desmosomes. In *Progress in Molecular Biology and Translational Science* pp 95–118.
- Kozlov G, Wong K, Wang W, Skubák P, Muñoz-Escobar J, Liu Y, Siddiqui N, Pannu NS & Gehring K (2018) Ankyrin repeats as a dimerization module. *Biochem. Biophys. Res. Commun.* **495**: 1002–1007
- Krendel M, Zenke FT & Bokoch GM (2002) Nucleotide exchange factor GEF-H1 mediates cross-talk between microtubules and the actin cytoskeleton. *Nat. Cell Biol.*
- Lansbergen G, Grigoriev I, Mimori-Kiyosue Y, Ohtsuka T, Higa S, Kitajima I, Demmers J, Galjart N, Houtsmuller AB, Grosveld F & Akhmanova A (2006) CLASPs Attach Microtubule Plus Ends to the Cell Cortex through a Complex with LL5 β . *Dev. Cell* **11**: 21–32
- Lau TL, Kim C, Ginsberg MH & Ulmer TS (2009) The structure of the integrin α IIb β 3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* **28**: 1351–1361
- Lee HS, Anekal P, Lim CJ, Liu CC & Ginsberg MH (2013) Two modes of integrin activation form a binary molecular switch in adhesion maturation. *Mol. Biol. Cell* **24**: 1354–1362
- Lee HS, Lim CJ, Puzon-McLaughlin W, Shattil SJ & Ginsberg MH (2009) RIAM activates integrins by linking talin to Ras GTPase membrane-targeting sequences. *J. Biol. Chem.* **284**: 5119–5122
- Lerer I, Sagi M, Meiner V, Cohen T, Zlotogora J & Abeliovich D (2005) Deletion of the ANKRD15 gene at 9p24.3 causes parent-of-origin-dependent inheritance of familial cerebral palsy.

Hum. Mol. Genet. **14**: 3911–3920

Leshchyns'ka I & Sytnyk V (2016) Reciprocal interactions between cell adhesion molecules of the immunoglobulin superfamily and the cytoskeleton in neurons. *Front. Cell Dev. Biol.* **4**:

Lesk AM & Chothia C (1982) Evolution of proteins formed by β -sheets. II. The core of the immunoglobulin domains. *J. Mol. Biol.* **160**: 325–342

Ley K, Laudanna C, Cybulsky MI & Nourshargh S (2007) Getting to the site of inflammation: The leukocyte adhesion cascade updated. *Nat. Rev. Immunol.* **7**: 678–689

Li G, Du X, Vass WC, Papageorge AG, Lowy DR & Qian X (2011) Full activity of the deleted in liver cancer 1 (DLC1) tumor suppressor depends on an LD-like motif that binds talin and focal adhesion kinase (FAK). *Proc. Natl. Acad. Sci. U. S. A.* **108**: 17129–17134

Li J, Mahajan A & Tsai MD (2006) Ankyrin repeat: A unique motif mediating protein-protein interactions. *Biochemistry* **45**: 15168–15178

Li J, Su Y, Xia W, Qin Y, Humphries MJ, Vestweber D, Cabañas C, Lu C & Springer TA (2017) Conformational equilibria and intrinsic affinities define integrin activation. *EMBO J.* **36**: 629–645

Li W, Yi P & Ou G (2015) Somatic CRISPR-Cas9-induced mutations reveal roles of embryonically essential dynein chains in *Caenorhabditis elegans* cilia. *J. Cell Biol.* **208**: 683–692

Liakouli V, Cipriani P, Di Benedetto P, Ruscitti P, Carubbi F, Berardicurti O, Panzera N & Giacomelli R (2018) The role of extracellular matrix components in angiogenesis and fibrosis: Possible implication for Systemic Sclerosis. *Mod. Rheumatol.* **28**: 922–932

Liang J, Jaffrey SR, Guo W, Snyder SH & Clardy J (1999) Structure of the PIN/LC8 dimer with a bound peptide. *Nat. Struct. Biol.* **6**: 735–740

Lietha D, Cai X, Ceccarelli DFJ, Li Y, Schaller MD & Eck MJ (2007) Structural Basis for the Autoinhibition of Focal Adhesion Kinase. *Cell* **129**: 1177–1187

Lightcap CM, Kari G, Arias-Romero LE, Chernoff J, Rodeck U & Williams JC (2009) Interaction with LC8 is required for Pak1 nuclear import and is indispensable for zebrafish development. *PLoS One* **4**:

- Linder S & Kopp P (2005) Podosomes at a glance. *J. Cell Sci.* **118**: 2079–2082
- Liu J, Wang Y, Goh WI, Goh H, Baird MA, Ruehland S, Teo S, Bate N, Critchley DR, Davidson MW & Kanchanawong P (2015) Talin determines the nanoscale architecture of focal adhesions. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E4864–E4873
- Lo KWH, Naisbitt S, Fan JS, Sheng M & Zhang M (2001) The 8-kDa Dynein Light Chain Binds to its Targets via a Conserved (K/R) XTQT Motif. *J. Biol. Chem.* **276**: 14059–14066
- Lock JG, Baschieri F, Jones MC, Humphries JD, Montagnac G, Strömblad S & Humphries MJ (2019) Clathrin-containing adhesion complexes. *J. Cell Biol.* **218**: 2086–2095
- Lock JG, Jones MC, Askari JA, Gong X, Oddone A, Olofsson H, Göransson S, Lakadamyali M, Humphries MJ & Strömblad S (2018) Reticular adhesions are a distinct class of cell-matrix adhesions that mediate attachment during mitosis. *Nat. Cell Biol.* **20**: 1290–1302
- Luo BH, Carman C V. & Springer TA (2007) Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.* **25**: 619–647
- Luo M, Mengos AE, Mandarino LJ & Sekulic A (2016) Association of liprin β -1 with kank proteins in melanoma. *Exp. Dermatol.* **25**: 321–323
- Lux SE, John KM & Bennett V (1990) Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**: 36–42
- Luxenburg C, Geblinger D, Klein E, Anderson K, Hanein D, Geiger B & Addadi L (2007) The architecture of the adhesive apparatus of cultured osteoclasts: From podosome formation to sealing zone assembly. *PLoS One* **2**:
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD & Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**: W636–W641
- Makokha M (2004) The solution structure of the pH-induced monomer of dynein light-chain LC8 from *Drosophila*. *Protein Sci.* **13**: 727–734 Available at: <http://www.proteinscience.org/cgi/doi/10.1110/ps.03462204>. [Accessed October 8, 2020]
- Makokha M, Hare M, Li M, Hays T & Barbar E (2002) Interactions of cytoplasmic dynein light

- chains Tctex-1 and LC8 with the intermediate chain IC74. *Biochemistry* **41**: 4302–4311
- Malabarba MG, Milia E, Faretta M, Zamponi R, Pelicci PG & Di Fiore PP (2001) A repertoire library that allows the selection of synthetic SH2ITALT with altered binding specificities. *Oncogene* **20**: 5186–5194
- Mana G, Clapero F, Panieri E, Panero V, Böttcher RT, Tseng HY, Saltarin F, Astanina E, Wolanska KI, Morgan MR, Humphries MJ, Santoro MM, Serini G & Valdembri D (2016) PPFIA1 drives active $\alpha 5\beta 1$ integrin recycling and controls fibronectin fibrillogenesis and vascular morphogenesis. *Nat. Commun.* **7**:
- Matsumoto S, Fumoto K, Okamoto T, Kaibuchi K & Kikuchi A (2010) Binding of APC and dishevelled mediates Wnt5a-regulated focal adhesion dynamics in migrating cells. *EMBO J.* **29**: 1192–1204
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**: 658–674
- McDonald JA, Kelley DG & Broekelmann TJ (1982) Role of fibronectin in collagen deposition: Fab' to the gelatin-binding domain of fibronectin inhibits both fibronectin and collagen organization in fibroblast extracellular matrix. *J. Cell Biol.*
- Mehrbod M, Trisno S & Mofrad MRK (2013) On the activation of integrin $\alpha 1\beta 3$: Outside-in and inside-out pathways. *Biophys. J.* **105**: 1304–1315
- Mercurio AM, Rabinovitz I & Shaw LM (2001) The $\alpha 6\beta 4$ integrin and epithelial cell migration. *Curr. Opin. Cell Biol.*
- Mimori-Kiyosue Y, Grigoriev I, Lansbergen G, Sasaki H, Matsui C, Severin F, Galjart N, Grosveld F, Vorobjev I, Tsukita S & Akhmanova A (2005) CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. *J. Cell Biol.* **168**: 141–153
- Mitchison T & Kirschner M (1984) Dynamic instability of microtubule growth. *Nature* **312**: 237–242
- Molony L, McCaslin D, Abernethy J, Paschal B & Burridge K (1987) Properties of talin from chicken gizzard smooth muscle. *J. Biol. Chem.* **262**: 7790–7795
- Montcouquiol M, Rachel RA, Lanford PJ, Copeland NG, Jenkins NA & Kelley MW (2003)

- Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* **423**: 173–177
- Mosavi LK, Cammett TJ, Desrosiers DC & Peng Z (2004) The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* **13**: 1435–1448
- Moser M, Nieswandt B, Ussar S, Pozgajova M & Fässler R (2008) Kindlin-3 is essential for integrin activation and platelet aggregation. *Nat. Med.* **14**: 325–330
- Muller WA (2013) Getting Leukocytes to the Site of Inflammation. *Vet. Pathol.* **50**: 7–22
- Murphy DA & Courtneidge SA (2011) The ‘ins’ and ‘outs’ of podosomes and invadopodia: Characteristics, formation and function. *Nat. Rev. Mol. Cell Biol.* **12**: 413–426
- Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**: 355–367
- Nielsen H & Wennström (2012) Cell adhesion molecules in Alzheimer’s disease. *Degener. Neurol. Neuromuscul. Dis.:* 65
- Nyarko A, Cochrun L, Norwood S, Pursifull N, Voth A & Barbar E (2005) Ionization of His 55 at the dimer interface of dynein light-chain LC8 is coupled to dimer dissociation. *Biochemistry* **44**: 14248–14255
- Orekhov VY & Jaravine VA (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**: 271–292
- Pan W, Sun K, Tang K, Xiao Q, Ma C, Yu C & Wei Z (2018) Structural insights into ankyrin repeat-mediated recognition of the kinesin motor protein KIF21A by KANK1, a scaffold protein in focal adhesion. *J. Biol. Chem.* **293**: 1944–1956
- Pankov R, Cukierman E, Katz BZ, Matsumoto K, Lin DC, Lin S, Hahn C & Yamada KM (2000) Integrin dynamics and matrix assembly: Tensin-dependent translocation of $\alpha 5\beta 1$ integrins promotes early fibronectin fibrillogenesis. *J. Cell Biol.*
- Panni S & Dente L (2002) In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *J. Biol. Chem.* **277**: 21666–21674

- Paoletti A & Phong TT (2007) Anchoring microtubules at the spindle poles. *Nat. Cell Biol.* **9**: 619–621
- Paradžik M, Humphries JD, Stojanović N, Nestić D, Majhen D, Dekanić A, Samaržija I, Sedda D, Weber I, Humphries MJ & Ambriović-Ristov A (2020) KANK2 Links α V β 5 Focal Adhesions to Microtubules and Regulates Sensitivity to Microtubule Poisons and Cell Migration. *Front. Cell Dev. Biol.* **8**:
- Parsons JT, Horwitz AR & Schwartz MA (2010) Cell adhesion: Integrating cytoskeletal dynamics and cellular tension. *Nat. Rev. Mol. Cell Biol.* **11**: 633–643
- Patil A, Kinoshita K & Nakamura H (2010) Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci.* **11**: 1930–1943
- Pfister KK, Shah PR, Hummerich H, Russ A, Cotton J, Annuar AA, King SM & Fisher EMC (2006) Genetic analysis of the cytoplasmic dynein subunit families. *PLoS Genet.* **2**: 11–26
- Ponti A, Machacek M, Gupton SL, Waterman-Storer CM & Danuser G (2004) Two distinct actin networks drive the protrusion of migrating cells. *Science (80-.).* **305**: 1782–1786
- Proszynski TJ & Sanes JR (2013) Amotl2 interacts with LL5 β , localizes to podosomes and regulates postsynaptic differentiation in muscle. *J. Cell Sci.* **126**: 2225–2235
- Pufall MA & Graves BJ (2002) Autoinhibitory domains: Modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* **18**: 421–462 Available at: www.annualreviews.org [Accessed March 12, 2019]
- Puthalakath H, Huang DCS, O'Reilly LA, King SM & Strasser A (1999) The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex. *Mol. Cell* **3**: 287–296
- Puthalakath H, Villunger A, O'Reilly LA, Beaumont JG, Coultas L, Cheney RE, Huang DCS & Strasser A (2001) Bmf: A proapoptotic BH3-only protein regulated by interaction with the myosin V actin motor complex, activated by anoikis. *Science (80-.).* **293**: 1829–1832
- Radhakrishnan I, Pérez-Alvarado GC, Parker D, Dyson HJ, Montminy MR & Wright PE (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions. *Cell* **91**: 741–752

- Radnai L, Rapali P, Hódi Z, Süveges D, Molnár T, Kiss B, Bécsi B, Erdödi F, Buday L, Kardos J, Kovács M & Nyitray L (2010) Affinity, avidity, and kinetics of target sequence binding to LC8 dynein light chain isoforms. *J. Biol. Chem.* **285**: 38649–38657
- Rafiq NBM, Nishimura Y, Plotnikov S V., Thiagarajan V, Zhang Z, Shi S, Natarajan M, Viasnoff V, Kanchanawong P, Jones GE & Bershadsky AD (2019) A mechano-signalling network linking microtubules, myosin IIA filaments and integrin-based adhesions. *Nat. Mater.* **18**: 638–649
- Ramot Y, Molho-Pessach V, Meir T, Alper-Pinus R, Siam I, Tams S, Babay S & Zlotogorski A (2014) Mutation in KANK2, encoding a sequestering protein for steroid receptor coactivators, causes keratoderma and woolly hair. *J. Med. Genet.* **51**: 388–394
- Reck-Peterson SL, Redwine WB, Vale RD & Carter AP (2018) The cytoplasmic dynein transport machinery and its many cargoes. *Nat. Rev. Mol. Cell Biol.* **19**: 382–398
- Richards JP, Bächinger HP, Goodman RH & Brennan RG (1996) Analysis of the structural properties of cAMP-responsive element-binding protein (CREB) and phosphorylated CREB. *J. Biol. Chem.* **271**: 13716–13723
- Rossier O & Giannone G (2016) The journey of integrins and partners in a complex interactions landscape studied by super-resolution microscopy and single protein tracking. *Exp. Cell Res.* **343**: 28–34 Available at: <http://dx.doi.org/10.1016/j.yexcr.2015.11.004>
- Roy BC, Kakinuma N & Kiyama R (2009) Kank attenuates actin remodeling by preventing interaction between IRSp53 and Rac1. *J. Cell Biol.* **184**: 253–267
- Rozario T & DeSimone DW (2010) The extracellular matrix in development and morphogenesis: A dynamic view. *Dev. Biol.* **341**: 126–140
- Sabatier L, Chen D, Fagotto-Kaufmann C, Hubmacher D, McKee MD, Annis DS, Osher DF & Reinhardt DP (2009) Fibrillin assembly requires fibronectin. *Mol. Biol. Cell*
- Saito M, Tucker DK, Kohlhorst D, Niessen CM & Kowalczyk AP (2012) Classical and desmosomal cadherins at a glance. *J. Cell Sci.* **125**: 2547–2552
- Sarkar S, Roy BC, Hatano N, Aoyagi T, Gohji K & Kiyama R (2002) A novel ankyrin repeat-containing gene (Kank) located at 9p24 is a growth suppressor of renal cell carcinoma. *J. Biol. Chem.* **277**: 36585–36591

- Schiemer J, Bohm A, Lin L, Merrill-Skoloff G, Flaumenhaft R, Huang JS, Le Breton GC & Chishti AH (2016) α 13 switch region 2 relieves talin autoinhibition to activate α 1b β 3 integrin. *J. Biol. Chem.* **291**: 26598–26612
- Schmoranzler J, Kreitzer G & Simon SM (2003) Migrating fibroblasts perform polarized, microtubule-dependent exocytosis towards the leading edge. *J. Cell Sci.* **116**: 4513–4519
- Schnorrer F, Bohmann K & Nusslein-Volhard C (2000) The molecular motor dynein is involved in targeting swallow and bicoid RNA to the anterior pole of *Drosophila* oocytes. *Nat. Cell Biol.* **2**: 185–190
- Scita G, Confalonieri S, Lappalainen P & Suetsugu S (2008) IRSp53: crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends Cell Biol.* **18**: 52–60
- Seddiki R, Narayana GHNS, Strale PO, Balcioglu HE, Peyret G, Yao M, Le AP, Lim CT, Yan J, Ladoux B & Mège RM (2018) Force-dependent binding of vinculin to α -catenin regulates cell-cell contact stability and collective cell behavior. *Mol. Biol. Cell* **29**: 380–388
- Seetharaman S & Etienne-Manneville S (2019) Microtubules at focal adhesions - a double-edged sword. *J. Cell Sci.*
- Senetar MA & McCann RO (2005) Gene duplication and functional divergence during evolution of the cytoskeletal linker protein talin. *Gene* **362**: 141–152
- Sepp KJ, Hong P, Lizarraga SB, Liu JS, Mejia LA, Walsh CA & Perrimon N (2008) Identification of neural outgrowth genes using genome-wide RNAi. *PLoS Genet.* **4**:
- Sharp DJ, Rogers GC & Scholey JM (2000) Cytoplasmic dynein is required for poleward chromosome movement during mitosis in *Drosophila* embryos. *Nat. Cell Biol.* **2**: 922–930
- Shen B, Delaney MK & Du X (2012) Inside-out, outside-in, and inside-outside-in: G protein signaling in integrin-mediated cell adhesion, spreading, and retraction. *Curr. Opin. Cell Biol.* **24**: 600–606
- Shen Y & Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**: 227–241
- Shimaoka M, Takagi J & Springer TA (2002) Conformational regulation of integrin structure and function. *Annu. Rev. Biophys. Biomol. Struct.* **31**: 485–516

- Skinner SP, Goult BT, Fogh RH, Boucher W, Stevens TJ, Laue ED & Vuister GW (2015) Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**: 154–161
- Song C, Wen Y, Rayala SK, Chen M, Ma J, Zhang M & Kumar R (2008) Serine 88 phosphorylation of the 8-kDa dynein light chain 1 is a molecular switch for its dimerization status and functions. *J. Biol. Chem.* **283**: 4004–4013
- Song X, Yang J, Hirbawi J, Ye S, Perera HD, Goksoy E, Dwivedi P, Plow EF, Zhang R & Qin J (2012) A novel membrane-dependent on/off switch mechanism of talin FERM domain at sites of cell adhesion. *Cell Res.* **22**: 1533–1545
- Song Y, Benison G, Nyarko A, Hays TS & Barbar E (2007) Potential role for phosphorylation in differential regulation of the assembly of jdynein light chains. *J. Biol. Chem.* **282**: 17272–17279
- Sottile J & Hocking DC (2002) Fibronectin polymerization regulates the composition and stability of extracellular matrix fibrils and cell-matrix adhesions. *Mol. Biol. Cell*
- Srivastava A, Pastor-Pareja JC, Igaki T, Pagliarini R & Xu T (2007) Basement membrane remodeling is essential for Drosophila disc eversion and tumor invasion. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 2721–2726
- Stehbens S & Wittmann T (2012) Targeting and transport: How microtubules control focal adhesion dynamics. *J. Cell Biol.* **198**: 481–489
- Stehbens SJ, Paszek M, Pemble H, Ettinger A, Gierke S & Wittmann T (2014) CLASPs link focal-adhesion-associated microtubule capture to localized exocytosis and adhesion site turnover. *Nat. Cell Biol.* **16**: 558–570
- Stepanenko O V., Stepanenko O V., Kuznetsova IM, Verkhusha V V. & Turoverov KK (2013) Beta-Barrel Scaffold of Fluorescent Proteins. Folding, Stability and Role in Chromophore Formation. In *International Review of Cell and Molecular Biology* pp 221–278.
- Stubb A, Guzmán C, Närvä E, Aaron J, Chew TL, Saari M, Miihkinen M, Jacquemet G & Ivaska J (2019) Superresolution architecture of cornerstone focal adhesions in human pluripotent stem cells. *Nat. Commun.* **10**:
- Sun Z, Tseng HY, Tan S, Senger F, Kurzawa L, Dedden D, Mizuno N, Wasik AA, Thery M, Dunn AR

- & Fssler R (2016a) Kank2 activates talin, reduces force transduction across integrins and induces central adhesion formation. *Nat. Cell Biol.* **18**: 941–953
- Sun Z, Tseng HY, Tan S, Senger F, Kurzawa L, Dedden D, Mizuno N, Wasik AA, Thery M, Dunn AR & Fssler R (2016b) Kank2 activates talin, reduces force transduction across integrins and induces central adhesion formation. *Nat. Cell Biol.* **18**: 941–953
- Suzuki S & Naitoh Y (1990) Amino acid sequence of a novel integrin β 4 subunit and primary expression of the mRNA in epithelial cells. *EMBO J.*
- Takito J, Inoue S & Nakamura M (2018) The sealing zone in osteoclasts: A self-organized structure on the bone. *Int. J. Mol. Sci.* **19**:
- Tamkun JW, DeSimone DW, Fonda D, Patel RS, Buck C, Horwitz AF & Hynes RO (1986) Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell* **46**: 271–282
- Theocharis AD, Skandalis SS, Gialeli C & Karamanos NK (2016) Extracellular matrix structure. *Adv. Drug Deliv. Rev.* **97**: 4–27
- Theriot JA & Mitchison TJ (1991) Actin microfilament dynamics in locomoting cells. *Nature* **352**: 126–131
- Thomson NC (1992) Anti-inflammatory therapies Available at: <http://www.bioscience>. [Accessed March 5, 2019]
- Tsien RY (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**: 509–544
- Uhlik MT, Temple B, Bencharit S, Kimple AJ, Siderovski DP & Johnson GL (2005) Structural and evolutionary division of phosphotyrosine binding (PTB) domains. *J. Mol. Biol.* **345**: 1–20
- van der Vaart B, van Riel WE, Doodhi H, Kevenaar JT, Katrukha EA, Gumy L, Bouchet BP, Grigoriev I, Spangler SA, Yu K Lou, Wulf PS, Wu J, Lansbergen G, van Battum EY, Pasterkamp RJ, Mimori-Kiyosue Y, Demmers J, Olieric N, Maly I V., Hoogenraad CC, et al (2013) CFEOM1-associated kinesin KIF21A is a cortical microtubule growth inhibitor. *Dev. Cell* **27**: 145–160 Available at: <http://dx.doi.org/10.1016/j.devcel.2013.09.010>
- Vallee RB, Williams JC, Varma D & Barnhart LE (2004) Dynein: An Ancient Motor Protein Involved in Multiple Modes of Transport. *J. Neurobiol.* **58**: 189–200

- Vanzo RJ, Martin MM, Sdano MR & South ST (2013) Familial KANK1 deletion that does not follow expected imprinting pattern. *Eur. J. Med. Genet.* **56**: 256–259
- Vasiliev JM, Gelfand IM, Domnina L V., Ivanova OY, Komm SG & Olshevskaja L V. (1970) Effect of colcemid on the locomotory behaviour of fibroblasts. *J. Embryol. Exp. Morphol.* **24**: 625–640
- Vicente-Manzanares M & Horwitz AR (2011) Adhesion dynamics at a glance. *J. Cell Sci.* **124**: 3923–3927
- Vinogradova O, Velyvis A, Velyviene A, Hu B, Haas TA, Plow EF & Qin J (2002) A structural mechanism of integrin $\alpha\text{IIb}\beta\text{3}$ 'inside-out' activation as regulated by its cytoplasmic face. *Cell* **110**: 587–597
- Vong S & Kalluri R (2011) The Role of Stromal Myofibroblast and Extracellular Matrix in Tumor Angiogenesis. *Genes and Cancer* **2**: 1139–1145
- Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T & Bricogne G (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**: 293–302
- Wai Wong C, Dye DE & Coombe DR (2012) The role of immunoglobulin superfamily cell adhesion molecules in cancer metastasis. *Int. J. Cell Biol.*
- Wang Y & McNiven MA (2012) Invasive matrix degradation at focal adhesions occurs via protease recruitment by a FAK-p130Cas complex. *J. Cell Biol.* **196**: 375–385
- Wang Y, Onishi Y, Kakinuma N, Roy BC, Aoyagi T & Kiyama R (2005) Alternative splicing of the human Kank gene produces two types of Kank protein. *Biochem. Biophys. Res. Commun.* **330**: 1247–1253
- Webb DJ, Parsons JT & Horwitz AF (2002) Adhesion assembly, disassembly and turnover in migrating cells - Over and over and over again. *Nat. Cell Biol.* **4**:
- Wegener KL & Campbell ID (2008) Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (Review). *Mol. Membr. Biol.* **25**: 376–387
- Weng Z, Shang Y, Yao D, Zhu J & Zhang R (2018) Structural analyses of key features in the KANK1 · KIF21A complex yield mechanistic insights into the cross-talk between microtubules and the cell cortex. *J. Biol. Chem.* **293**: 215–225

- Whitewood AJ, Singh AK, Brown DG & Goult BT (2018) Chlamydial virulence factor TarP mimics talin to disrupt the talin-vinculin complex. *FEBS Lett.* **592**: 1751–1760
- Wilson MJ, Salata MW, Susalka SJ & Pfister KK (2001) Light chains of mammalian cytoplasmic dynein: Identification and characterization of a family of LC8 light chains. *Cell Motil. Cytoskeleton* **49**: 229–240
- Winograd-Katz SE, Fässler R, Geiger B & Legate KR (2014) The integrin adhesome: From genes and proteins to human disease. *Nat. Rev. Mol. Cell Biol.* **15**: 273–288
- Witz IP (2008) The selectin-selectin ligand axis in tumor progression. *Cancer Metastasis Rev.* **27**: 19–30
- Worbs T, Hammerschmidt SI & Förster R (2017) Dendritic cell migration in health and disease. *Nat. Rev. Immunol.* **17**: 30–48
- Xie C, Zhu J, Chen X, Mi L, Nishida N & Springer TA (2010) Structure of an integrin with an α domain, complement receptor type 4. *EMBO J.* **29**: 666–679
- Yan J, Yao M, Goult BT & Sheetz MP (2015) Talin Dependent Mechanosensitivity of Cell Focal Adhesions. *Cell. Mol. Bioeng.* **8**: 151–159
- Yang J, Zhu L, Zhang H, Hirbawi J, Fukuda K, Dwivedi P, Liu J, Byzova T, Plow EF, Wu J & Qin J (2014) Conformational activation of talin by RIAM triggers integrin-mediated cell adhesion. *Nat. Commun.* **5**: 5880
- Yao M, Goult BT, Chen H, Cong P, Sheetz MP & Yan J (2014) Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Sci. Rep.* **4**: 4610
- Yao M, Goult BT, Klapholz B, Hu X, Toseland CP, Guo Y, Cong P, Sheetz MP & Yan J (2016) The mechanical response of talin. *Nat. Commun.* **7**: 11966
- Ye F, Kim C & Ginsberg MH (2012) Reconstruction of integrin activation. *Blood* **119**: 26–33
- Yu M, Le S, Ammon YC, Goult BT, Akhmanova A & Yan J (2019) Force-Dependent Regulation of Talin-KANK1 Complex at Focal Adhesions. *Nano Lett.* **19**: 5982–5990
- Yue J, Xie M, Gou X, Lee P, Schneider MD & Wu X (2014) Microtubules Regulate Focal Adhesion Dynamics through MAP4K4. *Dev. Cell* **31**: 572–585

- Zacharchenko T, Qian X, Goult BTT, Jethwa D, Almeida TBB, Ballestrem C, Critchley DRR, Lowy DRR & Barsukov ILL (2016) LD Motif Recognition by Talin: Structure of the Talin-DLC1 Complex. *Structure* **24**: 1130–1141
- Zaidel-Bar R, Ballestrem C, Kam Z & Geiger B (2003) Early molecular events in the assembly of matrix adhesions at the leading edge of migrating cells. *J. Cell Sci.* **116**: 4605–4613
- Zamir E, Katz M, Posen Y, Erez N, Yamada KM, Katz BZ, Lin S, Lin DC, Bershadsky A, Kam Z & Geiger B (2000) Dynamics and segregation of cell-matrix adhesions in cultured fibroblasts. *Nat. Cell Biol.* **2**: 191–196
- Zaytseva O, Tennis N, Mitchell N, Kanno SI, Yasui A, Heierhorst J & Quinn LM (2014) The novel zinc finger protein dASCIZ regulates mitosis in Drosophila via an essential role in dynein light-chain expression. *Genetics* **196**: 443–453
- Zhang Y, Zhang H, Liang J, Yu W & Shang Y (2007) SIP, a novel ankyrin repeat containing protein, sequesters steroid receptor coactivators in the cytoplasm. *EMBO J.* **26**: 2645–2657
- Zhu L, Yang J, Bromberger T, Holly A, Lu F, Liu H, Sun K, Klapproth S, Hirbawi J, Byzova T V., Plow EF, Moser M & Qin J (2017) Structure of Rap1b bound to talin reveals a pathway for triggering integrin activation. *Nat. Commun.* **8**:
- Zhu Y, Kakinuma N, Wang Y & Kiyama R (2008) Kank proteins: A new family of ankyrin-repeat domain-containing proteins. *Biochim. Biophys. Acta - Gen. Subj.* **1780**: 128–133

APPENDIX

Appendix 1:



Adhesions Assemble!—Autoinhibition as a Major Regulatory Mechanism of Integrin-Mediated Adhesion

Rejina B. Khan and Benjamin T. Goult*

School of Biosciences, University of Kent, Canterbury, United Kingdom

OPEN ACCESS

Edited by:

Filippo Frischi,
University of Essex, United Kingdom

Reviewed by:

Christopher Cooper,
University of Huddersfield,
United Kingdom
Stephane Rety,
INSERM U1210 Laboratoire de
Biologie et Modélisation de la Cellule
(LBMC), France

***Correspondence:**

Benjamin T. Goult
btg2@kent.ac.uk

Specialty section:

This article was submitted to
Structural Biology,
a section of the journal
Frontiers in Molecular Biosciences

Received: 27 September 2019

Accepted: 26 November 2019

Published: 17 December 2019

Citation:

Khan RB and Goult BT (2019)
Adhesions Assemble!—Autoinhibition
as a Major Regulatory Mechanism of
Integrin-Mediated Adhesion.
Front. Mol. Biosci. 6:144.
doi: 10.3389/fmolb.2019.00144

The advent of cell-cell and cell-extracellular adhesion enabled cells to interact in a coherent manner, forming larger structures and giving rise to the development of tissues, organs and complex multicellular life forms. The development of such organisms required tight regulation of dynamic adhesive structures by signaling pathways that coordinate cell attachment. Integrin-mediated adhesion to the extracellular matrix provides cells with support, survival signals and context-dependent cues that enable cells to run different cellular programs. One mysterious aspect of the process is how hundreds of proteins assemble seemingly spontaneously onto the activated integrin. An emerging concept is that adhesion assembly is regulated by autoinhibition of key proteins, a highly dynamic event that is modulated by a variety of signaling events. By enabling precise control of the activation state of proteins, autoinhibition enables localization of inactive proteins and the formation of pre-complexes. In response to the correct signals, these proteins become active and interact with other proteins, ultimately leading to development of cell-matrix junctions. Autoinhibition of key components of such adhesion complexes—including core components integrin, talin, vinculin, and FAK and important peripheral regulators such as RIAM, Src, and DLC1—leads to a view that the majority of proteins involved in complex assembly might be regulated by intramolecular interactions. Autoinhibition is relieved via multiple different signals including post-translation modification and proteolysis. More recently, mechanical forces have been shown to stabilize and increase the lifetimes of active conformations, identifying autoinhibition as a means of encoding mechanosensitivity. The complexity and scope for nuanced adhesion dynamics facilitated via autoinhibition provides numerous points of regulation. In this review, we discuss what is known about this mode of regulation and how it leads to rapid and tightly controlled assembly and disassembly of cell-matrix adhesion.

Keywords: talin, autoinhibition, integrin, vinculin, riam, cell adhesion, mechanobiology

INTRODUCTION

The Origin of Complex Cell Systems

The emergence of life is an enigmatic question that humankind has pondered for millennia. At some point millions of years ago, the last universal common ancestor (LUCA) appeared and all organisms on earth descended from this initial cell (Yutin et al., 2008), or so the story goes. The steps to produce such an organism are controversial and mystifying. However, evolution from this single cell seems somewhat easier to imagine. A key step in the formation of multicellularity and more complex organisms was the development of cell adhesion molecules. The ancient origins of integrin-mediated adhesions has been traced back to the genesis of multicellularity (Sebé-Pedrós et al., 2010; Brunet and King, 2017).

Initially, these cell adhesion molecules enabled cells to form interactions with other cells to form cell-cell junctions and, in animal cells, interactions with the newly-acquired extracellular matrix (ECM) to form cell-matrix junctions. The ability of cells to form sheets of cells and attach to an underlying matrix was central to the development of multicellular animal life.

These attachment points also developed into sensitive sensory modules, able to feel the mechanics of the microenvironment and adopt the role of mechanotransduction centers—enabling cells to monitor and respond to mechanical cues and convert them into biological signals to elicit different cellular responses.

Autoinhibition as a Regulator of Protein Activity

Key to the development of complexity is the ability to regulate adhesion and control the proteins that assemble together to form adhesive structures. One way that proteins can be dynamically regulated is via formation of an intramolecular interaction that maintains the protein in an inactive state until adhesion assembly is required. The regulation of autoinhibition, and the factors that enable regulation of the activity of the protein, provides regulatable checkpoints in the system (Figure 1A).

The concept of autoinhibitory domains can be found throughout biology—these are regions of proteins which can form intramolecular interactions which regulate behavior and activity. We refer the reader to the review on autoinhibitory domains by Pufall and Graves as it provides an excellent, still highly relevant description of the fundamentals of autoinhibition (Pufall and Graves, 2002).

CELL-MATRIX ADHESIONS: ATTACHMENT AND SIGNALING CENTERS

Integrin Adhesion Complexes

The affinity of integrins for the matrix is regulated by the process of integrin activation whereby the integrin that resides in a low-affinity, compact arrangement at rest is converted to, and stabilized in, an extended high-affinity conformation for ligand (Shattil et al., 2010; Sun et al., 2016). Integrin activation serves as a paradigm for how protein activity can be regulated by conformation (Figure 1B). Following activation, huge protein complexes coalesce onto the short cytoplasmic integrin tails,

ultimately giving rise to large signaling complexes. The order of this assembly is the focus of intense research.

Autoinhibition at the Heart of Integrin Adhesion Complexes

Many of the proteins that assemble to form integrin adhesion complexes are recruited to the adhesion site in an inactive state and, upon receiving the correct signals, are activated and set to work mediating adhesion and mechanotransduction. As we will see later, autoinhibition is not limited to the pre-adhesion phase of the protein's life; autoinhibition plays a central role in coordinating the dynamic assembly and disassembly required for cell migration—cells need to adhere and detach in perfect synchrony to enable progressive directional movement. Due to limitation in space, it is not possible to detail all of the adhesion proteins that are regulated by autoinhibition. Here, we will focus on some of the best characterized to serve as paradigms of autoinhibition regulating adhesion.

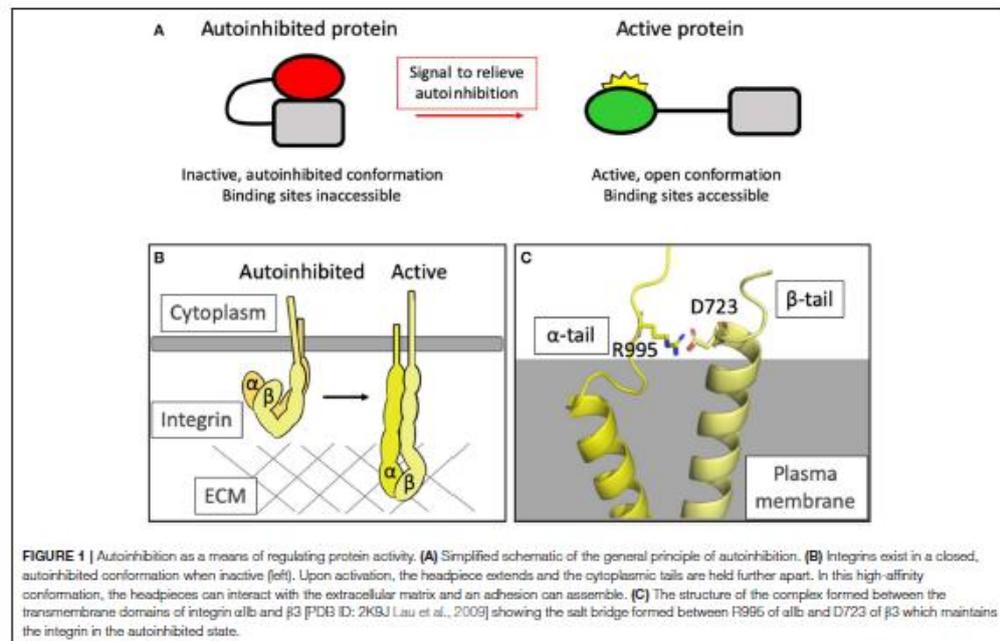
Integrins

Perhaps the best place to start is at the integrins themselves. The inside-out/outside-in activation of integrins represents a textbook example of autoinhibition and the regulation of ligand-binding affinity via allosteric effectors.

Whilst the concept of adhesions being formed between cells and the substratum had been appreciated since the work of Abercrombie (Abercrombie, 1961), integrins, as the receptors mediating these cell-ECM links, were first characterized in the 80s (Tamkun et al., 1986) and over the last 35 years have been the subject of extensive research effort (a PubMed search for the term "integrin" on 9th September 2019 yields 73,512 results). The integrin family of proteins consists of transmembrane, α/β -heterodimeric cell surface adhesion receptors that generally consist of a large ectodomain, a single transmembrane domain and a short cytoplasmic tail domain. This structure facilitates the characteristic bidirectional signaling ability of integrins wherein signals that regulate countless crucial cellular activities can be transmitted across the plasma membrane. In vertebrates, 18 α - and 8 β -subunits have been identified which form non-covalent links in different combinations to generate 24 integrin heterodimers (Hynes, 1992). Many great reviews have covered the complexity of bidirectional signaling through integrins (Qin et al., 2004; Luo and Springer, 2006; Shattil et al., 2010; Campbell and Humphries, 2011) but it is worth a quick summary here.

The concept of integrin activation has been around for over 30 years, and the first paper showing that conformational changes regulate integrins was published in 1990 (Frelinger et al., 1990). The concept of "outside-in" signaling, whereby extracellular ligands can trigger large conformational changes to the integrin structure leading to increased integrin activation, quickly followed (Du et al., 1991). The use of monoclonal antibodies that recognized epitopes on the integrins that are only accessible in certain conformations led to the notion of allosteric behavior of integrins (Mould et al., 1996; Askari et al., 2009) mediated by autoinhibition.

"Inside-out" signaling refers to integrins sensing intracellular signals, resulting in the binding of talin and kindlin to the



cytoplasmic tail of the β -subunit leading to activation of the integrin. In reality, it is likely that both of these bidirectional signaling axes work in tandem, with extra- and intracellular cues contributing constantly to orchestrate the overall dynamics of the integrin.

The structural changes in integrin conformation that occur upon activation, leading to the relief of autoinhibition and exposure of the ligand binding sites both extracellularly and intracellularly, are extensive and result in large-scale reorganization of the ectodomains (Ye et al., 2012) and separation of the transmembrane and cytoplasmic tail domains (Kim, 2003) (Figure 1B).

Integrins are generally thought to have three major conformations: a bent, closed form, an extended form with the headpiece of the ectodomain still closed, and an extended form with the headpiece open and the α - and β -cytoplasmic tails a greater distance apart (Shimaoka et al., 2002; Li et al., 2017). The binding sites for many integrin ligands are cryptic—integrin activation causes a switch to the extended, open conformation and triggers exposure of these surfaces.

Autoinhibition in the intracellular region is maintained via an electrostatic interaction between the two cytoplasmic tails: for example, in $\alpha_{IIb}\beta_3$ integrin, Asp723 in the β_3 tail binds Arg995 in the $\alpha_{IIb}\beta_3$ tail (Anthis and Campbell, 2011) (Figure 1C). This salt bridge is crucial for maintaining the low-affinity state, holding the legs together (Hughes et al., 1996; Vinogradova et al.,

2002; Kim et al., 2009; Lau et al., 2009) and preventing the separation of the legs that drives activation of the integrin.

Protein-protein interactions between the β -tail and the actin-binding protein talin (see next section) can relieve this autoinhibitory interaction and drive tail separation, propagating a conformational rearrangement of the ectodomains and unfurling to reveal the ligand-binding motifs (Harburger and Calderwood, 2009). The structure of the talin2 F2F3 domains bound to the cytoplasmic tail of the β_{1d} tail [PDB ID: 3G9W; Anthis et al., 2009] revealed that part of the activation process mediated by the talin head binding to integrins (Calderwood et al., 1999) was to not just break this salt bridge but to form an alternate salt bridge between the Asp723 and a conserved basic residue in talin [Lys327 in mouse talin2; Anthis et al., 2009]. Active talin, with a little help from the FERM domain protein kindlin (Rogalski et al., 2000; Moser et al., 2008), can therefore convert integrin from an autoinhibited to an active conformation.

Talin

Talins are ~270 kDa adaptor proteins involved in integrin-mediated adhesions and were first discovered in adhesion plaques (cell-ECM junctions) in fibroblasts (BurrIDGE and Connell, 1983). The two isoforms, talin1 and talin2, are encoded by separate genes (Senetar and McCann, 2005) but have the same domain structure consisting of a non-canonical linear FERM domain

(F0-F3) connected, via a linker, to a rod domain comprised of 13 α -helical bundles (R1-R13) followed by a C-terminal dimerization domain (DD) (Calderwood et al., 2013; Goult et al., 2013b) (Figure 2A).

Talin activates integrins by binding to the cytoplasmic tail of the β -integrin subunit via a phosphotyrosine-binding (PTB)-like region in F3—this domain is termed integrin binding site 1 (IBS1). As mentioned in the previous section, the interaction between the talin head domain with integrin is sufficient to relieve autoinhibition of the integrin, exposing the previously cryptic binding sites for integrin ligands. However, talin itself is regulated by multiple layers of autoinhibition.

The first evidence that talin was autoinhibited came from the initial structural and biochemical characterization of talin in 1987 (Molony et al., 1987). Electron microscopy analysis of purified talin from chicken gizzard smooth muscle revealed that talin was a flexible elongated molecule but could adopt a more globular compact form at low ionic strength. It has since been demonstrated that the autoinhibition of talin into this compact cytosolic conformation (Goult et al., 2013a; Dedden et al., 2019) occurs primarily via an interaction between the integrin binding site in F3 and the rod domain R9 (Goksoy et al., 2008; Goult et al., 2009; Song et al., 2012) with additional weaker interactions including that between F2F3 and R1R2 (Banno et al., 2012; Goult et al., 2013a). The compact autoinhibited structure (Molony et al., 1987; Goult et al., 2013a) is facilitated by the formation of a talin homodimer, formed via the dimerization domain (DD) at the very C-terminal helix (Gingras et al., 2008). This stabilizes the autoinhibited conformation via the various inter- and intramolecular interactions taking place. In this configuration, the integrin binding site in the F3 domain and the actin-binding sites in the rod are masked which implies that, in order for integrin activation and subsequent signaling to occur, a conformational change must be induced in order to relieve talin autoinhibition.

Layers of Talin Autoinhibition

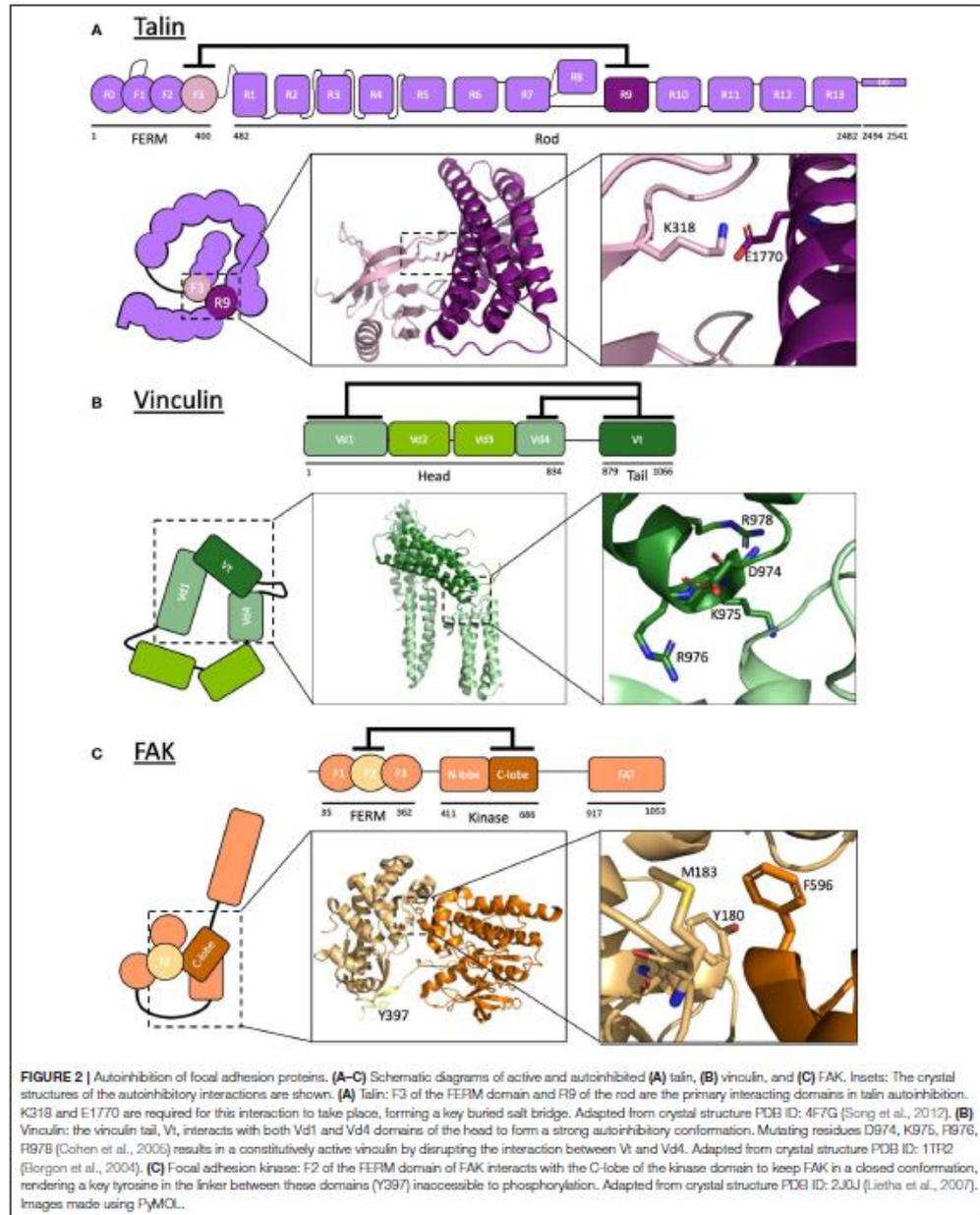
Many binding sites in talin are concealed as a result of these layers of autoinhibition (Gough and Goult, 2018)—most notably the sites for integrin, actin and vinculin—but also the sites in the rod with which the majority of talin binding partners interact. The folded conformation has the rod domains wrapped around the edge of the compact conformation with the two head domains of the dimer buried inside (Goult et al., 2013a). The rod domains have binding sites on their faces, some of which will be buried inside the closed conformation and inaccessible. However, some binding sites are still accessible—for example, Rap1-interacting adapter molecule (RIAM) is able to interact with autoinhibited talin, binding to a folded surface on R2R3 which is outward-facing. RIAM-mediated coupling of talin to Rap1's membrane-targeting motifs is a key event before integrin activation can occur (Lee et al., 2009, 2013; Shattil et al., 2010; Yao et al., 2014a; Lagarrigue et al., 2015).

R9 and integrin bind to the same site on F3, indicating that the autoinhibitory interaction of talin and its integrin binding are mutually exclusive. The specific sites for talin autoinhibition have been investigated: on R9, these include a negatively charged

surface comprised of residues Asp1676, Asp1763, Glu1770, Glu1798, and Glu1805 (Goult et al., 2009) which bind to the positively charged integrin "activation" loop on F3 including Lys316, Lys318, Lys320, Lys322, and Lys324 (Goksoy et al., 2008; Goult et al., 2009; Song et al., 2012). The crystal structure of the complex between F2F3 and R9 [PDB ID: 4F7G Song et al., 2012] and the recent cryo-EM structure of the autoinhibited form of talin (Dedden et al., 2019) have provided atomic detail of this interface and confirmed the key role of Glu1770 in mediating autoinhibition via a buried salt bridge with Lys318. Mutations to Glu1770 in R9 have been shown to reduce the autoinhibitory interaction (Goult et al., 2009), and this mutant has allowed detailed analysis of talin uncoupled from the upstream signaling pathways (see later section "Manipulation of autoinhibition by mutation").

The layers of autoinhibition in talin extend beyond the head-tail interaction. Once this "top layer" of autoinhibition is relieved, the binding sites on talin for integrin and actin are exposed and form the core of the adhesion complex. Talin has three actin-binding sites (Hemmings et al., 1996) and the regulation by autoinhibition of two actin-binding sites in the rod region, ABS2 and ABS3, demonstrates two further layers of talin regulation. The C-terminal ABS3 (McCann and Craig, 1997; Gingras et al., 2008) is comprised of the two R13 domains of the dimer linked together via the DD. ABS3 is inaccessible in the closed, compact form (Goult et al., 2013a) but, upon talin unfurling, is available to bind actin. The affinity of ABS3 for actin is further regulated via autoinhibition *within* the R13 domain, as the first helix (upstream helix, USH) of R13 limits actin binding (McCann and Craig, 1997). This domain-level autoinhibition requires force and/or changes in local pH (Srivastava et al., 2008) to allow maximal actin binding. Once ABS3 engages actin, it can capture the retrograde flow of actin that begins to exert forces onto the tethered talin molecule. These forces can cause conformational changes in talin and relieve autoinhibition or disrupt binding interactions of other domains. The second actin binding site in talin, ABS2, is in domains R4–R8 in the middle of the talin rod (Atherton et al., 2015). In the absence of force, ABS2 is cryptic and maintained in a low-affinity autoinhibited state via the adjacent domains R3 and R9 (Atherton et al., 2015). Forces exerted on talin via ABS3 relieve this autoinhibition and reveal ABS2 which can then form the high-affinity, tension-bearing cytoskeletal linkages with actin (Atherton et al., 2015; Kumar et al., 2016; Ringer et al., 2017). Here, mechanical force is the major driver relieving ABS2 autoinhibition, but this only occurs following relief of the layers of autoinhibition preceding it: i.e., talin head-tail autoinhibition and the changes in ABS3 that facilitate actin binding to talin.

Many of the binding partners of talin require sites which are not constitutively accessible. Rod domain partners may bind when domains are folded, unfolded, or at intermediate levels of folding (Goult et al., 2013b, 2018). These sites can be exposed by various stimuli; for example, talin has 11 cryptic vinculin-binding sites (VBS) which are buried in the hydrophobic core of the folded rod helical bundles. These VBS are exposed when talin is under force, causing sequential rod domain unfolding and allowing the first subdomain of the vinculin head, Vd1,



to bind (Hytönen and Vogel, 2008; del Rio et al., 2009; Yao et al., 2016). Vinculin binding to these VBS subsequently allows for stabilization and maturation of the adhesion (Yao et al., 2014a). Each VBS-containing rod domain has a different force threshold at which it unfolds, leading to different forces required to relieve autoinhibition of each VBS. Further, each VBS helix has a different mechanical stability, and so the VBS-vinculin interactions at each site also have different strengths (Wang et al., 2019).

It remains to be determined exactly which binding sites are available in each conformation, and the catalog of talin ligands is expanding constantly (Goult et al., 2018). Furthermore, the lifetimes of each conformation are regulated by autoinhibition modulated by many different signaling cues including PTMs, force and calpain cleavage (see later section "Control of autoinhibition").

This leads to a domino effect of autoinhibitory relief steps downstream of Rap1 activation: (i) RIAM is activated, which translocates talin to the plasma membrane where (ii) the talin head-tail autoinhibition is relieved. (iii) Once active, talin relieves the autoinhibition of integrin. (iv) By connecting integrins to the actin retrograde flow, talin autoinhibition is further relieved, first with activation of enhanced actin binding to ABS3. (v) As force increases, mechanical activation of vinculin-binding sites in talin occurs, initially by unfolding of the R3 domain, the least stable of the rod domains (Goult et al., 2013b; Yao et al., 2014a). (vi) Mechanical exposure of high-affinity actin binding takes place via relief of autoinhibition of ABS2. This is just one example of a simplified, linear route through the autoinhibitory landscape regulating adhesion assembly; many other factors can feed into and modulate these steps and the order in which they occur.

Vinculin

Like talin, vinculin was first discovered as a component of adhesion plaques (Geiger, 1979). Vinculin is a ~116 kDa actin-binding protein comprised of a large, globular head consisting of four α -helix-containing domains (Vd1–Vd4) linked to a tail domain (Vt) by a proline-rich hinge region (Figure 2B).

Vinculin interacts with various proteins involved in integrin-mediated cell-ECM adhesion and cadherin-mediated cell-cell adhesion including α -actinin, vinexin, and ARP2/3 (Wachsstock et al., 1987; Kioka et al., 1999; DeMali et al., 2002). However, its most notable binding partners are talin and filamentous actin (Burridge and Mangeat, 1984). Vinculin also has a similar function at cell-cell adhesions, stabilizing the junctions via interaction between actin and α -catenin (Bays and DeMali, 2017). This linkage is crucial for mechanotransduction (Huvneers and de Rooij, 2013; Zaidel-Bar, 2013; Yao et al., 2014b) and stabilization of cadherin at the cell surface (Peng et al., 2010).

When talin is subjected to force, sequential unfolding of helical bundles of the rod domain occurs which exposes up to 11 cryptic vinculin binding sites (VBS) to which Vd1 can bind (Izard et al., 2004; Papagrigoriou et al., 2004). Filamentous actin, on the other hand, interacts with Vt (Johnson and Craig, 1995).

Autoinhibition of vinculin, like talin, is mediated by a head-tail interaction that conceals the binding surfaces for most known ligands, including for talin and actin (Cohen et al.,

2005, 2006). This interaction occurs via two interfaces: between Vd1 and Vt and between Vd4 and Vt. Although these are individually low-affinity, the combined interaction amounts to a strong autoinhibitory interaction (Cohen et al., 2005). The crystal structure of full-length vinculin in its autoinhibited form (Bakolitsa et al., 2004; Borgon et al., 2004) shows the two head domains interacting with Vt. A constitutively active vinculin termed vinculin T12 has been developed wherein a cluster of four charged residues in Vt are mutated: Asp974Ala, Lys975Ala, Arg976Ala, and Arg978Ala. This mutant has significant loss of affinity to the vinculin head compared to wildtype vinculin (Cohen et al., 2005). More recently, an improved constitutively active vinculin has been developed, the T12K mutant, with Asp974 mutated to a lysine (Asp974Lys) to further destabilize the interaction with the head (Chorev et al., 2018). Constitutively active vinculin markedly reduces adhesion turnover (Humphries et al., 2007; Carisey et al., 2013), and locks talin into an extended conformation (Yao et al., 2014a). Not surprisingly, this loss of adhesion dynamics is lethal in flies (Maartens et al., 2016).

Focal Adhesion Kinase (FAK)

The mechanisms underlying autoinhibition of focal adhesion kinase (FAK) are well-characterized and, being a kinase, a striking example of coupling regulation of protein activity to downstream signaling cascades. FAK is another linear multidomain protein with numerous binding sites for ligands. As the name suggests, FAK contains an enzymatically active kinase domain able to phosphorylate tyrosine residues in target proteins. FAK comprises an N-terminal FERM domain, a kinase domain, a ligand-binding region and a C-terminal 4-helix bundle termed the focal adhesion targeting (FAT) domain (Figure 2C).

At rest, FAK adopts an autoinhibited, closed conformation where the FERM domain directly binds and occludes both the catalytic cleft and the activation loop of the kinase domain (Cooper et al., 2003; Lietha et al., 2007). The crystal structure of autoinhibited FAK allowed identification of the residues on each domain involved in the interaction (Lietha et al., 2007). In this closed conformation, the FAK activation loop is sequestered, preventing autophosphorylation of Tyr397. Interestingly, FERM-mediated autoinhibition of kinase activity has also been observed in the Janus Kinase (JAK) family. JAK proteins have a similar domain structure to FAK and share this common mode of regulation (Zhou et al., 2001).

Two mutants have been designed which significantly increase kinase activity of FAK: a double mutant in F2 of the FERM domain (Tyr180Ala and Met183Ala) and a mutation in the kinase domain (Phe596Asp) (Lietha et al., 2007).

Activation of FAK requires relief of autoinhibition and this can occur through the FAK-FERM domain engaging the integrin tail, leading to release of the activation loop and autophosphorylation. As a result, binding between FAK and the Src tyrosine kinase is enhanced, leading to an active FAK-Src kinase complex able to activate and regulate many other proteins as a major driver of adhesion signaling (Schlaepfer et al., 1999; Zhao and Guan, 2011; Horton et al., 2016).

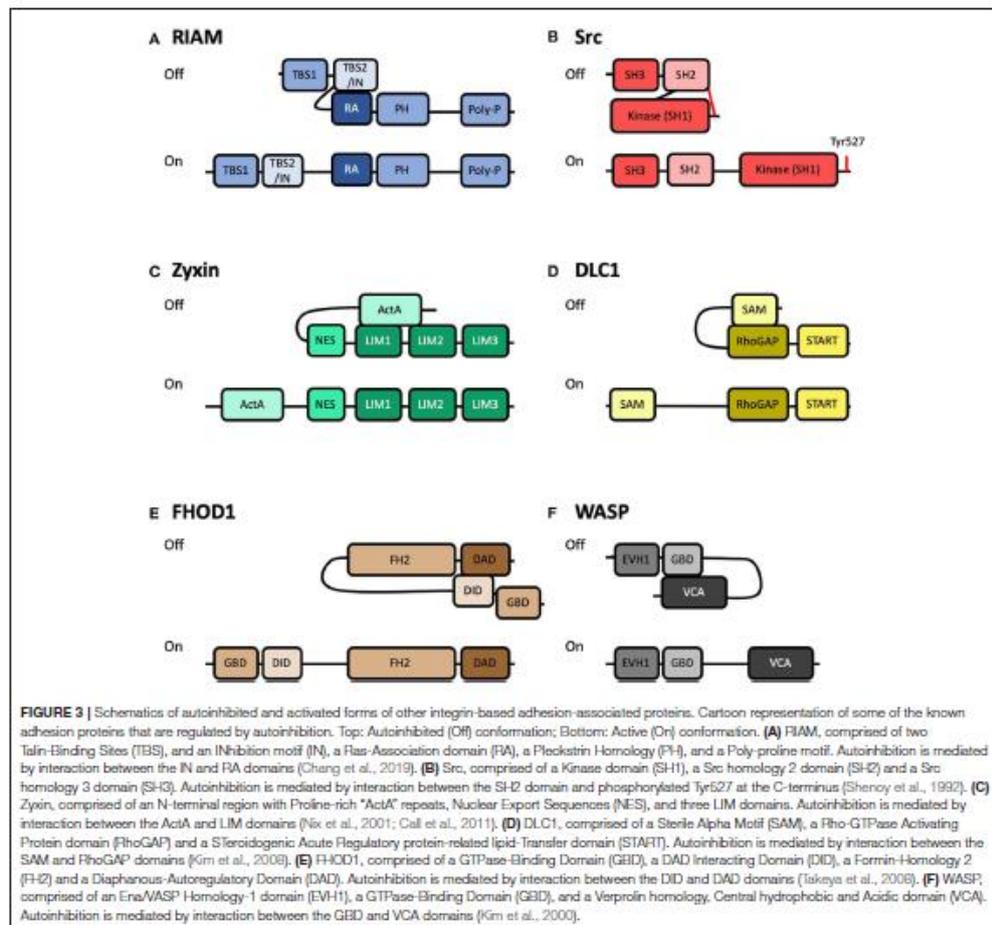
Further, upon release of autoinhibition, FAK can become part of a tethered linkage in force-transmission pathways, tethering to integrins via its FERM domain while the C-terminal region couples to cytoskeletal proteins. This suggests that mechanical force may also contribute to the lifetime of the active form of the protein (see later section: Mechanical regulation of autoinhibition).

Rap1-Interacting Adapter Molecule (RIAM)

An important regulator of the integrin adhesion complex assembly is Rap1-interacting adapter molecule (RIAM). RIAM is a Rap1 effector that can interact directly with talin and translocate it to the plasma membrane, and thus into close proximity to the

integrins (Han et al., 2006; Lee et al., 2013). The direct interaction of RIAM and talin is mediated via the N-terminus of RIAM which contains two talin binding sites (TBS) which interact with four of the talin rod domains (Goult et al., 2013b), with high-affinity binding occurring to the R2 and R3 talin rod domains. RIAM is another linear molecule, comprised of the two TBS and a long flexible linker connecting to Ras association (RA) and pleckstrin homology (PH) domains (Figure 3A).

The structural basis of RIAM autoinhibition was recently shown to be mediated by the region between the two TBS and overlapping with TBS2, termed the inhibitory region (IN) (Figure 3A). This IN region, mapped to residues 27–93, was shown to interact with, and occlude, the Rap1 binding site on



the RA domain (Zhang et al., 2014; Chang et al., 2019) via a switchblade-type autoinhibitory interaction. A tyrosine residue forming part of the binding site, Tyr45, was also shown to be a substrate for focal adhesion kinase (FAK). This suggests a novel regulatory axis whereby FAK-mediated phosphorylation of RIAM relieves RIAM autoinhibition, exposing both the RA and TBS domains, increasing the ability of RIAM to (i) co-localize with Rap1 at the leading edge of cells and (ii) to recruit talin to the same location. A Glu60Ala/Asp63Ala mutation was shown to render RIAM constitutively active (Chang et al., 2019).

This Rap1:RIAM:talin nexus is thus tightly controlled in a myriad of ways, and the dynamic balance of activity status of each protein is implicitly entangled. RIAM is likely activated by active Rap1 (Bos, 2005; Stefanini and Bergmeier, 2016), and the interaction with talin is likely to tilt the equilibrium balance further. The interaction between the PH domain and membrane phospholipids will further tip the equilibrium. The lifetime of the interaction can then be controlled by phosphorylation via FAK and dephosphorylation by the relevant phosphatase. All of these factors will control the activation state of RIAM, highlighting the complex balance of factors that determine the lifetime of interactions.

Src

Relief of autoinhibition of FAK following activation at adhesion sites leads to FAK autophosphorylation of Tyr397 (Frisch, 1996) which provides one mechanism for the recruitment of Src to adhesion complexes. Src, like FAK, is a non-receptor tyrosine kinase that associates with sites of adhesion. The cellular Src protein was originally discovered from its homology to the Rous sarcoma virus oncogene protein product, v-Src (Bishop et al., 1978). Determining the domain structure of Src transformed the field of cell signaling as it is comprised of three modular domains: a kinase domain (Src Homology 1), a phosphotyrosine-binding SH2 domain (Src Homology 2) and a polyproline-recognizing SH3 domain (Src Homology 3). Identification of these modular binding domains in Src (Pawson and Gish, 1992) set a paradigm for cell signaling pathways. Interestingly, Src is maintained in an inactive cytosolic autoinhibited state as the C-terminus of Src contains a tyrosine, Tyr527, which when phosphorylated interacts with Src's own SH2 domain (Shenoy et al., 1992) (Figure 3B). Interestingly, v-Src is constitutively active as it lacks this regulatory C-terminal autoinhibition motif. Phosphorylation of FAK Tyr397 can also enhance Src activation since the Src SH2 domain binds FAK, thus displacing the autoinhibitory tail. Relief of Src autoinhibition can be sustained by phosphatases that dephosphorylate Tyr527 and switch the molecule to its active form.

Src recruitment to, and phosphorylation of, downstream molecules allows many pathways to be regulated through Src, and drives crosstalk between integrin, Src and Rho-family GTPases (Huvencers and Danen, 2009). This provides one route through the complex activation process where a sequence of autoinhibitory interactions are relieved: integrin activation, leading to increased FAK activation, which leads to increased Src activation. All of these enhanced activities can be further modulated by post-translational modification.

And the Rest...

Many other adhesion components have also been shown to autoinhibit. The core adhesion proteins are all long linear molecules so it is possible that they all contain autoinhibitory domains. Table 1 contains a non-exhaustive list of adhesion proteins regulated in this way, and constitutively activating mutations that can be introduced.

The autoinhibited and active conformations of the proteins zyxin, WASP, FHOD1, and DLC1 are shown in Figure 3. These are just a small selection of the 200+ proteins associated with integrin adhesion complexes, but highlights the similarities between the regulatory mechanisms. In each case, the activity of the protein is controlled by a head-tail interaction which is mediated via an autoinhibitory domain binding to, and occluding, a major binding site. Very few proteins in cells are constitutively active.

TABLE 1 | Adhesion proteins regulated by autoinhibition.

Protein	Primary interacting regions	Constitutively active mutant	References
Integrin	Electrostatic interaction between the two cytoplasmic tails	R995A in α 5 or D723A in β 3	Hughes et al., 1996
Talin	F3 and R9	E1770A	Gout et al., 2009
Vinculin	Vd1 and Vd4 and Vt	*T12*: D974A, K975A, R976A, R978A *T12K*: D974K, K975A, R976A, R978A	Cohen et al., 2005 Chorev et al., 2018
RIAM	IN and IA	E60A, D63A	Chang et al., 2019
FAK	F2 and kinase C-lobe	Y397F	Lietz et al., 2007
Src	C-terminus and SH2	Y527F	Shenoy et al., 1992
DLC1	SAM and RhoGAP domains	delta SAM	Kim et al., 2008
α -actinin	CaM-like and neck-F1 domains	NEECK mutant	Young, 2000; Filibeiro et al., 2014
Filamin A	Interactions in its immunoglobulin repeats, inc. I20 and I21	I2144E	Lad et al., 2007
FHOD1	Diaphanous Inhibitory Domain (DIE) and Diaphanous Autoregulatory Domain (DAD)	S1131D, S1137D, T1141D	Takeya et al., 2008
WASP	GED and VCA domains	delta C	Kim et al., 2000
Zyxin	ActA and LIM region	S142D	Nik et al., 2001; Call et al., 2011

CONTROL OF AUTOINHIBITION

Studies into the regulatory complexity of proteins' activation status is revealing adhesions to be highly dynamic. With recent advances in microscopy, it is now possible to observe constant switching between inactive and active conformations of proteins within an adhesion, as exemplified by the mounting evidence that integrins segregate into distinct nanoclusters within focal adhesions (Rossier et al., 2012; Spiess et al., 2018; Orré et al., 2019). Mutants that perturb talin autoinhibition [E1770A Ellis et al., 2013; Haage et al., 2018] or vinculin autoinhibition [T12 Cohen et al., 2005; Carisey et al., 2013] also result in strong attenuation of adhesion dynamics. Together, these results suggest autoinhibition of adhesion components is critical to the tight, dynamic regulation.

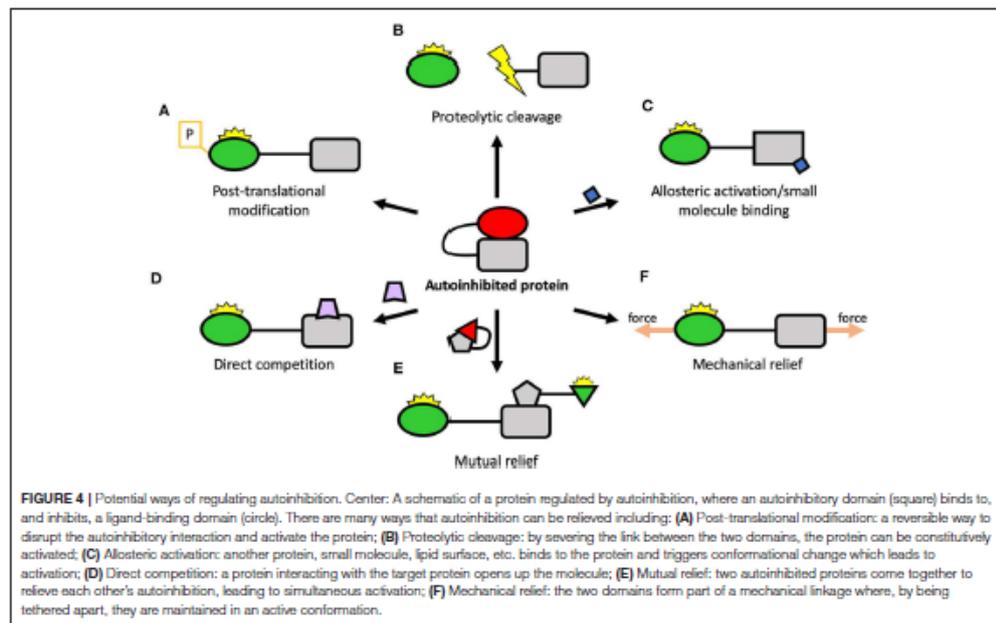
Methods of Relief of Autoinhibition

There are numerous ways that autoinhibition of proteins is controlled with multiple entry points for the regulation of protein activity, some of which are shown in Figure 4. The equilibrium between autoinhibited and active protein is exquisitely poised, such that a subtle shift toward a more active state of one protein can be sufficient to perturb the system and cause large-scale changes to cellular processes. Rather like the "butterfly effect" in chaos theory where a minor change in the initial conditions can lead to large changes in the system overall (Lorenz, 1962), these chaotic events give rise to order in the cellular world.

Upregulation of a protein in response to a signal of some kind, say activation of a GTPase or interaction of two proteins, can trigger rapid amplification and cascades of activation steps that result in large global changes in the adhesion and the cytoskeleton of a cell.

One common mechanism for relieving autoinhibition is **post-translational modifications** (PTMs) (Figure 4A). Here, a kinase or other enzyme chemically alters the target protein and relieves the autoinhibition, switching the protein "on" (or "off" in the case of proteins like Src). PTMs can also stabilize an active state once open and thus control the lifetime of the active state. The most well-studied PTM is phosphorylation but there are many others, including acetylation, methylation (Gunawan et al., 2015), sumoylation (Huang et al., 2018), etc. The converse of this is the removal of PTMs, where a second enzyme such as a phosphatase removes the PTM and reverses the switch.

As well as temporary PTMs, there are also irreversible modifications that can occur, such as **proteolytic cleavage** (Figure 4B). Here, the protein is cleaved, separating into two polypeptides. The best-characterized proteases linked to adhesion belong to the calpain family of calcium-dependent cysteine proteases which cleave many adhesion components. Talin has been shown to be a substrate for calpain (Franco et al., 2004), as have paxillin and FAK (Chan et al., 2010). Calpain cleaves talin in three sites: one site in the neck liberating the head from the rod (Franco et al., 2004), one site at the C-terminus immediately before the dimerization domain (Bate et al., 2012),



and a cryptic third site in the R10 domain (Zhang et al., 2012). Each of these cleavage sites will result in constitutively active talin fragments. Interestingly, whilst the initially-identified role of calpain cleavage was primarily to regulate adhesion disassembly (Franco et al., 2004), recent work has shown that calpain cleavage activation of talin is required at adhesion genesis to trigger proper adhesion formation, recruitment of further non-cleaved talin and adhesion maturation (Saxena et al., 2017), with the cleavage products playing key additional, presumably non-adhesive roles.

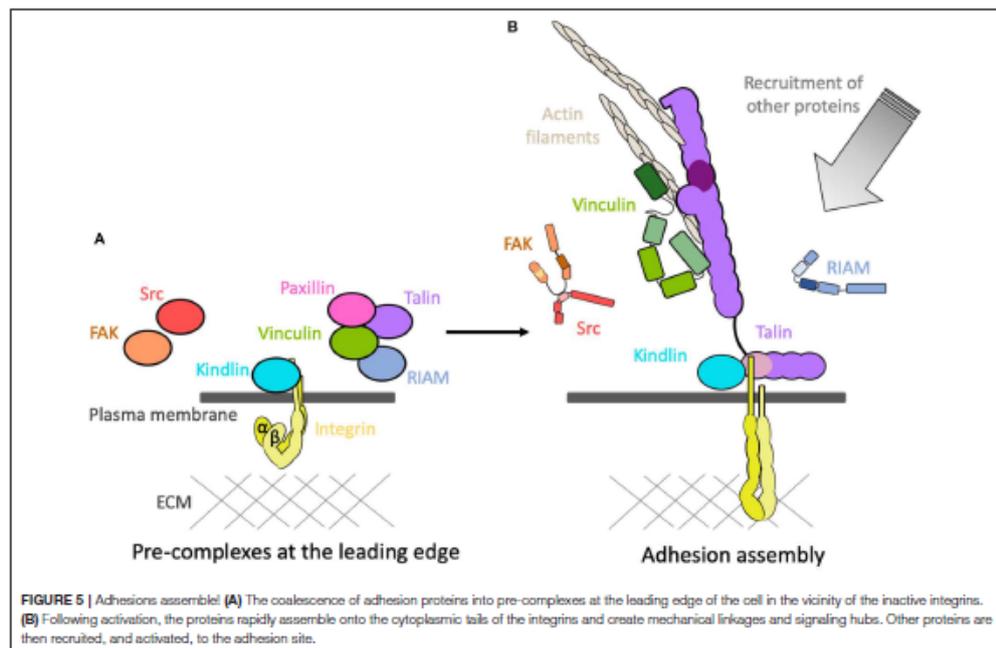
Many proteins are activated by interaction with an activating molecule—this can be a protein, the membrane, etc. Here, the activating molecule can activate the target via **allosteric activation** (Figure 4C) where binding causes conformational changes to the target leading to relief of autoinhibition. This can also be via a **small molecule binding** (Figure 4C): for instance, calpain proteases are activated by influx of calcium and, once active, can activate other proteins by cleavage. Another scenario is that binding of the activating moiety can activate by **direct competition** (Figure 4D) with the autoinhibitory interaction. For instance, talin can be activated by the plasma membrane, where the interaction between the plasma membrane and the F3 domain of talin is stronger than the autoinhibitory interaction holding talin in a closed conformation (Saltel et al., 2009; Song et al., 2012), causing talin to undergo large-scale reorganization upon membrane engagement.

A further scenario that is likely to be playing a major role in adhesion assembly is the **mutual relief of autoinhibition** (Figure 4E). Here, two proteins in an “off” conformation come together and mutually activate each other. Alternatively, one protein becomes activated and, in doing so, is able to activate another, triggering a “domino rally”-style activation cascade.

Mechanical Relief—Mechanotransduction

A common theme of adhesion regulation by autoinhibition is that the proteins are regulated by head-tail autoinhibition. Here, the protein folds in a manner that is mediated by autoinhibitory domains which, when interacting, mask binding sites for ligands. Upon relief of autoinhibition, the protein unfurls to reveal the active form and expose additional binding sites. The protein changes dramatically from a more compact form to an open and extended conformation. The protein will be in equilibrium between the closed and the open states, and the lifetime of each state will be controlled by the factors described in the previous section.

Many proteins in adhesion complexes form tethered linkages where the head and the tail are held apart (Figures 4F, 5B), for instance, vinculin when active forms a linkage between talin and actin. This dramatically reduces the affinity of autoinhibition as the two interacting domains are held apart and less able to interact. In these tethered systems, autoinhibition is a major mechanism enabling mechanotransduction as tethering alters the



dynamics of the protein maintaining it for longer in the active state. This effect can be further enhanced if the tether is under mechanical force: if the protein is stretched, it is even harder for autoinhibition to occur (Wang et al., 2019). Many proteins can adopt multiple conformational states where binding sites are masked via intramolecular interactions and autoinhibition is where the lowest energy state results in suppression of the protein's functional role. As mechanical forces can lower the energy of active states, they can thus relieve autoinhibition.

Autoinhibition in proteins that form part of mechanical linkages in the cell are particularly sensitive to mechanical forces. Whilst force might not be required to open up talin or vinculin from their head-tail autoinhibition, force will play a major role in coordinating the lifetime of the active state (Wang et al., 2019), in essence holding the molecule in an open, extended conformation. As such, the affinity of the protein with its ligands is directly correlated to the force on the system.

MANIPULATION OF AUTOINHIBITION BY MUTATION

Detailed biochemical characterization of the autoinhibitory domains in proteins enables the development of precise mutations that disrupt autoinhibition and render the protein constitutively active. Such mutations (examples of which are shown in Table 1) provide a way of uncoupling proteins from their activation processes and enable the detailed study of systems where the upstream signaling pathways are uncoupled. Such approaches complete a pipeline from biochemical to *in cellulo* to *in vivo* characterization and can enable detailed analysis of the protein in its active form, detached from the complex regulatory elements that control their activities.

CASE STUDY: E1770A—CONSTITUTIVELY ACTIVE TALIN

The autoinhibition of talin is mediated by intramolecular and intermolecular interactions, and the major autoinhibitory domain has been mapped to the R9 of the rod domain binding to the integrin-binding site in F3.

Biochemical characterization of talin autoinhibition showed that the integrin-binding F3 domain interacted with the talin rod (Goksoy et al., 2008). Further refinement showed that the major autoinhibitory interaction was between F3 and the R9 rod domain (Goult et al., 2009), an interaction mediated by a negative surface on R9 which interacts with and occludes the basic activation loop on F3 (Wegener et al., 2007; Song et al., 2012). This structural information enabled identification of a R9 mutation (E1770A) which disrupted the interaction with F3. Analysis of this E1770A mutant in the context of full-length talin in HUVEC cells (Kopp et al., 2010) showed that a constitutively active talin leads to adhesions forming much more rapidly and in significantly larger numbers.

Introduction of the E1770A mutant in *Drosophila* (E1777A in fly talin) led to insight into the loss of essential developmental phenotypes (Ellis et al., 2013). In a mouse model this mutation

results in various defects as a result of more mature and stable focal adhesions, which ultimately results in impeded wound healing (Haage et al., 2018).

This pipeline from *in vitro* to *in cellulo* to *in vivo*, encompassing structural and cellular methods, allows a broader understanding of autoinhibition and the consequences of disrupting such integral regulatory mechanisms. As our understanding of autoinhibition of integrin-mediated adhesion proteins develops, such a multi-disciplinary approach is likely to be an increasingly effective means of investigation.

PRE-COMPLEXES

The mechanically-sensitive interactions that occur following integrin activation have been the subject of intensive research and are reasonably well-understood. However, how all these proteins come to be in the right place and time prior to activation is much less clear. This coalescence of proteins together prior to the assembly of force-dependent linkages would suggest the formation of pre-complexes (Figure 5A). It is likely that autoinhibition maintaining the proteins in an "off" state plays a major role in the formation of these pre-complexes. The conditions prior to formation of mechanical linkages will be very different, suggesting that some of these proteins interact in different force-independent ways. The interactions between the core proteins in the pre-complex state prior to force are likely to be fundamentally different to those that form following activation.

The precise mechanism of how these proteins interact to give rise to pre-complexes is not well-understood. Using "knock-sideways" experiments where proteins are fused with mitochondrial-targeting motifs, pre-complexes between paxillin and autoinhibited talin and vinculin have been observed (Atherton et al., 2019). Furthermore, the use of constitutively active talin and vinculin mutants targeted to mitochondria enables the study of complexes following activation of one component. Fully autoinhibited talin and vinculin do not preassemble on the mitochondria but, by simulating the tethered state of either talin or vinculin using mutation, the release of the head-tail interaction of one was sufficient to trigger interaction of the two proteins. Such technologies are a powerful way to study the most basic of questions—in a cellular context, which proteins bind which, when, and where?

The rapid development of new sophisticated microscopy techniques is allowing the observation of complexes within adhesions to be seen with unprecedented levels of detail (Kanchanawong et al., 2010), facilitating detection of multiple populations of a molecule within an adhesion. For instance, vinculin has at least three distinct states within an adhesion (Case et al., 2015) dictated by interactions with paxillin, talin and actin. Inactive vinculin is initially recruited by paxillin to the plasma membrane at the integrin adhesion "ground zero." Subsequent vinculin activation, promoted by talin, leads to vinculin moving away from the plasma membrane and to a "signaling layer" and ultimately the "force transmission" layer.

Further, evidence for pre-complexation of inactive adhesion proteins comes from the use of fluorescent fluctuation analysis approaches (Bachir et al., 2014). Here, talin and vinculin are seen to associate before the formation of the integrin-talin complex in the nascent adhesion. More recently, new technologies utilizing machine learning trained on high-resolution traction force microscopy, coupled with single particle tracking and fluorescence fluctuation time-series analysis, is enabling the genesis of nascent adhesions to be visualized (Han et al., 2019). This visualization of the earliest stages of adhesion assembly is revealing that talin-vinculin-paxillin pre-complexes are a prerequisite for efficient and meaningful adhesion maturation. Talin and vinculin are required to be together at the moment of force generation to provide efficient maturation—if this is disrupted, maturation is limited. Nascent adhesions where talin and vinculin are recruited at different times fail to mature efficiently. Pre-complex formation appears to be required for efficient assembly of force-bearing linkages and for adhesions to mature following traction force.

It has recently been shown that many of the talin molecules at adhesion sites are non-force bearing, suggesting they are targeted to adhesion sites, ready for action, but are not all simultaneously engaged (Lemke et al., 2019). It will be interesting to explore if this is common to other adhesion molecules and what complexes maintain them at the adhesion site.

Altogether, this amounts to compelling evidence that adhesion proteins form pre-complexes prior to adhesion assembly. Having the key proteins localized together at time zero may be a requirement for rapid and productive adhesion assembly. For example, talin, vinculin and paxillin form a pre-complex (Bachir et al., 2014; Han et al., 2019) prior to association with integrins and prior to mechanical forces, and a second pre-complex of kindlin and integrin has been identified (Rossier and Giannone, 2016). Could these two pre-complexes coming together be sufficient to lead to mutual relief of autoinhibition and trigger adhesion assembly?

The use of FRET-based tension sensors (Austen et al., 2015; Kumar et al., 2016), traction force experiments, microscopy and the rapid advances in artificial intelligence, material science and microfluidics are enabling incredible insight into the interactions of proteins in mature adhesions, in 3D and nascent adhesions as the resolution improves. These technical advances are helping to answer the major open question in the field regarding how these proteins assemble and the interactions that mediate complex formation prior to activation of the integrin.

CONCLUSION AND PHILOSOPHICAL REFLECTIONS

Cells sense the chemical and mechanical properties of their environment through integrins, clustered in punctate adhesion complexes linked to the actin cytoskeleton. The precision and speed with which these adhesions assemble following activation

of an integrin is remarkable and highlights the need for incredible robustness in the process of assembling such huge multiprotein complexes.

The core of most integrin adhesions comprises talin and kindlin bound to and coordinating the activation state of integrin. Once engaged to integrin, the talin rod serves as a mechanosensitive signaling hub (MSH) and, onto this hub, hundreds of proteins are recruited (Goult et al., 2018). Each protein is regulated, many by autoinhibition, and the activation status of each protein provides a regulatable point in the process. A signal leading to activation of one protein (i.e., integrin-ligand interaction outside the cell, or Rap1 activation inside the cell, etc.) might trigger a cascade of activation where autoinhibition of many proteins is simultaneously or sequentially relieved. This enables the system to exact both subtle and sweeping changes to the adhesion and signaling outputs of the adhesion.

Tight regulation and control of adhesion is crucial during development and adult tissue homeostasis. Defects in this tight regulation are implicated in many pathological conditions. Out of the 60 core consensus integrin adhesion proteins, 32 have been shown to be involved in cancer development and progression (Winograd-Katz et al., 2014). Inherited gene mutations in adhesion components are a significant source of disease and disability. Misregulation of the activity of any protein in the network will disrupt the fine balance needed to ensure the exquisitely tuned adhesive and signaling balance. Mutations in adhesion proteins and regulators give rise to diverse malfunctions and diseases as they perturb this balance in different ways which tilt the system in different tissues.

Various intra- and extracellular changes in the environment of the cell trigger subtle readjustments and changes to the adhesive structures enabling appropriate responses. The whole adhesion machinery can be affected from afar by tweaks and changes via signaling pathways that alter the regulation of adhesion proteins. Likewise, external adjustments to the world outside the cell can be propagated into the cell through these adhesions, and subtle change in the lifetime of an interaction, or small reduction in the affinity of the autoinhibition of a protein (such as by increased force extending the lifetime of a linkage), will trigger changes to the system as it reorients to this new homeostasis, adopting an altered adhesion signaling complex which augments the programming of the cell.

AUTHOR CONTRIBUTIONS

BG and RK wrote the manuscript.

ACKNOWLEDGMENTS

We thank David Critchley and members of the Goult laboratory for critical reading of the manuscript. BG was funded by BBSRC grants (BB/N007336/1 and BB/S007245/1), and HFSP grant (RGP00001/2016). RK was funded by a University of Kent studentship.

REFERENCES

- Abercrombie, M. (1961). The bases of the locomotory behaviour of fibroblasts. *Exp. Cell Res.* 8, 188–198. doi: 10.1016/0014-4827(61)90348-2
- Anthis, N. J., and Campbell, I. D. (2011). The tail of integrin activation. *Trends Biochem. Sci.* 36, 191–198. doi: 10.1016/j.tibs.2010.11.002
- Anthis, N. J., Wegener, K. L., Ye, F., Kim, C., Goult, B. T., Lowe, E. D., et al. (2009). The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *Eur. Mol. Biol. Organ. J.* 28, 3623–3632. doi: 10.1038/emboj.2009.287
- Askari, J. A., Buckley, P. A., Mould, A. P., and Humphries, M. J. (2009). Linking integrin conformation to function. *J. Cell Sci.* 122, 165–170. doi: 10.1242/jcs.018556
- Atherton, P., Lausacker, F., Carisey, A., Gilmore, A., Critchley, D., Barsukov, I., et al. (2019). Force-independent interactions of talin and vinculin govern integrin-mediated mechanotransduction. *bioRxiv*. doi: 10.1101/629683
- Atherton, P., Stutchbury, B., Wang, D.-Y., Jethwa, D., Tsang, R., Meiler-Rodriguez, E., et al. (2015). Vinculin controls talin engagement with the actomyosin machinery. *Nat. Commun.* 6, 1–12. doi: 10.1038/ncomms10038
- Austen, K., Ringer, P., Mehlich, A., Chrostek-Grashoff, A., Kluger, C., Klingner, C., et al. (2015). Extracellular rigidity sensing by talin isoform-specific mechanical linkages. *Nat. Cell Biol.* 17, 1597–1606. doi: 10.1038/ncb3268
- Bachir, A. I., Zareno, J., Moissoglu, K., Plow, E. F., Gratton, E., and Horwitz, A. R. (2014). Integrin-associated complexes form hierarchically with variable stoichiometry in nascent adhesions. *Curr. Biol.* 24, 1–9. doi: 10.1016/j.cub.2014.07.011
- Bakolitsa, C., Cohen, D. M., Bankston, L. A., Bobkov, A. A., Dadwell, G. W., Jennings, L., et al. (2004). Structural basis for vinculin activation at sites of cell adhesion. *Nature* 430, 583–586. doi: 10.1038/nature02610
- Banno, A., Goult, B. T., Lee, H., Bate, N., Critchley, D. R., and Ginsberg, M. H. (2012). Subcellular localization of talin is regulated by inter-domain interactions. *J. Biol. Chem.* 287, 13799–812. doi: 10.1074/jbc.M112.341214
- Bate, N., Gingras, A. R., Bachir, A., Horwitz, R., Ye, F., Patel, B., et al. (2012). Talin contains a C-terminal calpain2 cleavage site important in focal adhesion dynamics. *PLoS ONE* 7:e34461. doi: 10.1371/journal.pone.0034461
- Bays, J. L., and DeMali, K. A. (2017). Vinculin in cell–cell and cell–matrix adhesions. *Cell. Mol. Life Sci.* 74, 2999–3009. doi: 10.1007/s00018-017-2511-3
- Bishop, J. M., Baker, B., Fujita, D., McCombe, P., Sheiness, D., Smith, K., et al. (1978). Genesis of a virus-transforming gene. *Natl. Cancer Inst. Monogr.* 48, 219–23.
- Borgon, R. A., Vornheim, C., Bricogne, G., Bois, P. R. J., and Izard, T. (2004). Crystal structure of human vinculin. *Structure* 12, 1189–1197. doi: 10.1016/j.str.2004.05.009
- Bos, J. L. (2005). Linking Rap to cell adhesion. *Curr. Opin. Cell Biol.* 17, 123–128. doi: 10.1016/j.cob.2005.02.009
- Brunet, T., and King, N. (2017). The origin of animal multicellularity and cell differentiation. *Dev. Cell* 43, 124–140. doi: 10.1016/j.devcel.2017.09.016
- Burridge, K., and Connell, L. (1983). A new protein of adhesion plaques and ruffling membranes. *J. Cell Biol.* 97, 359–367. doi: 10.1083/jcb.97.2.359
- Burridge, K., and Mangeat, P. (1984). An interaction between vinculin and talin. *Nature* 308, 744–746. doi: 10.1038/308744a0
- Calderwood, D. A., Campbell, I. D., and Critchley, D. R. (2013). Talins and kindlins: partners in integrin-mediated adhesion. *Nat. Rev. Mol. Cell Biol.* 14, 503–517. doi: 10.1038/nrm3624
- Calderwood, D. A., Zent, R., Grant, R., Rees, D. J., Hynes, R. O., and Ginsberg, M. H. (1999). The Talin head domain binds to integrin beta subunit cytoplasmic tails and regulates integrin activation. *J. Biol. Chem.* 274, 28071–28074. doi: 10.1074/jbc.274.40.28071
- Call, G. S., Chung, J. Y., Davis, J. A., Price, B. D., Primavera, T. S., Thomson, N. C., et al. (2011). Zyxin phosphorylation at serine 142 modulates the zyxin head-tail interaction to alter cell–cell adhesion. *Biochem. Biophys. Res. Commun.* 404, 780–784. doi: 10.1016/j.bbrc.2010.12.058
- Campbell, I. D., and Humphries, M. J. (2011). Integrin structure, activation, and interactions. *Cold Spring Harb. Perspect. Biol.* 3:a004994. doi: 10.1101/cshperspect.a004994
- Carisey, A., Tsang, R., Greiner, A. M., Nijenhuis, N., Heath, N., Nanzgiewicz, A., et al. (2013). Vinculin regulates the recruitment and release of core focal adhesion proteins in a force-dependent manner. *Curr. Biol.* 23, 271–281. doi: 10.1016/j.cub.2013.01.009
- Case, L. B., Baird, M. A., Shtengel, G., Campbell, S. L., Hess, H. F., Davidson, M. W., et al. (2015). Molecular mechanism of vinculin activation and nanoscale spatial organization in focal adhesions. *Nat. Cell Biol.* 17, 880–892. doi: 10.1038/ncb3180
- Chan, K. T., Bennin, D. A., and Huttenlocher, A. (2010). Regulation of adhesion dynamics by calpain-mediated proteolysis of focal adhesion kinase (FAK). *J. Biol. Chem.* 285, 11418–11426. doi: 10.1074/jbc.M109.090746
- Chang, Y.-C. C., Su, W., Cho, E. ah, Zhang, H., Huang, Q., Phillips, M. R., et al. (2019). Molecular basis for autoinhibition of RIAM regulated by FAK in integrin activation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 3524–3529. doi: 10.1073/pnas.1818880116
- Chorev, D. S., Volberg, T., Livne, A., Eisenstein, M., Martins, B., Kam, Z., et al. (2018). Conformational states during vinculin unlocking differentially regulate focal adhesion properties. *Sci. Rep.* 8:2693. doi: 10.1038/s41598-018-21006-8
- Cohen, D. M., Chen, H., Johnson, R. P., Choudhury, B., and Craig, S. W. (2005). Two distinct head-tail interfaces cooperate to suppress activation of vinculin by talin. *J. Biol. Chem.* 280, 17109–17117. doi: 10.1074/jbc.M414704200
- Cohen, D. M., Kutscher, B., Chen, H., Murphy, D. B., and Craig, S. W. (2006). A conformational switch in vinculin drives formation and dynamics of a talin–vinculin complex at focal adhesions. *J. Biol. Chem.* 281, 16006–16015. doi: 10.1074/jbc.M600738200
- Cooper, L. A., Shen, T.-L., and Guan, J.-L. (2003). Regulation of focal adhesion kinase by its amino-terminal domain through an autoinhibitory interaction. *Mol. Cell Biol.* 23, 8030–8041. doi: 10.1128/MCB.23.22.8030-8041.2003
- Dedden, D., Schumacher, S., Kelley, C. F., Zacharias, M., Biertmpfd, C., Fässler, R., et al. (2019). The Architecture of Talin1 reveals an autoinhibition mechanism. *Cell* 179, 120–131.e13. doi: 10.1016/j.cell.2019.08.034
- del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J. M., and Sheetz, M. P. (2009). Stretching single talin rod molecules activates vinculin binding. *Science* 323, 638–641. doi: 10.1126/science.1162912
- DeMali, K. A., Barlow, C. A., and Burridge, K. (2002). Recruitment of the Arp2/3 complex to vinculin: coupling membrane protrusion to matrix adhesion. *J. Cell Biol.* 159, 881–891. doi: 10.1083/jcb.200206043
- Du, X., Plow, E. F., Frelinger, A. L., O'Toole, T. E., Loftus, J. C., and Ginsberg, M. H. (1991). Ligands "activate" integrin α IIb β 3 (platelet GPIIb-IIIa). *Cell* 65, 409–416. doi: 10.1016/0092-8674(91)90458-B
- Ellis, S. J., Goult, B. T., Fairchild, M. J., Harris, N. J., Long, J., Lobo, P., et al. (2013). Talin autoinhibition is required for morphogenesis. *Curr. Biol.* 23, 1825–1833. doi: 10.1016/j.cub.2013.07.054
- Franco, S. J., Rodgers, M. A., Perrin, B. J., Han, J., Bennin, D. A., Critchley, D. R., et al. (2004). Calpain-mediated proteolysis of talin regulates adhesion dynamics. *Nat. Cell Biol.* 6, 977–983. doi: 10.1038/ncb1175
- Frelinger, A. L., Cohen, I., Plow, E. F., Smith, M. A., Roberts, J., Lam, S. C. T., et al. (1990). Selective inhibition of integrin function by antibodies specific for ligand-occupied receptor conformers. *J. Biol. Chem.* 265, 6346–6352.
- Frisch, S. M. (1996). Control of adhesion-dependent cell survival by focal adhesion kinase. *J. Cell Biol.* 134, 793–799. doi: 10.1083/jcb.134.3.793
- Geiger, B. (1979). A 130K protein from chicken gizzard: its localization at the termini of microfilament bundles in cultured chicken cells. *Cell* 18, 193–205. doi: 10.1016/0092-8674(79)90368-4
- Gingras, A. R., Bate, N., Goult, B. T., Hazelwood, L., Canestrelli, I., Grossmann, J. G., et al. (2008). The structure of the C-terminal actin-binding domain of talin. *EMBO J.* 27, 458–469. doi: 10.1038/gj.embj.7601965
- Goksoy, E., Ma, Y.-Q., Wang, X., Kong, X., Perera, D., Plow, E. F., et al. (2008). Structural basis for the autoinhibition of Talin in regulating integrin activation. *Mol. Cell* 31, 124–133. doi: 10.1016/j.molcel.2008.06.011
- Gough, R. E., and Goult, B. T. (2018). The tale of two talins—two isoforms to fine-tune integrin signalling. *FEBS Lett.* 592, 2108–2125. doi: 10.1002/1873-3468.13081
- Goult, B. T., Bate, N., Anthis, N. J., Wegener, K. L., Gingras, A. R., Patel, B., et al. (2009). The structure of an interdomain complex that regulates

- Talin activity. *J. Biol. Chem.* 284, 15097–15106. doi: 10.1074/jbc.M900078200
- Goult, B. T., Xu, X.-P., Gingras, A. R., Swift, M., Patel, B., Bate, N., et al. (2013a). Structural studies on full-length talin1 reveal a compact auto-inhibited dimer: implications for talin activation. *J. Struct. Biol.* 184, 21–32. doi: 10.1016/j.jstb.2013.05.014
- Goult, B. T., Yan, J., and Schwartz, M. A. (2018). Talin as a mechanosensitive signaling hub. *J. Cell Biol.* 217, 3776–3784. doi: 10.1083/jcb.2018.08061
- Goult, B. T., Zacharchenko, T., Bate, N., Tsang, R., Hey, F., Gingras, A. R., et al. (2013b). RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J. Biol. Chem.* 288, 8238–8249. doi: 10.1074/jbc.M112.438119
- Gunawan, M., Venkatesan, N., Loh, J. T., Wang, J. F., Berger, H., Neo, W. H., et al. (2015). The methyltransferase Ezh2 controls cell adhesion and migration through direct methylation of the extranuclear regulatory protein talin. *Nat. Immunol.* 16, 505–516. doi: 10.1038/ni.3125
- Haage, A., Goodwin, K., Whitewood, A., Camp, D., Bogutz, A., Turner, C. T., et al. (2018). Talin autoinhibition regulates cell-ECM adhesion dynamics and wound healing in vivo. *Cell Rep.* 25, 2401–2416.e5. doi: 10.1016/j.celrep.2018.10.098
- Han, J., Lim, C. J., Watanabe, N., Soriani, A., Ratnikov, B., Calderwood, D. A., et al. (2006). Reconstructing and deconstructing agonist-induced activation of integrin α IIb β 3. *Curr. Biol.* 16, 1796–1806. doi: 10.1016/j.cub.2006.08.035
- Han, S. J., Dean, K. M., Whitewood, A. J., Bachir, A., Gutierrez, E., Groisman, A., et al. (2019). Formation of talin-vinculin pre-complexes dictates maturation of nascent adhesions by accelerated force transmission and vinculin recruitment. *bioRxiv* 735183. doi: 10.1101/735183
- Harburger, D. S., and Calderwood, D. A. (2009). Integrin signalling at a glance. *J. Cell Sci.* 122, 1472–1472. doi: 10.1242/jcs.052910
- Hemmings, L., Rees, D. J., Ohanian, V., Bolton, S. J., Gilmore, A. P., Patel, B., et al. (1996). Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *J. Cell Sci.* 109, 2715–2726.
- Horton, E. R., Humphries, J. D., Stutchbury, B., Jacquemet, G., Ballestrem, C., Barry, S. T., et al. (2016). Modulation of FAK and Src adhesion signaling occurs independently of adhesion complex composition. *J. Cell Biol.* 212, 349–364. doi: 10.1083/jcb.201508080
- Huang, Z., Barker, D., Gibbins, J. M., and Dash, P. R. (2018). Talin is a substrate for SUMOylation in migrating cancer cells. *Exp. Cell Res.* 370, 417–425. doi: 10.1016/j.yexcr.2018.07.005
- Hughes, P. E., Diaz-Gonzalez, F., Leong, L., Wu, C., McDonald, J. A., Shattil, S. J., et al. (1996). Breaking the integrin hinge. *J. Biol. Chem.* 271, 6571–6574. doi: 10.1074/jbc.271.12.6571
- Humphries, J. D., Wang, P., Streuli, C., Geiger, B., Humphries, M. J., and Ballestrem, C. (2007). Vinculin controls focal adhesion formation by direct interactions with talin and actin. *J. Cell Biol.* 179, 1043–1057. doi: 10.1083/jcb.200703036
- Huveneers, S., and Danen, E. H. J. (2009). Adhesion signaling—crosstalk between integrins, Src and Rho. *J. Cell Sci.* 122, 1059–1069. doi: 10.1242/jcs.039446
- Huveneers, S., and de Rooij, J. (2013). Mechanosensitive systems at the cadherin-F-actin interface. *J. Cell Sci.* 126, 403–413. doi: 10.1242/jcs.109447
- Hynes, R. O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 69, 11–25. doi: 10.1016/0092-8674(92)90115-5
- Hytönen, V. P., and Vogel, V. (2008). How force might activate talin's vinculin binding sites: SMD reveals a structural mechanism. *PLoS Comput. Biol.* 4:e24. doi: 10.1371/journal.pcbi.0040024
- Izard, T., Evans, G., Borgon, R. A., Rush, C. L., Bricogne, G., and Bois, P. R. J. (2004). Vinculin activation by talin through helical bundle conversion. *Nature* 427, 171–175. doi: 10.1038/nature02281
- Johnson, R. P., and Craig, S. W. (1995). F-actin binding site masked by the intramolecular association of vinculin head and tail domains. *Nature* 373, 261–264. doi: 10.1038/373261a0
- Kanchanawong, P., Shtengel, G., Pasapera, A. M., Ramkio, E. B., Davidson, M. W., Hess, H. F., et al. (2010). Nanoscale architecture of integrin-based cell adhesions. *Nature* 468, 580–584. doi: 10.1038/nature09621
- Kim, A. S., Kakalis, L. T., Abdul-Maman, N., Liu, G. A., and Rosen, M. K. (2000). Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* 404, 151–158. doi: 10.1038/35004513
- Kim, C., Lau, T.-L., Ulmer, T. S., and Ginsberg, M. H. (2009). Interactions of platelet integrin α IIb and β 3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood* 113, 4747–4753. doi: 10.1182/blood-2008-10-186551
- Kim, M. (2003). Bidirectional transmembrane signaling by cytoplasmic domain separation in integrins. *Science* 301, 1720–1725. doi: 10.1126/science.1084174
- Kim, T. Y., Healy, K. D., Der, C. J., Scialy, N., Bang, Y.-J., and Juliano, R. L. (2008). Effects of structure of Rho GTPase-activating protein DLC-1 on cell morphology and migration. *J. Biol. Chem.* 283, 32762–32770. doi: 10.1074/jbc.M800617200
- Kioka, N., Sakata, S., Kawachi, T., Amachi, T., Akiyama, S. K., Okazaki, K., et al. (1999). Vinxin: a novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organization. *J. Cell Biol.* 144, 58–69. doi: 10.1083/jcb.144.1.59
- Kopp, P. M., Bate, N., Hansen, T. M., Brindle, N. P. J., Praskelt, U., Debrand, E., et al. (2010). Studies on the morphology and spreading of human endothelial cells define key inter- and intramolecular interactions for talin1. *Eur. J. Cell Biol.* 89, 661–673. doi: 10.1016/j.ejcb.2010.05.003
- Kumar, A., Ouyang, M., K. V. den D., McGhee, E. J., Tanaka, K., Anderson, M. D., et al. (2016). Talin tension sensor reveals novel features of focal adhesion force transmission and mechanosensitivity. *J. Cell Biol.* 213, 371–383. doi: 10.1083/jcb.201510012
- Lad, Y., Kiema, T., Jiang, P., Penttinen, O. T., Coles, C. H., Campbell, I. D., et al. (2007). Structure of three tandem filamin domains reveals auto-inhibition of ligand binding. *EMBO J.* 26, 3993–4004. doi: 10.1038/emboj.7601827
- Lagarigue, F., Vilas Ankal, P., Lee, H.-S., Bachir, A. L., Ablack, J. N., Horwitz, A. F., et al. (2015). A RIAM/lamellipodin-talin-integrin complex forms the tip of sticky fingers that guide cell migration. *Nat. Commun.* 6:8492. doi: 10.1038/ncomms9492
- Lau, T. L., Kim, C., Ginsberg, M. H., and Ulmer, T. S. (2009). The structure of the integrin α IIb β 3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* 28, 1351–1361. doi: 10.1038/emboj.7601827
- Lee, H.-S., Ankal, P., Lim, C. J., Liu, C.-Y., and Ginsberg, M. H. (2013). Two modes of integrin activation form a binary molecular switch in adhesion maturation. *Mol. Biol. Cell* 24, 1354–1362. doi: 10.1091/mbc.12-09-0695
- Lee, H.-S., Lim, C. J., Puzon-McLaughlin, W., Shattil, S. J., and Ginsberg, M. H. (2009). RIAM activates integrins by linking Talin to Ras GTPase membrane-targeting sequences. *J. Biol. Chem.* 284, 5119–5127. doi: 10.1074/jbc.M807117200
- Lenke, S. B., Weidemann, T., Cost, A.-L. L., Grashoff, C., and Schnorrer, F. (2019). A small proportion of Talin molecules transmit forces at developing muscle attachments in vivo. *PLoS Biol.* 17:e3000057. doi: 10.1371/journal.pbio.3000057
- Li, J., Su, Y., Xia, W., Qin, Y., Humphries, M. J., Vestweber, D., et al. (2017). Conformational equilibria and intrinsic affinities define integrin activation. *EMBO J.* 36, 629–645. doi: 10.15252/emboj.201695803
- Lietha, D., Cai, X., Ceccarelli, D. F. J., Li, Y., Schaller, M. D., and Eck, M. J. (2007). Structural basis for the autoinhibition of focal adhesion kinase. *Cell* 129, 1177–1187. doi: 10.1016/j.cell.2007.05.041
- Lorenz, E. N. (1962). Deterministic nonperiodic flow. *J. Atmos. Sci.* 20, 130–141.
- Luo, B. H., and Springer, T. A. (2006). Integrin structures and conformational signaling. *Curr. Opin. Cell Biol.* 18, 579–586. doi: 10.1016/j.cob.2006.08.005
- Maartens, A. P., Wellmann, J., Wittome, E., Klapholz, B., Green, H., and Brown, N. H. (2016). *Drosophila* vinculin is more harmful when hyperactive than absent, and can circumvent integrin to form adhesion complexes. *J. Cell Sci.* 129, 4354–4365. doi: 10.1242/jcs.189878
- McCann, R. O., and Craig, S. W. (1997). The I/LWEQ module: a conserved sequence that signifies F-actin binding in functionally diverse proteins from yeast to mammals. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5679–84. doi: 10.1073/pnas.94.11.5679
- Molony, L., McCaslin, D., Abernethy, J., Paschal, B., and Burridge, K. (1987). Properties of talin from chicken gizzard smooth muscle. *J. Biol. Chem.* 262, 7790–7795.
- Moser, M., Nieswandt, B., Ussar, S., Pozgajova, M., and Fässler, R. (2008). Kindlin-3 is essential for integrin activation and platelet aggregation. *Nat. Med.* 14, 325–330. doi: 10.1038/nm1722
- Mould, A. P., Akiyama, S. K., and Humphries, M. J. (1996). The inhibitory anti- β 1 integrin monoclonal antibody 13 recognizes an epitope that

- is attenuated by ligand occupancy. *J. Biol. Chem.* 271, 20365–20374. doi: 10.1074/jbc.271.34.20365
- Nix, D. A., Fradelizi, J., Bockholt, S., Menichi, B., Louvard, D., Friederich, E., et al. (2001). Targeting of zyxin to sites of actin membrane interaction and to the nucleus. *J. Biol. Chem.* 276, 34759–34767. doi: 10.1074/jbc.M102820200
- Orré, T., Rossier, O., and Giannone, G. (2019). The inner life of integrin adhesion sites: from single molecules to functional macromolecular complexes. *Exp. Cell Res.* 379, 235–244. doi: 10.1016/j.yexcr.2019.03.036
- Papagrigoriou, E., Gingras, A. R., Barsakov, I. L., Bate, N., Fillingham, I. J., Patel, B., et al. (2004). Activation of a vinculin-binding site in the talin rod involves rearrangement of a five-helix bundle. *EMBO J.* 23, 2942–2951. doi: 10.1038/sj.emboj.7600285
- Pawson, T., and Gish, G. D. (1992). SH2 and SH3 domains: from structure to function. *Cell* 71, 359–362. doi: 10.1016/0092-8674(92)90504-6
- Peng, X., Cuff, L. E., Lawton, C. D., and DeMali, K. A. (2010). Vinculin regulates cell-surface E-cadherin expression by binding to β -catenin. *J. Cell Sci.* 123, 567–577. doi: 10.1242/jcs.056432
- Pufall, M. A., and Graves, B. J. (2002). Autoinhibitory domains: modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* 18, 421–462. doi: 10.1146/annurev.cellbio.18.031502.133614
- Qin, J., Vinogradova, O., and Plow, E. F. (2004). Integrin bidirectional signaling: a molecular view. *PLoS Biol.* 2:e169. doi: 10.1371/journal.pbio.0020169
- Ribeiro, E. D. A., Pinotiss, N., Ghisleni, A., Salmazo, A., Konarev, P. V., Kostan, J., et al. (2014). The structure and regulation of human muscle α -Actinin. *Cell* 159, 1447–1460. doi: 10.1016/j.cell.2014.10.056
- Ringer, P., Weifl, A., Cost, A.-L., Freikamp, A., Sabass, B., Mehlich, A., et al. (2017). Multiplexing molecular tension sensors reveals piconewton force gradient across talin-1. *Nat. Methods* 14, 1090–1096. doi: 10.1038/nmeth.4431
- Rogalski, T. M., Mullen, G. P., Gilbert, M. M., Williams, B. D., and Moerman, D. G. (2000). The UNC-112 gene in *Caenorhabditis elegans* encodes a novel component of cell-matrix adhesion structures required for integrin localization in the muscle cell membrane. *J. Cell Biol.* 150, 253–264. doi: 10.1083/jcb.150.1.253
- Rossier, O., and Giannone, G. (2016). The journey of integrins and partners in a complex interactions landscape studied by super-resolution microscopy and single protein tracking. *Exp. Cell Res.* 343, 28–34. doi: 10.1016/j.yexcr.2015.11.004
- Rossier, O., Octeau, V., Sibarita, J.-B., Leduc, C., Tessier, B., Nair, D., et al. (2012). Integrins $\beta 1$ and $\beta 3$ exhibit distinct dynamic nanoscale organizations inside focal adhesions. *Nat. Cell Biol.* 14, 1057–1067. doi: 10.1038/ncb2588
- Saltel, F., Mortier, E., Hytönen, V. P., Jacquier, M.-C., Zimmermann, P., Vogel, V., et al. (2009). New PI(4,5)P₂- and membrane proximal integrin-binding motifs in the talin head control beta3-integrin clustering. *J. Cell Biol.* 187, 715–731. doi: 10.1083/jcb.200908134
- Saxena, M., Chagde, R., Hone, J. C., Wolfenson, H., and Sheetz, M. P. (2017). Force-induced calpain cleavage of Talin is critical for growth, adhesion development, and rigidity sensing. *Nano Lett.* 17, 7242–7251. doi: 10.1021/acs.nanolett.7b02476
- Schlaepfer, D. D., Hauck, C. R., and Sieg, D. J. (1999). Signaling through focal adhesion kinase. *Prog. Biophys. Mol. Biol.* 71, 435–478. doi: 10.1016/S0079-6107(98)00052-2
- Sebé-Pedrós, A., Roger, A. J., Lang, F. B., King, N., and Ruiz-Trillo, I. (2010). Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10142–10147. doi: 10.1073/pnas.1002257107
- Senetar, M. A., and McCann, R. O. (2005). Gene duplication and functional divergence during evolution of the cytoskeletal linker protein talin. *Gene* 362, 141–152. doi: 10.1016/j.gene.2005.08.012
- Shatil, S. J., Kim, C., and Ginsberg, M. H. (2010). The final steps of integrin activation: the end game. *Nat. Rev. Mol. Cell Biol.* 11, 288–300. doi: 10.1038/nrm2871
- Shenoy, S., Chackalampal, I., Bagrodia, S., Lin, P. H., and Shalloway, D. (1992). Role of p34cdc2-mediated phosphorylations in two-step activation of pp60^{src} during mitosis. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7237–7241. doi: 10.1073/pnas.89.15.7237
- Shimaoka, M., Takagi, J., and Springer, T. A. (2002). Conformational regulation of integrin structure and function. *Annu. Rev. Biophys. Biomol. Struct.* 31, 485–516. doi: 10.1146/annurev.biophys.31.101101.140922
- Song, X., Yang, J., Hirbawi, J., Ye, S., Perera, H. D., Goksoy, E., et al. (2012). A novel membrane-dependent on/off switch mechanism of talin FERM domain at sites of cell adhesion. *Cell Res.* 22, 1533–1545. doi: 10.1038/cr.2012.97
- Spies, M., Hernandez-Varas, P., Oddone, A., Olofsson, H., Blom, H., Waihe, D., et al. (2018). Active and inactive $\beta 1$ integrins segregate into distinct nanoclusters in focal adhesions. *J. Cell Biol.* 217, 1929–1940. doi: 10.1083/jcb.201707075
- Srivastava, J., Barreiro, G., Grosscurth, S., Gingras, A. R., Goult, B. T., Critchley, D. R., et al. (2008). Structural model and functional significance of pH-dependent talin-actin binding for focal adhesion remodeling. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14436–14441. doi: 10.1073/pnas.0805163105
- Stefanini, L., and Bergmeier, W. (2016). RAP1-GTPase signaling and platelet function. *J. Mol. Med.* 94, 13–19. doi: 10.1007/s00109-015-1346-3
- Sun, Z., Guo, S. S., and Päsler, R. (2016). Integrin-mediated mechanotransduction. *J. Cell Biol.* 215, 445–456. doi: 10.1083/jcb.201609037
- Takeya, R., Taniguchi, K., Narumiya, S., and Sumimoto, H. (2008). The mammalian formin FHOD1 is activated through phosphorylation by ROCK and mediates thrombin-induced stress fibre formation in endothelial cells. *EMBO J.* 27, 618–628. doi: 10.1038/emboj.2008.7
- Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horwitz, A. F., et al. (1986). Structure of fibronectin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell* 46, 271–282. doi: 10.1016/0092-8674(86)90744-0
- Vinogradova, O., Velyvis, A., Velyviene, A., Hu, B., Haas, T. A., Plow, E. F., et al. (2002). A structural mechanism of integrin $\alpha 1 \beta 3$ “inside-out” activation as regulated by its cytoplasmic face. *Cell* 110, 587–597. doi: 10.1016/S0092-8674(02)00906-6
- Wachsstock, D. H., Wilkins, J. A., and Lin, S. (1987). Specific interaction of vinculin with α -actinin. *Biochem. Biophys. Res. Commun.* 146, 554–560. doi: 10.1016/0006-291X(87)90564-X
- Wang, Y., Yan, J., and Goult, B. T. (2019). Force-dependent binding constants. *Biochemistry* 58, 4696–4709. doi: 10.1021/acs.biochem.9b00453
- Wegener, K. L., Partridge, A. W., Han, J., Pickford, A. R., Liddington, R. C., Ginsberg, M. H., et al. (2007). Structural basis of integrin activation by Talin. *Cell* 128, 171–182. doi: 10.1016/j.cell.2006.10.048
- Winograd-Katz, S. E., Päsler, R., Geiger, B., and Legate, K. R. (2014). The integrin adhesome: from genes and proteins to human disease. *Nat. Rev. Mol. Cell Biol.* 15, 273–88. doi: 10.1038/nrm3769
- Yao, M., Goult, B. T., Chen, H., Cong, P., Sheetz, M. P., and Yan, J. (2014a). Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Sci. Rep.* 4:4610. doi: 10.1038/srep04610
- Yao, M., Goult, B. T., Klapholz, B., Hu, X., Tosseland, C. P., Guo, Y., et al. (2016). The mechanical response of talin. *Nat. Commun.* 7:11966. doi: 10.1038/ncomms11966
- Yao, M., Qiu, W., Liu, R., Efremov, A. K., Cong, P., Seddiki, R., et al. (2014b). Force-dependent conformational switch of α -catenin controls vinculin binding. *Nat. Commun.* 5:4525. doi: 10.1038/ncomms5525
- Ye, F., Kim, C., and Ginsberg, M. H. (2012). Reconstruction of integrin activation. *Blood* 119, 26–33. doi: 10.1182/blood-2011-04-292128
- Young, P. (2000). The interaction of titin and alpha-actinin is controlled by a phospholipid-regulated intramolecular pseudoligand mechanism. *EMBO J.* 19, 6331–6340. doi: 10.1093/emboj/19.23.6331
- Yutin, N., Makarova, K. S., Mekhedov, S. L., Wolf, Y. I., and Koonin, E. V. (2008). The deep archaeal roots of eukaryotes. *Mol. Biol. Evol.* 25, 1619–1630. doi: 10.1093/molbev/mn108
- Zaidi-Bar, R. (2013). Cadherin adhesome at a glance. *J. Cell Sci.* 126, 373–378. doi: 10.1242/jcs.111559
- Zhang, F., Saha, S., and Kashina, A. (2012). Arginylation-dependent regulation of a proteolytic product of talin is essential for cell-cell adhesion. *J. Cell Biol.* 197, 819–836. doi: 10.1083/jcb.201112129
- Zhang, H., Chang, Y.-C., Brennan, M. L., and Wu, J. (2014). The structure of Rap1 in complex with RIAM reveals specificity determinants and

- recruitment mechanism. *J. Mol. Cell Biol.* 6, 128–139. doi: 10.1093/jmcb/mjt044
- Zhao, X., and Guan, J. L. (2011). Focal adhesion kinase and its signaling pathways in cell migration and angiogenesis. *Adv. Drug Deliv. Rev.* 63, 610–615. doi: 10.1016/j.addr.2010.11.001
- Zhou, Y.-J., Chen, M., Cusack, N. A., Kimmel, L. H., Magnuson, K. S., Boyd, J. G., et al. (2001). Unexpected effects of FERM domain mutations on catalytic activity of Jak3: structural implication for Janus kinases. *Mol. Cell* 8, 959–969. doi: 10.1016/S1097-2765(01)00398-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Khan and Goult. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix 2:



Chapter 9

Biochemical Characterization of the Integrin Interactome

Rejina B. Khan, Lorena Varela, Alana R. Cowell, and Benjamin T. Goult

Abstract

More than 250 proteins are associated with the formation of integrin adhesion complexes involving a vast number of complex interactions between them. These interactions enable adhesions to serve as dynamic and diverse mechanosignaling centers. Our laboratory focuses on the biochemical and structural study of these interactions to help unpick this complex network. Here, we describe the general pipeline of biochemical assays and methods we use. The chapter is split into two sections: (1) protein production and characterization and (2) biochemical assays for the characterization of binding between full-length proteins and/or specific regions of proteins with other proteins, peptides, and phospholipids. The suite of assays we use routinely includes circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy for sample quality assessment, prior to biochemical analysis using NMR, fluorescence polarization (FP), microscale thermophoresis (MST), size-exclusion chromatography multiangle light scattering (SEC-MALS), and pulldown/cosedimentation-based approaches. The results of our analysis feed into *in vivo* studies that allow for the elucidation of the biological role of each interaction.

Key words Integrin, Adhesome, Interaction, Biochemistry

1 Introduction

Cell–matrix adhesions are large dynamic macromolecule complexes that assemble on the cytoplasmic face of the integrin receptors. Over 250 different proteins have been detected in these integrin adhesion complexes (IACs) and new components are continually being discovered. Implicit in the formation of these large IACs are many interactions that together enable adhesions to serve as dynamic and diverse mechanosignaling centers. The goal of our laboratory is to understand these interactions and disrupt each interaction in turn in order to define approaches that enable us to unpick this complexity with exquisite detail.

Biochemical assays are analytical, *in vitro* procedures that are used to characterize biological molecules and their interactions.

Rejina B. Khan and Lorena Varela contributed equally.

Miguel Vicente-Manzanares (ed.), *The Integrin Interactome: Methods and Protocols*, Methods in Molecular Biology, vol. 2217, https://doi.org/10.1007/978-1-0716-0962-0_9, © Springer Science+Business Media, LLC, part of Springer Nature 2021

There are many biochemical assays that have been developed and each assay has its own advantages and disadvantages. The aim of this chapter is not to cover them all but instead to describe the biochemical pipeline we use in our laboratory for expression, purification and biochemical and biophysical investigation of integrin adhesion complex-associated proteins. The chapter starts with a brief overview of our common methods of protein production, including cloning and mutagenesis to generate the expression plasmids for recombinant protein expression. The generation of high-quality protein samples is key to biochemical studies. Circular dichroism (CD) and nuclear magnetic resonance (NMR) are powerful approaches for assessing protein quality and ensuring correct folding. We then introduce our “go-to” suite of biochemical assays for characterization of protein interactions which include CD, NMR, fluorescence polarization (FP), microscale thermophoresis (MST), size-exclusion chromatography multi-angle light scattering (SEC-MALS) and the pulldown/cosedimentation-based approaches including GST-pulldowns, actin cosedimentation and lipid cosedimentation.

We use these methods to characterize binding between full-length proteins and/or specific regions of proteins. One of the ultimate goals is to identify the specific residues involved in mediating interactions so as to design mutations that abolish binding. Our work goes hand-in-hand with *in vivo* studies; biochemical insight allows mapping of specific sites of interest which can then be tested in cells or animal models to validate the findings. Utilizing mutations which disrupt the interactions for *in vivo* studies allows for observation of a phenotype, ultimately allowing identification of the role for that specific interaction.

2 Materials

2.1 Molecular Cloning

2.1.1 Restriction Enzyme Method

1. Thermocycler.
2. Template DNA which contains the insert of interest.
3. dNTPs.
4. Oligonucleotides (primers) designed for your insert.
5. Polymerase (e.g., Promega GoTaq[®] DNA Polymerase) and associated buffer.
6. DNA clean up kit (e.g., QIAGEN QIAquick Gel Extraction Kit).
7. Vector DNA.
8. Restriction enzymes (based upon the designed oligonucleotides) and appropriate reaction buffer.
9. 1% agarose gels for gel electrophoresis analysis and gel extraction.

10. Appropriate ligase for your vector of choice (e.g., Promega T4 DNA Ligase) and associated buffer.
11. *E. coli* DH5 α /DH10 β or other appropriate bacterial strain.
12. Plasmid DNA miniprep kit (e.g., QIAGEN QIAprep Spin Miniprep Kit).
13. Appropriate sequencing oligonucleotides.

2.1.2 Gibson Assembly

1. Thermocycler.
2. Template DNA of vector and of insert.
3. dNTPs.
4. Oligonucleotides (primers).
5. Polymerase (e.g., Promega Pfu[®] DNA Polymerase) and associated buffer.
6. DNA clean up kit (e.g., QIAGEN QIAquick Gel Extraction Kit).
7. Gibson Master Mix (e.g., NEB Master Mix).
8. *E. coli* DH5 α /DH10 β or other appropriate bacterial strain.
9. Plasmid DNA miniprep kit (e.g., QIAGEN QIAprep Spin Miniprep Kit).
10. Appropriate sequencing oligonucleotides.

2.1.3 Site-Directed Mutagenesis

1. Thermocycler.
2. Template DNA.
3. dNTPs.
4. Oligonucleotides (primers).
5. DNA polymerase (e.g., Promega Pfu[®] DNA Polymerase) and associated buffer.
6. *DpnI* endonuclease (e.g., NEB *DpnI*).
7. DNA clean up kit (e.g., QIAGEN QIAquick Gel Extraction Kit).
8. *E. coli* DH5 α /DH10 β or other appropriate bacterial strain.
9. Plasmid DNA miniprep kit (e.g., QIAGEN QIAprep Spin Miniprep Kit).
10. Appropriate sequencing oligonucleotides.

2.2 Protein Expression

2.2.1 Buffers

1. LB media (per liter) 5 g tryptone, 10 g yeast extract, and 5 g sodium chloride in 1 L. distilled/deionized water. The final pH should be 7.0–7.5. Autoclave.
2. 2M9 minimal media (per liter) is comprised of two parts, solution A and B. Solution A: 12.5 g/L Na₂HPO₄·2H₂O, 7.5 g/L KH₂PO₄. The final pH should be 7.0–7.5. Autoclave.

118 Rejina B. Khan et al.

Solution B: 4.0 g glucose (D-Glucose- $^{13}\text{C}_6$ (>99% ^{13}C) for carbon labeling), 10.0 mL BME Vitamins (e.g., Sigma BioReagent BME vitamins 100 \times), 2.0 mL MgSO_4 (from autoclaved 1 M stock), 0.1 mL CaCl_2 (from autoclaved 1 M stock), 1.0 g NH_4Cl (^{15}N -labeled NH_4Cl for nitrogen labeling) diluted in 10.0 mL water. Use a 0.2 μm filter to sterilize and add to autoclaved solution A.

2.2.2 General Equipment/Reagents

1. Appropriate antibiotic(s): e.g., 100 mg/mL ampicillin and 50 mg/mL kanamycin for 1000 \times stock solutions.
2. Refrigerated shaking incubator (for 37 $^\circ\text{C}$ and 18 $^\circ\text{C}$ incubation).
3. His-tagged protein cloned into expression vector of choice (e.g., pET-151 TOPO) or GST-tagged protein cloned into expression vector of choice (e.g., pGEX-TEV).
4. *E. coli* BL21(DE3) or other appropriate bacterial strain.
5. Centrifuge.
6. Buffer in which to resuspend the harvested cell pellet (resuspend the pellet in the buffer of the first purification step, for example, for His-tagged protein resuspend pellet in Nickel buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole).
7. Protease inhibitor (e.g., Roche Complete[®] Protease Inhibitor Tablets).

2.3 Protein Purification

All buffers should be filtered and degassed.

2.3.1 Buffers

Protein Purification
for His-Tagged Proteins
Using Immobilized
Nickel-Affinity
Chromatography

1. HisTrap HP column, 5 mL (GE Healthcare).
2. Nickel buffer A: 20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole.
3. Nickel buffer B: 20 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazole.

Protein Purification by
Batch Method
for His-Tagged Proteins

1. Agarose beads charged with nickel (e.g., HisPur[™] Ni-NTA Superflow agarose).
2. Nickel buffer A: 20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole.
3. Wash buffer: 20 mM Tris pH 8.0, 500 mM NaCl, 50 mM imidazole.
4. Nickel buffer B: 20 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazole.

- | | |
|--|--|
| Protein Purification by Batch Method for GST-Tagged Proteins | <ol style="list-style-type: none"> 1. Glutathione beads (e.g., Pierce™ Glutathione Superflow agarose). 2. PBS: 137 mM NaCl, 27 mM KCl, 100 mM Na₂HPO₄, 18 mM KH₂PO₄, pH 7.4. |
| Ion-Exchange Chromatography | <ol style="list-style-type: none"> 1. Anion exchange (Q HP) column, 5 mL (GE Healthcare). 2. Q buffer A (20 mM Tris pH 8.0, 50 mM NaCl). 3. Q buffer B (20 mM Tris pH 8.0, 1 M NaCl). 4. Cation exchange (SP HP) column, 5 mL (GE Healthcare). 5. S buffer A (20 mM phosphate pH 6.5, 50 mM NaCl). 6. S buffer B (20 mM phosphate pH 6.5, 1 M NaCl). |
| 2.3.2 General Equipment/Apparatus | <ol style="list-style-type: none"> 1. Sonicator, French press, or cell disruptor. 2. Centrifuge (for low speed 2831 × <i>g</i> (4000 rpm in a Beckman JA-10 rotor), and 48,384 × <i>g</i> (20,000 rpm in a Beckman JA-25.5 rotor) at 4 °C). 3. Dialysis tubing (e.g., SnakeSkin™, ThermoFisher) with appropriate molecular weight cut off (usually 3, or 10 kDa MWCO) or PD-10 desalting column (GE Healthcare). 4. SDS-PAGE gels at a percentage appropriate to visualize proteins of interest. 5. ÄKTA™ purification system (GE Healthcare). 6. Spectrophotometer (e.g., Implen NanoPhotometer). |
| 2.4 Peptides | <ol style="list-style-type: none"> 1. Peptide resuspended in water or PBS to an appropriate concentration. 2. Thiol-reactive fluorescent dye (e.g., fluorescein-5-maleimide or BODIPY™ TMR C5-Maleimide, both from ThermoFisher), solubilized in DMSO according to manufacturer's instructions. 3. TCEP (tris(2-carboxyethyl)phosphine). 4. Triton X-100. 5. PD-10 desalting column (GE Healthcare). |
| 2.5 Biochemical Assays | <ol style="list-style-type: none"> 1. JASCO J-715 spectropolarimeter with JASCO power supply PS-150J equipped with JASCO Peltier thermoelectric cooler PTC-423S/15. 2. Quartz cuvette. 3. Protein (0.2–0.5 mg/mL). |
| 2.5.1 Circular Dichroism (CD) | |
| 2.5.2 Nuclear Magnetic Resonance (NMR) | <ol style="list-style-type: none"> 1. NMR spectrometer (e.g., a 600 MHz five channel Bruker Avance III spectrometer with a CryoProbe). |

120 Rejina B. Khan et al.

2. Unlabeled or isotopically labeled protein (150 μ M concentration is ideal for HSQC titrations).
3. Deuterated water.
4. NMR tubes: 3 or 5 mm tubes (Wilmad) or 5 mm susceptibility matched tube and plunger (Shigemi).
5. NMR software CcpNmr Analysis for spectra analysis.

2.5.3 Fluorescence Polarization (FP)

1. Black 96-well plate (e.g., Nunc™ F96 MicroWell™ Polystyrene Plate from ThermoFisher).
2. Microplate reader with appropriate filters for fluorophore (e.g., BMG CLARIOstar equipped with polarizing filters).
3. Fluorescently labeled peptide (*see* “coupling peptides”).
4. Protein at appropriate concentration (determined by affinity of the interaction).
5. PBS (100 mM Na_2HPO_4 , 18 mM KH_2PO_4 , pH 7.4, 137 mM NaCl, 27 mM KCl).
6. Curve fitting software (e.g., GraphPad Prism).

2.5.4 Microscale Thermophoresis (MST)

1. Microscale thermophoresis instrument (e.g., NanoTemper Monolith NT.115).
2. Monolith His-tag labeling kit (e.g., RED-Tris-NTA—this kit includes His-tag dye, buffers, capillaries and tubes).
3. His-tagged protein (target).
4. Ligand to study interaction with (typical ligands include protein without His-tag, peptide, nucleotide, macromolecule, lipid vesicle, or metal ion).
5. Benchtop centrifuge (up to 16,200 $\times g$), refrigerated.
6. MO.Control software for analysis.

2.5.5 Size-Exclusion Chromatography-Multiangle Light Scattering (SEC-MALS)

1. ÄKTA™ purification system (GE Healthcare).
2. Appropriate SEC columns (e.g., Superdex 200, Sephacryl 200, or Sephacryl 300).
3. MALS system (e.g., Malvern MALS system Viscotek SEC-MALS 9 with Modular RI detector Viscotek VE 3580).
4. OmniSEC software for analysis.

2.5.6 GST-Pulldowns

1. GST-tagged protein bound to Glutathione beads (e.g., Pierce™ Glutathione Superflow agarose).
2. SDS-PAGE gels at a percentage appropriate to visualize your proteins (e.g., NuPAGE™ 4–12% Bis-Tris Protein Gels).
3. ImageJ software for analysis.

**2.5.7 Actin
Cosedimentation Assay**

1. Polymerized F-actin at a stock concentration above the critical concentration (we use ~100 μ M).
2. Actin co-sed buffer (10 mM Tris pH 7.0, 2 mM MgCl₂, 0.2 mM ATP, 1 mM DTT, 50 mM NaCl).
3. Ultracentrifuge (up to 100,000 $\times g$, for example a Beckman TLA-100 rotor) for high-speed actin-binding assay.
4. Benchtop centrifuge (up to 16,200 $\times g$) for low-speed actin-bundling assay.
5. SDS-PAGE gels at a percentage appropriate to visualize your proteins (e.g., NuPAGE™ 4–12% Bis-Tris Protein Gels).
6. ImageJ software for analysis.

**2.5.8 Lipid
Cosedimentation Assay**

1. Chloroform for dissolving lipids to make lipid films.
2. Nitrogen cylinder.
3. Lipids: phosphatidylcholine (PC), phosphatidylserine (PS) from Sigma-Aldrich and Phosphatidylinositol 4,5-bisphosphate (PIP2) from Avanti Polar Lipids (Alabaster, AL).
4. Protein buffer (20 mM Tris pH 7.4, 0.1 mM EDTA, 15 mM β -mercaptoethanol).
5. Lipid co-sed buffer (20 mM HEPES pH 7.4, 0.2 mM EGTA).
6. Benchtop centrifuge (up to 16,200 $\times g$).
7. SDS-PAGE gels at a percentage appropriate to visualize your proteins (e.g., NuPAGE™ 4–12% Bis-Tris Protein Gels).
8. ImageJ software for analysis.

3 Methods

**3.1 Molecular
Cloning**

For biochemical and structural analysis, we predominantly use two expression vectors: (1) pET-151, which contains a 6 \times His-tag followed by a TEV cleavage site, and the base vector contains a multiple cloning site with BamHI, NotI, EcoRI, and XhoI restriction sites; (2) pGEX-TEV (a modified version of pGEX-4T), which contains a glutathione S-transferase (GST)-tag followed by a TEV cleavage site, and the base vector contains a broader multiple cloning site with BamHI, NheI, EcoRI, SalI, XhoI, and NotI restriction sites. Both of these vectors encode ampicillin resistance. The cloning methods described here are applicable to any vector but are written in this chapter as used for pET-151 and pGEX-TEV specifically.

**3.1.1 Restriction Enzyme
Digest and Ligation**

Restriction enzyme digest followed by ligation of the generated sticky ends is a classic approach to molecular cloning. We use this

method primarily for subcloning regions of interest, often from mammalian vectors or genomic DNA, into our expression vectors listed above. Here, we describe our general approach for this process and the reagents we use.

1. Design forward and reverse oligonucleotides using the sequence of the insert. Each oligonucleotide should: consist of ~15–20 bases from your insert, introduce required restriction sites, contain ~60–70% guanine/cytosine content, and terminate with a guanine or cytosine base at the 5' end.
2. Perform PCR according to manufacturer's instructions for your polymerase of choice and purify the product. To purify, use ethanol precipitation method or commercial kits such as QIAGEN QIAquick Gel Extraction Kit.
3. Carry out a restriction digest, usually at 37 °C for 2–4 h (check manufacturer's instructions for your restriction enzymes of choice), for your purified PCR product and for your vector of choice. Each reaction should be performed with identical/compatible restriction enzymes to generate compatible sticky ends. Use agarose gel extraction to purify the digested products.
4. Set up a ligation reaction. For pET-151 and pGEX-TEV vectors, T4 DNA ligase is a suitable enzyme. Follow manufacturer's instructions for your enzyme of choice.
5. Transform the ligation reaction into DH5 α /DH10 β *E. coli*. Perform minipreps for individual colonies and perform a test digest, if suitable.
6. Sequence the DNA to completion using appropriate sequencing oligonucleotides to ensure that the insert is present.

3.1.2 Gibson Assembly

Gibson assembly methodology is a flexible cloning strategy developed by Daniel G. Gibson, of the J. Craig Venter Institute [1]. This is a robust exonuclease-based method to assemble DNA seamlessly and in the correct order. The reaction is carried out under isothermal conditions using three enzymatic activities: a 5' exonuclease generates long overhangs, a polymerase fills in the gaps of the annealed single strand regions, and a DNA ligase seals the nicks of the annealed strands and fills in the gaps. If deletions or insertions in the same plasmid are required, it is only necessary to design two primers (forward and reverse) that introduce the desired variation in the original template DNA. If a change of plasmid is required, two sets of forward and reverse primers are required, one set for the plasmid and another set for the insert DNA fragments, to introduce the desired variations. A high-fidelity DNA polymerase and a thermocycler are required to perform separate PCR reactions to amplify the different noncircularized fragments with overlapping ends.

Table 1
PCR program used for Gibson assembly and for site-directed mutagenesis

Process	Time	Temperature (°C)	Cycles
Initial DNA denaturation	30 s	95	1
DNA denaturation	30 s	95	30
Primers annealing	30 s	$T = T_m - 5$	
DNA chain extension	Depends on polymerase used and plasmid or fragment size	Depends on polymerase used	
DNA elongation completion	5 min	Depends on polymerase used	1
Storage	End	4	1

1. Design the required primers. First, define the desired final assembled product and then select primer sequences for both sides of the joints with sufficient binding to both the templates and the overlapping region of the other fragment to allow the Gibson assembly reaction to proceed. Starting with either fragment, select a region of sequence that starts at the joint with a T_m of around 60 °C, making sure to include a G/C anchor at the 5' end. This is the binding region of the primer. Next, add the overlapping region by selecting some bp of the other fragment. In general, an overlap of 40 bp yields a sufficient T_m for the Gibson reaction, so if the primer is extended by 20 bp that will give 40 bp of overlap. This process needs to be repeated for each joint.
2. Perform PCR reactions to amplify each fragment following manufacturer's instructions for your polymerase of choice. The PCR program we normally use is described in Table 1.
3. Purify the PCR products using ethanol precipitation method or commercial kits (we use QIAGEN QIAquick Gel Extraction Kit).
4. Perform the Gibson assembly reaction by adding the Gibson Master Mix (we use the NEB Gibson Assembly Master Mix) to the recommended concentrations of vector and insert as described in the manufacturer's instructions. Perform the reaction in a thermocycler at 50 °C for 15 min.
5. Transform the now circularized vector into high insert stability competent cells (we use DH10β *E. coli*).
6. Perform minipreps for individual colonies and sequence the DNA to completion using appropriate sequencing oligonucleotides to confirm the correct plasmid assembly. Once checked, insert the plasmid into BL21(DE3) *E. coli* cells to overexpress the proteins.

124 Rejina B. Khan et al.

3.1.3 Site-Directed Mutagenesis

Site-directed mutagenesis (SDM) is used to make changes to DNA to modify the protein of interest. Such modifications we introduce include targeted mutation of key residues to disrupt binding, disease-causing mutations, small insertions or small deletions, introduction of stop codons to make shorter proteins, etc. SDM uses PCR with overlapping forward and reverse primers designed such that they introduce the desired variations into the sequence. A high-fidelity DNA polymerase and a thermocycler are required to amplify the mutated DNA plasmid.

1. Design forward and reverse oligonucleotides using the sequence of your insert.
2. The primers should be approximately 30 bp in length with the mutated site as close to the center as possible, with a minimum of 12 bp either side.
3. Perform PCR according to manufacturer's instructions for your polymerase of choice and purify the product. The PCR program used is described in Table 1.
4. Following temperature cycling, treat the product with *DpnI*. The *DpnI* endonuclease is used to digest the parental DNA template and to select for mutation-containing synthesized DNA. Purify the PCR products using ethanol precipitation method or commercial kits (we use QIAGEN QIAquick Gel Extraction Kit).
5. Transform the vector containing the desired mutations into high insert stability competent cells (we use DH10 β *E. coli*).
6. Perform minipreps for individual colonies and sequence the DNA to completion using appropriate sequencing oligonucleotides to confirm the presence of the mutation. Once checked, transform the plasmids into BL21 (DE3) *E. coli* cells to overexpress the proteins.

3.2 Protein Expression

Here, we describe the most common expression conditions for the talin constructs that we work with. These conditions give good expression for many proteins and are a good starting point. However, it is desirable to optimize the conditions for each protein to attain maximum yields. Expression trials at different temperatures, different IPTG concentrations and different induction times will enable the identification of optimal conditions for expression. For proteins that cannot be expressed in bacteria, or require chaperones, cofactors, or posttranslation modification, it can be necessary to use insect or mammalian cell expression systems, but these are not discussed here.

3.2.1 Unlabeled Protein Expression

Unlabeled proteins are expressed recombinantly in bacterial cultures using standard LB media. The required antibiotic will be dependent on the plasmid and strain used. Our usual strategy

uses the pET-151 vectors transformed into BL21 (DE3)⁺ *E. coli* which requires only ampicillin.

1. Transform BL21(DE3)⁺ *E. coli* with the plasmid encoding the protein of interest and streak onto LB agar plate containing the appropriate antibiotic(s).
2. Inoculate 10 mL of LB media containing the appropriate antibiotics with BL21 (DE3)⁺ *E. coli* culture containing the desired protein construct. Grow this starter culture overnight at 37 °C in a rotary shaker.
3. Use 5 mL of the overnight culture to inoculate 1 L of LB with the appropriate antibiotic and continue to grow the culture at 37 °C to an approximate OD₆₀₀ of 0.6–0.8. At this point, induce the cell culture with 1 mM IPTG for 3 h at 37 °C or 200 μM IPTG for overnight expression at 18 °C.
4. Harvest the bacteria by centrifugation at 2831 × *g* (4000 rpm in a Beckman JA-10 rotor) at 4 °C for 10 min and resuspend them in 30 mL of Nickel buffer A on ice. Add protease inhibitor tablet. At this point, the resuspended pellet can be frozen for further purification in the future or it can be immediately further purified.

3.2.2 Isotopically Labeled Protein Expression

For multidimensional NMR experiments, proteins need to be labeled with ¹⁵N and/or ¹³C. For isotopic labeling, 2M9 minimal media needs to be used.

1. Inoculate 10 mL of 2M9 minimal media containing the appropriate antibiotic(s) with BL21(DE3)⁺ *E. coli* culture containing the desired protein construct. Grow this starter culture overnight at 37 °C in a rotary shaker.
2. Use 5 mL of the overnight culture to inoculate 1 L of 2M9 minimal media with appropriate antibiotic and continue to grow the culture at 37 °C to an OD₆₀₀ of 0.6–0.8. At this point, induce the cell cultures with IPTG (200 μM final concentration) for 15–20 h at 18 °C.
3. Harvest the bacteria by centrifugation, as previously described, and resuspend in 30 mL of Nickel buffer A on ice. Add protease inhibitor tablet. At this point, the resuspended pellet can be frozen for further purification in the future or it can be immediately further purified.

3.2.3 Condensation Method of Isotopically Labeled Protein Expression

In order to boost the rate of cell growth and simultaneously minimize the required cost (by reducing the amount of isotopic labeling), the 4:1 condensation method can be used [2]. Here, the cells are grown to high density in LB media before being transferred into the labeled media for expression.

126 Rejina B. Khan et al.

1. Inoculate LB media containing the appropriate antibiotic (four times the volume of the 2M9 minimal media that will be used later) with 5 mL of starter culture per L.
2. Grow cells at 37 °C to an OD₆₀₀ of 0.6–0.8.
3. Next, wash the cells to remove the rest of LB media by centrifugation (3000 × *g*, 5 min at 4 °C) and resuspend them in 200 mL of 2M9 solution A and repeating the centrifugation.
4. Resuspend the cells in 2M9 minimal media final desired volume (solution A + solution B) with appropriate antibiotic(s) and incubate for 1 h at 18 °C to allow them to adapt to the new growth conditions. At this point, induce protein expression with 200 μM IPTG at 18 °C for 15–20 h.
5. Harvest the bacteria by centrifugation, as previously described, and resuspend in 30 mL of Nickel buffer A on ice. Add protease inhibitor tablet. At this point, the resuspended pellet can be frozen for further purification in the future or it can be immediately further purified.

3.3 Protein Purification

There are many approaches and variations of methodology for purifying recombinant proteins. Here we describe our common strategies which are optimized for speed and purified protein quality. Collect samples for SDS-PAGE analysis at each step of the purification for monitoring the purification process.

1. If resuspended bacterial pellet is frozen, thaw it at room temperature. Next, lyse the cells by sonication for 5 cycles of 30 s on followed by 30 s off at ~35% amplitude. Ensure to keep the sample on ice throughout the sonicating process and thereafter.
2. Centrifuge the lysed cells at ~48,000 × *g* at 4 °C for 30 min. Discard the pelleted cell membranes. Retain the supernatant and filter using a 0.45 μm filter to remove any remaining cell debris. Prepare a sample of the supernatant and pellet for SDS-PAGE analysis—this will allow validation that the protein of interest is soluble.

3.3.1 Protein Purification for His-Tagged Proteins Using Immobilized Nickel-Affinity Chromatography

1. Equilibrate a 5 mL HisTrap HP (*see Note 1*) column using Nickel buffer A and Nickel buffer B. At a flowrate of ~4 mL/min, wash through with buffer: 3 column volumes (cv) of 100% A followed by 3 cv of 100% B and finally 6 cv of 100% A.
2. Ensure that the peristaltic sample loading pump has been equilibrated in Nickel buffer A. Load the filtered supernatant onto the column at a flowrate of ~2.5 mL/min.
3. Following sample application, use the sample pump to wash through ~20 mL of Nickel buffer A. This ensures that the entirety of the sample is loaded onto the column and reequilibrates the sample loading pump for the next run.

4. Wash out unbound protein by washing the column with 5–8 cv of Nickel buffer A at 5 mL/min.
5. Use gradient elution at 5 mL/min—this strategy increases the amount of Nickel buffer B (high imidazole buffer) in a linear gradient. Set the starting %B concentration to 0% and the target %B concentration to 100% over 50 mL. Collect 1–5 mL fractions.
6. Finish by reequilibrating the column and the system using 5 cv of Nickel buffer A at 5 mL/min. The next sample can be loaded at this point.
7. Once complete, store the column and the system in 20% ethanol by flowing at 1 mL/min for 2 cv.
8. Prepare samples for SDS-PAGE analysis for fractions corresponding to a peak to confirm presence of protein of interest and to assess purity. Pool suitable fractions.

*3.3.2 Protein Purification
for His-Tagged Proteins by
Batch Method*

Purification of His-tagged proteins can also be performed by the batch method, a method useful for purification of proteins which are more sensitive or susceptible to aggregation/degradation. The batch method involves the use of centrifugation to separate the supernatant from the nickel-NTA beads to allow purification. This method is an effective alternative if purifying multiple proteins in parallel and/or there is limited access to an ÄKTA™ purification system. Collect samples for SDS-PAGE analysis at each step of the purification for monitoring the purification process.

1. Use ~1 mL Ni-NTA agarose slurry (e.g., Qiagen HisPur™ Ni-NTA Superflow) per liter of bacterial culture and wash the beads by adding 2 × 30 mL Nickel buffer A and mixing gently. Centrifuge at 700 × *g* (2510 rpm in a Hettich Rotanta 460 R centrifuge) for 2 min at 4 °C. Discard supernatant.
2. Add the filtered cell lysate to the washed beads and agitate at room temperature for 30 min. For some proteins, particularly those which are more unstable, it may be preferable to perform this step at 4 °C.
3. Centrifuge at 700 × *g* (2510 rpm in a Hettich Rotanta 460 R centrifuge) for 2 min at 4 °C. Remove supernatant.
4. Wash the beads with 40 mL Nickel buffer A and mix gently. Centrifuge at 700 × *g* (2510 rpm in a Hettich Rotanta 460 R centrifuge) for 2 min at 4 °C. Remove supernatant. Repeat this wash step 5–6 times total.
5. Perform the final wash step using 40 mL of Wash buffer.
6. After the final wash step, resuspend the beads in 5 mL of buffer and pour into an empty gravity flow column (Bio-Rad).

128 Rejina B. Khan et al.

7. Elute the His-tagged protein with Nickel buffer B in 1 mL fractions. Some proteins precipitate in high imidazole, so it may be preferable to dilute the eluted protein fractions with Nickel buffer A immediately following elution.
8. Prepare a sample of each fraction for SDS-PAGE analysis to confirm presence of protein of interest and to assess purity. Pool suitable fractions.

Following either method for purification of His-tagged proteins, the pooled protein sample should be 90–95% pure. If the His-tag is to be cleaved, add TEV protease at this stage. Exchange the pooled fractions (including TEV protease, if applicable) into an appropriate buffer for cleavage. If ion exchange chromatography will be performed, use dialysis to exchange the sample into S buffer A or Q buffer A, depending on whether cation or anion exchange chromatography are to be performed.

3.3.3 Protein Purification by Batch Method for GST-Tagged Proteins

1. Take ~1 mL glutathione agarose slurry (e.g., Pierce™ Glutathione Superflow) and wash the beads by adding 20 mL PBS and mixing gently. Centrifuge at $700 \times g$ for 3 min at 4 °C. Remove supernatant.
2. Add the filtered cell lysate to the washed beads and agitate at room temperature for 1–2 h. For some proteins, particularly those which are more unstable, it may be more suitable to perform this step at 4 °C.
3. Centrifuge at $700 \times g$ for 3 min at 4 °C. Remove supernatant.
4. Wash the beads with 30 mL PBS and mix gently. Centrifuge at $700 \times g$ for 3 min at 4 °C. Remove supernatant. Repeat this wash step 5–6 times.
5. After final wash step, remove the supernatant and resuspend the beads in a suitable buffer. Prepare a sample of this bead-bound protein slurry for SDS-PAGE analysis to ensure that the protein of interest is present. The protein is now ready to be used for GST-pulldown experiments, if applicable.
6. If the protein is to be cleaved from its GST-tag and thus the beads, add TEV protease and agitate overnight at room temperature or 4 °C as required.
7. Transfer to a column and collect the flow through. This fraction will contain the protein of interest, now cleaved from the GST-tag. Confirm this by preparing a sample for SDS-PAGE analysis.

3.3.4 Ion-Exchange Chromatography

We use ion-exchange chromatography to remove TEV protease and cleaved His-tags from our purified protein. Even if TEV cleavage is not performed, this is a useful step to increase the purity of the final protein sample generated. Use the isoelectric point of your

protein of interest to deduce whether a Q (anion exchange) or S (cation exchange) column is required. With the buffers listed, an isoelectric point of ~7 and below is generally suitable for a Q column; an isoelectric point above ~8 is suitable for a S column. For any proteins with an isoelectric point within/close to this range, the pH of the buffers could be altered (ensuring to remain within the buffering capacity) or the HisTrap HP column could be used again instead. If the latter is used, the cleaved protein of interest will be in the column flow through, separated from the TEV protease and cleaved His-tag which will both instead bind to the column.

1. Wash either a 5 mL HiTrap Q HP or a 5 mL HiTrap SP HP column (both from GE Healthcare, *see Note 2*) using the appropriate buffers, that is (1) Q or S buffer A and (2) Q or S buffer B. At a flowrate of ~4 mL/min, wash through with buffer: 3 cv of 100% A followed by 3 cv of 100% B and finally 6 cv of 100% A. This allows removal of any contaminants and then prepares the column for use.
2. Ensure that the sample pump has been equilibrated in Q/S buffer A. At a flowrate of ~2.5 mL/min, load the filtered sample onto the column.
3. Following sample application, use the sample pump to load ~20 mL of Q/S buffer A. This ensures that the entirety of the sample loaded has reached the column.
4. Wash out unbound protein by washing column with 5–8 column volumes of Q/S buffer A at 5 mL/min.
5. Use gradient elution at 5 mL/min. Set start %B concentration at 0 and target %B concentration to 100. Fractionate using settings of your choice (Fig. 1a). Stepwise elution can also be used.
6. Finish by reequilibrating the column using 4 column volumes of Q/S buffer A at 5 mL/min. The next sample can be loaded at this point and the process repeated.
7. Once complete, store the column and the system in 20% ethanol by flowing at 1 mL/min for 2 cv.
8. Prepare samples for SDS-PAGE analysis of fractions corresponding to a peak (Fig. 1b). Once fractions containing protein of interest and minimal contaminants have been identified, pool these together.
9. Use a PD-10 desalting column or dialysis to exchange the purified protein into a suitable buffer for the experiments to be performed. Measure the concentration using the absorbance at 280 nm (we use Implen NanoPhotometer to measure this) and the molar extinction coefficient (calculated using

130 Rejina B. Khan et al.

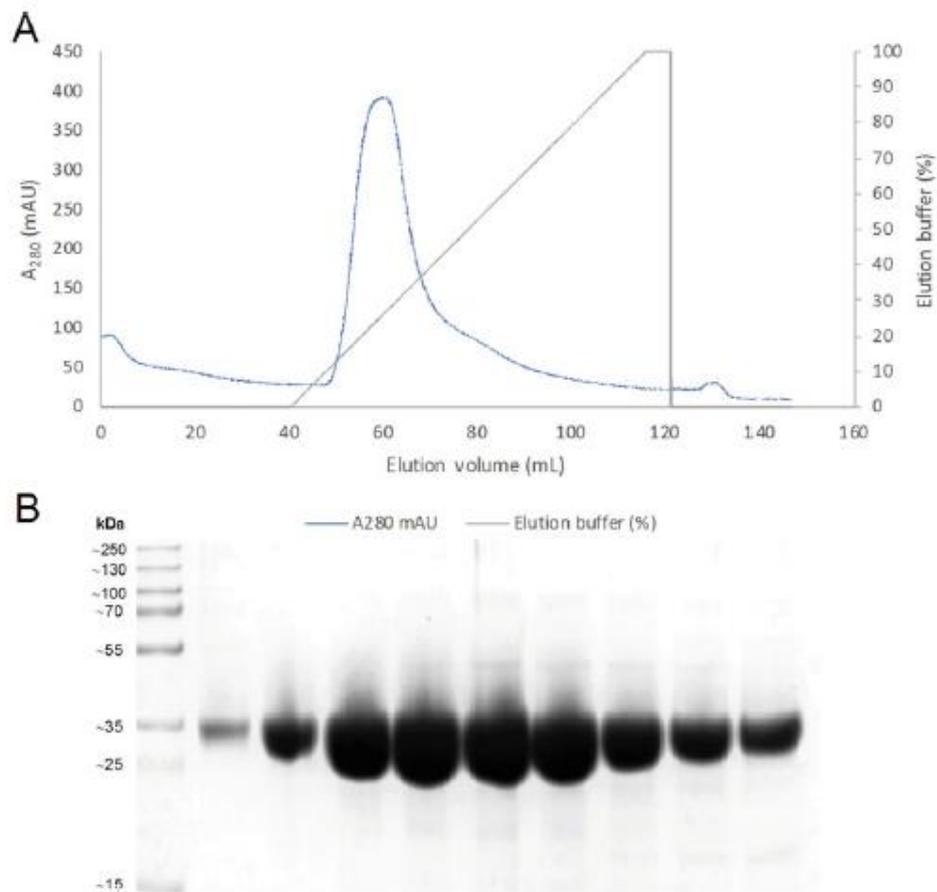


Fig. 1 (a) Example elution profile using anion-exchange chromatography (Q column) of talin1 R7-R8 G1404L mutant. The peak in A280 absorbance (blue) corresponds to the elution of the protein of interest at a given concentration of elution buffer (gray). (b) SDS-PAGE analysis of elution fractions corresponding to the peak in A280 absorbance

ProtParam <https://web.expasy.org/protparam/>) and concentrate or dilute as required.

10. Aliquot and flash freeze the final purified protein using liquid nitrogen. Note: the effects of flash freezing may vary between proteins—while most proteins we have worked with function identically before freezing and after thawing, some proteins are susceptible to precipitation, degradation or differences in activity.

3.4 Peptides

3.4.1 Synthetic Peptide Design

There are a number of factors to consider when designing synthetic peptides to use for biochemical and biophysical experiments. These will vary depending for each experiment, but we typically design peptides as summarized briefly here.

1. Identify the region of the protein to synthesize as a peptide. We use a combination of deletion mapping to identify the region of the protein involved in the interaction, secondary structure prediction (using PSIPRED [3]), amino acid conservation (regions of functional importance are highly conserved across species), and motif identification (i.e., if it is a vinculin binding site, or an LD-motif then both of these have well defined consensus sequences).
2. If the identified region is part of a secondary structure element such as a helix then it is important that the peptide spans the entire helix and has a couple of nonhelical residues either side to enable the secondary structure to form.
3. The solubility of each peptide is dependent on the specific sequence and the net charge of the peptide. These factors can be calculated using various online servers for predicting peptide properties. If poor solubility is predicted then these can be dissolved in DMSO, but it can also be useful to adjust the sequence to increase the percentage of charged residues.
4. If fluorescence polarization experiments will be carried out (or any experiments where coupling to a thiol-reactive moiety will be required), the peptide design should include a terminal cysteine residue. The choice of N- or C-terminus can depend on whether the putative interacting region is located close to the termini. If so, add the terminal cysteine residue to the opposite end.
5. If the target peptide already contains a cysteine in its sequence then maleimide coupling is not possible. Ordering peptides precoupled to the required dye may be more suitable.

Ordering synthetic peptides up to ~35 amino acids is reasonably cost effective. Above 10 mg, the cost tends to get prohibitive and it might be better to consider making them recombinantly. We recommend to always order the peptide in 2 mg aliquots.

3.4.2 Coupling Peptides

For fluorescence polarization experiments (*see* following section), synthetic peptides must be coupled to a fluorescent dye. While the peptides can be synthesized with these, we generally couple them in-house. We use fluorescein-5-maleimide or BODIPY™ TMR C5-Maleimide diluted according to manufacturer's instructions. An important consideration is to choose the thiol-reactive fluorophore so that it is compatible with the equipped polarizing filters in your plate reader.

132 Rejina B. Khan et al.

1. Mix 100 μM peptide (resuspended in water or PBS), 0.1% Triton X-100, 5 mM TCEP, and 25 μL of fluorescent dye. Make up to 1 mL using PBS.
2. Protect from light and leave stirring at room temperature for 2 h.
3. Using a PD-10 desalting column equilibrated with PBS, remove excess dye. Elute using 2.5 mL PBS and collect the entirety of the flow through. The coupled peptide will elute after \sim 1 mL.
4. Prepare 50 μL aliquots, flash-freeze using liquid nitrogen and store at -20°C until needed.

3.5 Biochemical Assays

3.5.1 Circular Dichroism (CD)

CD is an excellent tool for rapid determination of the secondary structure and folding and binding properties of proteins. CD also provides insight early on in the project to the behavior of the protein of interest including its thermal stability. This method is a critical control test for validating point mutations and ensuring they do not have off-target effects such as protein destabilization. CD works by measuring the wavelength dependence of the differential absorption of the left and right-handed circularly polarized light. The left- and right-handed components of a polarized beam of light interact differently with the chiral centers of an optically active chromophore (present in the amides of the polypeptide backbone of proteins).

Far-UV CD spectroscopy has high sensitivity to changes in the different secondary structure elements in proteins including α -helices, β -sheets, β -turns and random coils, whereas near-UV CD spectroscopy gives information about changes in the tertiary structure of the protein. Melting curves can also be acquired by measuring the change of CD signal at a fixed wavelength over a temperature gradient. Figure 2 shows the typical melting curve data set obtained using this method. In this experiment (taken from ref. 4), the melting curve of the talin rod domain 3 (R3) was measured for the wild type (WT) and a stabilizing mutant (IVVI).

3.5.2 Measurement of Far-UV CD Spectra

1. Prepare a protein sample at 20–50 μM .
2. Place the sample in a quartz cuvette using an appropriate path length depending on the concentration of the sample.
3. Collect the far-UV spectra between 260 and 190 nm. We normally average 4–8 scans at 100 nm/min, 0.5 nm step resolution, 1 s of response and 0.5 nm of band width (*see Note 3*).

3.5.3 Measurements of Melting Curves

1. To monitor the time course of unfolding, prepare a protein sample at 20–50 μM .

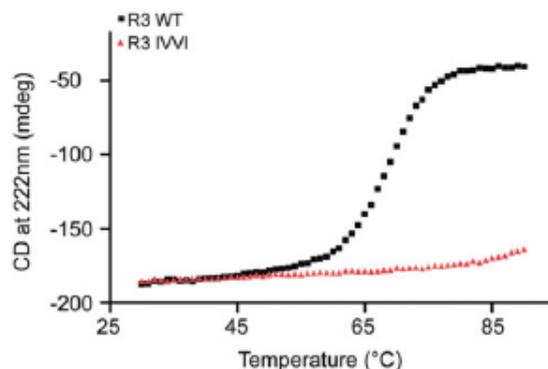


Fig. 2 Example of a dataset collected using the CD assay. The melting curve for wildtype talin R3 domain (black) and a stabilizing IVI mutant (red) were measured as described in the text by monitoring the change in CD at 222 nm with increasing temperature. The melting temperature (T_m) is the point halfway between the transition of the folded and unfolded state; for WT it is 69.6 °C, the IVI mutant >90 °C. (Reproduced from ref. 4 under a Creative Commons CC-BY-NC license)

2. Place the sample in the quartz cuvette and ensure a lid is in place to avoid evaporation of the sample.
3. Set the temperature ramp while keeping a constant wavelength. We normally measure the CD signal at 222 nm (maximal signal for an α -helix) as a function of temperature between 25 and 90 °C with 20 s step resolution, 4 s of response and 1.0 nm of band width.
4. Far-UV CD spectra can be collected at different temperatures during the melting curve course (*see Note 3*).

3.5.4 Nuclear Magnetic Resonance (NMR)

NMR spectroscopy is a powerful technique to obtain structural and dynamic information on proteins and protein complexes. NMR provides insight into the behavior of the protein and how homogeneous the sample is. It has many uses and is a central part of our repertoire: we use it for structural studies, small molecule screening, protein–protein interactions, protein–ligand interactions, assessing the effects of truncations and mutations on domain folding, and many other contexts. We have discussed its role in structural determination in our previous review [5]. Our focus here is on the use of NMR as a biochemical assay for studying interactions.

During an NMR experiment, the sample is placed in a strong magnetic field and irradiated with pulses of radio frequency electromagnetic radiation which cause the NMR-active nuclei to resonate at characteristic frequencies. Chemical shifts are very sensitive to the chemical environment constituting a very sensitive probe of

134 Rejina B. Khan et al.

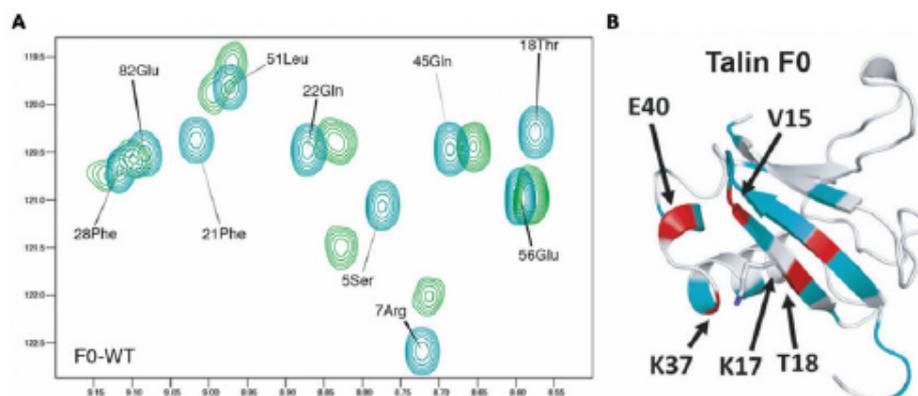


Fig. 3 Example of a 2D ¹⁵N HSQC NMR experiment showing binding of unlabeled Rap1b to 50 μM talin F0. (a) The chemical shift assignments of the F0 spectra are shown; each peak pertains to a specific residue. On addition of Rap1, the residues in close proximity to the binding site experience an altered chemical environment which results in the peak moving. (b) The peaks that move can be mapped onto the protein structure to define the binding site. (Reproduced from ref. 18 under a license from the Company of Biologists (<https://jcs.biologists.org/content/131/24/jcs225144>))

the conformation of the protein in solution and its interactions with ligands. Proteins can be unlabeled for 1D NMR, or they need to be singly labeled (¹⁵N) or doubly labeled (¹⁵N and ¹³C) in order to make those nuclei active under the magnetic field. Each amide group (NH) of the amino acids (with the exception of proline) of a protein will have a particular chemical shift that will be represented as a single peak with a specific location in an NMR spectrum. These chemical shifts can be easily measured using 2D ¹H-¹⁵N HSQC (Heteronuclear Single-Quantum Correlation) based experiments [6]. For fast time-resolved measurements, 2D ¹H-¹⁵N SOFAST-HMQC (band-Selected Optimized Flip-Angle Short Transient Heteronuclear Multiple-Quantum Correlation) spectra [7] can be measured in a few minutes providing similar information as the standard 2D HSQC. Binding of ligands to proteins result in localized alterations to the chemical environment of the residues in close proximity to the binding site. These chemical shift displacements produced by ligand binding allow the biochemical and structural characterization of protein interactions. Figure 3 shows part of the HSQC spectra of ¹⁵N-labeled talin F0 domain in the absence (teal) and presence (green) of the small GTPase Rap1b.

To follow biochemical features of the protein (folding, aggregation, binding, etc.) by two-dimensional NMR:

1. Prepare a sample of ^{15}N -labeled protein at approximately 0.1–1 mM in 500 μL of buffer for standard 5 mm NMR tubes or 350 μL for Shigemi tubes, containing 10% v/v D_2O .
2. Place the sample into the NMR tube. Take care to avoid bubbles.
3. Set the probe to the desired temperature (usually 298 K).
4. Once the spectrometer is set up (tuned, locked, shimmed, and pulse lengths determined) in accordance with the Bruker manual, collect 1D proton and 2D ^1H - ^{15}N HSQC spectra.
5. Process spectra (we use NMRPipe [8] and/or Bruker Topspin).
6. Analyze processed spectra using appropriate NMR software. In our laboratory, we use CcpNmr Analysis [9]. We have described our CCPN pipeline for talin NMR previously [5].
7. If more specific structural information is required, the assignment of the ^1H - ^{15}N cross peaks of the HSQC spectrum of the protein needs to be performed. This identifies the residue that each amide peak corresponds to. To easily assign a protein, double-labeling is recommended (^{15}N and ^{13}C) in order to be able to use three-dimensional, triple resonance NMR experiments. A normal set of 3D NMR experiments to assign a protein includes HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, HNCO, and HN(CA)CO.
8. NMR chemical shift assignments for each assigned protein are deposited into the Biological Magnetic Resonance Data Bank (BMRB) repository [10].

3.5.5 Fluorescence Polarization (FP)

We use FP experiments to investigate putative interactions between proteins and fluorescently labeled peptides. When a fluorescent molecule is excited by polarized light, it will emit light that is polarized. This assay utilizes the rapid tumbling that occurs when fluorescent peptides are in solution; as the molecule tumbles, it will lead to loss of polarization. If an interaction occurs between the fluorescently labeled peptide and the protein, the fluorescently labeled peptide will tumble more slowly with increasing levels of protein, causing an increase in the polarization of the light in one direction. This change in polarization as a function of protein concentration can be quantified, allowing determination of binding affinity. Figure 4 shows a typical data set obtained using this method. In this experiment, synthetic peptides of the talin binding site of KANK proteins were labeled with the fluorophore BODIPY-TMR [11]. This peptide was then added to a serial dilution of purified talin R7-R8 domains as follows.

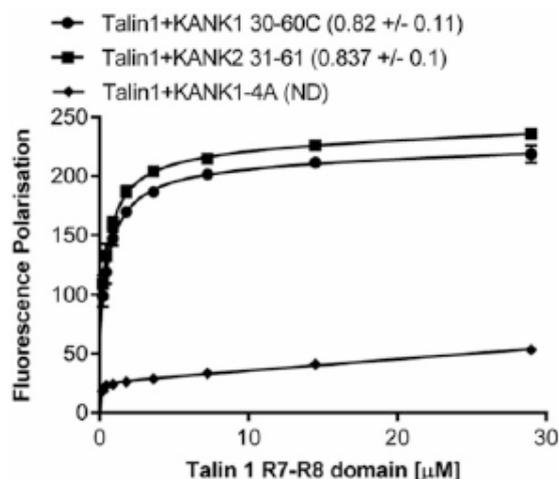


Fig. 4 Example of a dataset collected with the FP assay. Binding of BODIPY-labeled KANK1 (30–60), KANK2 (31–61), and KANK1-4A peptides to talin1 R7–R8. Dissociation constants \pm SE (μM) for the interactions calculated with the one-site binding equation are indicated in the legend. *ND* not determined. (Reproduced from ref. 11 under a Creative Commons CC-BY license)

1. Set up an FP program with appropriate parameters for your filters as described in the plate reader software manual. We generally perform experiments at 25 °C.
2. Each experiment uses one lane of a 96-well Nunc plate. Add 50 μL PBS into wells 1–11.
3. Add 50 μL of protein to well 12 (talin1 R7-R8 stock solution was 60 μM).
4. Add 50 μL of protein to well 11 and mix. Take 50 μL of this and add to well 10. Mix and transfer 50 μL of this into well 9. Repeat until well 2 and discard the final 50 μL .
5. Make 750 μL stock of 2 μM fluorescently labeled peptide. Add 50 μL of the peptide stock to wells 1–12 and mix well. Note: the fluorescently labeled peptide is susceptible to photobleaching so it is advised to protect from light.
6. Place the plate into the plate reader and run the appropriate FP program for the filters to be used.
7. Repeat in triplicate.
8. To calculate the binding constant, we use nonlinear curve fitting and a one site total binding model on GraphPad Prism.

FP assays can be used for multiple experiments, including competition assays where a labeled peptide is displaced from a complex using a competitor ligand.

3.5.6 Microscale Thermophoresis (MST)

MST is a flexible assay that allows for the analysis of interactions of proteins with ligands of any size and is well suited for binding analysis of proteins which can only be expressed in small amounts. We use MST to measure protein–protein interactions as, unlike the FP assay, it does not require a difference in size between the two species. This also makes the technique ideal for measuring affinities of protein dimerization. The MST assay measures temperature-induced changes in directed movement of a fluorescently labeled component against increasing concentrations of a putative ligand. A laser is used to create a temperature gradient (typically a few degrees), which induces the movement of particles due to a phenomenon called thermophoresis. Thermophoresis results in movement of the fluorescent moiety which can be quantified. The fluorescently labeled protein is titrated with an unlabeled ligand which upon binding alters the amount of movement in a concentration-dependent manner. This can be quantified using this assay, allowing determination of binding affinity. Figure 5 shows a typical data set obtained using this method. In this experiment, the dimerization of Talin rod domain containing protein 1 (TLNRD1) was measured using MST [12]. Any fluorescent labeling strategy will work for this technique, but we find that using NTA-dyes that bind with high affinity to the His-tag provides a simple, rapid, and efficient method for studying protein interactions.

Sample preparation includes the following steps:

1. Mix 1.5 μL of His-tag dye (diluted according to manufacturer's instructions) with 200 nM of His-tagged protein. Make up to 150 μL total volume using PBS. Leave incubating at room temperature for 30 min and then centrifuge at $16,200 \times g$

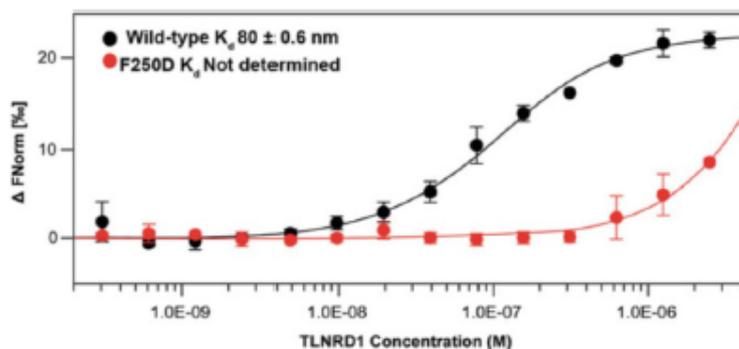


Fig. 5 Example of a dataset collected with the MST assay. Unlabeled TLNRD1 was titrated into a fixed concentration (50 nM) of fluorescently labeled TLNRD1 (black). The fitted curve yielded a dimerization K_D of 80 ± 0.59 nM. A monomeric mutant, F250D (red), resulted in a K_D not determined. (Reproduced with permission from ref. 12)

(13,000 rpm in a standard benchtop centrifuge) for 10 min at 4 °C to remove any aggregates. Note: protein concentrations may need adjusting according to affinities being measured.

2. Set up 14 × 200 µL tubes. Add 10 µL of PBS-Tween to tubes 2–14.
3. Add 10 µL ligand (unlabeled protein) to tube 1.
4. Add 10 µL ligand to tube 2 and mix. Transfer 10 µL of this to tube 3. Mix and take 10 µL of this and add to tube 4. Repeat until well 14 and discard the final 10 µL. Thus, tube 1 contains the highest concentration of ligand and tube 14 contains the lowest.
5. Add 10 µL of protein–dye solution to each tube and mix uniformly.
6. Place a capillary into each tube. This will draw up the solution. Place each capillary into the corresponding space on the stage within the NanoTemper Monolith NT.115.

Steps to set up the titration protocol.

1. Using MO.Control software, open a new session using Nano-RED settings (a different option may need to be selected according to the dye used).
2. Select “binding affinity.”
3. The “target” is the His-tagged protein which has been coupled to the His-dye. Select the His-tag option and input information for the other parameters, for example stock concentrations for each component.
4. For MST power and LED power, these can be set to autodetect if unsure. As a starting point, we generally use 50 and 40, respectively.
5. Ensure that a temperature has been set and keep this consistent for all experiments. We generally use 25 °C.
6. Start cap scan and MST. The cap scan should be consistent for each sample. If any potential issues such as sample inhomogeneity or concentration-dependent aggregation are suspected, the software will highlight this.
7. Generate a triplicate dataset.

*3.5.7 Size-Exclusion
Chromatography
Multiangle Light Scattering
(SEC-MALS)*

SEC-MALS is a technique that combines size exclusion chromatography (SEC) with multiangle light scattering (MALS) analysis. A MALS detector is a form of static light scattering detector which allows the absolute molecular weight (MW) and potentially the radius of gyration (R_g) of a sample to be measured. In SEC-MALS instruments, a MALS detector is connected to an

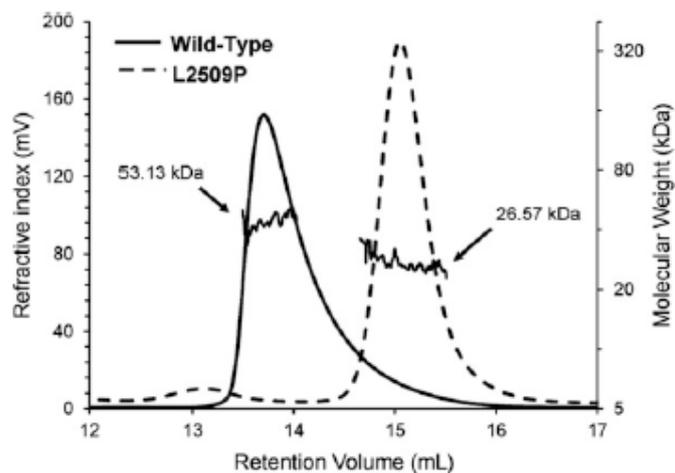


Fig. 6 Example of a dataset collected with the SEC-MALS binding assay. Figure shows the SEC-MALS analysis of talin R13-DD (dimerization domain) which constitutes actin-binding site 3 (ABS3). The WT protein elutes as a dimer whereas the L2509P mutant elutes as a monomer. The molecular weight of each peak obtained from the MALS is given. (Reproduced from ref. 13 under a Creative Commons CC-BY-NC-ND license)

HPLC/FPLC SEC column and a concentration, refractive index or UV detector is also connected.

The SEC-MALS detector measures the light scattered by the protein sample at different angles and a Debye plot is created. SEC-MALS can be used to acquired information about your protein size and also about aggregation state and binding by determining the absolute molecular weight, independent of the protein conformation, size and elution position. Figure 6 shows a typical data set obtained using this method. In this experiment, the SEC-MALS elution profiles of talin WT R13-DD (dimerization domain) and the L2509P mutant show the difference in molecular weight between them indicating dimerization in the wild type and not in the mutant [13].

To perform a SEC-MALS experiment:

1. Chose the appropriate SEC column for your proteins' size and equilibrate with the protein buffer until the MALS signal is stable. Most of the experiments we perform in our laboratory use either a Superdex-75 or Superdex-200 column (GE Healthcare).
2. Load the filtered protein sample, ideally 100 μ L of protein at 100–150 μ M.

140 Rejina B. Khan et al.

3. Run method appropriate for column size and type. Do not change the flow at any point to keep a stable signal. We normally use a flow rate of 0.75 mL/min.
4. The eluted peaks from the SEC column are measured by the MALS machine (we use a Viscotek Sec-Mals 9 and Viscotek RI detector VE3580 (Malvern Panalytical)) which is placed in line downstream of the column.
5. The elution profile is then analyzed to determine the absolute molecular weight of the protein/proteins of interest (we use the OmniSEC software from Malvern Panalytical).

3.5.8 GST Pulldowns

GST-tagged proteins bound to glutathione beads can be an excellent tool for pull-down experiments. Due to the difficulty in measuring the concentration of a purified GST-tagged protein while it is bound to glutathione beads, it may be useful to prepare samples of this and of the protein you are wishing to perform the pull-down experiment against for SDS-PAGE analysis. An image processing program such as ImageJ [14] can be used to quantify the bands on the gel for each protein, allowing an estimation of the concentration of the glutathione beads-bound protein.

1. Use the purified glutathione bead-bound GST-tagged protein (from **step 5** of Subheading 3.3.3). Mix 50 μ L of this with the putative ligand at the concentration required, to a final volume of 200 μ L. Using a 1:1 ratio of the GST-tagged protein to putative ligand is standard but it may be useful to have a different ratio depending on the particular interaction being investigated. Incubate for 1 h at room temperature.
2. After incubation, apply to a gravity flow column. It is useful to collect the flow through and prepare a sample for SDS-PAGE analysis.
3. Wash the beads with 1 mL of buffer and repeat this for a total of 2–3 washes. Allow the buffer to pass through the column.
4. Resuspend the beads in 500 μ L buffer. Transfer to a tube and centrifuge at $700 \times g$ (2510 rpm in a Hettich Rotanta 460 R centrifuge) for 3 min at 4 °C. Remove the supernatant.
5. Prepare a sample of the supernatant and of the pellet for SDS-PAGE analysis. If an interaction occurs, the bead-bound GST-tagged protein and the putative ligand should both be present in the pellet.
6. Important controls include: (1) Washed glutathione beads without the GST-tagged protein present and (2) glutathione beads with the GST-tag present but the protein removed via cleavage with TEV protease. These eliminate false-positive results generated from the putative ligand interacting with the beads or with the GST-tag.

7. Optimization of conditions may be required. These could include concentrations of the GST-tagged protein or the putative ligand, incubation temperatures, amount loaded onto SDS-PAGE gels, among others.
8. Generate a triplicate dataset. An image processing program such as ImageJ [14] can be used to quantify the bands on the gel.

3.5.9 Actin Cosedimentation Assay

Many adhesion proteins engage the actin cytoskeleton either directly or indirectly as part of actin-binding complexes. As a result, the biochemical characterization of protein interactions with actin is important for understanding these cytoskeletal linkages and can be achieved using an actin cosedimentation assay. By varying the speed at which the samples are spun it is possible to observe actin binding (high speed actin-binding assay) and actin bundling (low speed actin-bundling assay).

Actin Polymerization

All actin was isolated from rabbit muscle acetone powder (ours was kindly gifted by Prof Mike Geeves) using a previously described protocol [15]. Purified G-actin (monomeric) is stored at 10 mg/mL in G-buffer (5 mM Tris pH 7.5, 0.2 mM CaCl₂, 0.2 mM ATP) at -80 °C until required. Actin is then polymerized by diluting to ~50 μM with actin polymerization buffer (50 mM KCl, 2 mM MgCl₂, 0.2 mM ATP, 1 mM DTT, 1 mM NaN₃, pH 7), which is then stored at 4 °C and used for up to 1 month after polymerization.

High-Speed Actin-Binding Assay

The interaction between a protein of interest and actin can be measured using a simple high-speed actin cosedimentation assay. When spun at high speed such as 100,000 × *g*, polymerized actin filaments will form a pellet. If the protein of interest binds to the F-actin, it will go into the pellet with the actin whereas, if no interaction occurs, the protein will remain in the supernatant.

1. Dilute prepolymerized actin to 20 μM in polymerization buffer and incubate with 20 μM of protein for 1 h at room temperature. Protein-actin ratios can be altered according to experiment requirements. Optimal final volume is 100 μL.
2. Spin samples at 100,000 × *g* for 20 min at 4 °C using an ultracentrifuge. It helps to mark the outside facing edge of the centrifugation tube to identify pellet position.
3. Transfer 50 μL of supernatant to a clean tube and add 10 μL of 6× gel loading buffer. Discard the remaining supernatant, taking care not to disturb the pellet.
4. Gently wash the pellet with 50 μL polymerization buffer and resuspend in 60 μL 1× gel loading buffer (Co-sed buffer plus 6× gel loading buffer).

142 Rejina B. Khan et al.

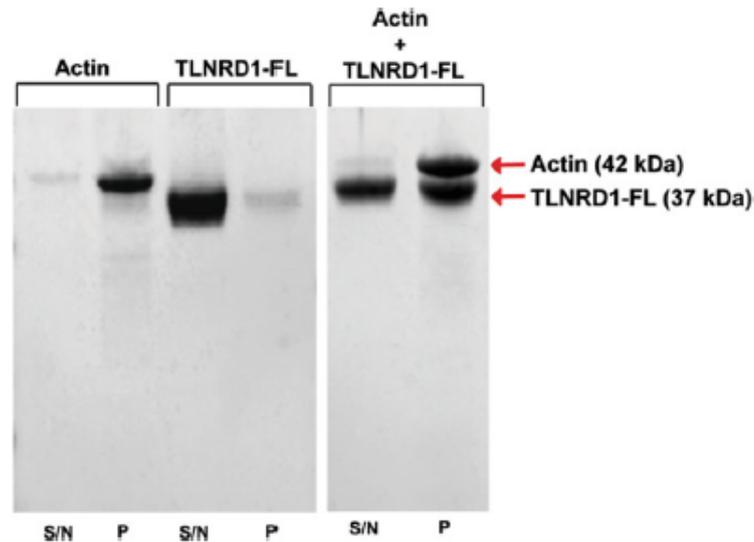


Fig. 7 Example of a dataset collected with the high speed actin-binding assay. This experiment shows TLNRD1 binding to actin. Left panel: Control experiment of the high-speed spin of the actin and TLNRD1 protein alone. Actin is seen in the Pellet fraction (P) whereas the TLNRD1 protein remains in the Supernatant fraction (S/N). Right panel: When TLNRD1 is incubated with actin TLNRD1 protein is seen in the Pellet fraction. (Reproduced with permission from ref. 12)

5. Load equal volumes of pellet and supernatant onto SDS-PAGE gels. We analyze the band densities using ImageJ software [14].

Figure 7 shows a typical data set obtained using this method. When spun at high speed, F-actin filaments will pellet. If proteins interact with the actin, they will cosediment and be detected in the pellet fraction. In this experiment, TLNRD1 was incubated with F-actin and shown to be an actin-binding protein, as seen by the TLNRD1 band in the pellet fraction only in the presence of F-actin [13].

Low-Speed Actin-Bundling Assay

Some proteins are able to bind to more than one actin filament simultaneously and so can induce actin-bundling. This actin-bundling functionality can be tested using a similar approach but using a low-speed actin cosedimentation assay. At low speed centrifugation (e.g., $16,200 \times g$), F-actin does not pellet and remains in the supernatant. If the protein of interest induces bundling of the actin filaments, the F-actin will form a pellet. By comparing the amount of actin in the pellet vs. supernatant, whether the protein is bundling actin filaments can be determined. Figure 8 illustrates how a low speed centrifugation can identify proteins with actin-bundling capabilities. When spun at low speed, single F-actin

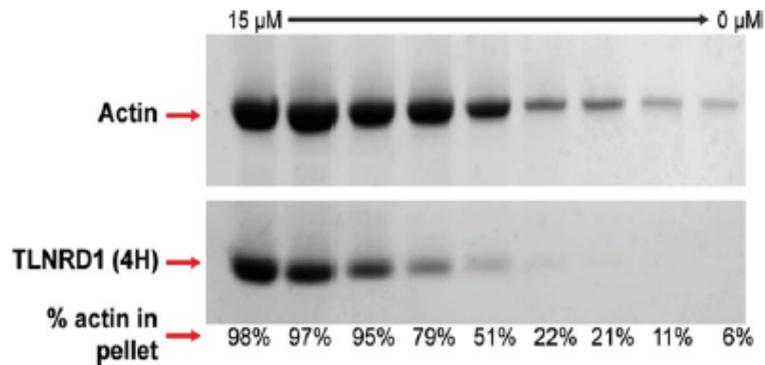


Fig. 8 Example of a dataset collected with the low speed actin-bundling assay. This experiment shows TLNRD1 bundling actin. In the absence of TLNRD1 the actin is mostly in the solution. As the amount of actin-bundling protein is increased the amount of bundled actin in the pellet increases. (Reproduced with permission from ref. 12)

filaments do not pellet while bundled F-actin filaments (many filaments crosslinked together) do. In this assay, actin is spun at low speed; by itself (right side of gel) only a small amount of actin is found in the pellet. However, as the actin-bundling protein, TLNRD1 is added the amount of actin in the pellet increases as the filaments are bundled together.

1. Prepare samples using the same approach as the high-speed binding assay. Protein-actin ratios can be altered according to experiment requirements.
2. Spin samples at $16,200 \times g$ (13,000 rpm in a standard benchtop centrifuge) for 20 min at 4 °C using a benchtop centrifuge.
3. As previously, transfer 50 μ L of supernatant to a clean tube and add 10 μ L of $6\times$ gel loading buffer. Carefully remove and discard the remaining supernatant.
4. Resuspend the pellet in 60 μ L $1\times$ gel loading buffer (Co-sed buffer plus $6\times$ gel loading buffer).
5. Analyze supernatant and pellets on SDS-PAGE using the same approach as the high-speed assay.

3.5.10 Lipid Cosedimentation Assays

Integrin adhesion complexes assemble on the cytoplasmic face of the integrin. As a result, many proteins are in close proximity to, and engage with, the plasma membrane. The biochemical characterization of protein interactions with lipids is important for understanding the role of the membrane in mediating these interactions and can be probed using lipid cosedimentation assays. In these experiments, we prepare phospholipid vesicles comprised of different combinations of lipids. Figure 9 shows a typical data set obtained using this method. When spun at low speed, lipid vesicles

144 Rejina B. Khan et al.

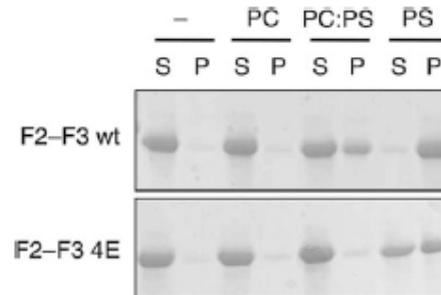


Fig. 9 Example of a dataset collected with the lipid cosedimentation assay. This experiment shows the talin F2–F3 domain (0.15 mg/mL) binding to lipid vesicles. Either WT or a non–membrane binding mutant (4E) F2–F3 is mixed with vesicles (0.5 mg/mL) consisting of phosphatidylcholine (PC), phosphatidylserine (PS), or a 4:1 ratio of PC:PS. The Supernatant (S) and Pellet (P) lanes are shown. The vesicles are found in the Pellet (P) fractions and proteins that bind the vesicles are also found in the pellet. (Reproduced from ref. 16 under license from John Wiley and Sons (<https://www.embopress.org/doi/full/10.1038/emboj.2009.287>))

will pellet. If proteins interact with the lipids, they will cosediment and be detected in the pellet fraction. In this experiment, the talin head domains F2–F3 were incubated with different lipid vesicles [16]. The talin head cosediments with vesicles with negatively charged head groups as seen by the F2–F3 band in the pellet fraction only in the presence phosphatidyl serine.

Preparation of Large Multilamellar Vesicles

We prepare large multilamellar vesicles essentially as described previously [17].

1. Dissolve lipids in chloroform and evaporate under nitrogen to produce films of dried phospholipids.
2. Swell films of dried phospholipids at 5 mg/mL in 20 mM HEPES pH 7.4, 0.2 mM EGTA at 42 °C for 3 h.
3. Centrifuge at $16,200 \times g$ (13,000 rpm in a standard benchtop centrifuge) for 20 min at 4 °C.
4. Resuspend pellet at 5 mg/mL in 20 mM HEPES pH 7.4, 0.2 mM EGTA.
5. For large multilamellar vesicles (LMV) the resuspended lipids can be used as is.
6. To generate small unilamellar vesicles (SUV), the resuspended lipids can be passed $11\times$ through a lipid extruder (we use Avanti Mini-Extruder). The extrusion process makes the vesicles smaller and more uniform in size.

- | | |
|------------------------|--|
| Proteins | <ol style="list-style-type: none"> 1. Dilute protein into 20 mM Tris pH 7.4, 0.1 mM EDTA, 15 mM β-mercaptoethanol at 1 mg/mL. 2. Centrifuge at $16,200 \times g$ (13,000 rpm in a standard benchtop centrifuge) for 20 min at 4 °C. 3. Transfer and keep supernatant. |
| Interaction Experiment | <ol style="list-style-type: none"> 1. Incubate protein sample at 0.15 mg/mL for 30 min at 25 °C in the absence or presence of phospholipid vesicles (0.5 mg/mL) in 200 μL total volume. 2. Make up the following mixture: <ul style="list-style-type: none"> 20 μL lipid (final conc. 0.5 mg/mL). 30 μL protein (final conc. 0.15 mg/mL). 150 μL 20 mM Tris pH 7.4, 0.1 mM EDTA, 15 mM β-mercaptoethanol. 3. Incubate at 25 °C for 30 min. 4. Centrifuge at $16,200 \times g$ for 20 min at 4 °C. 5. Remove supernatant and add 40 μL of 5\times sample buffer. 6. Resuspend pellet in 120 μL of 1\times sample buffer. 7. For optimal visualization using SDS-PAGE analysis, we load 14 μL for supernatant fractions and 7 μL for pellet fractions. 8. Stain with Coomassie blue. 9. The percentage of protein bound (protein in pellet/total protein) can now be calculated by measuring band density in ImageJ [14]. |

3.5.11 Structural Studies A key part of our pipeline is to elucidate the structure of the proteins and protein complexes. Investigating proteins and their interactions at the atomic level provides insight into their biological function and enables the elucidation of the molecular determinants of their interactions. This information allows the design of targeted mutations to disrupt the interactions which, following validation of the mutations' efficacy using the biochemical pipeline outlined here, allows the interaction to be studied in fine detail in cell biology experiments. We use NMR, small-angle X-ray scattering (SAXS), X-ray crystallography, and cryo-electron microscopy (cryo-EM) to explore the structural aspects of integrin adhesion complexes. A detailed discussion of these structural biology approaches is beyond the scope of this chapter and will be the subject of a subsequent chapter.

4 Notes

1. HisTrap HP columns (5 mL) have a pressure limit of 0.3 mPa. For each step of this process, ensure that a high-pressure alarm

is set to warn if the pressure is reaching/exceeding this limit. The maximum flow rate for this column is 5 mL/min.

2. Q HP and SP HP columns (5 mL) have a pressure limit of 0.5 mPa. For each step of this process, ensure that a high-pressure alarm is set to warn if the pressure is reaching/exceeding this limit. The maximum flow rate for this column is 5 mL/min.
3. A nitrogen flow of 9 L/min must be used to purge and refrigerate the system to protect the optics.

Acknowledgments

We thank members of the Goult laboratory past and present for help in the development of these assays. We also thank Alex Gingras and Neil Bate for helping develop the protocols for protein expression and purification. The Goult lab is funded by BBSRC grants (BB/N007336/1 and BB/S007245/1), and HFSP grant (RGP00001/2016). R.K. is funded by a University of Kent studentship.

References

1. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6 (5):343–345. <https://doi.org/10.1038/nmeth.1318>
2. Marley J, Lu M, Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20(1):71–75. <https://doi.org/10.1023/a:1011254402785>
3. McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16(4):404–405. <https://doi.org/10.1093/bioinformatics/16.4.404>
4. Goult BT, Zacharchenko T, Bate N, Tsang R, Hey F, Gingras AR, Elliott PR, Roberts GC, Ballestrem C, Critchley DR, Barsukov IL (2013) RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J Biol Chem* 288 (12):8238–8249. <https://doi.org/10.1074/jbc.M112.438119>
5. Skinner SP, Goult BT, Fogh RH, Boucher W, Stevens TJ, Laue ED, Vuister GW (2015) Structure calculation, refinement and validation using CcpNmr analysis. *Acta Crystallogr D Biol Crystallogr* 71(Pt 1):154–161. <https://doi.org/10.1107/S1399004714026662>
6. Bax A, Morris GA (1981) An improved method for heteronuclear chemical shift correlation by two-dimensional NMR. *J Magn Reson* (1969) 42(3):501–505. [https://doi.org/10.1016/0022-2364\(81\)90272-9](https://doi.org/10.1016/0022-2364(81)90272-9)
7. Schanda P, Kupce E, Brutscher B (2005) SOF AST -HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33(4):199–211. <https://doi.org/10.1007/s10858-005-4425-x>
8. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293. <https://doi.org/10.1007/bf00197809>
9. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59 (4):687–696. <https://doi.org/10.1002/prot.20449>
10. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL (2008) BioMagResBank. *Nucleic Acids Res* 36(Database issue):

- D402–D408. <https://doi.org/10.1093/nar/gkm957>
11. Bouchet BP, Gough RE, Ammon YC, van de Willige D, Post H, Jacquemet G, Altelaar AM, Heck AJ, Goult BT, Akhmanova A (2016) Talin-KANK1 interaction controls the recruitment of cortical microtubule stabilizing complexes to focal adhesions. *Elife* 5. <https://doi.org/10.7554/eLife.18124>
 12. Cowell AR, Jacquemet G, Singh AK, Paatero I, Brown DG, Ammon Y-C, Akhmanova A, Ivaska I, Goult B (2020) Talin Rod Domain Containing Protein 1 (TLNRD1), a novel actin bundling protein which promotes filopodia formation. *BioRxiv*. <https://doi.org/10.1101/2020.05.19.103606>
 13. Azizi L, Cowell AR, Mykuliak VV, Goult BT, Turkki P, Hytonen VP (2020) Cancer associated talin point mutations disorganise cell adhesion and migration. *bioRxiv*: 2020.2003.2025.008193. <https://doi.org/10.1101/2020.03.25.008193>
 14. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675. <https://doi.org/10.1038/nmeth.2089>
 15. Spudich JA, Watt S (1971) The regulation of rabbit skeletal muscle contraction. I. Biochemical studies of the interaction of the tropomyosin-troponin complex with actin and the proteolytic fragments of myosin. *J Biol Chem* 246(15):4866–4871
 16. Anthis NJ, Wegener KL, Ye F, Kim C, Goult BT, Lowe ED, Vakonakis I, Bate N, Critchley DR, Ginsberg MH, Campbell ID (2009) The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO J* 28(22):3623–3632. <https://doi.org/10.1038/emboj.2009.287>
 17. Niggli V, Kaufmann S, Goldmann WH, Weber T, Isenberg G (1994) Identification of functional domains in the cytoskeletal protein talin. *Eur J Biochem* 224(3):951–957. <https://doi.org/10.1111/j.1432-1033.1994.00951.x>
 18. Camp D, Haage A, Solianova V, Castle WM, Xu QA, Lostchuck E, Goult BT, Tanentzapf G (2018) Direct binding of Talin to Rap1 is required for cell-ECM adhesion in *Drosophila*. *J Cell Sci* 131(24):jcs225144. <https://doi.org/10.1242/jcs.225144>