UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

School of Biological Sciences

Dissecting interactions within focal adhesions: studies on Vinculin

by

Clare Louise Cox

Thesis for the degree of Doctor of Philosophy

September 2008

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
SCHOOL OF BIOLOGICAL SCIENCES

Doctor of Philosophy

DISSECTING INTERACTIONS WITHIN FOCAL ADHESIONS: STUDIES ON VINCULIN

by Clare Louise Cox

Vinculin and paxillin are proteins that localise to focal adhesions. Focal adhesions are specialised sites of cell attachment between the cytoplasmic side of the cell membrane and the extracellular matrix. They consist of receptors that link extracellular matrix ligands to the actin cytoskeleton via various protein assemblies. They act as mechanical links and as sites of signal transduction to transduce signals for cell locomotion, cell attachment and detachment, apoptosis and gene expression. A range of cellular responses depend critically on the composition and regulation of focal adhesions.

Vinculin and paxillin will interact *in vitro*. The tail domain of vinculin interacts with a motif on paxillin called an LD motif. The interaction between LD motifs and their target proteins is important for regulation of focal adhesion signalling, yet little is known regarding recognition mechanisms between LD domains and interaction partners.

In this thesis, the molecular nature of the vinculin tail (Vt) and paxillin LD motif interaction has been studied using purified Vt, synthesised LD peptide mimics and recombinant paxillin His-LD1/LD2. The ^{1}H-^{15}N-HSQC spectrum of a Vt/I997S mutant has been assigned and the NH assignments transferred to a wild-type Vt spectrum. Chemical shift perturbation studies have subsequently been undertaken using wild-type Vt and paxillin.

The data presented here is consistent with specificity for LD motifs and points to an interaction between Vt and LD1 and LD2. In contrast there is little or no interaction with LD4. Two binding models are proposed; a single binding site on Vt face 3-4 where LD motifs bind in an extended conformation, or a two-site binding on Vt face 3-4 with two LD motifs bound as α-helices. Both models need further analysis. NMR data for the paxillin His-LD1/LD2 construct suggests a predominantly unstructured molecule in solution that can catalyse precipitation of Vt when Vt is added to excess.

It is clear that the mechanism of interaction between Vt and paxillin LD motifs is distinct from that of focal adhesion kinase interacting with paxillin. Further investigation is required to elucidate the precise mechanism of binding. A comparison of this data with other LD-protein interactions suggests there are little similarities between the target sequences that LD motifs recognise and that target proteins can be structurally different.

Table of Contents

Tables and Figures

DECLARATION OF AUTHORSHIP

I, Clare Louise Cox, declare that the thesis entitled 'Dissecting interactions within focal adhesions: studies on Vinculin' and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself or jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed:

Date:

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Jörn Werner for help and guidance throughout this challenging project. I would also like to thank the BBSRC for funding.

Thank you to the NMR group past and present for your help, especially Rob Hagan who helped me during the early stages of my project. Thanks also go to Dr. Geoff Kelly at the MRC Biomedical NMR Centre at the National Institute for Medical Research (NIMR), and to Dr. Sara Whittaker at the Henry Wellcome Building for Biomolecular NMR Spectroscopy at the University of Birmingham for assistance in acquiring valuable NMR data. Also, thanks to Neville Wright for carrying out mass spectrometry for me and for chats in the corridor!

A very large thank you goes to Dr. Stuart Findlow for all the help and support during my PhD. I also want to thank Andrea Roberts and Halina Mikolajek (although at time of writing Halina is now a Bainbridge). Without both of you I do not think I would have made it! Thanks for all the cups of tea and other beverages of a stronger nature.

The largest thanks has to go to Robin Tirrell -thanks for putting up with me for the last four years and for supporting me when things were tough. Here it is Rob, all in one piece. Remember-you promised you would read the whole thing! And to my family-Mum, Dad, Sarah and Lisa, and to Alison and Jim-thanks for being supportive about the whole thing.

A final thanks goes to vinculin for being the most frustrating protein I have ever worked on!

Abbreviations

1D	one dimensional
2D	two dimensional
3D	three dimensional
A_{260}	absorbance at 260nm
A_{280}	absorbance at 280nm
Akt1	protein kinase B
ANK	ankyrin
Arf	ADP-ribosylation factor
Arp2/3	actin-related protein 2/3
ASAP1	arf GAP containing SH3, ANK repeat and PH domains
CD	circular dichroism
CEF	chick embryo fibroblasts
CH	calponin homology
Crk	C10-regulator of kinase
CrkL	crk-like
CSA	chemical shift anisotropy
Csk	C-terminal Src kinase
CV	column volume
δ	chemical shift
D1	domain 1 (vinculin)
D2	domain 2 (vinculin)
D_2O	deuterium oxide
D3	domain 3 (vinculin)
D4	domain 4 (vinculin)
D5	domain 5 (Vt)
ddH_2O	double distilled water
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DOCK180	180-kDa protein downstream of Crk
E.coli	Escherichia coli
E6	E6 oncogene of human papilloma virus
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
ELMO	engulfment and cell motility
ERK	extracellular signal-regulated kinases (MAPK)
ERK1	extracellular signal-regulated kinase 1 (MAPK3)
ERK2	extracellular signal-regulated kinase 2 (MAPK1)
ESI-MS	electrospray ionisation mass spectrometry

FA	focal adhesion
FAK	focal adhesion kinase
FAT	focal adhesion targeting
FB	fibrillar adhesion
FERM	band 4.1 and ezrin, radixin, and moesin homology
FRAP	fluorescence recovery after photobleaching
FRET	förster (fluorescence) resonance energy transfer
FX	focal complex
GAP	GTPase-activating protein
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factors
Git1	G protein-coupled receptor kinase interacting protein 1
GRAF	GAP for Rho associated with focal adhesion kinase
Grb	growth factor receptor-bound protein
GST	glutathione-S-transferase
GTP	guanosine-5'-triphosphate
HIS	histidine tag
HSQC	heteronuclear single quantum correlation
HTI	head-tail interaction
ILK	integrin-linked kinase
INEPT	insensitive nuclei enhanced by polarisation transfer
IpaA	invasion plasmid antigen A
IPTG	isopropyl ß-D-1-thiogalactopyranoside
JNK	c-Jun N-terminal kinase
K_d	dissociation constant
kDa	kilodaltons
K_{ex}	exchange rate
kinase	phosphorylation enzyme
L	ligand
LB	luria broth
LD	leucine/aspartate
LIM	lin11, isl-1 & mec-3
MAPK	Mitogen-activated protein kinase
MEK	MAPK/ERK kinase
MEKK1	MEK kinase 1
MKK4/7	MAPK kinase 4/7
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhancement spectroscopy
N-WASP	N-Wiskott Aldrich syndrome protein

OD_{600}	optical density at 600nm
P	protein
p120RasGap	p120 Ras GTPase activating protein
p130Cas	crk-associated substrate
PABA1	poly(A)-binding protein 1
PAK	p21-activated serine-threonine kinase
PBD	paxillin binding domain
PBS	paxillin binding subdomain
PCR	polymerase chain reaction
pD	measure of acidity/alkalinity in $^{2}H_2O$
PDB	protein data bank
pH	measure of acidity/alkalinity in $^{1}H_2O$
PH	pleckstrin homology
PI3K	phosphoinositide 3-kinase
PIP2	phosphatidylinositol bisphosphate
PIX	pak-interactive exchange factors
PKCa	protein kinase C alpha
PKL	paxillin kinase linker
PL	protein-ligand
ppt	precipitate
PTK	protein tyrosine kinase
PTP-PEST	protein tyrosine phosphatase-PEST
Rac	rac small GTPase
Rac1	rac small GTPase
Raf	raf non receptor protein kinase (ser/thr)
Ras	ras small GTPase
RF	radiofrequency pulse
Rho	rho small GTPase
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SH2	src homology 2 domain
SH3	src homology 3 domain
SOC	super optimal broth plus glucose
Sos	son of sevenless
Src	src tyrosine kinase
τc	calculated correlation time
τm	measured correlation time
TALOS	torsion angle likelihood obtained from shifts and sequence similarity
TBE	tris/borate/EDTA

TEMED	tetramethylethylenediamine
Tm	melting temperature
TOCSY	total correlation spectroscopy
TROSY	transverse relaxation optimised spectroscopy
UV	ultraviolet
VASP	vasodilator-stimulated phosphoprotein
VBS	vinculin binding site
Vh	vinculin head
Vt	vinculin tail
WT or wt	wild-type

1 Introduction

1.1 Aims

Vinculin, paxillin and focal adhesion kinase are proteins that localise to regions of the cell called focal adhesions. Focal adhesions are discrete sites that mediate cell adhesion, motility, differentiation, development and apoptosis (1-4). This study aims to determine the nature of the interaction between the tail domain of vinculin and paxillin LD motifs. A comparison to other known protein-LD motif interactions will be made in an attempt to elucidate common interaction motifs. At the beginning of this study, the only characterised interactions were using vinculin and focal adhesion kinase with paxillin LD motifs (5-9). Since then the field has advanced significantly with the structure of Git1 (10) and actopaxin (11) in complex with the LD motifs of paxillin.

The main technique used in this study is nuclear magnetic resonance (NMR) however other biophysical techniques have been employed.

1.2 Cell-extracellular matrix connections: Focal Adhesions

Cells such as platelets, fibroblasts and epithelial cells make numerous contacts with the extracellular matrix (ECM) (12). Observations that fibroblast cells do not adhere uniformly to a substrate, but at discrete focal points (13) led to the identification of regions in the cell that were named focal adhesions (14). Since then, an increasing number of components have been identified within focal adhesions and their spatial and temporal organisation has been realised (2;15-20).

A classical focal adhesion is a specialised site of cell adhesion between the cell membrane and the ECM (18). They consist of receptors that link ECM ligands to the cytoskeleton via various protein complexes. Since their initial identification, focal adhesions have

subsequently been divided into various forms depending on their protein composition, method of formation and position in the cell (16;21). The term focal complex is used to determine newly forming dot-like adhesions at the edge of lamellipodia that either turnover or become focal adhesions (18;22-25). Focal adhesions are larger mature adhesions associated with actin (18;25) and fibrillar adhesions develop from focal adhesions and define an elongated form typically containing fibronectin receptors and tensin (25;26).

The reason for these different forms is a consequence of the dynamic and flexible nature of these adhesions. They act as signalling hubs that can integrate many complex signals to control key processes in the cell (27). Consequently, continual regulation and rearrangement of adhesion to the ECM and changes in the cytoskeleton is necessary. Thus focal adhesion components are highly regulated and turn over in response to mechanical and chemical information to maintain the correct functioning of the cell.

As a result, adhesion sites between the cell and the ECM can be highly variable and heterogeneous (18;19;28). For the purpose of this report the term focal adhesion (FA) will be used when describing connections between the cell and the ECM. Where appropriate the distinction will be made as to early, immature FAs or late, more mature FAs, but it is important to remember that at any moment in time a FA will be dynamically regulated to respond to the constantly changing cellular environment.

FA formation is primarily integrin mediated and indirectly links ligands of the ECM to the actin cytoskeleton (29;30). Integrins are heterodimeric transmembrane receptors that transmit signals across the membrane in both directions; outside-in signalling between the ECM and the cytoplasm, and inside-out signalling between the cytoplasm and the ECM (3;31;32). Integrin binding to the ECM and to intracellular components can lead to integrin activation and clustering with a concomitant recruitment of cytoskeletal proteins to the adhesion site (33).

The initial weak link between small clusters of integrins and the cytoskeleton is thought to be mediated by the binding of the cytoskeletal protein talin1 (34)(see Figure 1.2.1.B). Initial adhesions also contain proteins such as actin (35), paxillin, vinculin and focal

adhesion kinase (FAK) and show high levels of tyrosine phosphorylation (18). Figure 1.2.1C shows a snapshot of a FA and some of the possible interactions.

Currently there are over 90 proteins that have been reported to associate with FAs (36) and their complex interactions regulate many cellular processes such as cell motility, cell survival, transcription, cytoskeletal organisation and cell proliferation (3) and so play roles in embryonic development, immunity, wound healing and attachment and detachment from the extracellular matrix (1;2). If this regulation functions abnormally, altered rates of cell motility and resistance to apoptosis can occur which can lead to cancer (4). Other effects of disregulation include abnormally migrating inflammatory cells such as in chronic inflammation and asthma (37). Developmental abnormalities can also occur making many FA mutants embryonic lethal (36).

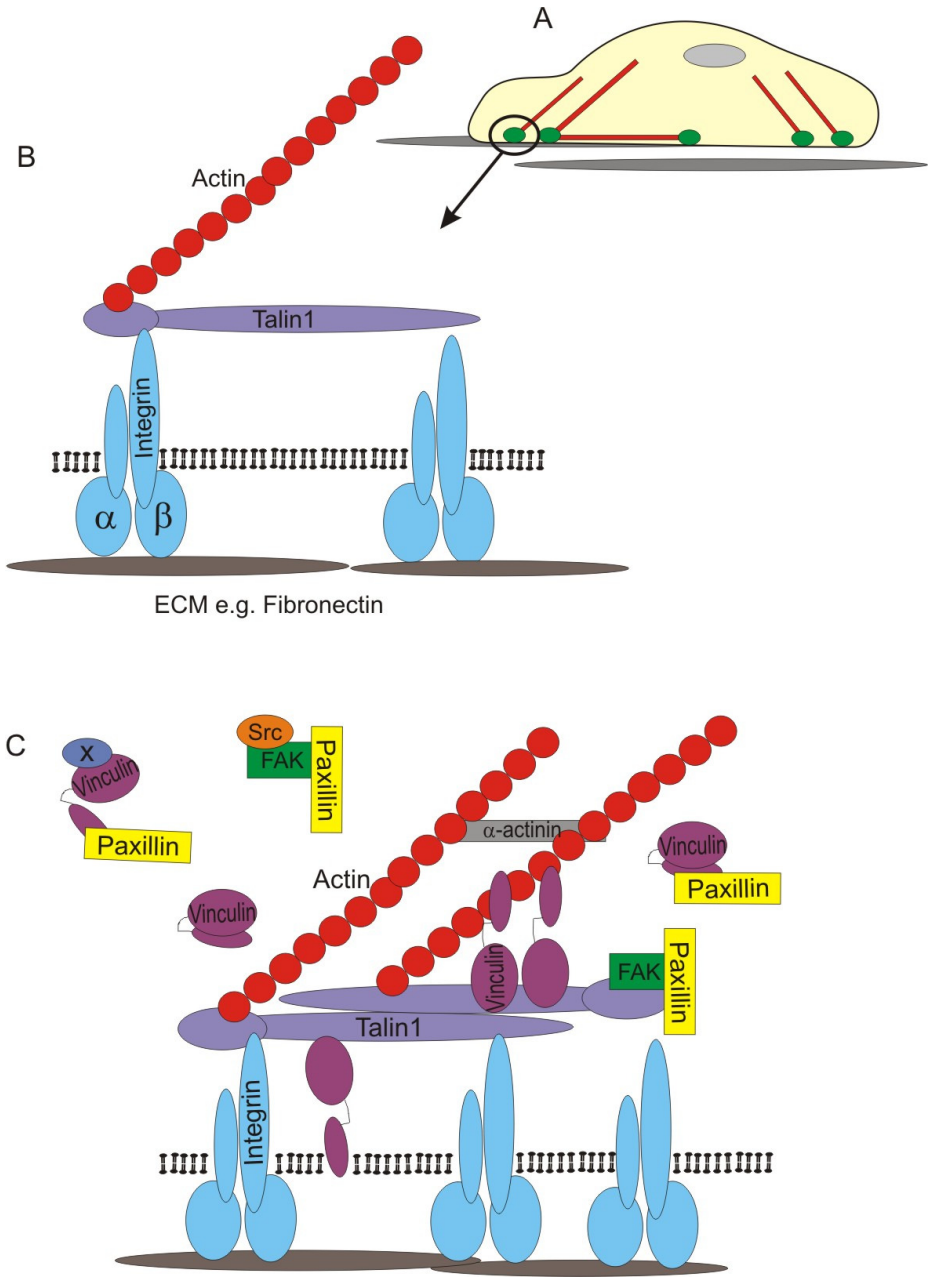


Figure 1.2.1 Potential interactions occurring at focal adhesions

A: representation of a fibroblast cell plated onto ECM ligands (such as fibronectin). FAs are sites between the cell membrane and the ECM, shown as green circles. They are attached to actin stress fibres, shown in red. B: Likely composition of the initial weak link in early formation of FAs (34). The cell membrane is shown in black. C: Possible interactions in a maturing FA. While many of these interactions have been shown in vitro, less is known about their cellular context. The interaction between vinculin and paxillin is poorly understood so these are possible but speculative in this picture. Many of the interactions can be found in references (3;38-49).

1.2.1 Maturation and turnover of focal adhesions

Once an adhesion contact is formed it can either turn over or mature into a more organised structure (50;51). The fate of a FA is largely determined by the environmental cues that the cell is experiencing. Maturation and turnover can be regulated by the phosphorylation state of proteins in the FA such as FAK, Src and paxillin (50;52), tyrosine phosphatases such as PTP-PEST (53) and proteases such as calpains (31). The dynamics are thought to be controlled by the small GTPases, Rac and Rho; Rac activity is required for formation of early FAs, and Rho activity is required for maturation (26). Force will also control FA dynamics as FAs can function as mechanosensors. When a force is applied from within the cell (for example via the actin cytoskeleton) or externally from the ECM, FAs will increase in size and maturate as more proteins are recruited. FAs will also shrink if this force is relaxed and this occurs proportionally to the size of the force (54-56).

Migrating cells are continuously forming and disassembling FAs at the leading edge of a cell protrusion (50). Some of these remain fixed as other proteins are recruited and the cell migrates over them while others turnover and are recycled (57). At the rear of the cells FAs are being disassembled and released by the cell causing contraction of the cell and motion forward, the mechanisms of which are distinct from turnover and disassembly at the leading edge (21).

Once a FA is formed in motile cells they do not appear to move relative to the substrate (15;29). In contrast, stationary cells appear to have motile FAs that move within the cell. One suggestion is that motile FAs prevent the cell from generating the traction required for movement, but enables a connection that is ready to be remodelled when migratory signals are received, thus the cell is in an 'idling' state, ready to respond to environmental cues (15). Clearly both states require tight regulation to ensure correct functioning.

Understanding the interactions between the proteins involved in FA dynamics and the complex interplay between signalling pathways that control cell migration is challenging. Many of the pathways are not linear and involve feedback loops and interconnections

between pathways. For example FAK, which is an important player in FA dynamics, can promote cell migration via multiple pathways involving interactions with Src (58), PI3K (59), Grb7(60;61), p130Cas (62) and N-WASP (63;64), and each of these interactions appears to be required for the promotion of cell migration (52).

Consequently, understanding the regulation of FA formation and turnover is of key importance. FAs are dynamically assembled and disassembled within cells, and this turnover is partly governed by the transient interactions between the proteins involved (2). So while the composition of the FAs is important, the binding events that take place between the proteins is equally important to understand. This requires an intimate understanding of the molecular nature of the individual reactions, their relative affinities and the stoichiometries of binding. At present it is difficult to study direct protein-protein interactions *in vivo*, although methods are emerging that are beginning to make this feasible, such as proximity of proteins using FRET (e.g. (19)) and binding kinetics using FRAP (e.g. (2)). *In vitro* experiments that identify binding partners have been useful in exploring potential interactions, and now that the dynamic and heterogeneous nature of FAs is fully realised it is now becoming feasible to ask very specific questions about the function of binding partners within FAs.

This thesis builds on *in vitro* observations about the interactions of the cytoskeletal protein paxillin with vinculin and FAK. The following sections introduce the three proteins and explore the current knowledge about the roles and interactions within FAs.

1.3 Introducing selected cytoskeletal components

A key feature of many proteins involved in FAs is that they are made of different modules that can provide a platform for binding a number of other proteins and signalling molecules (65). The following section introduces paxillin, FAK and vinculin which are three important cytoskeletal proteins involved in FA signalling.

1.3.1 Paxillin

Paxillin is a 68kDa cytoskeletal protein identified in 1989 as a phosphotyrosine containing protein in Src-transformed cells (66). It was characterised as a vinculin binding partner and named paxillin after the Latin *paxillus*, meaning a small stake or peg (67). The name arose from the belief that paxillin acted as a passive structural link between actin and the ECM (68), but it is now realised that paxillin is an active signalling molecule within FAs, not just a passive tether.

Paxillin is recruited early in the formation of FAs (69). It contains no intrinsic enzyme activity but facilitates multiple protein-protein interactions that can drive the formation of protein complexes, hence it is termed an adaptor protein (5). Paxillin null cells form FAs but they have delayed spreading and delayed cell migration when plated on fibronectin (70). This is consistent with the idea that paxillin is involved in the regulation of FA turnover rather than formation of FAs.

The domain structure of paxillin is shown in Figure 1.3.1. It is a modular protein that contains a number of domains that are interaction sites for other proteins. The C-terminus contains four tandem LIM domains and a number of phosphorylation sites while the N-terminus contains five LD domains, several phosphorylation sites and a proline rich region that may serve as a potential SH3 docking motif (68).

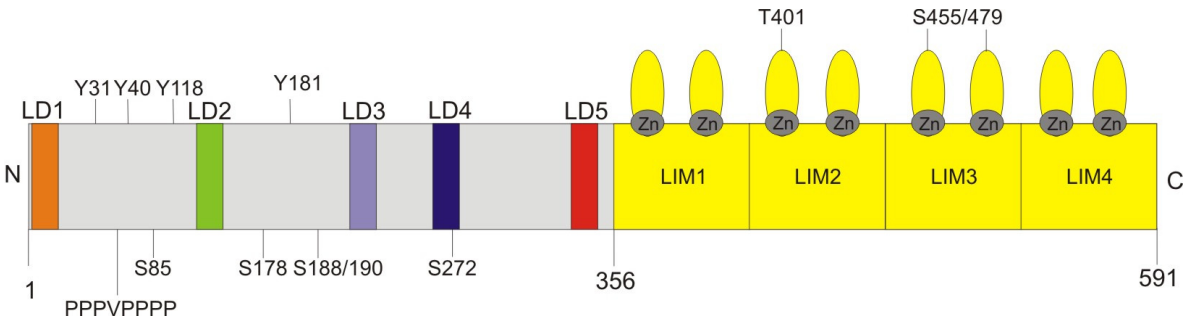


Figure 1.3.1 Domain structure of paxillin (human numbering)

Paxillin contains five LD domains, four LIM domains, multiple phosphorylation sites (not all sites are represented here) and a proline-rich region (PPPVPPPP)

There are many known interaction partners of paxillin, some of the best characterised can be found in Table 1.3.1. The C-terminal LIM domains are double zinc finger motifs that are known to mediate protein-protein interactions (71). LIM3 is required for targeting of paxillin to focal adhesions (5) and phosphorylation of LIM2/3 regulates this process (72), but the protein that functions to target paxillin to FA remains elusive (68). Proteins that can bind to LIM domains of paxillin include PTP-PEST (73;74) and α- and γ-tubulin (75).

Table 1.3.1 Interaction partners of paxillin

Paxillin domain	Binding partner	Reference
LIM1	None identified	
LIM2	α-tubulin, γ-tubulin unidentified kinase	(75;76) (72)
LIM3	α-tubulin, γ-tubulin FA localisation PTP-PEST unidentified kinase	(75;76) (5) (73;74) (72)
LIM4	PTP-PEST	(73;74)
LD1	Actopaxin (α-parvin) ILK Vinculin E6 protein	(11;77) (78) (5;79;80) (81)
LD2	Vinculin FAK Git1	(5;79) (5;6;79) (10;80;82)
LD3	None identified	
LD4	Actopaxin (α-parvin) FAK Vinculin PKL Git1 Bcl-2	(11;77) (5;6;79) (5;79;80) (80) (10;80;82) (83)
LD5	None identified	
Proline-rich region 48-57	Src Crk	(84) (84)

The N-terminus of paxillin contains five leucine rich sequences called LD motifs. LD motifs can mediate binding to proteins such as vinculin and FAK (5;79), Git1(10;80;82), ILK (78), actopaxin (77) and PKL (80) (Table 1.3.1). They were identified by sequence

analysis of similar regions of paxillin that bound to FAK and vinculin and have a consensus sequence of LDxLLxxL (5;81;85). Paxillin LD motifs 1 to 5 are shown in Figure 1.3.2.

The LD motifs in paxillin are predicted to form amphipathic α-helices with each of the leucine residues being positioned on one face of the helix (85). This would form a hydrophobic binding interface down one face of the helix that could interact with other proteins. The interaction domain on molecules that interact with paxillin has been named the paxillin binding subdomain (PBS) (86), these regions are expected to share similarities that direct LD binding (85).

Whether LD motifs always adopt α-helical arrangements is still under some debate as subsequent studies have provided contradictory results. An LD2 peptide was shown to be predominantly α-helical in aqueous solution (as judged by circular dichroism and NMR data) (7;8), yet LD4 was unstructured in solution and forms a helix on binding (7). Hoellerer *et al.* (6) showed that LD2 and LD4 were helical in the crystal structure of LD bound to the FAT domain of FAK, but suggest that the peptides were unstructured alone (65). It is likely that phosphorylation plays a major role in the structure of LD motifs as recent data has shown that phosphorylation of LD4 at serine S272 reduces the binding of the FAT domain of FAK by destabilising the helicity of LD4 (87).

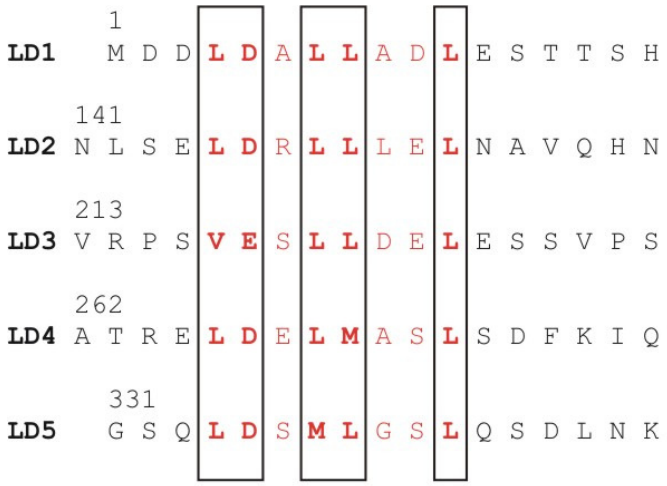


Figure 1.3.2 Paxillin LD motifs (human numbering)

The LD motif is shown in red, the consensus residues in bold.

Paxillin phosphorylation is vital for its correct function. A large number of potential phosphorylation sites have been identified (88) and it is important to determine their function in signalling, the proteins that phosphorylate them and any interaction partners. Two well-characterised phosphorylation sites are tyrosine Y31 and tyrosine Y118 (89;90), mutations in these cause slower FA disassembly (91) and the integrity of these sites is important to promote adhesion turnover (51). The SH2 domains of Crk and CrkL, PI3K and p120RasGap can interact with phospho-tyrosine pY31 and pY118 to regulate cell migration and FA turnover (68;92;93). Other less well-characterised tyrosine phosphorylation sites include tyrosine Y40, Y88 and Y181 (94;95).

There are important serine and threonine phosphorylation sites within the LIM domains (threonine T396/401 in LIM2, serine S455/479 in LIM3) that are involved in regulating paxillin localisation within FAs (72), other sites include threonine T344 and serine S361 within LIM1 and serine S272 and S308 within LD4 and LD5 respectively (88).

There are increasing numbers of proteins being implicated as interaction partners of paxillin. These include the integrin tails of α4, α9, β1 and β3 (96-99) and PABA1 (100;101). For a thorough review of paxillin interactions, the reader is directed to papers by Brown and Turner (68) and Deakin and Turner (102).

1.3.2 Focal adhesion kinase

Focal adhesion kinase (FAK) is a 125kDa cytoplasmic non-receptor protein tyrosine kinase (42;103;104). It is a multidomain protein consisting of a central catalytic kinase domain with non-catalytic N- and C-terminal domains (Figure 1.3.3). One role of the N-terminal FERM domain is to bind the cytoplasmic domain of β1-integrins (97;105) and there is evidence to suggest the FERM domain plays a role in the regulation of localisation and catalytic activity of FAK (106;107). The C-terminus consists of a FAT domain which targets the protein to FAs (108;109) and can bind to paxillin (110) and talin (48). There are three proline-rich sequences, PRR1-3; PRR2 can bind to the SH3 domain of p130Cas (111)

and PRR3 can bind the SH3 domains of GRAF and ASAP1, which are GTPase activating proteins (112;113). In addition there are multiple phosphorylation sites within the domains of FAK (103;105).

FAK kinase-dead mutants retain most of the functions of FAK (114), and its modular structure implicates FAK as an adaptor protein as well as a kinase (115).

FAK localises to early FAs. FAK null fibroblasts show excessive formation of focal adhesions and reduced cell motility (116). Inhibition of endogenous FAK results in reduced motility and cell death (116;117). Conversely, enhancing FAK signaling increases cell motility and can promote cell survival (58;118). Thus, FAK has been implicated in regulating FA turnover, cell motility and cell survival.

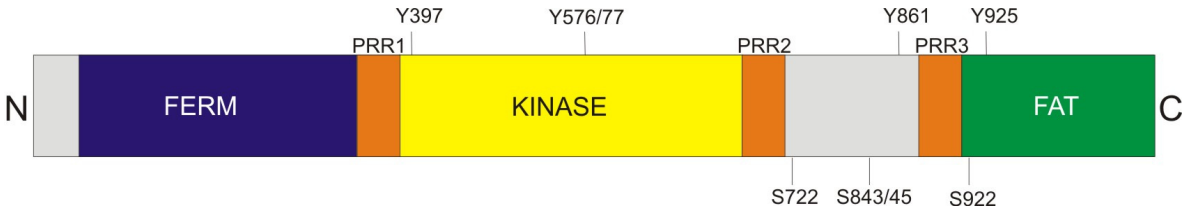


Figure 1.3.3 Domain structure of focal adhesion kinase

Consists of a FERM domain, a central kinase domain and a focal adhesion targeting domain (FAT). There are three proline-rich motifs, PRR1-3 and several phosphorylation sites are shown. FAK autophosphorylates tyrosine Y397 which creates a docking site for the SH2 domain of Src.

FAK is a complex protein to investigate as it can couple to many signalling pathways (Figure 1.3.4). The best characterised of these is the Raf-MEK-ERK signal transduction pathway (42;119), but FAK is also proposed to couple to p130Cas-Crk-DOCK-ELMO (105) to activate Rac1 as well as to PI3K-Akt to promote cell survival (103). FAK also has a role in signalling to the JNK pathway (120;121).

FAK can interact with a number of different partners. FAK autophosphorylation on tyrosine Y397 creates a docking site for SH2 domains of Src-family PTKs. Src binding to FAK promotes further FAK activity by phosphorylating FAK in the kinase domain (109). This FAK-Src signalling complex can phosphorylate proteins like p130Cas and paxillin thus affecting FA dynamics (62;89;90). Phosphorylation of tyrosine Y861 and Y925 in the C-terminus of FAK creates a Grb2 docking site which couples FAK signalling to the Ras-MAPK pathway (103;122). As previously stated, FAK can also couple to other signalling pathways (94) so is an important mediator in regulation of FA dynamics.

FAK is also an important target for cancer therapy. FAK is over-expressed in many tumours and over-expression is related to increased tumour malignancy (123). It is believed that FAK plays a role in tumour proliferation, invasion and enabling anchorage free growth (115). Thus there is great interest in elucidating the signalling pathways and interactions that FAK mediates.

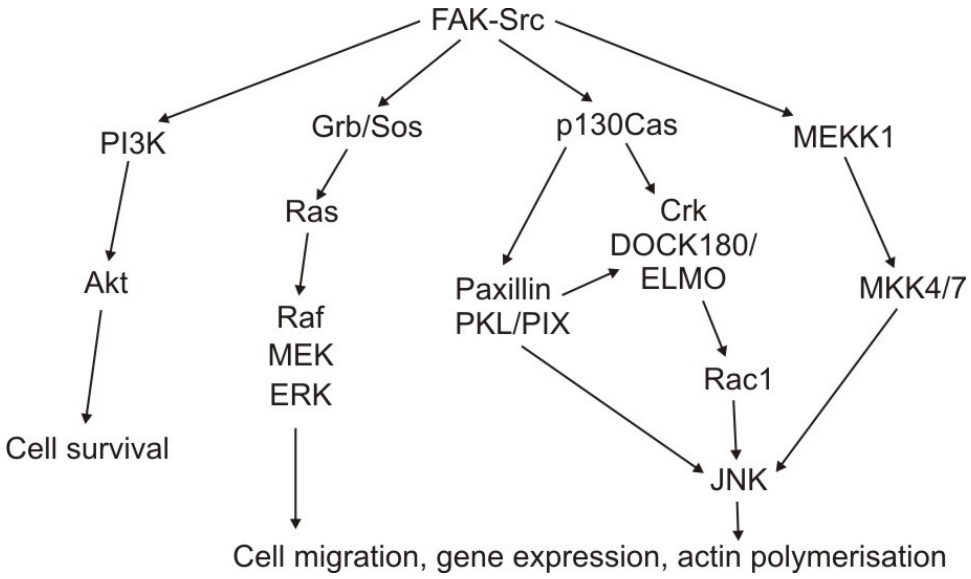


Figure 1.3.4 Summary of integrin signalling pathways in which FAK is involved
FAK plays a role in these signalling pathways after integrin stimulation. The Raf-MEK-ERK pathway is best characterised, the others are less well defined. For more detail on the signal transduction cascades and references to the primary data, refer to (42;103;105;124-127).

1.3.3 Vinculin

Vinculin is a cytoskeletal protein of 116kDa that is found in FAs and cell-cell junctions (128). It was first isolated in the mid-to-late 1970s, with several groups identifying a 130kDa protein from chicken gizzard (129-133), but it appears in the literature as vinculin from 1980 (134).

Vinculin is not absolutely required for FA formation as vinculin null cells still assemble FAs (135). Cells lacking vinculin are less adherent and less well spread when plated on fibronectin than wild-type cells and exhibit increased cell motility and wound closure (136;137). They also show increased FAK activity and increased phosphorylation of FAK, paxillin and p130Cas (128;138). They also have a resistance to apoptosis induced by serum withdrawal and detachment from the cellular substrate (anoikis) (4). Vinculin over-expression reduces cell motility (139). The lack of a defined phonotype for vinculin knockout cells is consistent with the multiple roles vinculin appears to possess. It is a negative regulator of cell motility (137), acts to strengthen FAs (140), plays a role in mechanotransduction (141), regulates FA dynamics (137;142) and can regulate FAK-paxillin signalling and apoptosis (4).

The domain structure of vinculin (Figure 1.3.5) shows that it consists of five domains. D1-D4 form the head domain of vinculin, Vh; and D5 is the vinculin tail, Vt. Vinculin binds to many ligands and is an important regulatory protein within FAs.

The first structural investigations into vinculin showed vinculin as a compact globular molecule (143) or as a roughly globular head attached to an elongated extended tail (144;145). High-resolution imaging by Winkler *et al.* (146) suggested that there were multiple conformations of vinculin; in some molecules an extended tail could be seen that appeared to have four spherical regions that are likened to pearls on a string while in others, the tail could be seen packing close to the head domain. The proline rich linker, which had been predicted from the primary sequence of chick vinculin (147;148) was suspected to provide the flexibility to create these multiple forms.

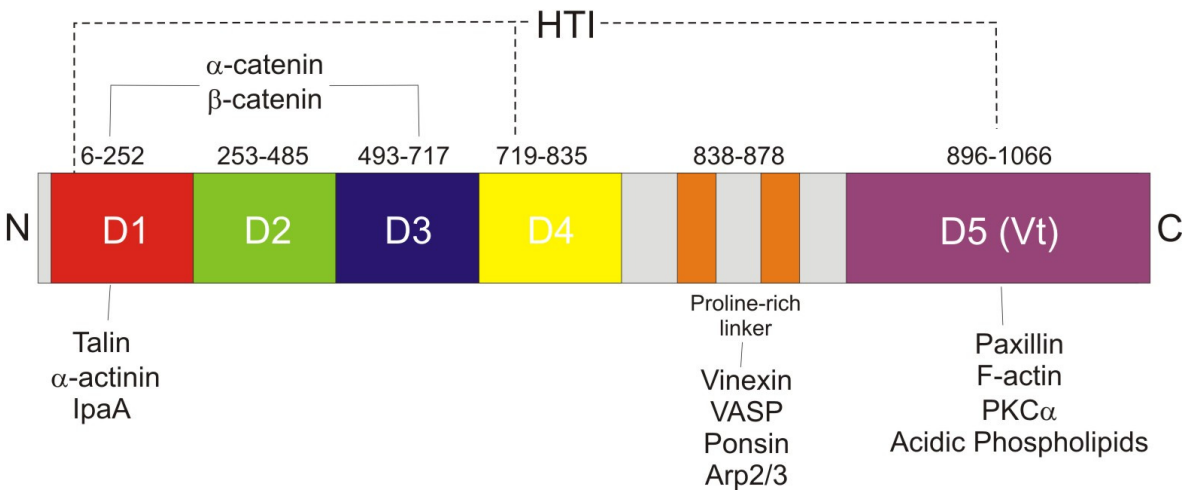


Figure 1.3.5 Domain structure of vinculin (1-1066)

There are five domains, D1 (red), D2 (green), D3 (blue), D4 (yellow) and D5 (magenta). D1 to D4 make up the head domain of vinculin, Vh. D5 comprises the tail domain, Vt. Several ligand binding sites are indicated. An intra-molecular head-tail interaction (HTI) occurs between D1 and D4 with Vt, shown as a dotted line.

Ligand binding studies with vinculin show that there is competition between the globular vinculin head (Vh), vinculin tail (Vt) and their ligands. For example, talin and vinculin tail will compete for binding to the head domain of vinculin (149), α-actinin will bind more strongly to isolated vinculin head than to intact vinculin (150), and vinculin tail will co-sediment and cross link with F-actin, but neither co-sedimentation or cross linking occurred with intact vinculin (151). Consequently, a hypothesis has been proposed where an intra-molecular interaction between the head and tail domain will affect how vinculin can interact with its ligands (149).

Crystal structures of the full-length molecule provide a structural basis for the regulation of vinculin binding to its ligands (152;153). All five domains of vinculin are α-helical, and the tail domain is held between two 'pincers' that are formed by D1 and D3 (Figure 1.3.6). The interaction sites for PIP_2 and F-actin are partially occluded in this structure suggesting that it represents a closed 'autoinhibited' form of vinculin where the ligand binding sites are sterically blocked (41;152). Binding of talin and α-actinin were shown to require

conformational changes in D1; a mutant of D1 which retained the α-helical structure of that in the crystal bound strongly to Vt but not to α-catenin or talin. A conformational change is required to reverse this scenario, showing that ligand binding to vinculin is regulated sterically and allosterically (152).

The K_d for this head-tail interaction (HTI) has been estimated to be below 1nM (152). The principle intra-molecular binding site is between D1 (residues 1-258) and the top of the Vt helical bundle (152-154), yet in solution this D1 domain binds only weakly to Vt, with a K_d of 11.5μM being reported (155). The autoinhibitory site in vinculin was found to be bipartite with a second interaction determined between D4 and Vt (155). The K_d D1-Vt was measured as 10^{-5}M, for D4-Vt was estimated as 10^{-2}M, yet together they create a Vh-Vt complex with a K_d of 10^{-7}M. Once Vt is bound to D1, there will be a large local concentration of residues at the D4-Vt interface to act as a molecular latch that stabilises the autoinhibited state. Tethering of the two domains by the proline linker probably enhances the affinity further.

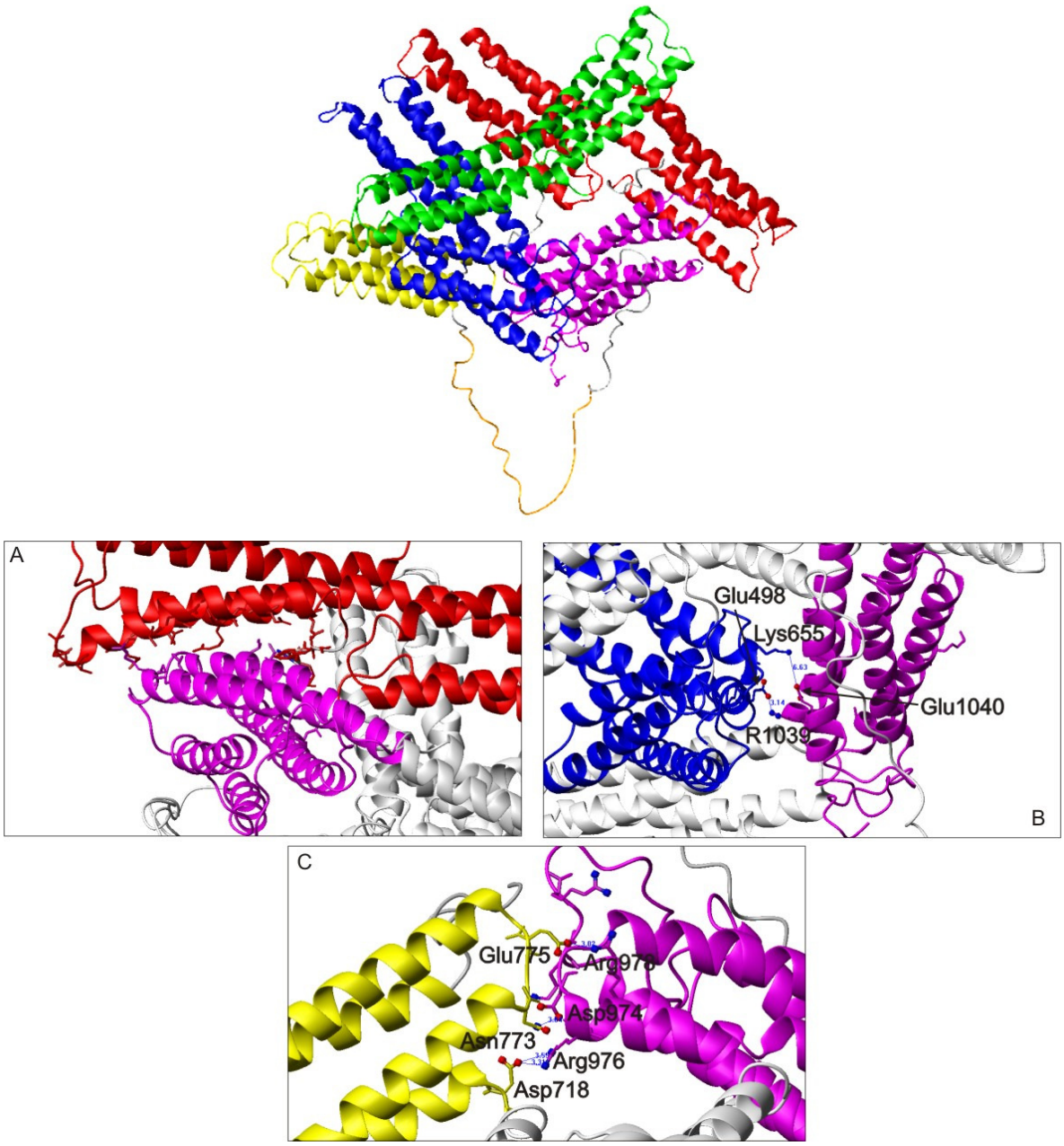


Figure 1.3.6 Crystal structure of full-length autoinhibited vinculin

Top: The five domains are shown, D1 (red), D2 (green), D3 (blue), D4 (yellow) and D5(Vt) (magenta). The proline rich linker region is in orange. Vt is held between D1 and D3 with additional connections from D4. A: there are extensive connections between Vt and D1, 8 hydrogen bonds and 26 hydrophobic interactions between M1, P2, V3, F4, E10, E14, A17, I20, V24, I25, E28, V32, D33, I109, R113 of C1 and G942, K944, R945, I948, T993, Q994, I997, L998, Q1021, M1022, H1025 and N1029 (154). B: the interactions between D3 and Vt (153). C: the interaction between D4 and Vt (155).

The closed, autoinhibited form represents a low affinity binding form where the head interacts with the tail to prevent ligand binding. The flexible elongated form is suggested to represent a high affinity binding form of vinculin where binding sites are available that were sterically and allosterically blocked in the closed form. This head-tail interaction (HTI) hypothesis presents a picture of a dynamic vinculin molecule that exists in multiple forms, a closed autoinhibited form, and active forms where the HTI is disrupted and ligand binding ensues. A cartoon of this is shown in Figure 1.3.7.

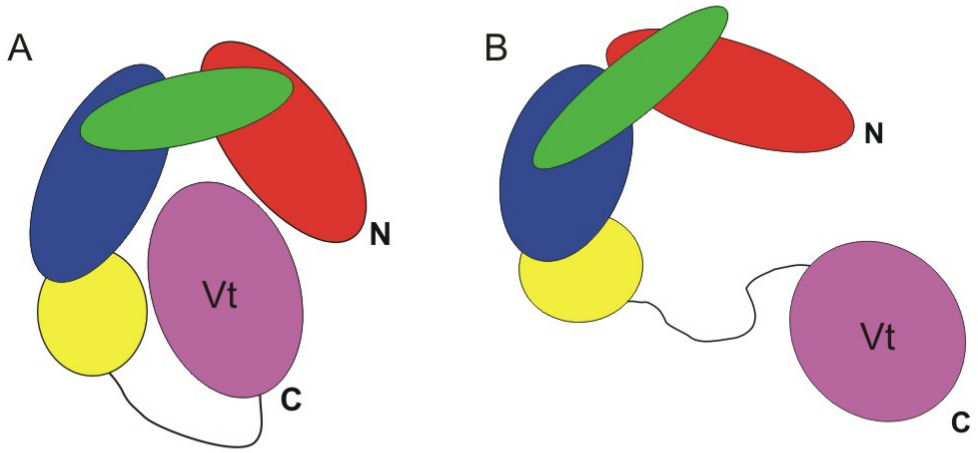


Figure 1.3.7 Cartoon of vinculin in its closed, autoinhibited form (A) and in an open, activated form (B)

In A, interactions between D1 (red) and D4 (yellow) hold Vt in and autoinhibitory state. Upon activation, Vt is released and binding sites are exposed in Vh and Vt.

Direct evidence for an open, activated form and closed, autoinhibited form of vinculin has been provided using FRET (156). This study shows that vinculin exists in an autoinhibited form in the cytoplasm and in an activated form in focal adhesions. Interestingly, there was heterogeneity in vinculin active states amongst different focal adhesions, vinculin can be observed in FAs but be in a closed form; thus vinculin recruitment and activation are distinct processes.

An important aim is to determine the ligands that can recruit and activate vinculin. The hypothesis goes that the vinculin HTI could be regulated by vinculin ligands (149). In

theory, any domain-specific vinculin binding partner could be a candidate for disrupting the head-tail interaction. Table 1.3.2 shows many of the binding partners of vinculin.

The mechanism of activation of vinculin is controversial and not fully understood. The HTI is predicted to have a K_d of $<10^{-9}$M (152) which needs to be disrupted to activate vinculin. One attractive mechanism is that a combinatorial input of ligands may be required to do this (152;155;157). Several molecules may need to bind to the closed form before vinculin is fully activated. This presents a model in which vinculin is able to respond to the spatial co-localisation of its binding partners (152), and means that the activation of vinculin would only occur when the environmental signals were exactly correct, thus providing a tight regulatory mechanism. Ligands that have been implicated in activating vinculin include talin (154;158), F-actin (41) and PIP_2 (159), although the latter now appears to be required for focal adhesion turnover rather than activation of vinculin (142).

The combinatorial input hypothesis proposes that binding of a singular ligand is not sufficient to activate vinculin (152;155;157;160). An alternative model is that the vinculin binding sites of both talin and α-actinin are sufficient to relieve the autoinhibitory property of vinculin (158;161;162). The vinculin binding sites (VBS) within talin are buried and have a low affinity for Vh (152;163;164) but when activated are proposed to be sufficient to activate vinculin (161). Talin VBSs can bind to Vh via a mechanism called helical bundle conversion (154) which causes massive conformational changes within vinculin D1 that displaces Vt (161). Instead of directly competing with Vt, the new D1 conformation is incompatible with binding to Vt and binds talin with higher affinity (165). This alone could be sufficient to activate vinculin. Consequently, it is of great importance to elucidate the mechanism by which vinculin is activated.

Table 1.3.2 Binding partners of vinculin

Ligand	Domain of vinculin	Binding residues on vinculin	Key references
F-actin	Vt	893-985 1016-1066 upper patch : 925, 927, 931, 934, 935, 938-942, 944, 945, 948 lower patch : 962, 965, 966, 969, 970, 977, 1049-1055	(151;166;167) (41)
Paxillin	Vt	PBS1: 952-971 PBS2: 978-1000	(4;86;110)
Talin	Vh (D1)	1-258 two sites? C-terminal helical bundle and N-terminal helical bundle	(154;157;168) (169) (170)
Acidic phospholipids	Vt	916-970 Basic ladder: helix3 C-terminal arm: 1052-1066	(40;171) (142) (172)
PIP_2	Vt	893-985 1016-1066 C-terminal arm	(173;174) (137;152)
α-actinin	Vh (D1)	1-107	(150;162;175)
VASP	Proline rich linker	842-846	(176;177)
Vinexin	Proline rich linker	849-893	(178;179)
Arp2/3	Proline rich linker	850-881	(180)
IpaA	Vh (D1)	1-265	(181;182)
Ponsin	Proline rich linker	837-878	(183)
PKCα	Vt	1016-1066 S1033/S1045 phosphorylated	(184;185)
α-catenin	Vh (D1-D3)	1-718	(152;186)
β-catenin	Vh	1-718	(187)

Vinculin therefore acts as an important regulatory protein that can function as a molecular switch to regulate downstream signalling events (165). With so many diverse interaction partners it is clear that vinculin is a multifunctional protein. Interacting with Arp2/3 suggests a link to actin polymerisation and membrane protrusion (180), binding to F-actin

has been shown to be a crucial link between FAs and the cytoskeleton and may act to stabilise FAs (188), and vinculin also plays a role outside of FAs via its interaction with α- and β-catenin (186).

1.3.3.1 The crystal structure of Vt contains two molecules in the unit cell

The crystal structure of wild-type Vt (172) contains two Vt molecules in the asymmetric unit (Figure 1.3.8) where the N-terminus of one molecule packs against the C-terminus of another molecule. There is also crystal packing between molecules from apposing unit cells, and the authors discuss a crystallographic dimer interface which is centred around the most hydrophobic portion of the molecule at residues isoleucine I997 and methionine M1022 (these residues are indicated in Figure 1.3.8). This interaction is distinct from the N- and C- terminal interactions that occur within the unit cell. An example of the crystal packing can be seen in Figure 1.3.9 where isoleucine I997 and methionine M1022 are involved in a hydrophobic interaction *between* unit cells. These two interactions could potentially impact on solution NMR studies.

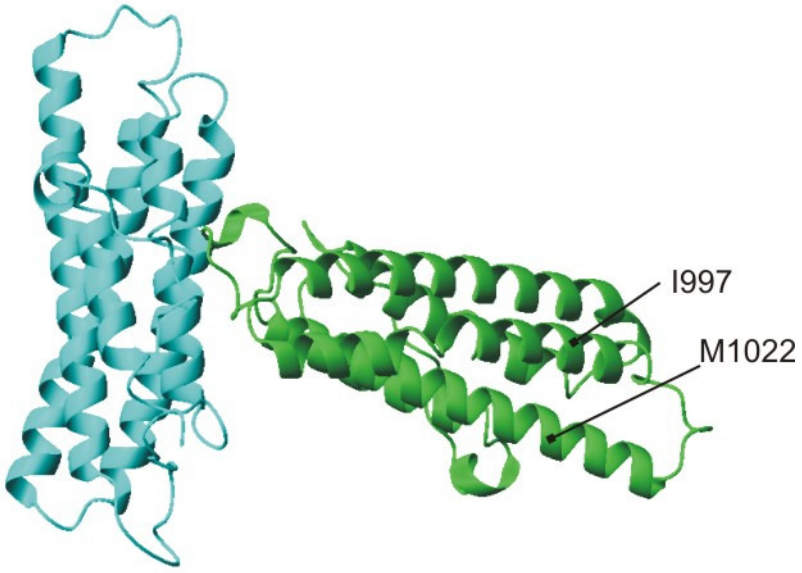


Figure 1.3.8 The crystal structure of Vt has two Vt molecules in the asymmetric unit
There is a potential interaction between the C-terminus of one Vt molecule (green) and the N-terminus and helix 2 of a second Vt molecule (cyan). Residues I997 and M1022 are indicated.

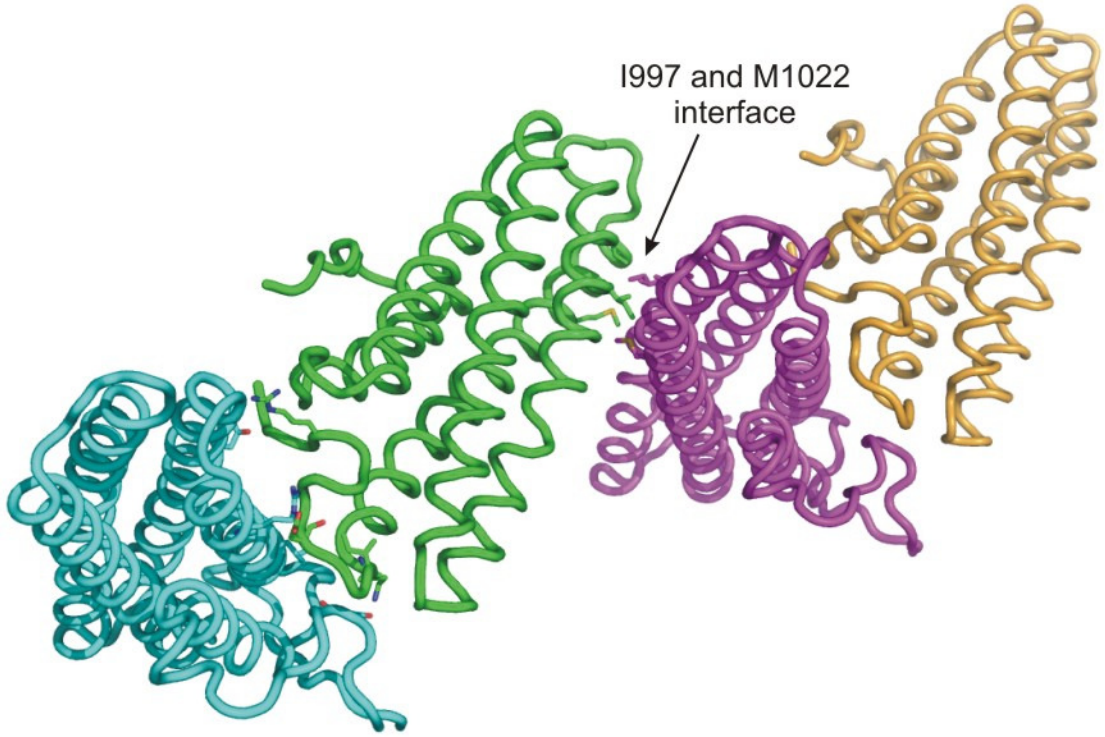


Figure 1.3.9 Crystal packing of Vt molecules in the lattice

The cyan and green molecules are from one unit cell, the magenta and orange are from another. Potential salt bridges are indicated between the cyan and green molecules within a unit cell. The I997 and M1022 hydrophobic interaction occurs between two molecules from different unit cells and is indicated between the green and magenta molecules. Picture kindly prepared by Dr. Stuart Findlow (University of Southampton) using Pymol (DeLano Scientific)..

1.4 Some current knowledge regarding vinculin, paxillin and FAK

In vitro data shows that vinculin and FAK can bind to paxillin LD motifs (5;79;80;110). Interestingly, the vinculin and FAK binding sites overlap, and recent data suggests that FAK and vinculin may compete for binding to paxillin (4) which would affect the downstream signalling events within FAs. The following section attempts to cover some of the knowledge about the three proteins from both cellular and *in vitro* observations.

FAK, paxillin and vinculin are all recruited into early FAs (18;28) and phosphorylation of FAK and paxillin also occurs at early stages in adhesion (19). Both FAK null and paxillin null cells show delayed cell spreading (70;189) implicating them in FA turnover. Expressing FAK in paxillin null cells shows that FAK disassembly from FAs is slower than in control cells (51). Table 1.4.1 summarised the cellular properties of paxillin, FAK and vinculin null cells. Many of the phenotypes seen for paxillin -/- cells are similar to that seen in FAK -/- cells, whereas the vinculin null phenotypes are generally different (with the exception of cell spreading). Whilst these observations provide clues to the function of these proteins in FAs, their precise roles are difficult to elucidate as the molecular basis behind these phenotypes are complex.

Table 1.4.1 Phenotypes of null and knockout cells

Paxillin null cells or paxillin deficient cells	FAK null cells or FAK deficient cells	Vinculin null cells or vinculin deficient cells
Delayed cell spreading (70;190)	Delayed cell spreading (105)	Delayed cell spreading (136;137)
Less motile (70)	Less motile (116;117)	More motile (137)
Abnormal FA formation (70)	Excessive FA formation (116)	Fewer, smaller FAs (137)
Unknown	Promotes cell death (116;117)	Promotes cell survival (4)
Adhere same as wt (70)	Adhere same, poorly spread (116)	Less adherent and poorly spread (136;137)
Embryonic lethal (70)	Embryonic lethal (103)	Embryonic lethal (128)

1.4.1 The FAK-paxillin interaction

The precise role of the FAK-paxillin interaction is not fully understood. A direct interaction between FAK and paxillin was detected *in vitro* using GST pull-down experiments and immunoblotting (79;110). The N-terminus of paxillin was shown to interact with the C-terminal FAT domain of FAK, and there were two discrete regions in FAK that supported this binding, PBS1 (FAK 919-935) and PBS2 (FAK 1034-1039) (110). Within paxillin, regions in the N-terminus containing LD2 and LD4 were shown to bind to FAK and mutations in both LD2 and LD4 are necessary to decrease binding of paxillin to

FAK (5;191). Subsequently, the binding sites of LD peptides have been mapped onto the structure of the FAT domain of FAK; this has shown that LD motifs will interact at two sites in the FAT domain (6;7;9). This will be discussed in more detail in Section 1.5.

One potential role of the FAK-paxillin interaction is to direct tyrosine phosphorylation of paxillin (192). The paxillin binding site in FAK has been found to be necessary for cell spreading (91). In this study, inhibition of paxillin tyrosine phosphorylation could be rescued by both wild-type FAK and catalytically inactive FAK, showing that the kinase activity of FAK was not responsible for tyrosine phosphorylation of paxillin. However, when FAK was expressed with a deficient Src kinase binding site (Y397F), the phosphorylation of paxillin was not restored. This suggests that the interaction of Src with FAK is necessary for phosphorylating paxillin. As both Src binding and paxillin binding to FAK is necessary, a model was suggested where FAK simultaneously binds to Src and paxillin to bring them into close proximity for phosphorylation to occur which promotes cell spreading.

Tyrosine phosphorylation of paxillin through a FAK/Src complex is important in its function as an adaptor molecule (43;89;90;94), and while FAK binding is not essential for tyrosine phosphorylation of paxillin, association with FAK may enhance its level of tyrosine phosphorylation (191). Two important phosphorylation sites on paxillin are tyrosine Y31 and tyrosine Y118 (89;90). If paxillin null cells are transfected with paxillin that has mutated tyrosine Y31 and tyrosine Y118 phosphorylation sites, disassembly of FAs is slower compared to control cells (91). It has been shown that paxillin is phosphorylated in dynamic adhesions and that the integrity of the tyrosine Y31 and Y118 paxillin phosphorylation sites was important to promote adhesion turnover (51).

A study by Zaidel-Bar *et al.* (25) studied adhesion complexes at different stages of maturity; early forming focal contacts (FX), focal adhesions (FA) and fibrillar adhesions (FB). They found that phosphorylation states of paxillin between adhesions varied, with the highest levels of phospho-paxillin found in early forming FXs, variable levels in FAs and no phospho-paxillin in FBs. It was shown that phospho-paxillin was required for FA

assembly, where it can then bind to FAK to initiate turnover. Phospho-paxillin bound more strongly to FAK than paxillin that was not phosphorylated. A constitutively phosphorylated paxillin (where Y31 and Y118 had been mutated to glutamate) enhanced the number of FX and increased the turnover of FXs, confirming the role of phospho-paxillin in assembly and turnover. In contrast the number of FAs was not changed but turnover of these was increased.

Thus, the phosphorylation state of paxillin is important in FA dynamics and this is in part due to the interaction with FAK. Downstream consequences of phosphorylating paxillin Y31 and Y118 generates SH2 binding sites for Crk adaptor proteins which can couple to the protein p130CAS. Alternatively, phospho-paxillin can interact with raf-MEK-ERK or with Grb/Sos-ras to couple to the MAPK/ERK signalling pathway. Cell migration, cell invasion and cell survival has been shown to be regulated by these two pathways (124).

The interaction between FAK-Src-paxillin-Crk-p130CAS may also couple to the rac-PAK-JNK signalling cascade which is important in mediating cell motility signals (93;193). MAPK/ERK signalling is equally important as ERK can act to serine phosphorylate paxillin on serine S83 (126;127;194;195) which may be important for increased FAK-paxillin interactions (196). Phosphorylated paxillin can also couple to Csk which is a negative regulator of Src (197).

1.4.2 The vinculin-paxillin interaction

The role of the vinculin-paxillin interaction within cells is poorly understood. Observations in vinculin null cells indicate that the paxillin and FAK interaction and the level of paxillin and FAK phosphorylation is enhanced compared to wild-type cells (4). Vinculin null cells have more dynamic FAs (137), and vinculin is involved in stabilising FAs and preventing turnover (140;141) although this may be an effect unrelated to or distinct from the vinculin-paxillin interaction. Vinculin recruitment to FAs has been shown to increase the strength of an adhesion in response to force (140). One mechanism of relieving this may be due to

PIP_2 (137;142), and it also appears necessary to disrupt the talin-vinculin interaction to cause FA turnover (157).

Paxillin was identified as a vinculin binding protein in 1990 by Turner and colleagues (67). The paxillin binding site was found to reside in the tail domain of vinculin (Vt), and using iodinated GST-fusions the region was delineated to Vt 978-1000 (86). GST pull-down experiments between GST-paxillin and chicken gizzard crude lysate show that both vinculin and FAK can be pulled down (79).

A novel motif on paxillin was characterised by Brown *et al.* (5) that binds to vinculin. C-terminal truncated GST-paxillin constructs were incubated with smooth muscle lysate and western blots were carried out with antibodies to vinculin. Paxillin regions spanning 143-164 were shown to be necessary for binding and deletion of this region from full length paxillin results in no binding to vinculin. Analysis of this region in paxillin identified 13 amino acid stretches of conserved residues which were named LD motifs. LD motifs 1-4 were originally identified (5), a fifth LD motif was later predicted by virtue of its binding to papillomavirus E6, and was renamed LD3 (85). The sequences of the LD motifs can be seen in Figure 1.3.2. Vinculin has been shown *in vitro* to interact with LD1, LD2 and LD4 (5;80).

Cellular evidence of the role of paxillin-vinculin interaction is scarce. They can co-localise with one another in FAs (18;67;138;198), but a direct interaction between the two *in vivo* has not been identified to date. Vinculin appears not to affect the residency time of paxillin in FAs, and vinculin conformation may not be important in paxillin dynamics (157). It is also not responsible for recruitment of paxillin to FAs (5).

In a study by Bubeck *et al.* (199), when the binding activities of vinculin were studied by targeting Vt to mitochondria, Vt recruited F-actin but not paxillin. The failure of Vt to recruit paxillin was unexpected. More recently, Humphries *et al.* (188) found that while vinculin head and paxillin co-localised in protruding lamellipodia, vinculin tail and paxillin did not. Some co-localising was seen in retracting areas between Vt and paxillin. Again,

while this result was unexpected, it provides evidence contrary to a vinculin-paxillin interaction *in vivo*. In the study by Humphries *et al.* (188) much of the Vt in the cells was associated with actin, and this may explain why the paxillin-vinculin interaction could not be identified as the two binding sites may be mutually exclusive. Similarly, it may be that the cellular situation was not appropriate for detecting this interaction. But it remains a possibility that while paxillin and Vt will interact *in vitro*, within cells the interaction may not be a functional one.

An interesting study by Subauste *et al.* (4) provides convincing evidence that vinculin can modulate the FAK-paxillin interaction to control cell motility and apoptosis. Vinculin null cells are resistant to apoptosis induced by serum withdrawal and detachment from the cellular substrate (anoikis). Vt (811-1066) was shown to restore cell death signals to wild-type levels. In vinculin null cells, ERK1/2 activity was elevated in response to detachment from the substrate and Vt rescued these cells and restored ERK1/2 activity to wild-type levels. ERK antisense oligonucleotides and the MEK inhibitor PD98059 could also restore apoptosis. This data suggests that Vt is modulating the ERK1/2 pathway (4) .

In the same study, the paxillin-FAK interaction was elevated in vinculin null cells and Vt rescue cells returned the interaction to a wild-type level. The FAK-paxillin interaction is important for phosphorylation of FAK at Y397 and paxillin at Y118 and Y31 (192;200). Vinculin null cells have increased phosphorylation of these residues, and upon detachment the levels of phospho-tyrosine pY397 and pY118 rapidly decreased. In contrast, vinculin null cells rescued with Vt show wild-type levels of phosphorylation both constitutively and during anoikis (4). Overexpression of paxillin mutated at residues phosphorylated upon FAK interaction, Y31F/Y118F (192), restored anoikis in vinculin null cells and also inhibited ERK1/2 activity during anoikis (4).

Similar effects were seen for cell motility, vinculin null cells were more motile and the MEK inhibitor decreased the rate of cell motility. The paxillin Y31F/Y118F mutant also inhibited the enhanced migration in vinculin null cells. Taken together, the data points to a role of Vt in modulating the interaction of FAK and paxillin to regulate ERK1/2 signalling

cascades. The hypothesis goes that FAK and Vt will compete for binding of the LD motifs in paxillin (4).

Paxillin will interact with FAK via LD2 and LD4, while vinculin will interact with LD1, LD2 and LD4. The above hypothesis presumes that binding to the LD motifs is mutually exclusive. It is feasible to imagine that the competition between FAK and vinculin for paxillin could be due to a number of factors, such as relative affinities for binding sites, conformation of proteins involved, phosphorylation states of proteins and relative local concentrations of binding partners.

1.5 Structural basis of study

Great effort has been made to investigate the interaction between vinculin and talin (149;154;158;163;164;169;201-203) and vinculin and F-actin (41;151;166;167;204), but the vinculin-paxillin interaction has seen much less attention. This thesis builds on previous structural data regarding the FAT domain of FAK binding to paxillin (6-9). Paxillin LD motifs have been shown to bind to the FAT of FAK and Vt using biophysical binding assays (5;79;80;110), and alongside the structural knowledge of the binding event between FAK and LD motifs, predictions have been made as to the binding event occurring between Vt and LD motifs (6).

1.5.1 FAT of FAK and Vt are structural homologues that both bind to LD motifs

The crystal structures of the FAT domain of FAK (205) and Vt (172) are shown in Figure 1.5.1. For simplicity, the FAT domain of FAK will be referred to from now on as FAT. FAT and Vt have been shown to be structural homologues (205). The α-helices in FAT are straight and antiparallel connected by short turns which makes the structure very compact and symmetrical (205). In contrast the α-helices in Vt are less parallel to one another and helix 4 is kinked resulting in a less compact form. Vt helices 1-4 are basic and helix 5 is acidic (172). The structure of the Vt domain alone has subsequently been shown to be very

similar to the structure of Vt in the full length molecule (152;153). Recently the structure of another structural homologue, Git1 has been solved (10) which can also be seen in Figure 1.5.1 and shares the antiparallel helix topology.

Git1 can bind to paxillin LD motifs LD2 and LD4 (10;80;82), FAT will bind to LD2 and LD4 (5-9) and Vt can bind to LD1, LD2 and LD4 (5;80). Since all these proteins share structural features it is possible that the mode of binding to LD motifs is similar.

This thesis aims to determine how Vt interacts with paxillin LD motifs and then compare this to what is known from the FAT binding data. A comparison will then be made between Vt, FAT and the new data from Git1 binding in an attempt to understand the features that determine binding and specificity for LD motifs.

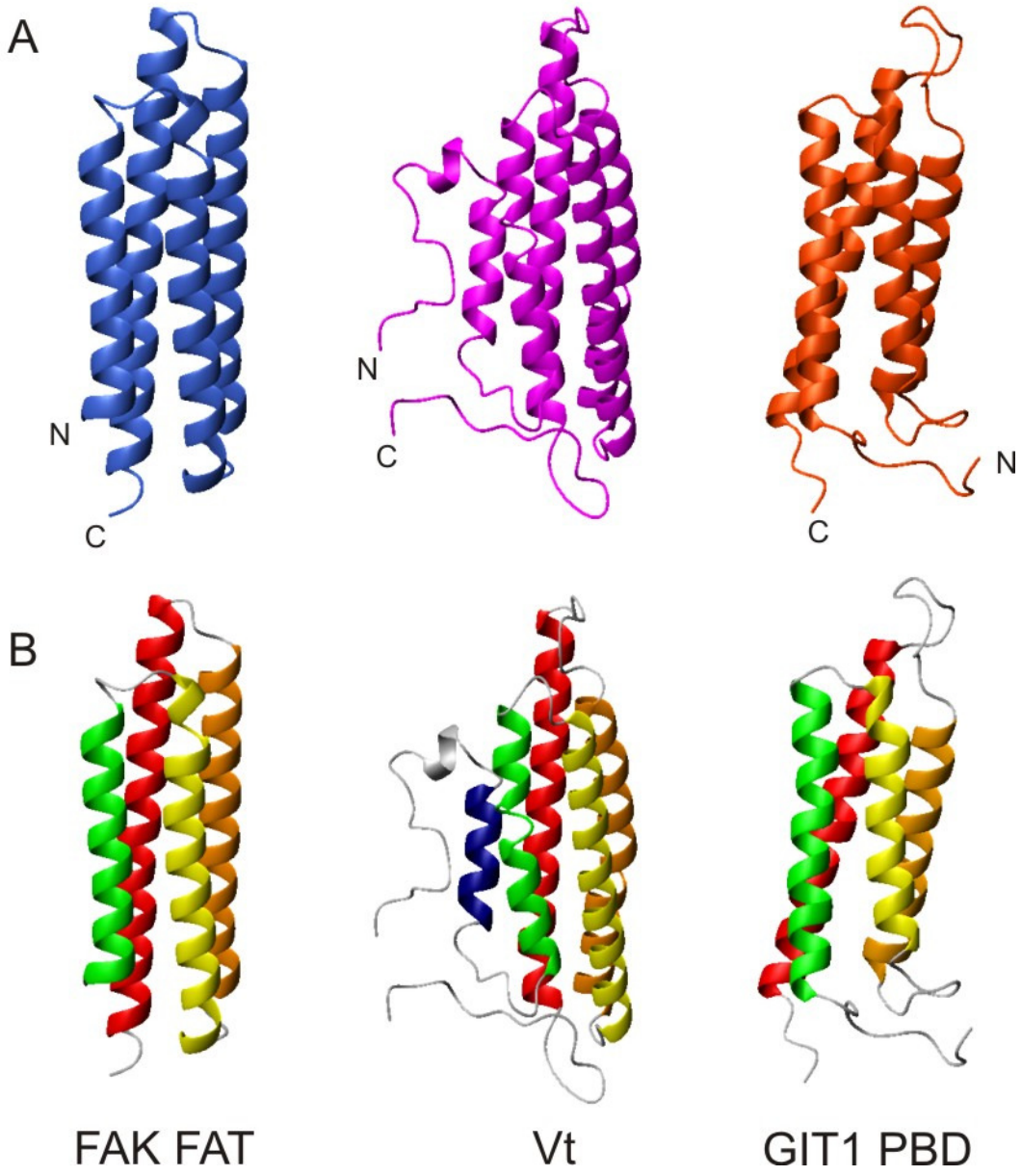


Figure 1.5.1 Crystal structures of the FAT domain of FAK, Vt and the PBD of Git1

A: crystal structure of FAT (blue) showing 4-helix up-down antiparallel topology (205), crystal structure of Vt (magenta) showing 5-helix up-down antiparallel topology (172) and solution structure of GIT1 PBD (molecule 0001-orange-red) showing 4-helix up-down antiparallel topology (10). B: comparison of the helices in the two structures. Green: FAT H1, Vt H2, PBD H1; yellow: FAT H2, Vt H3, PBD H2; orange: FAT H3, Vt H4, PBD H3; red: FAT H4, Vt H5, PBD H4. Vt has an extra helix H1 shown in dark blue.

1.5.2 LD motif binding and specificity

While the individual domains of paxillin have been intensively studied, there is no structural data on the full length molecule. The LD motifs have been predicted to be α-helical (85) and the LIM domains are double zinc finger motifs, but it is possible that paxillin is relatively unstructured and becomes more organised upon binding and/or modifications such as phosphorylation (206).

The LD motif has been defined as a novel binding module involved in protein-protein interactions (5;85), but it remains unclear precisely how they interact with their target proteins. Analysis of GST- LD fusions and their propensity for binding FAT and vinculin show that LD2 binds to vinculin and LD2 and LD4 binds to FAT (5) (Table 1.5.1.). Further analysis reveals that there is strong binding between vinculin and GST-LD2, low binding to GST-LD4 and binding to GST-LD1 (80). These combined results tell us that FAT will bind to paxillin LD2 and LD4 while vinculin will bind to paxillin LD1, LD2 and LD4. As previously stated, GIT1 can bind to paxillin LD2 and LD4 motifs (10;80;82).

Table 1.5.1 Summary of paxillin LD interaction with FAT domains

Paxillin LD motifs	Vinculin binding[(a)]	FAK binding[(a)]
LD2-LD5	++++	++++
LD2-LD3	++	+ (++)[(b)]
LD2	++	+
No LD	-	-
LD3-LD5	-	+++
LD3	-	-

(c)	Vt	FAT	Git1
LD1	✓		
LD2	✓	✓	✓
LD3			
LD4	✓	✓	✓
LD5			

(a) results from (5). (b) two separate fusions spanning LD2 and LD3 were tested for FAT, the first paxillin 54-265 showed + binding and the second paxillin 133-265 showed ++. The reason for this is unclear. (c) summary of LD binding for Vt, FAT and Git1.

The region of FAT that binds to paxillin was initially suggested to consist of two regions, called PBS1 (FAK 919-935) and PBS2 (FAK 1034-1039) (110). These sites create a binding interface on helix 1 and helix 4 of FAT (face 1-4). PBS1 and PBS2 were also

suggested in Vt by virtue of sequence alignment with the experimentally determined regions of FAT. PBS2 had previously been determined by Wood *et al.* (86) as Vt residues 978-1000, and PBS1 was predicted to reside in Vt 952-971 (110). PBS1 and PBS2 create a binding interface on Vt face 3-4 (Figure 1.5.2).

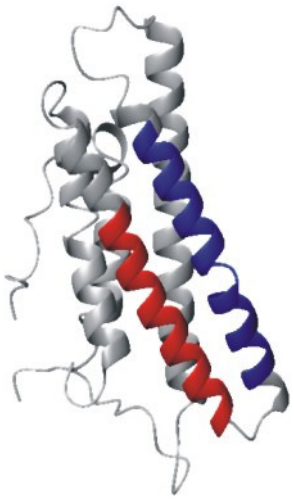


Figure 1.5.2 Crystal structure of Vt showing paxillin binding subdomains 1 (PBS1:red) and 2 (PBS2:blue) (86;110)

PBS2 is experimentally determined by in vitro binding assays, PBS1 is predicted by virtue of sequence alignment with FAT

More recently there have been several studies that map the regions of FAT that bind to paxillin using NMR and X-ray crystallography (6-9). Figure 1.5.3 shows a cartoon of the FAT domain bound to two LD peptides taken from (6). Two sites for binding are shown, on opposite faces of the FAT domain. One site is formed on FAT face 1-4 and a second site is on FAT face 2-3. Both peptides were shown to bind simultaneously.

There is some debate as to whether the LD motif is promiscuous. Hoellerer *et. al*. (6) showed that both LD2 and LD4 would bind equally to either site thus the LD peptides were promiscuous binding motifs. In contrast, Bertolucci *et al.* (7) suggests that LD2 will bind to both faces of FAT, but LD2 favours face 1-4 and has a higher affinity for this site. LD4 would only bind to face 2-3, and LD4 has a higher affinity for this site than LD2. A recent

study appears to confirm the latter which suggests that there is specificity between LD motifs (207).

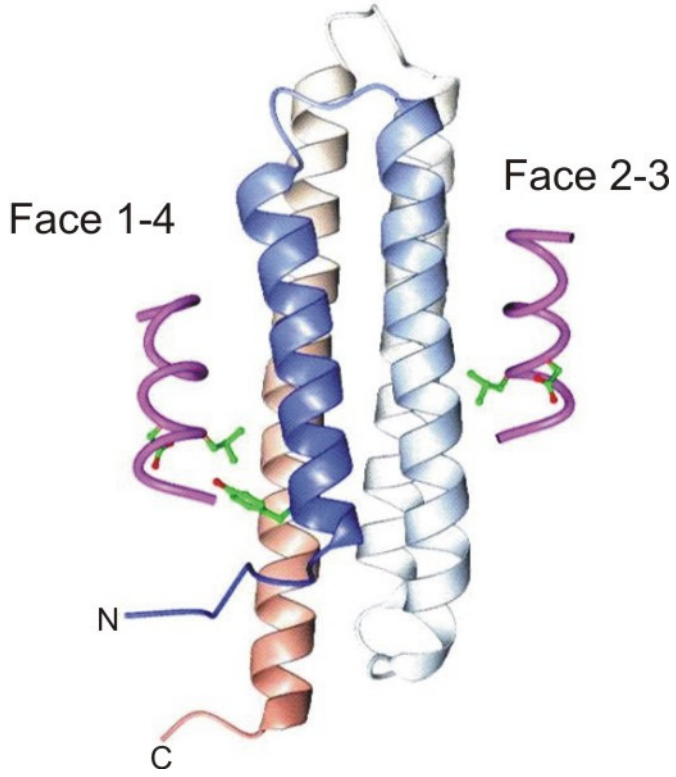


Figure 1.5.3 The FAT of FAK with two LD4 peptides bound

FAT of FAK is shown, helix1 in blue, helix 2 light blue, helix 3 white and helix 4 red. LD4 peptides are shown in magenta, with the first LD of the motif shown in green on both peptides and Y925 (Grb site) shown on FAK. Picture reproduced from (6).

Both FAT and Vt will bind to LD2 and LD4, with Vt also showing binding to LD1. Structure based sequence alignment suggested conservation of a number of residues between FAT and Vt that were involved in the FAT-LD interaction (6). These structural similarities between FAT and Vt make it possible that Vt would bind to LD motifs in the same manner, by forming two separate binding sites on opposite sides of the Vt molecule. The alignments made by Hoellerer *et al.* (6) are shown in Figure 1.5.4. and reveal the predicted residues for the Vt-paxillin LD interaction.

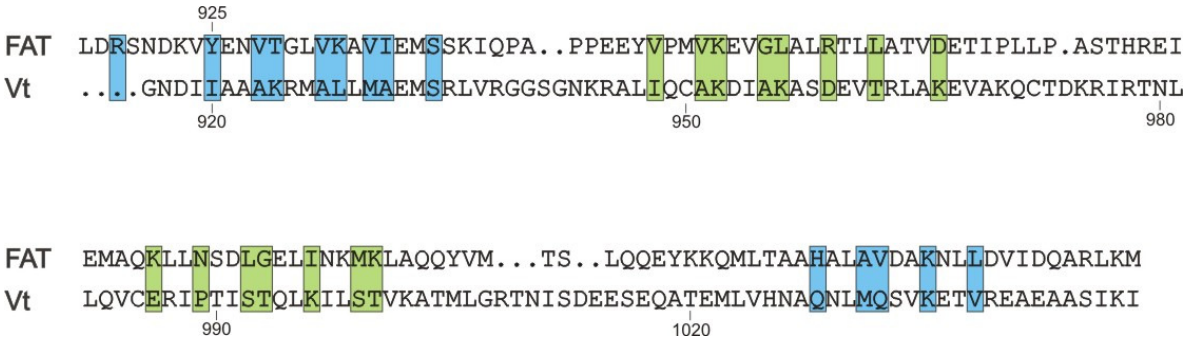


Figure 1.5.4 Sequence alignment of FAT and Vt (human numbering) aligned based on their crystal structures as in (6)

Shown in boxes are the experimentally determined residues affected by LD peptides binding to FAT (6) and the corresponding predictions for the sequence of Vt. Blue boxes refer to residues on FAT face 1-4 and green refer to FAT face 2-3.

Mapping these residues onto the crystal structure of Vt reveals the two faces that could be affected by LD motifs (Figure 1.5.5). The binding interface formed by helix 3-4 is potentially accessible for binding, yet the face formed by helices 5-2 is occluded by helix1. *Hoellerer et al.* (6) state that Vt face 5-2 resembles the binding interface of FAT whereas Vt face 3-4 does not.

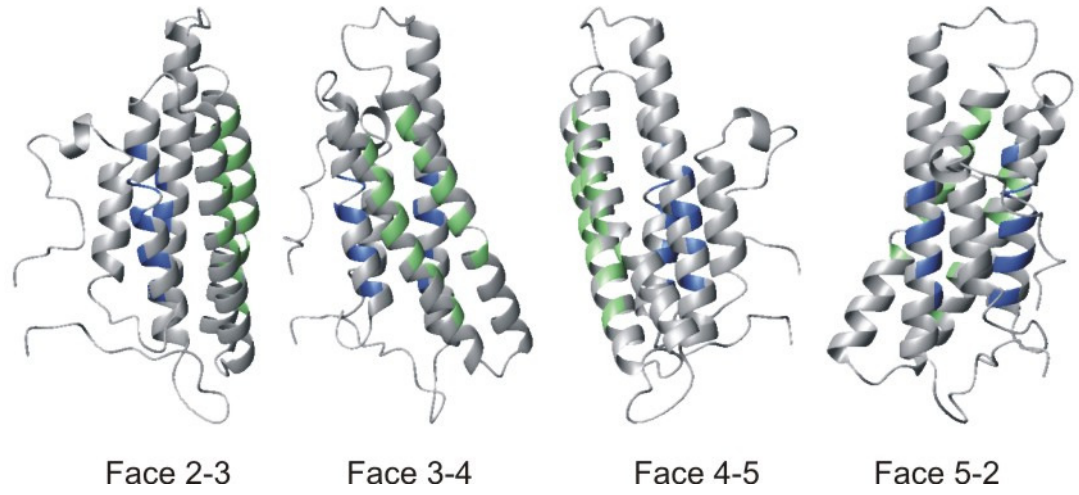


Figure 1.5.5 Predicted binding residues on Vt based on experimental data using the FAT domain of FAK binding to LD2 and LD4 (6)

Pale green residues make up binding interface Vt 3-4 (FAK 2-3) and pale blue make up face Vt 5-2 (FAK 1-4).

This thesis aims to verify the interaction between Vt and paxillin LD motifs *in vitro*. The molecular basis of the interaction between Vt and LD1, LD2 and LD4 will be analysed and compared to the experimentally determined PBS1 and PBS2, (which create a one site binding interface on Vt face 3-4), and the predicted binding residues from Hoellerer *et al.* (6) (which create two binding interfaces on opposite faces of Vt, face 2-3 and 3-4).

The interaction studies will primarily be carried out using high resolution NMR to study the precise molecular nature of this interaction. After this, it should be possible to carry out a detailed comparison between FAT and Vt, and also using the recent GIT1 data (10) to analyse the mechanism of binding to paxillin LD motifs.

Being able to dissect these interactions *in vitro*, and on a molecular scale, is highly useful as it facilitates directed studies of *in vivo* functions, for example specific domains can be produced with pre-determined mutations and then the effects within FAs and on cells can be analysed explicitly. Because of the complexity of FA signalling and mechanotransduction, *in vivo* studies can be difficult to interpret fully as they can be looking at averages of interaction partners and at defined time-points within the cell. To this end *in vitro* experiments can be very useful in exploring all possible interactions which can then be specifically searched for in an appropriate form of FA or in particular cell types.

2 Materials and Methods

2.1 Recipes

2.1.1 Protein expression and detection

Luria Broth (1L) pH7.5

10g Bactotryptone

5g Yeast Extract

10g NaCl

2mL Glycerol

0.27g $MgCl_2.7H_2O$

Minimal M9 Media (1L) pH7.4

6g Na_2HPO_4

3g KH_2PO_4

0.5g NaCl

0.5g $MgSO_4.7H_2O$

15mg $CaCl_2.2H_2O$

After autoclaving:

Add ^{15}N NH_4Cl at 1g/L

Add ^{13}C Glucose at 2-4g/L, add unlabelled glucose at 4g/L

Triple Labeled Minimal M9 Media (1L) pH7.4

Made up in 70% D_2O, 30% H_2O

6g Na_2HPO_4

3g KH_2PO_4

0.5g NaCl

0.5g $MgSO_4.7H_2O$

15mg $CaCl_2.2H_2O$

After autoclaving:

Add ^{15}N NH_4Cl at 1g/L

Add ^{13}C Glucose at 2-4g/L

1mL/L of 1mg/mL biotin

1mL/L of 1mg/mL thiamine

Make final solution 10% Silantes CND medium

IPTG

1M stock: 4.76g IPTG made up to 20mL using double distilled water. Sterile filter using a 0.22µm syringe filter. Freeze at -20°C in 1mL aliquots.

Ampicillin

100mg/mL stock: 2g ampicillin in 20mL of double distilled water. Sterile filter using a 0.22µm syringe filter. Freeze at -20°C in 1mL aliquots.

LB-agar

To 100mL LB add 1.5g powdered agar. Autoclave.

Disruption buffer/reducing sample buffer (10mL)

2.5mL 1M Tris-HCl pH6.8

4.0mL 20% SDS

0.5mL 0.4% Bromophenol blue

2.0mL Glycerol

1.0mL ddH2O

860µl β-mercaptoethanol

SDS-PAGE

10% gel

	Stacking gel	Running gel
Protogel 33% acrylamide soln.	0.66mL	3.33mL

Water	2.90mL	2.03mL
Glycerol	-	1.30mL
3M Tris, 0.3% SDS pH 8.45	1.24mL	3.33mL
TEMED	10µl	10µl

Just prior to pouring, make up ammonium persulphate at 100mg/mL in about 1mL of water. To the running gel add 50µl and to the stacking gel add 25µl.

Running buffers

Cathode (1L)

100mM Tris pH8.2

100mM Tricine

0.1% SDS

Anode (1L)

200mM Tris pH8.9

Run SDS-PAGE gels at 120V for 1 hour

Coomassie Blue (1L)

0.2% Brilliant Blue

40% methanol

10% acetic acid

50% ddH_2O

Dissolve the brilliant blue into methanol first, then add the other ingredients

2.1.2 Molecular biology

Pfu polymerase and buffer, dNTPs, Dpn1 and buffer B were obtained from Promega Primers for PCR were designed by R.Hagan and obtained from MWG Biotech. Sequences can be found in Appendix 5.3.

1% Agarose gel

40mL TBE + 0.4g agarose. Microwave for 1 minute. Add 3µl Ethidium Bromide (10mg/mL). Run at 120V for 30 minutes.

For samples, add 14µl DNA loading buffer, 6µl sample.

1x TBE (1L)

10.8g Tris pH8.3

5.5g Boric acid

0.93g EDTA

10x DNA loading buffer

50% Glycerol

0.4% Bromophenol Blue

0.4% Xylene Cyanol

SOC media (100mL)

2% Bactotryptone (2g)

0.5% Yeast extract (0.5g)

10mM NaCl (0.6g)

2.5mM KCl (0.02g)

10mM $MgCl_2$ (0.2g)

10mM $MgSO_4$ (0.25g)

20mM Glucose (0.36g)

2.1.2.1 Fusion protein constructs

Wild-type vinculin (residues 879-1066) in pET-15b was a kind gift from Professor D. Critchley (University of Leicester). The plasmid contains the wild-type gene with an N-terminal histidine affinity tag and a thrombin cleavage site. After thrombin cleavage, the protein contains an extra GSHM at the amino terminus. Mutants have been produced from this wild-type template. Protein sequences can be found in Appendix 5.2.

Constructs of pGEX-4T-1-LD1/LD2 and pET-15B-LD1/LD2 were obtained from Prof D. Critchley (University of Leicester) and were transformed into *E.coli* BL21 (DE3) cells. Sequence of paxillin His-LD1/LD2 containing the histidine tag and extra N-terminal residues can be found in Appendix 5.4.

2.2 Expression and purification methods

2.2.1 Standard expression of wild-type and mutant vinculin tail

The plasmid was expressed in *E.coli* BL21 (DE3) cells. Wild-type Vt was streaked onto an agar plate containing ampicillin at 100μg/mL and incubated overnight at 37°C. Overnight cultures were prepared by taking 15mL LB with added ampicillin at 100μg/mL and inoculating this with one colony from the agar plate. This was grown overnight at 37°C, 160rpm.

2.2.1.1 Standard culture growth

1.2L of LB media was prepared and 400mL aliquotted into three 2L baffled flasks. The flasks were covered with aluminium foil and autoclaved. Prior to use, ampicillin at 100μg/mL was added. 5mL of overnight culture was used to inoculate each 2L flask, and the resulting cultures were grown at 37°C, 160rpm. When the OD_{600} reached ~0.8, flasks were treated with IPTG (final concentration 1mM) and the cells grown for a further 3-4 hours at 30°C, 160rpm. The cells were pelleted using a Beckman J2-21 centrifuge with a JLA 10.500 rotor at 7000rpm (6037g) for 25 minutes, resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use.

2.2.1.2 ^{15}N Minimal media growth

1.2L of M9 minimal media was prepared and 400mL aliquotted into 3 x 2L baffled flasks. The flasks were covered with aluminium foil and autoclaved. Prior to use, ^{15}N ammonium

chloride was added at 1g/L and glucose was added at 4g/L. 1.2g of ^{15}N ammonium chloride was dissolved in 15mL ddH_2O, 0.22µm-filtered and 5mL dispensed into each flask. 4.8g of ^{12}C glucose was dissolved in 90mL ddH_2O, 0.22µm-filtered and 30mL dispensed into each flask. Ampicillin was added to each flask at 100µg/mL. 30mL of LB overnight culture was pelleted at 900g for 15mins and the cells resuspended in 30mL M9 minimal media. 10mL of the cell suspension was used to inoculate each flask, and the cultures grown at 37°C, 160rpm until OD_{600} ~0.6-0.8. At this point, flasks were treated with IPTG (final concentration 1mM) and the cells grown for a further 16-18 hours at 30°C. The cells were pelleted at 6037g for 25 minutes, resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use.

2.2.2 ^{15}N, ^{13}C, ^{2}H Triple labelled expression of Vt/I997S

2.2.2.1 Partial deuteration

In order to adapt the BL21 (DE3) cells to $^{2}H_2O$, overnight cultures were prepared in the following way. 10mL of LB with ampicillin at 100µg/mL were inoculated using a colony from a Vt/I997S agar plate and the culture grown at 37°C, 160rpm, until OD_{600} ~0.8. Five sterilins containing 10mL LB made with 50% $^{2}H_2O$ (with ampicillin at 100µg/mL) were inoculated using 1mL of the culture and incubated overnight at 37°C, 160rpm.

2L of M9 minimal media were prepared using 70% $^{2}H_2O$, 30% $^{1}H_2O$ split equally between five 2L baffled flasks, covered with aluminium foil and autoclaved. Prior to use, ^{15}N ammonium chloride and ^{13}C glucose were added at 1g/L and 4g/L respectively. (2g of ^{15}N ammonium chloride was dissolved in 5mL ddH_2O, 0.22µm-filtered and 1mL added to each flask; 8g of ^{13}C glucose was dissolved in 50mL ddH_2O, 0.22µm-filtered and 10mL was added to each flask). Ampicillin was added at 100µg/mL. In addition to this, the following supplements were added to enhance growth and expression: 1mL of 1mg/mL biotin per litre medium; 1mL of 1mg/mL thiamine per litre medium; 100mL of Silantes CND medium per litre medium.

One 10mL overnight culture was used to inoculate each flask containing 400mL media using the following procedure. The overnight cultures were pelleted at 2500rpm in a bench-top centrifuge for 15 minutes and each pellet was resuspended using 10mL of the minimal M9 media taken from a flask. The resuspended cells were then used to inoculate the media which were subsequently incubated at 37°C, 160rpm until the OD_{600} ~0.6-0.8. At this point, IPTG was added to each flask (final concentration 1mM) and the grown at 30°C, 160rpm for 16-18 hours. The cells were then pelleted at 6037g for 25 minutes. They were resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use.

2.2.2.2 Perdeuterated, ^{15}N ^{13}C Vt/I997S

When attempting to fully deuterate Vt/I997S it was important to minimise the presence of any ^{1}H in the sample. All of the chemicals used in this preparation are dissolved in deuterium oxide and freeze dried to exchange any ^{1}H with ^{2}H. To ensure sufficient protein is obtained, 2L of M9 minimal media is required. The following chemicals were weighed and dissolved in deuterium oxide.

Chemical	Mass/g	Volume of Deuterium Oxide added/mL	To make Solutions
Na_2HPO_4	12	60	2L M9
KH_2PO_4	6	20	2L M9
NaCl	1	10	2L M9
$MgSO_4.7H_2O$	1	10	2L M9
$CaCl_2.2H_2O$	0.03	10	2L M9
$15NH_4Cl$	2	10	2L M9
Ampicillin	0.5	5	100mg/mL
IPTG	1.2	5	1M
Biotin/Thiamine	0.005 of each	5	1mg/mL

The D_2O solutions were sterile filtered using a 0.2µm syringe filter and then freeze dried for two days to remove any residual ^1H.

The following media were prepared to enable the BL21 (DE3) cells to adapt to growing in ^2H_2O.

Media Type	% ^1H_2O	% ^2H_2O	Volume made/mL
Luria broth	100	0	50
Luria broth	50	50	50
M9	50	50	100
M9	0	100	100

These were autoclaved with an additional 5 empty 2L baffled flasks. These flasks are for making up the final deuterated M9 media. It is important not to autoclave the final media once it is prepared as H_2O/D_2O exchange will occur which will reduce the percentage of ^2H in the solution.

The BL21 (DE3) cells were adapted in the following way. 50ml 100% ^1H_2O LB (ampicillin 100µg/ml) were inoculated with cells containing the Vt/I997S plasmid. The culture was grown overnight at 37°C, 160rpm. The following morning, two 1mL aliquots were used to inoculate two sterelins containing 10mL 50% ^2H_2O LB. These were grown at 30°C, 160rpm for 8 hours after which 1mL was used to inoculate a further 10mL 50% ^2H_2O. These were grown at 37°C, 160rpm overnight. After about 20 hours of growth, the cells were pelleted and the LB media discarded. Cell pellets were resuspended in 10mL of M9 media containing 50% ^2H_2O which was subsequently added to the remaining 90mL of the same media. This was grown for 24 hours to a high cell density (OD_{600}>1), before 10mL was used to finally inoculate 100mL of M9 made with 100% ^2H_2O.

Under sterile conditions, 2L of M9 minimal media was made using the freeze dried salts and using ^2H_2O in the place of water. The measured pH of the media was 6.9 and was not adjusted to minimise addition of ^1H to the media. As pD is pH + 0.41 (208), then the

media is at equivalent pH7.3. This media was split between the five 2L baffled flasks, 400mL in each.

Prior to induction, the freeze-dried ampicillin was made up to 100mg/mL using $^{2}H_2O$ and was added to the flasks at 100µg/mL. 2g freeze-dried $^{15}NH_4Cl$ was dissolved in 10mL $^{2}H_2O$, and 2mL added to each flask. Deuterated, ^{13}C-glucose (Cambridge Isotope Labs) was then added at 2g/L. Finally, the biotin/thiamine mix was dissolved in $^{2}H_2O$ to 1mg/mL and 1mL was added per litre media. To stabilise the bacterial growth, Silantes CND media was added to a final concentration of 10% vol/vol.

10mL of the 100% $^{2}H_2O$ overnight culture was added to each flask. Cultures were grown at 37°C, 160rpm until OD_{600} 0.6-0.8. At this point, the freeze-dried IPTG was redissolved in $^{2}H_2O$ to a final concentration of 1M. This was added to the flasks to a final concentration of 1mM to induce the culture.

Cultures were incubated at 30°C, 160rpm for 16 hours. After this time the cells were pelleted using a Beckman J2-21 centrifuge with a JLA 10.500 rotor at 9000rpm. These cell pellets were softer than observed for a $^{1}H_2O$ growth. The cell pellet was then stored in 20mM Tris pH7, 150mM NaCl at -20°C.

2.2.3 Selective ^{14}N-unlabelling of arginine residues in Vt/I997S

1L of M9 minimal media was aliquotted equally into two 2L baffled flasks and autoclaved. Prior to use, ^{15}N ammonium chloride was added at 1g/L and unlabeled glucose at 4g/L. 1g of ^{15}N ammonium chloride was dissolved in 10mL ddH_2O, 0.2µm-filtered and 5mL added to each flask. 4g of ^{12}C glucose was dissolved in 100mL ddH_2O, 0.2µm-filtered and 50mL added to each flask. Ampicillin was added at 100µg/mL to each flask. 20mL of LB overnight culture was pelleted at 900g for 15mins and resuspended in 20mL M9 media. 10mL of the cell suspension was used to inoculate each flask, and the cultures grown at 37°C, 160rpm to OD_{600} ~0.4. At this point an excess amount of unlabeled arginine was added to 0.1g/L. Growth was continued at 37°C, 160rpm until OD_{600} 0.6-0.8. IPTG was

added to each flask to a final concentration of 1mM and the flasks incubated at 30°C, 160rpm for 16-18 hours. The cells were then pelleted at 6037g for 25 minutes. They were resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl, and frozen at –20°C prior to use.

2.2.4 Selective ^{15}N-labelling of amino acids in Vt/I997S

2L of M9 minimal media were aliquotted equally into two 2L baffled flasks and autoclaved. To each flask, 4g/L ^{12}C-glucose, 1µg/mL biotin and thiamine, and ampicillin at 100µg/mL were added. Each flask was labelled with leucine, lysine, valine or isoleucine. Prior to inoculation, 0.1g/L of each unlabelled amino acid was added, *except* for the amino acid that the flask is labelled with. Each flask was inoculated with a 10mL LB overnight culture cell pelleted resuspended in M9 minimal media. Cultures were grow at 37°C, 160rpm to OD_{600} ~0.4. At this point, 0.1g/L of ^{15}N labelled amino acid, ^{15}N-leucine, ^{15}N-lysine, ^{15}N valine or ^{15}N-isoleucine was added to the appropriately labelled flask. Growth was continued at 37°C, 160rpm to OD_{600} 0.6-0.8. IPTG was added to a final concentration of 1mM and the flasks incubated for 16 hours at 30°C, 160rpm. The cells were pelleted at 6037g for 25 minutes and stored in 20mM Tris pH7, 150mM NaCl at -20°C.

2.2.5 Expression of GST-LD1/LD2 and His-LD1/LD2

2.2.5.1 Standard growth

Glycerol stocks of pGEX-GST-LD1/LD2 and pET-15b-His-LD1/LD2 were streaked onto agar plates containing ampicillin at 100µg/mL and grown at 37°C overnight. Colonies were used to inoculate two 10mL LB media containing ampicillin at 100µg/mL which were grown at 37°C, 160rpm for 16 hours. 1.2L of LB media was prepared and 400mL aliquotted into three 2L baffled flasks, covered with aluminum foil and autoclaved. Ampicillin was added to each flask to a final concentration of 100µg/mL. Each flask was inoculated with 5mL culture and grown at 35°C, 160rpm until OD_{600}=0.5. At this point, the temperature was dropped to 16°C and the cells cooled for 30min. IPTG was then added to

a final concentration of 1mM and the cells grown for 21 hours at 16°C, 160rpm. The cells were pelleted at 6037g for 25 minutes, resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use.

2.2.5.2 ^{15}N Minimal media growth for His-LD1/LD2

1.2L of M9 minimal media was prepared and aliquotted equally into three 2L baffled flasks. The flasks were covered with aluminium foil and autoclaved. Prior to use, ^{15}N ammonium chloride and ^{12}C glucose was added at 1g/L and 4g/L respectively. (1.2g of ^{15}N ammonium chloride was dissolved in 15mL ddH_2O, 0.2µm-filtered and 5mL added to each flask. 4.8g of ^{12}C glucose was dissolved in 90mL ddH_2O, 0.2µm-filtered and 30mL added to each flask). Ampicillin was added to each flask at 100µg/mL. 30mL of LB overnight culture was pelleted at 900g for 15mins and resuspended in 30mL M9 media. Each flask was inoculated with 10mL of cell suspension and were grown at 35°C, 160rpm until OD_{600}=0.5. At this point, the temperature was dropped to 16°C and the cells cooled for 30 minutes. IPTG was then added to a final concentration of 1mM and growth continued for 21 hours at 16°C, 160rpm. The culture was pelleted at 6037g 25 minutes, resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use.

2.2.6 Expression of other proteins: Calexcitin

Glycerol stocks were streaked onto an agar plate containing ampicillin at 100µg/mL and grown at 37°C overnight. Colonies were used to inoculate two 10mL LB media containing ampicillin at 100µg/mL which were grown at 37°C, 160rpm for 16 hours. 1.2L of LB media was prepared and 400mL aliquotted into three 2L baffled flasks, covered with aluminum foil and autoclaved. Ampicillin was added to a final concentration of 100µg/mL. Each flask was inoculated with 5mL culture and grown at 35°C, 160rpm to OD_{600}=0.5. At this point, the temperature was dropped to 16°C and the cells cooled for 30 minutes. IPTG was then added to a final concentration of 1mM and the cultures grown for 21 hours at 16°C, 160rpm. The cultures were pelleted at 6037g for 25 minutes, resuspended in 10mL of 20mM Tris-HCl pH 7.0, 150mM NaCl and frozen at –20°C prior to use

This method of expression was used so that calexcitin could be expressed simultaneously with GST-LD1/2. Calexcitin expression can also be carried out at 30°C for ~6 hours after induction with IPTG.

2.2.7 Purification of wild-type and mutant vinculin tail

Buffer A: 20mM Tris pH7.0, 150mM NaCl
Buffer B: 20mM Tris pH7.0, 150mM NaCl, 250mM Imidazole
Buffer C: 20mM Tris pH7.0, 25mM NaCl
Buffer D: 20mM Tris pH7.0, 1M NaCl

Wild-type *E.coli* cell pellets from 1.2L growths were thawed and resuspended in 50mL of buffer A. The suspension was sonicated using an XL-2020 sonicator at power 9 for a total of twenty 30 second bursts, with cooling in an ice bath between bursts for 50 seconds. The sonicated cells were centrifuged at 16000rpm in a Beckman J2-21 centrifuge (Rotor JA-20) for 25 minutes. The supernatant was collected.

2.2.7.1 Nickel-NTA affinity chromatography

25mL of Ni-NTA beads (Qiagen) were packed into an Amersham XK26 column and equilibrated using 4CV of buffer A. The sonicated supernatant was loaded into a 50mL superloop, pumped onto the column at 1mL/min and washed to UV280nm baseline. The column was eluted using the following percentages of buffer B: 0%B 2CV, 18%B 4CV, 100%B 4CV. An SDS-PAGE gel was run to visualise the elution profile and the appropriate peak was pooled. The approximate protein concentration was measured by an A_{280} reading against a buffer B blank.

2.2.7.2 Thrombin cleavage

5mM $CaCl_2$ was added to the pooled protein and a 200µl sample was taken as a non-enzyme control. This control was incubated at 37°C, 65rpm, alongside the cleavage

reaction to monitor spontaneous cleavage and degradation. A 50µl sample was taken as time-point zero, 50µl of disruption buffer was added and was boiled for 5 minutes, ready for SDS-PAGE. Thrombin was added to the main pool at 1U/1-2mg protein and was incubated overnight at 37°C, 65rpm. After 21 hours, the cleavage was checked on an SDS-PAGE gel and should be >99% complete.

2.2.7.3 Cation exchange

This step removes thrombin and any remaining impurities. A 50mL SP Sepharose (HiLoad 26/10 HP) was equilibrated using buffer D followed by buffer C. The cleavage pool was diluted ½ using buffer C to reduce the NaCl concentration. The sample was loaded into 50mL superloop, loaded at 1mL/min and washed to UV baseline. The protein is eluted using 0-100%D over 4.5CV. The elution was held at 100%D for 1.5CV and then washed back to 0%. The peak fractions were run on an SDS-PAGE gel and the appropriate fractions were pooled. The protein was concentrated to ~1mg/mL in the same buffer as it elutes from the column, and snap frozen and stored at –80°C.

Dialysis into lower salt concentrations was performed thus: Vt was concentrated to 1mg/mL in the elution buffer (high salt~400mM). Next dialysis was performed overnight into the desired buffer for 16 hours with one change. Typically, Vt was dialysed into 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. After dialysis the protein was concentrated to the desired concentration. However, if a high concentration of Vt is required without the addition of 50mM arginine, 50mM glutamate, the protein was concentrated in the presence of the elution buffer containing high NaCl concentrations before dialysis. This reduced the chance of Vt precipitating.

2.2.8 Purification of paxillin His-LD1/LD2

Buffer A: 20mM Tris pH7.0, 150mM NaCl, 6M Urea
Buffer B: 20mM Tris pH7.0, 150mM NaCl, 250mM Imidazole
Buffer C: 20mM Tris pH7.0, 25mM NaCl, 2mM DTT

Buffer D: 20mM Tris pH7.0, 1M NaCl

1.2L of *E.coli* cell pellets containing expressed paxillin His-LD1/LD2 were thawed and resuspended in 50mL of buffer A. The suspension was sonicated using an XL-2020 sonicator at power 8 for a total of twelve 30 second bursts, with cooling between bursts in an ice bath for 50 seconds. The sonicated cells were centrifuged at 16000rpm in a Beckman J2-21 centrifuge (Rotor JA-20) for 25 minutes. The supernatant was collected.

2.2.8.1 Nickel-NTA affinity chromatography

10mL of Ni-NTA beads (Qiagen) were packed into an Amersham XK16 column and equilibrated using 4CV of buffer A. The sonicated supernatant was loaded into a 50mL superloop, loaded onto the column at 0.5mL/min and washed to UV280nm baseline. The column was eluted using the following percentages of buffer B: 0%B 2CV, 10%B 4CV, 100%B 4CV. An SDS-PAGE gel was run to visualise the elution profile and the appropriate peak was pooled.

2.2.8.2 Anion exchange

This allows the construct to flow through the column and impurities bind to the column. A 25mL Q-Sepharose column was equilibrated using buffer D followed by buffer C. The cleavage pool was diluted ½ using buffer C to reduce the NaCl concentration. The sample was loaded into 50mL superloop, loaded at 1mL/min and washed to UV280nm baseline. The impurities were eluted using a 100%D step until the UV returns to baseline. The unbound flow-through fraction was retained and pooled.

2.2.8.3 Desalting and freeze-drying

The protein cannot be concentrated to a lower sample volume using spin concentrators as the protein passes through the membrane. The flow-through fraction was desalted using a

25 mL sephadex G25 superfine column. The column was equilibrated into double distilled water and 5mL aliquots of the pooled protein was added each run. A water isocratic gradient was carried out and the construct eluted first and the salts second. The protein peak was pooled each time and an A280 reading is taken. The protein pool was separated into 10mL aliquots and frozen in dry ice for 1 hour prior to freeze drying.

2.3 Protein characterisation theory and methods

2.3.1 Mass spectrometry

Electrospray ionization mass spectrometry (ESI MS) was carried out to determine the molecular weight of proteins in this study. In this method, proteins are buffer exchanged into a highly volatile solvent and then passed through a highly charged needle which disperses the protein solution into a fine mist. The solvent evaporates resulting in highly charged protein ions. Protons are added when the protein travels through the needle to give additional charge. The mass-to-charge ratio (m/z) is then analysed in a vacuum chamber.

The m/z spectrum shows a range of species with different charges and mass differences, each successive peak differs by a charge of 1 and a mass of 1 (from the proton). The mass can then be determined from any two neighboring peaks, but in practice the peaks are averaged to produce a more accurate result.

Mass spectrometry in this project was carried out by Neville Wright at the University of Southampton.

2.3.2 Analytical size exclusion chromatography

All analytical SEC was carried out using a Superdex-75 column (24ml volume). Column was calibrated and run using 50mM Na_2HPO_4, 150M NaCl. Samples were injected onto the column using a 200µl sample loop. 100µl of 1mg/ml protein samples were used.

2.3.3 Far UV circular dichroism

Circular dichroism (CD) can be used to study the secondary structure of proteins and measures the optical activity of proteins in solution. The CD phenomenon occurs by passing left and right circularly polarised light through an optically active molecule. These two beams can be absorbed differentially so that after they have passed through the same sample, their sum is no longer plane polarised but elliptically polarised (209).

Figure 2.3.1 shows characteristic spectra of different conformations of poly-L-lysine (210). α-helices are characterised by minima at 208nm and 222nm, β-sheets have positive peaks around 195nm and a negative peak at 216nm, and random coil produce minima below 200nm and a positive peak at around 218nm (211).

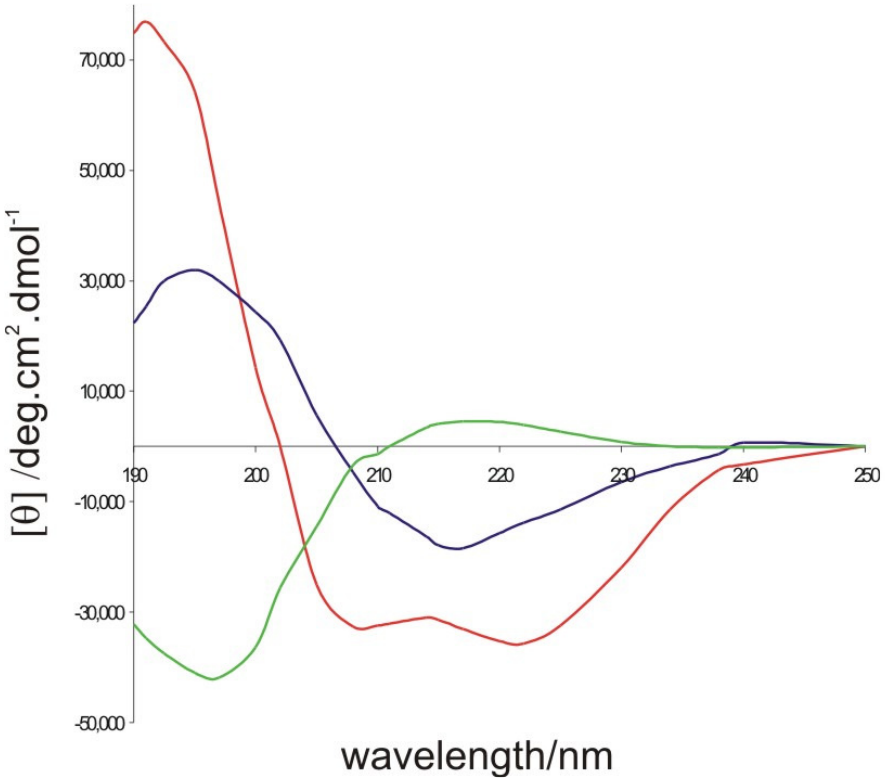


Figure 2.3.1 Far UV circular dichroism spectra of the α-helical (red), β-sheet (blue) and random coil (green) forms of poly-L-lysine (210)

Many proteins are a mixture of different types of structural motifs and a given CD spectrum contains information on the proportions of each. The problem is how to calculate these proportions. In this thesis, CD spectra have been analysed empirically, noting where minima and maxima occur, and some of the data has been submitted to an online server running SOMCD (212), which will calculate the predicted contribution of α-helical, β-sheet, β-turn and random coil structure.

Thermal denaturation can be followed using far UV circular dichroism. By measuring the ellipticity at a set wavelength, a denaturation or 'melting' curve can be measured. The midpoint of the transition will give the melting temperature, Tm. The effect on the melting temperature can be monitored by adding ligands to the desired protein and seeing if the stability is affected.

2.3.4 Far UV CD wavelength scans and thermal melting

Far UV circular dichroism experiments were carried out on a JASCO-J-720 spectrometer. Wavelength scans were recorded in the far UV range, 190-250nm with the following parameters set: pathlength 1cm, band width 1nm, sensitivity 200mdeg, resolution 0.2nm, response 1s, speed 100nm/min. Scans were repeated 9 times. Thermal denaturation time scans were recorded once for 65 minutes with the following parameters: pathlength 1cm, band width 2nm, sensitivity 50mdeg, resolution 10s, response 0.125s.

Wavelength scans to determine secondary structure were carried out on wild-type Vt and mutants. All proteins were at 500μM in a starting buffer containing 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. They were diluted into double distilled water (at pH7) to 0.5μM which gave a satisfactory CD signal. His-LD1/LD2 was diluted into water from a 224μM stock that was in the same starting buffer to a final concentration of 0.5μM. LD peptides were measured at 50μM in water (diluted from the same buffer) which was the highest concentration achievable without compromising the HT value. All final samples were between pH7-7.1.

Thermal melting experiments were carried out using the same protein concentrations as above. Samples were made in a 2mL volume. The spectrometer was thermally insulated and a heated waterbath was used to control the temperature. The temperature cannot be recorded directly using this setup, so it was necessary to record the temperature independently of the CD measurements. The first stage was to calibrate the temperature difference between the temperature reading on the water bath and the reading in the cuvette. This was done by placing a digital thermometer in the 2mL cuvette containing the buffer and starting a temperature drive from 25ºC to 85ºC. Every 2 minutes the temperature was recorded for the waterbath and the cuvette and these data were used to make a calibration curve. A graph of cuvette thermometer reading (y) vs. the heatbath reading (x) showed that the data was linear which meant that for a recorded heatbath temperature the actual cuvette temperature could be calculated. The lag phase between the two readings was negligible due to efficient thermal insulation. This graph produced the equation $y = 0.9776x - 2.9606$ which was used to calculate the temperature in the cuvette from a given heatbath reading (value A). A second factor to account for was that when the digital thermometer was placed in the waterbath it did not read the same temperature as the heatbath. Consequently the heatbath was shown to be the correct reading. The relationship between the digital reading (x) and the real heatbath temperature (y) was also linear and gave rise to an equation $y = 1.0114x - 2.1146$ which is used to convert value A into the real temperature value (value B). This means that experimental data collected was subjected to these two equations to create the temperature vs. CD data.

After this calibration, protein melting experiments could be carried out. A complete experiment is carried out thus: the desired sample was placed into the spectrometer with a sealed cap over the cuvette. A wavelength scan was recorded at 25ºC to monitor the conformation. A stop clock was set to 65 minutes and the countdown was started and the time scan was begun (measuring at 220nm). The sample was left to equilibrate at 25ºC for 10 minutes before a temperature gradient from 25-85ºC was started. Every two minutes since the start of the countdown the temperature was recorded from the waterbath. Once at 85ºC the sample was left until the time runs out (usually 5-10mins). A wavelength scan was measured at 85ºC. The waterbath was cooled back to 25ºC using ice and after 15

minutes a wavelength scan was repeated at 25ºC. Experiments were carried out each time with new Vt samples made from the same stock. For each set, both the free ligand and the complex Vt plus ligand was measured.

Wavelength scans were analysed by subtracting the free ligand spectra from the spectra recorded of the complex. Thermal denaturing data was imported into Excel (Microsoft) using the program Standard Analysis (JASCO). It was subjected to calibration equations in Excel and the data was exported into the program Origin7 (OriginLab) where the data was smoothed using 10 point adjacent averaging and the differential was calculated. This was fitted using the non-linear curve fitting algorithms in Origin7 to find the Tm (Figure 2.3.2).

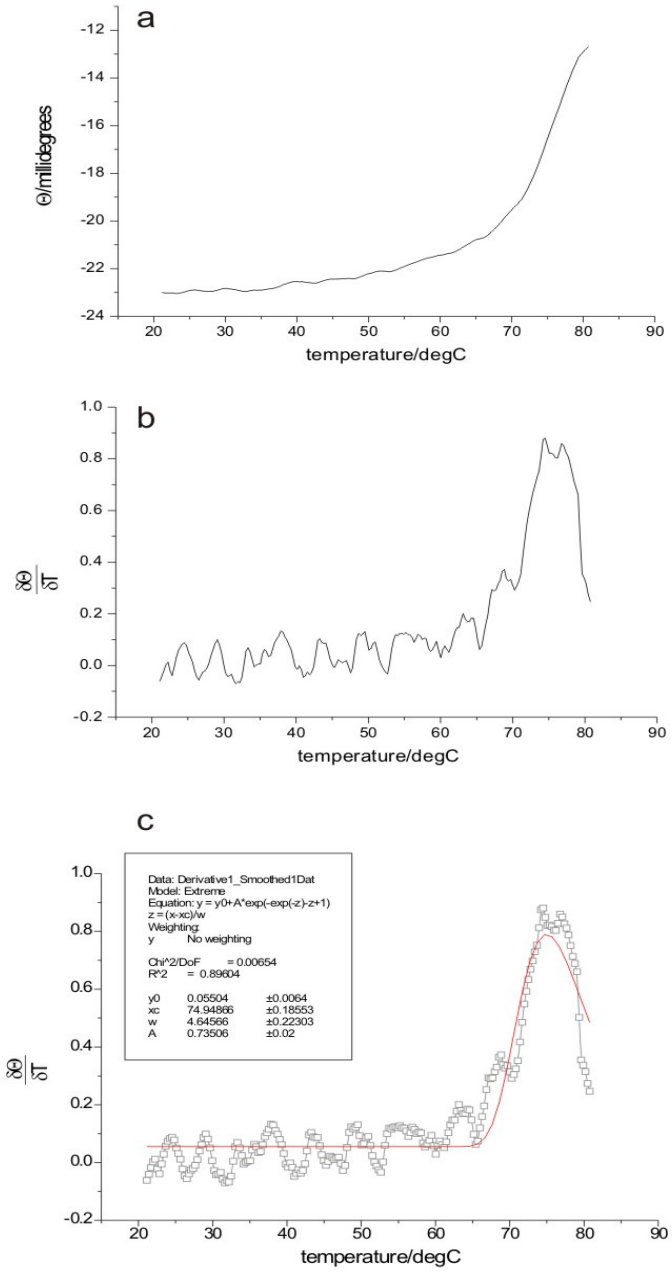


Figure 2.3.2 Finding the midpoint of the melting curve for wild-type Vt

a: the melting curve for Vt. b: The derivative of this melting curve. c: Non-linear curve fit to the derivative to find the Tm.

2.4 Site directed mutagenesis of Vt

The method of mutagenesis used was the Stratagene Quikchange method. It is a non-PCR based method as the primers only amplify the parental strand and not the mutated strand. The primers are designed to anneal to the parent plasmid and are complementary to opposite strands of the template. The primers for SDM can be found in Chapter 3.2.3 (Figure 3.2.2).

To create Vt/I997S, Vt/M1022S and Vt/ΔC mutant, the following procedure was followed. Three 10mL overnight cultures in LB with ampicillin (100μg/mL) were prepared and inoculated using wild-type Vt. The entire 10mL was used to produce a DNA miniprep using the Promega Wizard plus SV miniprep kit. The final miniprep concentration was calculated by diluting the DNA 1/100 in double distilled water and taking an A260nm reading. This value was used to calculate the concentration in the following manner:

A260nm x 100 x 50 = DNA concentration (ng/μl)

Site directed mutagenesis on the wild-type template was carried out in order to create the desired mutation. The following amplification reaction mixture was set up (all volumes are in μl):

	I997S	M1022S	ΔC
10x Pfu buffer	5	5	5
100ng template	determined	by miniprep	concentration
125ng forward primer	0.934	0.968	0.892
125ng reverse primer	0.914	0.968	0.915
dNTP mix	1	1	1
Sterile water	up to 50	up to 50	up to 50

The reaction mixes were put into a PCR block at 94°C for 1min to ensure that the DNA was completely denatured, then 1μl Pfu polymerase was added to the mixtures. The following amplification cycle was carried out:

Step	Temperature	Time	Cycles
Denaturation	94°C	1min	18 cycles
Annealing	55°C	1min	
Extension	68°C	13min	
Final extension	68°C	10min	1 cycle
Hold	4°C	∞	

At this time overnight cultures of BL21 (DE3) cells in LB without ampicillin were prepared and left to incubate at 30°C, 160rpm for 18 hours.

After the amplification was complete, a 1% agarose gel was run to ensure amplified DNA bands were present. Next, 1µl Dpn1 and 5µl of the Dpn1 buffer B was added to each reaction mix and incubated for 1hour at 37°C to digest the methylated parental DNA.

Next, the overnight culture of BL21 (DE3) cells were made competent. The following provides cells for two transformations: 500µl of overnight culture was used to innoculate 10mL of LB with no antibiotic added and growth was carried out at 37°C, 160rpm to OD_{600} ~0.6. The cells were pelleted in a benchtop centrifuge at 3000rpm for 15 minutes and the supernatant discarded. The pellet was resuspended in 5mL ice-cold 50mM $CaCl_2$ and chilled for 20 minutes on ice. The centrifugation was repeated and the pellet resuspended in 500µl ice-cold $CaCl_2$. The resuspended pellet was then aliquotted in 250µl volumes into sterile ependorfs and stored on ice until needed.

The competent cells were transformed with the amplified DNA. 25µl of the amplification reaction mixture was added to 250µl competent cells, chilled on ice for 20mins and heat shocked at 42°C for 90 seconds. 200µl of SOC media containing no antibiotic was added to the transformation mixture and incubated at 37°C, 160rpm for 1 hour. This mixture was spread on an LB-agar plate containing 100µg/mL ampicillin and grown overnight at 37°C, 160rpm.

Transformants were picked from the plate and used to inoculate 10mL LB aliquots containing 100µg/mL ampicillin which were grown overnight at 37°C, 160rpm. 1mL of culture was taken and used to inoculate 10mL LB with 100µg/mL ampicillin and grown to OD_{600} ~0.6. At this point the cells were pelleted, resuspended in 800µl LB and 200µl sterile glycerol was added. This was placed into a cryovial and stored at –80°C as a glycerol stock. The remaining 9mL of overnight culture was pelleted and a mini-prep was prepared using the Promega Wizard plus SV miniprep kit. This mini-prep was then ethanol precipitated ready for sequencing reactions. 2.5 volumes of ice-cold ethanol was added to the mini-prep DNA. Next, 1/10 of the original DNA volume of sodium acetate pH5.2 was added. The samples were frozen at –80°C for 20 minutes. Next, they were spun in a benchtop centrifuge at maximum speed for 10min and the supernatant was discarded. 1mL of ice-cold 70% ethanol was added to the pellet and the centrifugation was repeated. The supernatant was removed again and the pellets left to air-dry overnight.

To make the mutant Vt/I997S/M1022S the template was changed and Vt/I997S was used with the M1022S primers.

2.5 Nuclear magnetic resonance (NMR)

Solution-state NMR is a powerful tool that can be used in the determination of molecular structure, intra-molecular dynamics, protein folding and ligand-protein interactions. NMR enables the observation of individual nuclei in a protein thus reports with atomic resolution. It is a rapidly advancing field with the first application of NMR to a biological sample being reported in 1954 (213).

One useful feature of NMR is that it is possible to utilise the phenomenon without requiring a full quantum-mechanical understanding or description of the processes involved. There are a number of excellent textbooks that describe in further detail the experiments described here, including pulse sequences and quantum mechanics (214-218). In this thesis I have attempted to introduce the techniques that I have used and discussed some of the theory behind the important aspects of my work.

NMR occurs due to the property of spin. ^{1}H (protons) are spin ½ nuclei and are thus magnetically active. Naturally abundant ^{12}C is not magnetically active and ^{14}N has unsuitable properties in solution, so it is necessary to incorporate isotopes such as ^{15}N or ^{13}C which are spin ½ and magnetically active. This enables the NMR spectra of biological molecules to be obtained.

A spin ½ nuclei can orient itself in two ways when placed in an external magnetic field, either parallel to the field (α-state) or antiparallel (β-state). This creates two energy levels with a small excess of spins being in the parallel, low energy orientation. The small excess parallel to the magnetic field causes net magnetisation in this direction (this is z-magnetisation). Application of a radiofrequency pulse perpendicular to the external field causes transitions between energy levels and an NMR spectrum can be recorded. As there is only a small excess of molecules in the low energy state, NMR is a relatively insensitive technique. To counteract this, large sample concentrations are used to increase this apparent excess of ground state molecules so that transitions can be measured.

When a macromolecular sample is placed in an external magnetic field, the nuclei will align with the magnetic field and *z*-magnetisation is created. NMR experiments are based on manipulating this *z*-magnetisation using radiofrequency pulses to reorient nuclear spins and then detect the return to equilibrium (*z*-magnetisation). The position of a nuclei in an NMR spectrum is dependant on the effective magnetic field at the nucleus. This is known as the chemical shift and is determined by shielding effects of electrons and influences by surrounding atoms.

2.5.1 Multidimensional NMR

One dimensional (1D) NMR spectra are plots of intensity vs. frequency. A 1D-NMR experiment consists of a preparation stage and a detection stage (Figure 2.5.1). During the preparation stage, magnetisation is transferred to the desired nuclei in a predetermined manner (for example a 90º pulse that rotates the magnetisation from the *z*-axis and places it

along the *x*-axis). During the detection stage the signal is recorded. This signal will eventually decay back to equilibrium along the *z*-axis via different relaxation mechanisms (see Section 2.5.8.1) and is thus called the Free Induction Decay (FID). In order for this FID to be recorded and stored by the computer the FID must be digitised. In practice this requires regular sampling of the signal during the detection period by incrementing the t_1 delay.

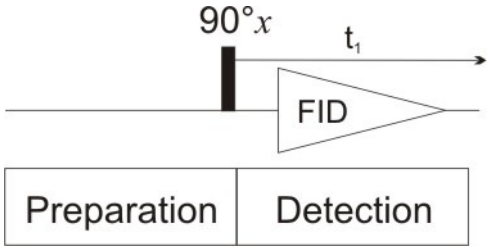


Figure 2.5.1 Basic building blocks for a one-dimensional NMR spectrum
During preparation, equilibrium magnetisation is manipulated to transfer magnetisation onto the desired nuclei, in this case by a 90º pulse that places it on the x-axis. The signal is then detected during the detection stage during time t_1 as it decays back to the z-axis.

2.5.2 Two-dimensional NMR

For simple molecules and organic compounds, 1D-NMR is sufficient for many purposes. In proteins however, 1D-NMR gives rise to complicated spectra comprising many, heavily overlapped peaks that makes interpretation almost impossible. In order to resolve some of this overlap a second dimension can be added to record spectra, known as 2D-NMR. 2D-NMR spectra consist of two separate frequency axes and one intensity axis. To build a 2D-NMR experiment there are four stages, preparation, evolution, mixing and detection. Preparation is the same as for a 1D experiment, to create magnetisation in a desired form (see Figure 2.5.2 where it is a 90º pulse). Next there is an evolution time where the spins are allowed to precess freely and relax back to equilibrium. The mixing time follows where magnetisation is transferred between spins and manipulated into a signal that can be detected. Detection follows in time t_2.

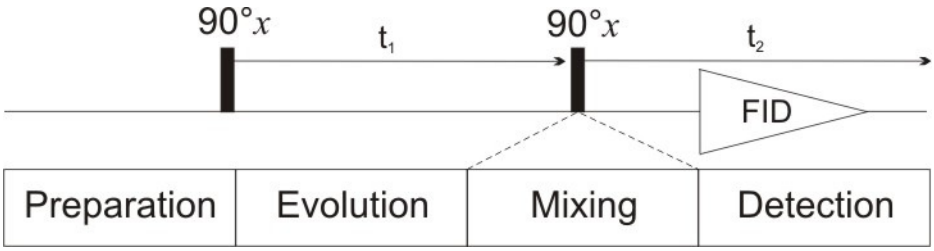


Figure 2.5.2 Basic building blocks for a two-dimensional NMR spectrum
During preparation, magnetisation is transferred onto the desired nuclei, in this case by a 90º pulse that places it on the x-axis. Evolution allows frequency labelling of the precession spin, and mixing enables transfer of magnetisation between spins. The signal is then detected during the detection stage during time t_2 as it decays back to the z-axis.

The evolution time t_1 is not fixed but can be varied incrementally to produce a series of separate data. During this time the magnetisation is labelled with the chemical shift frequency of the first nucleus. Importantly, no observations are made during t_1 so complex magnetisation can be created. During the mixing time, magnetisation is transferred between different nuclei and manipulated into an observable signal. The mixing period is important in establishing the relationship between the two dimensions and the pulses in the mixing period will determine the ultimate content of the spectrum.

The t_2 detection period is direct detection in real time and is sampled at regular intervals as for a 1D experiment. By incrementing the evolution period the collection of data in a second, 'indirect' dimension is enabled and is the basis of 2D-NMR. The evolution time t_1 is first set to zero and an FID recorded during t_2. The spins are allowed to return to equilibrium and then the pulse sequence is executed again but with a t_1 delay and an FID recorded as before. This is repeated incrementally thus building up a series of data that can be Fourier transformed to give two frequency axes.

As molecular weight increases past 20kDa, it becomes difficult to study molecules by 2D methods alone. Larger proteins contain more resonances and because they have longer rotational correlation times (see Section 2.5.9) and therefore faster transverse relaxation the linewidths are larger, leading to increased overlap of resonances. Increased transverse

relaxation rate also decreases overall sensitivity due to less magnetisation surviving through the pulse sequence. The overlap problem can be overcome by using three dimensional NMR.

2.5.3 Three dimensional NMR

A 3D-NMR experiment is a further extension of the 2D scheme. It can be constructed by joining two, 2D experiments together, leaving out the first detection period and the second preparation pulse. An additional evolution and mixing time is added (Figure 2.5.3) and t_1 and t_2 are incremented independently with direct detection during t_3.

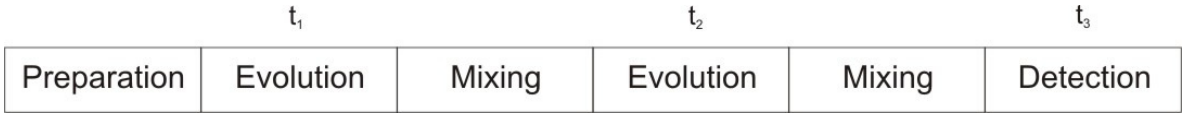


Figure 2.5.3 Basic building blocks for a three-dimensional NMR spectrum
This is an extension of the 2D scheme, with an extra mixing and evolution period added.

In this thesis, two and three dimensional experiments have been carried out. NMR experiments can be homonuclear (considering one type of spin such as protons) or heteronuclear (considering more than one type of spin). This thesis only uses heteronuclear experiments observing correlations between ^{1}H, ^{15}N and ^{13}C.

2.5.4 Heteronuclear 2D-NMR: ^{1}H-^{15}N-HSQC

Heteronuclear NMR allows the detection of protons correlated to other spin active nuclei such as ^{13}C and ^{15}N. The natural abundance of these isotopes in proteins is low, and they have low gyromagnetic ratios which impacts on their relative sensitivity. Therefore isotopic enrichment is required of ^{13}C and ^{15}N, and to overcome the lower sensitivity of carbon and nitrogen, NMR experiments are carried out by transferring magnetisation from the more sensitive proton, to the heteroatom (such as ^{15}N) then back onto proton where it is detected.

The 2D-heteronuclear experiment used throughout this thesis is the ^{1}H-^{15}N Heteronuclear Single Quantum Correlation Spectroscopy (^{1}H-^{15}N-HSQC) (219;220). This experiment records one-bond correlations between protons and ^{15}N by transferring magnetisation from protons onto nitrogen using an INEPT transfer (insensitive nuclei enhanced by polarisation transfer)(221). During the evolution period the ^{15}N nuclei acquire a frequency label specific to nitrogen (the ^{15}N chemical shift) and this labelled magnetisation is transferred back to protons where it is detected. This experiment is an important building block in 3D heteronuclear and triple resonance experiments and the basic HSQC experiment has been developed in various ways (218). The experiment used in this thesis is the gradient enhanced ^{1}H-^{15}N-HSQC (222).

Proton detection is useful as it increases the sensitivity of the experiment. Protons have a higher larmor frequency (the frequency of precession) and will therefore have larger starting equilibrium magnetisation that ensures a maximum starting signal. A larger amount of magnetisation can therefore be transferred to less sensitive ^{15}N, effectively increasing the signal obtained from this nuclei.

The ^{1}H-^{15}N-HSQC experiment produces a 2D spectrum with the $^{1}H^{N}$ measured in the direct dimension and ^{15}N measured in the indirect dimension. Thus resonances arise that have a proton and nitrogen resonance and correspond to each backbone NH in the protein, i.e. one resonance for each amino acid in the protein. The exception to this is proline, which has no free NH. Also observed in this spectrum are NH_2 groups from the sidechains of asparagine and glutamine and the indole H^{N} protons of tryptophan. An example of a ^{1}H-^{15}N-HSQC spectrum can be seen in Section 3.1.2.4, Figure 3.1.9.

^{1}H-^{15}N-HSQC spectra can be used to determine whether a protein is folded or whether it shows unstructured properties. Unfolded proteins produce overlapping chemical shifts centered in the middle of the spectrum as all residues are experiencing a similar environment due to the aqueous surroundings. In folded proteins the amino acids experience different chemical environments, reflected in the dispersion of chemical shifts.

The chemical shifts of NH resonances in a ^{1}H-^{15}N-HSQC spectrum are also very sensitive to changes in local environment, so addition of ligands and small molecules can cause changes in the peak positions and can be used to map specific interaction sites on proteins. This will be explored further in Section 2.5.13.

The ^{1}H-^{15}N-HSQC spectrum of vinculin tail has been assigned in this thesis and used for chemical shift perturbation experiments. In order to assign the spectrum, three-dimensional NMR experiments were required. These are described in the following sections.

2.5.5 3D-^{1}H-^{15}N-TOCSY-HSQC

TOCSY stands for total correlation spectroscopy (223;224). Experiments making use of TOCSY transfers act to transfer magnetisation through bonds via strong J-couplings (also known as scalar couplings). The J-couplings for bonds within amino acid residues are shown in Figure 2.5.4. Atoms with a coupling constant over four bonds (^{4}J) are extremely weak, so TOCSY transfers are restricted to atoms within an amino acid. Crosspeaks are generated between atoms in a spin system in a pattern which is specific to the amino acid type. Magnetisation is transferred during the mixing time, generally crosspeaks due to direct couplings will build up maximum intensity in 1/2J while larger couplings will need longer mixing times in the region of 100-200ms. Relaxation also occurs during the mixing time which can lead to loss of signal intensity; losses due to relaxation are larger when going through larger couplings.

The 3D ^{1}H-^{15}N-TOCSY-HSQC (225) provides correlations between atoms in a spin system, but they are edited according to the ^{15}N amide so that the protons that are observed are the ones directly bound to the nitrogen, thus reducing the overlap seen in 2D-TOCSY experiments. It will give amino-acid specific information but no sequential connectivities. It is usually used in concert with the 3D ^{1}H-^{15}N-NOESY-HSQC (see next section).

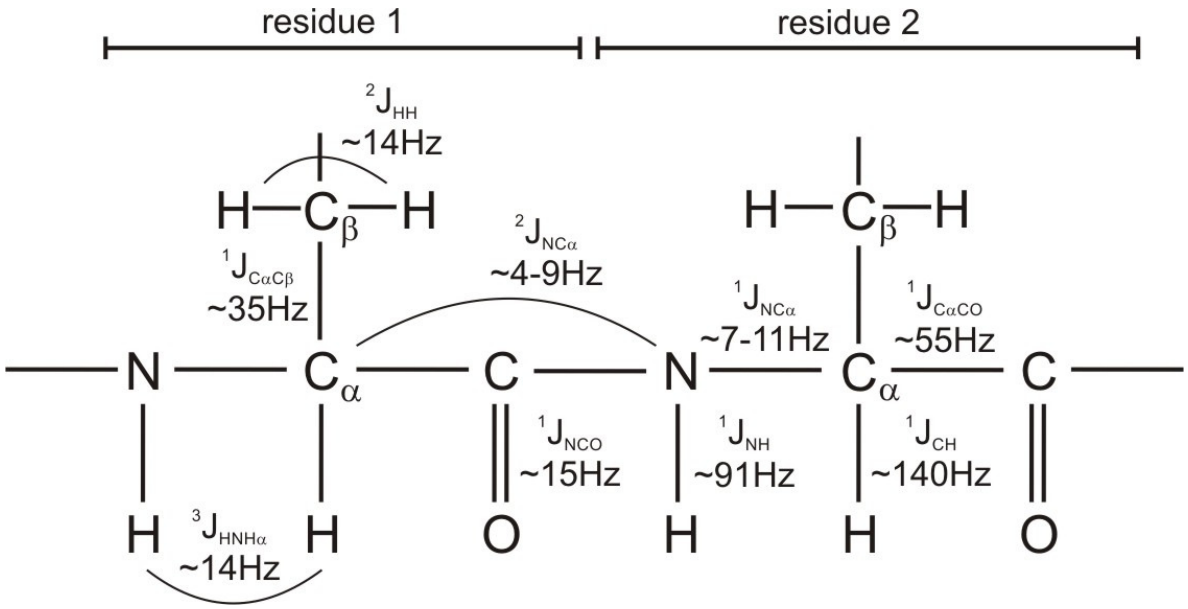


Figure 2.5.4 J-coupling constants between nuclei in a protein

2.5.6 3D-^{1}H-^{15}N-NOESY-HSQC

NOESY stands for nuclear Overhauser enhancement spectroscopy (226;227). The nuclear Overhauser effect is caused by dipolar relaxation which can give rise to transfer of magnetisation through space. When two spins come together in space they will 'sense' the magnetic field of the other and cause splitting of a signal. In solution, molecular tumbling averages these splittings but generates a fluctuating magnetic field which enables spins to transfer magnetisation between them. The efficiency of transfer through space depends on distance such that NOEs can only be observed between atoms less than 5Å apart. Short mixing times (50-150ms) are required to minimise spin diffusion in larger proteins with long correlation times. The size of an NOE is inversely correlated to nuclear distance.

The 3D-^{1}H-^{15}N-NOESY-HSQC (225;228-230) is used to resolve crosspeaks between ^{1}H spins by using the ^{15}N heteronucleus that they are directly bound to. Only magnetisation on protons attached to ^{15}N is selected for detection, despite all protons being frequency labelled. This results in only crosspeaks that are crossrelaxing with amide protons.

The difference between the 3D-TOCSY and the 3D-NOESY experiment is that while the TOCSY relies on through bond transfers and only gives amino acid specific information, NOESY data can provide information on the order of amino acids in a protein as it transfers magnetisation through space. The NOESY experiment should provide the connectivities between different amino acids. The two experiments are usually used in concert to assign protein spectra or for structural determination.

In this thesis, TOCSY transfers were inefficient so the ^{1}H-^{15}N-NOESY-HSQC experiment has been utilised in a different manner. The vinculin tail is a highly α-helical protein, in an α-helix the $^{1}H^{N}$-$^{1}H^{N}$ distance (dNN) is such that sequential NOEs can be detected between the amides of neighbouring amino acids (for a useful description of this see (215). Thus, one can sequentially walk along the protein backbone via the NOEs between NH groups.

2.5.7 Triple resonance experiments

Three dimensional triple resonance experiments have been carried out that correlate backbone NH, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$ and ^{13}CO. They have been used to assign the protein backbone in addition to the ^{1}H-^{15}N-NOESY-HSQC experiment. Figure 2.5.5 shows the correlations that are observed for the experiments used in this thesis, namely the HNCA (231-233), HNCACB (234), HN(CO)CA (233;235), HN(CO)CACB (236) and HNCO (231;233;237). These experiments correlate the backbone NH resonance with the carbons through one and two bond couplings and give rise to sequential information. The HNCA experiment will show a Cα resonance peak for a specific amino acid (linked to its NH) and a weaker Cα peak corresponding to the amino acid that precedes it in the sequence (i-1 peak). Therefore you can walk along the sequence using the Cα resonances. The HNCACB provides further specificity by giving the $C\beta_i$ and $C\beta_{i-1}$. The pattern of Cα and Cβ chemical shifts for a given amino acid can provide amino acid specific information. Figure 2.5.6 shows the patterns that can be expected for each amino acid.

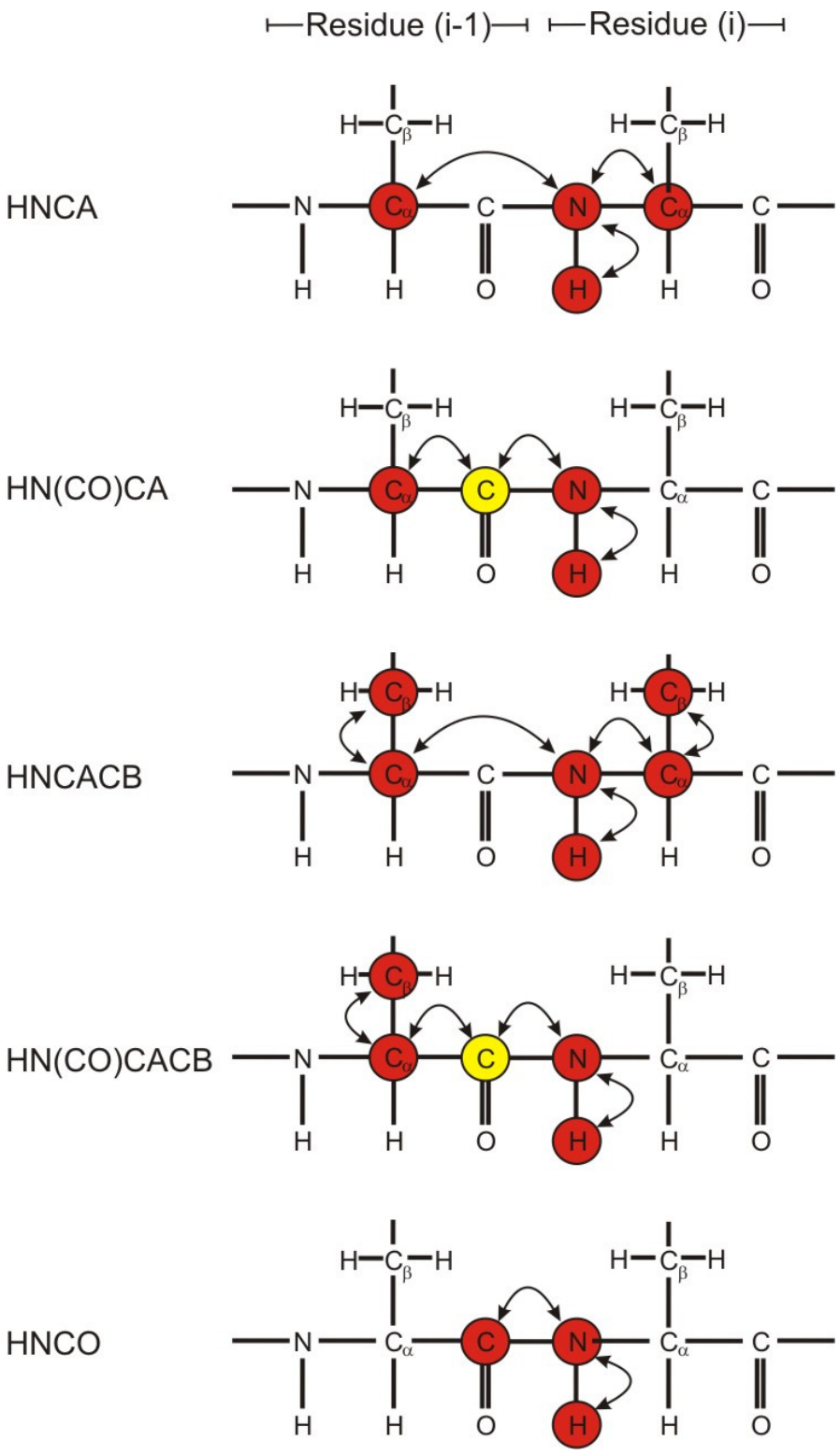


Figure 2.5.5 Triple resonance experiments used in this thesis

Red atoms are those which are frequency labelled and detected, yellow are those through which magnetisation passes but the atoms are not labelled. Arrows denote the scalar couplings used for the magnetisation transfers.

Amino acid	Carbon alpha (ppm)	Carbon beta (ppm)
Alanine	53.19	19.04
Arginine	56.82	30.71
Aspartate	54.65	40.89
Asparagine	53.50	38.74
Cystine	57.80	33.64
Glutamate	57.42	30.05
Glutamine	56.58	29.22
Glycine	45.38	-
Histidine	56.51	30.29
Isoleucine	61.56	38.67
Leucine	55.65	42.31
Lysine	56.94	32.80
Methionine	56.18	33.07
Phenylalanine	58.14	40.01
Proline	63.29	31.90
Serine	58.68	63.79
Threonine	62.19	69.63
Tryptophan	57.63	30.22
Tyrosine	58.09	39.34
Valine	62.47	32.73

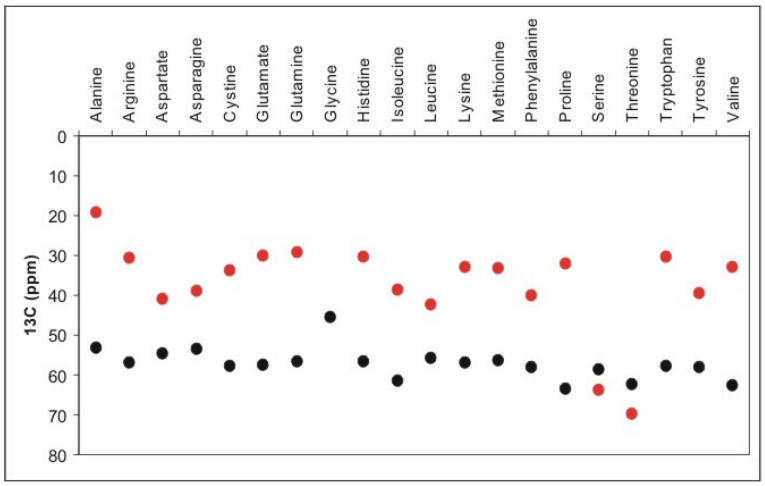


Figure 2.5.6 Expected carbon chemical shifts for amino acids in the protein backbone

2.5.8 TROSY

TROSY stands for transverse relaxation optimised spectroscopy (238). The TROSY effect can be exploited to improve the lineshapes in NMR spectra at high magnetic field strengths. In this thesis, TROSY has been used in many experiments at 800MHz and 900MHz and produced much improved spectra. In order to understand TROSY it is necessary to briefly introduce some concepts of relaxation in NMR.

2.5.8.1 T_1 and T_2 relaxation

Relaxation is the process by which magnetisation returns to equilibrium. At equilibrium there is no transverse magnetisation (along *x* or *y* axes) but there is longitudinal magnetisation along the *z*-axis caused by alignment of the spins in the external magnetic field. RF pulses can be applied to create transverse magnetisation and relaxation occurs that drives transverse magnetisation to zero and longitudinal magnetisation to a steady state value.

The bulk magnetisation is an average of the magnetic moments of all the individual spins. At equilibrium the *z*-magnetisation is in the longitudinal plane while the *x/y* contributions average to zero in the transverse plane. Application of the RF pulse changes this and

causes net magnetisation in the transverse plane and a reduction in the z-magnetisation. After the RF pulse the transverse magnetisation precesses and returns to equilibrium along *z*. This is how the FID is created and measured.

Relaxation back to equilibrium is a natural process and occurs due to naturally occurring magnetic fields in the sample. Longitudinal relaxation (or T_1 relaxation) causes *z*-magnetisation to return to equilibrium through interaction with the local fields in the environment. Each spin has its own magnetic moment that can act as a local field to affect neighboring spins. Due to thermal motion in the sample, the local fields vary in size and orientation and a fluctuating local field is produced. If the fluctuation should be close to that of the larmor frequency of the precessing spins then this can cause the magnetisation to be transferred to the surroundings. This relaxation is inefficient due to the fluctuating field being very weak and the variety of frequencies present in the local field.

Transverse relaxation (or T_2 relaxation) causes the transverse magnetisation to return to zero. After application of the RF pulse to create transverse magnetisation the spins precess freely at their larmor frequency. The size of this transverse magnetisation decays slowly because the individual spins cannot remain exactly synchronous with one another. After a certain time the individual spins become out of phase with one another, they lose coherence due to the local fields affecting each spin slightly differently. Eventually the transverse relaxation will return to zero as the spins become completely dephased.

2.5.8.2 Dipolar and CSA relaxation mechanisms

Two dominant relaxation mechanisms for spin ½ nuclei are dipolar and chemical shift anisotropy (CSA). The dipolar mechanism occurs via a local field produced by the magnetic moment of one spin affecting a second spin. This local field depends on the distance between the two spins, the gyromagnetic ratio of the spins and the orientation of the vector joining the spins relative to the applied magnetic field. The field of a neighboring spin can act to transfer magnetisation between spins sensing this field and can

act to dissipate the energy of excited spins, thus making important contributions to T_1 and T_2 relaxation.

CSA relaxation involves only one spin, in this case the local field is caused by electron currents in the nucleus in response to the external magnetic field. The nucleus experiences both the applied external field and the local field which will alter the larmor frequency by a small amount (this is the basis for the chemical shift). As the molecules tumble in solution the size and direction of this local field changes. Fluctuations in this local magnetic field as the molecule rotates is a source of relaxation.

Usually the dipole-dipole interaction is the dominant source of relaxation, however as the magnetic field strength increases, CSA becomes more important .

2.5.8.3 The TROSY effect

A nucleus in an NMR sample is likely to experience dipolar interactions from other nuclei and CSA from its own electrons. If these two effects are dependent on each other then cross-relaxation (or cross-correlation) is observed.

In the example of the ^{15}N-^{1}H spin pair, cross-relaxation will lead to affects on the doublets that arise due to dipolar couplings. In a situation where there is little or no cross-relaxation between 2 spins it is normal to observe a doublet of peaks where the splitting is equal between the two and relaxation occurs similarly for both spins (Figure 2.5.7a). Cross-relaxation between two spins such as in the case of ^{15}N-^{1}H produces an asymmetric doublet where one peak has a small linewidth and one has a very large linewidth (Figure 2.5.7c). In an NMR experiment it is normal to decouple these peaks to produce an average peak in the centre. As shown in Figure 2.5.7b and d, averaging the first situation keeps the linewidth the same but doubles the intensity, while in situation two the linewidth and intensity is an intermediate between the two. The line narrowing effect observed for one line of the doublet is particularly pronounced at higher fields.

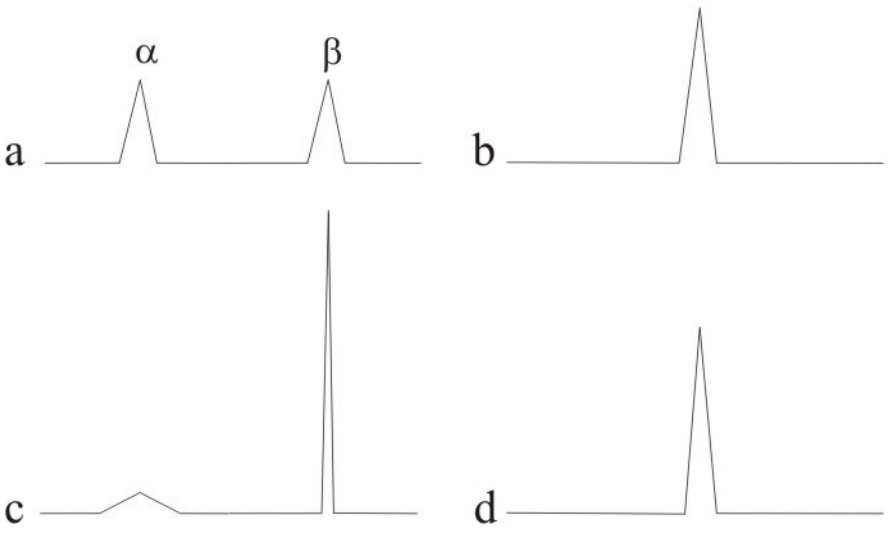


Figure 2.5.7 Cross-relaxation effect between two nuclei

a: little or no cross-relaxation between 2 spins presents as an equal splitting between two nuclei. Decoupling these two peaks keeps the linewidth the same but doubles the intensity (b). c: Cross-relaxation between two spins such as in the case of ^{15}N-^{1}H produces an asymmetric doublet where one peak has a small linewidth and one has a very large linewidth, decoupling leads to the linewidth and intensity becoming an intermediate between the two (d). Image adapted from (214).

In a conventional ^{1}H-^{15}N-HSQC spectrum it is usual to collapse the doublets (of which there are four) to produce one central peak (Figure 2.5.8a). Where cross-relaxation is occurring and CSA is having a large effect, some of these lines are much broader than the others (Figure 2.5.8b) and when they are averaged there will be a considerable reduction in peak height compared to the narrow intense peak. Therefore, where this effect is very large, such as in the case of large molecules at high fields, it is best not to remove the splittings due to the heteronuclear couplings and retain the signal from the large narrow peak. This is the basis of TROSY (238) and as a result, the spectra of larger proteins can be vastly improved using this technique.

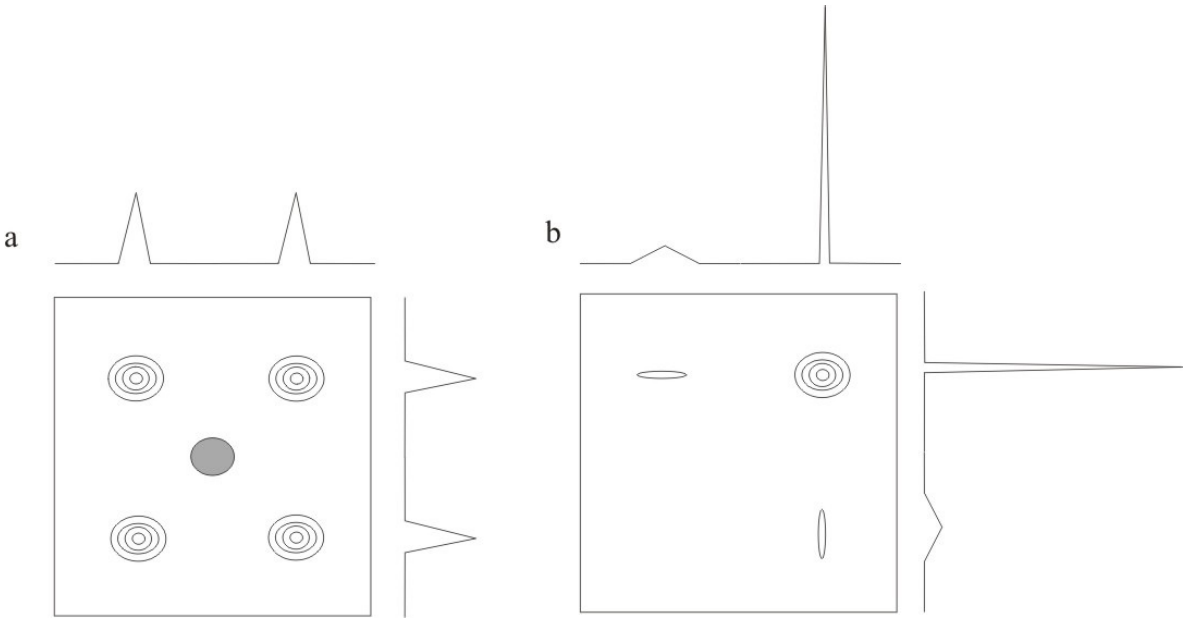


Figure 2.5.8 Example of the TROSY effect in a 2D-^{1}H-^{15}N HSQC.

Resonances are often decoupled in a 2D-^{1}H-^{15}N HSQC spectrum to produce one central peak (a: central grey peak). Where cross-relaxation is occurring, the linewidths can vary between the four splittings and it is beneficial to maintain the narrow intense peak (b, top right peak).

2.5.9 Rotational correlation time

The rotational correlation time of Vt and mutants was measured in this thesis as a gauge of aggregation effects and/or oligomersation in solution. The rotational correlation time (τ_c) of a protein in solution is defined as the average time for all molecules in a solution to achieve an orientation one radian away from the starting position (214). For a protein of 20kDa, a correlation time of around 10ns would be expected.

τ_c is determined by the size, shape, and dynamics of the molecule (239). As the correlation time is determined in part by the molecular weight, measuring τ_c can give clues to aggregation occurring in solution and oligomerisation states of proteins (240). For protein

less than 30kDa, τ_c is usually measured as a ratio of longitudinal and transverse relaxation rates (241-243) and by relating experimental values for τ_c to the solvent accessible area of a molecule, estimates of aggregation can be made as in Section 2.6.6 (239;240;244).

As represented in Figure 2.5.9, small molecules have fast rotational tumbling and $T_1 = T_2$. As the correlation time increases, T_1 decreases to a minimum and then increases again while T_2 continually decreases. The linewidth is related to T_2, such that as molecular weight increases, correlation time also increases and T_2 gets smaller. Therefore linewidths get larger as τ_c increases (215;217).

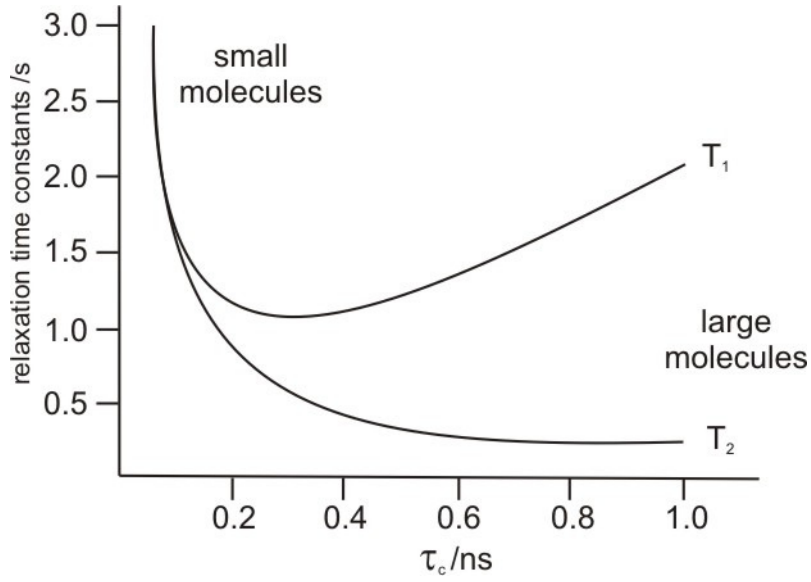


Figure 2.5.9 Variation of T_1 and T_2 with correlation time

As correlation time increases, T_1 decreases to a minimum and then increases again while T_2 decreases. This decrease in T_2 with increasing τ_c results in line broadening in NMR spectra.

2.5.10 Partial deuteration and perdeuteration

For proteins with large molecular weights, NMR spectra contain a large numbers of signals and increased overlap of the resonances. Also due to the longer correlation time of larger molecules in solution, relaxation of the magnetisation on the spins is faster meaning that the

signal lasts for a shorter time. As there is more efficient transverse T_2 relaxation the linewidths of the resonances are increased which can compound the overlap problem and reduce the signal to noise of the experiment (245).

As mentioned previously, dipolar relaxation is a dominant mechanism by which magnetisation is lost, this is especially due to ^{1}H-^{1}H (245) and ^{13}C-^{1}H spin pairs (246). An effective way of making relaxation occur more slowly, thus allowing the signal to remain for a longer period, is by replacing protons attached to carbons with deuterons (^{2}H). The substitution of deuterons for protons depletes the number of protons available for dipole-dipole interactions, and as deuterons have approximately a 6-fold lower gyromagnetic ratio relative to protons (247), the transfer of magnetisation is much less efficient.

Deuteration of proteins can be complete, where around 99% of all carbon-bound atoms are replaced (known as perdeuteration), or partial, where the percentage of deuteration is an average distribution throughout the molecule. Deuteration can also be site-specific where ^{2}H is incorporated at known sites within a protein. Protons are reintroduced into the amide NH positions to enable NMR detection, this is often from water buffers during the purification stage.

Deuteration has been employed in this thesis to improve data acquired from triple resonance experiments (such as those in Section 2.6.4.1). These experiments encounter problems as the molecular weight of proteins studied approaches 20-25kDa. The long correlation times for these molecules causes efficient (and fast) relaxation of the spins involved. These relaxation effects can lead to an absence of detectable Cβ resonances in the HNCACB experiment. As the HNCACB is an 'out-and-back' experiment, the success of the magnetisation transfer requires Cα T_2 relaxation time long enough to establish the Cα to Cβ correlation (248). Without the Cβ resonances, sequential assignment becomes very difficult as both the amino acid specific information and the sequential connectivity is lost.

Deuteration is therefore useful for reducing the relaxation of magnetisation in larger molecules. The result of this on the spectrum is an increase in sensitivity due to

magnetisation surviving for longer and improved resolution (especially in carbon dimensions). The lack of resolution is in part due to the short acquisition times required because of the normally rapid transverse relaxation rates (246).

Deuteration also helps to suppress spin diffusion. As stated earlier, the NOE acts to transfer magnetisation between hydrogen atoms through space. This transfer may be more efficient through multiple short steps (e.g. through bonds) rather than a one-step transfer over the longer, direct distance (249). During experiments such as the ^{1}H-^{15}N-NOESY-HSQC spin diffusion can spread the signal over many protons so that the distance and connectivity information is lost. Proteins contain large networks of protons, and for proteins with a long correlation time, spin diffusion can cause problems if not accounted for. Deuteration of the carbon bound protons prevents them from transferring magnetisation from the solvent and reduces spin diffusion.

2.5.11 Selective labelling and unlabelling

Selective labelling has been used in this thesis to enable the assignment of the protein backbone (for method see Section 2.2.4). The procedure allows specific ^{15}N-labelled amino acids to be incorporated into proteins so that only the labelled amino acids are detectable using NMR. When a ^{1}H-^{15}N-HSQC spectrum is recorded, only the amino acids with the ^{15}N label will be observed, hence the amino acid identity of these resonances is immediately known. During assignment, this information can be used unambiguously to place connected residues from triple resonance experiments into the protein sequence.

Conversely, selective unlabelling inserts a non-magnetic, amino acid specific ^{14}N in a protein that is labelled with ^{15}N at all other amino acid positions. The selected amino acids are then missing from the ^{1}H-^{15}N-HSQC spectrum. For method see Section 2.2.3.

2.5.12 Secondary structure from Cα and Cβ shifts

Secondary structure can be identified in folded proteins from Cα and Cβ chemical shifts (250-252). The values for the carbon shifts for amino acids measured in urea (the random coil chemical shifts) can be subtracted from the experimental carbon shifts for the protein in question. This will produce a chemical shift deviation from the random coil value which is plotted against the residue number. Large positive deviations for Cα are characteristic of α-helices and negative deviations are seen for extended structures like β-sheets. Conversely, Cβ shifts tend to be negative for α-helices and positive for extended structures.

2.5.13 Chemical shift perturbation

Solution NMR is a powerful tool for evaluating protein-ligand interactions. It relies on the chemical shifts of the residues in the protein being sensitive to subtle changes in chemical environment (253). It is especially useful as it can be used to detect weak interactions, with K_ds in the millimolar range being detectable (254).

The most widely used NMR method to determine binding interfaces is chemical shift mapping or chemical shift perturbation (254). This is achieved by recording a ^{1}H-^{15}N-HSQC spectrum of an isotopically labelled protein, and then monitoring the changes in the chemical shifts as a binding partner is titrated in. These changes occur if the magnetic environment of the individual nuclei is changing.

In this thesis, chemical shift perturbation has been used to explore protein interactions. The manner in which the interaction occurs will be suggested by the exchange regime in the system. The NMR properties of the free and bound states are different. These differences can be observed between the chemical shifts, relaxation rates and diffusion rates (253), and the differences are determined by the chemical exchange rate, K_{ex} (255). When ligand is titrated into a protein solution, the chemical shift change, as well as the quality of the resulting spectra, is determined by the chemical exchange regime that the system is under (218).

2.5.13.1 Exchange regimes

Chemical exchange refers to any process where a nucleus exchanges between two or more environments which results in a change in its NMR parameters. Examples of factors resulting in chemical exchange include the making or breaking of chemical bonds, conformational changes within the molecule or pH changes within the sample. NMR provides an excellent method for monitoring the exchange of a nucleus between different environments as long as they are magnetically distinct.

The simplest manner in which a ligand will bind to a protein can be described as a second order exchange process,

Equation 2.i

$$P + L \underset{k_r}{\overset{k_f}{\rightleftharpoons}} PL$$

where K_f is the forward (association) rate constant and K_r is the reverse (dissociation) constant. The chemical exchange regime for the system will depend on the relationship between the exchange rate (K_{ex}) and the difference in chemical shifts between the free (P) and bound (PL) forms. For the above system, if the exchange rate is greater than the difference in chemical shift between P and PL, then fast exchange will be observed. If the exchange rate is less that the difference between the chemical shifts for the free and bound forms then slow exchange is occurring. When the exchange rate approaches the same as the difference between the two states then the system is in the intermediate exchange regime (Equation 2.ii).

Equation 2.ii Chemical exchange regimes

Fast Exchange: $K_{ex} >> \delta_P - \delta_{PL}$

Intermediate Exchange: $K_{ex} = \delta_P - \delta_{PL}$

Slow Exchange: $K_{ex} << \delta_P - \delta_{PL}$

K_{ex} is the exchange rate, δ_P is the chemical shift of the protein, δ_{PL} is the chemical shift of the protein-ligand complex

The exchange regime will affect the appearance of the protein resonances in the spectrum. During fast exchange, when $K_{ex} >> \delta_P - \delta_{PL,}$ one peak is observed which is an average signal from both the free and bound forms. Slow exchange manifests as two peaks, one from the free form and one from the bound form due to the difference in chemical shift being larger than K_{ex}. Intermediate exchange causes severe line broadening and coalescence into one broad peak (Figure 2.5.10).

During the course of a titration, several ^{1}H-^{15}N-HSQC spectra are recorded and overlaid. If a nucleus is in fast exchange, as ligand is titrated in there will be a peak for the free protein which tracks to a new position as ligand concentration increases. This peak will track until it reaches the final position for the fully bound protein. This occurs because the peak is an average of the two different populations of free and bound forms. It is easy to monitor this by following the assigned peak across the spectrum (Figure 2.5.11a). If slow exchange is occurring then the position of a peak remains the same but it reduces in intensity as the population of free form changes into the bound form. A new peak corresponding to the bound form will reappear at a new position in the spectrum (Figure 2.5.11b). It is usually necessary to reassign the ^{1}H-^{15}N-HSQC spectrum of the bound form to determine the identity of the new peaks. If the k_{ex} becomes the same as the chemical shift difference between states P and PL then the system in intermediate exchange. When this occurs the two signals from the free and bound forms merge into one and severe line broadening can occur. If this line broadening occurs too dramatically, then the resonances may disappear from the spectrum completely.

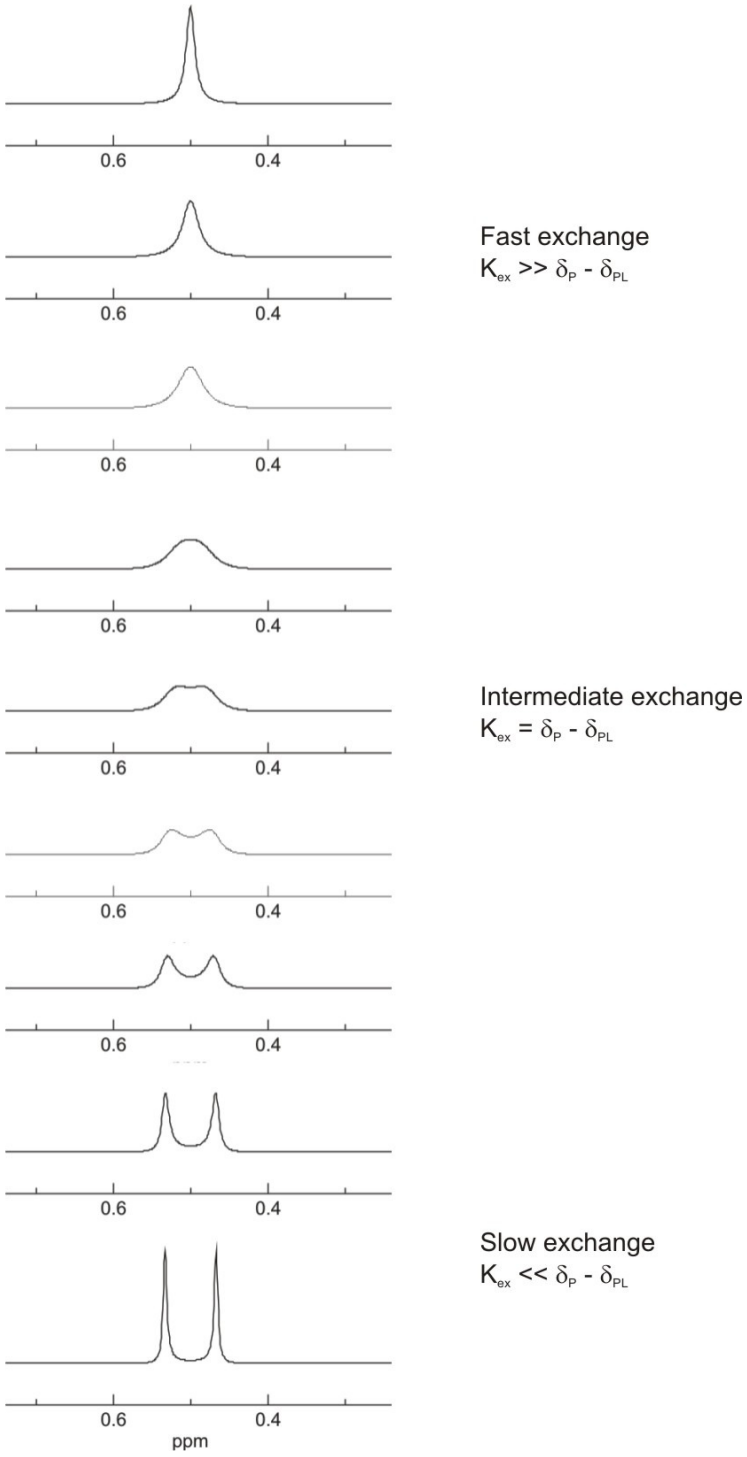


Figure 2.5.10 Effects of chemical exchange

In fast exchange, K_{ex} is > the difference in chemical shifts between two resonance peaks and a single peak is observed. In slow exchange, K_{ex} < the difference between chemical shifts and two peaks are observed. When the K_{ex} is the same as the difference in chemical shifts, intermediate exchange is observed and line broadening occurs. (Figure prepared using WINDNMR-Pro).

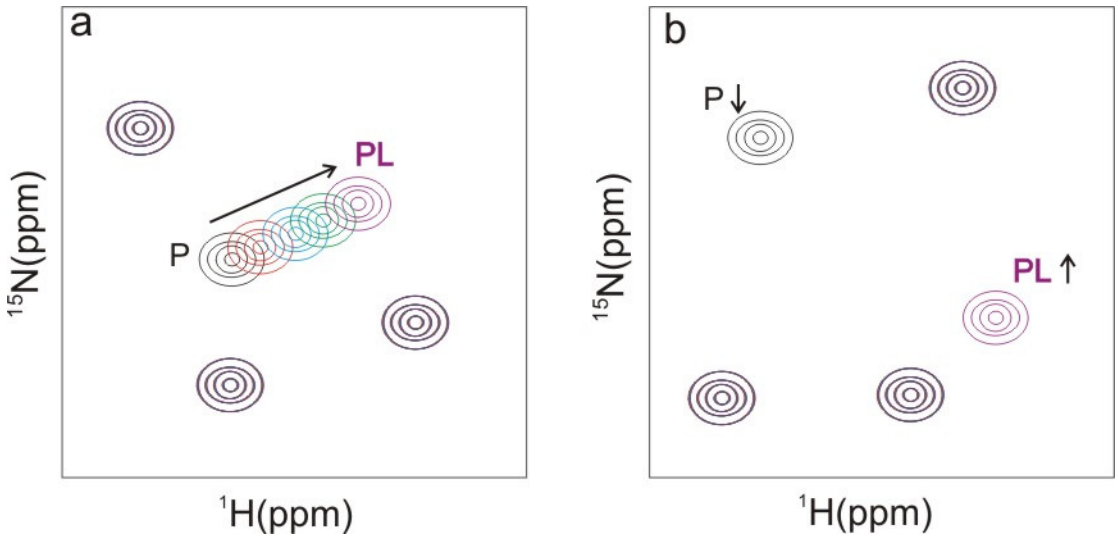


Figure 2.5.11 Effect of fast and slow exchange on a 1H-15N-HSQC spectrum

a: fast exchange manifests as one peak that tracks across the spectrum, starting from free protein (P) and ending at saturated protein-ligand complex (PL). b: slow exchange manifests as one peak that remains at the same chemical shift but loses intensity while a second new peak appears at a different chemical shift and increases in intensity

Fast exchange is often observed for weaker interactions and slow exchange can be observed in stronger interactions where complex dissociation is slow (254). For fast exchange, affinities can be obtained from plotting the chemical shift change vs. the ligand concentration, however slow exchange processes can be accompanied by line broadening effects and can make estimating affinities difficult (256).

Resonances that show large chemical shift changes upon binding are assumed to be involved in the binding interface between the protein and the ligand, however, smaller changes can occur for residues close to the binding site, so some care needs to be taken when interpreting chemical shift perturbation data.

2.5.14 Hydrogen-deuterium exchange

Main chain and side chain hydrogens in proteins are in constant exchange with the hydrogens in the solvent (257;258). Exchangeable groups within proteins include the NH hydrogens and hydrogens in polar side chains attached to N, O and S atoms. Hydrogens attached to carbon do not exchange easily (258). The mechanism of hydrogen exchange in folded proteins is complex due to the different environments of core and surface hydrogens and factors such as steric inaccessibility to solvent, charge distributions in neighboring amino acids and hydrogen bonding that will all affect the rate of hydrogen exchange (259).

Hydrogen exchange between the backbone NH groups and solvent occurs in a chemical reaction between the peptide group and an OH^- ion or H_3O^+ ion. Hydrogens that are more accessible will undergo this reaction more readily than stable hydrogens such as those in hydrogen bonds. However, when left for long enough, even deeply buried hydrogens will exchange at a rate that can be measured (260).

The rate of exchange between the molecule and the solvent is determined by the dynamic nature of protein molecules. Several different exchange mechanisms have been proposed (see for example reference (259)), the two main models being the solvent penetration model and the unfolding model. The solvent penetration model assumes that the ions can gain entrance to the protein core via transiently formed small pores and cavities. These could appear due to rapid fluctuations of atoms in the interior or could arise due to rearrangements in hydrogen bonding. The unfolding model proposes that hydrogens can only exchange when they are not in a hydrogen bonded form. This can only occur in a small fraction of time where a high energy unfolding reaction occurs that transiently separates donor and acceptor molecules. This is a dynamic opening and closing reaction and can be a local unfolding where only a few residues unfold at a time, a global whole molecule unfolding or some intermediate subglobal form in between (257;261).

Hydrogen exchange can be measured by looking at rates of exchange with deuterium oxide. A popular model for hydrogen exchange with deuterium is shown in Equation 2.iii.

According to this model, an exchangeable hydrogen that is slow to exchange most often resides in the closed state where exchange cannot occur (NH_{cl}). Occasionally the molecule experiences a dynamic motion that causes it to open (NH_{op}). Once an amide hydrogen reaches the open state, the balance between closing without exchange and moving forward with exchange dictates the two limits, EX1 and EX2 (262).

Equation 2.iii

$$NH_{cl} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} NH_{op} \xrightarrow{k_{rc}} ND$$

The EX2 limit occurs when the return to the closed form is rapid compared to the exchange step and as a consequence the exchange is base catalysed. This can give clues to thermodynamic behavior. The EX1 limit occurs when return to the closed form is slow compared to the exchange step and is therefore not base catalysed. EX1 provides information about kinetics and can monitor rates of unfolding or of local opening (263).

In this thesis hydrogen-deuterium exchange has been used to study the protection of proteins to exchange when ligands are bound. Ligand binding to a protein is predicted to protect the NH groups that are involved in the interaction by stabilising them so they do not exchange as readily. Using NMR, a protein can be dissolved in D_2O, consecutive ^{1}H-^{15}N-HSQC spectra can be recorded and the rate of exchange of individual NH groups measured. When used in this manner the deuterium atom cannot be detected. Therefore any NH groups that are undergoing exchange with D_2O will be missing from a ^{1}H-^{15}N-HSQC spectrum. Protection and stabilisation of NH groups is observed as a decrease in the exchange rate, NH resonances would persist longer than in the free form. Conversely, any destabilisation or change in dynamic nature of the protein upon ligand binding would be observed as an increase in exchange rate with D_2O.

Hydrogen-deuterium exchange can be a powerful method of confirming binding sites in proteins when used in concert with chemical shift perturbation experiments. Hoellerer *et*

al. (6) used chemical shift perturbation experiments to elucidate the binding sites for LD peptides on FAK and confirmed these by showing the same sites were protected in the presence of ligand, observed as increased protection from exchange with deuterium. This method has been employed in this thesis.

2.6 NMR Methods

2.6.1 Instrumentation

Three different field strengths were used in this study. 600MHz data was acquired on the spectrometer in the School of Biological Sciences at the University of Southampton. 800MHz data was acquired at the MRC Biomedical NMR Centre at the National Institute for Medical Research (NIMR), the experiments were conducted by Dr. Geoff Kelly. 900MHz data was acquired at the Henry Wellcome Building for Biomolecular NMR Spectroscopy at the University of Birmingham. Data was acquired with the assistance of Dr. Sara Whittaker. Some 800MHz and 900MHz data was acquired with a cryoprobe.

2.6.2 Sample preparation

2.6.2.1 Vinculin tail

Wild-type Vt and all mutant proteins were concentrated to approximately 1-2mg/mL in the high concentration NaCl buffer that the protein elutes from the cation exchange column, typically ~400mM NaCl. Spin concentrators (Vivaspin) with either 15mL or 2mL volume were used to concentrate Vt (10,000 molecular weight cut-off with a regenerated cellulose membrane). The proteins were dialysed into the appropriate buffer using slide-a-lyser dialysis cassettes (Pierce) with a 10,000 molecular weight cut off. Dialysis was carried out at 4ºC for at least 16 hours. Typically 2 x 5L buffer was used to dialyse protein volumes from 0.5mL to 10mL. Buffer choice was dependant on the desired experiment and can be found in relevant sections. The concentration was measured post dialysis by absorbance at 280nm and further concentration was carried out as required in new buffer. A280nm

readings were divided by the estimated absorbance at 1g/L to get a more accurate estimation of the concentration.

Typically, 200µM Vt was used for ^{1}H-^{15}N-HSQC spectra at 600MHz. ^{15}N-relaxation experiments were carried out using 100µM Vt, and 500-600µM Vt was used for assignment experiments.

The NMR sample was prepared in a volume of 520µl or 320µl. For experiments at 600MHz, 5% deuterium oxide and 0.5µl of 2M sodium azide was added. The pH was checked and adjusted as necessary.

For experiments on spectrometers with an attached cryoprobe, 10% deuterium oxide was added. For samples at 900MHz that were not in a shigemi tube, 600µl sample volume was required to ensure good shimming.

2.6.2.2 Paxillin His-LD1/LD2

The purified paxillin protein was stored as a freeze-dried pellet at -20ºC prior to use. Known quantities of protein were dissolved into the necessary buffer to ensure the correct final protein concentration. Remembering that paxillin His-LD1/LD2 cannot be concentrated in the normal manner using spin concentrators, it was important to calculate the correct volume of buffer to use, the protein pellet will make up some of the final volume so it was necessary to add the buffer in small amounts to make the correct final volume. The sample was dissolved by pippetting up and down into the solution, care was needed to ensure there was no frothing of the viscous protein sample. The protein concentration was checked using an A280nm reading and corrected for the predicted absorbance at 1g/L.

2.6.3 Data acquisition and processing

All experiments were carried out at a temperature of 25°C. At 600MHz the proton carrier frequency was set to the water resonance (4.7791ppm).

Gradient-enhanced ^{1}H-^{15}N-HSQC experiments (222) for Vt were typically acquired with 32 scans and 84 increments, higher signal-to-noise was achieved when required with 128 scans and 128 increments. At 600MHz the proton spectral width was 10006.3Hz and the proton acquisition time was 0.064s. The nitrogen acquision time was 0.054s and 0.083s for ni=84 and ni=128 respectively. The ^{15}N spectral width was 1550Hz and the ^{15}N carrier was set to 155.044ppm. The decoupler offset frequency (dof2) for ^{15}N was 1580Hz.

All 3D data was acquired using standard pulse sequences from the BioPack software (Varian). 600MHz ^{1}H-^{15}N-NOESY-HSQC data was acquired with a mixing time of 120ms, 600MHz ^{1}H-^{15}N-TOCSY data was acquired with a 30ms mixing time, 800MHz ^{1}H-^{15}N-NOESY-HSQC data was acquired with a mixing time of 90ms and 900MHz ^{1}H-^{15}N TROSY-NOESY data was acquired using a mixing time of 100ms.

All processing was carried out in NMRpipe format (264). Phasing was corrected using NMRDraw (264) and spectra were visualised using NMRview5 (265).

2.6.4 Methods for the assignment of Vt/I997S

2.6.4.1 Data collection

Table 2.6.1 shows all triple resonance and ^{15}N-edited data that was collected on Vt/I997S to enable assignment.

Table 2.6.1 Data collected for assignment of Vt/I997S

Experiment	TROSY	Field MHz[(a)]	Buffer Conditions	Protein label & Concentration
HNCACB	No	600	20mM Tris pH7, 300mM NaCl	15N, 13C, 60-70% D_2O, 0.5mM
HNCA	No	600	20mM Tris pH7, 300mM NaCl	15N, 13C, 60-70% D_2O, 0.5mM
HN(CO)CA	No	600	20mM Tris pH7, 300mM NaCl	15N, 13C, 60-70% D_2O, 0.5mM
15N-TOCSY-HSQC	No	600	20mM Tris pH7, 300mM NaCl	15N, 0.5mM
15N-NOESY-HSQC	No	600	20mM Tris pH7, 300mM NaCl	15N, 0.5mM
HNCACB	Yes	800	20mM Tris pH7, 300mM NaCl	^{15}N, ^{13}C, 60-70% D_2O, 0.5mM
HNCA	Yes	800	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, 60-70% D_2O, 0.5mM
HNCA	Yes	900	20mM Tris pH7, 300mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, 60-70% D_2O, 0.5mM
HN(CO)CA	Yes	900	20mM Tris pH7, 300mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, 60-70% D_2O 0.5mM
HNCA	No	900	20mM Tris pH7, 150mM NaCl	15N, 13C, 60-70% D_2O, 0.55mM
HNCACB	No	900	20mM Tris pH7, 150mM NaCl	15N, 13C, 60-70% D_2O, 0.55mM
15N-NOESY-HSQC	No	800	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 0.6mM

HNCACB	Yes	800 Cryo	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, >95% D_2O, 0.5mM
HN(CO)CACB	Yes	800 Cryo	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, >95% D_2O, 0.5mM
15N-TROSY-NOESY	Yes	900 Cryo	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, >95% D_2O, 0.5mM
HN(CO)CA	Yes	800 Cryo	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, >95% D_2O, 0.4mM
HNCO	Yes	800 Cryo	20mM Tris pH7, 150mM NaCl, 50mM Arginine, 50mM Glutamic acid	15N, 13C, >95% D_2O, 0.4mM

(a) Cryo: data acquired with cryoprobe

2.6.4.2 Assignment procedure

The method used to enable assignment of Vt/I997S began with initial NH connectivities being identified using amide NOE data from ^{1}H-^{15}N-NOESY-HSQC spectra. When a set of NH resonances were linked, the triple resonance experiments were used to confirm the connectivities and to possibly extend them further. Any amino acid specific information was noted from the triple resonance experiments and the carbon α and β chemical shift information submitted to the 'type-prob' and 'seq-prob' programs (266) that returns the most likely amino acid type and position in the sequence for these connected residues. These probabilities were confirmed and the correct position chosen using the known lysine and arginine residues from selective labelling and unlabelling. A final check was carried out using the triple resonance data to ensure that all connectivties were correct.

The Cα and Cβ shifts were recorded and the deviation from random coil calculated using values from (252). These were plotted against residue number to predict secondary structure and to verify that assignment was correct.

2.6.4.3 TALOS

The secondary chemical shift data was used to predict the ϕ and φ torsion angles using the program TALOS (Torsion Angle Likelihood Obtained from Shifts and sequence similarity) (267). TALOS contains a database of 186 proteins whose x-ray crystal structures are known to 2.2Å, and searches the database to find the 10 best matches for a given amino acid triplet. It uses the triplet to give data about the central residue. If these matches have consistent values then their averages and standard deviations are used as an average for a prediction for the given data.

The predictions for Vt/I997S were compared to the .PDB file for wild-type Vt (1QKR) to determine similarities and differences.

2.6.5 Estimation of rotational correlation time

Relaxation experiments were carried out using 100μM Vt in 20mM Tris pH7, 400mM NaCl. ^{15}N relaxation time constants were calculated from a series of one-dimensional ^{15}N T1 and T2 relaxation spectra recorded with relaxation delays of 0.02s, 0.08s, 0.16s, 0.32s, 0.60s, 1.00s, 1.40s for T1 and 0.01s, 0.03s, 0.05s, 0.07s, 0.09s, 0.11s for T2.

To extract the correlation time, the amide regions recorded for both T1 and T2 at each delay were integrated from 7.62ppm-12ppm, avoiding the NH_2 peaks. This provided the intensity as the area under the peaks. The first integral value was normalised to 1 and the remaining values were measured compared to this. These intensities (I/I_0) were plotted against relaxation delay (s) using the xmgrace program. The results of this plot were fitted using the non-linear curve fitting algorithm in xmgrace with the parameters in Equation 2.iv and the a1 value was extracted.

Equation 2.iv $y=a0*\exp(-a1*x)$.

The y value corresponds to I/I_0 and the a1 value corresponds to either $1/T_1$ or $1/T_2$. T_1/T_2 was calculated for each data set and was inputted into a program called '$t_m$est' which estimates correlation time based on the approach by Kay *et.al* (242).

2.6.6 Hydrodynamic modelling

The Vt sequence for wild-type and mutants 892-1061 was submitted to the hydrodynamic modelling program HYDRONMR (240;244). The residues 878-891 and 1062-1066 were removed as they are expected to be flexible as judged by the crystal structure. Structure files containing monomeric protein and dimeric protein found in the asymmetric unit were submitted. The suite of HYDRO programs models behavior of macromolecules by producing a shell model that creates the surface of the protein by using small beads. This 'bead model' can be used to predict correlation times for a macromolecule in solution. By submitting a structure file of a monomer and dimer, correlation times for monomeric and dimeric protein can be predicted and from this it is possible to estimate the proportion of dimer in solution using Equation 2.v. This estimation of proportion of dimer assumes that the dimer found in the crystal interface is also occurring in solution.

Equation 2.v $1/\tau_m = (\propto D/\tau_{dim}) + (\propto M/\tau_{mon})$

τ_m = measured correlation time

$\propto D$ = proportion of dimer

$\propto M$ = proportion of monomer

τ_{dim} = predicted dimer correlation time

τ_{mon} = predicted monomer correlation time

2.6.7 Chemical shift perturbation

Chemical shift perturbation experiments have been carried out at 600MHz and 900MHz.

Table 2.6.2 Chemical shift perturbation experiments that have been carried out

	Field (MHz)	[NaCl]/mM	50mM Arg, 50mM Glu
LD1 peptide	600	400	None
	600	150	None
	600	50	Yes
LD2 peptide	600	400	None
	600	150	None
	600	50	Yes
LD4 peptide	600	400	None
	600	150	None
	600	50	Yes
Paxillin His-LD1/LD2	900	150	None
	900	50	Yes

2.6.7.1 Wild-type Vt plus paxillin peptides LD1, LD2 and LD4

2.6.7.1.1 Sample preparation

The peptides used for the titrations were synthesised by Peptide Protein Research Ltd. These 13mer peptides represent the paxillin LD motifs LD1 residues 1-13, LD2 residues 141-153 and LD4 residues 262-274 (human numbering). His-LD1/LD2 was expressed and purified in the lab. Figure 2.6.1 shows the LD peptides and the longer His-LD1/LD2 construct used.

LD1: MDDLDALLADLES M_M: 1420.537
LD2: NLSELDRLLLELN $M_{M:}$ 1541.746
LD4: ATRELDELMASLS M_M: 1435.598

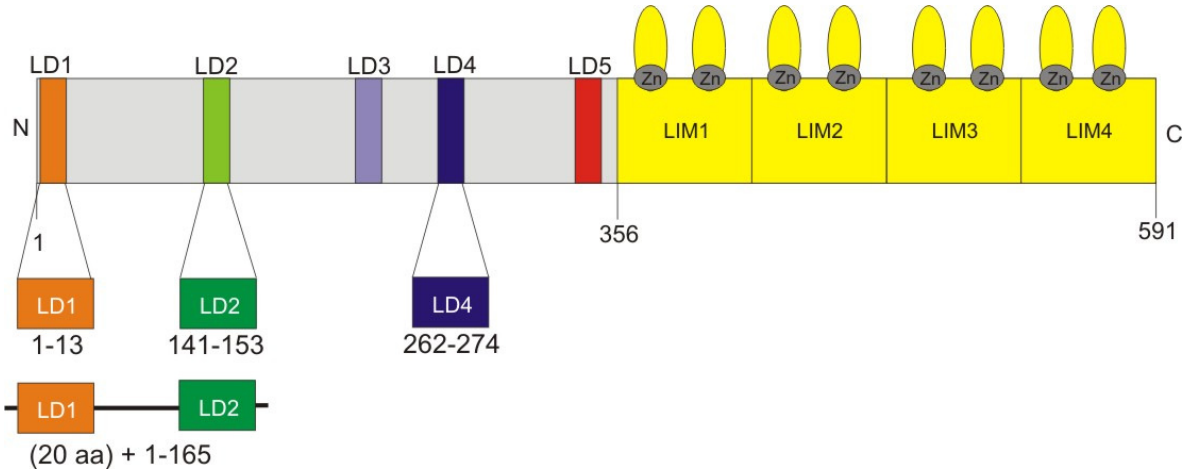


Figure 2.6.1 LD constructs used in this thesis

Wild-type Vt was at 200µM in one of the following buffers: 20mM Tris pH7, 400mM NaCl; 20mM Tris pH7, 150mM NaCl; or 20mM Tris pH7, 50mM NaCl, 50mM arginine, 50mM glutamate.

To carry out the titration, two protein samples were required at the same concentration and volume. One contained Vt at the desired concentration (solution A), the other contained both Vt and one LD peptide at a molar excess of 20x (solution B). Solution B was used to titrate into solution A which ensures that the concentration of Vt does not change throughout the experiment.

In order to overcome large pH changes on addition of excess peptide to Vt, Vt and peptide were prepared at a concentration that is double the final desired concentration. For example, to prepare 1mL of Vt at 200µM, 500µl at 400µM would be made in the appropriate buffer. To prepare 500µl of LD peptide at 2mM, a 4mM solution of LD peptide would be made in 250µl of the identical buffer and the pH adjusted to 7.

The 250µl of LD peptide was added carefully to 250µl of Vt solution producing a 200µM protein solution containing 2mM LD peptide. To the remaining 250µl Vt, the same volume of buffer was added to make a 500µl Vt solution at 200µM. The pH of both solutions was

adjusted so that they are identical (pH7). 5% deuterium oxide and 0.5µl of 2M sodium azide was added to both samples.

2.6.7.1.2 Experimental procedure

A titration plan was made prior to conducting the experiments. A ^{1}H-^{15}N-HSQC spectrum was run on the free Vt (solution A) and on the Vt plus 20x molar excess LD peptide (solution B). The titration was performed by removing a calculated volume from solution A and then replacing it with the identical volume of solution B. To calculate the amount of solution B to add, a titration planner created by Dr. Stuart Findlow (University of Southampton) was used. This calculates the volumes by using Equation 2.vi.

Equation 2.vi

$$\delta = \frac{\left([L_f] - [L_i]\right) \times vol}{\left([L] - [L_i]\right)}$$

δ = volume changed (amount to remove and add back into sample)
$[L_f]$ = final ligand concentration in sample
$[L_i]$ = initial ligand concentration in sample
vol = total sample volume
[L] = concentration of stock ligand

Molar ratios of Vt:LD peptide chosen for titration were 0, 0.25, 0.5, 0.75, 1.0, 2.0, 5.0, 10.0. For each point in the titration series, a ^{1}H-^{15}N-HSQC spectrum was acquired with 32 scans and 84 increments. Before each titration point the spectrometer was reshimmed.

2.6.7.2 Wild-type Vt plus paxillin His-LD1/LD2

2.6.7.2.1 Sample preparation

Paxillin His-LD1/LD2 was stored as a freeze-dried solid at -20ºC. It cannot be concentrated using spin concentrators, so it needs to be dissolved into solutions to a known final concentration.

For titration at 900MHz, 1.2mL of ^{15}N-wild-type Vt was prepared at 200µM in 20mM Tris pH7, 50mM NaCl, 50mM arginine, 50mM glutamate. 600µl was taken for the start point of the titration and 10% deuterium oxide and 0.5µl of 2M sodium azide was added. The pH was checked and adjusted to 7 as necessary. The remaining 600µl was used to dissolve ^{14}N-His-LD1/LD2 that has been freeze dried in known quantities into glass vials. Vt was added into the vials and the solid was dissolved carefully so that bubbles were not created in the Vt solution. His-LD1/LD2 was dissolved into Vt at a molar excess of 6x.

2.6.7.2.2 Experimental procedure

A titration plan was made prior to conducting the experiments as shown in section 2.6.7.1.2. A ^{1}H-^{15}N-TROSY spectrum was run on the free Vt (solution A) and then on the Vt plus 6x molar excess paxillin His-LD1/LD2 (solution B). The titration series and method was conducted for titration points 0, 0.25, 0.5, 1, 2 and 4x molar excess. TROSY experiments were acquired with 16 scans, 128 increments.

2.6.8 Hydrogen-deuterium exchange experiments

Hydrogen-deuterium exchange experiments were carried out on 500µL volumes of wild-type Vt alone, wild-type Vt plus a 10-fold excess of paxillin peptides LD1, LD2 and LD4, and with a 4-fold excess of paxillin His-LD1/LD2. All Vt samples were at 200µM, 20mM Tris pH7, 50mM NaCl, 50mM arginine, 50mM glutamate and Vt plus ligand samples were prepared as in section 2.6.7.1.1.

A ^{1}H-^{15}N-HSQC spectrum was recorded of 200µM wild-type Vt at pH7 with 32 scans and 84 increments. The precise volume of the solution was measured and then the entire sample was freeze-dried for 16 hours. The structure of Vt is intact after freeze-drying. Immediately prior to running the second ^{1}H-^{15}N-HSQC experiment, the identical volume of deuterium oxide was added to re-dissolve the freeze-dried Vt. The pH was checked to ensure it complied with pH = pD – 0.41 (208). The second ^{1}H-^{15}N-HSQC experiment was run and the time between adding deuterium oxide and starting the experiment was measured (this time needs to be minimised). In these experiments the time was 13 minutes. A series of consecutive ^{1}H-^{15}N-HSQC spectra were collected for the next 24 hours and a final spectrum was recorded 50 hours post D_2O addition. Each spectrum is acquired in a period of 2 hours.

This was repeated for all Vt plus ligand samples, and the time between re-dissolving in D_2O and beginning the series of HSQC experiments was maintained at 13 minutes.

2.7 GST Pulldown and GST Catch-up experiments

2.7.1 GST-pulldown assay

For the paxillin-GST-LD1/LD2 only control: 500µl of cell lysate from paxillin-GST-LD1/LD2 was bound to 200µl of glutathione 4B fastflow resin by incubating for 1hour at room temperature. The resin was then spun down at 5000rpm in a microfuge and the supernatant removed. The resin was washed once with 20mM Tris pH7, 150mM NaCl and the resin was spun down and the wash removed. The paxillin GST-LD1/LD2 construct was next eluted using 10mM glutathione and the elution fraction run on an SDS-PAGE gel to visualise the protein. For the paxillin-GST-LD1/LD2 plus wild-type Vt: 500µl of cell lysate from paxillin-GST-LD1/LD2 was incubated together with 1mL of purified wild-type Vt at 1mg/mL for 1 hour. The method was followed as above, and the interaction was visualised on an SDS-PAGE gel. For the GST control: Purified GST was bound to glutathione 4B fastflow resin and then washed to a UV baseline. Purified Vt at 1mg/mL was loaded onto the resin containing the GST fusion protein and washed to a UV baseline.

The resin was eluted using 20mM reduced glutathione and interactions observed on an SDS-PAGE gel.

2.7.2 Catch-up comparative chromatographic retention assay

2.7.2.1 Principles

A method developed by Charbonnier *et al.* (268) has been used in this study, the Holdup (or Catch-up) Comparative Chromatographic Retention Assay. This assay is based on the principles of reversible binding of fusion proteins to affinity resins, but is distinct from a GST-pulldown experiment as no washing steps are included.

GST-pulldown experiments are based on the fact that GST-fusion proteins will bind reversibly to an affinity matrix. When bound, the fusion protein can be used to 'fish' for interaction partners. Any protein that elutes with the GST-fusion protein is considered to be an interaction partner. A major problem with GST-pulldown experiments is that washing steps are employed between binding of the partners and observing the final complex. This means that the detection of an interaction is dependent on its ability to survive these washing steps. In other words, strong interactions will survive, but weaker more elusive interactions may not be observed due to the complex being washed away (268).

In this assay, the results obtained are observing interactions at equilibrium, and can be used both as an interaction detection tool and also has a way of evaluating K_d. The authors describe two versions of the assay, the hold-up and catch-up experiments. The difference is that hold-up experiments use affinity resin with ligand pre-bound to the beads, the catch-up experiment pre-incubates both ligands together before binding. In this thesis, the catch-up version has been utilised.

The principle of the experiment is shown in Figure 2.7.1. Ligand fusion protein is mixed with resin and analyte protein in two equal reactions. A competitor molecule is added to

one of the reactions and the liquid phases of both reactions are extracted. Comparing the two extracts reveals the proportion of free and bound analyte at equilibrium. Selective disappearance of the analyte in the competitor-free reaction indicates an interaction.

2.7.2.2 Method

20mL expression grows were carried out of GST-LD1/2 and calexcitin. Purified Vt was used at 1-2mg/mL concentration in 20mM Tris, pH7, 300mM NaCl. The cell pastes were thawed and resuspended in 2.5mL of 20mM Tris pH7, 100mM NaCl and sonicated on ice using a XL-2020 sonicator at power 3 with a microprobe for 12 x 5sec on, 30sec off. Care was taken to ensure no heating or frothing occurred in the sample to allow efficient sonication and prevent damage to the proteins. Fusion protein cell lysate were mixed together and left on ice with gentle agitation for 1 hour according to Table 2.7.1.

Table 2.7.1 Incubation reactions

Reaction	45µl	15µl
1	GST	Pure Vt (1-2mg/mL)
2	GST-LD1/2	Pure Vt (1-2mg/mL)
3	GST-LD1/2	Calexcitin (control)
4	GST-LD1/2	Tris Buffer (control)

Meanwhile, glutathione beads are prepared. For each reaction, two microfuge tubes need to be prepared, one labelled plus (+) and one labelled minus (–). This will give rise to 8 x 1.5mL microfuge tubes. Into each tube, glutathione 4B beads are aliquotted to make 15µl of bed volume. The beads are washed with 1mL of 20mM Tris pH7, 100mM NaCl and spun down at 2000 rpm in a benchtop microcentrifuge for 10mins to remove the supernatant. This is repeated twice with 200µl buffer. When the 15µl bed volume is achieved in each tube, 15µl of 20mM Tris pH7, 100mM NaCl is added to each bed volume.

When the incubation time is complete, each reaction is split in half into the appropriate tubes, adding 30µl of reaction to each. Each microfuge tube is vortexed for 1 minute each. The vortex should be set at a speed so that the beads form a gentle 'swirl' in the reaction mixture. Beads should not be allowed to jump up the side of the tubes else they will be lost. This vortex procedure is repeated three times. The samples are next left for 30 minutes on ice with gentle agitation.

Two buffers are required; the 'minus' buffer (-) containing 20mM Tris pH7, 100mM NaCl; and the 'plus' buffer (+) containing 20mM Tris pH7, 100mM NaCl, 20mM reduced glutathione. To all tubes with the (+) condition, 30µl of the 'plus buffer' is added. To all tubes with the (–) condition, 30µl of the 'minus buffer' is added. The vortex procedure is repeated twice and the reactions are left for 30minutes.

Next, glass wool filters are prepared. Label 8 x 1.5mL microfuge tubes to match the reaction tubes. For each tube, take a 0.5mL microfuge tube and pierce a hole in the tip using a syringe needle. Stuff the small tube with glass wool to form a filter. The corresponding reaction mixture and beads is added to the 0.5mL tubes containing the filter. The small tubes are capped and placed into the 1.5mL microfuge tubes. These are then spun down at 2000 rpm for 5 minutes. The liquid phase will be filtered into the microfuge tube and will contain the proteins of interest.

An SDS-PAGE gel is run of all reactions, the (+) and (-) for each reaction is run adjacent to one another. To load the gel, take 20µl of sample, add 20µl of reducing sample buffer and boil for 5 minutes. Load 20µl into the wells, ensuring that no sample spills into the next wells. Stain and destain.

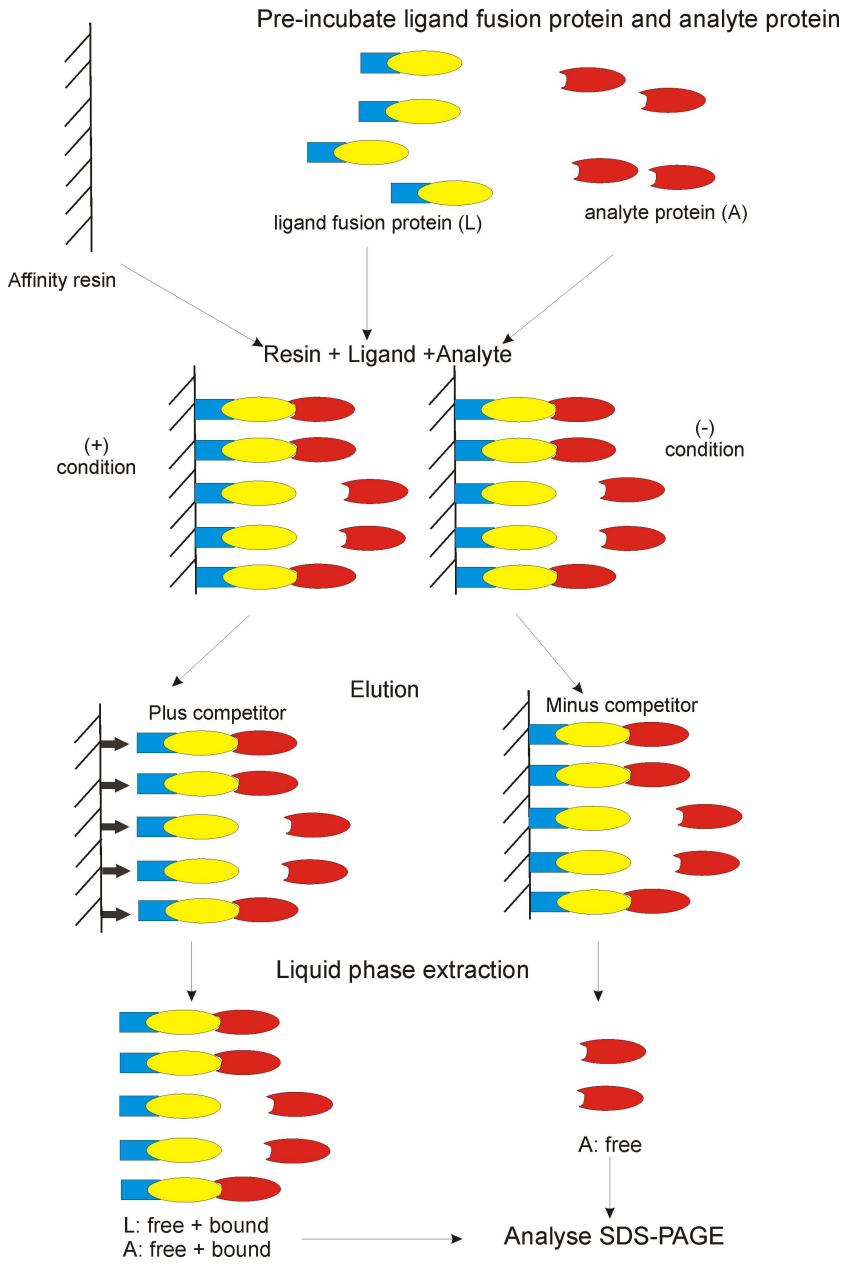


Figure 2.7.1. Principle of the Catch-up assay

The ligand fusion protein is shown in cyan and yellow (L) and the binding partner or analyte is shown in red (A). L & A are pre-incubated and then batch mixed with affinity resin. Half of the reaction is mixed with a buffer containing a competitor molecule for elution, and half is mixed with buffer containing no competitor. The liquid phase is extracted and the resin and any associated proteins are 'caught up' on the resin. Analysis of the liquid phase from both the competitor fraction and non-competitor fraction will give an indication of binding affinity. Figure adapted from (268).

2.8 Preparation of figures

All molecular visulisation has been completed using the program MOLMOL (269) with the exception of Figure 1.3.9 which was created by Dr. Stuart Findlow (University of Southampton) using the program Pymol (DeLano Scientific). NMR spectra were created using NMRview5 (265). All images were manipulated using CorelDRAW12 before importing into Microsoft Word.

3 Results

3.1 Expression, Purification and Characterisation of wild-type Vt

Wild-type vinculin tail (residues 879-1066) has been expressed and purified for use in this study. The construct was provided in the pET-15b construct and contains an N-terminal histidine tag. The plasmid was a kind gift from Professor D. Critchley (University of Leicester) and the sequence can be found in Appendix 5.1.

3.1.1 Expression and purification

Wild-type vinculin tail (Vt) was over-expressed in both Luria Broth and M9 minimal media to produce an induction band well above the *E.coli* background (Figure 3.1.1). The expression band for Vt runs higher than expected for its molecular weight of 21586.8. Expression in LB takes four hours from induction at 30ºC and typically yields 8mg/L. Expression levels of wild-type Vt in M9 minimal media is comparable, although the length of expression is extended to 16 hours at 30°C.

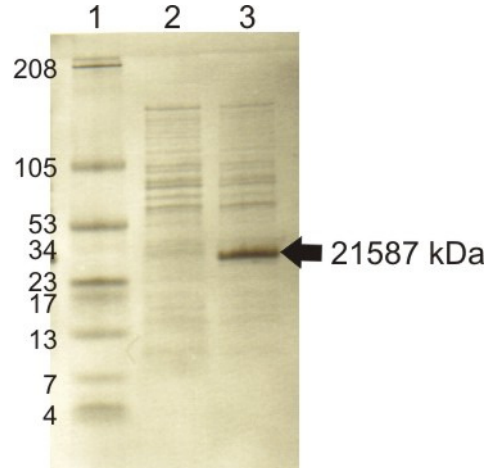


Figure 3.1.1 Polyacrylamide SDS-Page gel showing induction of wild-type Vt using LB medium after induction with 1mM IPTG for 4 hours.

Lane 1: molecular weight standard-multimark, with molecular weights of products alongside (in kDa); lane 2: pre induction sample; lane 3: 4 hour post induction. Induction bands can be seen higher than expected for a protein of 21kDa.

The purification of wild-type Vt has been adapted from Bakolitsa *et.al*. (172) (see Section 2.2.7). The first stage of purification is a Nickel-NTA affinity column, a typical chromatogram is shown in Figure 3.1.2. The 100% elution fraction (250mM imidazole) is pooled and the 280nm absorbance measurement is recorded which enables determination of the amount of thrombin to add for histidine tag removal. The total amount of Vt in this elution pool is inaccurate due to the presence of 250mM imidazole which also absorbs at 280nm, however this value is sufficient for determining the concentration of thrombin required.

Thrombin is added at 1 unit per 1-2mg of protein, depending on the amount of thrombin required and the time that is available for cleavage. The rate at which thrombin cleaves the histidine tag from Vt can vary; generally 1 unit of thrombin per mg of protein, plus the addition of 5mM $CaCl_2$, will remove >99% of the tag after 21 hours at 35°C. It is necessary to test the extent of cleavage on SDS-PAGE gel prior to loading onto the next column. Vt is very robust as thrombin appears to only cleave Vt at the cleavage site, so leaving the reaction longer than required is not detrimental to the sample.

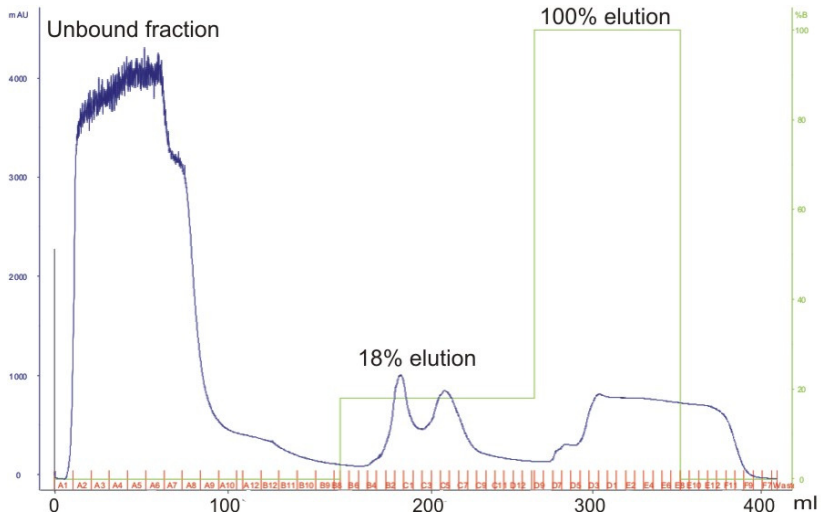


Figure 3.1.2 Wild-type Vt Ni-NTA loading and elution profile.

The blue trace is the absorbance at 280nm and is used to monitor the concentration of protein. The green trace shows the percentage of imidazole being used to elute, the first elution is 18% and the second is 100%. Fractions are shown in red along the x-axis.

A typical chromatogram from the cation column is shown in Figure 3.1.3. Vt elutes between 400-600mM NaCl and is greater than 95% pure after this step.

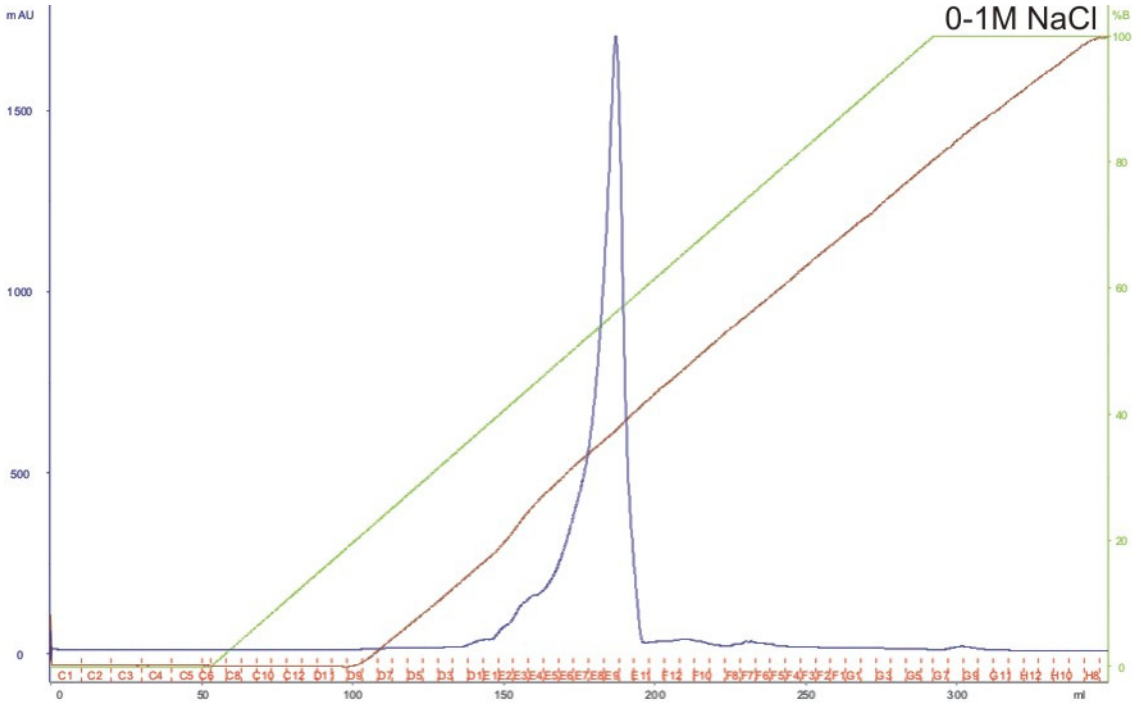


Figure 3.1.3 Wild-type Vt SP Sepharose (cationic) elution profile.

The blue trace is the absorbance at 280nm and is used to monitor the protein. The green trace shows the percentage of 1M NaCl being used to elute, the gradient is run from 0-100%. The brown line is a measure of the conductivity of the elution fractions. Fractions are shown in red along the x-axis.

Figure 3.1.4 shows an SDS-PAGE gel of ^{15}N labelled wild-type Vt as it progresses through the purification. The ^{15}N-labelled protein has very little impurities remaining after the Ni-NTA step in contrast to the ^{14}N-labelled preparation. After thrombin cleavage, the cation exchange step is utilised to remove the histidine tag and the thrombin from the sample, although this is never detectable on SDS-PAGE gels.

After purification, Vt is concentrated to 1mg/mL directly from the cation column for storage at -20^{o}C. Vt benefits from being concentrated close to the desired final concentration in the high NaCl elution buffer or in a buffer with 50mM arginine and 50mM

glutamate as additives before dialysis into lower NaCl buffers. Section 3.3.5 discusses the buffer choice for Vt.

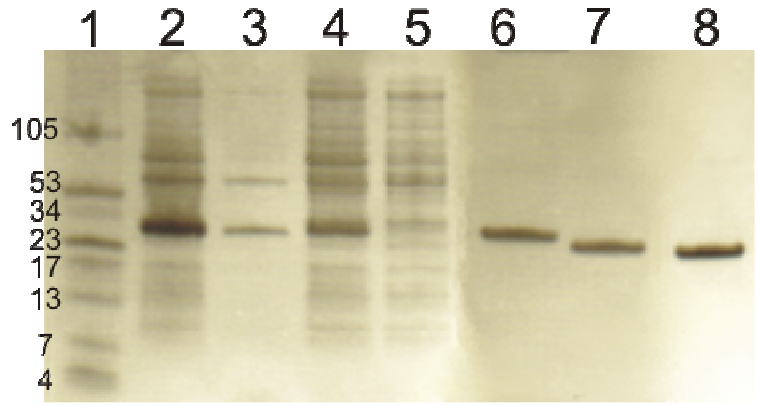


Figure 3.1.4 Polyacrylamide SDS-Page gel showing wild-type Vt at various stages of the purification.

Lane 1: molecular weight standard-multimark, with molecular weights of products alongside (in kDa); lane 2: sonicated homogenate-the major band is His-Vt; lane 3: insoluble material post sonication; lane 4: Ni-NTA load, His-Vt is the major band; lane 5: Ni-NTA flow-through showing that His-Vt has bound to the affinity column; lane 6: pooled His-Vt from Ni-NTA column, pre thrombin cleavage; lane 7: Vt post thrombin cleaved/cation load; lane 8: final wild-type Vt at 1mg/mL.

Analysis of the different methods of storage show that freezing at -20ºC does not appear detrimental to the sample. Figure 3.1.5 shows the amide region of a one-dimensional proton spectrum for wild-type Vt that has been stored by different methods. It was interesting to note that leaving Vt at 4ºC for four days in solution leads to a reduction in intensity of signal. This was important to realise as it means some urgency is required when handling these samples in NMR solution studies. Freezing at -20ºC was deemed an acceptable method of storage as the intensity and appearance of the spectrum was good, and freeze-drying also appears non-detrimental to the sample but was not chosen as the preferred method.

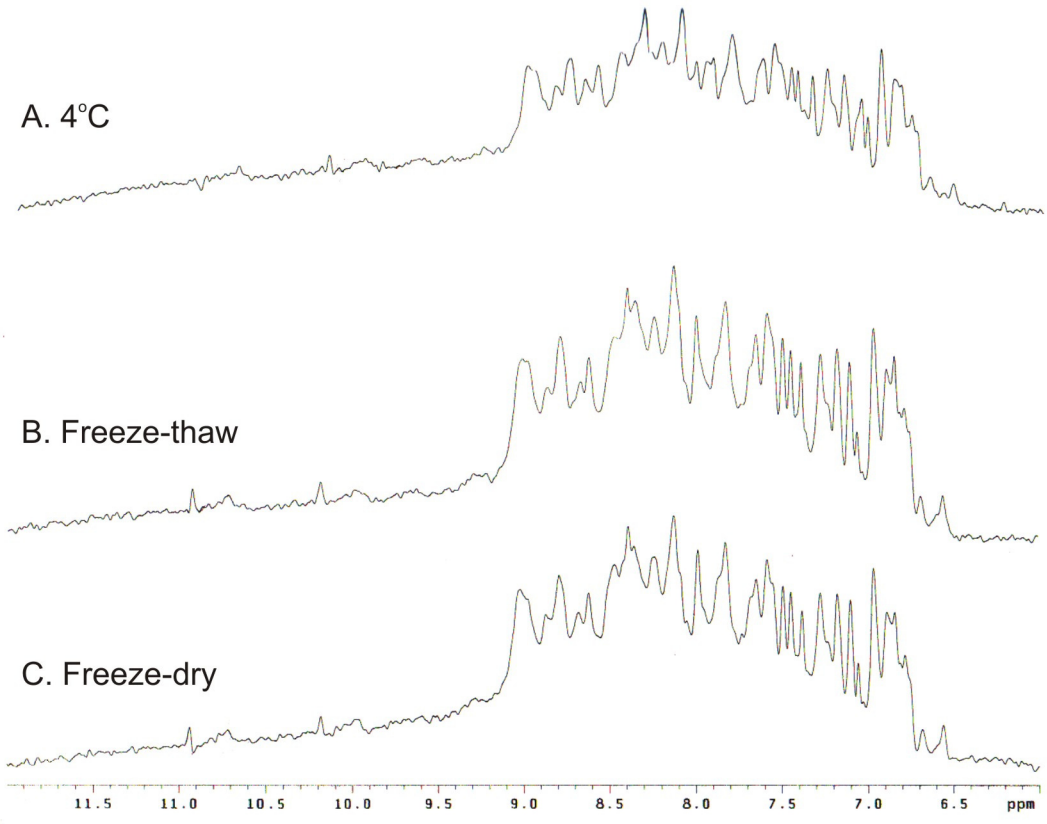


Figure 3.1.5. The amide region of a one-dimensional proton spectrum of Wt-Vt measured at 600MHz.

A: Storage at 4ºC; B: After freezing at -20ºC; C: After freeze-drying and resuspension in water. Vt was at 200µM in 20mM Tris pH7, 400mM NaCl.

3.1.2 Characterisation of wild-type Vt

3.1.2.1 Mass spectrometry

The mass spectrum of ^{15}N-labelled wild-type Vt shows that the molecular weight obtained from the purified protein matches that which is expected from the amino acid sequence and shows high incorporation of ^{15}N (Figure 3.1.6 and Table 3.1.1).

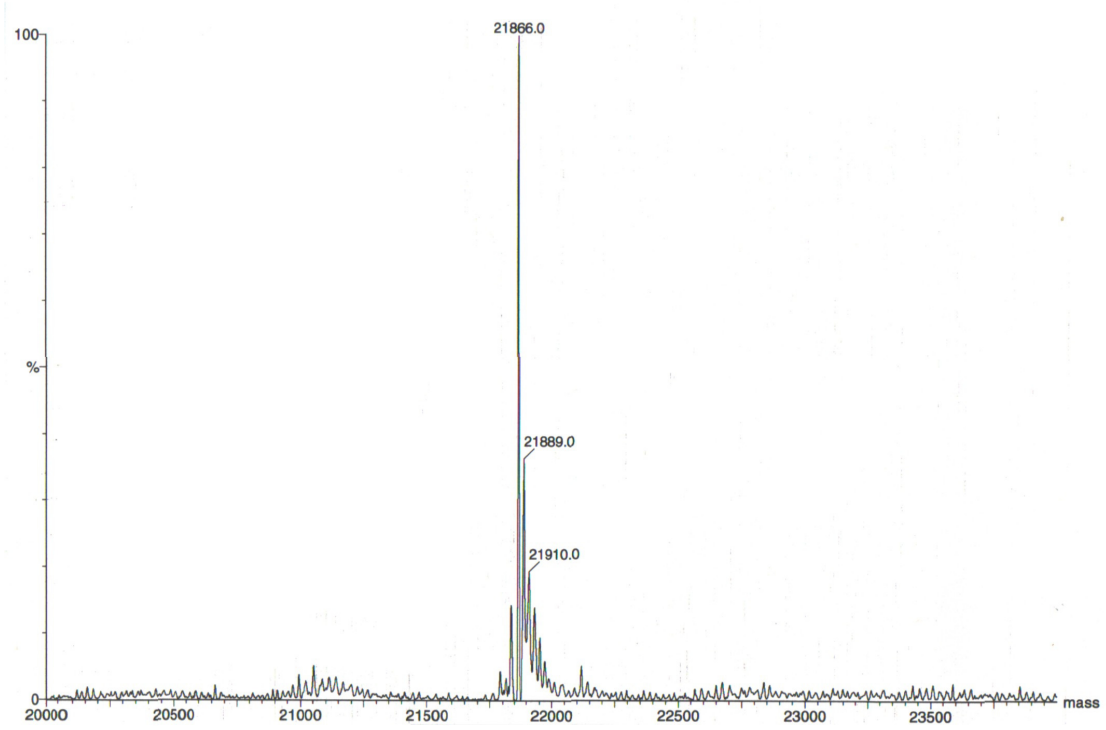


Figure 3.1.6 Electrospray ionization mass spectrometry (ESI MS) of ^{15}N wild-type Vt

Major peak at 21866 corresponds to wild-type Vt, peaks at 21889 (+23) and 21910 (+21) are likely to be sodium adducts

Table 3.1.1 Predicted molecular weight of ^{14}N and ^{15}N Wt-Vt and the measured molecular weight by mass spectrometry

	^{14}N-Wt-Vt	^{15}N-Wt-Vt
Predicted molecular weight[(a)]	21586.8	21866.8
Measured molecular weight	Not done	21866.0

(a) predicted using protparam (270)

3.1.2.2 Analytical size exclusion chromatography

Initial studies using analytical size exclusion chromatography suggests that wild-type Vt exists as a monomer (Figure 3.1.7). The molecular weight calculated from the column

calibration appears smaller than expected, with 17kDa being calculated from an expected 21kDa.

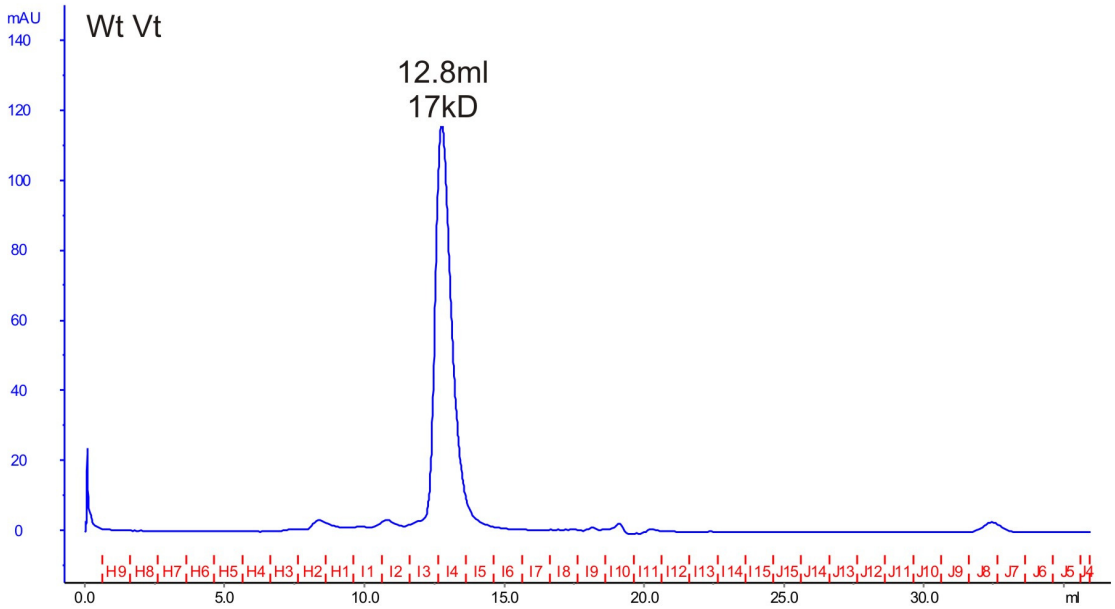


Figure 3.1.7 Analytical gel filtration (superdex-75) of Wt-Vt.

The blue trace is the protein absorbance at 280nm. Column was calibrated and run using 50mM Na_2HPO_4, 150M NaCl. 100µl of 1mg/ml Vt was injected.

3.1.2.3 Secondary structure by circular dichroism

The crystal structure of wild-type Vt shows it to be highly α-helical with no other secondary structure (588). Far UV circular dichroism has been carried out to determine whether the purified Vt protein assumes an α-helical conformation.

Wild-type Vt gives rise to a far UV spectrum characteristic of α-helical proteins, with CD minima at 208nm and at 223nm (Figure 3.1.8). Submission of this data to an online server that predicts secondary structure (SOMCD (212)) predicts that 98.7% of the secondary structure is α-helical, 0.6% is due to β-turn and 0.7% is due to random/other structural elements. None is β-sheet. This prediction is a little higher than expected, but is consistent with a high α-helical content.

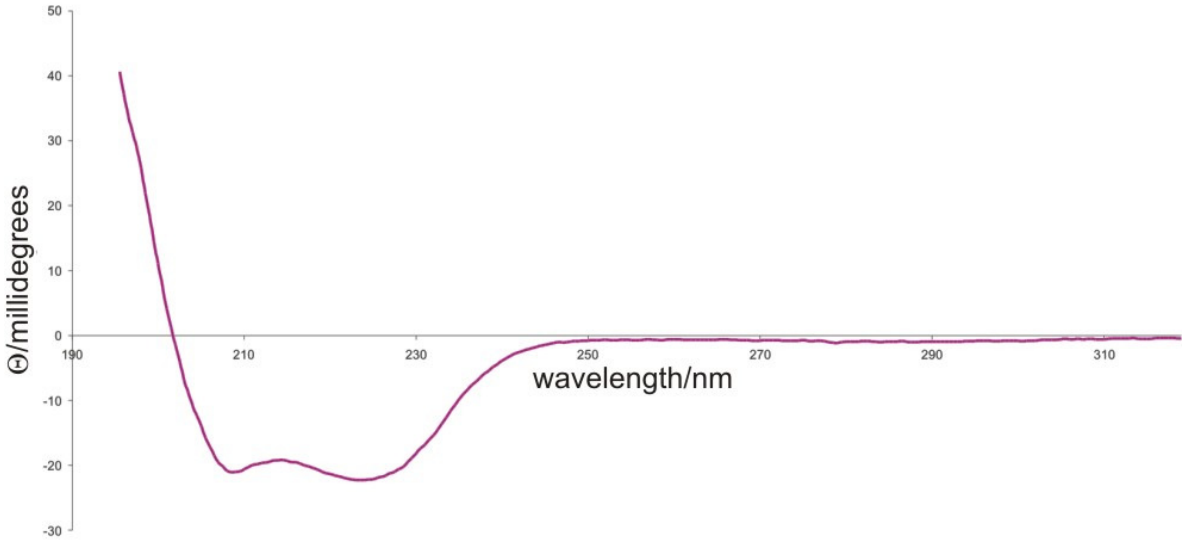


Figure 3.1.8 Far UV circular dichroism spectrum of wild-type Vt at 0.5µM in water.

3.1.2.4 ^{1}H-^{15}N-HSQC spectrum

A typical ^{1}H-^{15}N-HSQC is shown for wild-type vinculin tail in Figure 3.1.9. The spectrum shows well dispersed NH resonances between ~6.8ppm-10.7ppm in the proton dimension which is indicative of a folded protein. The wild-type Vt construct contains 192 residues. Since the four prolines do not give rise to a signal in an ^{1}H-^{15}N-HSQC spectrum and the first residue is not usually seen because it exchanges too quickly with the solvent, 187 NH resonances are expected to be seen from the backbone amides. In addition 3 peaks are expected from the indole NH of the tryptophan side chains. Also expected are 18 resonances arising from the NH_2 groups of asparagines and glutamines. These are easily identified as they have two distinct proton shifts corresponding to the same nitrogen shift. Discounting the signals arising from NH_2 groups, an initial peak count in the spectrum produces 180 peaks from an expected 190. The difference could be due to resonance degeneracy of some residues or to the fact that some residues are in intermediate exchange leading to undetectable, broad peaks.

At 600 MHz, the resonances at the centre of the spectrum are not well dispersed and there is some overlap with other peaks where there are two (or more) peaks with the same chemical shift that appear as one large peak in the spectrum.

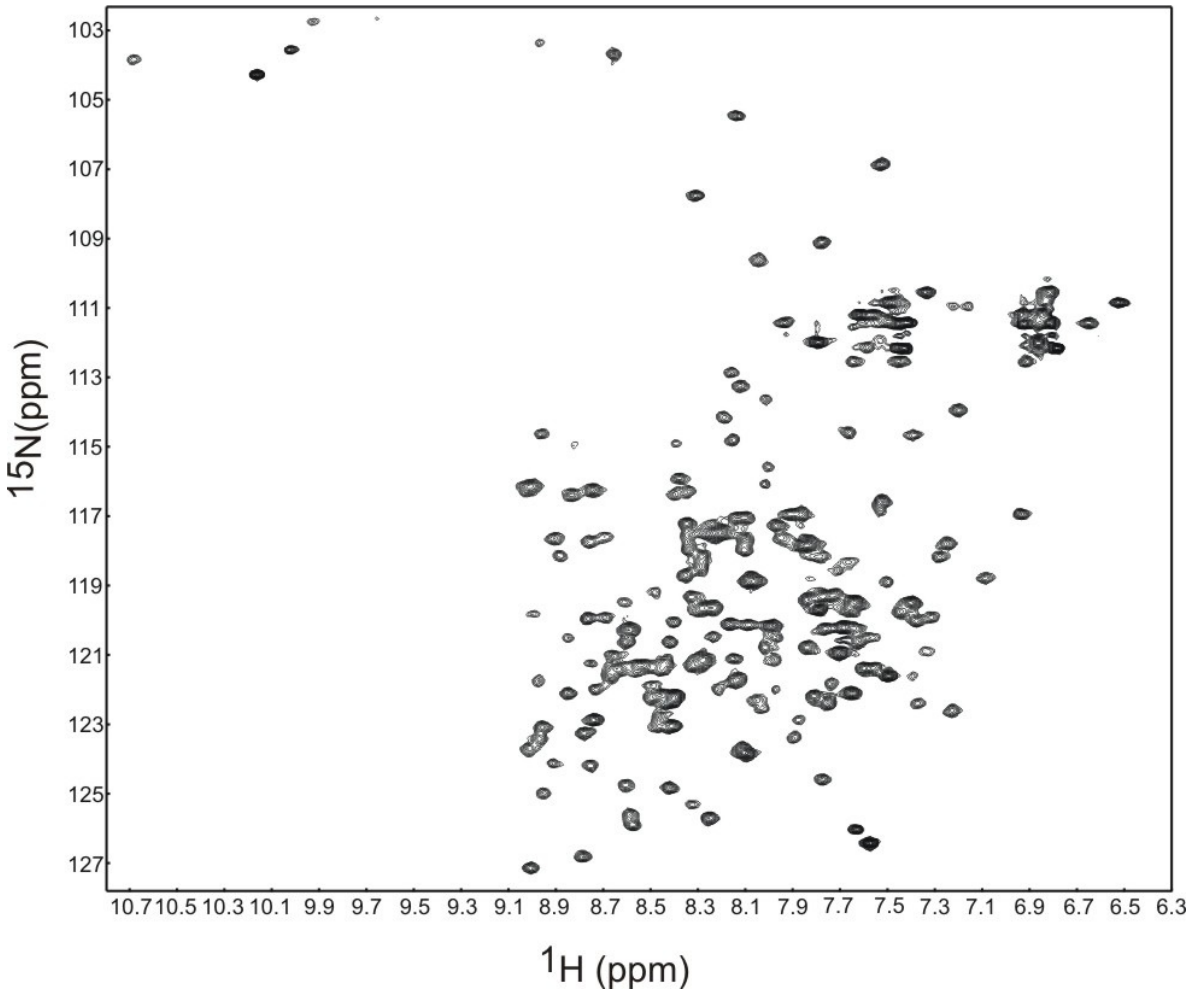


Figure 3.1.9 ^{1}H-^{15}N-HSQC spectrum of ^{15}N wild-type Vt

Vt concentration 200µM in 20mM Tris pH7, 300mM NaCl. Spectrum acquired for 12hrs at 600MHz, spactral folding results in the indole peaks appearing at the top of the spectrum.

3.1.2.5 Rotational correlation time

Rotational correlation times of proteins can be extracted from ^{15}N relaxation measurements and can be related to the molecular weight of a protein. The overall correlation time for wild-type Vt has been determined from averaged ^{15}N-relaxation times.

Figure 3.1.10 shows the first block from a 2D-^{15}N relaxation experiment for both T_1 and T_2 together with the integrated intensities. Intensities were plotted for increasing relaxation delay times and used to determine T_1 and T_2 values. From the data in Figure 3.1.10, values for T_1 and T_2 relaxation time constants were 1.054 and 0.062 respectively. These were used to calculate an experimentally measured correlation time (τ_m) of Vt in 20mM Tris pH7, 400mM NaCl at 100μM of 13.0ns.

As the crystal structure of Vt contains two molecules in the asymmetric unit, the monomer and dimer can be used to estimate the correlation time using a hydrodynamic bead model (244). For Vt residues 892-1061 the correlation time for a monomer was estimated to be 8.2ns and for a dimer it was estimated to be 20ns. Assuming a simple monomer-dimer equilibrium one can estimate the mole fraction of molecules that are in a monomeric state and that are in a dimeric state. The calculation predicts that wild-type Vt at 100μM is 63% dimer. This result gives an indication that Vt may exist as an oligomer or that aggregation may be occurring in solution between Vt molecules.

As previously mentioned in Section 1.3.3.1, there is a crystallographic dimer interface centered at residues isoleucine I997 and methionine M1022 which is located on the most hydrophobic part of the Vt molecule. It is possible that this interface is having an effect in solution. Similarly, there is a potential interaction between the N-and C-termini of Vt molecules which could also be affecting the oligomerisation state of Vt in solution.

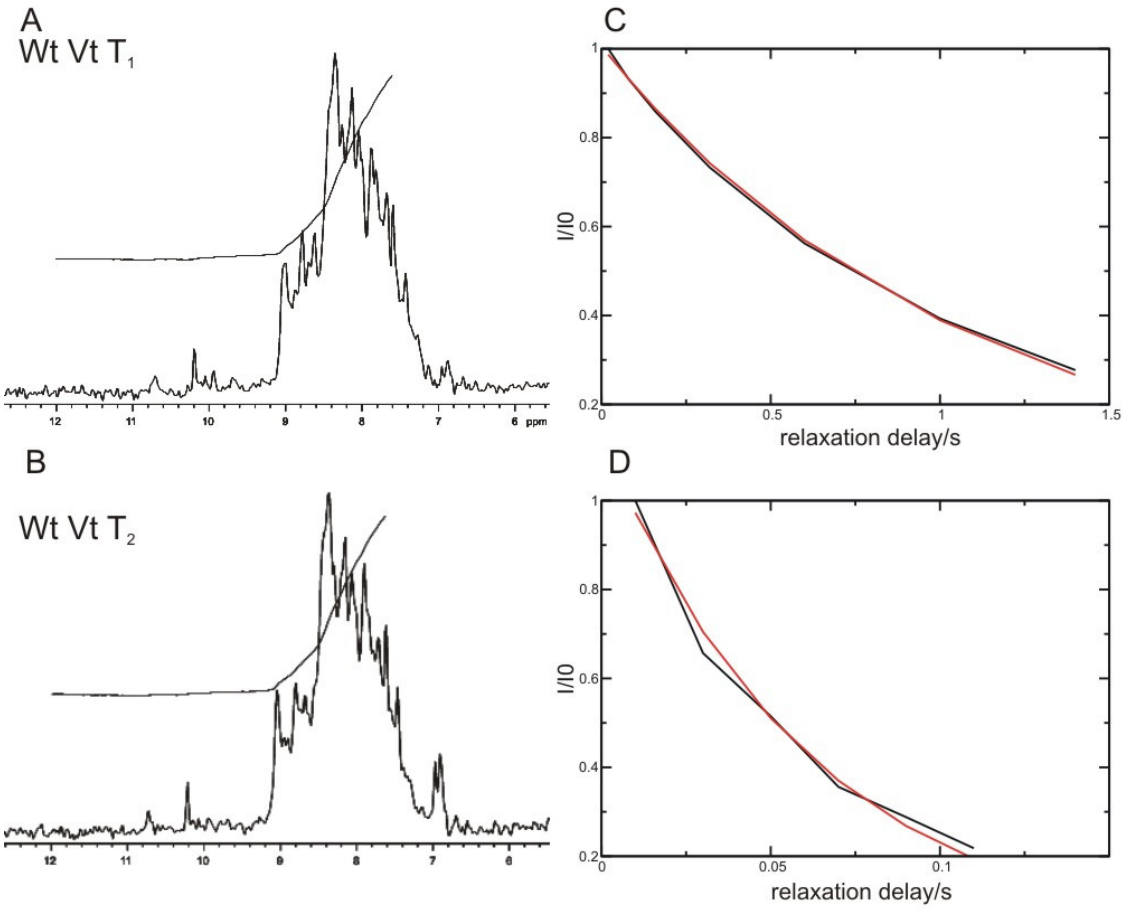


Figure 3.1.10 Results of ^{15}N relaxation experiments on wild-type Vt

A: amide region of the first relaxation delay for T_1=0.020ms, B: amide region of the first relaxation delay for T_2=0.010ms. The integration curve that was used is shown as a black line. C and D: graph used to determine T_1 and T_2 values, black line represents actual data and red line represents the curve fit.

3.1.2.6 Wild-type Vt summary

Recombinant wild-type Vt has been successfully expressed and purified and can be labelled with ^{15}N. Far-UV circular dichroism studies confirm that the protein is α-helical, as predicted from the crystal structure. Whilst gel filtration suggests that Vt exists as a monomer, measurement of the correlation time suggest that there could be transient higher order oligomers in solution.

3.2 Mutagenesis of Vt

Site directed mutagenesis is a technique used to introduce specific mutations in DNA molecules at predetermined positions. There are several methods of site directed mutagenesis available, the method chosen for this project is the Quikchange method from Stratagene, and while strictly speaking is a non-PCR method, does use PCR–type denaturation and annealing steps to introduce mutations.

The polymerase chain reaction (PCR) was originally developed by Kary Mullis and co-workers (271). The PCR method of site directed mutagenesis uses PCR to insert and amplify specific mutations into a template DNA. There are commonly three steps to PCR; the denaturation step where the DNA template and mutagenic primers are melted so that they separate from each other, the annealing step where the primers bind to the template DNA, and the extension step where a DNA polymerase extends the primers to make a new strand of DNA. This is repeated for 20-30 cycles to amplify the mutated DNA.

3.2.1 QuikChange site directed mutagenesis

Mutagenesis of Vt has been carried out using the 'Quikchange' method from Stratagene. This method allows site-specific mutations to be incorporated into double-stranded plasmids, which removes the need for subcloning into M13 based vectors and for single stranded DNA rescue (272). It does not require specialised vectors, restriction sites or multiple transformations which makes it a rapid method for introducing point mutations. It can also be used to delete or insert multiple amino acids. It is considered to be a 'non-PCR' method of SDM as the primers are designed to only amplify the parental strand and not the mutated strand, thus is not strictly true PCR

The primers used in this method are complementary to opposite strands of the vector. A high fidelity DNA polymerase replicates both strands of the plasmid without displacing the mutant primers. The primers are extended to produce a mutated plasmid. After the PCR, DpnI endonuclease is used to digest methylated and hemimethylated parent DNA, thus

selecting the unmethylated mutated plasmids. Precise details of this can be found in section 2.3.4.

There are a few variations to the method as proposed in the Quikchange manual. *Pfu* polymerase was used in the place of *pfu* turbo. This polymerase has the same fidelity as *pfu* turbo but needs a longer extension time. Also, it was found that the mutated DNA could be transformed directly into *E.coli* BL21 (DE3) cells that had been made calcium competent, meaning that a transformation step into XL1-blue cells was not required.

3.2.2 Rationale for mutations

A number of Vt mutants have been created during this study, a list can be seen in Table 3.2.1. Single point mutants have been made in an attempt to improve the solution behavior of Vt by disrupting a putative self-association interface. The double mutant was produced in the hope it would maximise this effect. A Vt deletion mutant with the final 21 residues removed has also been produced that removes some of the second putative self-association interface.

Table 3.2.1 Mutants created for use in this study

Mutant name	Mutation	Vt residues[(a)]
Vt/I997S	Ile997 to Ser	879-1066
Vt/M1022S	Met1022 to Ser	879-1066
Vt/I997S/M1022S	Ile997 to Ser Met1022 to Ser	879-1066
Vt/ΔC	C-terminus removed	879-1045

(a) all constructs have an additional four residues, GSHM, at the beginning of the sequence from the thrombin cleavage site.

The rationale for the point mutations comes from the crystal structure of Vt (172). As stated previously (Section 1.3.3.1), the second half of helix four of Vt is the most

hydrophobic region, and that residues within this region are involved in a crystallographic dimer interface. The residues that this interface is centered around are residues I997 and M1022 (Figure 3.2.1). The measured rotational correlation time for wild-type Vt suggests that wild-type Vt may exist in solution as an oligomer (Section 3.1.2.5), so by making this region less hydrophobic Vt monomer formation may be promoted. This would in turn reduce the correlation time and improve signals obtained in NMR experiments. Isoleucine 997 and methionine 1022 have been mutated to serine residues to introduce a hydrophilic charged residue that should disrupt the proposed dimer interface.

The Vt/ΔC mutation is based on a similar mutant from Bakolitsa *et al.* (172). This mutation deletes the C-terminal hairpin which could be a second self-association site for Vt, but has been implicated in binding to acidic phospholipids. It is designed to retain the 5 helix topology but lack self-association properties and acidic phospholipid binding propensity.

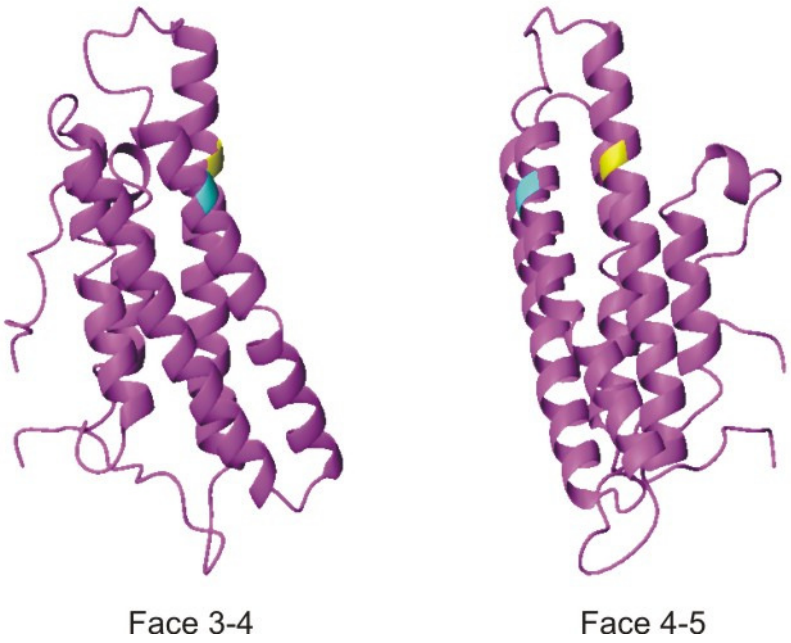


Figure 3.2.1 Wild-type Vt face 3-4 and face 4-5 showing sites of the point mutations. *The Vt crystal structure is coloured magenta, I997 is coloured cyan and M1022 is coloured yellow.*

3.2.3 Results of mutagenesis

The first step in the mutagenesis method is a 'PCR-style' step where the mutagenic primers are annealed to the template DNA to create new DNA containing the desired mutation. For single point mutations, wild-type double stranded miniprep DNA was used as the template, for double mutants the mutated Vt/I997S DNA was used as the template and the M1022S primer was used to insert the second mutation. Figure 3.2.2 shows the mutagenic primers used.

Vt/I997S For:	5'-CC ATC AGC ACG CAG CTC AAA **AGC** CTT TCC ACA GTG AAA GCT ACC-3'
Vt/I997S Rev:	5'-GGT AGC TTT CAC TGT GGA AAG **GCT** TTT GAG CTG CGT GCT GAT GG-3'
Vt/M1022S For:	5'-CA GAA CAG GCA ACT GAG **AGC** TTG GTT CAT AAC GCC CAG AAC C-3'
Vt/M1022S Rev:	5'-G GTT CTG GGC GTT ATG AAC CAA **GCT** CTC AGT TGC CTG TTC TG-3'
Vt/ΔC For:	5'-G AGA GAA GCT GAA GCA GCA TCC **TAG** AAG ATA AGA ACA GAT GCC GG-3'
Vt/ΔC Rev:	5'-CC GGC ATC TGT TCT TAT CTT **CTA** GGA TGC TGC TTC AGC TTC TCT C-3'

Figure 3.2.2 Mutagenic primers for wild-type Vt.
Forward and reverse primers are designated 'For' and 'Rev' respectively. The mutated codons are shown in bold.

After the PCR, an agarose gel was run to ensure that the DNA was amplified. The results in Figure 3.2.3 show a clear single band that represents the amplified DNA. At this point both mutant and template DNA will be present.

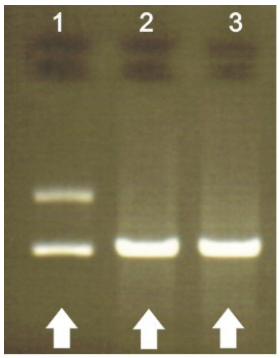
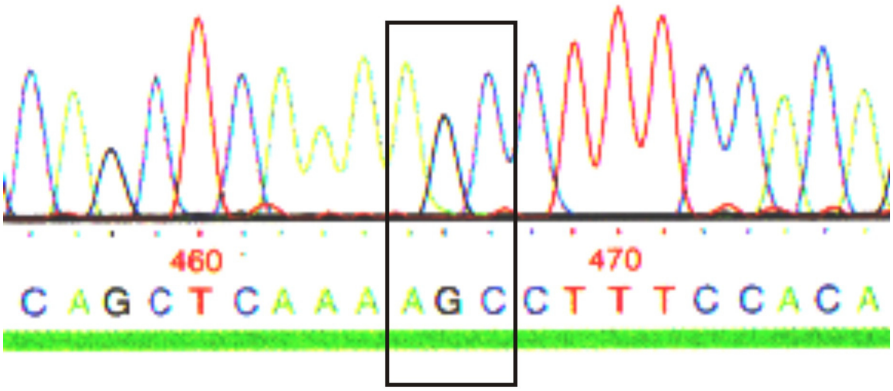


Figure 3.2.3 Ethidium bromide-stained 1% agarose gel showing the results of the PCR to make mutant Vt/I997S.

Lane 1: wild-type Vt miniprep used as a marker; lane 2 & 3: PCR products of wild-type Vt plus I997S mutagenic primers.

The template DNA was next digested by *DpnI* endonuclease and the mutated DNA is transformed into *E.coli* BL21 (DE3) cells. These are grown on LB-agar plates with ampicillin for selection and then colonies are picked and the DNA extracted and sent for sequencing. All mutants have been sequenced and are correct, and example of a sequencing reaction for Vt/I997S is shown in Figure 3.2.4.



5′-CAGCTCAAA**AGC**CTTTCCACA-3′

Figure 3.2.4 Segment of DNA sequencing profile received from MWG-biotech.

Black box confirms incorporation of the AGC codon to make the Vt/I997S mutant. The coloured peaks correspond to the base and the wide green bar represents the good quality >30, which means the probability of a wrong base is greater than 1 in 1000.

3.3 Characterisation of Vt mutants

All of the Vt mutants can be over-expressed in LB and M9 media (Figure 3.3.1). They have been purified using the same method for wild-type Vt, and all proteins behaved similarly to the wild-type throughout the purification. The exception was Vt/M1022S which behaved differently during the final concentration step where 65% of the material was lost. This suggests that the Vt/M1022S mutation alters the properties more than the I997S mutation.

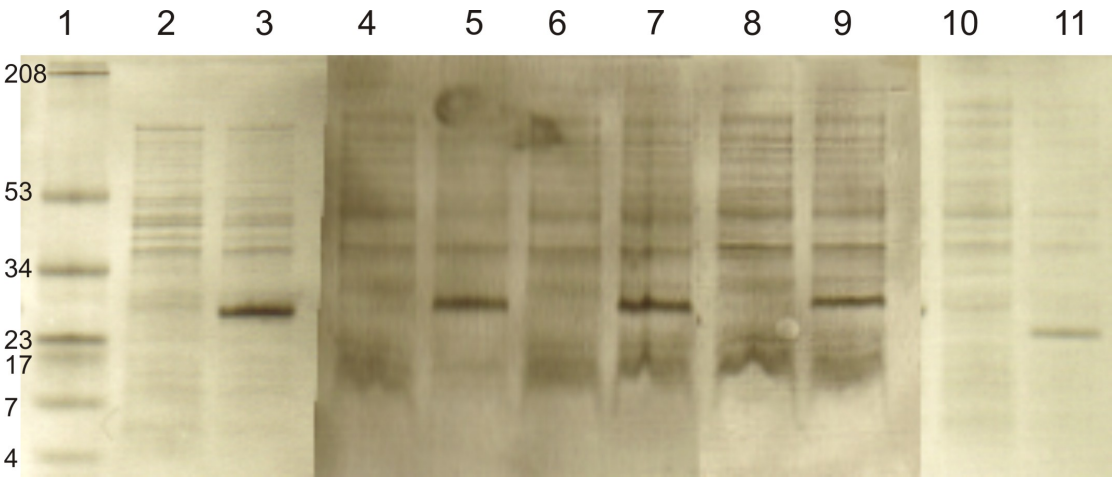


Figure 3.3.1 Composite polyacrylamide SDS-PAGE gel showing expression of Wt Vt and mutants in LB.

All cells were induced using 1mM IPTG and expression was at 30°C for 5 hours. Lane 1: molecular weight standard with molecular weights of products alongside (in kDa); lane 2: Wt Vt pre-induction; lane 3: Wt Vt post induction showing expression band; lane 4: Vt/I997S pre-induction; lane 5: Vt/I997S post induction; lane 6: Vt/M1022S pre-induction; lane 7: Vt/M1022S post induction; lane 8: Vt/I997S/M1022S pre-induction; lane 9: Vt/I997S/M1022S post induction; lane 10: Vt/ΔC pre induction; lane 11: Vt/ΔC post induction.

The calculated molecular weights for each Vt mutant are shown in Table 3.3.1. Mass spectra for the ^{15}N-labeled proteins reveal a high percentage of ^{15}N incorporation. The yields of wild-type Vt, Vt/I997S and Vt/I997S/M1022S are comparable, the Vt/M1022S mutant has a yield 50% less than wild-type Vt. By comparison, the Vt/ΔC yield is low.

Table 3.3.1 Average yields, calculated and measured molecular weights of ^{15}N labeled Vt constructs

Construct	Yield mg/L	Calculated ^{15}N Molecular weight/Da[(a)]	Measured ^{15}N Molecular weight/Da[(b)]	Difference/Da
Wt	8.0	21867	21866	1
Vt/I997S	8.2	21841	21839	2
Vt/M1022S	4.7	21823	21819	4
Vt/I997S/M1022S	7.9	21797	21792	5
Vt/ΔC	2.1	19214	19210	4

(a) Molecular weight calculated using the protparam tool (270). (b) measured by ESI-MS

3.3.1 Analytical size exclusion chromatography

Size exclusion chromatography was carried out on the mutants to judge whether the proteins were monomeric. Figure 3.3.2 shows that Vt/I997S and Vt/M1022S elute as monomers in a volume that suggests a molecular weight lower than expected but not dissimilar to the wild-type protein.

Figure 3.3.3 shows the elution profile for Vt/I997S/M1022S and Vt/ΔC. The absorbance trace at 280nm does not yield a single peak as expected. On inspection of the SDS-PAGE gel, there is no protein present in the early peak at 8.1 and 8.2mL. It is not clear what this is attributed to but is at the limit of the void volume for this column. Vt/I997S/M1022S is present in the later peak and predicts a molecular weight of 20kDa which suggests a monomeric protein. The Vt/ΔC SDS-PAGE gel shows that the peak at 22kDa corresponds to the mutant protein but there is degradation in the sample. It suggests that the Vt/ΔC is less stable at room temperature than the wild-type protein.

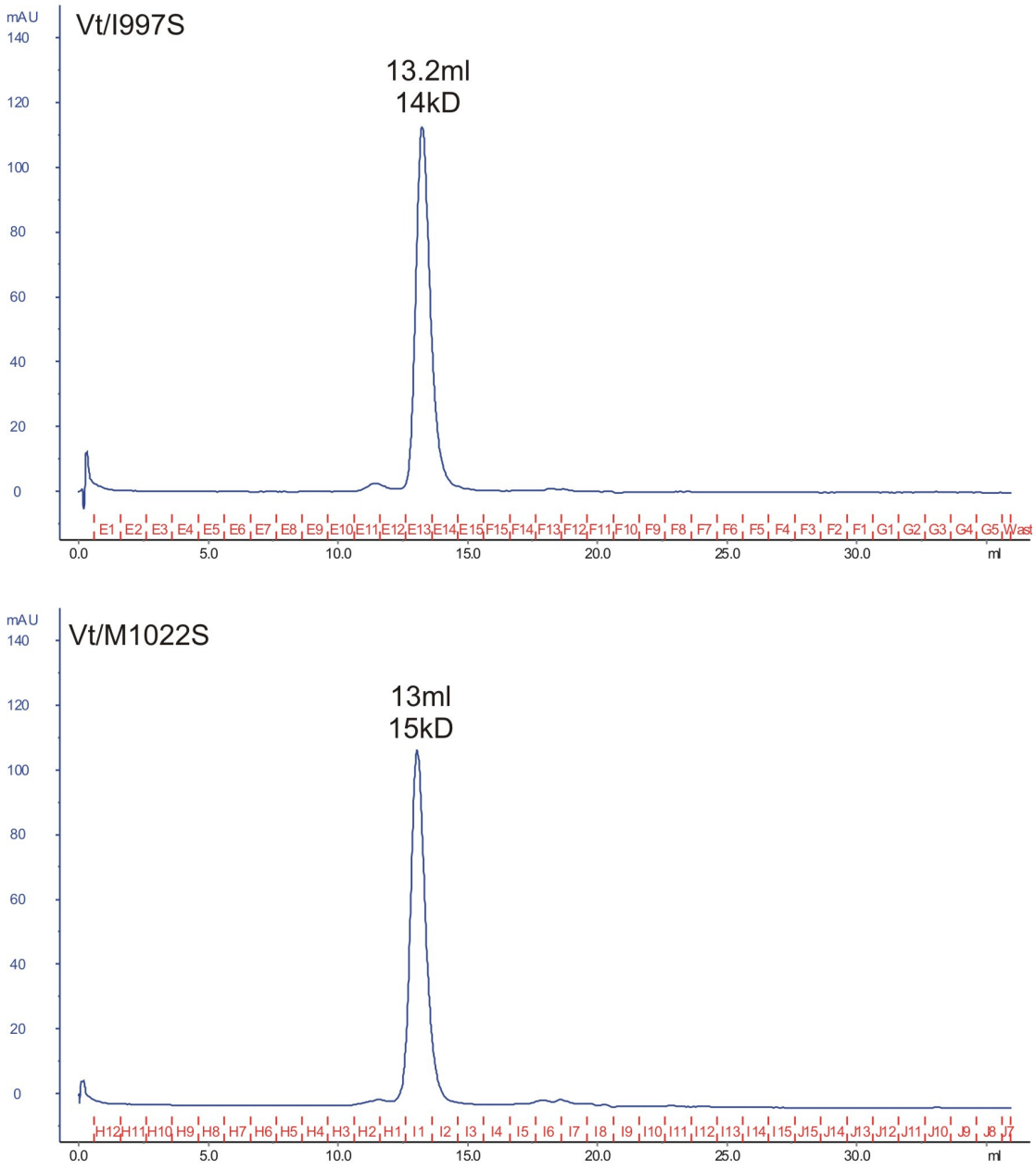


Figure 3.3.2 Analytical size exclusion chromatography (superdex-75) of Vt/I997S (top) and Vt/M1022S (bottom)

The blue trace is the protein absorbance at 280nm. Column was calibrated and run using 50mM Na_2HPO_4, 150M NaCl. 100µl of 1mg/ml Vt mutant was injected.

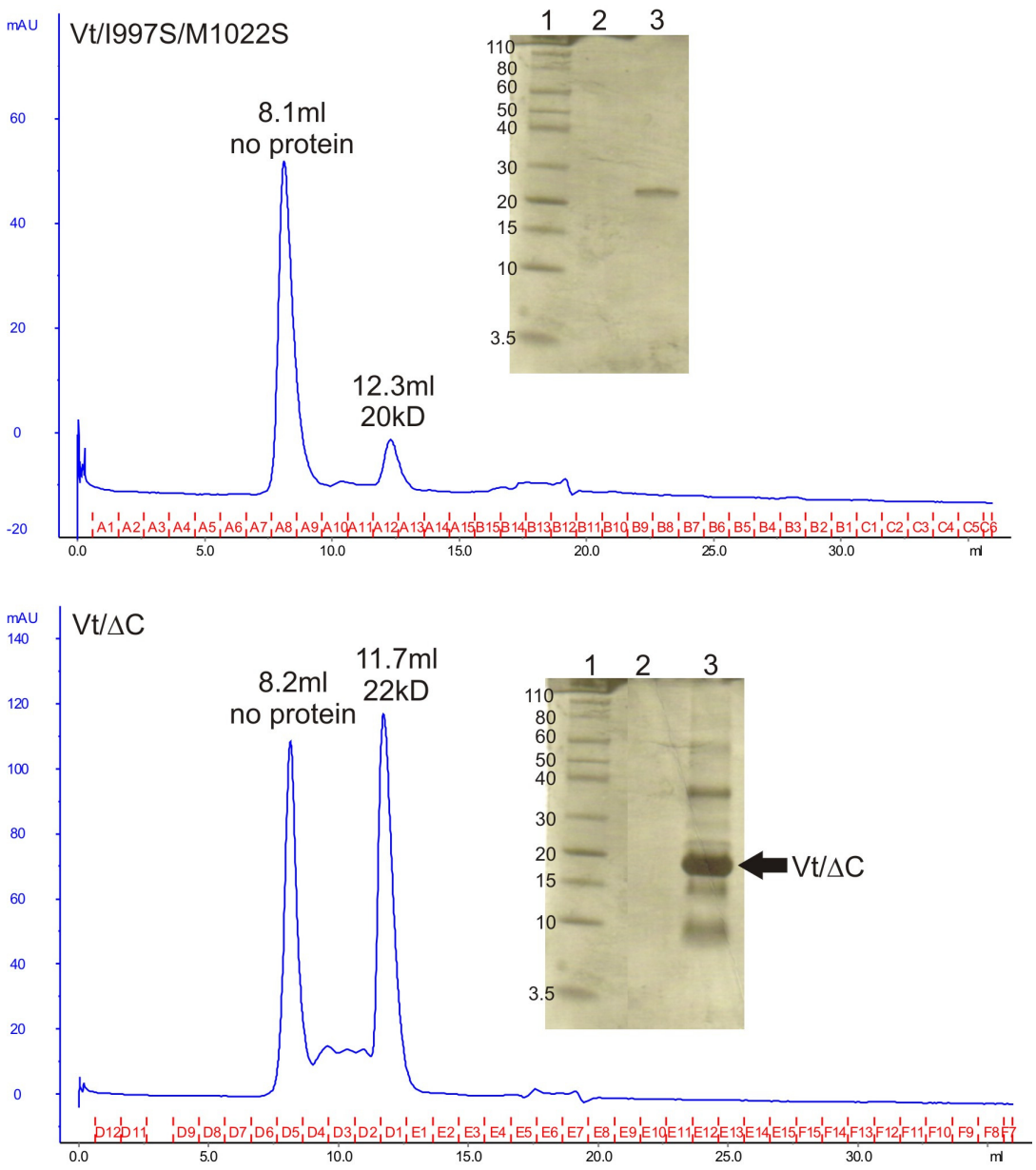


Figure 3.3.3 Analytical gel filtration (superdex-75) of Vt/I997S/M1022S and corresponding SDS-PAGE gel(top) and Vt/ΔC and corresponding SDS-PAGE gel (bottom).

The blue trace is the protein absorbance at 280nm. Polyacrylamide gel for Vt/I997S/M1022; lane1: molecular weight marker showing weights in kDa; lane 2: 8.1mL fraction; lane 3: 12.3mL fraction showing protein of interest. Polyacrylamide gel for Vt/ΔC; lane1: molecular weight marker showing weights in kDa; lane 2: 8.2mL fraction; lane 3: 11.7mL fraction showing protein of interest. Column was calibrated and run using 50mM Na_2HPO_4, 150M NaCl. 100µl of 1mg/ml Vt mutant was injected.

3.3.2 Secondary structure by circular dichroism

Far UV circular dichroism spectroscopy has been used to determine whether the mutations introduced into wild-type Vt have altered the native secondary structure.

While the CD minima for all mutants are similar suggesting a high content of α-helical structure, the magnitude of the CD signal is different between the mutants. Some explanation for this could be down to slight differences in protein concentrations or due to variations in the lengths of the α-helices (211). Visual inspection of the spectra suggest a high helical content and that this is supported by a more quantitative analysis using a CD prediction program (212). All spectra are similar to the wild-type Vt, which suggests that any structural differences are too subtle to be detected by CD measurements. Vt/M1022S is the most different from the wild-type, this may be a reflection of the observed protein losses during concentration. It may be that the helices are affected by this mutation more than the I997S mutation. The Vt/ΔC spectrum shows that the C-terminal 21 residues are not required for the formation of Vt helices.

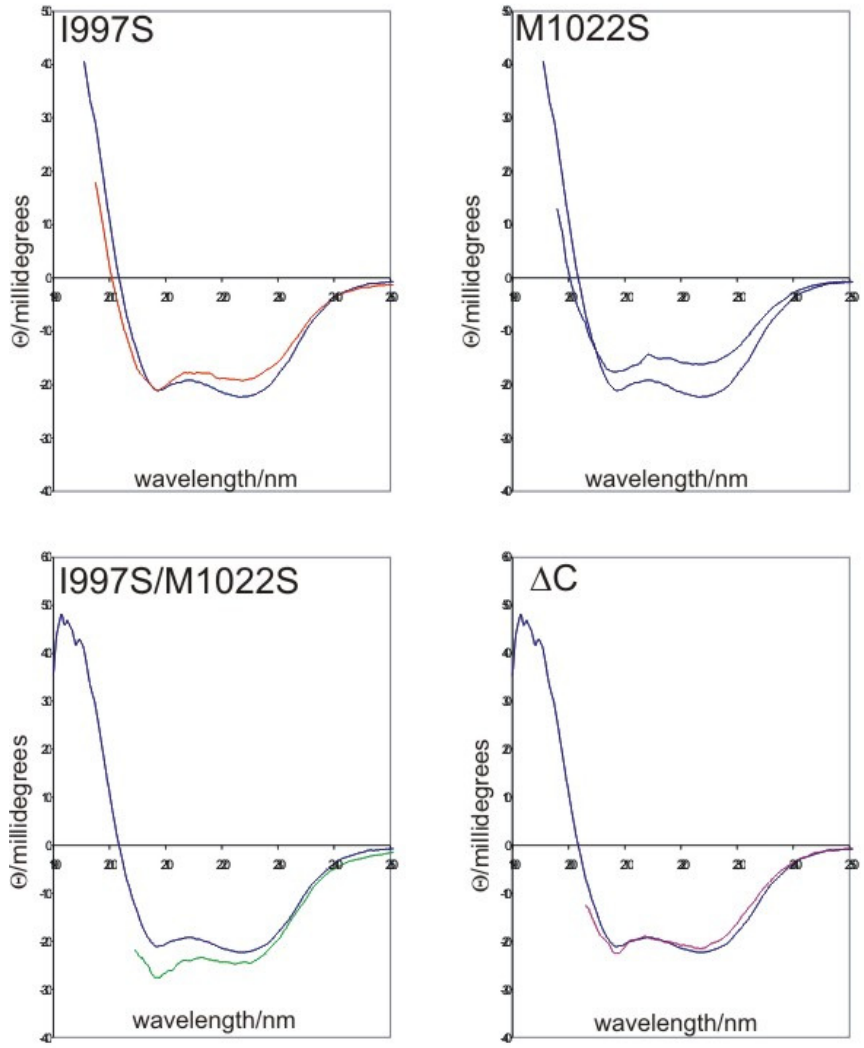


Figure 3.3.4 Far UV circular dichroism spectrum of all mutant Vt proteins compared to wild-type Vt

All proteins at 0.5μM in water. The dark blue trace in all spectra is Wt-Vt, red is Vt/I997S, blue is Vt/M1022S, green is Vt/I997S/M1022S, magenta is Vt/ΔC.

3.3.3 ^{1}H-^{15}N-HSQC spectrum of mutant Vt

^{1}H-^{15}N-HSQC spectra have been recorded for all mutants and can be seen in Figure 3.3.5-8. Globally, the spectra for the point mutants show good dispersion in the proton and nitrogen dimension which suggests that they are folded. The Vt/ΔC mutant shows evidence of

degradation and the dispersion is poorer. All spectra were recorded using identical acquisition parameters and processed identically on a sample of 200μM protein concentration. Overall the spectra for the point mutants are similar but there are subtle differences in many of the chemical shifts.

The intensity of the Vt/I997S spectrum (Figure 3.3.5) is increased compared to the wild-type Vt spectrum. The Vt/I997S peaks appear more rounded and better dispersed in the central region. An initial peak count produces 160 peaks suggesting that 20 are missing compared to the wild-type (the wild-type initial peak count produced 180 peaks). This does not take into account peak degeneracy, there may be multiple peaks in the same place in the spectrum.

The Vt/M1022S spectrum (Figure 3.3.6) also has well rounded peaks with a well dispersed central region, however it gives a markedly less intense spectrum than wild-type Vt. There are 176 peaks in this spectrum. As stated previously, concentrating this mutant protein resulted in a loss of 65% of the material, the HSQC provides further evidence that the Vt/M1022S mutation is adversely affecting the solution properties of Vt.

The double mutant, Vt/I997S/M1022S (Figure 3.3.7) spectrum has 175 well rounded peaks, but similarly to the Vt/M1022S single mutant, the spectrum is much weaker in intensity compared to wild-type Vt.

The ^{1}H-^{15}N-HSQC spectrum for the Vt/ΔC mutant in Figure 3.3.8 has been acquired for 2 hours and shows a marked difference to wild-type Vt. The resonances in the proton dimension are less well dispersed and the quality of the spectrum looks lower than wild-type Vt, as judged by the small sharp peaks in the region of 8.3-7.7ppm. The quality of the 12 hour data was poor-there were many more sharp peaks in the 8.3-7.7ppm region and the peaks overlapped greatly. It seems that the Vt/ΔC mutant is unstable and degrades rapidly at 25ºC. This would agree with observations in the size exclusion chromatography where the SDS-PAGE gel showed other bands apart from the desired protein (Figure 3.3.3). Comparing this data to the CD spectroscopy shows that while the C-terminal 21 residues

are not required for formation of Vt helices, they do play a role in stability of the Vt structure over time.

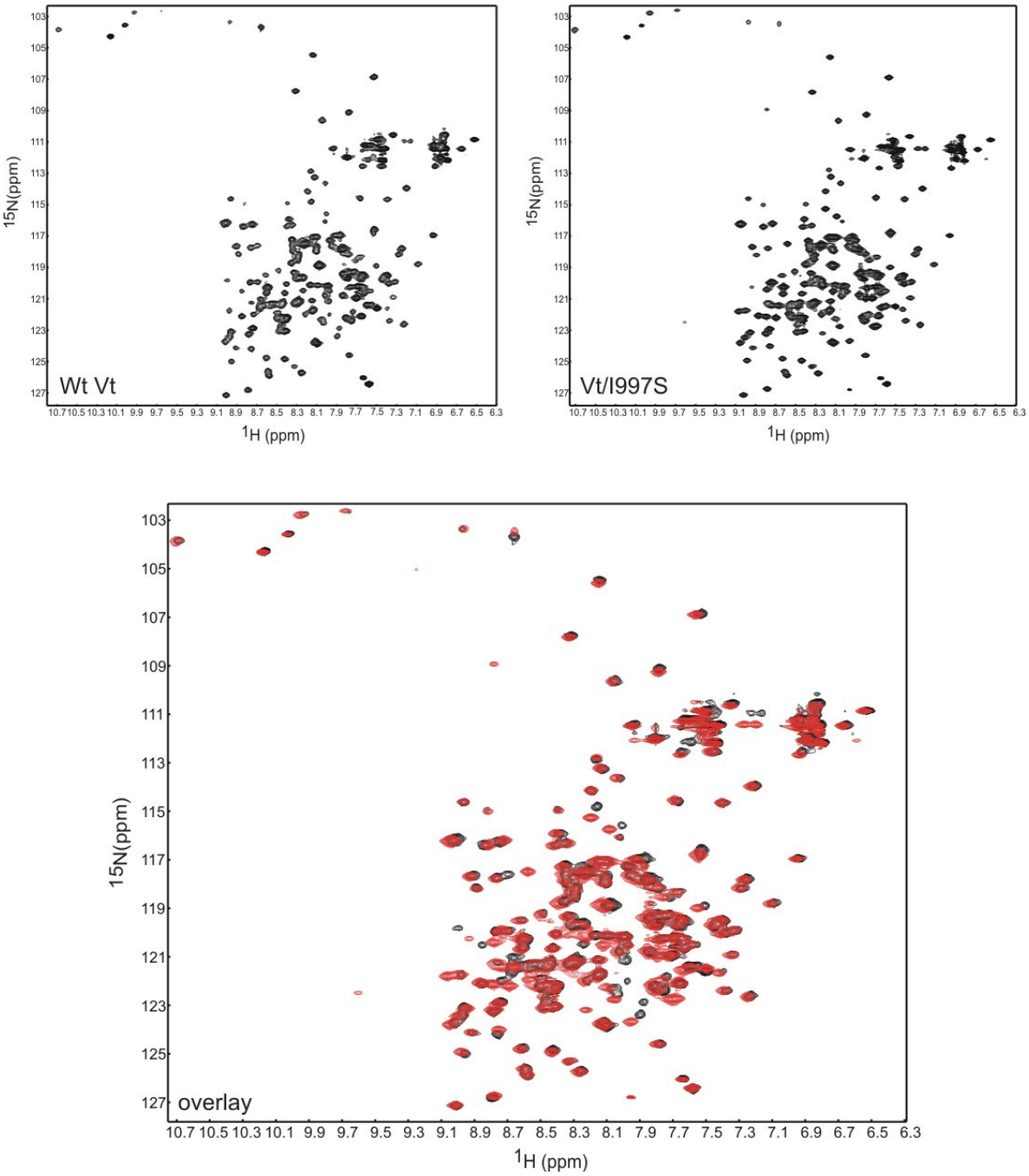


Figure 3.3.5 ^{1}H-^{15}N-HSQC spectra of Vt/I997S

Top: ^{1}H-^{15}N HSQC spectrum for Wt Vt (left) and Vt/I997S (right). Bottom: Overlay of the two spectra, Wt Vt in black and V/I997S in red. All spectra acquired for 12 hours in 20mM Tris pH7, 400mM NaCl.

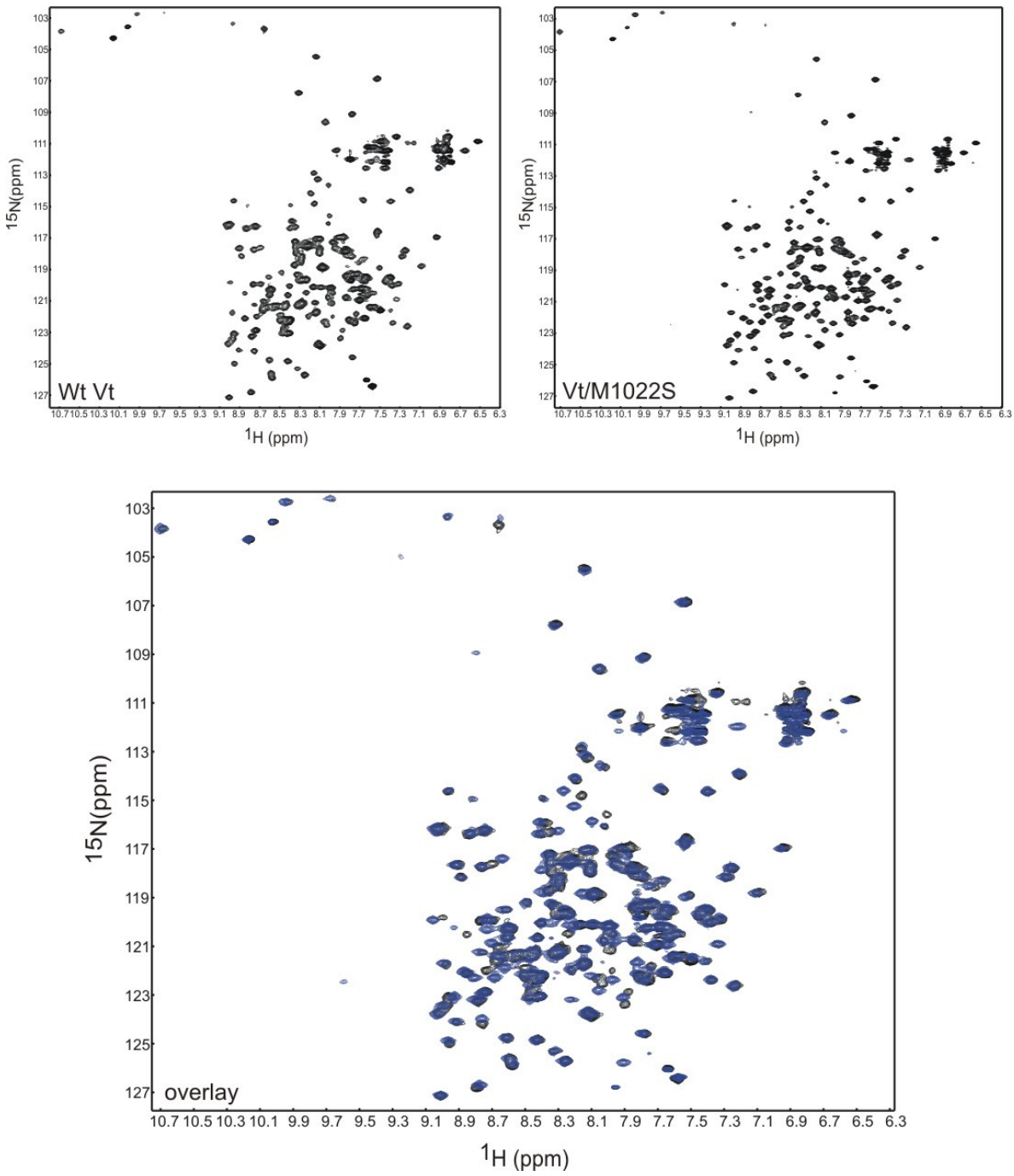


Figure 3.3.6 ^{1}H-^{15}N-HSQC spectra of Vt/M1022S

Top: ^{1}H-^{15}N HSQC spectrum for Wt Vt (left) and Vt/M1022S (right). Bottom: Overlay of the two spectra, Wt Vt in black and Vt/M1022S in blue. All spectra acquired for 12 hours in 20mM Tris pH7, 400mM NaCl.

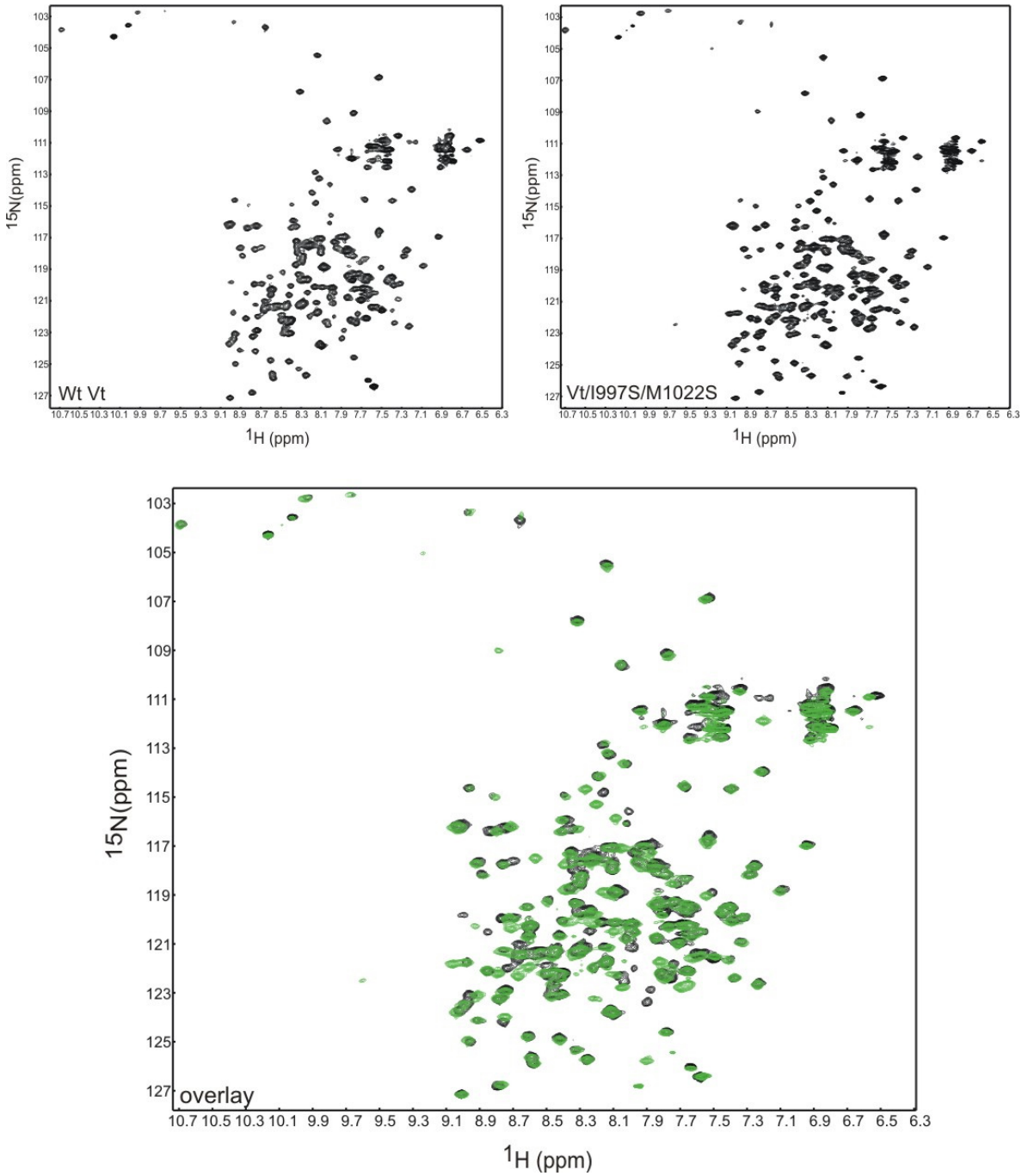


Figure 3.3.7 ^{1}H-^{15}N-HSQC spectra of Vt/I997S/M1022S

Top: ^{1}H-^{15}N HSQC spectrum for Wt Vt (left) and Vt/I997S/M1022S (right). Bottom: Overlay of the two spectra, Wt Vt in black and Vt/I997S/M1022S in green. All spectra acquired for 12 hours in 20mM Tris pH7, 400mM NaCl.

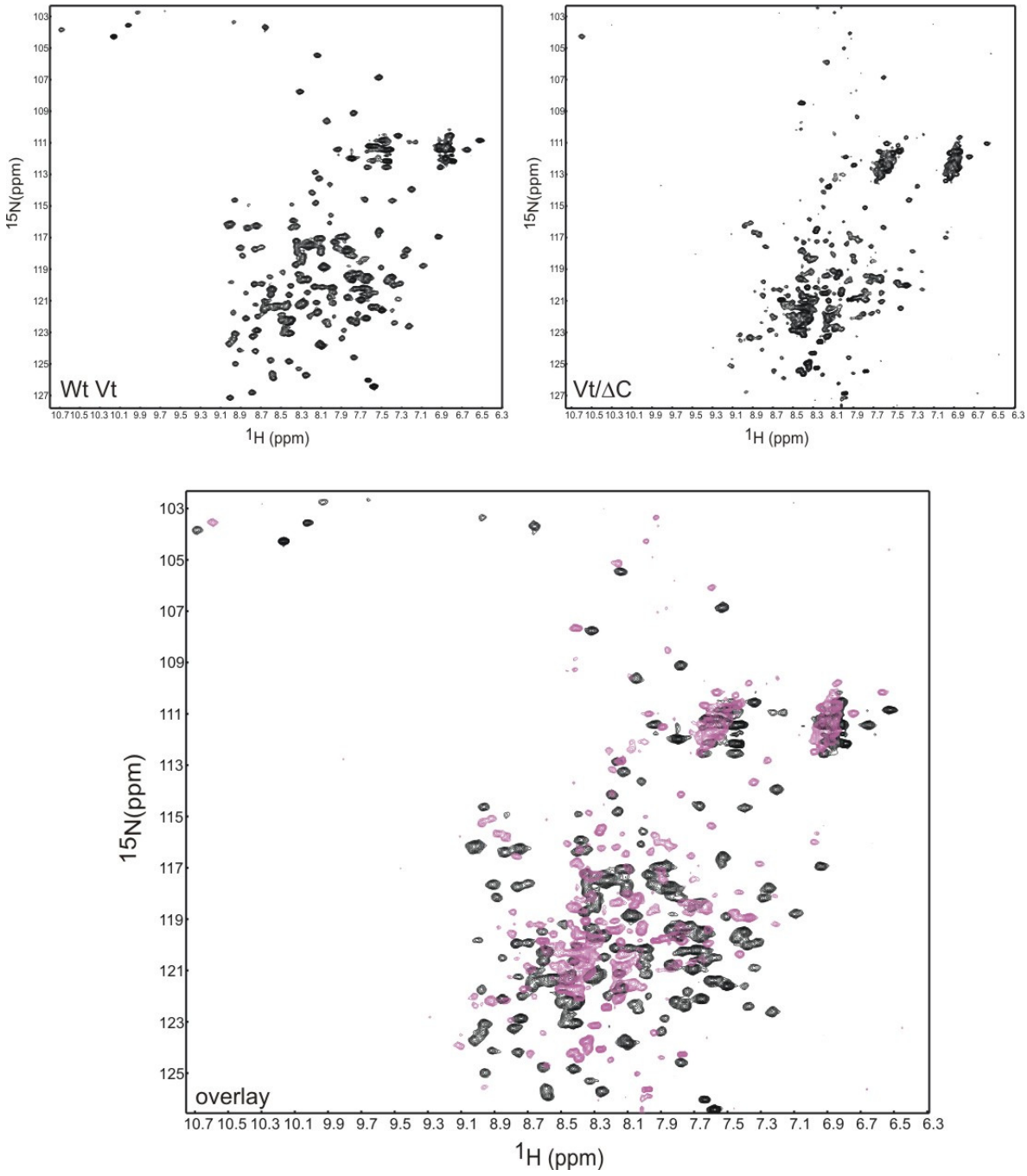


Figure 3.3.8 ^{1}H-^{15}N-HSQC spectra of Vt/ΔC

Top: ^{1}H-^{15}N HSQC spectrum for Wt Vt (left) and Vt/ΔC (right). Bottom: Overlay of the two spectra, Wt Vt in black and Vt/ΔC in magenta.

3.3.4 Rotational correlation times for Vt mutants

Rotational correlation times have been measured for Vt/I997S and Vt/M1022S to determine whether the mutations have promoted monomer formation. Wild-type Vt had a τ_m of 13.0ns and the hope was that the mutants would have a lower τ_m. Figure 3.3.9 shows an example of the amide regions that were integrated and the data plots that were used to estimate the τ_m. The fitted data gave T_1 and T_2 values of 1.465 and 0.090 respectively for Vt/I997S and 1.36 and 0.104 for Vt/M1022S. The calculated correlation times were 12.6ns for Vt/I997S and 11.3ns for Vt/M1022S. The proportion of dimer in solution was estimated as described in Section 2.5.9. The comparison with the wild-type protein can be seen in Table 3.3.2. Both mutants have reduced the correlation time and the apparent percentage of dimer as judged by hydrodynamic modelling. The reduction is not as marked as was expected, it may be that the rate at which this association is occurring has been affected which would account for the improved lineshapes in the HSQC spectra, but the association is still present. It is also possible that both of the self association interfaces as shown in Figure 1.3.8 and Figure 1.3.9 are both occurring, and these point mutations are only reducing the self association of one of these interfaces.

Table 3.3.2 Measured rotational correlation times for mutant proteins and proportion of dimer predicted in solution

Vt mutant	Correlation time/ns	% dimer estimated
Wt	13.0	63
Vt/I997S	12.6	59
Vt/M1022S	11.3	47
Vt/I997S/M1022S	Not determined	Not determined
Vt/ΔC	Not determined	Not determined

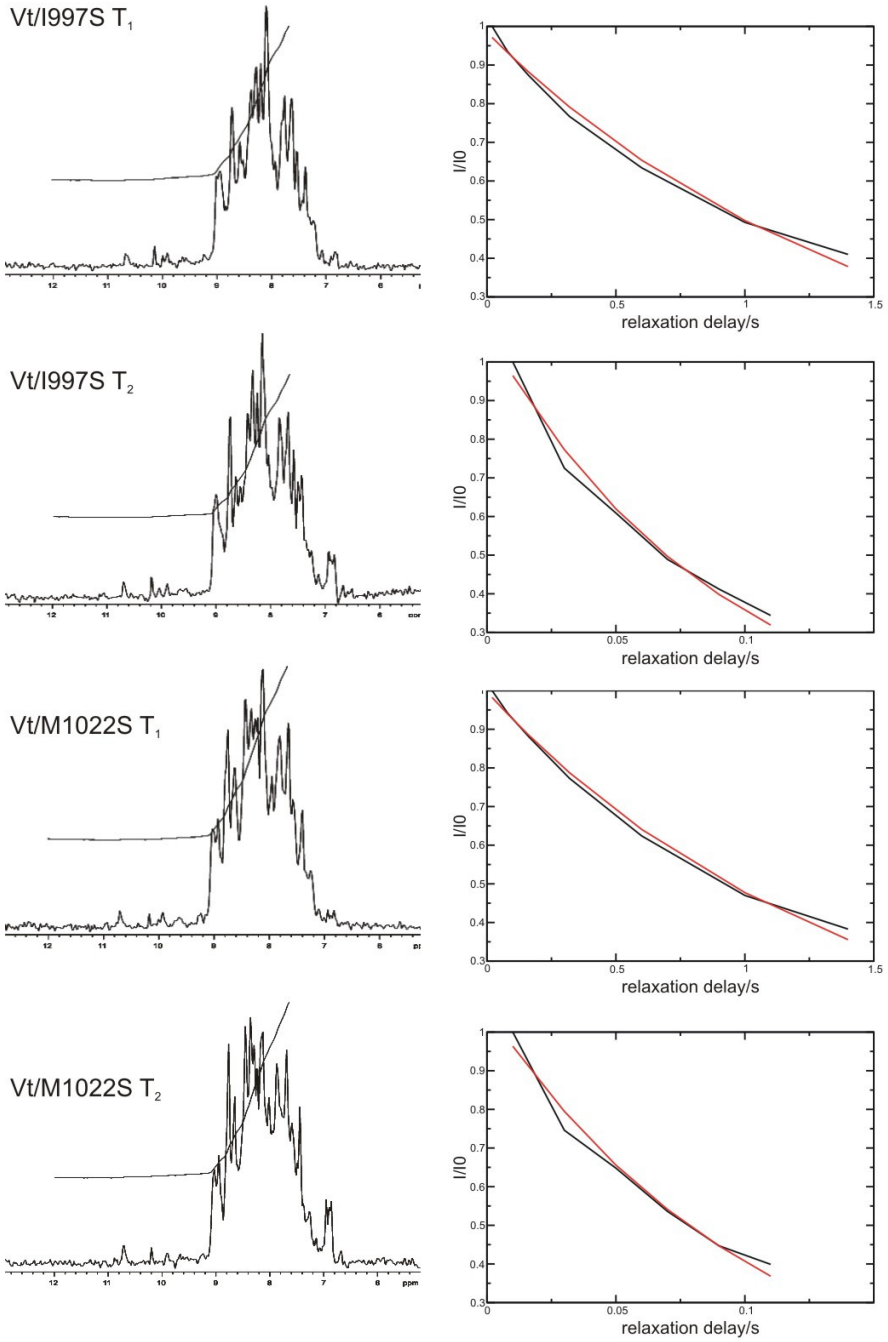


Figure 3.3.9 Results of ^{15}N relaxation experiments on Vt/I997S and Vt/M1022S

Left: amide region of the first relaxation delay, for T1=0.020ms, for T2=0.010ms. The integration curve that was used is shown as a black line. Right: graph used to determine T1 and T2 values, black line represents actual data and red line represents the curve fit.

3.3.5 Wild-type Vt and Vt/I997S buffer choice

The inherent low sensitivity of NMR experiments generally requires high protein concentrations typically in the order of hundreds of μM. The signal to noise of NMR experiments is also degraded by high sample conductivity, for example brought about by the addition of salt to the sample, an issue that is particularly significant when using cryogenic probes.

Attempts to concentrate both wild-type Vt and Vt/I997S to above 500μM resulted in precipitation of the protein in buffers containing 20mM Tris pH7 with 400mM NaCl. If the salt concentration is lowered to 150mM NaCl this precipitation seems to occur at lower concentrations, although this has not been measured quantitatively. An additional observation is that Vt/I997S at 500μM suffers from degradation over 2-3 days which causes problems in data collection for assignments as the experiments typically last for several days.

It was therefore necessary to optimise the buffer conditions to balance the salt concentration, protein concentration and stability. The solubility and stability of Vt in 400mM NaCl is better than in 150mM NaCl, yet 150mM NaCl is better for NMR studies and closer to physiological conditions. 50mM NaCl ultimately produced the best chemical shift perturbation data (see Section 3.6.3).

In order to achieve this reduction in salt without compromising on stability and solubility, 50mM arginine and 50mM glutamate was added to buffer solutions. Adding these amino acids to a 50mM NaCl buffer maintains the ionic strength to that of a 150mM NaCl solution, yet the mobility of these ions are reduced compared to NaCl. Since signal-to-noise (S/N) in an NMR experiment is determined in part by ionic strength and ion mobility, the reduction in mobility produces a better S/N for a given experiment without compromising on ionic strength. Therefore Vt is still as stable as in a buffer containing 150mM NaCl, but produces a better NMR spectrum. Adding 50mM arginine and 50mM glutamate to buffers has also been shown to protect proteins against degradation thus aid in

long term stability, both beneficial for NMR. This is why they were also added to 150mM NaCl containing buffers. Fortunately, the amino acids do not adversely affect any protein-protein or protein-nucleic acid interactions (273).

When Vt is dialysed into 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate it can be concentrated to a higher concentration than without (800µM has been achieved without significant aggregation). Qualitatively, the 50mM arginine/50mM glutamate buffer is preserving the protein for longer at high Vt concentrations, as judged by ^{1}H-^{15}N-HSQC spectra, but ultimately does not prevent degradation of the spectral quality. At Vt concentrations of 200µM in 150mM NaCl this buffer is not always necessary as the spectral quality is consistent over the time course of the experiments. Experiments carried out at 50mM NaCl also contained 50mM arginine and 50mM glutamate to ensure that Vt is stable.

3.3.6 Choice of mutant and isotopic labelling of Vt/I997S

To enable further studies on Vt it was necessary to assign the ^{1}H-^{15}N-HSQC spectrum. There were concerns that the wild-type Vt spectrum was not of sufficient quality to produce useful data to achieve full assignment. Consequently, Vt/I997S was chosen to use for assignment purposes on the basis that the protein yield is the same as for the wild-type protein yet the resonances in the ^{1}H-^{15}N-HSQC spectrum are sharper and more resolved.

Assignment of the Vt/I997S ^{1}H-^{15}N-HSQC spectrum is described in Chapter 3.4. Vt/I997S has been isotopically labelled with ^{15}N, ^{13}C and deuterium. Initially, deuterium labelling was carried out with the aim of 60-70% incorporation; however, only perdeuteration improved the acquired NMR data sufficiently to obtain good quality spectra for backbone assignment. Mass spectrometry of the perdeuterated protein shows that incorporation of deuterium into Vt/I997S was efficient, with greater than 99% incorporation (Figure 3.3.10).

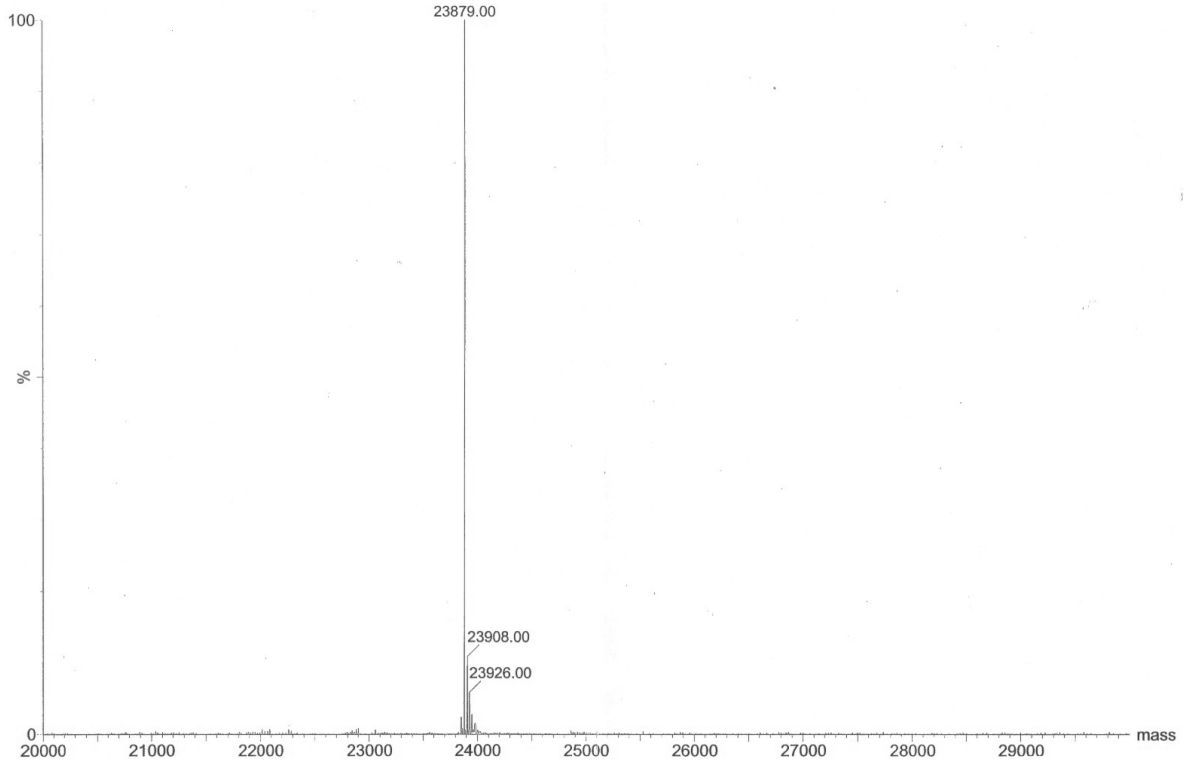


Figure 3.3.10 ESI-Mass spectrometry results for perdeuterated Vt/I997S

Measured mass is 23879, expected results 23911.7. This suggests deuterium incorporation is greater than 99%.

3.4 Sequential Resonance Assignment of the Vt/I997S backbone

For any detailed studies on proteins using NMR, it is necessary to identify the chemical shift of nuclei along the backbone which will then act as an 'atomic signature' that identifies each site in the protein (274). Assignment of the Vt/I997S backbone has been carried out so that chemical shift perturbation experiments can be interpreted. To date there is no published backbone resonance assignments for wild-type Vt or the mutant Vt/I997S.

The Vt/I997S protein used in this study has a molecular weight of 21587 Da and contains 192 amino acids. Backbone assignment using state-of-the-art NMR technology should be feasible for a protein of this size. However, for several reasons Vt assignment has been a challenge and has required the use of ^{15}N, ^{13}C and ^{2}H labelling together with TROSY spectroscopy at high fields. Lack of solubility, the need for high ionic strength buffers and limited chemical shift dispersion due to the α-helical nature of Vt have caused problems with quality of data collected, and the degenerate amino acid sequence has made it difficult to identify where in the sequence a series of connected resonances should lie.

As previously stated in Section 3.3.6, the Vt/I997S protein has been used for assignment purposes as the quality of the ^{1}H-^{15}N-HSQC spectrum shows improved dispersion and lineshapes compared to the wild-type protein. Several approaches have been taken to enable assignment. Heteronuclear triple resonance experiments on partially and fully deuterated protein, as well as ^{15}N-edited NOESY experiments were used with optional relaxation optimisation using TROSY spectroscopy. Additionally, amino acid selective labeling and unlabelling have all been employed. A list of experiments can be found in Chapter 2.6.4.1.

3.4.1 Assigned Vt/I997S ^{1}H-^{15}N-TROSY spectrum

The NH resonance assignments determined to date are shown in Figure 3.4.1 and Figure 3.4.2. Of the 192 residues in Vt/I997S, four are proline residues which will not be present in a ^{1}H-^{15}N-HSQC spectrum so cannot be assigned. This leaves 188 potential backbone NH resonances that are available for assignment. Of these, 147 backbone resonances have been assigned (78%). Within helices the proportion of assigned residues is higher (Table 3.4.1) and overall 88% of the α-helical regions are assigned. The crystal structure of wild-type Vt is shown in Figure 3.4.3 with the Vt/I997S assignments mapped onto it to show the distribution of assignments. Of the 147 backbone NH resonances assigned, all of the corresponding Cα shifts have been assigned, and for residues other than glycine, there are 9 residues that have no Cβ shift assigned to them, many of which are serine and threonine residues. There are also 109 initial CO assignments. An assignment table can be seen in Appendix 5.5.

Table 3.4.1 Proportion of assigned residues in the helical regions of Vt/I997S

Helix number	Number of assigned residues	Number of residues in helix	Percentage assigned
1	12	15	80
2	18	18	100
3	28	28	100
4[(a)]	21	27	78
5	27	32	84

(a) In helix four there is one proline residue that cannot be assigned

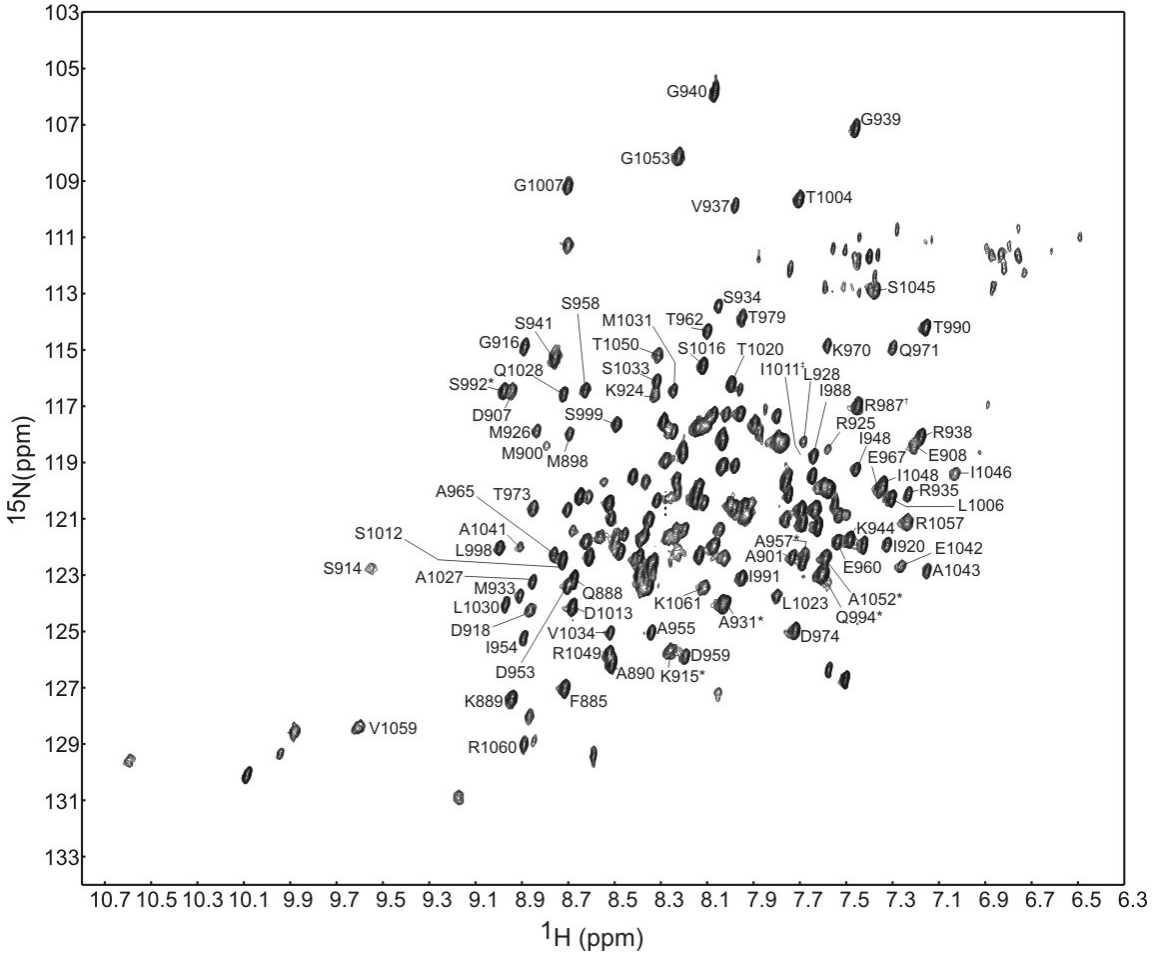


Figure 3.4.1 ^{1}H-^{15}N-TROSY spectrum of ^{15}N, ^{13}C, ^{2}H uniformly labelled Vt/I997S showing assigned NH resonances.

Spectrum acquired at 800MHz. Vt/I997S was at 500µM, perdeuterated, ^{15}N and ^{13}C labelled in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. Full spectrum showing the assignments. Resonances marked with '' are degenerate. The resonance labelled R987 (marked with †) corresponds to one residue in Vt/I997S but may correspond to two residues in wild type protein, and ‡ is a resonance that is only seen at low contour levels. . The assignment of the central region between 8.8-7.4 ppm and 116-124ppm in the proton and nitrogen dimension respectively are shown in Figure 3.4.2.*

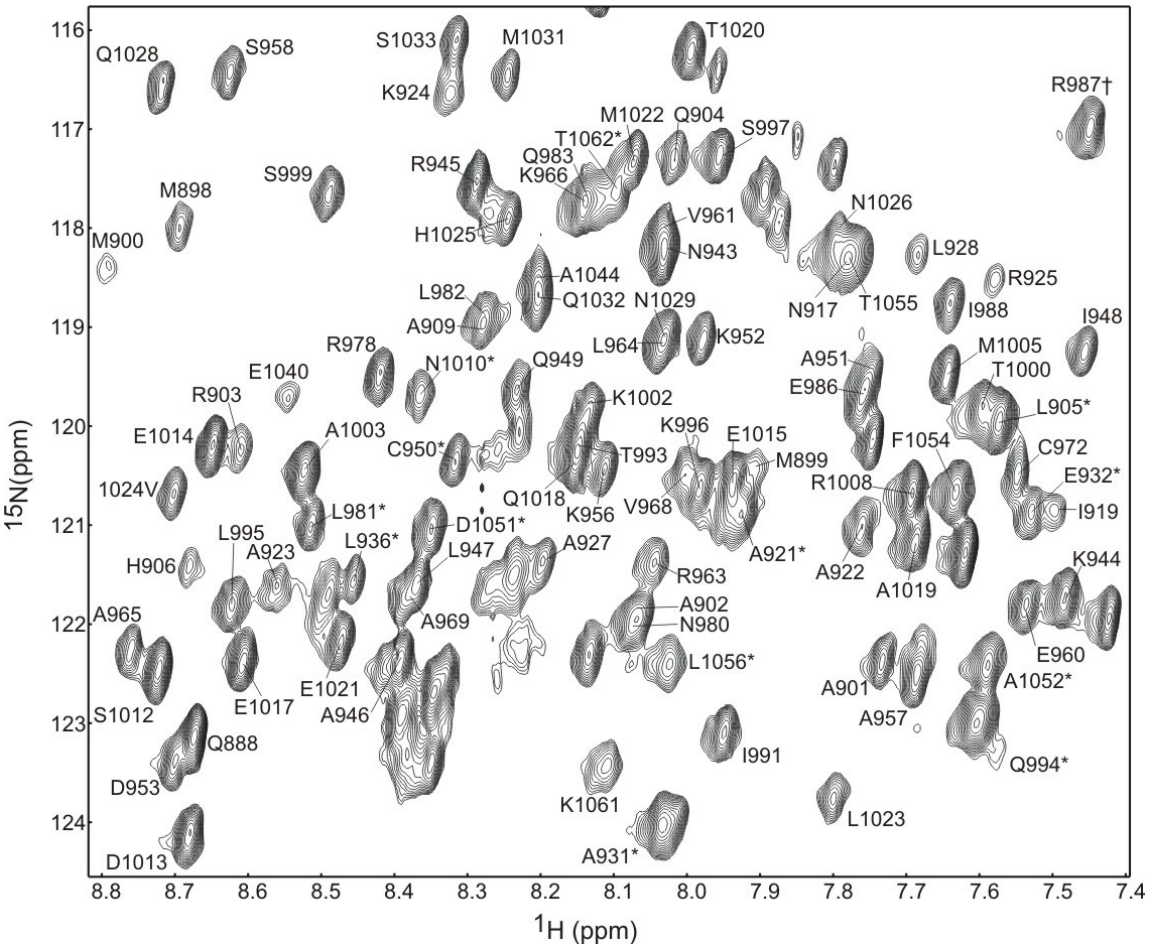


Figure 3.4.2 Expansion of central region of the ^{1}H-^{15}N-TROSY spectrum of ^{15}N, ^{13}C, ^{2}H uniformly labelled Vt/I997S showing assigned NH resonances.

Spectrum acquired at 800MHz. Vt/I997S was at 500µM, perdeuterated, ^{15}N and ^{13}C labelled in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. Close up of the central region showing assignments. Residues marked with '' are degenerate. R987 (marked with †) is one residue in Vt/I997S but may be two residues in wild type protein.*

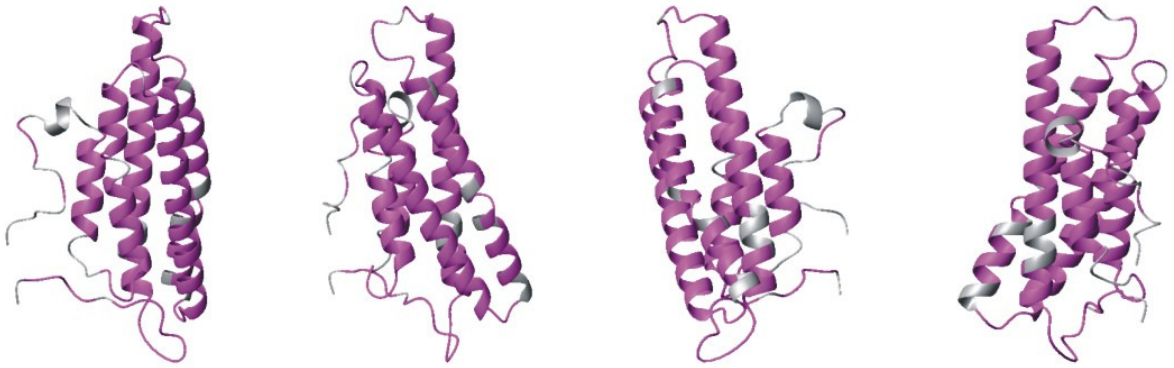


Figure 3.4.3 Assigned residues from Vt/I997S mapped onto the crystal structure of Wild-type Vt.

Grey residues are unassigned, magenta residues are assigned in the Vt/I997S spectrum.

3.4.2 Assignment strategy

The assignment of Vt/I997S has been approached from a variety of angles. Triple resonance experiments that correlate $^{1}H^{N}$, ^{15}N, $^{13}C\alpha$ and $^{13}C\beta$ have been used to provide sequential connectivities between amino acids. The HNCA, HNCACB, HN(CO)CA and HN(CO)CACB experiments have been conducted on partial and fully deuterated Vt/I997S. The data gathered from triple resonance experiments alone was not sufficient to complete the assignment. Therefore, in addition to the triple resonance experiments, ^{15}N-edited NOESY-HSQC and the equivalent TROSY experiment have been carried out to exploit the NOEs that arise between the amide groups of amino acids in an α-helix. Finally, in order to enable the identification of stretches of connected spin systems within the Vt/I997S sequence, amino acid selective labelling and unlabelling was conducted.

3.4.3 Triple resonance experiments

Several sets of triple resonance experiments have been recorded on ^{15}N, ^{13}C , 60-70% ^{2}H Vt/I997S as well as on ^{15}N, ^{13}C, 99% ^{2}H Vt/I997S. Using ^{1}H-^{15}N-TROSY versions of these experiments has improved the signal-to-noise of the acquired data.

The primary triple resonance experiments that have been used to assign the backbone are TROSY-HNCACB and TROSY-HN(CO)CACB on ^{15}N, ^{13}C, 99% ^{2}H Vt/I997S, measured at 800MHz with a cryoprobe. Figure 3.4.4 shows sequentially assigned ^{1}H-^{13}C sections at defined ^{15}N resonances from the TROSY-HNCACB spectrum. These represent the best quality data recorded of Vt/I997S. There are clear carbon resonances from both the α and β carbons for the amino acid at position i and for the preceding amino acid at position i-1.

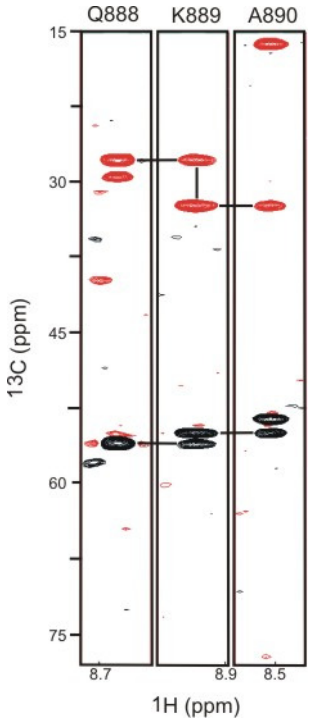


Figure 3.4.4 ^{1}H-^{13}C sections at defined ^{15}N resonances from TROSY-HNCACB spectrum recorded at 800MHz with a cryoprobe.

Vt/I997S was at 500μM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. In this spectrum Cα resonances are shown in black and Cβ resonances in red because their signals are 180º out of phase leading to positive and negative intensities. Sequential connectivities are shown as black lines and corresponding assignments shown across the top.

However, for the majority of carbon resonances, both the HNCACB and the HN(CO)CACB were required to confirm the presence of the resonances to the preceding amino acid. Figure 3.4.5 shows ^{1}H-^{13}C sections taken from the TROSY- HNCACB and the

TROSY- HN(CO)CACB and illustrates that for many sections, the i-1 peak was not visible in the HNCACB, and was connected using the HN(CO)CACB, for example the sections containing A1041, E1042 and I1046 are missing the i-1 information in the HNCACB.

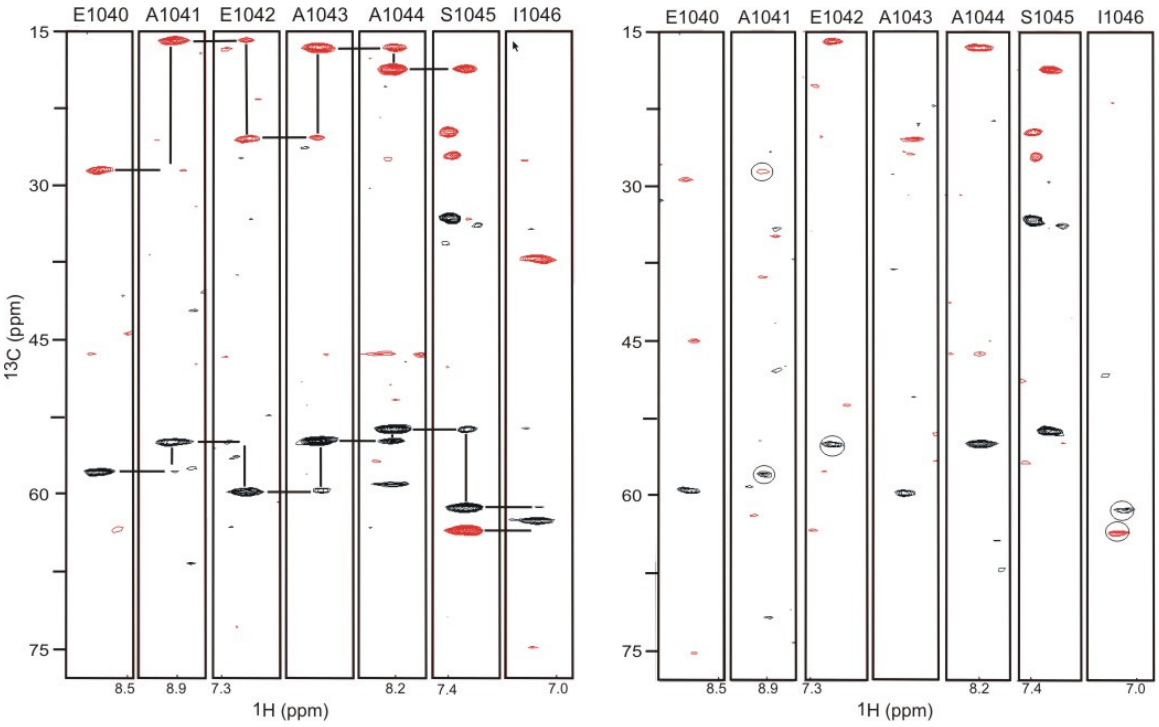


Figure 3.4.5 ^{1}H-^{13}C sections at defined ^{15}N resonances from a TROSY-HNCACB spectrum (left) and TROSY-HN(CO)CACB spectrum (right) recorded at 800MHz with a cryoprobe.

Vt/I997S was at 500µM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. In this spectrum Cα resonances are shown in black and Cβ resonances in red because their signals are 180º out of phase leading to positive and negative intensities. Sequential connectivities are shown as black lines and corresponding assignments shown across the top. The strips on the right were required to determine the connectivities, the circles represent resonances missing in the TROSY-HNCACB spectrum.

Whilst these two triple resonance experiments were successfully used in most cases, sometimes it was necessary to refer to an HNCA to verify either the presence of a Cα resonance of the preceding amino acid (i-1) or to decide which of the two Cα resonances had the strongest intensity (to decide which resonance corresponded to the Cα residue itself

(i)). The HNCA experiment used has a TROSY sequence element applied and was carried out on ^{15}N, ^{13}C , 60-70% ^{2}H Vt/I997S at 800MHz without a cryoprobe. The HNCA has improved sensitivity compared to the HNCACB due to the magnetisation transfer only having to go as far as the Cα atoms. Figure 3.4.6 is a good example of how it has been used to complement the other experiments. The ^{1}H-^{13}C section for M933 from the TROSY-HNCACB spectrum shows a very weak peak at 58.6ppm that could be a Cα i-1 peak, and this can be verified by inspecting the same section at 58.6ppm in the TROSY-HNCA. S934 had two Cα resonances at 60.3ppm and 61.7ppm but it was unclear which was the Cα i and which was the Cα i-1. The HNCA confirms that the resonance at 61.7ppm which is of lesser intensity in the HNCACB corresponds to the Cα i.

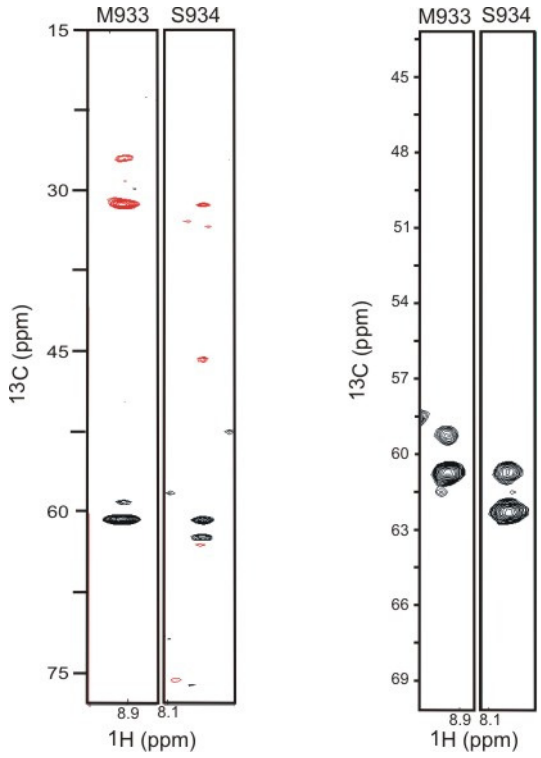


Figure 3.4.6 ^{1}H-^{13}C sections at defined ^{15}N resonances from a TROSY-HNCACB spectrum (left) and a TROSY-HNCA spectrum (right).

Vt/I997S was at 500μM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. In this spectrum Cα resonances are shown in black and Cβ resonances in red because their signals are 180º out of phase leading to positive and negative intensities.

3.4.4 ^{1}H-^{15}N NOESY-HSQC experiments

The triple resonance experiments alone were not sufficient to assign the backbone. The most significant problem was that the number of spin systems that were linked together representing a stretch of amino acids was insufficient and the chemical shifts not diverse enough to unambiguously place them in the sequence. In order to overcome this, the α-helical nature of Vt was exploited. In an α-helix the $^{1}H^{N}$-$^{1}H^{N}$ distance (dNN) is such that sequential NOEs can be detected between the amides of neighboring amino acids (215). The data collected on Vt/I997S only enables NOEs between residues i, i-1 and i+1 to be detected with certainty, NOEs to i+2 have occasionally been detected and are probably due to spin diffusion effects.

Initially a ^{1}H-^{15}N-NOESY-HSQC spectrum was acquired at 600MHz on ^{15}N, ^{13}C , ^{1}H Vt/I997S, connectivities were improved by acquiring at 800MHz. A ^{1}H-^{15}N-TROSY-NOESY was run ^{15}N, ^{13}C , 99% ^{2}H Vt/I997S at 900MHz to visualise only the NH and to see transfer to more distant amino acids in the helix, only in some cases was an NOE to an amino acid 2 residues away seen. A ^{1}H-^{15}N-TOCSY-HSQC to complement the above did not produce sufficient data to use in the assignment.

For the majority of the assignments, the connections were first formed using the ^{15}N-edited NOESY experiments and were then confirmed by inspecting the connections in the HNCACB and HN(CO)CACB. Figure 3.4.7 shows an example of this. The NOESY is shown at the top where the connections through the helices have been made, and the corresponding strips from the TROSY-HNCACB and TROSY-HN(CO)CACB are shown below with the connectivities confirmed.

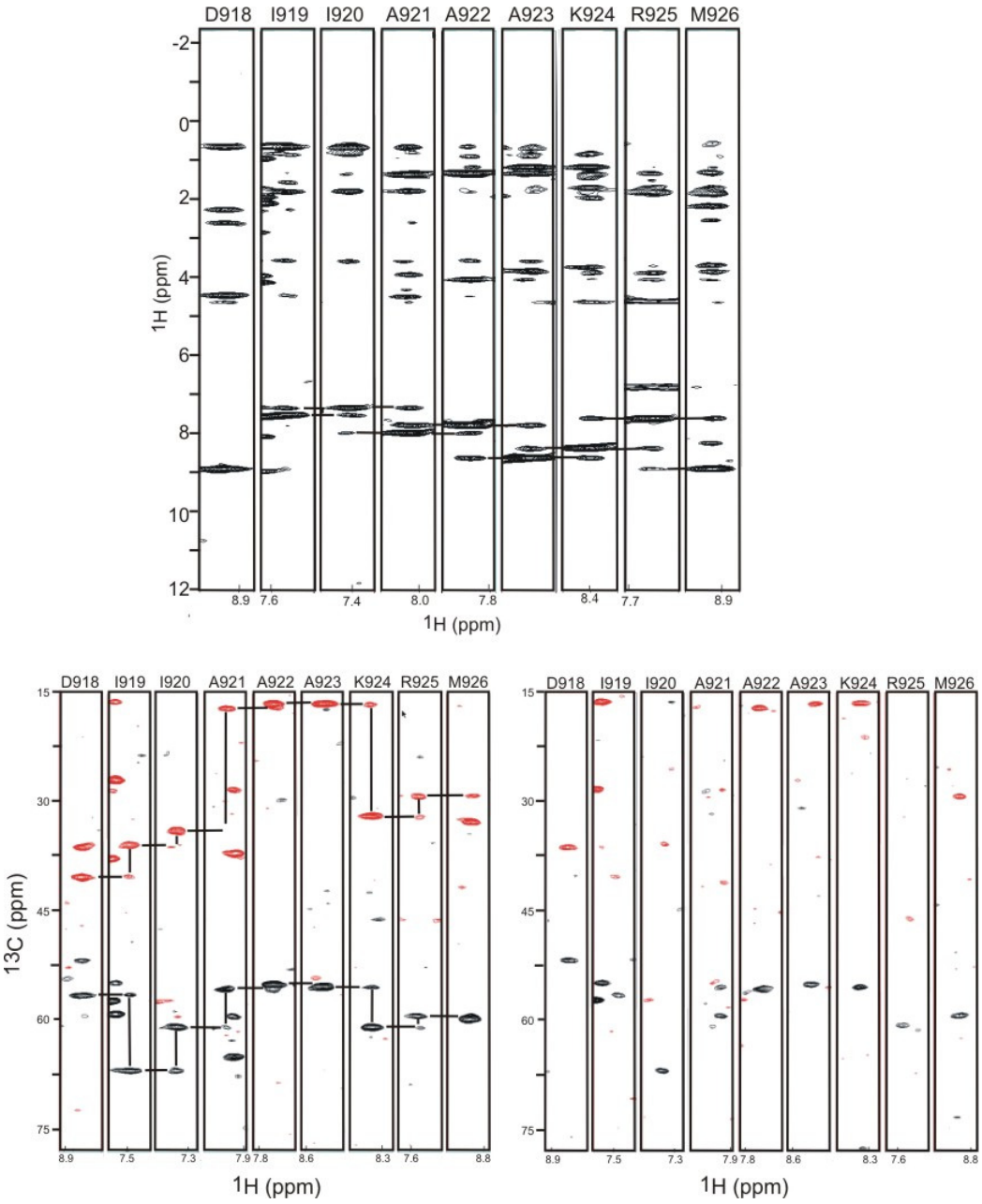


Figure 3.4.7 ^{1}H-^{13}C sections at defined ^{15}N resonances from a ^{1}H-^{15}N-NOESY-HSQC spectrum (top) showing NH connectivities through the helices. Bottom: TROSY-HNCACB & TROSY-HN(CO)CACB confirming the connectivities. All spectra recorded at 800MHz

Vt/I997S at 500µM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. Sequential connectivities are shown as black lines and corresponding assignments shown. Top: the ^{1}H-^{15}N-NOESY-HSQC was carried out on ^{15}N, ^{13}C , ^{1}H Vt/I997S, black peaks correspond to intra- and inter-residual NOEs between protons. Bottom: the triple resonance spectra are from ^{15}N, ^{13}C, 99%^{2}H Vt/I997S. In this spectrum Cα resonances are shown in black and Cβ resonances in red.

Once stretches of sequential resonances were determined, the amino acid specific information contained in the Cα and Cβ shifts should enable them to be placed into the Vt sequence. However, the Vt sequence was found to be degenerate and most assigned stretches could not be placed. Therefore it was necessary to find some 'anchor residues,'- residues whose identity could be determined so that connected resonances could be placed with confidence. This was facilitated by amino-acid selective labelling techniques.

3.4.5 Amino acid selective labelling and unlabelling

Amino acid selective ^{15}N labeling has been carried out on Vt/I997S for the amino acids lysine, valine, isoleucine and leucine, and selective 'unlabelling' of arginines has been completed (Section 2.2.3 and 2.2.4).

Selective labeling of lysine produced a ^{1}H-^{15}N-HSQC spectrum showing 14 NH resonances from an expected 16. Figure 3.4.8 shows this spectrum and an overlay showing which resonances are lysines. There are no chemical shift differences between the position of the peaks in the ^{15}N-labelled and the lysine labelled spectrum but resonance degeneracy has been exposed where a green resonance can be seen slightly to one side of a larger black resonance corresponding to a degenerate peak that contains resonances from at least two amino acids.

The remaining two lysine residues were believed to be missing from the spectrum, but on closer inspection, two potential lysine residues can be identified at very low contour levels in the ^{1}H-^{15}N-HSQC spectrum. The position on these are at 6.9ppm, 118.5ppm and 8.9ppm, 128ppm in the proton and nitrogen dimension respectively (Figure 3.4.10 and Figure 3.4.8). There are 11 residues that have been assigned as lysines from the data, the remaining five resonances must arise from unassigned residues K881, K911, K975, K1035 and K1047.

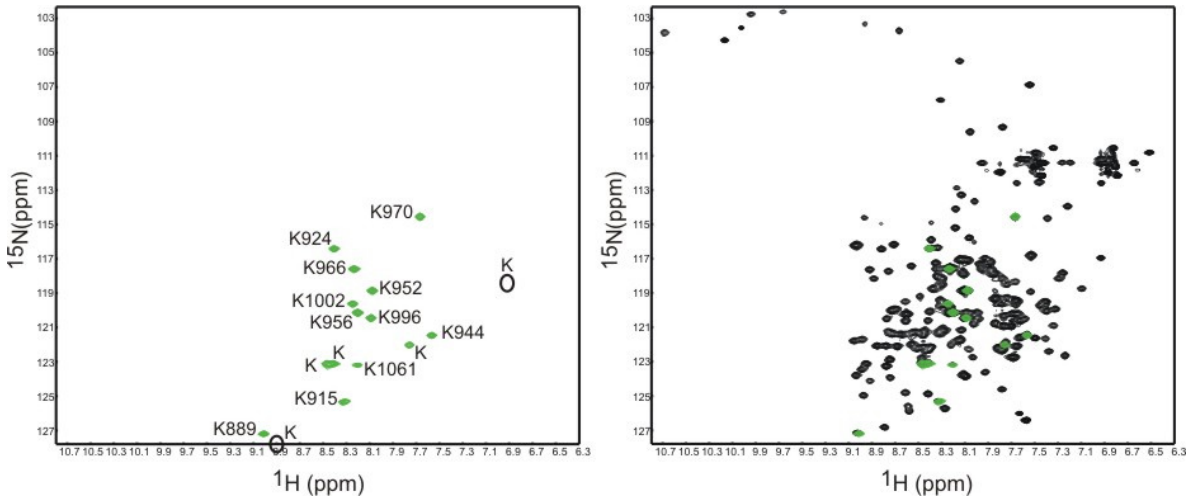


Figure 3.4.8 1H-15N-HSQC spectrum of Vt/I997S selectively labelled with 15N labelled lysine.

Left: recorded NH resonances shown in green with subsequent assignments shown. Black circles represent the two potential lysines that can be observed at low contour levels. Right: overlay of fully labelled ^{15}N Vt/I997S (black) and ^{15}N labelled lysine (green). Spectra acquired at 600MHz on 200μM Vt/I997S in 20mM Tris pH7, 300mM NaCl.

Selective unlabelling of arginine residues was also carried out. In this experiment, ^{14}N arginine is added in excess to an expression grow containing ^{15}N-labeled ammonium chloride. Resonances that originate from arginine will be missing from the ^{1}H-^{15}N-HSQC spectrum. From 15 expected missing peaks, 13 are clearly missing from the spectrum. Figure 3.4.9 shows the results of the selective unlabelling of arginines. Of the 13 resonances identified as arginines, 11 are assigned and two are unknown. This leaves two resonances missing from the spectrum. One of these, R903 did not initially appear due to the fact it was overlapped with another residue (E1014) so was not clearly identified. This was assigned independently of the selective labelling. The remaining peak cannot be found, the reason for this is unknown. R910, R976 and R1039 remain unassigned

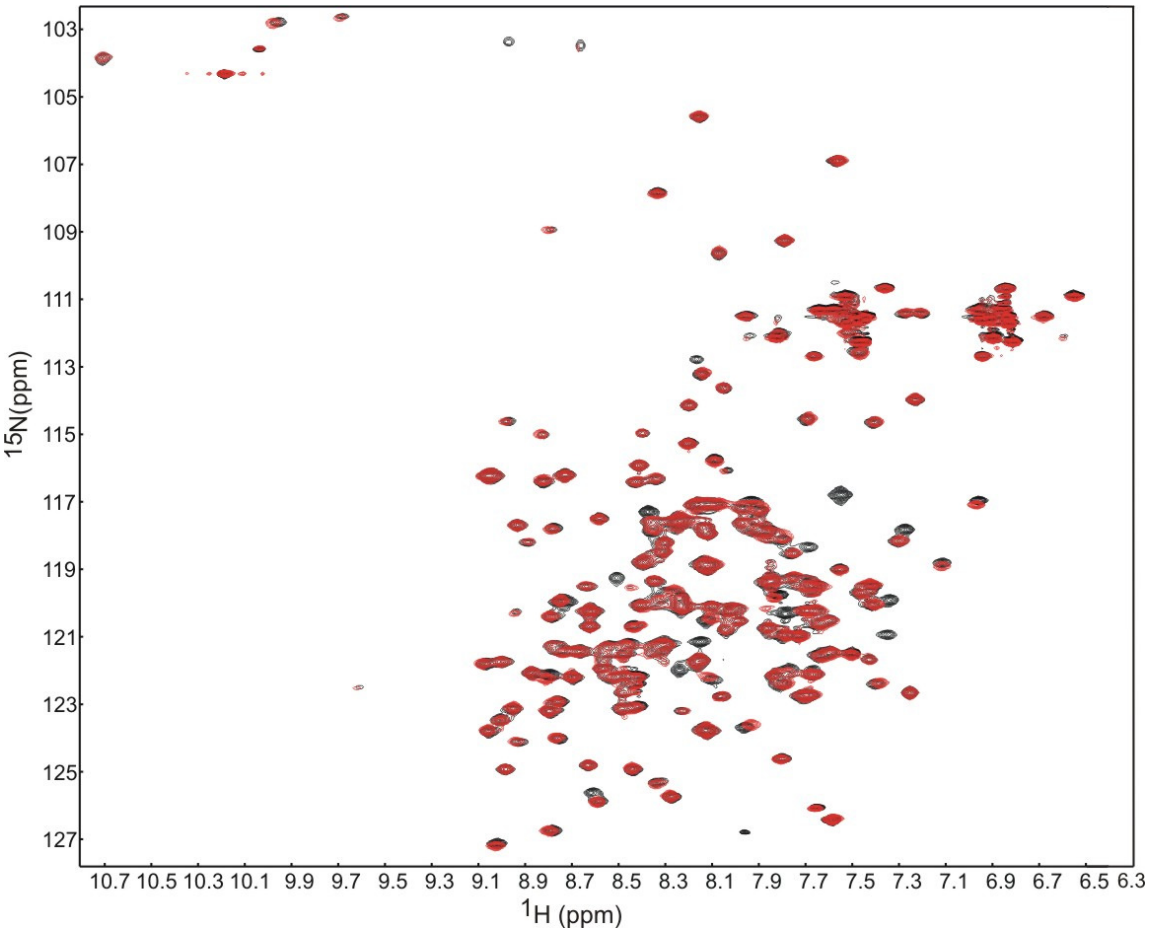


Figure 3.4.9 Overlay of ^{1}H-^{15}N-HSQC spectra for Vt/I997S selectively unlabelled with arginine.

Black is the fully ^{15}N-labelled protein, red is the same protein with arginines selectively unlabeled. Missing arginine resonances are represented as black peaks where there is no red peak on top of it. Spectra acquired at 600MHz on 200µM Vt/I997S in 20mM Tris pH7, 300mM NaCl.

As a consequence of these two experiments, the amino acid type of the resonances coming from 14 arginines and 14 (possibly 16) lysines is determined. This information was vital to facilitate the triple resonance based assignments as they helped to overcome uncertainty when attempting to place sequentially connected resonances into the Vt/I997S sequence. Without this information, the assignment of the Vt/I997S ^{1}H-^{15}N-HSQC spectrum would not have been possible.

Figure 3.4.10 shows all of the lysine and arginine resonances that have now been assigned. There are three arginine resonances and five lysine resonances that remain unassigned.

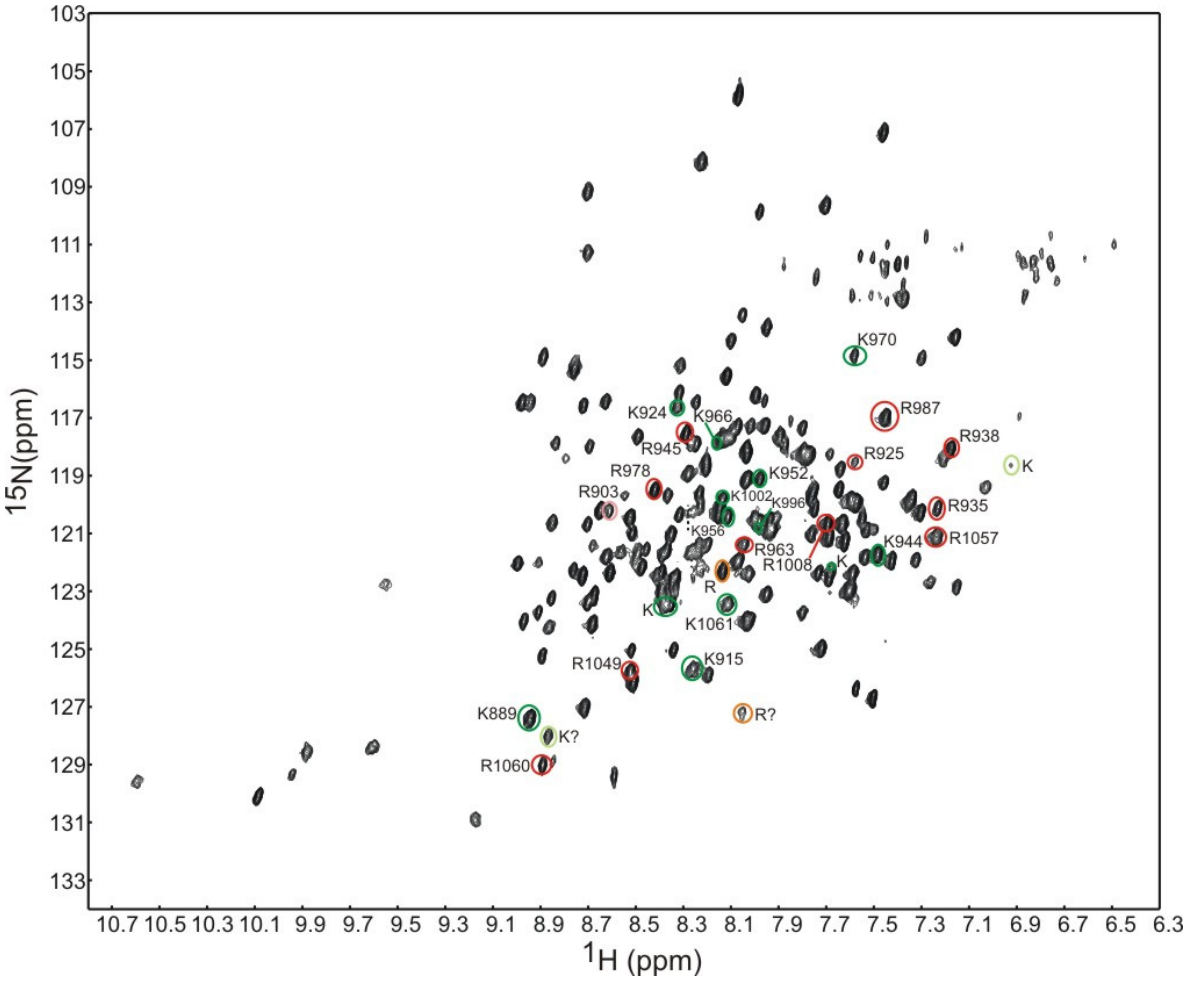


Figure 3.4.10 ^{1}H-^{15}N-TROSY spectrum of Vt/I997S showing assigned lysine (green) and arginine (red) NH resonances. Spectrum acquired at 800MHz

Lysines found at low contour levels are shown in pale green and unassigned arginines are shown in orange. Pink denotes R903 that was not identified by selective unlabelling experiments.

Unlike the lysine selective labelling experiments, selective labelling of valine, leucine and isoleucine did not produce a discrete set of peaks corresponding to the expected number of amino acids. Figure 3.4.11 shows the results of the valine selective labelling experiment. In this spectrum, 110 peaks can be identified at differing intensities (at the contour level

shown in Figure 3.4.11 there are 60). There are 9 valine residues in Vt/I997S so there are many more resonances in the spectrum than expected.

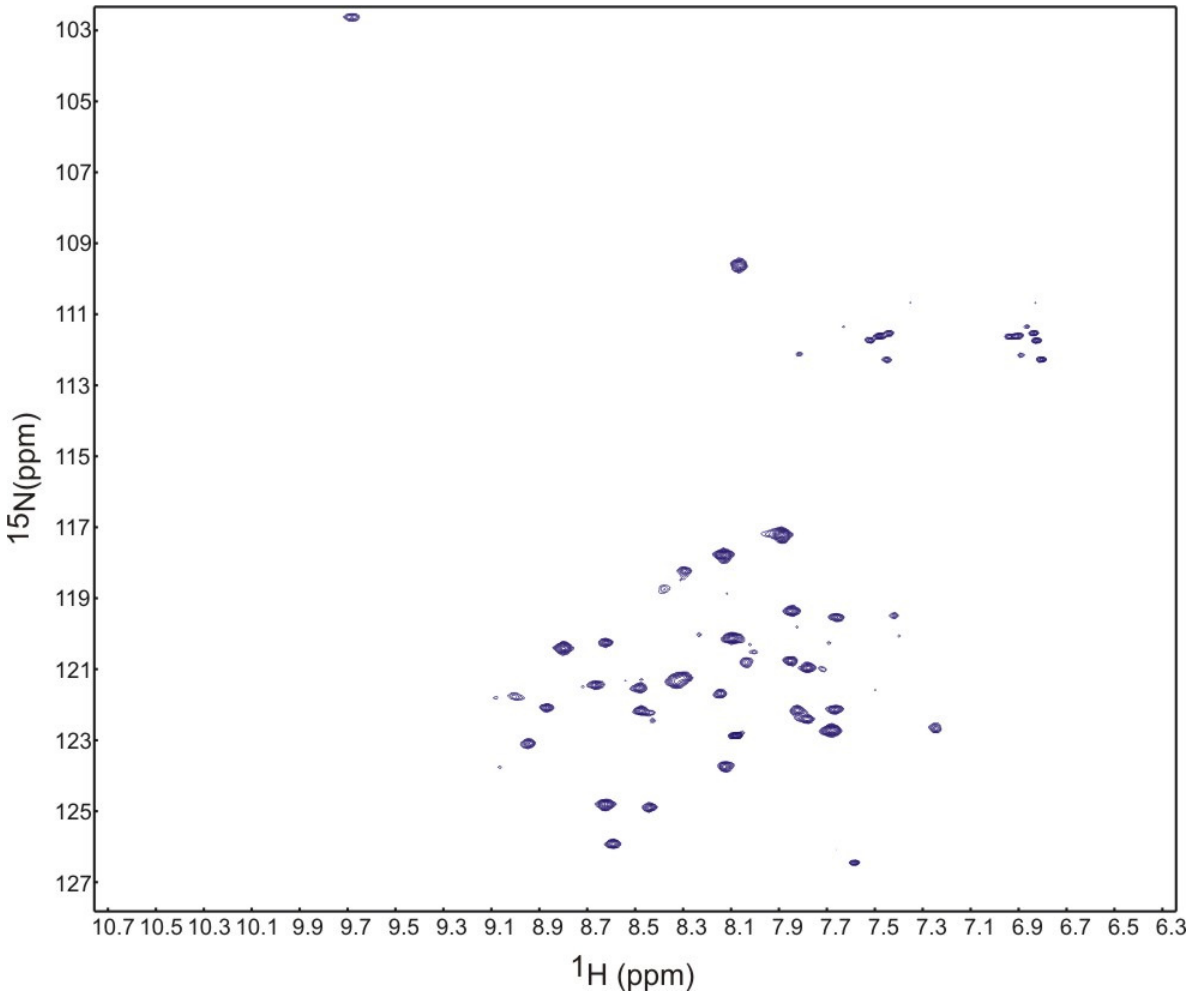


Figure 3.4.11 ^{1}H-^{15}N-HSQC spectrum of Vt/I997S selectively labelled with ^{15}N labelled valine.

The relative intensity of all resonances in the valine, isoleucine and leucine spectra were compared to fully ^{15}N-labelled protein and a plot was made to see if a set of intense peaks could be identified that would correspond to the expected amino acids (Figure 3.4.13). While there are some resonances that are much more intense than others, this data was not used for the purpose of assignment for fear of it being misleading.

In retrospect, this information has been checked to see whether it corresponds to the assignments that have now been completed. Valine and isoleucine resonances that had a relative intensity (relative to the ^{15}N fully labeled sample) greater than 1.5 and leucine resonances greater than 1 and 0.5 were picked and cross checked to see whether these corresponded to the appropriate residue.

Eleven resonances for valine met the above criteria (Figure 3.4.12). Of these, six had been identified as valines in the assignment. Four were unassigned and one was assigned as I991. This isoleucine also met the criteria in the isoleucine experiments.

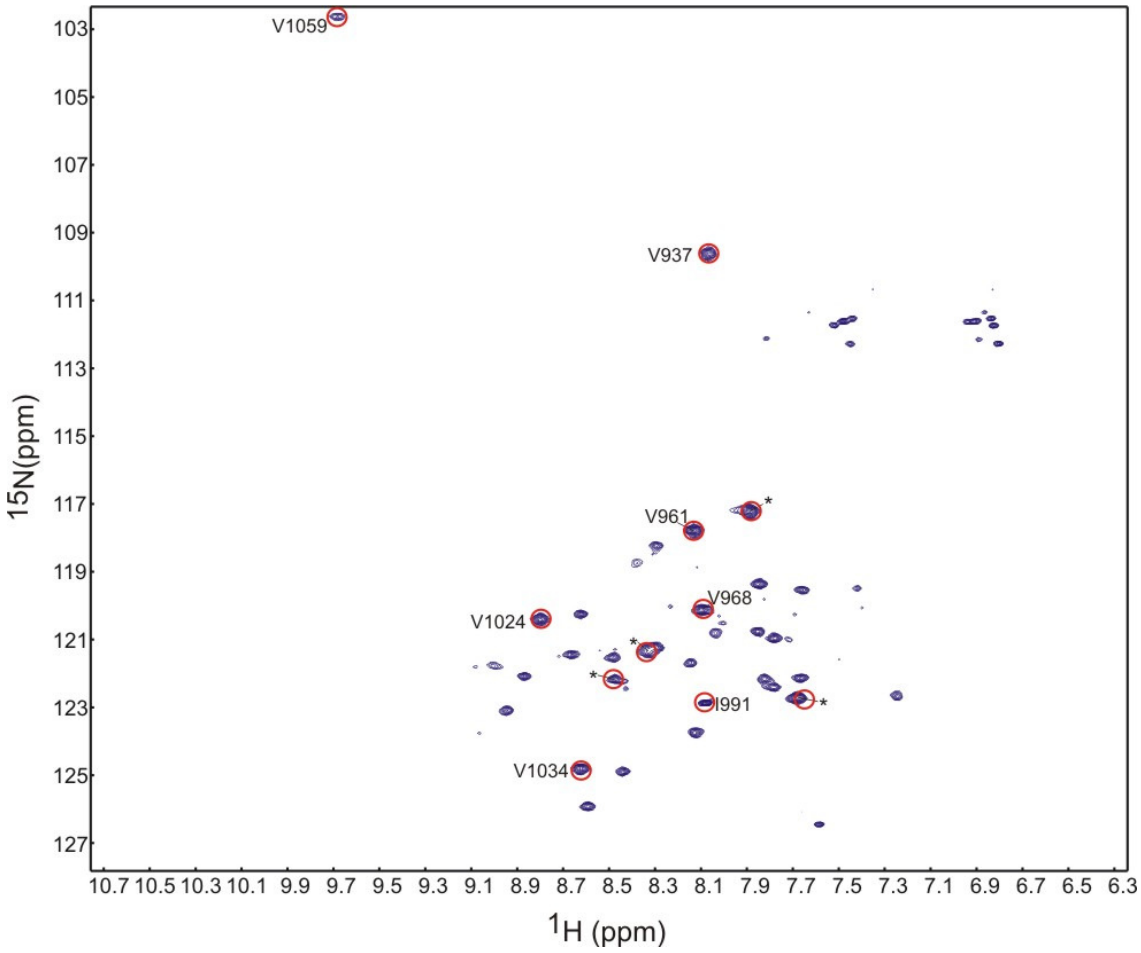


Figure 3.4.12 ^{1}H-^{15}N-HSQC spectrum of Vt/I997S selectively labelled with ^{15}N labelled valine showing eleven residues with the greatest relative intensity

*Blue peaks are all of the resonances that appeared in the spectrum, the most intense are marked with a circle. * represent unassigned resonances.*

While this data was not specific enough to aid in the assignment, it was used to verify the assignments of the valine residues in Vt/I997S. The isoleucine resonances also corresponded with many of the already assigned amino acids, although there were more resonances picked that were not isoleucine compared to the valine experiment. Many of these wrongly picked residues have been assigned as leucines. The leucine selective labelling was the least specific with many picked peaks coming from isoleucine and in some cases alanine. There were many peaks that did however correspond with assigned leucines.

It appears that isotope scrambling has occurred between valine, isoleucine and leucine. This is likely due to the action of transaminases (275). Future experiments of this sort might benefit from shorter induction times and addition of transaminase inhibitors such as α-aminooxyacetic acid (276) or β-chloro-L-alanine (277).

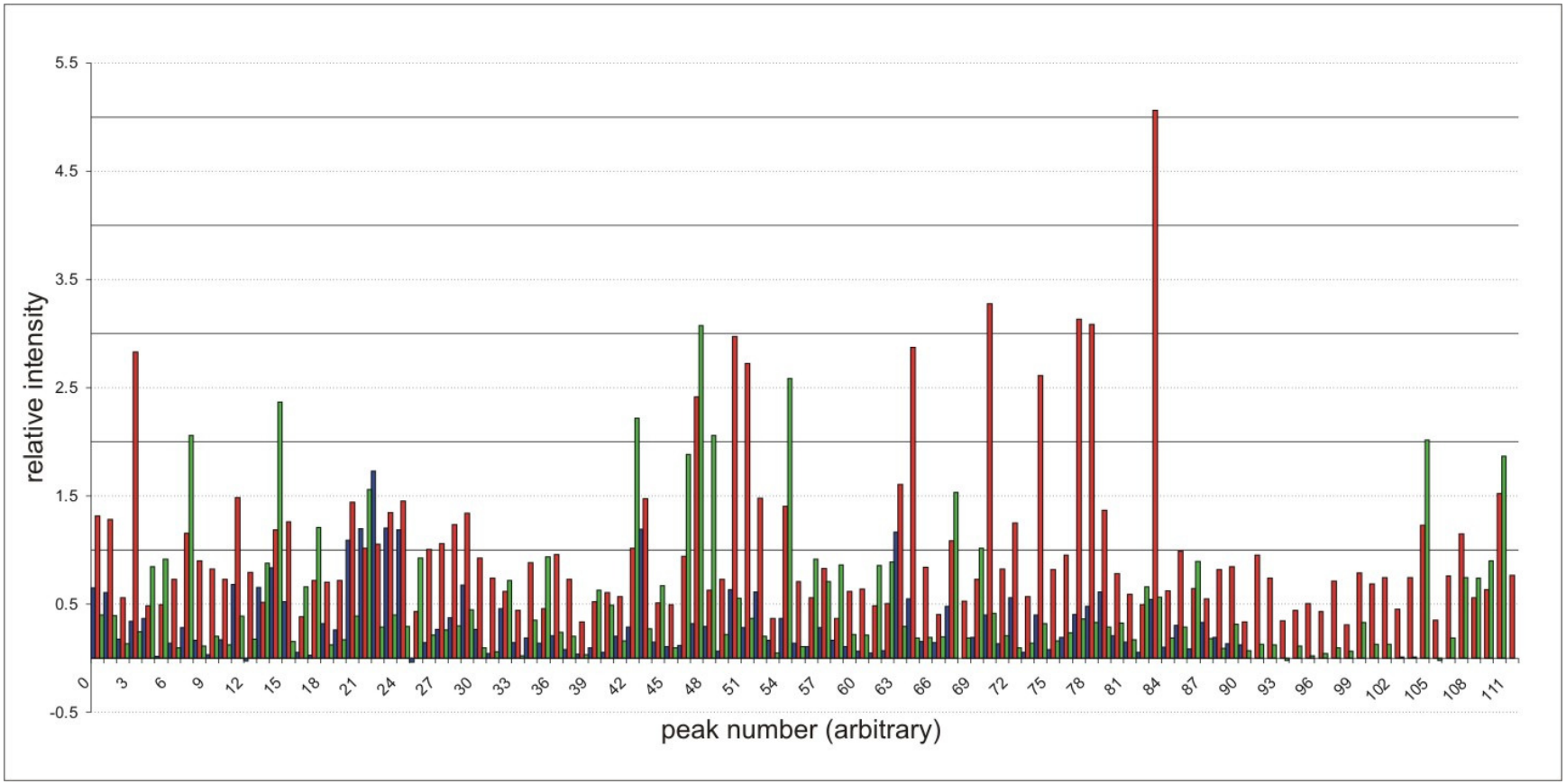


Figure 3.4.13 Relative intensity of resonances originating from valine (green), leucine (blue) and isoleucine (red) selective labelling experiments

Upper limits were set at relative intensity >1.5 for valine and isoleucine, >1 and >0.5 for leucine.

3.4.6 Secondary structure of Vt/I997S

The secondary ^{13}C chemical shift values for Vt/I997S in solution have been used to investigate the secondary structure of the mutant. Figure 3.4.14 shows the chemical shift deviation from random coil for the assigned Cα and Cβ resonances. Helical elements of the wild-type Vt protein from the crystal structure have been indicated as a comparison.

Circular dichroism has already been used to show the α-helical nature on Vt/I997S (Section 3.3.2), chemical shift deviations should give information on the helical nature of individual residues. Globally, the Cα chemical shift deviations suggest that there are 5 helices in Vt/I997S and these correlate with the helical regions found in the wild-type structure. Gaps in the data are due to missing assignments. The Cβ data is less informative, the helical regions from the wild-type correspond with the Cβ chemical shift deviations for Vt/I997S, yet there are not clear breaks in the helices as for the Cα data.

As there appear to be some subtle differences between Vt/I997S and the wild-type, the secondary chemical shift data was used to predict the backbone ϕ– (phi) and φ– (psi) torsion angles using TALOS (Torsion Angle Likelihood Obtained from Shifts and sequence similarity) (267). These predictions were then compared to the known crystal data for the wild-type protein (for methods see Section 2.6.4.3). For most of the Vt/I997S residues, the TALOS-predicted dihedral angles were close to that of the crystal structure. Some of the predictions from the Vt/I997S data were ambiguous in the fact that they did not all fall in the same regions in the Ramachandran plot, and some were also different from the crystal structure. These regions have been represented in Figure 3.4.15 in yellow. There were also some regions where the helicity seemed to extend past that of the expected crystal data, these have been represented in red. One of these residues R987 is a residue whose NH resonance appears as one in the Vt/I997S ^{1}H-^{15}N-HSQC spectrum but as two in the wild-type spectrum. Further evaluation of the significance of these differences between wild-type and Vt/I997S will require relaxation data and residual dipolar couplings to be measured to obtain dynamic and structural data.

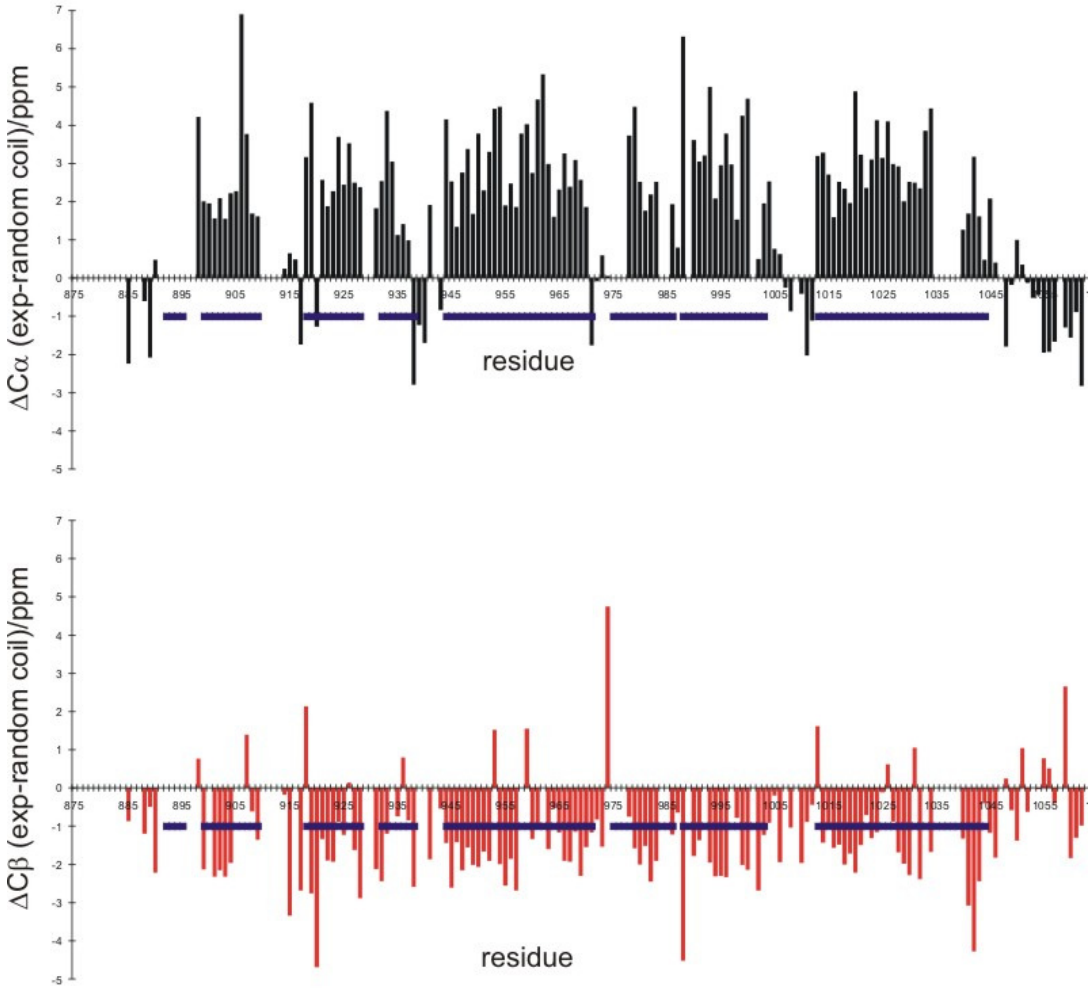


Figure 3.4.14 Chemical shift deviations for assigned residues in Vt/I997S

Black: Vt/I997S Cα chemical shift deviations from random coil peptides in urea. Red: Vt/I997S Cβ chemical shift deviations. Helical regions from the crystal structure of wild-type Vt are shown as blue lines.

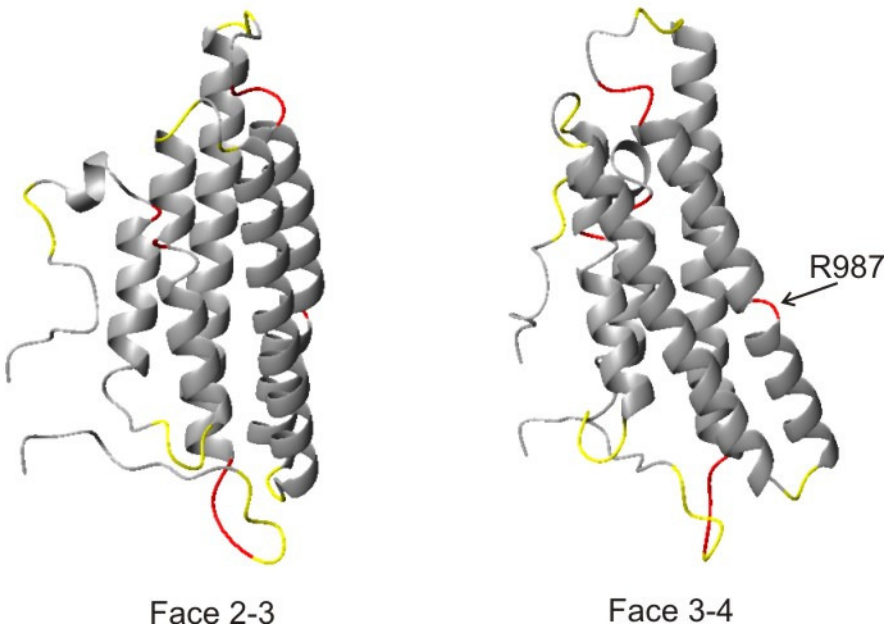


Figure 3.4.15 Regions in Vt where the ϕ- and φ– angles differ between the measured crystal structure and predicted torsion angles in solution

Red regions are where the helicity extends past what is expected from the crystal structure; yellow regions are where predictions are not good and/or differ from the crystal structure.

The region around the I997S mutation gives predictions in the α-helical region and are all similar to the isoleucine in the crystal structure, and the predictions for residues that have been implicated in a crystallographic dimer interface between asymmetric units (Figure 1.3.8 and Figure 1.3.9) all agree with the crystal structure. The only unassigned residue in this region is V1001.

3.4.7 Exchange effects on Vt/I997S

Several residues in Vt/I997S are exchanging on an NMR timescale in the ^{1}H-^{15}N-HSQC spectrum. These were identified with the help of Dr. Stuart Findlow, and were identified as different NH chemical shifts possessing identical ^{13}C chemical shift patterns.

Three of these residues include D1051, A1052 and G1053. On inspection of the crystal structure of wild-type Vt (172) these residues are involved in the self-association crystal interface between the N-and C-terminus of two Vt molecules, distinct from the

hydrophobic interface centred around I997 and M1022. This occurs within the asymmetric unit between the C-terminus of one Vt molecule and the N-terminus and helix 2 in another (Figure 3.4.16). D1052 can make a salt bridge to R935, and it can also hydrogen bond to G1053 (Figure 3.4.17). It is possible that this interface between two Vt molecules could occur in solution and would create an equilibrium between Vt dimers and monomers.

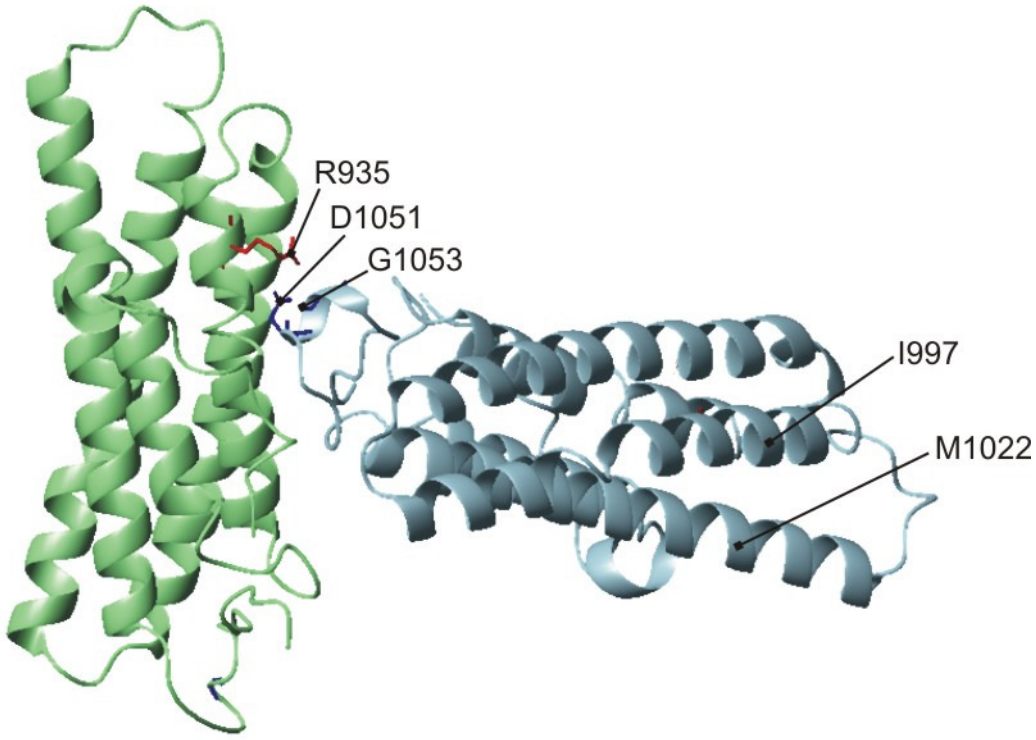


Figure 3.4.16 Two wild-type Vt molecules packing in the asymmetric unit

The C-terminal domain of one Vt molecule (blue) interfaces with helix 2 of a second molecule (green). A salt bridge can form between D1051 and R935, and D1051 can hydrogen bond to G1053. The hydrophobic interface residues I997 and M1022 are shown for reference.

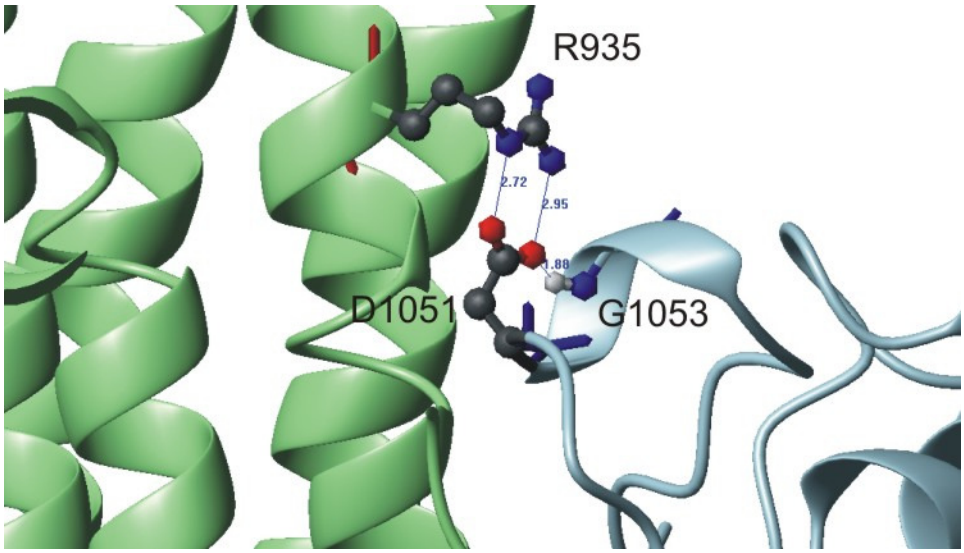


Figure 3.4.17 Close up of the region involved in the crystal interface between two Vt molecules

D1051 can form a salt bridge to R935 (distances 2.72Å and 2.95Å) and can hydrogen bond to G1053 (1.80Å).

Therefore, two potential interaction sites between wild-type Vt molecules could be acting in solution, the first between the hydrophobic regions centred at I997 and M1022, and the second between the C-terminus and helix 2. This may in part explain the larger than expected correlation time of 13.0ns for wild-type Vt, and may also explain why the single point mutations at I997S only reduced the τ_m to 12.6ns. Even with one self-association interface mutated, the second may still be having an effect in the Vt/I997S mutant.

3.4.8 Transferring Vt/I997S assignments to the wild-type ^{1}H-^{15}N-HSQC and TROSY spectrum

The assignments from Vt/I997S were transferred from the ^{1}H-^{15}N-TROSY spectrum measured at 800MHz onto a 600MHz ^{1}H-^{15}N-HSQC spectrum and a 900MHz ^{1}H-^{15}N-TROSY spectrum of wild-type Vt so that subsequent experiments using wild-type protein can be easily interpreted. The resonances have been transferred to the 900MHz spectrum

as well as the 600MHz as there is more resolution in the former so that individual peaks can be observed where at 600MHz they are overlapped.

While many of the NH resonances in the Vt/I997S spectrum coincide with the resonances in the wild-type spectra, there are some that cannot be transferred with complete confidence. A confidence level was assigned arbitrarily, resonances that occurred in an identical chemical shift in both spectra was assigned a value of '5' and the resonances that could not be identified with confidence were assigned the value '1'. When the assignments were being transferred there were certain regions with the same number of peaks but the patterns were different. These can be seen in Appendix 5.6, and show how confidence of 1 does not mean that they should be disregarded, just that care should be taken when interpreting data from these assignments.

In total, 138 of the possible 147 assignments from Vt/I997S were transferred to the wild-type spectrum. Of these, 82 were assigned with a confidence of 5, 27 with a confidence of 4, 13 with a confidence of 3, 5 with a confidence of 2 and 9 with a confidence of 1. The eleven residues assigned in Vt/I997S that could not be transferred are M889, H906, L928, K956, T993, S997, S999, K1002, I1011, Q1018 and A1019 (S997 is the mutated residue). The assigned 600MHz spectrum can be seen in Figure 3.4.18 and Figure 3.4.19, and the 900MHz spectrum in Figure 3.4.20 and Figure 3.4.21. The crystal structure of wild-type Vt with the corresponding Vt/I997S assignments and confidence levels is shown in Figure 3.4.22.

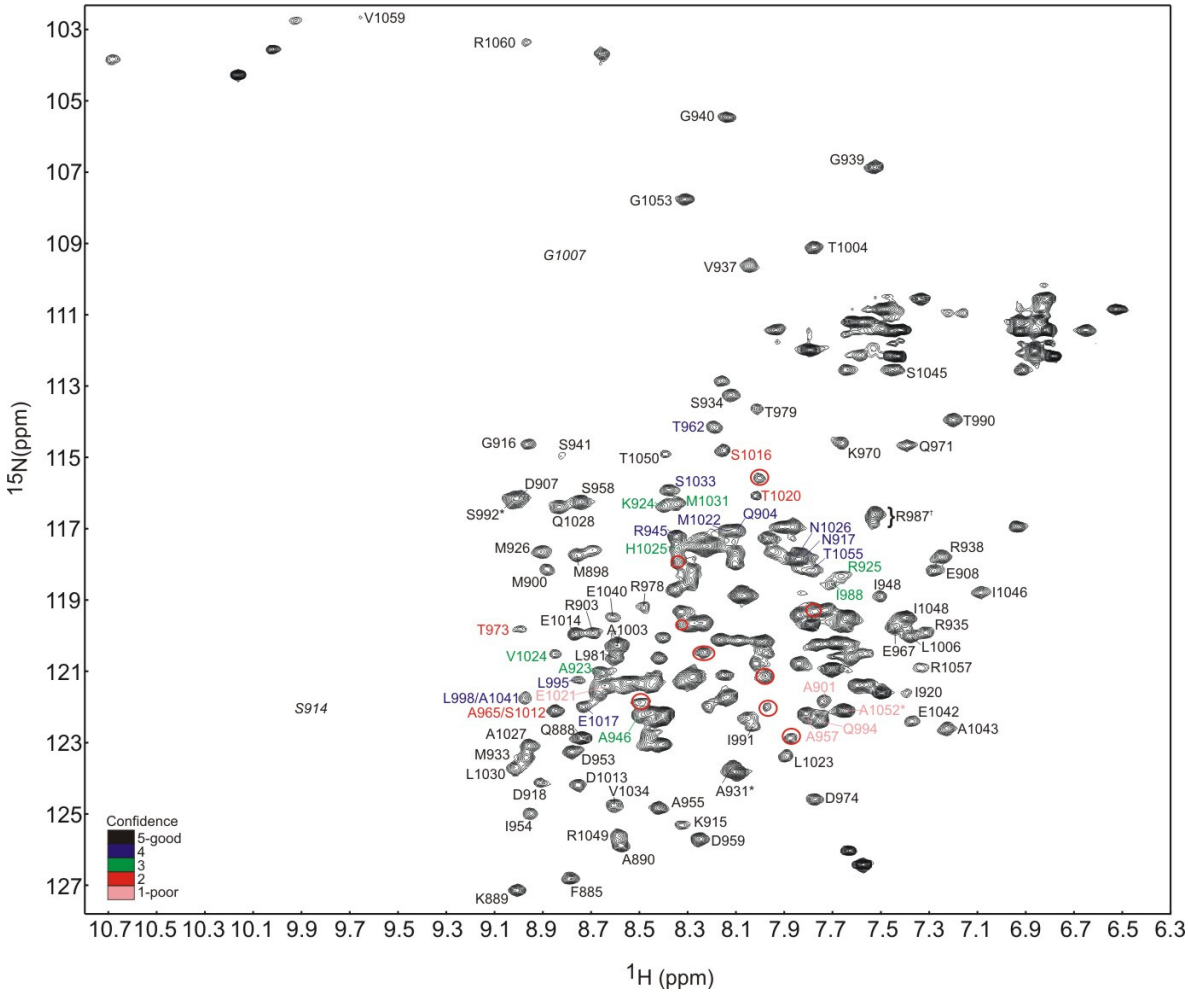


Figure 3.4.18 ^{1}H-^{15}N-HSQC spectrum of wild-type Vt showing transferred NH resonance assignments from Vt/I997S spectrum

Spectrum acquired at 600MHz. Wild-type Vt was at 200µM and labelled with ^{15}N in 20mM Tris pH7, 400mM NaCl. Full spectrum showing the assignments and confidence level, 5(black) meaning high confidence and 1(pink) meaning poor. Confidence scale s shown in bottom left hand corner of spectrum. Resonances marked with '' are degenerate. R987 (marked with †) is one residue in Vt/I997S but may be two residues in wild type protein, and ‡ is a resonance that is only seen at low contour levels. Red circles are resonances that are present in the wild-type spectrum but not in the Vt/I997S spectrum.*

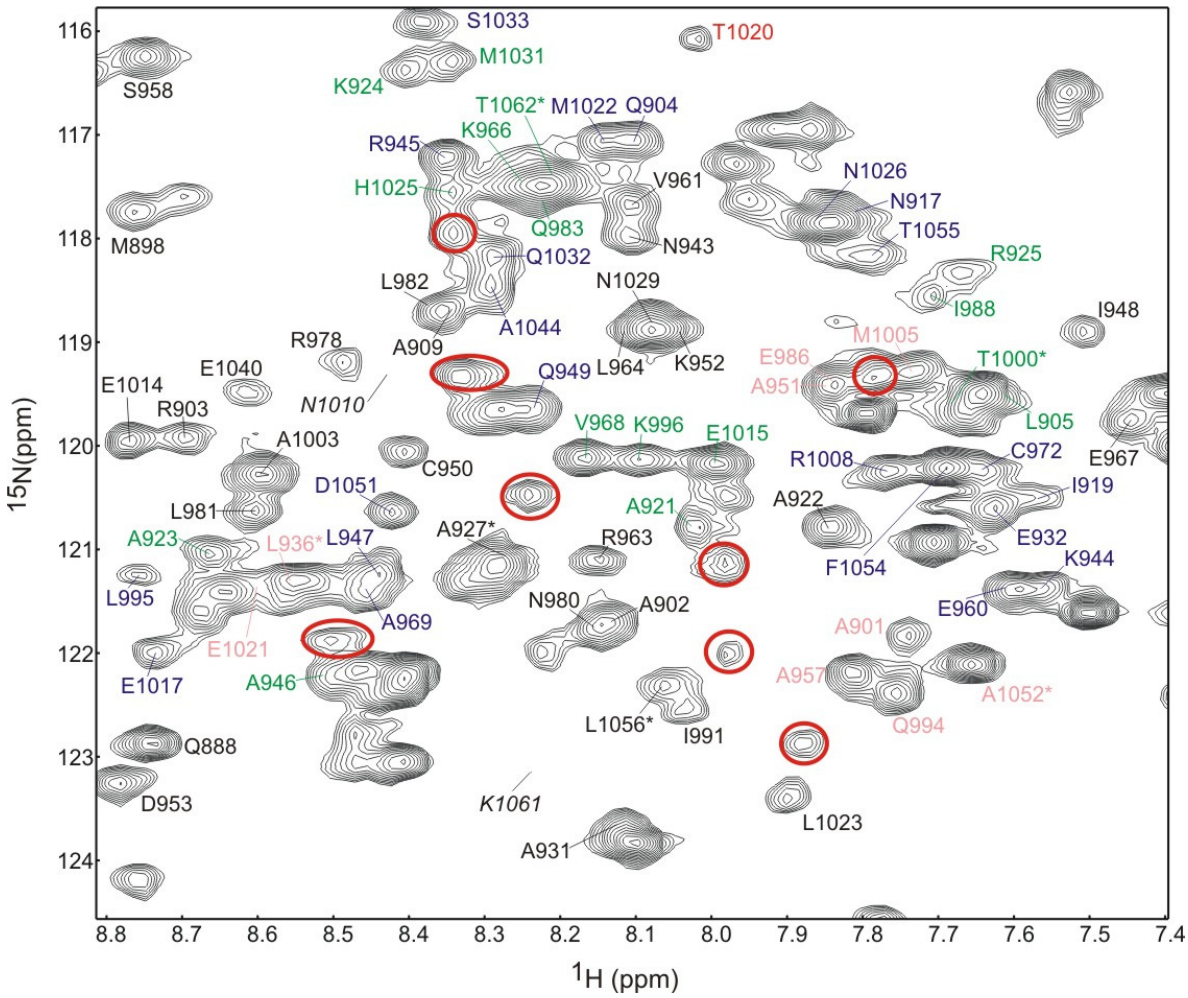


Figure 3.4.19 Expanded section of the ^{1}H-^{15}N-HSQC spectrum of wild-type Vt showing transferred NH resonance assignments from Vt/I997S spectrum

Confidence level, 5(black) meaning high confidence and 1(pink) meaning poor. Confidence scale is shown in bottom left hand corner of spectrum.. Resonances marked with '' are degenerate. Red circles are resonances that are present in the wild-type spectrum but not in the Vt/I997S spectrum*

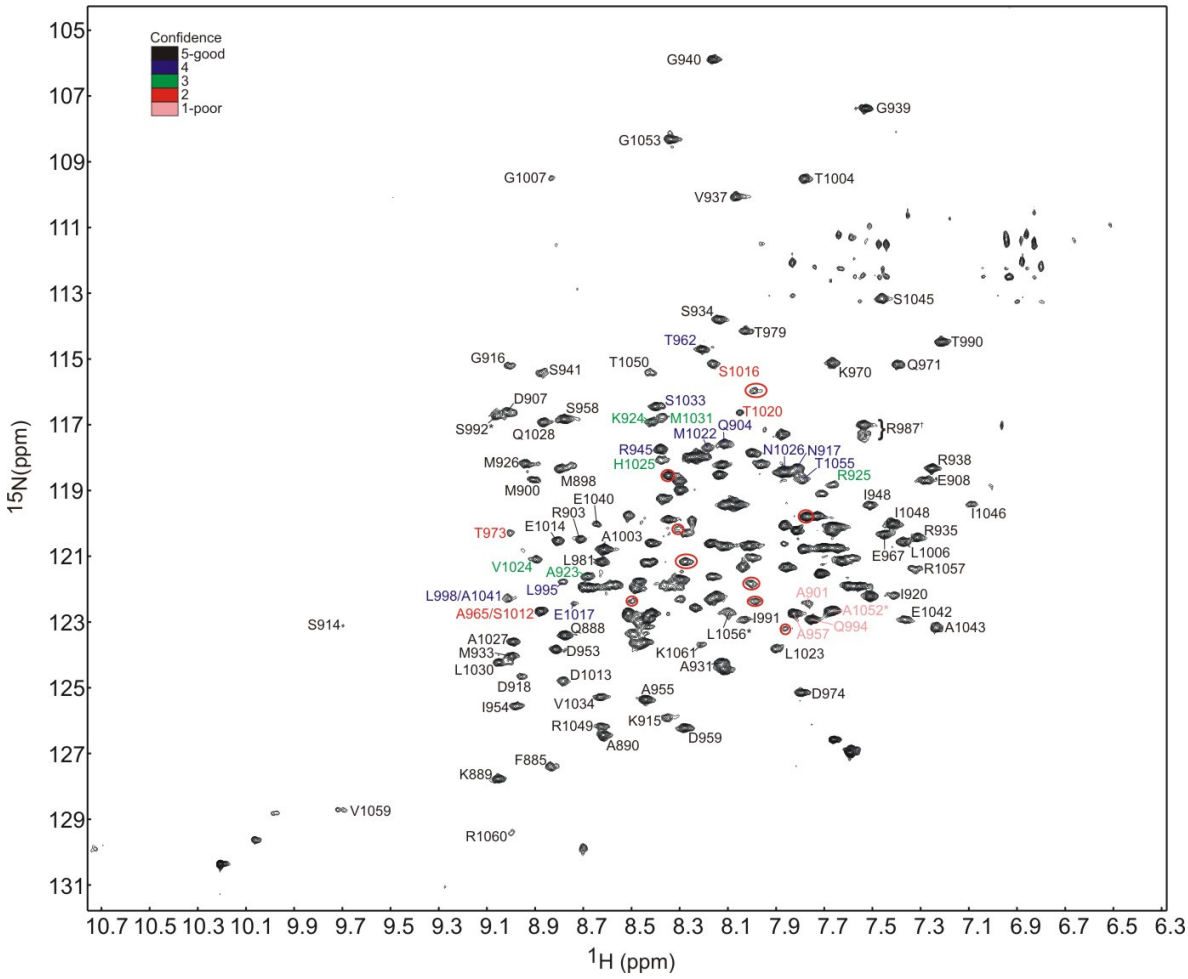


Figure 3.4.20 ^{1}H-^{15}N-TROSY spectrum of wild-type Vt showing transferred NH resonance assignments from Vt/I997S spectrum

Spectrum acquired at 900MHz with attached cryoprobe. Wild-type Vt was at 200µM and labelled with ^{15}N in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate.. Full spectrum showing the assignments and confidence level, 5(black) meaning high confidence and 1(pink) meaning poor. Confidence scale shown in bottom left hand corner of spectrum. Red circles are resonances that are present in the wild-type spectrum but not in the Vt/I997S spectrum

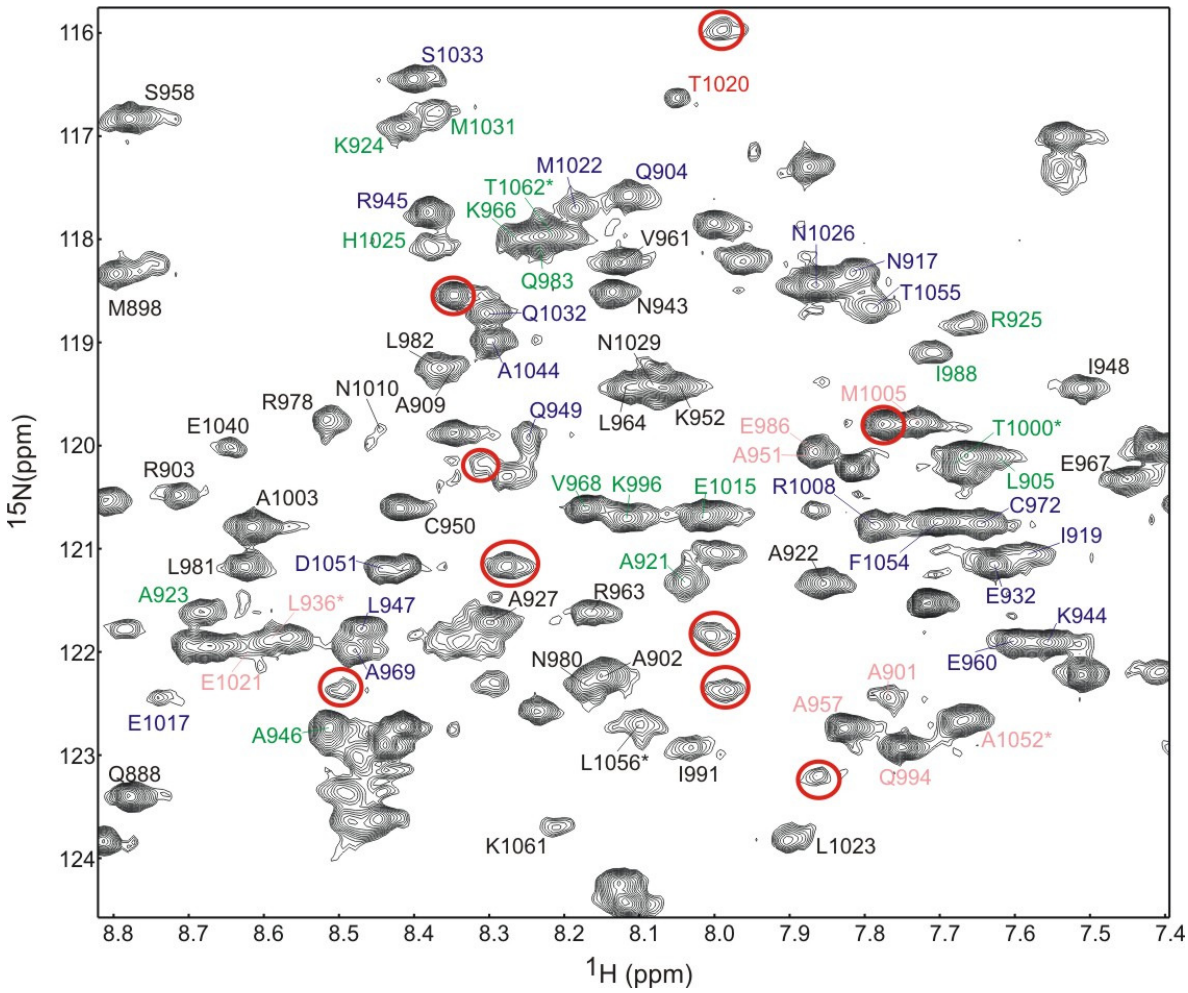


Figure 3.4.21 Expanded section of the ^{1}H-^{15}N-TROSY spectrum of wild-type Vt showing transferred NH resonance assignments from Vt/I997S spectrum

Confidence levels, 5(black) meaning high confidence and 1(pink) meaning poor. Red circles are resonances that are present in the wild-type spectrum but not in the Vt/I997S spectrum

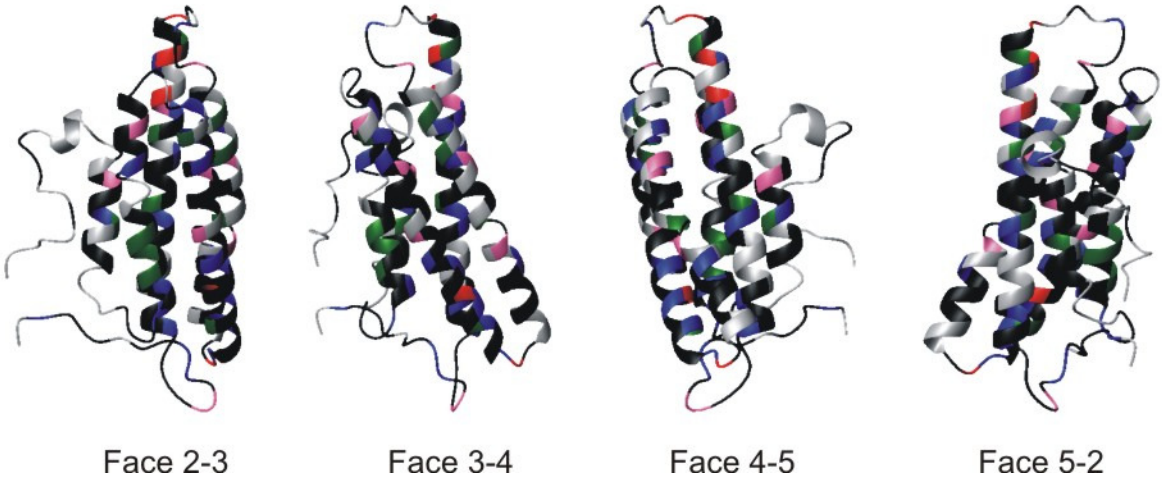


Figure 3.4.22 Wild-type Vt crystal structure shown at 90º rotations with assigned Vt/I997S NH resonances mapped on

Coloured residues correspond to the confidence of assignment. Black 5-good, blue 4, green 3, red 2 and pink 1-poor. Grey residues are either unassigned or not transferred to wild-type.

3.4.9 Summary

Assignment of the Vt/I997S backbone NH resonances has been challenging and using multiple NMR experimental data, 76% of the residues in Vt/I997S have been assigned. The secondary structure of Vt/I997S in solution appears very similar to the wild-type protein as judged by secondary Cα chemical shift deviations and from torsion angle predictions. There may be some subtle structural but further data is required to determine the significance of these differences.

The Vt/I997S NH assignments have been mapped onto the wild-type Vt ^{1}H-^{15}N-HSQC spectrum; of the 147 Vt/I997S NH assignments, 138 have been transferred to the wild-type spectrum. Using these wild-type Vt assignments, it will now be possible to interpret chemical shift perturbation data between Vt and paxillin LD motifs.

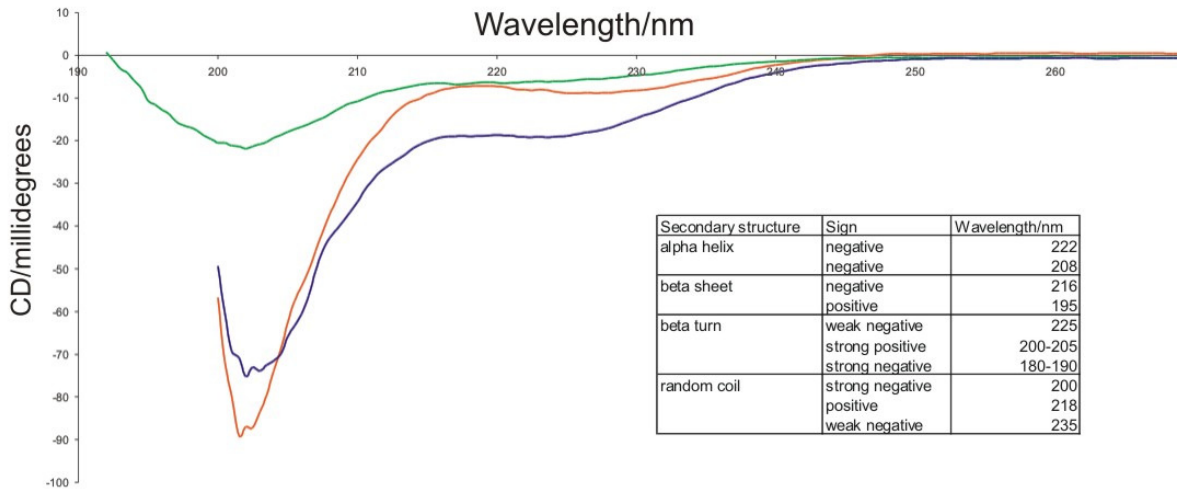
3.5 Characterisation of Paxillin

In order to study the interaction between Vt and paxillin, different paxillin regions have been utilised. Small 13mer peptides containing paxillin LD motifs LD1, LD2 and LD4 have been synthesised and longer paxillin constructs have been recombinantly expressed (Figure 2.6.1). Paxillin constructs containing paxillin LD1 and LD2 in pGEX-LD1/LD2 and pET-15B-LD1/LD2 were obtained from Professor D. Critchley (University of Leicester) and have been purified for use in interaction experiments. Paxillin constructs containing paxillin GST-LD3/LD4 and GST-LD4/LD5 (also obtained from Professor D. Critchley, University of Leicester) have been expressed but at low levels so purification has not been possible.

3.5.1 Secondary structure of paxillin LD1, LD2 and LD4 peptides

The paxillin 13mer peptides have been synthesised for use in chemical shift perturbation studies using NMR. The sequences of LD1, LD2 and LD4 can be found in Chapter 2.6.7.1. As there is some debate in the literature about the secondary structure of LD motifs, far UV circular dichroism was used in an attempt to discover clues about the secondary structure present in these short motifs. Figure 3.5.1 shows the results of the far UV CD on peptides LD1, LD2 and LD4 and a table showing typical values from known secondary structure.

All of the LD peptides have a negative signal at 202nm, LD1 and LD4 have a stronger signal than LD2. They also have a weak negative signal at 225nm. Using the table in Figure 3.5.1, it can be seen that these peptides are not consistent with one singular type of secondary structure. The negative value at 202nm is often seen for random coil while a weak negative value at 225 would be consistent with β-turn structure. CD spectra for oligopeptides notoriously encounter difficulties (211) and it is evident that there is not enough structural information here to predict with confidence the secondary structure of the LD peptides. More structural information will be required to determine the nature of the secondary structure of these peptides.



Secondary structure	Sign	Wavelength/nm
alpha helix	negative	222
	negative	208
beta sheet	negative	216
	positive	195
beta turn	weak negative	225
	strong positive	200-205
	strong negative	180-190
random coil	strong negative	200
	positive	218
	weak negative	235

Figure 3.5.1 Far UV spectra for peptides LD1 (orange) LD2 (green) and LD4 (blue) *All of the peptides have a minima at 202nm and a weak negative signal at 225nm. Peptides at 50μM in water. Typical values are from Rodger and Nordén (211).*

3.5.2 Expression and purification of paxillin containing LD1 and LD2

Paxillin GST-LD1/LD2 and His-LD1/LD2 were over-expressed in LB media. The GST construct produces an induction band above the *E.coli* background when expression is carried out at 16°C for 16 hours. (Figure 3.5.2), induction at 30°C does not produce an induction band. Paxillin His-LD1/LD2 was expressed in LB media and in M9 minimal media at 16°C for 16 hours, however an induction band cannot be observed on an SDS-PAGE gel. Induction can only be confirmed after capture on a Ni-NTA affinity column.

Attempts to purify GST-LD1/LD2 were made but the protein degrades when the GST tag is removed by thrombin cleavage. This construct was used for binding studies from the crude lysate (see 3.6.1 and 3.6.1). Initial attempts at purification of His-LD1/LD2 were hampered by large protein losses as monitored by SDS-PAGE gel. Attempts to remove the histidine tag also resulted in rapid protein loss.

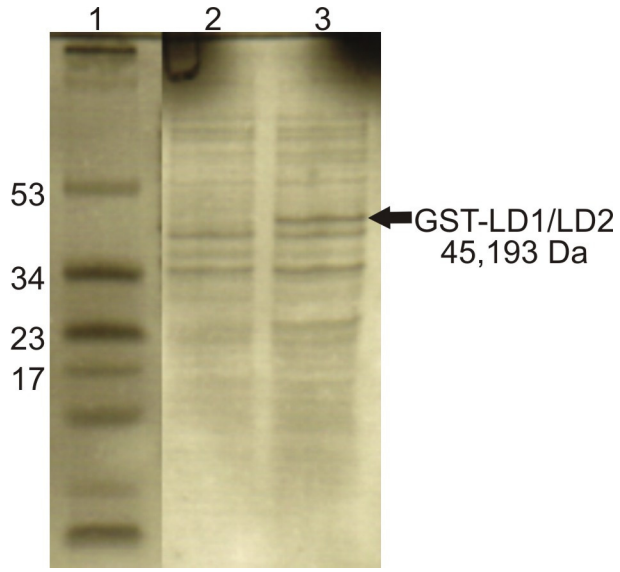


Figure 3.5.2 Composite polyacrylamide SDS-PAGE gel showing induction of paxillin GST-LD1/LD2 using LB medium after induction with 1mM IPTG for 16 hours at 16ºC.

Lane 1: molecular weight standard-multimark, with molecular weights of products alongside (in kDa); lane 2: pre induction sample; lane 3: 16 hour post induction. Induction bands can be seen corresponding to the molecular weight of 45kDa.

The N-terminus of paxillin is suggested to be natively unstructured (206) hence expression of N-terminal fragments may be plagued by unspecific proteolysis. Successful purification of the His-tagged LD1/LD2 paxillin construct has been enabled using 6M urea for cell lysis and affinity capture. The details of the purification are outlined in Chapter 2.2.8. The *E.coli* cell pellet is resuspended in a tris buffer containing 6M urea. Sonication is performed in this buffer, the cell lysate is clarified and the soluble fraction loaded onto a Ni-NTA column. The eluent does not contain urea, so the protein is buffer exchanged on the column. Treating the protein in this harsh manner results in strong protein band on an SDS-PAGE gel and also increases the integrity of the construct throughout purification. The chromatogram from this column is shown in Figure 3.5.3. The 100% fraction is pooled and continues through purification.

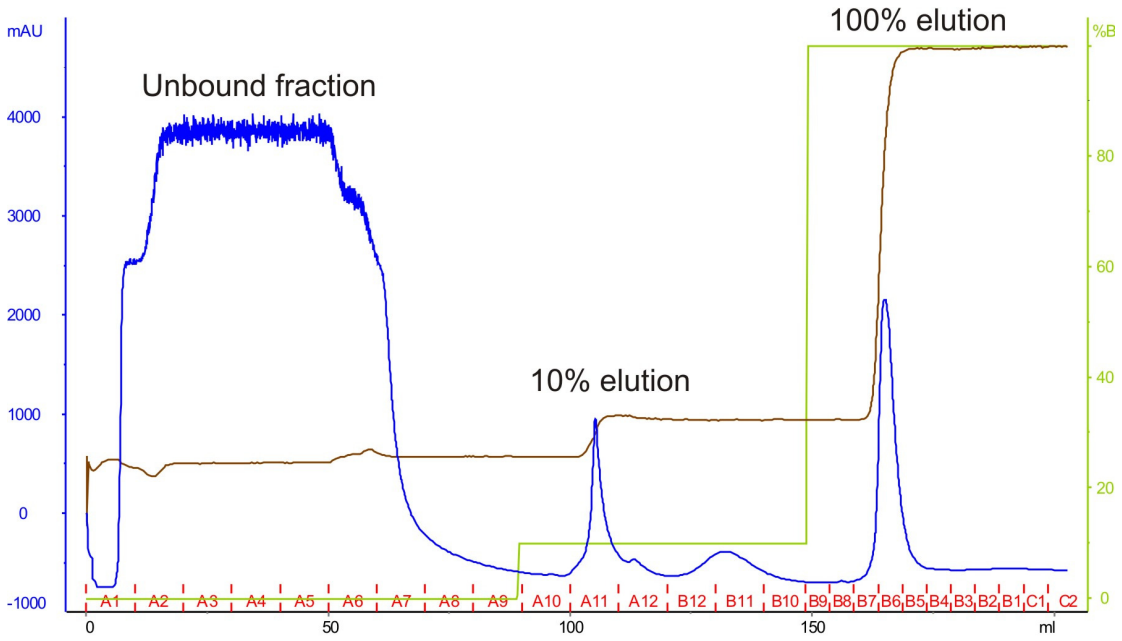


Figure 3.5.3 Paxillin His-LD1/LD2: Ni-NTA loading and elution profile.

The blue trace is the absorbance at 280nm and is used to monitor the proteins. The green trace shows the percentage of imidazole being used to elute, the first elution is 10% and the second is 100%. The brown trace monitors the conductivity. Fractions are shown in red along the x-axis.

Attempts at removal of the histidine tag have resulted in loss of the protein throughout the duration of the purification, after thrombin cleavage a singular band can be observed by SDS-PAGE, but the protein subsequently disappears. Due to the small size of the tag (6x histidines) it has not been removed from the paxillin construct in this study.

The second column is an anion exchange column, however the construct does not bind to the column, it is collected in the unbound flow-through fraction. Figure 3.5.4 shows the elution profile. Rather than attempting to bind his-LD1/LD2 to a cation column, it was collected as the unbound sample as it resulted in a low concentration of salt in the sample which would help during later desalting steps.

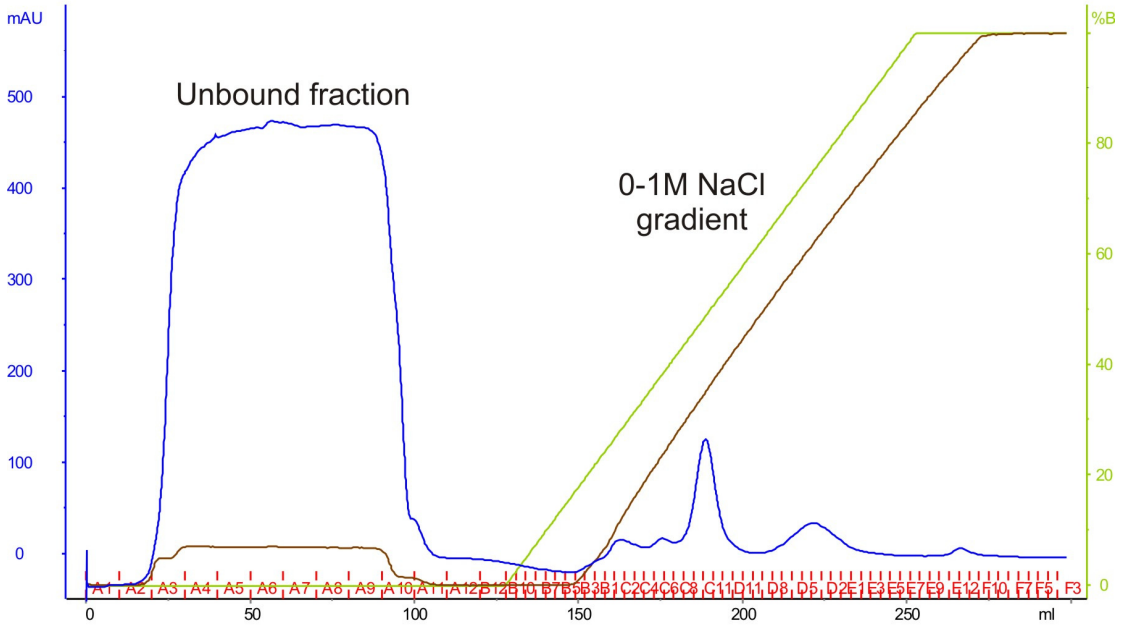


Figure 3.5.4 Paxillin His-LD1/LD2: Q-Sepharose (anionic) elution profile.
The blue trace is the absorbance at 280nm and is used to monitor the protein. The green trace shows the percentage of 1M NaCl being used to elute, the gradient is run from 0-100%. The brown line is a measure of the conductivity of the elution fractions. Fractions are shown in red along the x-axis.

The protein is lost through the membrane of spin concentrators with a 5,000 molecular weight cut-off, despite the molecular weight of paxillin his-LD1/LD2 being 19769 Da. For NMR studies, it is necessary that the protein concentration can be adjusted for different experiments, so the protein is next buffer exchanged into water and then freeze dried. Dialysis was not attempted due to worries that it may be lost through the membrane, instead the protein is run over a sephadex G25 column to desalt into water. The desalting profile is shown in Figure 3.5.5. The absorbance at 280nm is recorded and then the protein is then freeze dried in known quantities for 2-4 days.

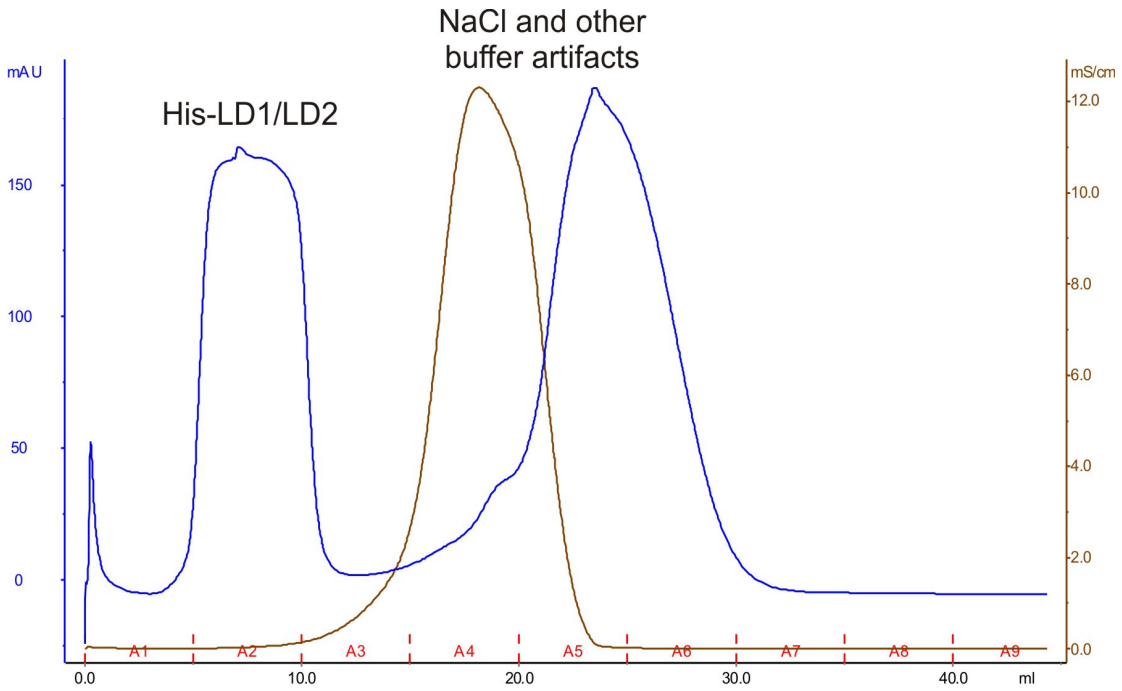


Figure 3.5.5 Paxillin His-LD1/LD2: Sephadex G25 elution profile.

The blue trace is the absorbance at 280nm and is used to monitor the protein. The brown line is a measure of the conductivity of the elution fractions. Fractions are shown in red along the x-axis.

Figure 3.5.6 shows a gel of the construct as it is purified. The final product is >95% pure and the yield is approximately 7mg/L.

It became apparent that care needs to be taken with the desalting step, as NMR data acquired at 900MHz showed presence of both urea and imidazole in a batch of His-LD1/LD2. Figure 3.5.7 shows how the imidazole peaks can interfere with the resonances in a ^{1}H-^{15}N-TROSY spectrum. Future protein preparations should be desalted several times to ensure complete removal of these contaminants from the sample.

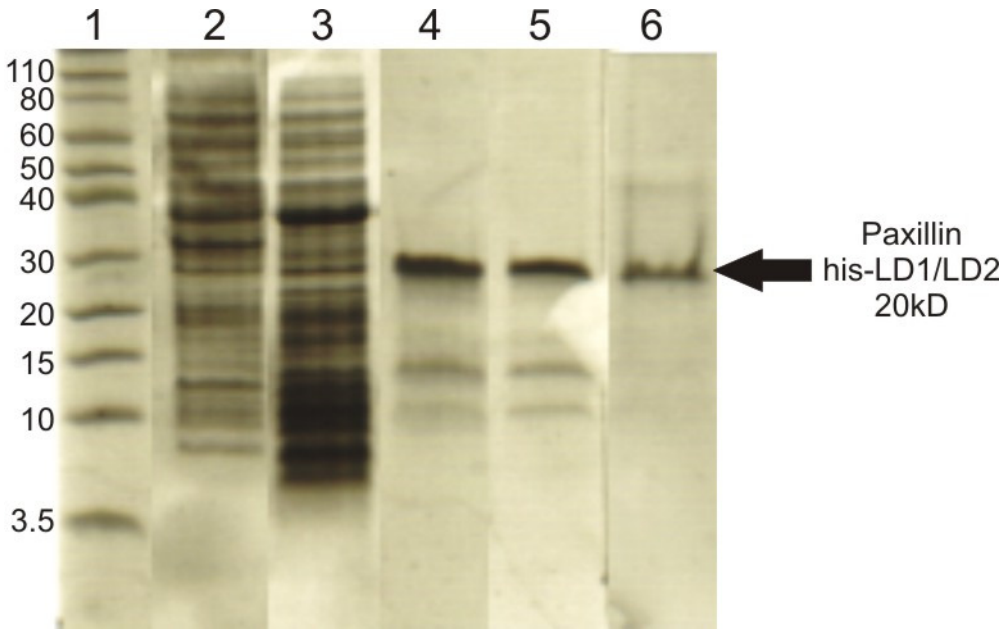


Figure 3.5.6 Composite polyacrylamide SDS-PAGE gel showing paxillin his-LD1/LD2 at various stages of the purification.

Lane 1: molecular weight standard-multimark, with molecular weights of products alongside (in kDa); lane 2: Ni-NTA load, paxillin cannot be observed; lane 3: Ni-NTA flow-through showing that no major band can be seen corresponding to paxillin; lane 4: pooled paxillin his-LD1/LD2 from Ni-NTA column ready to be loaded onto the anion column; lane 5:pooled unbound fraction from ion exchange ready for loading onto the desalting column; lane 6: Paxillin his-LD1/LD2 post desalt & post freeze-dry, resuspended in 20mM Tris pH7, 150mM NaCl.

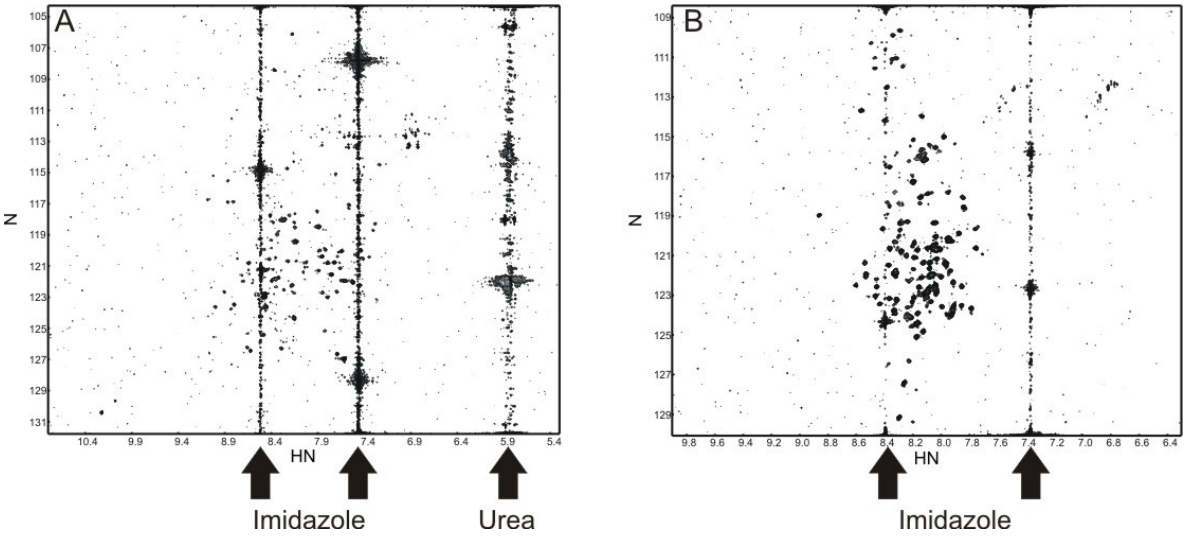


Figure 3.5.7 ^{1}H-^{15}N-TROSY spectra of wild-type Vt (200µM) plus an excess of paxillin (A); and paxillin His-LD1/LD2 at 100µM (B) at 900MHz with a cryoprobe showing that urea and imidazole have contaminated the sample

3.5.3 Mass spectrometry

The mass spectra of unlabelled ^{14}N and labelled ^{15}N-paxillin His-LD1/LD2 have been carried out (Figure 3.5.8 and Table 3.5.1). The mass spectrum results show that the construct is slightly smaller than expected. The difference in mass could be due to Met being removed from the N-terminus (see sequence in Appendix 5.4), inputting the sequence minus the first methionine into a molecular weight prediction program(270) returns the predicted molecular weight 19639 which is identical to the data obtained. Removal of this methionine would not interfere with the binding motifs.

Table 3.5.1 Results of mass spectrometry for ^{14}N and ^{15}N labelled His-LD1/LD2

	^{14}N-His-LD1/LD2	^{15}N- His-LD1/LD2
Predicted molecular weight	19769	20013
Measured molecular weight	19639	19876
Difference	130	137

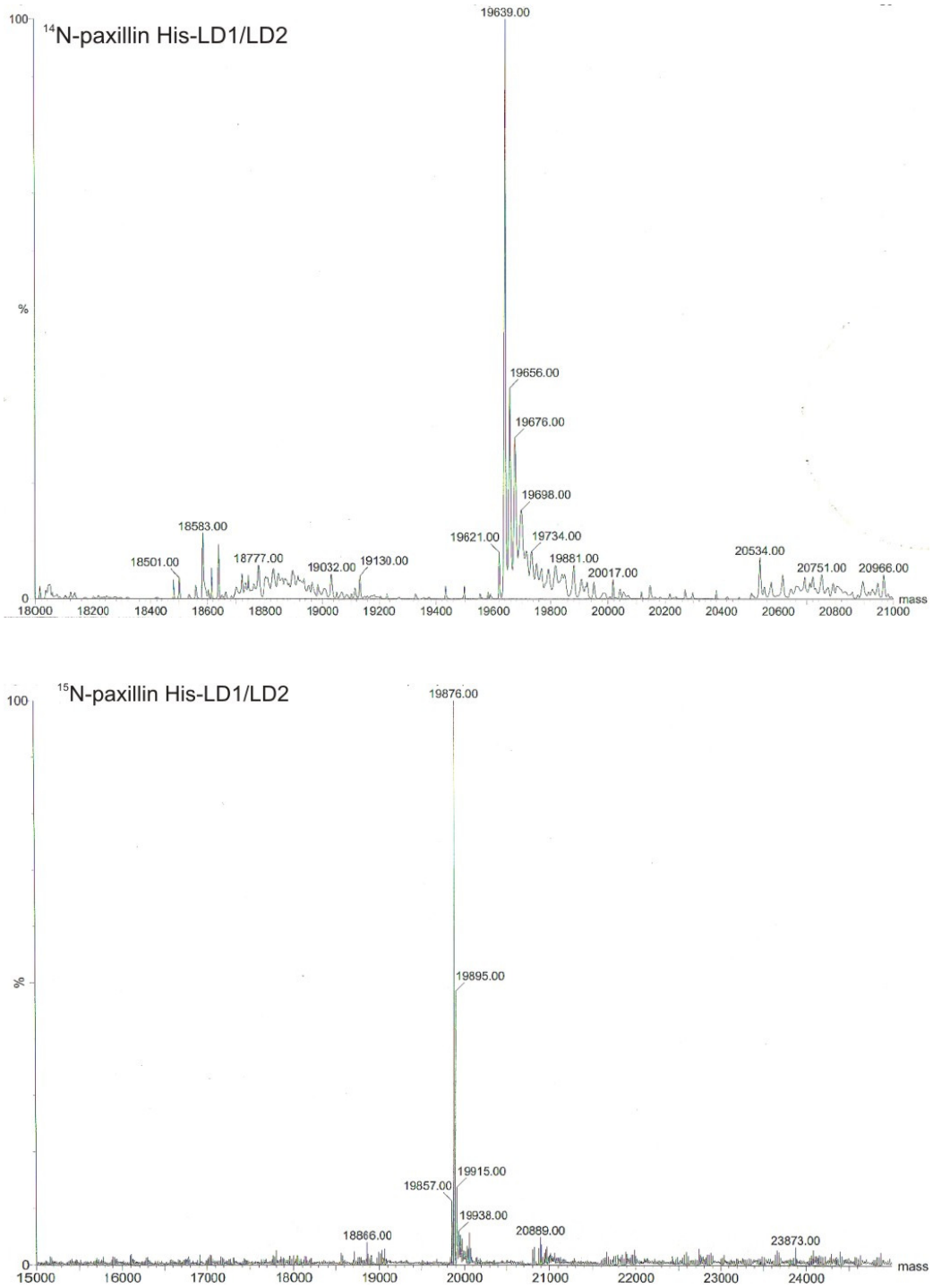


Figure 3.5.8 Electrospray ionization mass spectrometry of ^{14}N and ^{15}N His-LD1/LD2

Major peaks at 19639Da (^{14}N) and 19876Da (^{15}N) correspond to paxillin His-LD1/LD2 with ~130Da missing. ^{14}N: peaks at 19656 (+17), 19676 (+20), 19698 (+22) and 19734 (+36) are likely to be NH_4^+, Na^+ and Cl^- adducts respectively. ^{15}N: peaks at 19895 (+19), 19915 (+20) and 19938 (+23) are likely to be NH_4^+ and Na^+ adducts.

3.5.4 Analytical size exclusion chromatography

Size exclusion chromatography was carried out on paxillin His-LD1/LD2 to assess the oligomeric state of the protein. The results show a major peak at 42kDa and suggest that His-LD1/LD2 may exist as a dimer in solution.

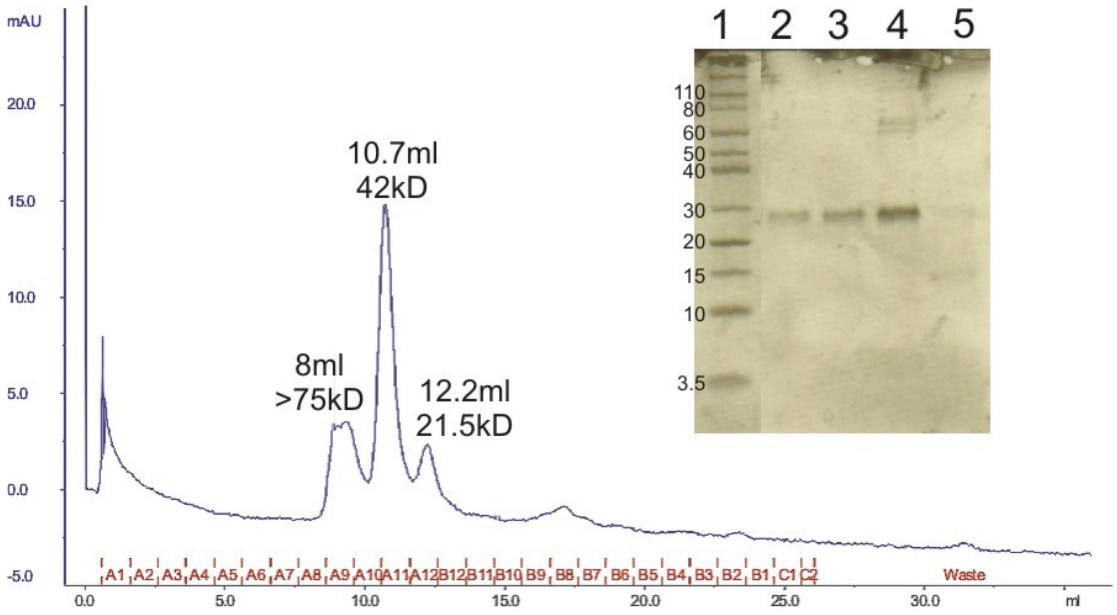


Figure 3.5.9 Analytical gel filtration (superdex-75) of paxillin His-LD1/LD2 in 20mM Tris pH7, 150mM NaCl and polyacrylamide gel of fractions.

The blue trace is the protein absorbance at 280nm. Lane 1: molecular weight marker showing weights in kDa; lane 2: protein peak from fraction A9 (8ml); lane 3: peak from fraction A10; lane 4: fraction A11 (10.7ml); lane 5: fraction A12 (12.2ml).

3.5.5 Secondary structure of paxillin His-LD1/LD2 using circular dichroism

The far-UV circular dichroism spectrum of paxillin His-LD1/LD2 in water is shown in Figure 3.5.10. The spectrum appears to lack the usual indicators of protein secondary structure, there is a weak negative at 230nm and there is a positive maxima at 200nm. For a random coil a net CD value of a strong negative just below 200nm, a positive band at

218nm and a weak negative band at 235nm is expected (211). While it is unclear exactly what the nature of the secondary structure is, there does appear to be elements of random coil.

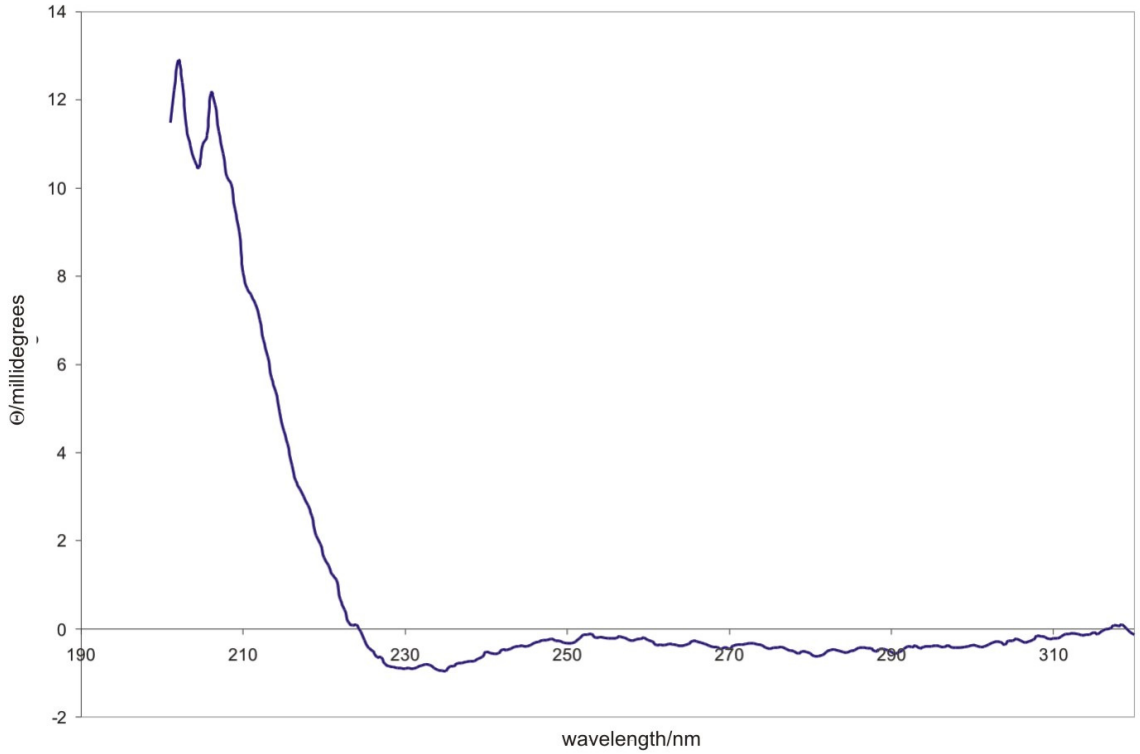


Figure 3.5.10 Far UV circular dichroism spectrum of paxillin His-LD1/LD2 at 0.5µM in water

3.5.6 ^{1}H-^{15}N-HSQC spectrum

A ^{1}H-^{15}N-HSQC spectrum for paxillin His-LD1/LD2 is shown in Figure 3.5.11. The observed NH resonances appear between 7.7-8.9ppm and are not well dispersed which is indicative of unstructured protein. There are 184 amino acids in His-LD1/LD2 of which 21 are prolines thus 163 resonances are expected. A peak count estimate shows that there are 159 resonances arising from NH groups. There is one tryptophan in the protein, the indole NH can clearly be seen at N129.5ppm, H^{N}10.1ppm, and there are 11 glycines, they may be the cluster of peaks between 113-109ppm. The 900MHz spectrum in Figure 3.5.12 shows

the same number of peaks but the overlap is reduced. It may be possible to assign some of this spectrum at 900MHz.

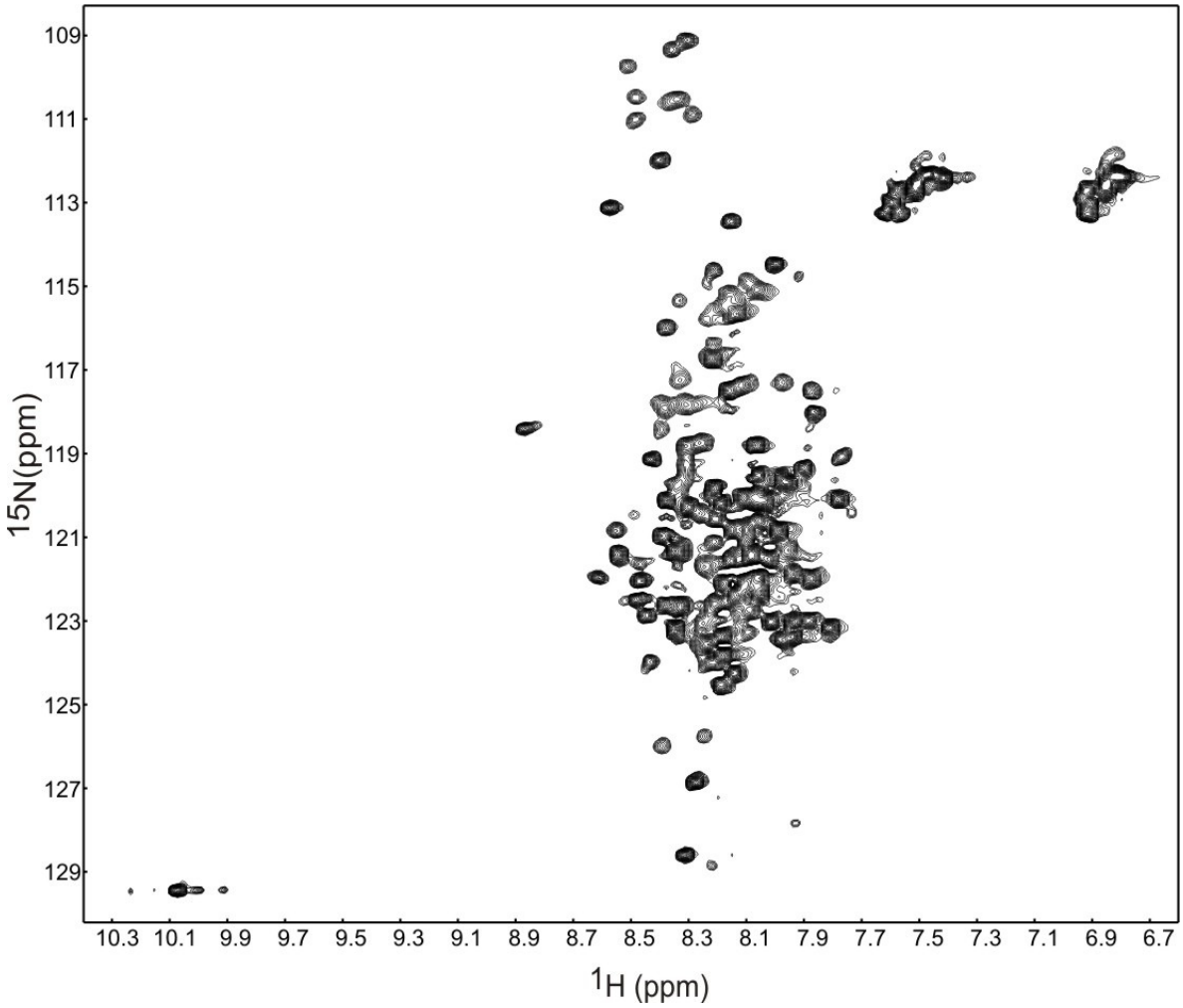


Figure 3.5.11 ^{1}H-^{15}N-HSQC spectrum of ^{15}N paxillin His-LD1/LD2 at 600MHz

Paxillin concentration 100µM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. Spectrum acquired for 2hrs at 600MHz.

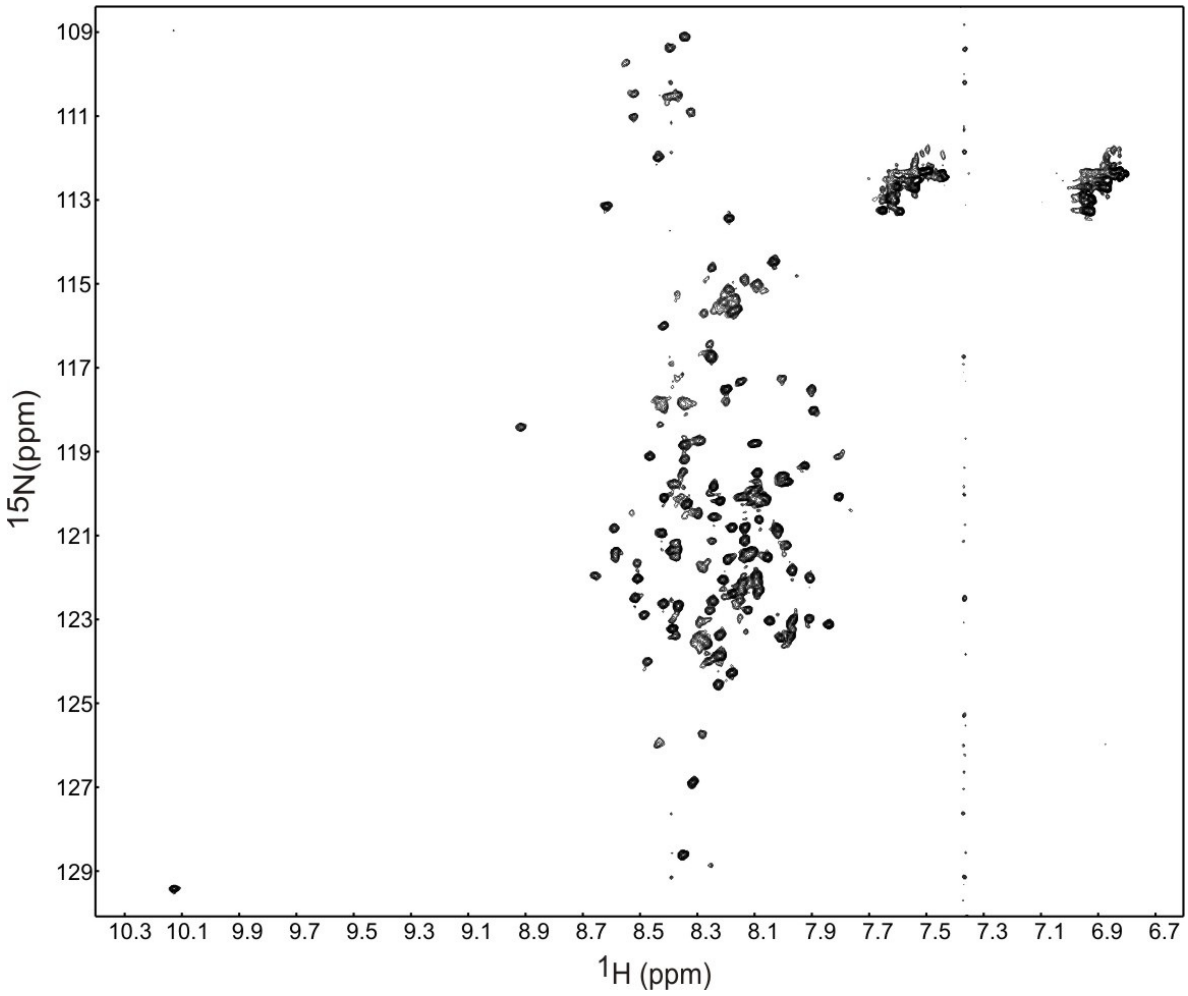


Figure 3.5.12 ^{1}H-^{15}N-HSQC spectrum of ^{15}N paxillin His-LD1/LD2 at 900MHz

Paxillin concentration 100μM in 20mM Tris pH7, 150mM NaCl, 50mM arginine, 50mM glutamate. Spectrum acquired for 1hr at 900MHz using a cryoprobe.

3.6 Vt- Paxillin Interaction Studies

The interaction between vinculin and paxillin was first identified in 1990 by Turner and colleagues (67). In this study the authors determined that paxillin interacted with the tail domain of vinculin. Further experiments (86) using truncated GST-vinculin polypeptides revealed that the region of Vt that is required to bind paxillin as Vt 978-1028.

In order to confirm that the Vt construct used in this study will interact with paxillin, two *in vitro* binding experiments were carried out, the GST-pull-down experiment and the catch-up comparative retention assay (268). These have been carried out using wild-type Vt and paxillin GST-LD1/LD2.

3.6.1 GST-pulldown experiments with Vt and paxillin

Initial attempts to observe an interaction between Vt and a synthetic paxillin peptide containing the LD1 motif did not show apparent chemical shifts of wild-type Vt in an NMR titration experiment (data not shown). Thus a GST-pulldown interaction assay was performed on larger paxillin constructs and wild-type Vt.

GST-pulldown experiments take advantage of the selective binding of gluthathione-S-transferase affinity tag to glutathione resins. A protein of interest which contains a GST-affinity tag is bound to a glutathione resin. Proteins can be passed over this resin and then washed off, any interacting proteins would remain bound to the GST-protein. This method only preserves interacting partners that survive multiple washing steps, so tends to select strong binding events where the exchange rate is slow or irreversible.

Figure 3.6.1 shows that when paxillin GST-LD1/LD2 and wild-type Vt are incubated together, they elute together from glutathione resin after washing steps to remove any other proteins. This can be seen from the extra band corresponding to Vt in lane 3 which is

missing from lane 2. The control experiment in Figure 3.6.2 show that Vt does not interact with GST, and Vt does not bind non-specifically to glutathione resin (data not shown).

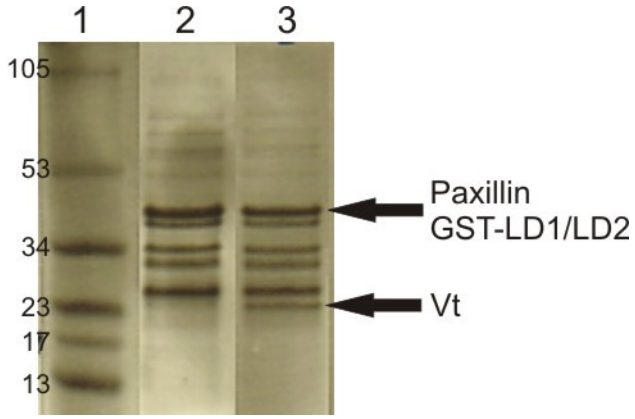


Figure 3.6.1 Composite polyacrylamide gel showing the results of the pulldown experiment.

Lane 1: molecular weight marker showing weights in kDa; lane 2: Paxillin GST-LD1/LD2 cell lysate bound, washed and eluted from glutathione beads not in presence of Vt; lane 3 Paxillin GST-LD1/LD2 cell lysate in presence of Vt that has been bound, washed and eluted from glutathione beads.

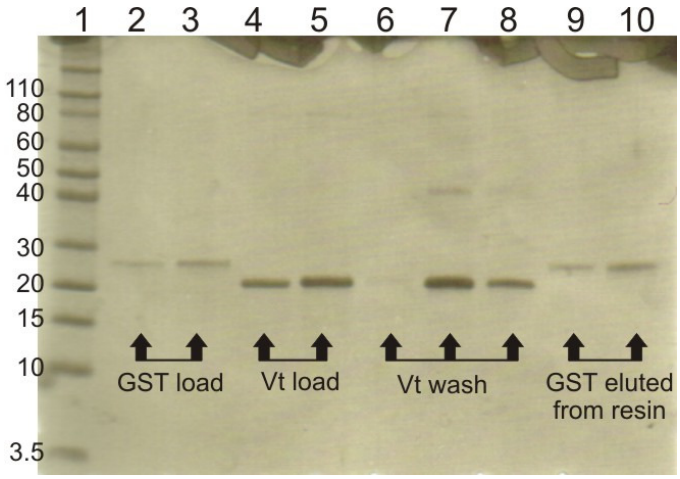


Figure 3.6.2 Polyacrylamide gel showing results of the GST control pulldown experiment.

Lane 1: molecular weight marker showing weights in kDa; lane 2 & 3: GST loaded onto the column, 5µl and 10µl loads; lane 4 & 5: Vt loaded onto the column, 5µl and 10µl loads; lanes 6, 7 & 8: washing unbound Vt from column; lane 9 & 10: GST eluted from the column-no Vt is present.

3.6.2 Catch-up comparative retention assay

The Catch-up Comparative Chromatographic Retention Assay is a new assay developed by Charbonnier *et al.* (268) and is based on the principles of reversible binding of fusion proteins to affinity resins. It is distinct from a GST-pull-down experiment as no washing steps are included and is therefore better suited to detect transient or fast exchanging interactions.

An example of results obtained from a catch-up experiment is shown in Figure 3.6.3. During the experiment, interaction partners are incubated together for 1 hour and then incubated with an affinity resin such as glutathione beads. One interaction partner contains an affinity tag, in this case GST. The mixtures are split in half and are treated with either a buffer containing 20mM reduced glutathione (the (+) condition) or a buffer containing no glutathione (the (-) condition). GST-containing proteins will be eluted from the resin when reduced glutathione is added but will remain on the resin when a buffer minus glutathione is used.

If the other interaction partner binds to the GST-containing protein, then both proteins will be observed on a reducing SDS-PAGE gel when the resin is treated with reduced glutathione. In the (-) glutathione condition, the GST-containing protein will remain on the resin and will not be seen on an SDS-PAGE gel. The interaction partner will be observed at a level that is determined by the strength of binding to the GST-protein and is judged by comparison to the (+) condition. No binding would be seen as a band on the (-) condition gel at the same intensity as on the (+) condition, the difference being that the GST-protein would not be observed (Figure 3.6.3(3). Strong binding would appear as no band on the gel compared to the (+) condition (Figure 3.6.3(1). Weak or transiently interacting proteins would appear at a level that was related to the binding equilibrium (Figure 3.6.3(1).

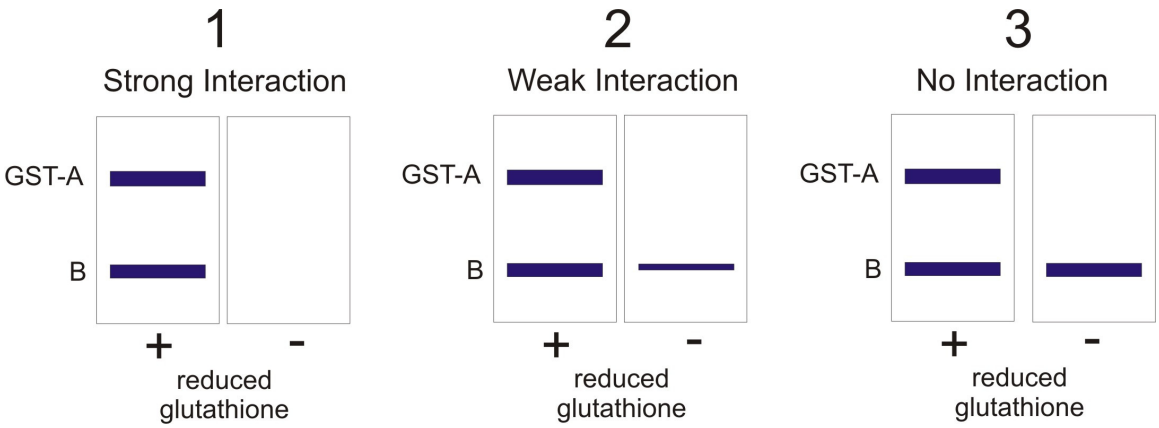


Figure 3.6.3. Example results from the catch-up assay.

Examples of a strong interaction (1), a weak interaction (2) and no interaction (3). Blue bands represent proteins on a polyacrylamide gel stained with coomassie blue.

This assay was carried out using the recombinant paxillin LD construct, GST-LD1/LD2. Figure 3.6.4 shows the results of the experiment. Lanes 1 and 2 show a control where a non-interacting partner, calexcitin, is added. Calexcitin does not interact with GST-LD1/LD2 as calexcitin can be seen at the same level in both lanes. Lanes 3 and 4 are controls to ensure that the assay is working as GST-LD1/LD2 can be seen in lane 3 but not in lane 4. Lane 5 and 6 show that Vt is interacting with GST-LD1/LD2 as Vt is present at a much lower level in lane 6 than in lane 5. The results are clarified in Figure 3.6.5. Control experiments to establish that the interaction was not GST-mediated were hampered by poor resolution on SDS-PAGE between Vt and GST, and experimental problems when using an unrelated GST-fusion protein. Observations suggested that the GST-fusion control protein was not binding to the resin making the results of the control inconclusive.

The catch-up experiment suggests an interaction between Vt and GST-LD1/LD2 that is of high affinity as very little Vt is in the liquid phase of the minus condition. Since the concentration of the GST-LD1/LD2 is unknown the strength of the interaction cannot be quantified.

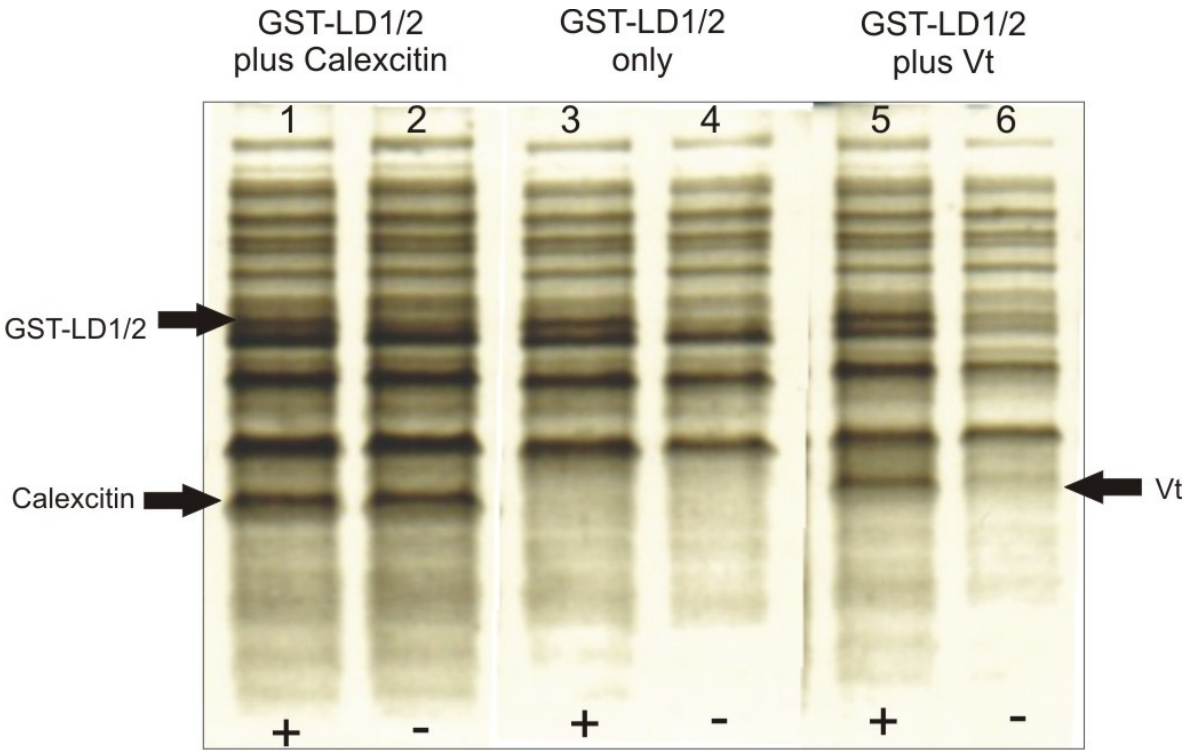


Figure 3.6.4 Composite polyacrylamide SDS-PAGE gel showing the results of the catch-up assay between GST-LD1/LD2 and wild-type Vt.

Three sets of experiments are shown. The plus (+) condition contains 20mM reduced glutathione for elution, the minus (-) condition does not contain 20mM reduced glutathione. Lanes 1&2: GST-LD1/LD2 plus calexcitin (control), both cell lysates. Lanes 3&4: GST-LD1/LD22 only (control). Lanes 5&6: GST-LD1/LD2 cell lysate plus purified Vt.

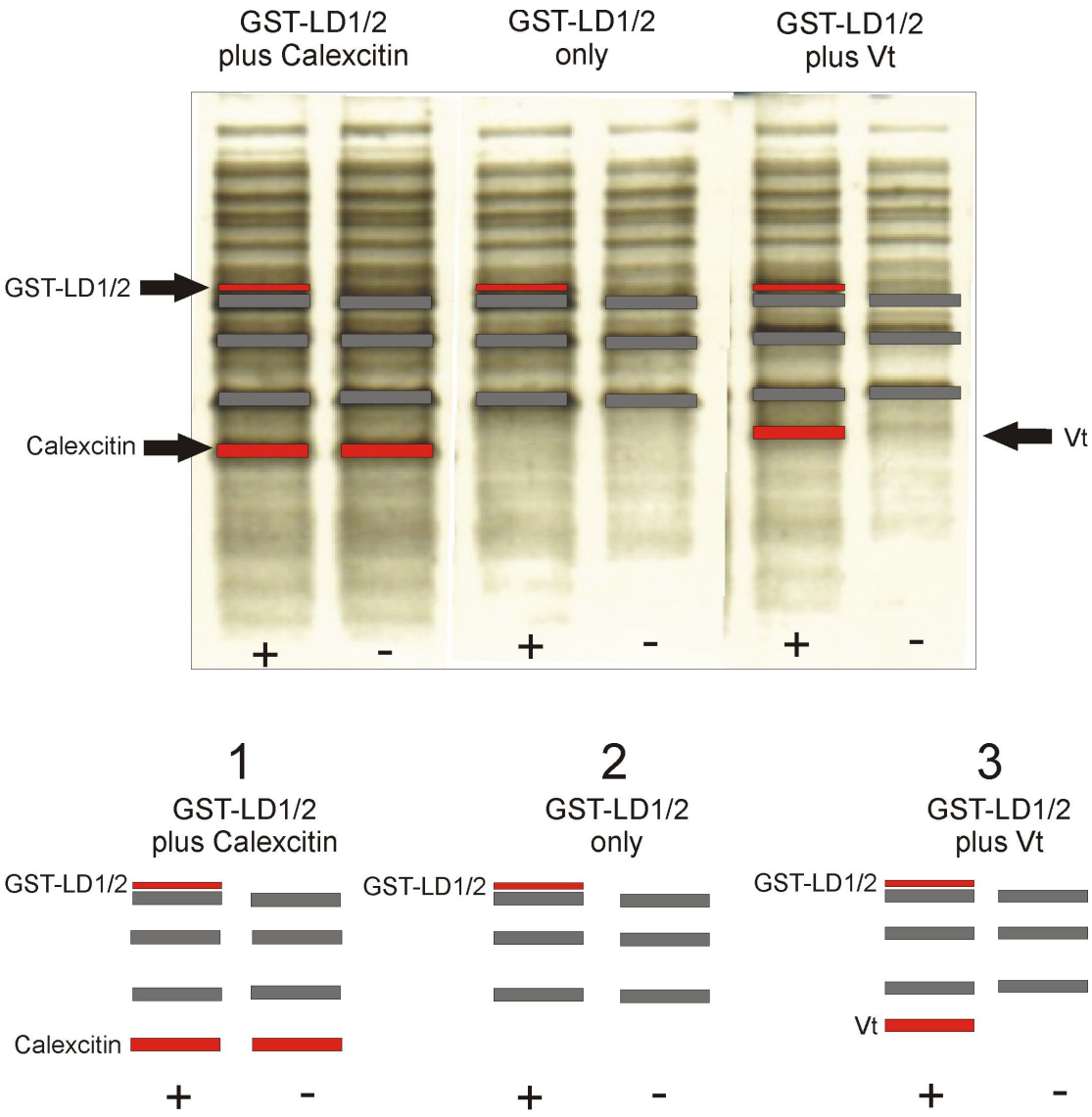


Figure 3.6.5 Results from Figure 3.6.4 with gel patterns clarified

Blocks shown are representative of protein bands on a gel. Red blocks are the bands of interest, grey blocks are placed there for reference only and come from the results gel. The plus condition contains 20mM reduced glutathione for elution, the minus condition does not contain 20mM reduced glutathione.

These results suggest an interaction between GST-LD1/LD2 and wild-type Vt, and agrees with the interaction previously observed (5;67;80;86). Catch-up experiments between Vt and GST-LD3/LD4 and GST-LD4/LD5 have been hampered by poor expression of the LD constructs.

3.6.3 Interaction studies of Vt and paxillin by NMR

An important aim of this thesis is to investigate the nature of the interaction between Vt and paxillin. Elucidating the location of the interaction site for paxillin on Vt, the specificity of LD motifs for Vt and strength of the interaction is important to enable the understanding of how the two molecules interact within FAs.

Solution NMR is a powerful tool for evaluating protein-ligand interactions. It relies on the chemical shifts of the residues in the protein being sensitive to subtle changes in chemical environment (253). It is especially useful as it can be used to detect weak interactions, with K_ds in the millimolar range being detectable (254).

The most widely used NMR method to determine binding interfaces is chemical shift perturbation (254). This can be achieved by recording a ^{1}H-^{15}N-correlation spectrum of an isotopically labelled protein, and then monitoring the changes in the chemical shifts as a binding partner is titrated in.

3.6.3.1 Chemical shift perturbation experiments

Titration series were carried out using wild-type Vt at 200µM protein concentration. Table 2.6.2 lists the titration experiments that have been carried out in this study. At 600MHz, ^{1}H-^{15}N-HSQC spectra were recorded for each titration point. At 900MHz, TROSY spectra were recorded. Ligands used for the titration were either synthesised paxillin 13mer peptides LD1, LD2 or LD4; or a longer purified paxillin construct His-LD1/LD2. The sequences of the peptides can be found in Chapter 2.6.7.1, and the sequence of paxillin His-LD1/LD2 can be found in Appendix 5.4.

Overall, chemical shift perturbations were small and did not produced a simple, discrete set of affected residues that formed a continuous interaction surface. However, intensity changes as a function of ligand concentration were notable, in some cases leading to

disappearance of resonances. As this may be indicative of intermediate exchange (Section 2.5.13.1), intensity changes were included in the analysis. Accordingly the analysis of the data has been approached in the following manner. For each assigned residue, three parameters were analysed and the results represented as block diagrams as a function of sequence. The first is the intensity change between the first and the last titration point (corresponding to equivalent molar ratios for all peptides except for the His-LD1/DL2 construct). This has been measured as a percentage intensity decrease, so the results are the percentage of the signal that remains at a 10-fold excess of ligand (4-fold for His-LD1/LD2). This manifests on a graph as a positive value for intensity decreases and a negative value for intensity increases. The more positive the bar then the more signal has been lost. The second and third parameters are the chemical shift differences between the first and last titration points in the proton and nitrogen dimension respectively.

3.6.3.2 Interaction of wild-type Vt with LD1, LD2, LD4 and His-LD1/LD2

All of the data analysis in this section was carried out using the titration data measured in 20mM Tris pH7, 50mM NaCl with 50mM arginine, 50mM glutamate buffer. Great care was taken to maintain the pH and buffer composition and the concentration of Vt throughout the titrations (see methods, Section 2.6.7.1). Criteria were set to distinguish whether a perturbation was a result of direct interaction with a ligand. Percentage intensity decreases were set to ≥50%, chemical shift differences in the proton and nitrogen dimension were set as 0.03ppm and 0.2ppm respectively. This criteria was judged to be acceptable, a more stringent cut-off of 60% intensity decrease and higher chemical shift differences resulted in most effects being lost. The following descriptions are summarised in Table 3.6.1.

LD1 plus wild-type Vt (Figure 3.6.6)

When a 10-fold excess of LD1 peptide is added to wild-type Vt there is a global intensity drop for most chemical shifts that are assigned along the backbone. There does not appear to be any distinct regions where the intensity loss is more localised. Of the assigned

residues, 64 have an intensity decrease larger than 50%. The average intensity decrease across all assigned chemical shifts is 44.0%

The largest proton chemical shift changes are contained in helices 1, 3 and 5. There are 15 residues that have chemical shift changes of greater than 0.03ppm in the proton dimension. The largest chemical shift changes in the nitrogen dimension are contained in helices 3, 4 and 5, and there are some changes in the linker region between helices 4 and 5. There are 10 residues that shift greater than 0.2ppm. Combining both proton and nitrogen chemical shift changes gives a total of 21 residues affected.

Figure 3.6.7 shows these affected residues mapped onto the crystal structure of Vt. The intensity decreases are spread throughout the molecule, although helix 3 seems to be affected more than other helices. The chemical shift changes appear to be less global, with face 3-4 and 4-5 being affected more than the other two faces. Helix 3 appears to be most affected.

LD2 plus wild-type Vt (Figure 3.6.8)

When a 10-fold excess of LD2 peptide is titrated into wild-type Vt there is a global intensity drop, the average decrease across all of the assigned chemical shifts is 45.7%. There are 54 residues that have an intensity decrease of greater than 50%.

The largest proton chemical shift changes are contained in helices 1 and 5, with some large shifts in helices 2 and 3 also. There are 11 chemical shifts that titrate more than 0.03ppm in the proton dimension. The largest chemical shift changes in the nitrogen dimension are in helices 1, 3, 4 and 5. There are 7 that are greater than 0.2ppm. Combining both proton and nitrogen shifts gives a total of 14 residues affected.

Figure 3.6.9 shows these affected residues mapped onto the crystal structure of Vt. The intensity decreases are spread throughout the molecule, with clusters of affected residues on helix 3, face 4-5 and face 5-2. The chemical shift changes appear to be less global, with

most of the shifts found near the top of the helical bundle. Faces 3-4, 4-5 and 5-2 have small clusters of affected residues.

LD4 plus wild-type Vt (Figure 3.6.10)

When a 10-fold excess of LD4 is titrated into Vt, the average chemical shift intensity decrease is 22.2%. There are only two chemical shifts that decrease in intensity more than 50% (C950 and M1005), and there is one residue that shifts in the proton dimension greater than 0.03ppm (A1044). No chemical shift changes are seen in the nitrogen dimension greater than and 0.2ppm.

Figure 3.6.11 shows these affected residues mapped onto the crystal structure of Vt. There are only three residues affected and they are on helix 3, on the linker between helix 4 and 5, and at the end of helix 5. There are no global affects or discrete regions that are being affected.

His-LD1/LD2 plus wild-type Vt (Figure 3.6.12)

For these experiments a 4-fold excess of paxillin His-LD1/LD2 was added to wild-type Vt, this excess causes a global reduction in intensity and an average intensity decrease for all assigned chemical shifts of 51.9% There are 85 chemical shifts that decrease in intensity more than 50%.

There are three chemical shifts that change in the proton dimension greater than 0.03ppm and only one that shifts in the nitrogen dimension greater than 0.2ppm. These are in helices 1 and 5.

Figure 3.6.13 shows these affected residues mapped onto the crystal structure of Vt. The intensity decreases are spread throughout the molecule, with affected residues clustering on all faces with helix 3 being affected the most. The chemical shift changes are found on helix 1 and 5 but do not seem to be clustered.

Table 3.6.1 Summary table of Vt perturbation upon LD titration

	Molar excess of ligand	Average decrease (%)	No. residues that decrease ≥ 50%	ΔH ≥0.03	ΔN ≥0.2	Total residues affected ΔH and ΔN	No. of residues affected in total
LD1	10	44.0	64	15	10	23	70
LD2	10	45.7	54	11	7	14	56
LD4	10	22.2	2	1	0	1	3
His-LD1/LD2	4	51.9	85	3	1	4	86

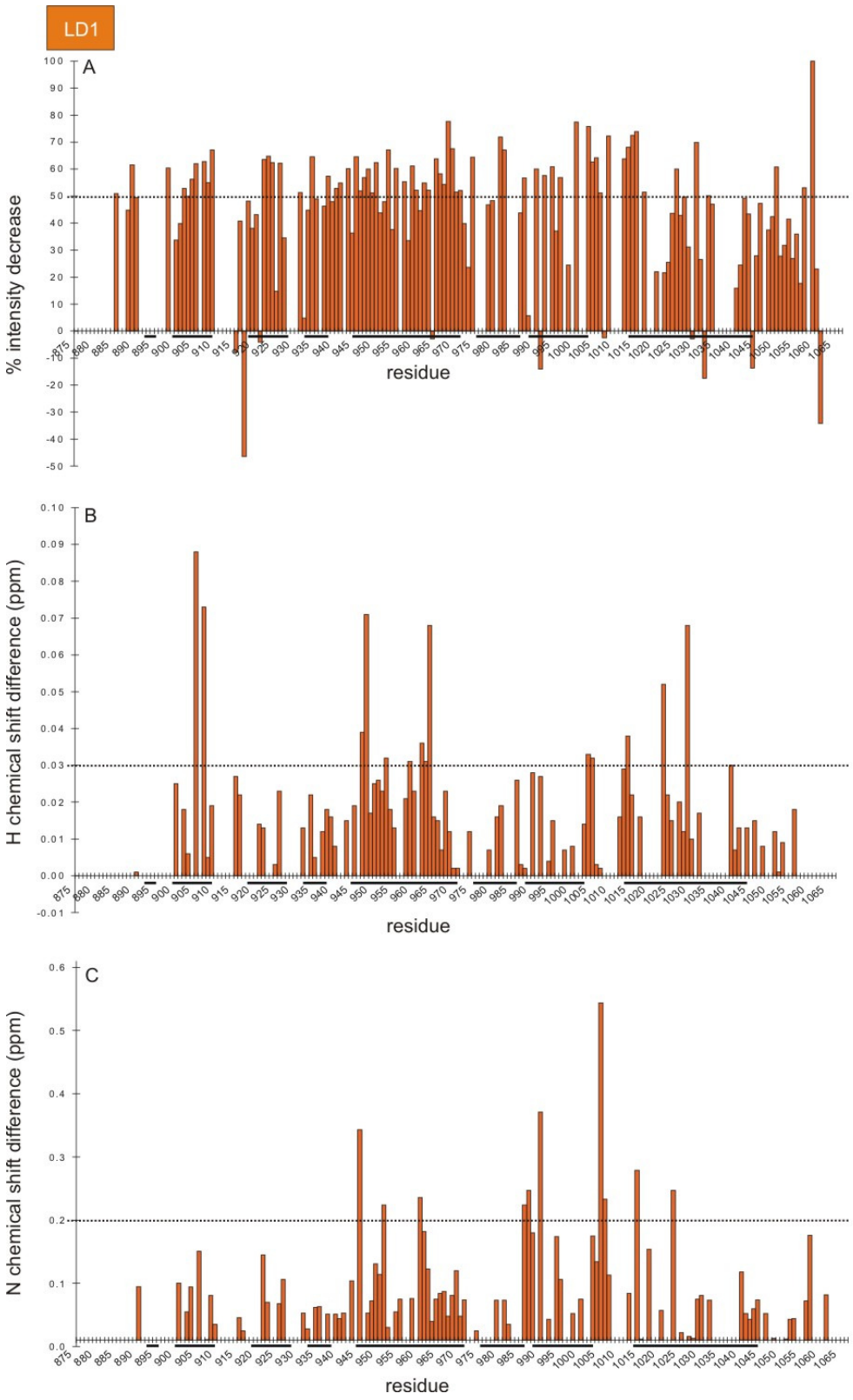


Figure 3.6.6 Wt-Vt plus LD1. Bar charts showing percentage intensity decrease (A), proton chemical shift difference (B) and nitrogen chemical shift difference (C).

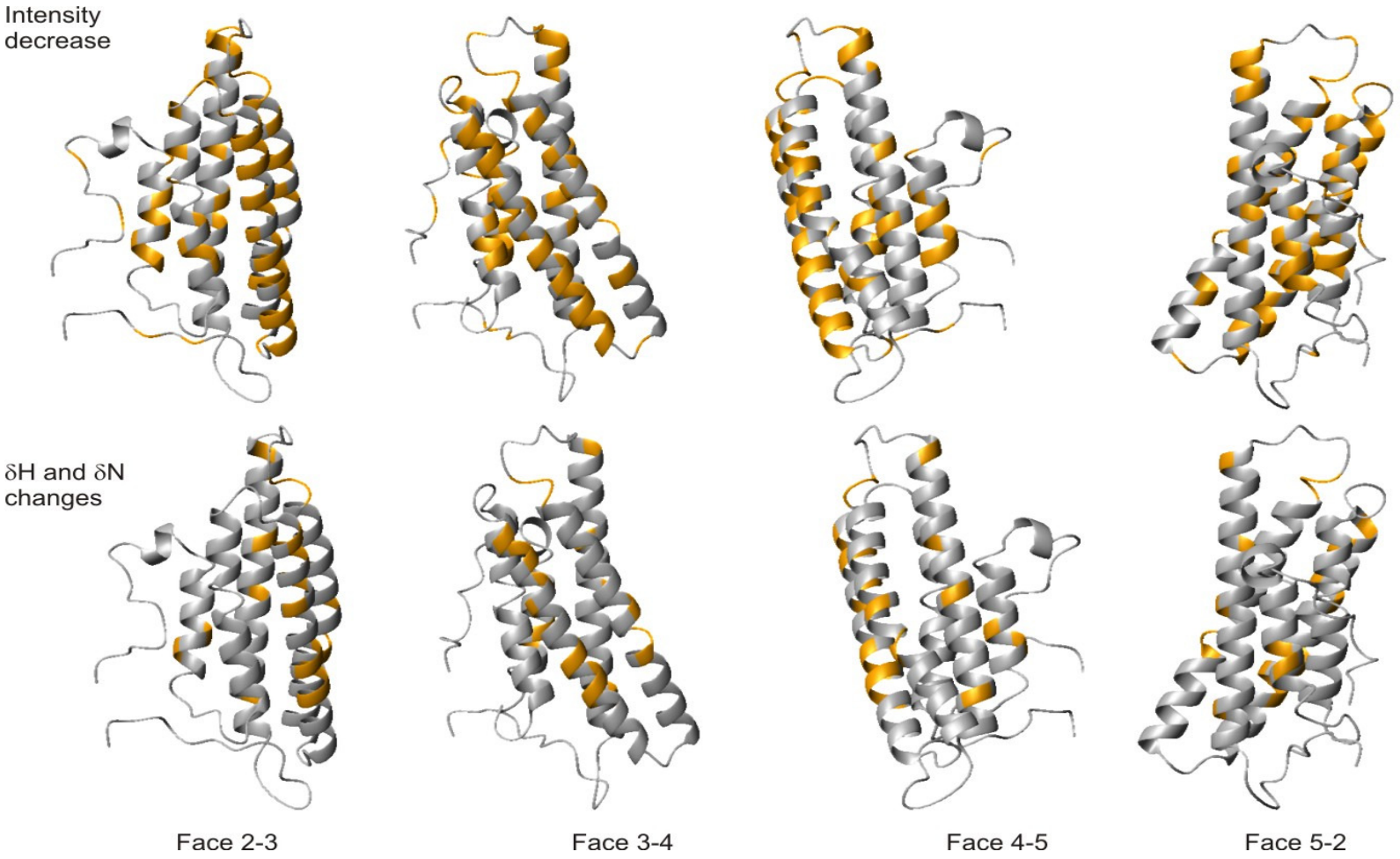


Figure 3.6.7 Chemical shift perturbations caused by LD1 peptide mapped onto the crystal structure of Wt-Vt. Perturbed NH resonances shown in orange, unperturbed NH resonances shown in grey. Criteria: intensity decrease ≥ 50% or ΔH ≥0.03ppm or ΔN ≥0.2ppm.

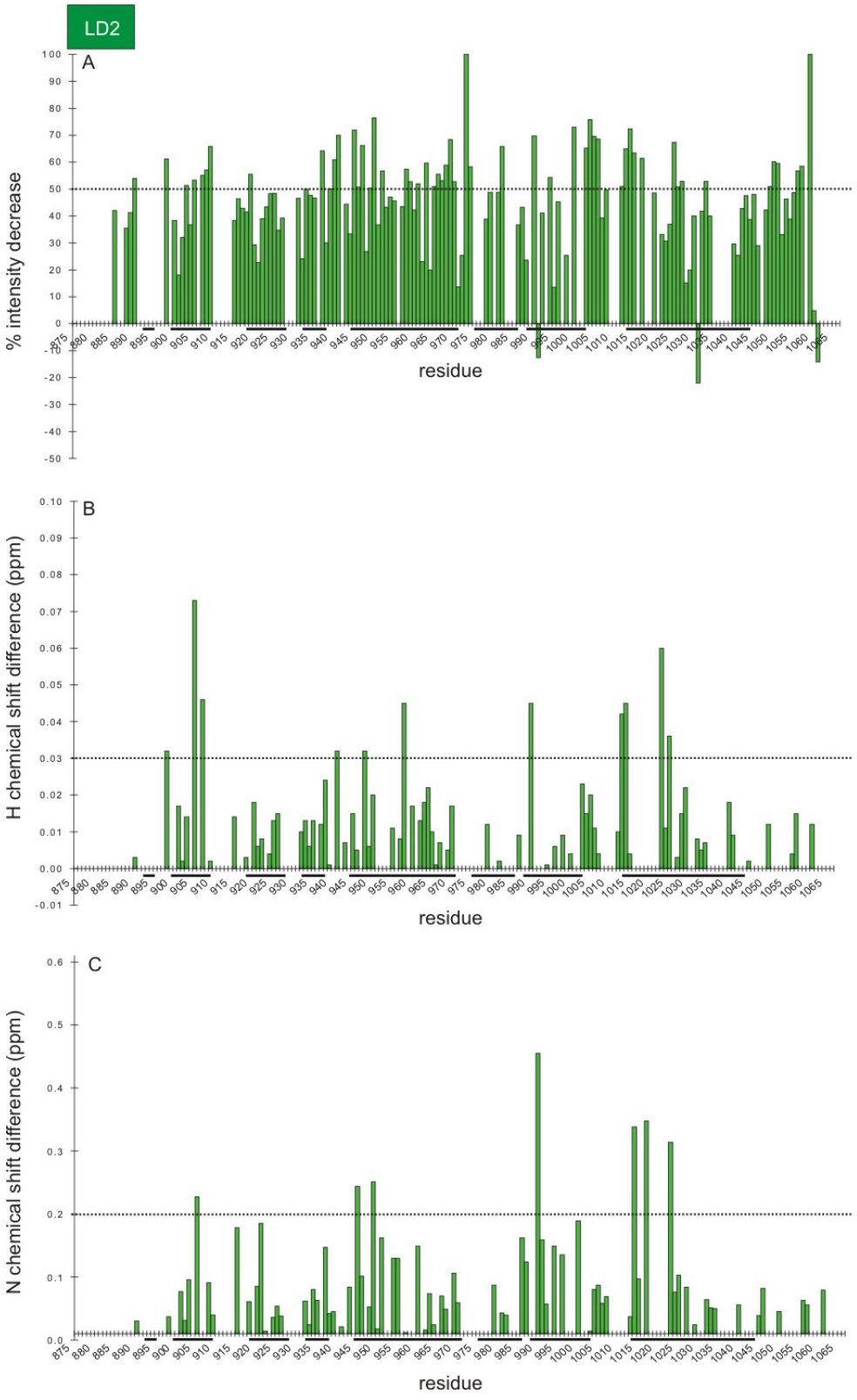


Figure 3.6.8 Wt-Vt plus LD2. Bar charts showing percentage intensity decrease (A), proton chemical shift difference (B) and nitrogen chemical shift difference (C).

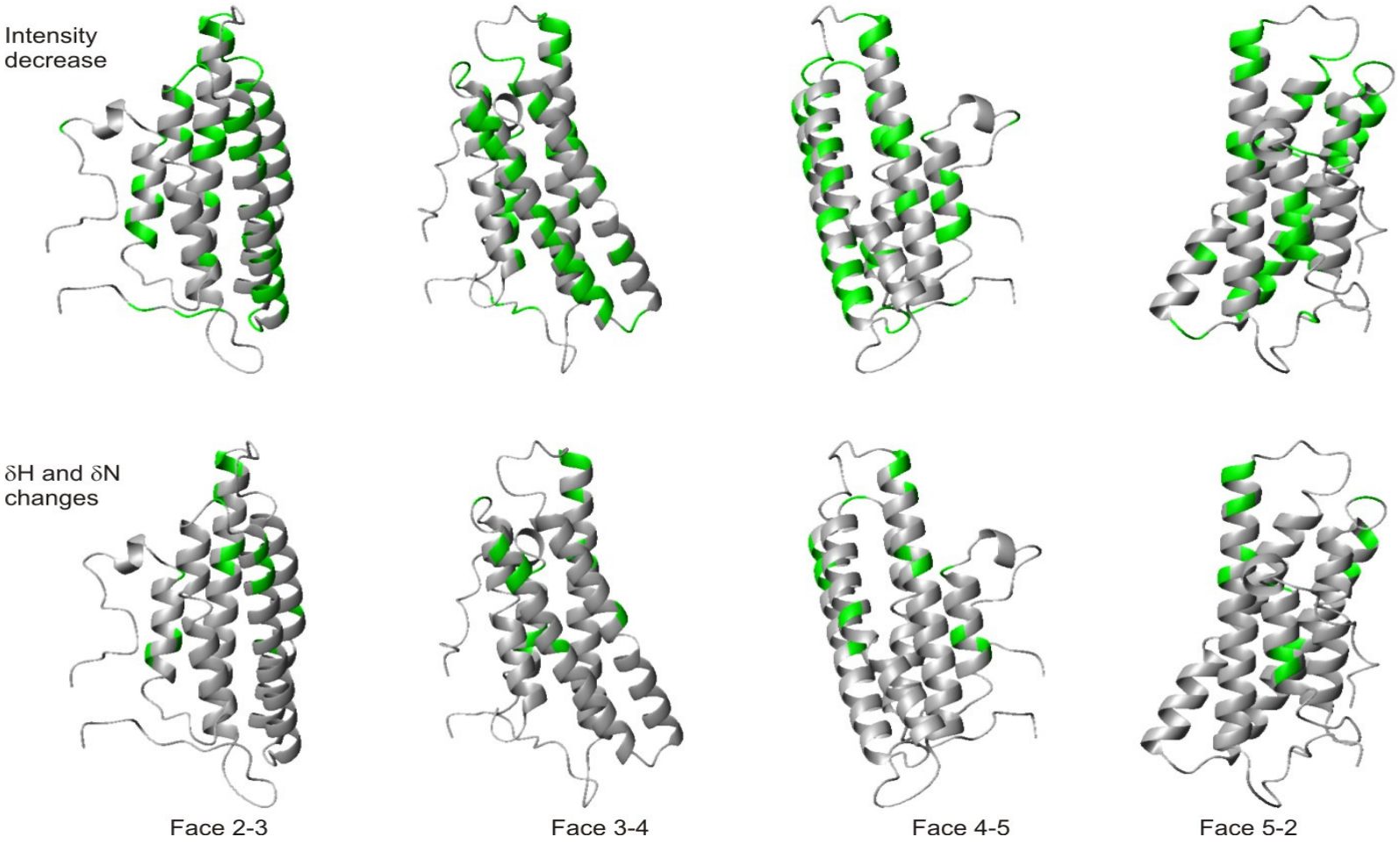


Figure 3.6.9 Chemical shift perturbations caused by LD2 peptide mapped onto the crystal structure of Wt-Vt. Perturbed NH resonances shown in green, unperturbed NH resonances shown in grey. Criteria: intensity decrease ≥ 50% or ΔH ≥0.03ppm or ΔN ≥0.2ppm.

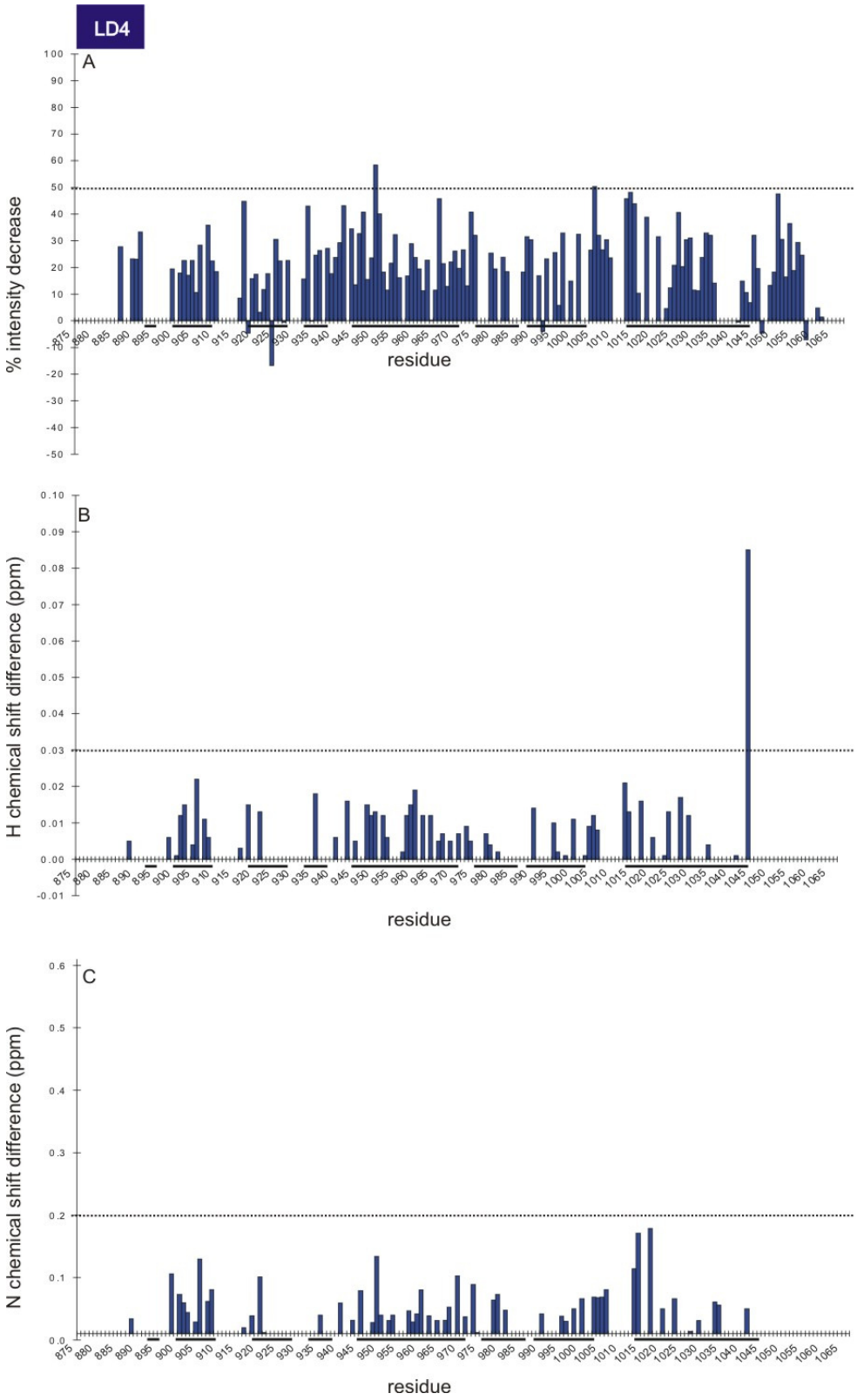


Figure 3.6.10 Wt-Vt plus LD4. Bar charts showing percentage intensity decrease (A), proton chemical shift difference (B) and nitrogen chemical shift difference (C).

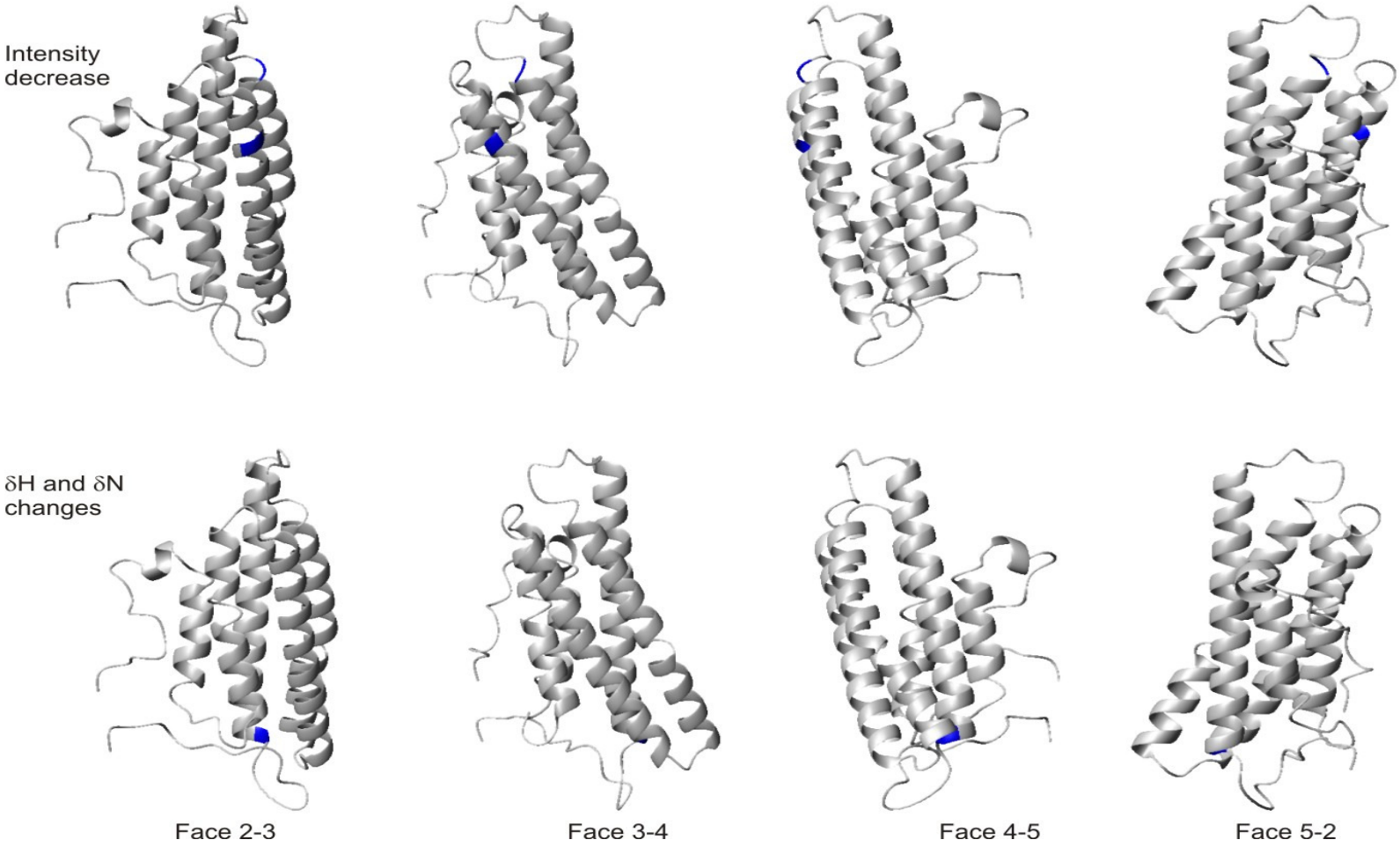


Figure 3.6.11 Chemical shift perturbations caused by LD4 peptide mapped onto the crystal structure of Wt-Vt. Perturbed NH resonances shown in blue, unperturbed NH resonances shown in grey. Criteria: intensity decrease ≥ 50% or ΔH ≥0.03ppm or ΔN ≥0.2ppm.

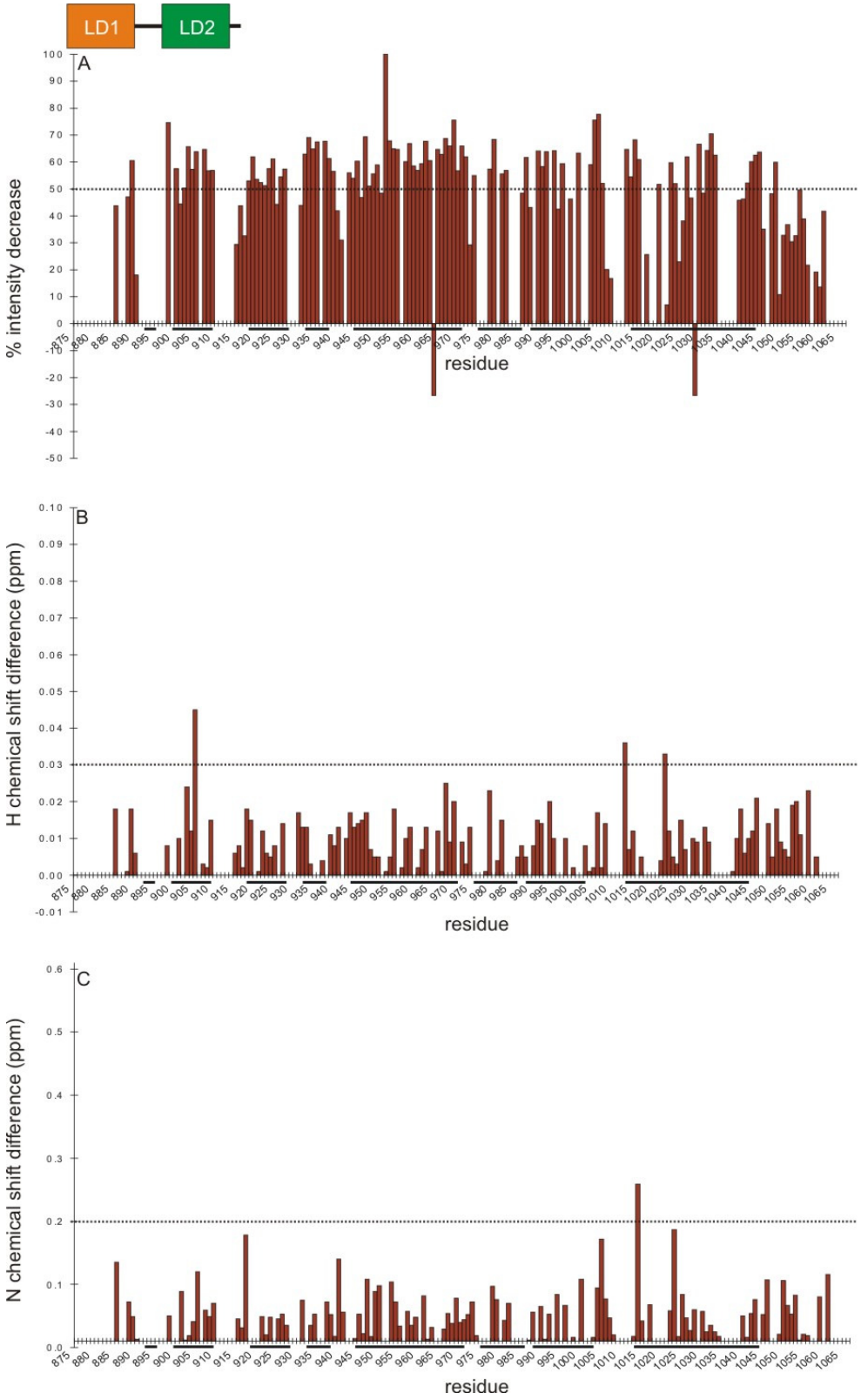


Figure 3.6.12 Wt-Vt plus His-LD1/LD2. Bar charts showing percentage intensity decrease (A), proton chemical shift difference (B) and nitrogen chemical shift difference (C).

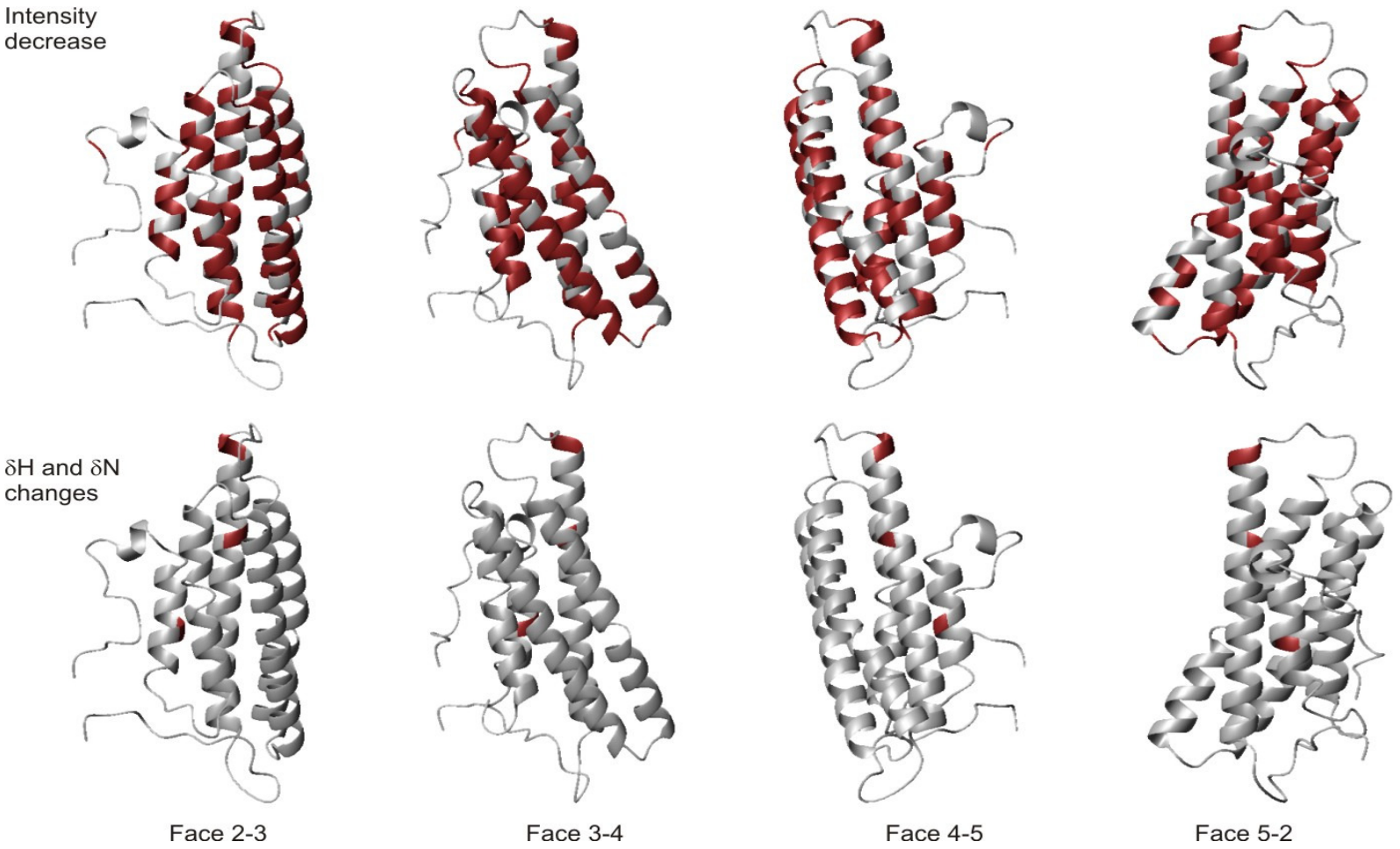


Figure 3.6.13 Chemical shift perturbations caused by His-LD1/LD2 mapped onto the crystal structure of Wt-Vt. Perturbed NH resonances shown in brown, unperturbed NH resonances shown in grey. Criteria: intensity decrease ≥ 50% or ΔH ≥0.03ppm or ΔN ≥0.2ppm.

3.6.3.3 Comparison of LD1, LD2, LD4 and His-LD1/LD2 data

A list of all the perturbed residues for LD1, LD2 and His-LD1/LD2 that meet the above criteria with an 'or' condition can be found in Appendix 5.7. All of the titration experiments at 600MHz were carried out under identical conditions so can be directly compared to one another, the His-LD1/LD2 titration carried out at 900MHz may be affected by different exchange regimes but is still comparable as the buffer conditions were the same.

The global reduction in intensity for the chemical shifts when titrating with LD1 and LD2 suggests that the system is experiencing intermediate exchange processes. Figure 3.6.14A and B shows all Vt residues that are perturbed when LD1 and LD2 are titrated in. As the figure shows, many of these are the same for LD1 peptide as for LD2. In contrast, the LD4 peptide only perturbs 3 residues (see Figure 3.6.11). The average global reduction in intensity is very similar for LD1 and LD2 (44.0% and 45.7%), while the average intensity decrease for the LD4 peptide is half, only 22.2%. LD4 is not perturbing residues in Vt in the same manner as the LD1 and LD2 peptides.

There are differences between the perturbations of the chemical shifts when titrating with LD1 and LD2. The list in Appendix 5.7 highlights the residues that are affected specifically by LD1 compared to LD2 and vice versa.

For LD1, the perturbed shifts that are titrating in the proton and nitrogen dimension and broadening are being affected on a fast-to-intermediate timescale. There are more perturbations affected in this manner with LD1 than for LD2. In the LD2 titration, there are residues in fast-to-intermediate exchange, but some of the perturbations seem to remain in the same position in the spectrum but reduce in intensity. This situation is ambiguous as it could suggest intermediate exchange where the free and bound complex shifts coincide, or in slow exchange. To identify slow exchange the chemical shifts of the saturated complex would need to be assigned.

Examples of the different types of perturbations can be seen in Figure 3.6.15. Threonine T990 is undergoing a fast-to-intermediate exchange process for both LD1 and LD2. The NH resonance shifts in the proton and nitrogen dimension whilst the intensity is simultaneously decreasing (Figure 3.6.15a & b). In contrast, the same resonance with LD4 does not move or decrease in intensity (c). Arginine R963 and glutamate E960 are examples of fast-to-intermediate exchange with LD1 peptide (d, g) and slow or intermediate exchange for LD2 (e, h). In both cases the resonances do track in the proton and nitrogen dimensions with LD2, but this is below the threshold for an affected peak, the resonance is perturbed dominantly by an intensity drop. In both cases LD4 peptide is not perturbing the resonance.

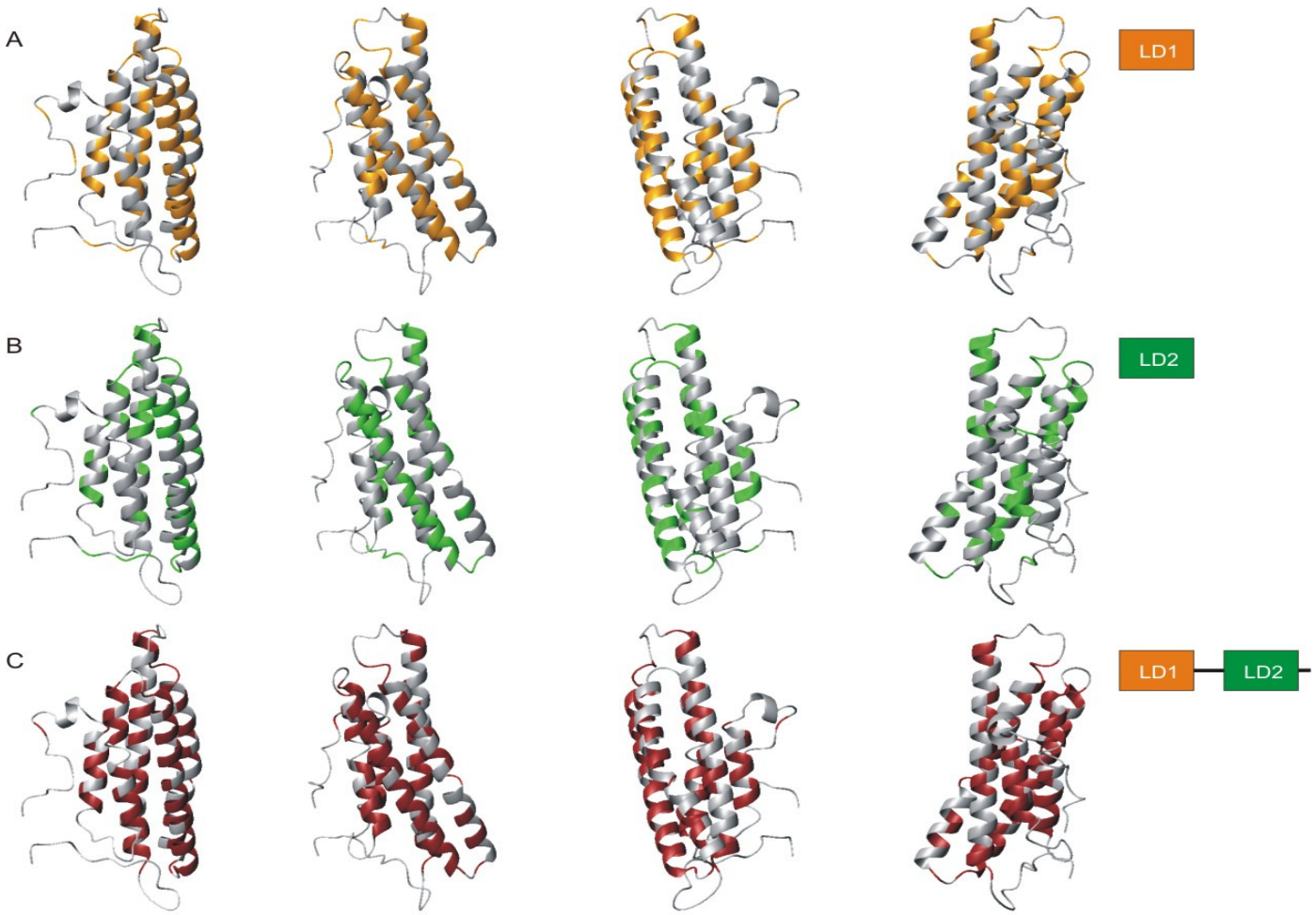


Figure 3.6.14 Residues that are affected by both slow and fast exchange in wild-type Vt.

A: when LD1 is added to a 10-fold excess, B: when LD2 is added to at 10-fold excess, C: when His-LD1/LD2 is added to a 4-fold excess.

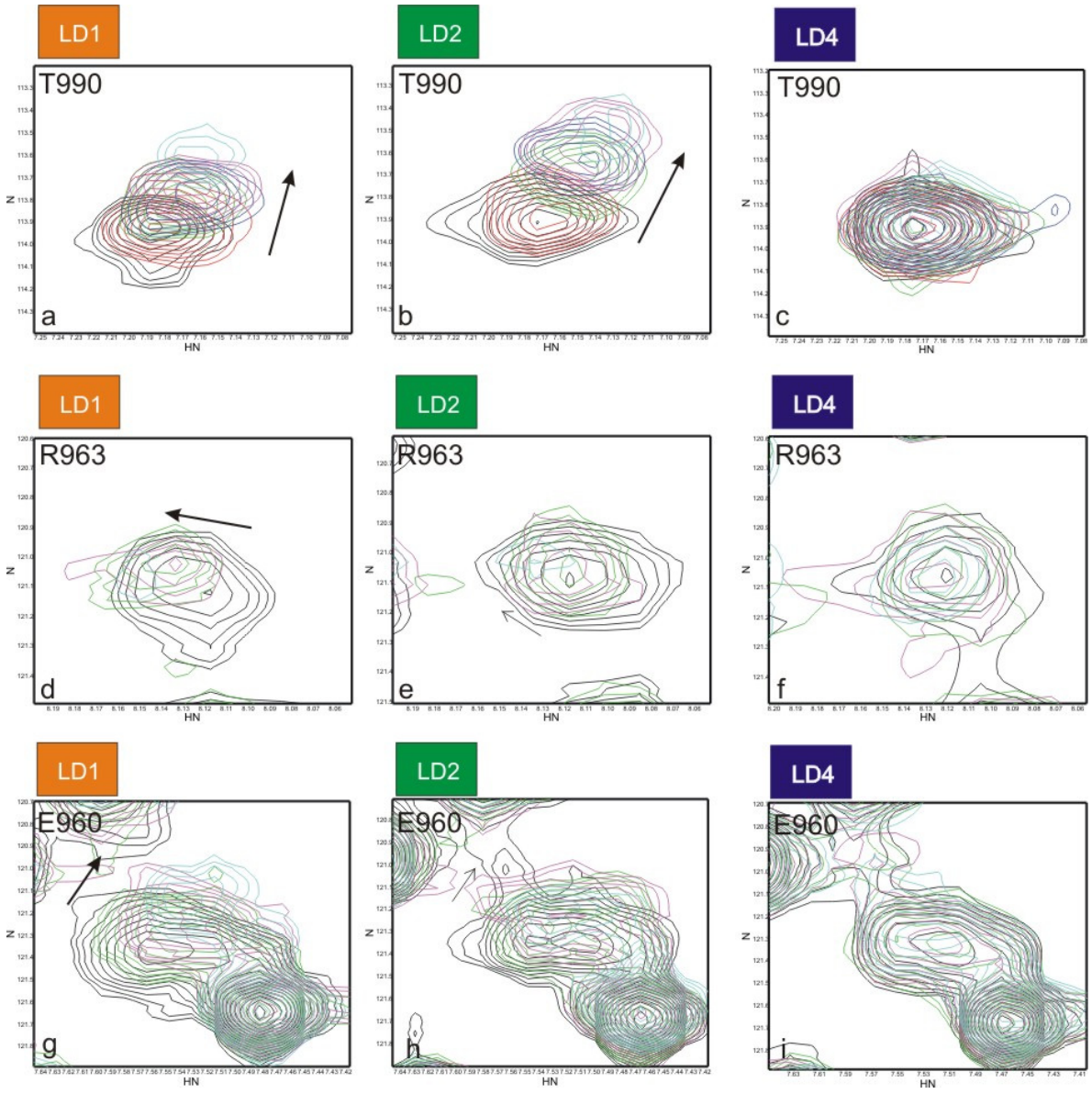


Figure 3.6.15. Examples of perturbed NH resonances when LD peptides are titrated into wild-type Vt.

Each picture is an overlay of several spectra with increasing molar ratios of peptide added. Molar ratio of peptide: 0 (black), 0.25 (red), 0.75 (green), 1.0 (blue), 2.0 (magenta), 10.0 (cyan). Arrows show direction and distance of fast exchanging peaks, bold arrows represent shifts greater than the criteria, unbold arrows show shifts that are below the criteria.

The His-LD1/LD2 titrations were carried out under identical conditions at 900MHz. Perturbed residues are also being affected by intermediate exchange processes during this titration. As Figure 3.6.14C shows, many of the resonances perturbed by LD1 and LD2 are also perturbed by His-LD1/LD2. However there are more resonances perturbed by His-LD1/LD2 (86 residues) compared to LD1 (70) and LD2 (56). Of the perturbed resonances, there are 20 that are specific to His-LD1/LD2 compared to the LD1 and LD2 independently. There are very few resonances that are exchanging on a fast-to-intermediate timescale with His-LD1/LD2, most are in an intermediate or slow exchange regime.

More chemical shifts with His-LD1/LD2 are in intermediate or slow exchange in the His-LD1/LD2 titration compared to the individual peptides. Figure 3.6.16 shows how the NH resonances from threonine T990, arginine R963 and glutamate E960 are being affected on a slower timescale (as suggested by the intensity decreases) compared to the individual LD1 and LD2 peptides

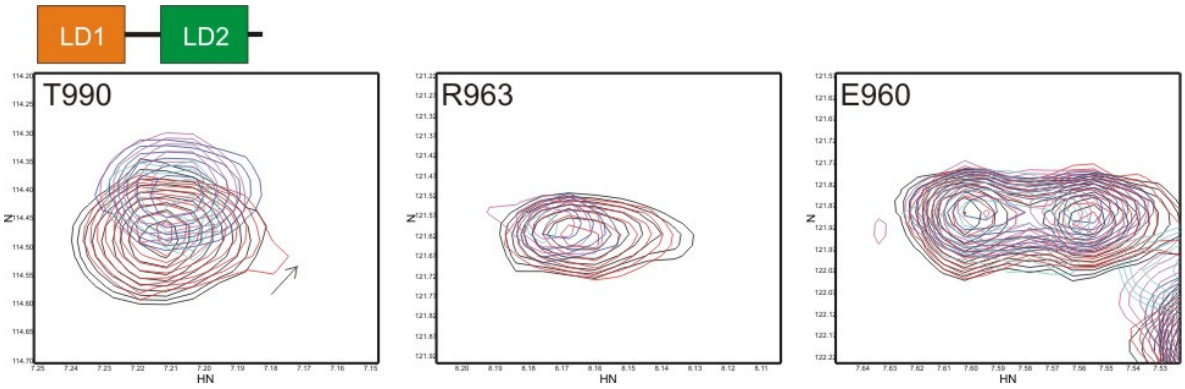


Figure 3.6.16 Examples of resonances that are affected on a slower timescale when His-LD1/LD2 is added to wild-type Vt compared to the individual LD1 and LD2 peptides.

Each picture is an overlay of several spectra with increasing molar ratios of His-LD1/LD2 added. Molar ratios: 0 (black), 0.25 (red), 1.0 (blue), 2.0 (magenta), 4.0 (cyan). Arrows show direction and distance of fast exchanging peaks, bold arrows represent shifts greater than the criteria, unbold arrows show shifts that are below the criteria.

3.6.3.4 Effect of salt on the interaction of Vt with LD peptides

The titrations with Vt and LD peptides have been conducted at various salt concentrations. Initially, 400mM NaCl was used as it was known that Vt has a stability problem at lower NaCl concentrations. In this initial study it was difficult to observe many fast exchanging resonances and there were global intensity drops and degradation of ^{1}H-^{15}N-HSQC spectral quality. The salt was subsequently dropped to 150mM which improved the quality of spectra obtained and also appeared to show more fast exchanging peaks. Finally it was dropped to 50mM with the addition of 50mM arginine and 50mM glutamate which has ultimately produced the best quality data.

In order to check that the salt concentration was not affecting the perturbations in a different manner between experiments, the LD1 titrations at different salt concentrations were subject to the same analysis as above. For NaCl concentrations of 400mM, 150mM and 50mM, the average intensity decrease across all assigned residues is 48.8%, 45.1% and 44.0% respectively.

Figure 3.6.17 presents chemical shift data for residues in fast-to-intermediate exchange with LD1, conducted at different NaCl concentrations. The major effect of lowering NaCl concentration has been to increase the size of the chemical shift changes undergoing fast-to-intermediate exchange. The effect is particularly marked for the shifts in the nitrogen dimension. The total number of chemical shift changes identified at 400mM NaCl was 15, at 150mM was 20 and at 50mM was 23. The 50mM NaCl has the advantage over the 150mM because it increases the observed shift perturbations so that the results can be analysed with more confidence.

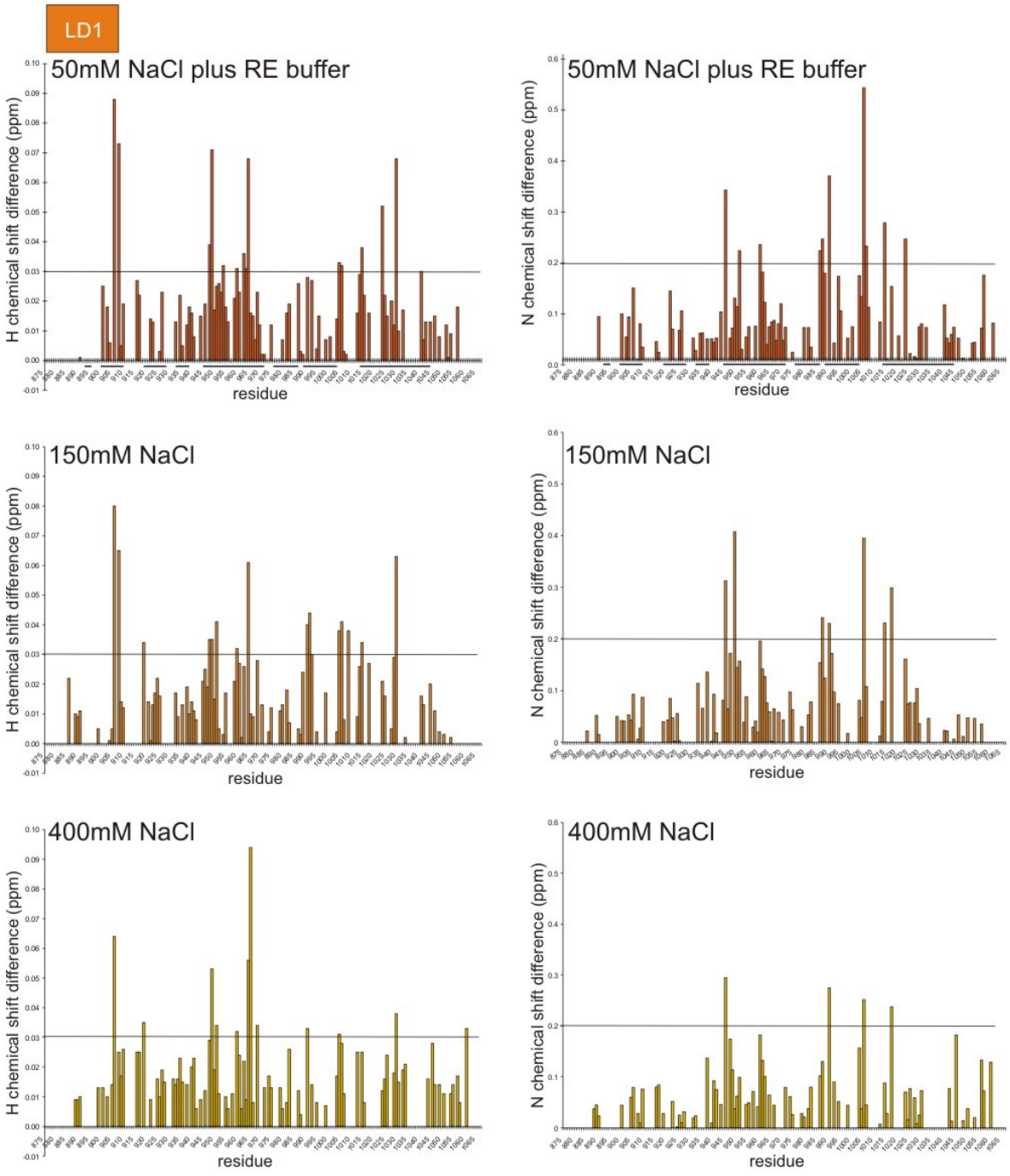


Figure 3.6.17 Bar graphs showing proton and nitrogen chemical shift differences at varying NaCl concentrations.

Line shows cut-off used in the analysis of 50mM NaCl titrations

3.6.3.5 Hydrogen-deuterium exchange for wild-type Vt in the presence of paxillin

Hydrogen-deuterium exchange experiments have been carried out to determine whether addition of LD peptides to wild-type Vt altered the protection of NH groups to exchange with D_2O. Previous studies with the FAT domain of FAK used this method to verify the proposed binding sites by observing that the residues suggested for the binding interface were also protected from exchange (6).

A ^{1}H-^{15}N-HSQC spectrum was recorded of wild-type Vt in 20mM Tris pH7, 50mM NaCl, 50mM arginine, 50mM glutamate, dissolved in water. This spectrum was acquired in 2 hours, which is the minimum time to get an adequate signal from a 200µM sample of Vt. It was then freeze-dried and re-dissolved in an identical volume of deuterium oxide. Immediately, a second identical ^{1}H-^{15}N-HSQC spectrum was recorded and then spectra were recorded consecutively with 2 hour intervals for 24 hours.

This method was repeated for wild-type Vt that had been prepared to contain a 10-fold excess of either LD1, LD2 or LD4.

3.6.3.6 Description of the data for wild-type Vt plus LD1, LD2, LD4 and His-LD1/LD2

Figure 3.6.19 shows the deuterium exchange data. When Vt is measured in water there are 164 resonances in this ^{1}H-^{15}N-HSQC spectrum that come from backbone NH atoms. After freeze-drying and re-dissolving in D_2O, 87 resonances remain (53%), these have resisted exchange with D_2O after 2 hours. The NH resonances that have been unambiguously assigned and which can be distinguished in the spectrum have been mapped onto the Vt crystal structure, the NH atoms have been represented (Figure 3.6.18). Most of the NH resonances that remain lie within the centre of the helices. This suggests that they may be involved in hydrogen bonding in the α-helix thus are protected from D_2O exchange. Helix 3 appears well protected against exchange, but there are NH resonances in all helices that

are protected. There is one NH that appears to be protected in the loop section near the top of the helix bundle, this is surprising and the reason for this is unknown.

When LD1 is added to Vt at a 10-fold molar excess, this ^{1}H-^{15}N-HSQC spectrum has 168 resonances when measured in water and 17 remain (10%) when measured in D_2O after 2 hours, a dramatic decrease. For LD2, there are 168 resonances when measured in water and 13 remain (8%) and for LD4, there are 164 resonances when measured in water and 43 remain (26%) when measured in D_2O. It is not possible to precisely identify the resonances that remain in the Vt plus LD peptide experiments. These results show that there is a decrease in the amount of protection to deuterium exchange when LD peptides are added to Vt compared to Vt alone.

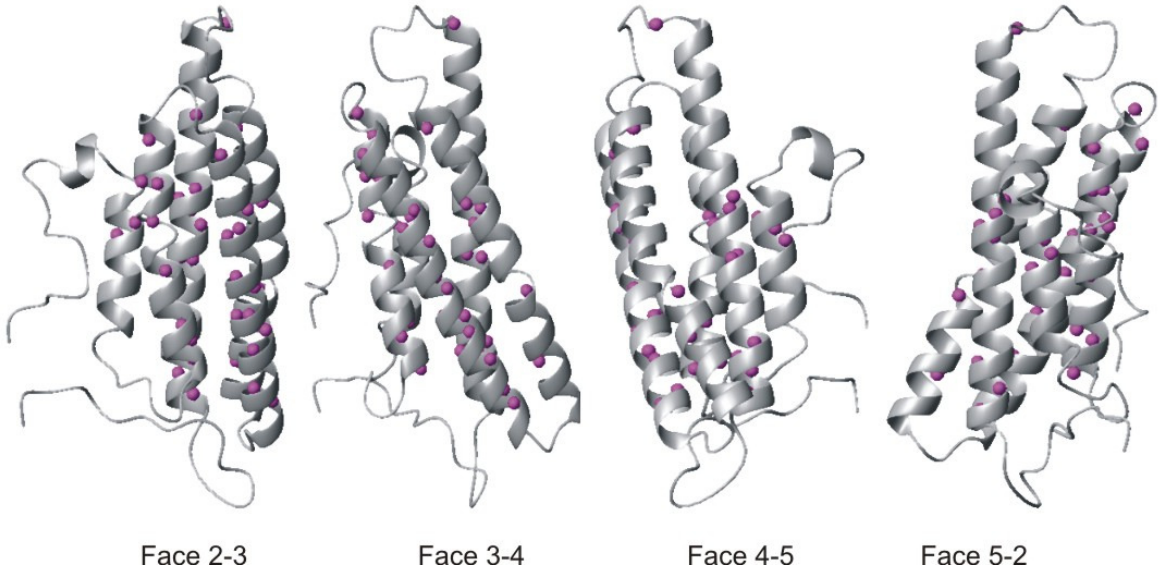


Figure 3.6.18 NH resonances in wild-type Vt that are protected from exchange with deuterium

Purple spheres represent NH groups that remain after 2 hours in D_2O

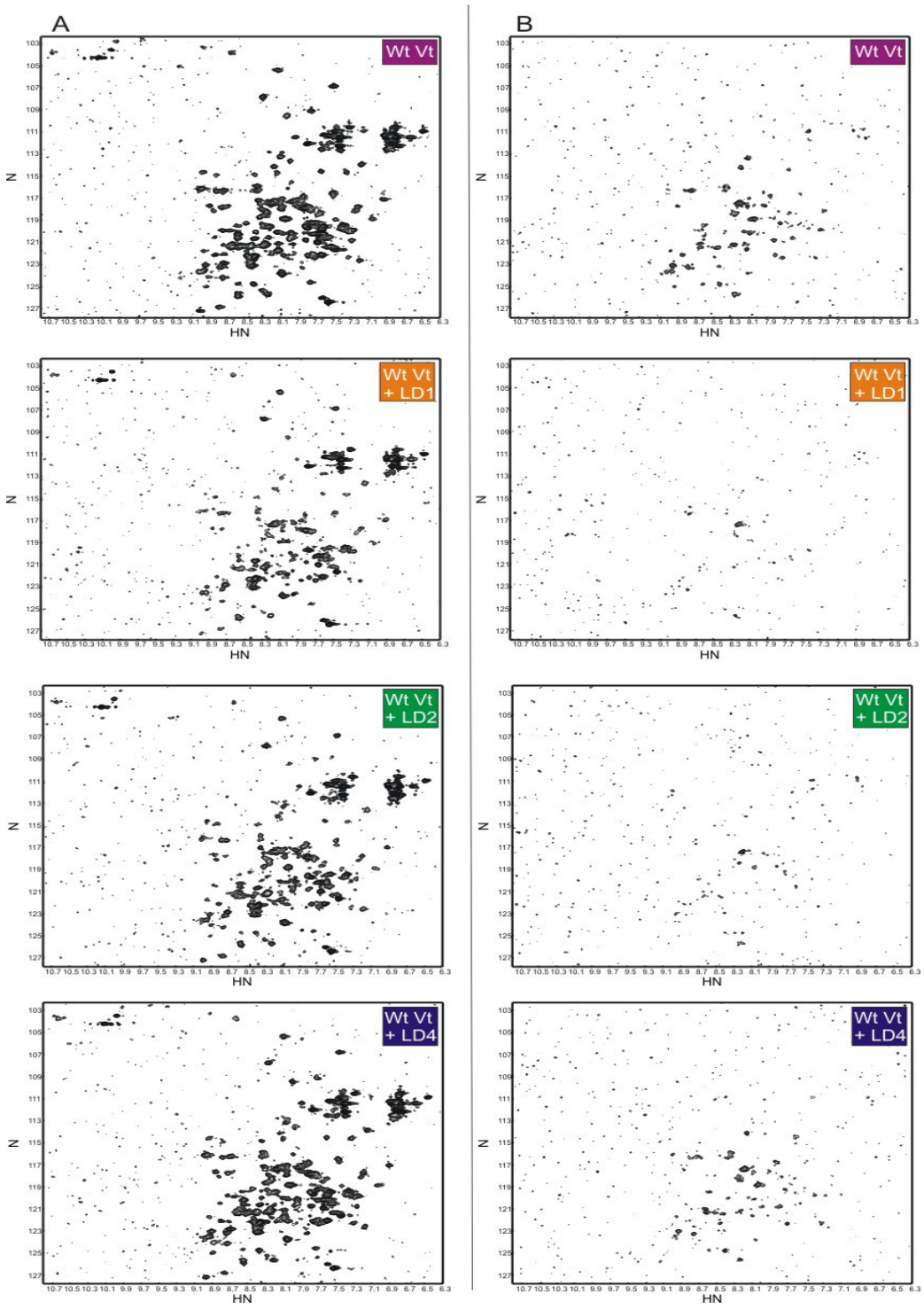


Figure 3.6.19 Results of the deuterium exchange experiments on wild-type Vt with different LD peptides added.

Spectra in column A are measured in water, spectra in column B are measured in D_2O. Peptide excess was 10-fold. The data were deliberately plotted at similar noise levels to allow a more accurate visual representation of the differences in the spectra.

Deuterium exchange experiments were carried out using Vt with His-LD1/LD2 and a 4-fold excess (Figure 3.6.20). There are 190 Vt resonances in this spectrum when measured in water, after freeze-drying and re-dissolving in D_2O, 37 resonances remained (19%).

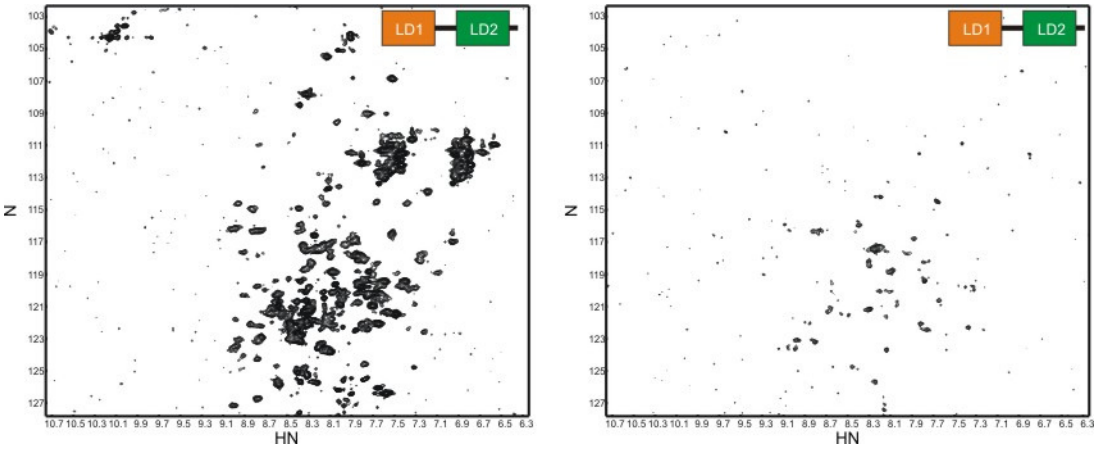


Figure 3.6.20 Results of the deuterium exchange experiments on wild-type Vt with His-LD1/LD2 added.

Spectrum on left is measured in water, spectrum on right is measured in D_2O. Ligand excess was 4-fold.

Table 3.6.2 Summary of results from the deuterium exchange experiments

	% resonances remaining after 2 hours
Wt Vt	53
Wt Vt + LD1	10
Wt Vt + LD2	8
Wt Vt + LD4	26
Wt Vt + His-LD1/LD2	19

The hydrogen-deuterium exchange data suggest that LD peptides are changing a global property of Vt. Table 3.6.2 summarises these results. From the data it can be seen that addition of LD1 and LD2 to Vt causes the amide groups to become more susceptible to exchange with D_2O. When LD4 is added the exchange is greater than for Vt alone, but not as drastic a change as for the other peptides. His-LD1/LD2 is an intermediary between the two. Care needs to be taken in interpreting the His-LD1/LD2 data as it is a large protein so may be having other effects on the spectrum.

3.6.3.7 Discussion

The titration data are complex to analyse and hydrogen-deuterium exchange results are not typical of this type of experiment. The chemical shift perturbations have not provided a discrete set of residues that could constitute an unambiguous binding face, and the deuterium exchange has shown that exchange is faster when paxillin is added to Vt. The following sections attempt to summarise the observations without over-interpreting the data that has been obtained.

Comparison of the three peptides LD1, LD2 and LD4 suggest that LD1 and LD2 will perturb the chemical shifts in the Vt spectrum whilst LD4 does not. Paxillin His-LD1/LD2 also perturbs similar residues to the individual LD1 and LD2 peptides.

Globally, there is a reduction in intensity of the chemical shifts as the LD1, LD2 and His-LD1/LD2 are titrated in to Vt. This is not due to changes in Vt concentration in the sample as this was carefully maintained. This intensity decrease was much less for LD4, and this is taken as evidence that the global intensity drop is caused by adding either LD1, LD2 or His-LD1/LD2. There are several reasons that could contribute to a global intensity decrease.

1. Increasing linewidth due to slower tumbling caused by increasing size
2. Contribution to line shape by intermediate exchange
3. Changing the folding state/affecting secondary structure of Vt

Vt could be interacting with the paxillin constructs which would increase the size of the whole molecule. This would increase the correlation time causing faster relaxation rates which would lead to line broadening and loss of signal-to-noise. Alternatively, Vt could be induced to interact with itself. There is evidence to show that Vt can dimerise (41;172;204), and it is possible that paxillin is promoting oligomerisation of Vt producing a molecule with a slower tumbling time.

Related to this is the possibility that adding paxillin puts the Vt chemical shifts into intermediate exchange leading to line broadening. This would apply to a paxillin-Vt interaction or Vt oligomerisaion.

Finally, if paxillin was affecting the structure of Vt then this could affect the spectrum. It is possible that Vt could become more loosely packed or slightly unfolded, this could affect the shape of the molecule and may also affect the correlation time.

Assuming for the moment that it is the paxillin interacting with Vt the following observations can be made. Both LD1 and LD2 peptides perturb the chemical shifts on a fast-to-intermediate exchange rate and an intermediate or slow exchange rate. LD1 affects more resonances in a fast-to-intermediate exchange manner than LD2. His-LD1/LD2 is mainly in intermediate or slow exchange with Vt. Interestingly, the His-LD1/LD2 data does not indicate a dramatically higher affinity than for the individual peptides. In other words paxillin His-LD1/LD2 does not appear to use aviditiy effects of two binding sites. The LD4 data suggests little interaction.

Paxillin LD motifs have been shown to bind to the solvent accessible surfaces of FAK (6-9), GIT1 (10) and actopaxin (α-parvin) (11). With this in mind, the data from the perturbation experiments was used to identify residues that appear on the solvent accessible surfaces of the molecule. This was judged by visual inspection, all perturbed residues that were on an accessible surface are potential candidates for binding. Table 3.6.3 shows all assigned residues that meet the criteria for an affected resonance that are also on the external surfaces of the α-helices of Vt. They were chosen by mapping all affected residues onto the crystal structure of Vt and picking the surface accessible residues. Residues that are affected in the core of the protein have not been picked and have been assumed to be non-specific effects.

Table 3.6.3. Potential candidates for binding interfaces on wild-type Vt

LD1							
Face 2-3		Face 3-4		Face 4-5		Face 5-2	
Ala	946	Arg	945	Gln	994	Ala	923
Cys	950	Ile	948	Thr	990	Lys	924
Asp	953	Gln	949	Arg	987	Ser	934
Ala	957	Ala	955	Glu	1014	Arg	938
Glu	960	Asp	959	Glu	1015	Asp	1013
Leu	964	Thr	962	His	1025	Glu	1017
Glu	967	Arg	963	Asn	1029		
Val	968	Lys	966	Ser	1033		
Gln	971	Lys	970	Glu	1040		
		Leu	982				
		Glu	986				
		Ser	992				
		Lys	996				
		Thr	1000				
LD2							
Face 2-3		Face 3-4		Face 4-5		Face 5-2	
Ala	946	Arg	945	Gln	994	Asp	1013
Cys	950	Ile	948	Thr	990	Glu	1017
Glu	967	Gln	949	Glu	1014		
Val	968	Lys	952	Glu	1015		
		Asp	959	His	1025		
		Arg	963	Asn	1026		
		Lys	966	Ser	1033		
		Lys	970				
		Leu	982				
		Thr	1000				
His-LD1/LD2							
Face 2-3		Face 3-4		Face 4-5		Face 5-2	
Asp	918	Arg	945	Arg	987	Ile	920
Ala	921	Ile	948	Thr	990	Ala	923
Ala	922	Gln	949	Gln	994	Lys	924
Glu	932	Lys	952	Glu	1014	Ala	927
Gln	949	Ala	955	Glu	1015	Ser	934
Cys	950	Asp	959	Asn	1026	Arg	935
Asp	953	Arg	963	Gln	1032	Arg	938
Ala	957	Lys	966	Ser	1033	Asp	1013
Glu	960	Lys	970			Thr	1020
Glu	967	Arg	978			Val	1024
Gln	971	Thr	979			Gln	1032
		Leu	982			Glu	1042
		Ser	992				
		Lys	996				
		Thr	1000				

From this it can be seen that the highest number of perturbed residues are on Vt face 3-4 with many of these being on helix 3. Mapping these residues onto the crystal structure of Vt presents a potential interaction site on face 3-4 that covers most of helix 3 and part of helix 4 (Figure 3.6.21). This region is similar when LD1, LD2 and His-LD1/LD2 is titrated into Vt suggesting that both LD1 and LD2 would interact at the same site on face 3-4, but there may be subtle differences between how LD1 or LD2 binds.

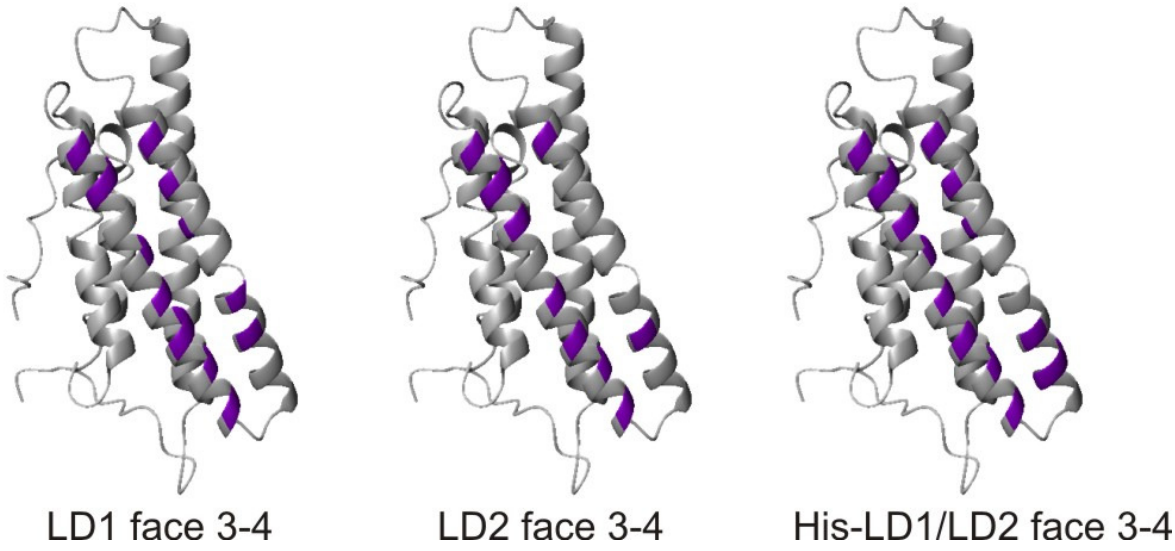


Figure 3.6.21 Perturbed surface residues of Vt face 3-4 (purple)

The affected residues on the face of helices 2-3 and 4-5 are shown in Figure 3.6.22. For face 2-3, LD1 and LD2 perturb residues on helix 3, His-LD1/LD2 perturbs some on helix 2 also. For face 4-5 there are few perturbations but they are the same for all three paxillin constructs.

Figure 3.6.23 shows the perturbed surface residues on face 5-2, plus all affected residues in the first helix, helix1. These residues can be found in Appendix 5.7. LD1 is perturbing a few residues on this face, LD2 perturbs less. In contrast there are a 12 residues perturbed by His-LD1/LD2.

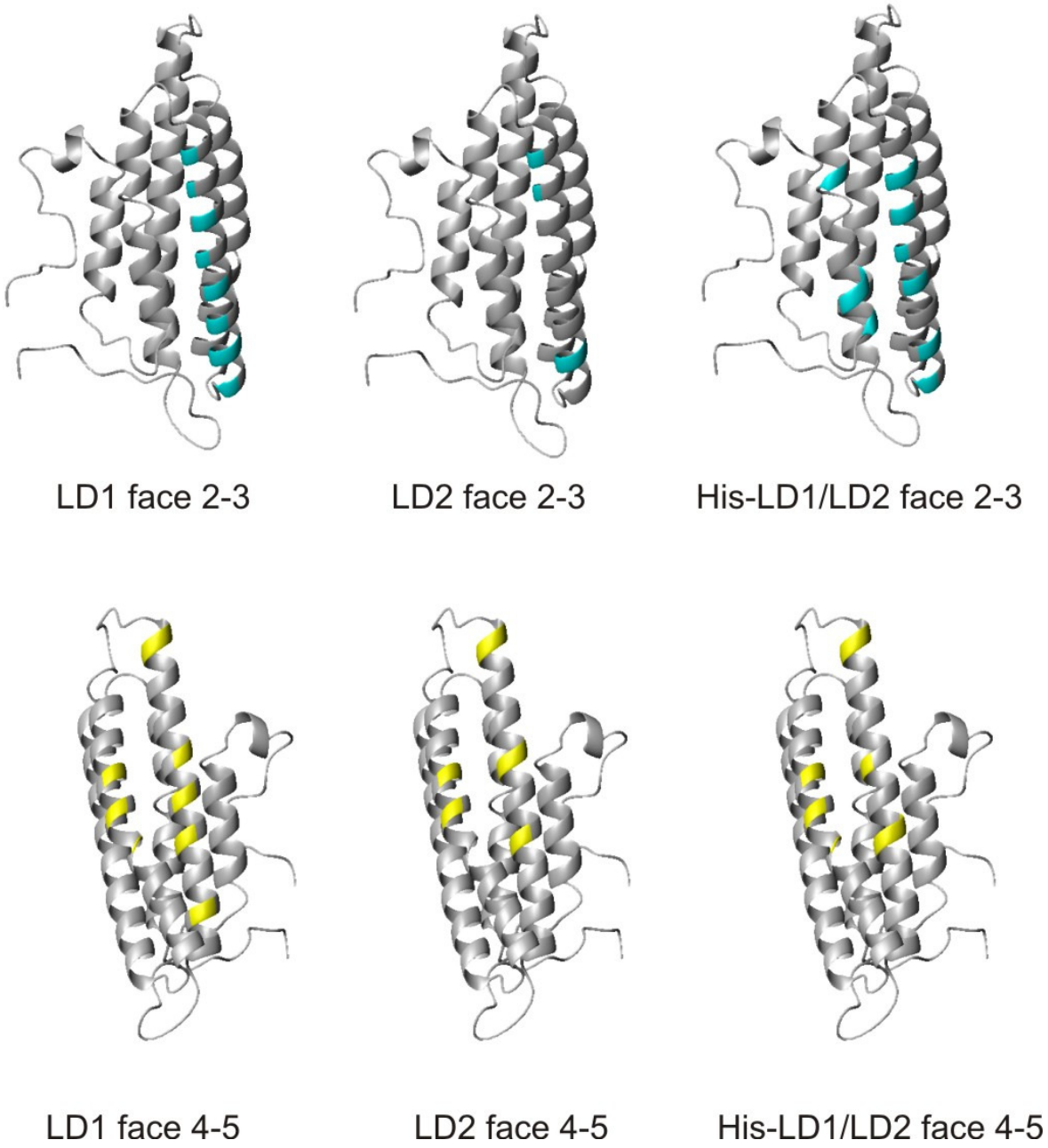


Figure 3.6.22 Perturbed residues on Vt face 2-3 (top-cyan) and 4-5 (bottom-yellow)

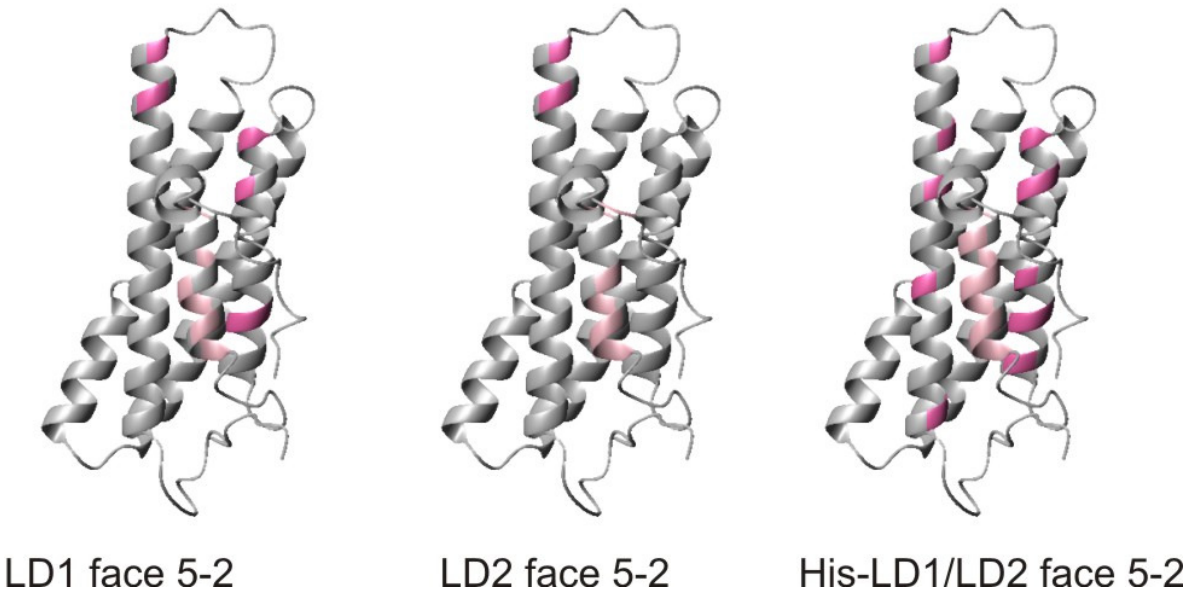


Figure 3.6.23 Potential interaction site of LD motifs with face 5-2 of Wt-Vt.
Dark pink are perturbed residues on face 5-2, light pink are all perturbed residues on helix1

The above figures present the perturbed residues that are on the exterior surfaces of the α-helices of Vt. Assuming that the interaction is a simple 'face-to-face' binding where the surface of one helix contacts the helix face of the ligand, these could be potential interaction sites. Interesting possibilities arise from these pictures, for instance there appears to be an attractive interface on face 3-4 (Figure 3.6.21) where subtle differences between the specific residues could be involved in determining whether LD1 or LD2 binds.

A second interesting observation is the perturbation of face 5-2 with His-LD1/LD2 compared to the individual LD peptides. LD2 peptides perturb only 2 residues on this face, LD1 perturbs 5, yet the longer construct containing both LD1 and LD2 perturbs 12 residues on face 5-2. This could suggest that His-LD1/LD2 may gain access to this face where the peptides cannot. This could occur if His-LD1/LD2 exhibits a higher affinity than the two motifs separately by having increased avidity, although the perturbation data does not fully support this. However, the GST-pulldown and catch-up data argue for an interaction with higher affinity between Vt and His-LD1/LD2 than suggested by the NMR data alone. One

limitation of the NMR data for the paxillin construct is that a molar ratio of 1:10 was not achieved, if this molar ratio is achieved then increased affinity and avidity effects may be observed. The structure of Vt suggests that face 5-2 could be occluded by helix1, His-LD1/LD2 may have a property that will enable access to this site.

It was hoped that hydrogen-deuterium exchange would shed light on the proposed binding interfaces by showing areas of increased protection in the regions of ligand interaction. Contrary to expectation, the data suggests that exchange is actually increased when Vt is in the presence of paxillin. There are several interesting possibilities for this result. If the NH groups are now more accessible to exchange it could suggest that the Vt molecule has been destabilised. It is also possible that Vt in solution is in a self-association equilibrium and that paxillin shifts the equilibrium to a more solvent accessible state. Either way it seems that addition of paxillin is affecting the dynamic properties of Vt by increasing its susceptibility to exchange with deuterium.

This deuterium data gives rise to the possibility that LD peptides are changing a global property of Vt and that the interaction between paxillin and Vt may be more complex than a simple 'face-face' helix interaction. There are large scale global changes occurring, and it is difficult to imagine how a 13mer peptide would induce all of these by binding to one small region on Vt. Residues are being affected in the same molecule by different rates of chemical exchange which could suggest that both a binding and unfurling/destabilising event is occurring simultaneously.

Because the hydrogen-deuterium exchange data show global perturbations, the background changes in the chemical shift perturbation data is perhaps unsurprising, so by setting a criteria that searches for a small number of the most perturbed chemical shifts and intensities it is hoped to identify a subset of residues that are reporting primarily on residues that bind LD peptides directly.

3.6.3.8 Summary of results

The data suggest that the Vt-LD interaction has a global and a local aspect, consequently LD binding is more complex than a face-to-face binding with preformed interaction surfaces. Dissection of these effects is difficult but the following observations seem to be possible

The data are consistent with a specificity between LD motifs; LD1 and LD2 both perturb residues in wild-type Vt, LD4 does not. The affinity can be estimated by looking at whether Vt residues undergo fast or slow exchange with LD peptides, using this rationale, LD2 would have a slightly higher affinity than LD1. His-LD1/LD2 would have a higher affinity than the individual LD peptides.

Both peptides LD1 and LD2 cause chemical shift perturbations that have been grouped into regions based on residues that are accessible on the helix faces of Vt. This data may be interpreted as potential interaction sites, the most attractive site being face 3-4 for LD1, LD2 and His-LD1/LD2. Face 5-2 may be a second site accessible to His-LD1/LD2, with LD1 interacting weakly and LD2 not interacting at all. Face 2-3 and 4-5 would be the least attractive interaction sites.

The deuterium exchange data shows that the global properties of Vt is altered when paxillin is added and suggests that the interaction between the two may be more complex than previously thought.

3.6.3.9 How do the current experiments compare to previous studies of Vt-LD interactions?

As the ligand binding data are not a clear-cut case producing a well defined binding site, there is merit in evaluating the previously published results in the light of this new data. It has previously been shown that Vt will bind to LD1, LD2 and LD4 (5;80) with LD2 being the strongest interaction and LD4 binding only weakly in GST pulldown studies.

Prior to this, the two regions of Vt suggested to be paxillin binding subdomain (PBS) were defined as PBS1 and PBS2. PBS1 was suggested due to sequence similarity to FAK as Vt residues 952-971 (110), and PBS2 was determined through truncation mutagenesis and pulldown studies to include Vt residues 978-1000 (86). PBS1 lies on helix 3 and PBS 2 lies on helix 4 of Vt (Figure 3.6.24).

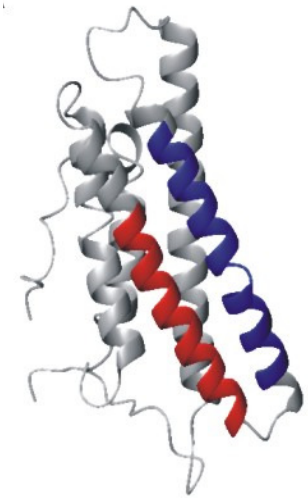


Figure 3.6.24 Paxillin binding subdomains 1 (PBS1:red) and 2 (PBS2:blue) (86;110)

These PBSs had also been predicted for the FAT of FAK (residing on helix 1 and helix 4). A study by Hoellerer *et al.* (6) determined that the binding interface overlapped partially with this PBS, in addition they determined the specific residues that LD2 and LD4 would interact with on FAT helices 1-4 and 2-3. The residues involved in the interface can be found in Table 3.6.4.

Based on these findings, Hoellerer *et al.* (6) predicted two new structure based interaction sites on Vt on opposite faces of Vt (Table 3.6.4). The Vt domain adopts a fold similar to the FAT domain, but FAT was found to interact with LD2 and LD4 (5-9) while Vt can interact with LD1, LD2 and LD4 (5;80). Hoellerer *et. al* (6) speculated that the residues involved in FAT-LD binding can be superimposed onto Vt to create two predicted binding sites for LD motifs (Figure 3.6.25). The two proposed sites make up face 3-4 and face 5-2

on Vt, the specific residues can be found in Table 3.6.4 and in the introduction, Figure 1.5.4. The authors state that Vt face 5-2 has the pre-requisites for binding LD motifs (based on their findings with the FAT domain), namely a hydrophobic patch between two helices surrounded by basic residues that can interact with negative charges in the peptide. However, Vt face 5-2 is occluded by helix 1 and may require displacement of helix 1 to enable binding.

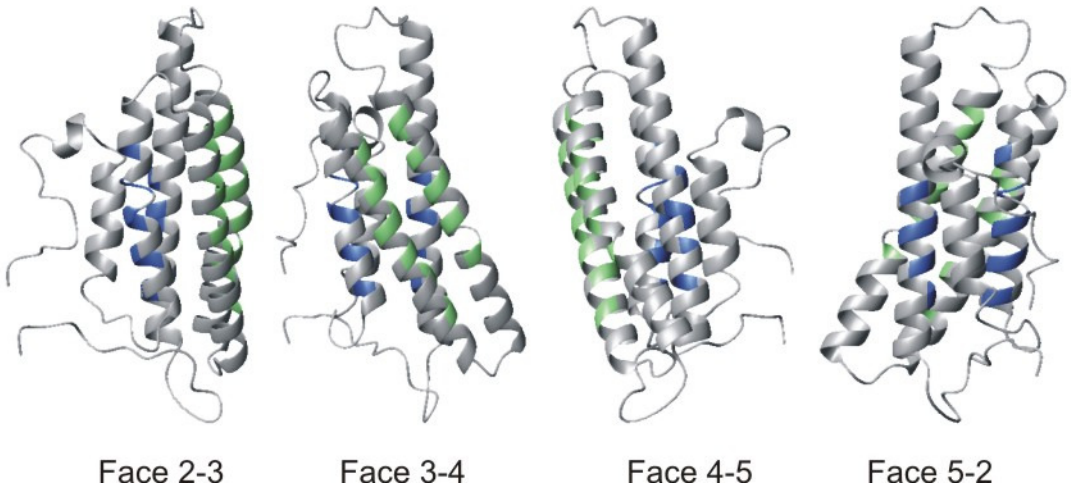


Figure 3.6.25 Predicted binding sites in Vt based on the FAT domain of FAK.
Pale green residues make up binding interface Vt 3-4 (FAK 2-3) and pale blue make up face Vt 5-2 (FAK 1-4).

With this knowledge a comparison can be made with the data presented in this thesis. The attractive binding face 3-4 on Vt is compared to the PBS (86;110) and the new predicted two-face binding sites (6) in Figure 3.6.26A-C. There are similarities between the experimentally derived data for Vt and both PBS1 and PBS2 (B) and the predicted two-site model (C).

The second potential site on Vt face 5-2 is compared in Figure 3.6.26D-E. While the region is similar between the two there are more residues affected by His-LD1/LD2 than predicted. This is likely a consequence of the extra amino acids in the His-LD1/LD2 protein compared to 13mer peptide.

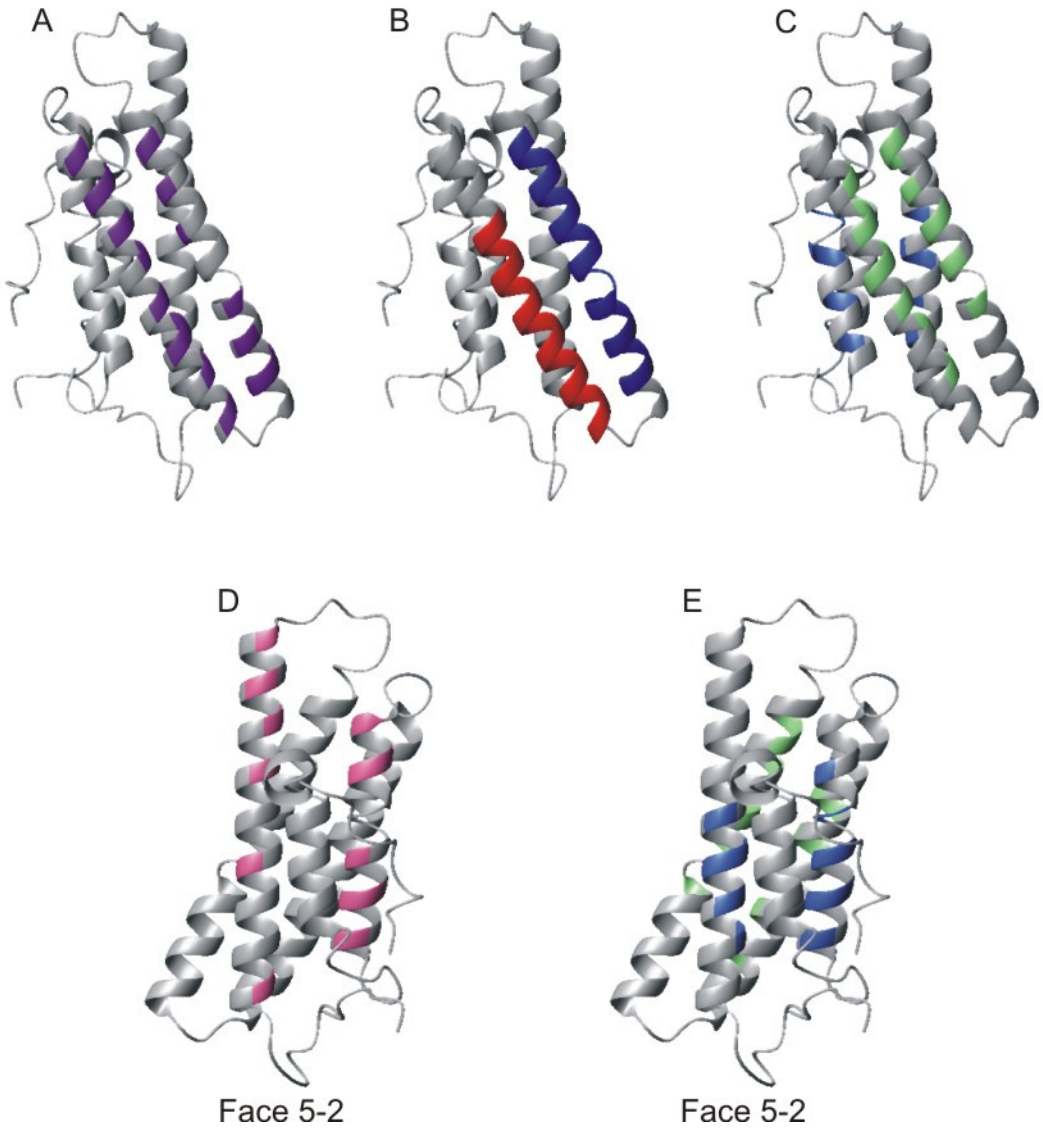


Figure 3.6.26 Comparison of proposed binding interface on Vt with previous predictions

A: Proposed binding interface 3-4 from the titration data (purple); B: PBS1 (blue) and PBS2 (red); C: Predicted binding interface 3-4 from FAK data; D: Proposed binding interface 5-2 from the titration data (pink); D: Predicted binding interface 5-2 from FAK data (light blue).

A list of the residues in FAT affected by LD2 and LD4 and the predicted residues in Vt are shown in Table 3.6.4. A comparison has been made to determine which paxillin LD motifs perturbed these predicted residues in the titration experiments. Because the assignment of Vt was conducted using Vt/I997S, some assignments could not be transferred to the wild-

type spectrum as peaks were ambiguous in their position or they were missing. This means that some are missing from the analysis, these are labeled accordingly in the table. As a result there are 10 out of 15 residues that can be analysed on face 3-4, and 9 out of 13 that can be analysed on face 5-2.

For the residues predicted by Hoellerer *et al.* (6) to be on the Vt face 3-4 interaction site, chemical shift perturbations above the criteria set for analysis were seen for nine residues by LD1, LD2 caused perturbations to five residues, and His-LD1/LD2 perturbed eight. On Vt face 5-2, LD1 perturbed three residues but LD2 did not perturb any of the predicted residues. His-LD1/LD2 perturbed six.

In addition, for face 3-4 there were and extra five, five and seven residues in Vt perturbed by LD1, LD2 and His-LD1/LD2 respectively that were not identified in the FAK binding study. For face 5-2 there are and extra three perturbed residues with LD1, two residues for LD2 and six residues for His-LD1/LD2.

This comparison is summarised in Table 3.6.5. for Vt and FAK Assuming that these sites do represent an interaction site in Vt, the data suggests that both LD1 and LD2 bind to Vt face 3-4, and in addition LD1 possibly may interact with Vt face 5-2 while this does not appear to be the case for LD2. The data is consistent with no interaction for either site with LD4.

It cannot be ruled out that LD2 and LD4 could interact with Vt face 5-2 as the affinity for this site may be too low to overcome potential steric hindrance from helix1.

Table 3.6.4 Comparison of the residues determined experimentally for FAK, the two site prediction for Vt and the experimental data from this thesis

Experimental data[a]		Predicted residues[b]		Vt residues affected[c]		
FAK	face 2_3	Vt	face 3_4	LD1	LD2	LD1/LD2
Val	951	Ile	948	+	+	+
Val	954	Ala	951	-	-	-
Lys	955	Lys	952	-	+	+
Gly	958	Ala	955	+	-	+
Leu	959	Lys	956	-	-	-
Arg	962	Asp	959	+	+	+
Leu	965	Thr	962	+	-	-
Asp	969	Lys	966	+	+	+
Lys	988	Glu	986	+	-	-
Asn	991	Pro	989	-	-	-
Leu	994	Ser	992	+	-	+
Gly	995	Thr	993	-	-	-
Thr	998	Lys	996	+	-	+
Met	1001	Ser	999	-	-	-
Lys	1002	Thr	1000	+	+	+

Experimental data[a]		Predicted residues[b]		Vt residues affected[c]		
FAK	face 1_4	Vt	face 5_2	LD1	LD2	LD1/LD2
Arg	919					
Tyr	925	Ile	920	-	-	+
Val	928	Ala	923	+	-	+
Thr	929	Lys	924	+	-	+
Val	932	Ala	927	-	-	+
Lys	933	Leu	928	-	-	-
Val	935	Met	930	-	-	-
Ile	936	Ala	931	-	-	-
Ser	939	Ser	934	+	-	+
His	1025	Gln	1028	-	-	-
Ala	1028	Met	1031	-	-	-
Val	1029	Gln	1032	-	-	+
Lys	1032	Lys	1035	-	-	-
Leu	1035	Val	1038	-	-	-

(a) experimentally determined binding residues for LD2 and LD4 (6). (b) predicted residues based on FAK data (6). (c) experimentally determined chemical shift perturbations from this study. '+' means perturbation observed above the criteria set in the analysis, '-' means perturbation not observed above the criteria Residues in blue are unassigned, residues in red cannot be unambiguously determined in the wild-type Vt spectrum when transferring assignments from Vt/I997S to wild-type ^{1}H-^{15}N-HSQC spectra.

Table 3.6.5 Comparison of the Vt data with FAK data from Hoellerer *et al.* (6)

	LD1		LD2		LD4		His-LD1/LD2	
	FAK	Vt	FAK	Vt	FAK	Vt	FAK	Vt
Face 3-4	-	+	+	+	+	-	nd	+
Face 5-2	-	+[(a)]	+	-	+	-	nd	+

Plus signs represent an interaction, minus represents no interaction. nd=not done. (a) possible interaction, few residues affected.

Deuterium exchange experiments carried out on the FAT domain of FAK showed that the same regions that were affected by chemical shift perturbations were also protected from exchange with D_2O, confirming that the LD peptides were interacting and giving those regions more stability. In contrast, D_2O exchange experiments with Vt do not show regions of increased protection when Vt and LD peptides are added together, in fact it suggests that Vt is becoming more susceptible to exchange in the presence of LD1 or LD2. This suggests that a global mechanism may be operating at the same time as local binding of LD peptides. What is clear is that the mechanism of interaction of Vt and paxillin is distinct from that of FAK and paxillin and it may not be a simple 'face-face' interaction between the helices.

The chemical shift perturbation data suggests that the Vt residues perturbed by LD peptides are subtly different to that of LD peptides with FAK. LD2 and LD4 were initially shown to be promiscuous recognition sequences when interacting with FAK (6), now there is some evidence that there may be site specificity (7;207). The data presented here would be consistent with site specificity of LD peptides for Vt. LD4 does not perturb Vt face 3-4 whilst LD1 and LD2 do. LD1 can affect residues on the occluded Vt face 5-2, whereas LD2 and LD4 cannot. And even though LD2 and LD4 perturbations to this face cannot be completely excluded, there is a difference between them and LD1 in affecting this face. His-LD1/LD2 possesses properties that will allow perturbations to Vt face 5-2 which under the conditions used here the individual peptides do not.

With the observations of global mechanisms causing increased hydrogen-deuterium exchange properties of Vt, thermal melting experiments have been carried out to observe the effects of LD peptides on the stability of Vt (Section 3.6.5).

3.6.4 The His-LD1/LD2 ^{1}H-^{15}N-HSQC spectrum in the presence of excess wild-type Vt

Since paxillin His-LD1/LD2 can be recombinantly expressed and isotope labelled it was possible to observe the perturbations on paxillin His-LD1/LD2 when titrated with unlabeled Vt. The ^{1}H-^{15}N-HSQC of paxillin at 900MHz with and without Vt at 4-fold excess is shown in Figure 3.6.27. Paxillin assignments were not yet available, so the perturbations cannot be assigned to specific amino acids, but the overlay shows that there are a number of resonances that are missing from the spectrum and there are some new peaks that have appeared.

In the spectrum, 3 peaks are new and there are 18 resonances that have completely disappeared. This suggests that only a small number of residues in paxillin are involved in the interaction with Vt. Tantalisingly, the number of residues contained in the LD1 and LD2 motif is 16; corresponding well to the number of resonances lost in the His-LD1/LD2 spectrum. A titration carried out at 600MHz under the same buffer conditions suggest that there are fast-to-intermediate and slow exchange effects occurring (data not shown). Attempts to identify whether one set of resonances may be affected before another (corresponding to one LD motif binding before another) was not possible to determine from the data. Clearly identification of the interacting residues must await assignment.

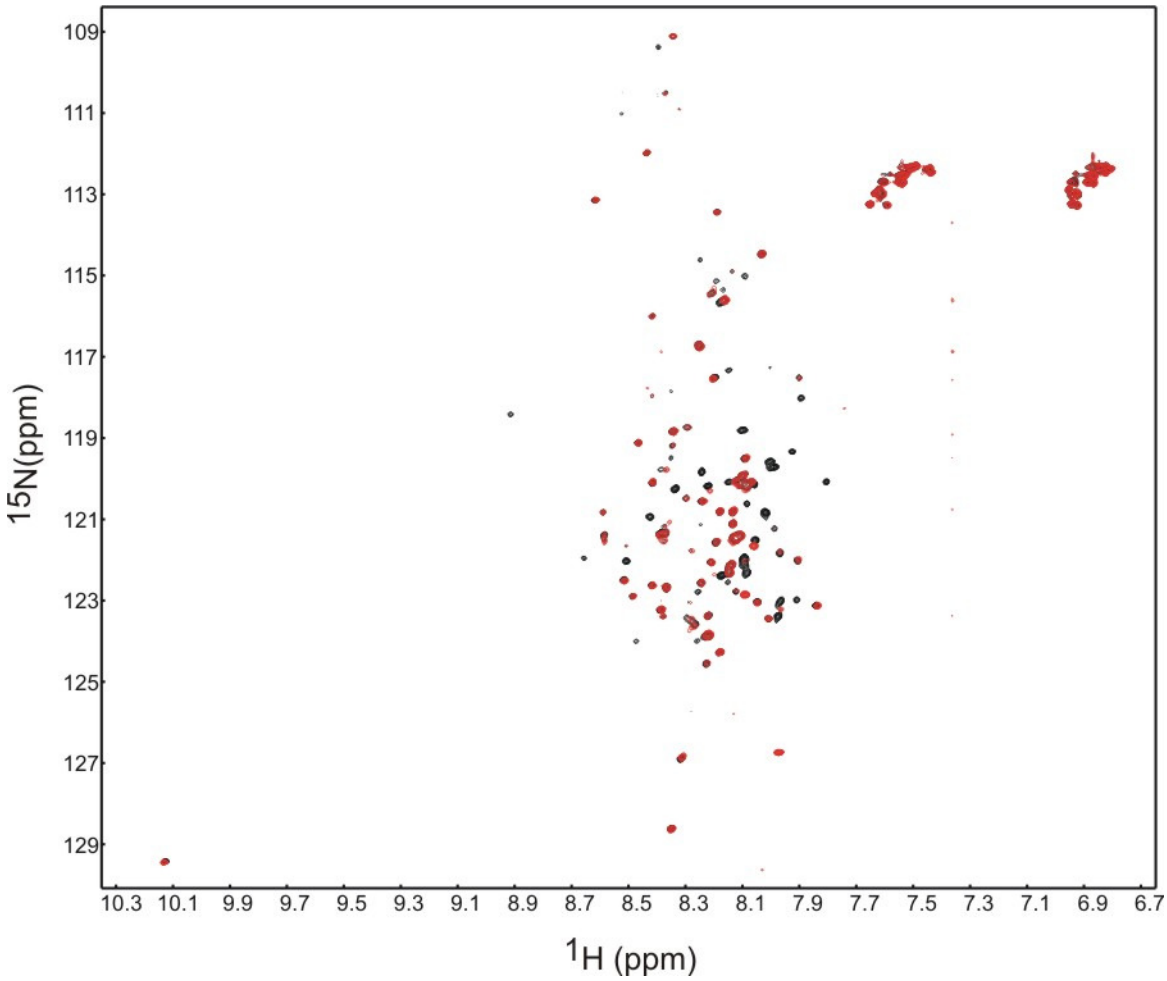


Figure 3.6.27 Overlay ^{1}H-^{15}N-HSQC spectrum of paxillin with excess Vt

Measured at 900MHz with a cryoprobe. Paxillin alone shown in black, paxillin with 4-fold excess of Vt is shown in red. Paxillin His-LD1/LD2 at 100µM in 20mM Tris pH7, 50mM NaCl, 50mM arginine, 50mM glutamate. Vt was made to 600µM and added to 4-fold excess in the same buffer.

An interesting observation was noted when an excess of Vt was added to paxillin. On addition of a 4-fold excess of Vt to the solution of paxillin (at 100µM), the sample is clear. After leaving the sample overnight a massive precipitation event occurs. Figure 3.6.28 shows the contents of various samples from this experiment. In the initial mixture of 4-fold excess Vt with paxillin (lane 3), a large band can be seen which coincides with the control Vt band. Interestingly, a band corresponding to His-LD1/LD2 cannot be observed, despite being detected in a ^{1}H-^{15}N-HSQC spectrum. There is also a protein band at 80kDa whose

identity is unknown. Mass spectrometry of this sample produces a peak of molecular mass matching that of Vt (data not shown). The sample is pelleted after 12 hours to remove the precipitate, lane 4 shows the remaining soluble fraction. Mass spectrometry of this sample also produces a peak corresponding to Vt, and the 80kDa band is still present. Lane 2 shows the precipitate. A mass spectrum of this sample was unsuccessful, but the major gel band would be consistent with the molecular weight of Vt.

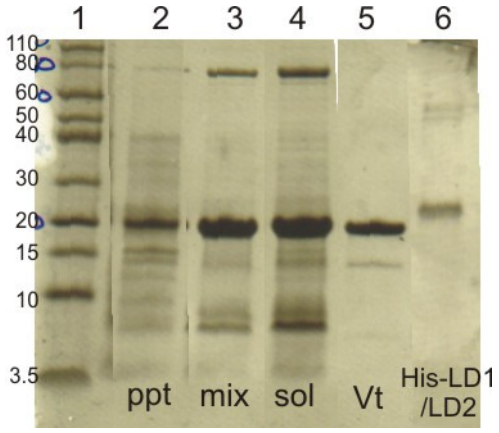


Figure 3.6.28 Composite polyacrylamide gel showing different fractions from a His-LD1/LD2 sample containing excess Vt

A 4-fold excess of ^{14}N-Vt was added to ^{15}N-His-LD1/LD2 (100μM). After 12 hours a precipitation event is observed. The contents of the precipitate is shown in lane 2 (ppt). Lane 3 shows the initial mixture of Vt and His-LD1/LD2 (mix), and lane 4 shows the remaining soluble fraction after removal of the precipitate. Lane 5 is Vt at 600μM that was left at room temperature for twelve hours, and lane 6 is a sample of His-LD1/LD2.

A comparison of paxillin resonance intensity from ^{1}H-^{15}N-HSQC spectra before and after a precipitation event showed that paxillin can still be seen in solution at a similar intensity, looking at the relative intensities of resonances that remain from a sample that has been spun down to remove precipitate show a decrease in intensity by an average of 20 %. This has been taken as evidence that paxillin is remaining in solution and that Vt may be precipitating out of solution. The small loss of paxillin intensity may infer that the precipitate was made predominantly from Vt. It is intriguing to speculate that paxillin may act as a catalyst for Vt aggregation or unfolding.

3.6.5 Conformation of Vt in the presence of paxillin and thermal denaturation of Vt

An idea that arose from the titration data is that the integrity of the structure of Vt could be affected by paxillin peptides LD1 and LD2 and the longer construct His-LD1/LD2. Hydrogen-deuterium exchange experiments show dramatically increased exchange rates with solvent in the presence of LD1, LD2 and His-LD1/LD2. With this in mind, far-UV CD spectra were acquired on Vt in the presence of paxillin to observe any conformational differences and the stability of Vt was measured in the presence of paxillin by thermal denaturation monitored by circular dichroism. The CD data was reported as ellipticity as the binding constant for the interaction is not known. For His-LD1/LD2, ratios of 1:1 and 1:4 have been used, 1:4 being the ratio used in the chemical shift perturbation experiments. For the LD peptides is was not possible to achieve the excess used for chemical shift perturbation, so LD peptides at 50μM were used. If one assumes that the Kd was in the region of 10^{-6}M for LD peptides, when Vt is at 0.5μM and LD is at 50μM the expected proportion bound is 33%.

3.6.5.1 Wild-type Vt

Figure 3.6.29a shows the circular dichroism spectrum of wild-type Vt at 25ºC and 85ºC. At 25ºC the protein has a typical spectrum for an α-helix. At 85ºC it is evident that the degree of helicity has reduced. Figure 3.6.29b shows the thermal denaturation curve for wild-type Vt. The melting temperature (Tm) can be calculated from the midpoint of the transition. This is found by finding the first derivative of this curve and finding the maximum value (see Methods Figure 2.3.2b-c). The results of this experiment suggest that wild-type Vt has a Tm of 74.9ºC ± 0.19. This melting curve is not ideal as it is lacking a baseline at the end of the experiment, this was unexpected and is in part due to the high Tm for Vt.

3.6.5.2 Wild-type Vt plus LD peptides

The difference spectra from the wavelength scans for wild-type Vt plus LD peptides are shown in Figure 3.6.30a. The addition of all LD peptides at 25ºC appears to reduce the helicity by a small amount, with the reduction by LD2 and LD4 being the same and LD1 being reduced the most. At 85ºC the reduction in helicity is marked and does appear to be greater than when Vt is alone. Figure 3.6.30b shows the melting profile for Vt plus LD peptides. The melting temperatures for wild-type Vt plus LD1, LD2 or LD4 is 69.8ºC ± 0.30, 78.4ºC ± 0.46 and 75.0ºC ± 0.44 respectively. Together this data suggests that LD1 peptide reduces the helicity of Vt and lowers the Tm by 5ºC, LD2 lowers the helicity of Vt but increases the Tm by 3ºC and LD4 lowers the helicity but does not affect the Tm.

3.6.5.3 Wild-type Vt plus His-LD1/LD2

The difference spectra from the wavelength scans for wild-type Vt plus paxillin His-LD1/LD2 is shown in Figure 3.6.31a. The addition of paxillin at a 1:1 and 1:4 ratio at 25ºC appears to reduce the helicity of Vt by a small amount. At 85ºC the reduction in helicity is marked and is greater than when Vt is alone. The 1:4 ratio has a very large change in the CD spectrum which is greater than the 1:1.

Figure 3.6.31b shows the melting profile for Vt plus His-LD1/LD2 at 1:1 and 1:4 ratio. The melting temperature for the 1:1 and 1:4 ratios is 72.8ºC ± 0.28 and 72.3ºC ± 0.38 respectively. Together this data suggests that His-LD1/LD2 reduces the helicity of Vt with the 1:4 ratio being more efficient at this than the 1:1. His-LD1/LD2 at 1:1 lowers the Tm by 2ºC, and at 1:4 lowers the Tm by 3ºC.

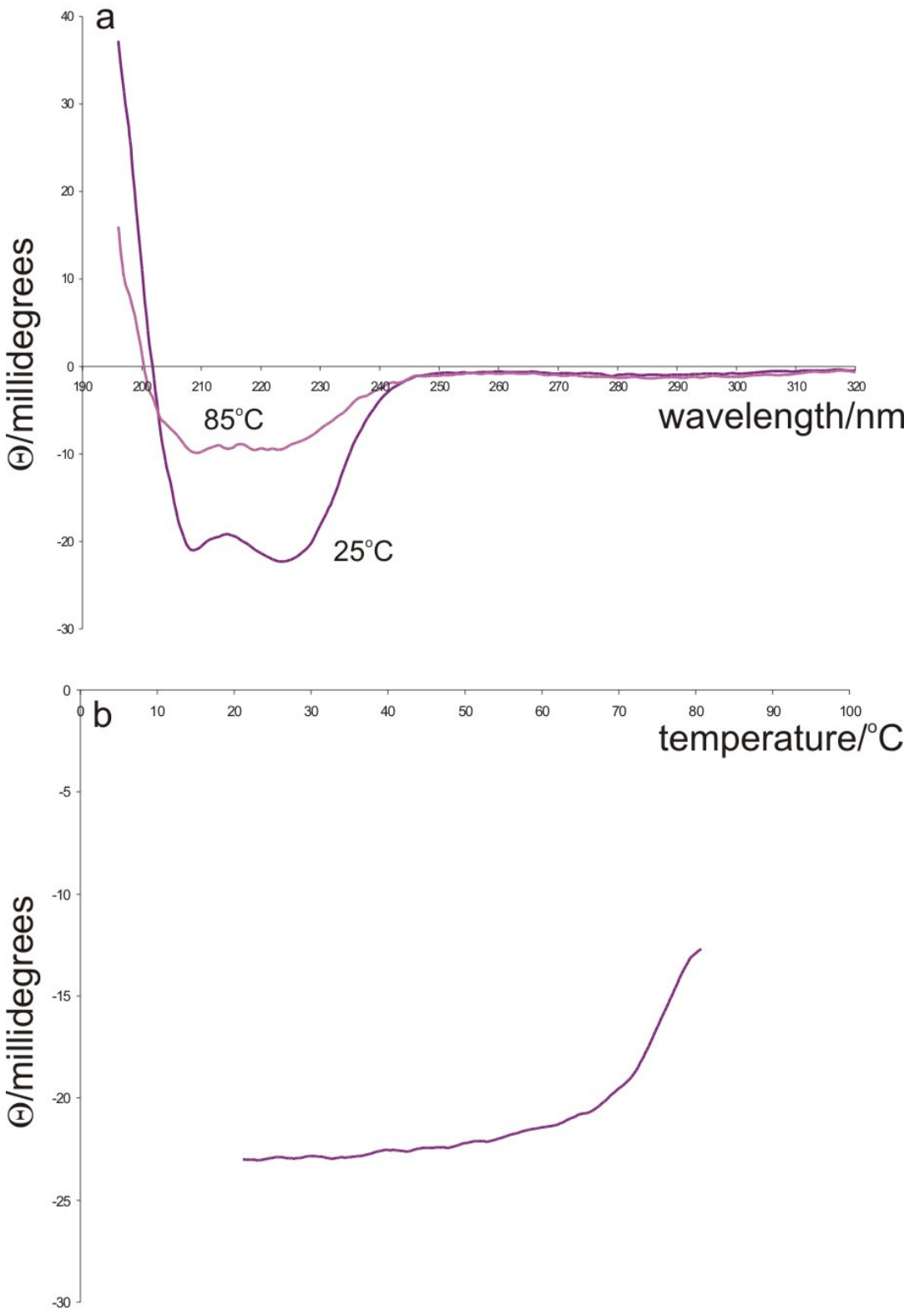


Figure 3.6.29 Far UV difference wavelength scan at 25ºC and 85ºC and thermal denaturation curve for wild-type Vt at 0.5µM

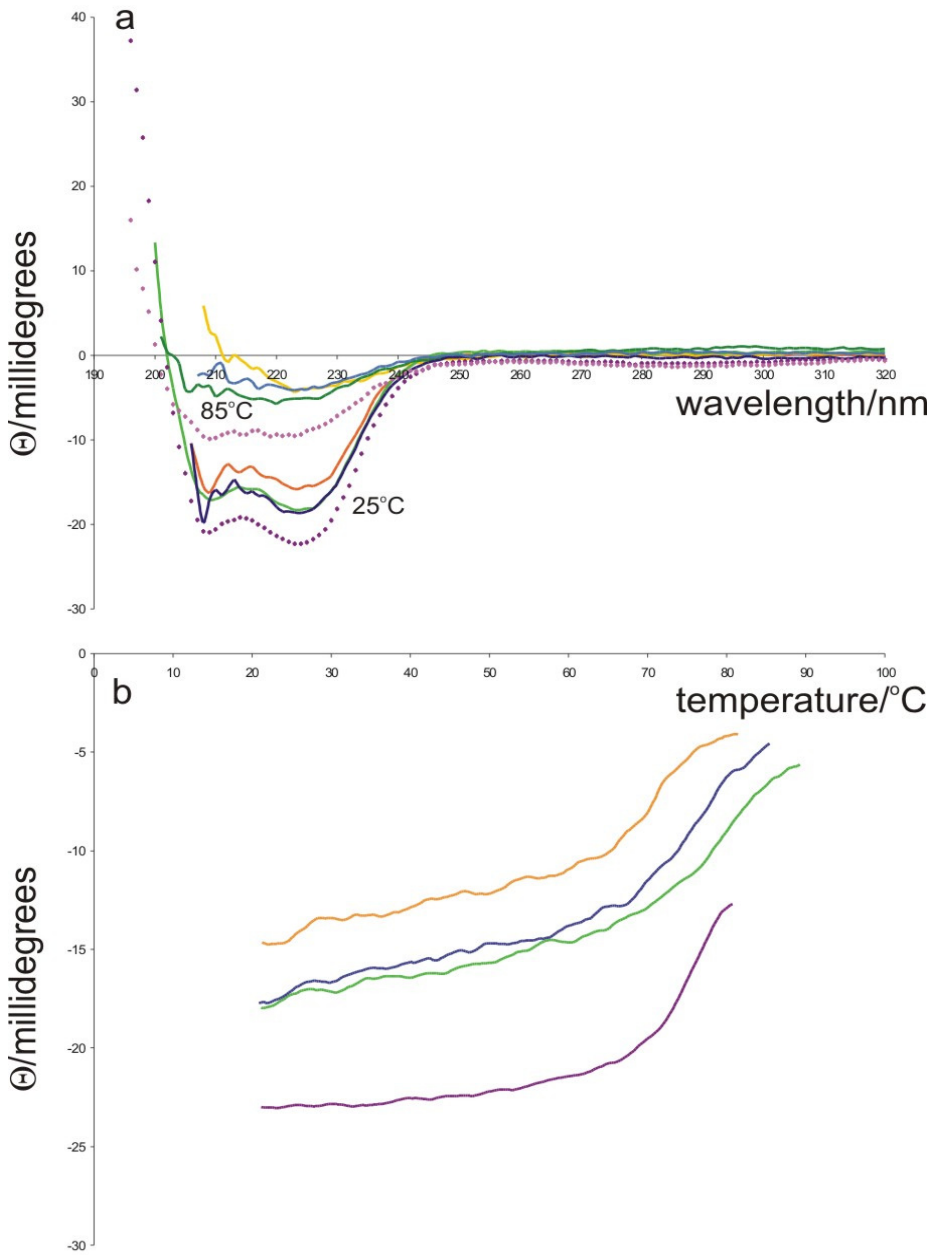


Figure 3.6.30 Far UV difference wavelength scans at 25ºC and 85ºC and thermal denaturation curve for wild-type Vt plus LD peptides

Vt was at 0.5µM, LD peptides at 50µM. A: LD1 plus Vt at 25ºC (dark orange) and 85ºC (light orange); LD2 plus Vt at 25ºC (light green) and 85ºC (dark green); LD4 plus Vt at 25ºC (dark blue) and 85ºC (light blue; Vt alone at 25ºC (purple circle) and 85ºC (pink circle). B: melting curves for Vt plus LD1 (orange), LD2 (green), LD4 (blue) and Vt only (purple).

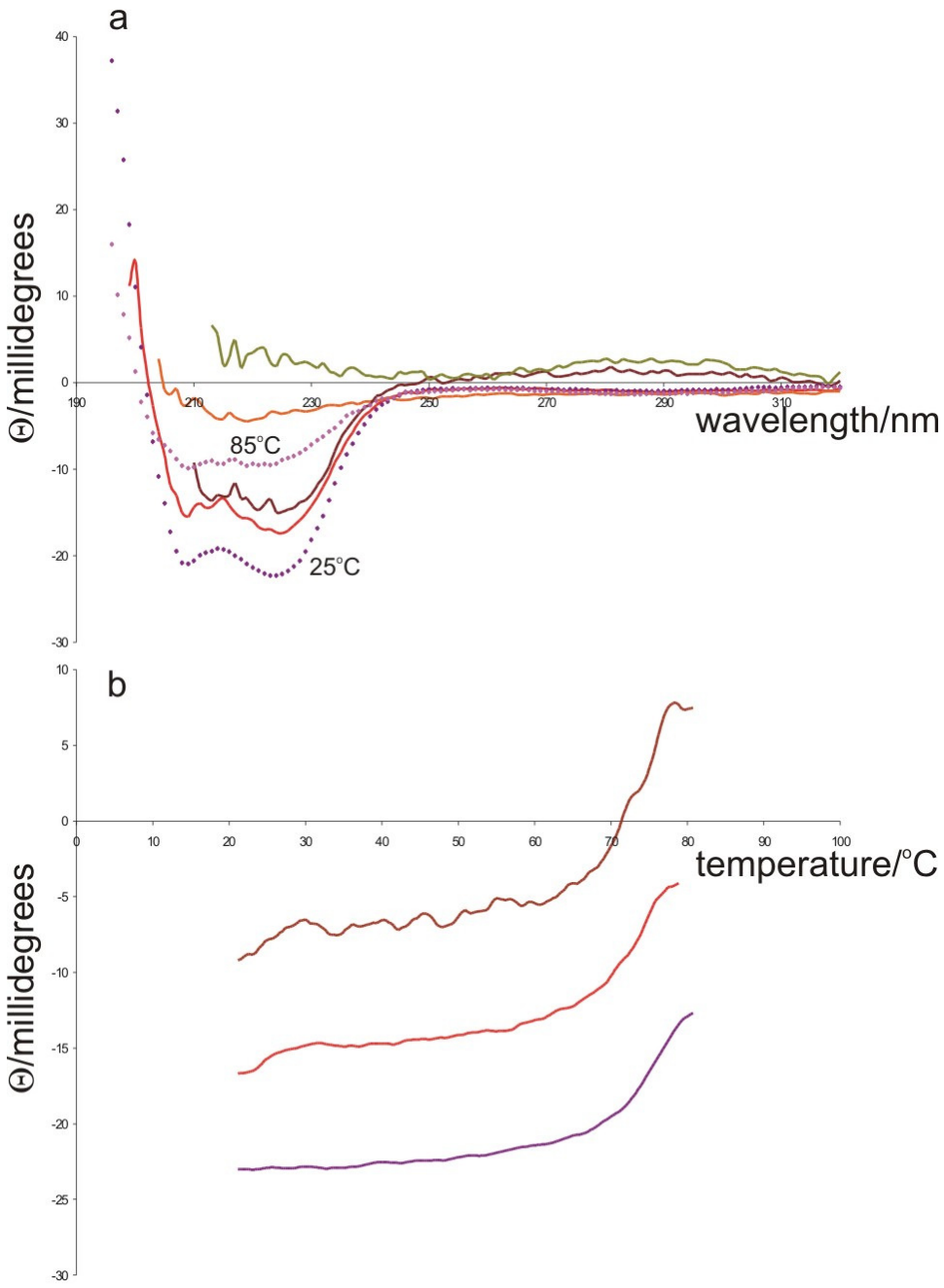


Figure 3.6.31 Far UV difference wavelength scans at 25ºC and 85ºC and thermal denaturation curve for wild-type Vt plus His-LD1/LD2

Vt was at 0.5µM. A: 1:1Vt:His-LD1/LD2 at 25ºC (red) and 85ºC (orange); 1:4 Vt:His-LD1/LD2 at 25ºC (brown) and 85ºC (khaki); Vt alone at 25ºC (purple circle) and 85ºC (pink circle). B: melting curves for 1:1 Vt:His-LD1/LD2 (red), 1:4 Vt:His-LD1/LD2 (brown) and Vt only (purple).

Table 3.6.6 Melting temperatures for all experiments

	Tm	% resonances remain after 2hours in D_2O
Wt Vt	74.9 ± 0.19	53
Wt Vt + LD1	69.8 ± 0.30	10
Wt Vt + LD2	78.4 ± 0.46	8
Wt Vt + LD4	75.0 ± 0.44	26
Wt Vt + His-LD1/LD2 1:1	72.8 ± 0.28	nd
Wt Vt + His-LD1/LD2 1:4	72.3 ± 0.38	19

The results for the Tm for all experiments are shown in Table 3.6.6. The high melting temperature for Vt was unexpected. Two other instances of Tm for Vt have been cited in the literature and these are conflicting. Saunders *et al.* (137) calculated the Tm for Vt using DSC and found the Tm to be 63.3ºC. In contrast, Bakolitsa *et al.* (152) show that the Tm is greater than 70ºC. This high melting temperature could reflect the fact that helices are more stable than the overall structure, or it could reflect aggregation of Vt at higher temperatures (278).

Unfortunately these results have not unambiguously answered the question of whether the Vt structural integrity is affected by LD peptides. Table 3.6.6 presents the Tm and deuterium exchange results together. LD4 peptide has the least effect on the D_2O exchange and appears to affect the helicity of Vt to a small degree but does not affect the Tm. LD1 leads to the loss of virtually all resonances in the D_2O spectrum and CD shows that the 2º structure is affected and the Tm is lowered, LD2 also leads to loss of most resonances in the D_2O exchange and the CD results show effects on the 2º structure while the Tm is increased. These results would be consistent with higher order aggregates occurring when paxillin is added to Vt and would also agree with the observation that excess Vt will precipitate when added to paxillin. These data would benefit from being carried out at higher LD peptide excesses.

4 Discussion

The nature of the interaction between Vt and paxillin has been under debate for some time. Vt was shown *in vitro* to interact with paxillin LD1, LD2 and LD4 (5;80). The region on Vt proposed to bind to paxillin was shown by truncation mutagenesis and sequence alignment to reside on helix 3 and 4 of Vt (see Section 1.5, Figure 1.5.2); this region was called the paxillin binding subdomain and consists of two regions, PBS1 and PBS2 (86;110). An alternative prediction for the binding region on Vt was made by Hoellerer *et al.* (6) which comprised of two binding sites on opposite faces on Vt (see Section 1.5, Figure 1.5.5)

This thesis set out to describe the molecular basis of the interaction between Vt and paxillin LD motifs. The original aims were to locate the binding site for paxillin LD motifs on Vt and to understand the specificity, stoichiometry and affinity of this interaction. A comparison of this data would then be completed using data from the interaction between the FAT domain of FAK with paxillin (6) in an attempt to further understand the molecular recognition involved when LD motifs interact with a target sequence, and also to understand how FAK and vinculin might compete for binding of paxillin (4). Since this original aim, the structures of two more paxillin binding partners, Git1 (10) and actopaxin (α-parvin) (11) in complex with LD motifs have been solved and have been included in this comparison.

The interaction between Vt and LD motifs has proven more complex than originally anticipated, but the body of data presented in this thesis has provided several new avenues of investigation which can now be explored.

4.1 Vinculin tail and paxillin can be purified for use in NMR studies

In this study, recombinant wild-type and mutant Vt proteins have been purified from *E.coli* expression to a purity >95% (Section 3.1.1 and 3.3). All purified proteins adopt a similar highly α-helical fold as judged by far-UV circular dichroism measurements and produce monomeric proteins when analysed by size exclusion chromatography (Section 3.1.2. and 3.3). Wild-type Vt and Vt/I997S have subsequently been used throughout this thesis. Buffers containing 50mM arginine and 50mM glutamate aided in the solubility and stability of Vt in solution at high concentrations (Section 3.3.5).

A recombinant paxillin construct, pET-15B-LD1/LD2 , containing LD motifs LD1 and LD2 has also been purified to >95% purity, called His-LD1/LD2. This purification is challenging and is still being refined to produce the best quality protein possible (Section 3.5). At present, the oligomeric state of His-LD1/LD2 is unclear; size exclusion chromatography suggests that it may exist in multiple forms. Circular dichroism suggests that it is predominantly unstructured and this is in agreement with the poor dispersion of proton signals in the ^{1}H-^{15}N-HSQC spectra.

4.2 The Vt/I997S backbone amides have been assigned and the assignments transferred to wild-type Vt

78% of the amino acids in Vt/I997S have been assigned in the ^{1}H-^{15}N-HSQC spectrum. 88% of the α-helical regions have been assigned, many of the missing regions are in extended loops at the N-and C-terminus. It was an especially challenging assignment problem, and was enabled using a multi-faceted approach involving 3D-heteronuclear editing techniques, amino acid selective labelling and triple resonance experiments in conjunction with high levels of perdeuteration. There is evidence in the ^{1}H-^{15}N-HSQC spectrum that Vt has multiple conformations for some backbone amides, as judged by slow exchange effects for some resonances in the spectrum (Section 3.4.7).

Of the Vt/I997S NH assignments, 93% could be transferred to the wild-type spectrum, 78% of these were transferred with a high level of confidence. There are still some unknown resonances that are unique to the wild-type spectrum, and some that have been transferred at a lower confidence level that need to be confirmed by recording triple resonance experiments on wild-type Vt.

4.3 Elucidating the nature of the interaction between Vt and paxillin has proven difficult

4.3.1 Catch-up experiments show an interaction but chemical shift perturbation data is more complex than anticipated

It has been previously observed that Vt and paxillin interact *in vitro* (5;67;79;80;86). GST-pulldowns and catch-up experiments carried out in this thesis confirm this interaction and show that GST-LD1/LD2 will interact with Vt in a manner that is not mediated by unspecific GST binding (Section 3.6.1 and 3.6.2). Based on the catch-up experiment, the interaction between the two appears to be tight, although it was not possible to estimate a dissociation constant.

Chemical shift perturbation studies using NMR have focused on specific LD motifs and the effect that they have on Vt (Section 3.6.3). When peptide mimics of LD1 and LD2 are added to Vt, chemical shift perturbations occurred throughout the entire Vt molecule rather than at discrete sites. These perturbations manifest themselves as intensity decreases as well as positional changes which may suggest that these interactions are in the fast-to-intermediate exchange regime. Due to this fact, adaption of the analysis methodology was required to combine the information from chemical shift changes and intensity changes. The wide distribution of perturbations suggest that there may be a global process affecting Vt in addition to LD binding. In an attempt to separate global and local changes, a threshold was set for chemical shift changes and intensity decreases. Resonances below this threshold were not considered significant for the determination of a LD binding site.

When analysed in this manner, peptides containing LD1 and LD2 cause perturbations that can be attributed to specific regions on Vt, and are consistent with an interaction between Vt and LD1 and LD2. When compared to LD1 and LD2, the effects on the spectrum from LD4 appear very small and were considered to be non-specific. This result is contrary to previous observations where LD4 interacts with Vt (5;80), the data presented here would be consistent with little interaction between Vt and LD4.

When paxillin His-LD1/LD2 is titrated with Vt to a 4-fold excess, significant aggregation is observed after 12 hours and a large precipitate is formed. It is believed that the precipitate consists of Vt but this needs to be confirmed. It is possible that paxillin is catalysing Vt aggregation under these conditions.

4.3.2 Hydrogen-deuterium exchange shows decreased protection against exchange

Hydrogen-deuterium exchange experiments between Vt and LD peptides suggest that addition of LD1 and LD2 increases the exchange rate of amides with D_2O, such that after 2 hours in D_2O, over 90% of resonances have disappeared from the spectra (Section 3.6.3.5). This is contrary to expected and is also different to the result observed with the FAT of FAK with LD2 and LD4 (6).

This suggests that LD peptides affect a global property of Vt and is consistent with the global chemical shift perturbations seen when Vt is titrated with LD1 and LD2 (Section 3.6.3). The reason for this is unknown, but it could be that the structural stability or oligomerisation state of Vt is altered leading to increased exchange.

4.3.3 Thermal denaturation is inconclusive but shows that Vt has a high melting temperature

Far-UV circular dichroism scans show that thermal denaturation reduces the helicity of Vt, and that this is possibly enhanced with the addition of paxillin. Thermal denaturation experiments show that Vt has a high melting temperature (75ºC), and that LD1, LD2 and

His-LD1/LD2 may alter this but not as much as expected (Section 3.6.5). This high melting temperature could be down to stability of Vt helices, or due to higher order aggregates forming as Vt denatures which could mask the effects of paxillin.

4.3.4 There is a potential interaction site on Vt face 3-4

It is clear from the data presented in this thesis that the interaction between Vt and paxillin is not a simple one. Chemical shift perturbation data suggests that there are underlying effects on Vt other than a simple face-face binding event. It is possible that Vt is becoming destabilised or unfolding when LD motifs are present, and it is possible that aggregation of Vt may be occurring. With these caveats in place, the results from the titrations suggest a potential binding interface on Vt.

The binding interface is proposed to reside on Vt face 3-4 (Figure 4.3.1 for LD1, Figure 4.3.2 for LD2 and Figure 4.3.3 for His-LD1/LD2). There are two clusters of affected residues that consist of a hydrophobic residue surrounded by polar and charged residues. It is interesting that when mapped onto the molecule viewed as the surface potential, two separate regions appear, one centred around I948 and one centred around L982. For the LD1 peptide, I948 is surrounded by R945, Q949, K996 and T1000, and L982 is surrounded by T962, R963, K966, E967, K970 and E986. For LD2, K996 is replaced by K952 and only has K966, K970, E967, R963 surrounding L982. For His-LD1/LD2 the affected residues include all of the previously stated amino acids plus two unique residues, R978 and T979.

This potential interface covers a large area of the 3-4 face, the size of this area is estimated as 2709Å^2 using the molecular graphics program MOLMOL (269). The same analysis predicts the entire Vt face 3-4 to be 3760Å^2, so the interaction interface covers 72% of this Vt face. The interface predicted by Hoellerer *et al.* (6) (see Figure 1.5.5) is estimated to cover 1960Å^2, which is 52% of the entire Vt face 3-4. The data presented in this thesis would be consistent with a proposed interface that is larger than the predicted data by Hoellerer *et al.* (6).

Assuming that LD motifs bind to this interface, it is difficult to imagine how one LD motif containing 13 residues in a helical conformation would make all of these contacts. Analysis of the LD motif binding to the FAT domain of FAK (6) reveals that the helical peptide is 16.6Å in length, this corresponds to approximately three turns of α-helix. The entire length of the interface proposed here is 33.2Å (from Cα R945 to Cα L982) corresponding to approximately six helical turns of Vt.

It is possible that a single LD peptide binds to this entire region in an extended conformation, alternatively there could be two separate binding sites for LD peptides on face 3-4 of Vt (Figure 4.3.4), one surrounding L982 and the other surrounding I948. An estimation of the length of these two sites individually produce a length of 16.0Å and 16.3Å for site I948 and site L982 respectively. While it is conceivable that a binding site of these dimensions could be occupied by an extended conformation of LD motif, all of the evidence so far indicates that LD motifs adopt a helical structure in the target bound state (6;7;10;11;87). It has been shown previously that increasing the helicity of LD4 increases the affinity of binding to FAT (87), so LD binding to Vt in an extended conformation would constitute an entirely new mode of interaction.

In the remainder of this section the implications of a two site binding model vs. an extended binding model will be considered. In the two site model, the site surrounding L982 will be called site 1 and the site surrounding I948 will be called site 2 (Figure 4.3.4). These are named according to their position in the full length Vt molecule (Figure 4.5.2).

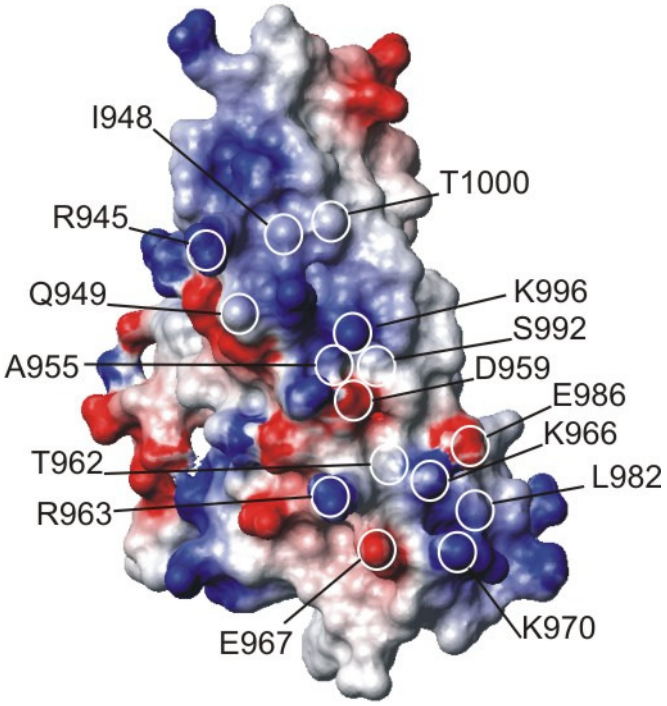


Figure 4.3.1 The proposed binding interface between Vt and LD1 on Vt face 3-4

The surface potential of Vt face 3-4 is shown, blue regions are positively charged and red regions are negatively charged. E967 was previously placed on the Vt 2-3 interface but has been moved to this face in retrospect. To indicate the approximate position of the binding sites the most exposed atoms of residues that show significant perturbations in the NMR spectra are circled and labelled with their respective amino acid name. Figure prepared using MOLMOL (269).

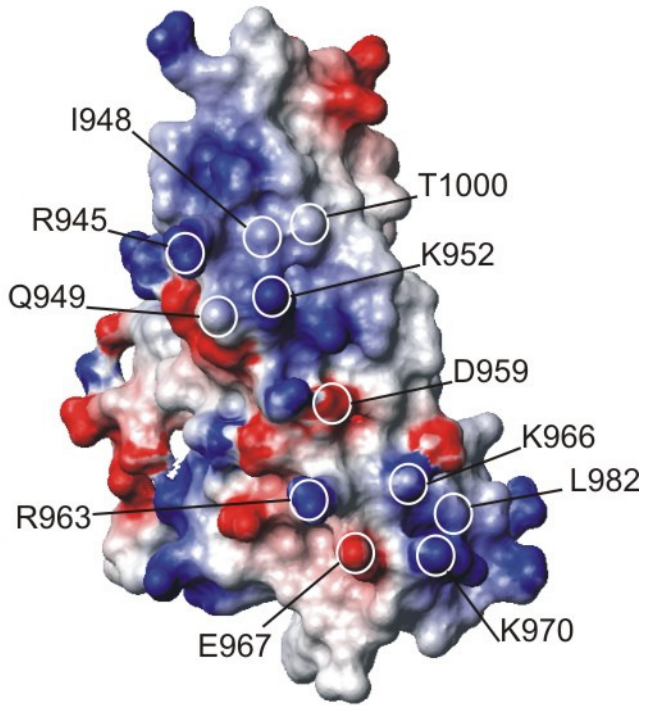


Figure 4.3.2 The proposed binding interface between Vt and LD2 on face 3-4

The surface potential of Vt face 3-4 is shown, blue regions are positively charged and red regions are negatively charged. E967 was previously placed on the Vt 2-3 interface but has been moved to this face in retrospect. To indicate the approximate position of the binding sites the most exposed atoms of residues that show significant perturbations in the NMR spectra are circled and labelled with their respective amino acid name. Figure prepared using MOLMOL (269).

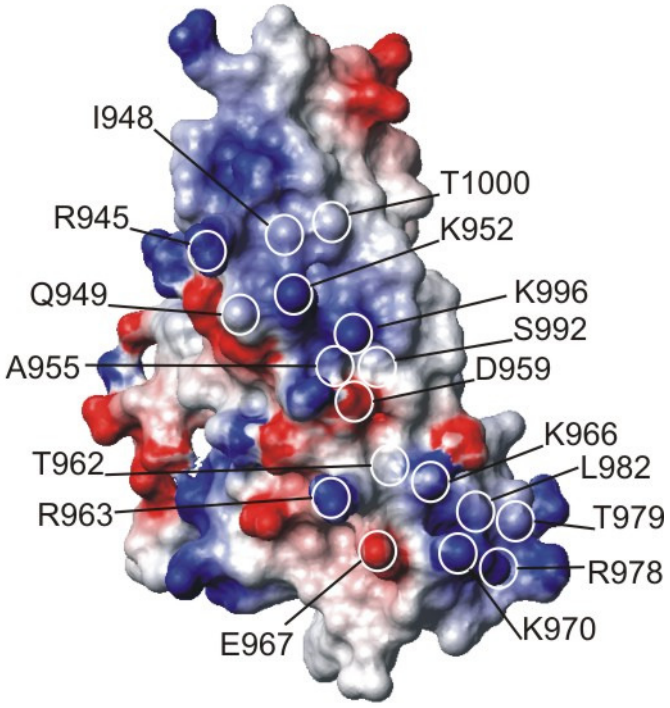


Figure 4.3.3 The proposed binding interface between Vt and His-LD1/LD2 on face 3-4
The surface potential of Vt face 3-4 is shown, blue regions are positively charged and red regions are negatively charged. E967 was previously placed on the Vt 2-3 interface but has been moved to this face in retrospect. To indicate the approximate position of the binding sites the most exposed atoms of residues that show significant perturbations in the NMR spectra are circled and labelled with their respective amino acid name. Figure prepared using MOLMOL (269).

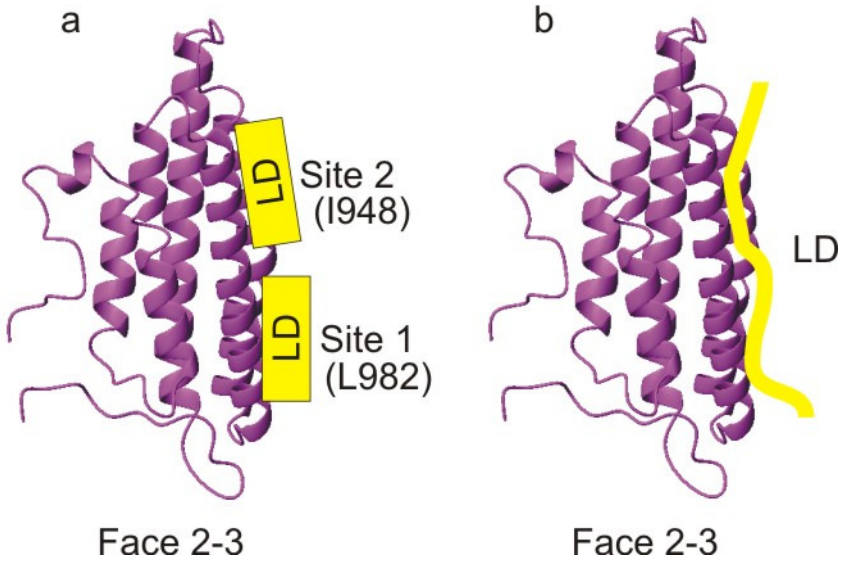


Figure 4.3.4 Two site binding hypothesis vs. extended conformation

(a) Model showing paxillin LD motifs binding to two sites in Vt as α-helices. (b) model showing one LD motif interacting in an extended conformation. Vt is shown from face 2-3 with the binding sites residing on face 3-4. Yellow boxes represent LD motifs in a helical conformation, yellow line represents one LD motif in an extended conformation.

4.3.5 Vt face 5-2 does not provide a convincing binding site

Based on the features determined in FAT binding to LD2 and LD4 (6), it was proposed that Vt face 5-2 may also bind to LD motifs. For this binding surface to become available it was recognised that helix 1 of Vt would need to be displaced (6). The presence of helix 1 on face 5-2 makes the analysis of this region complex. The data presented here suggests that overall the NH chemical shifts changes were small on Vt face 5-2 with LD1, LD2 and LD4 as well as with His-LD1/LD2. However, some chemical shift perturbations above the determined criteria have been observed both on the 5-2 interface and on helix 1. For LD1 there were six and for LD2 there were two affected chemical shifts above the criteria on this face. In contrast, for His-LD1/LD2 there were twelve affected chemical shifts (see Results, Figure 3.6.14 and Discussion Figure 4.4.1). This data would suggest that peptides LD1 and LD2 alone cannot occupy Vt face 5-2.

Figure 4.3.5 shows Vt face 5-2 after removing helix 1 (E880-D917) to show the residues that were perturbed by His-LD1/LD2. Interestingly, the affected residues line the perimeter of the proposed binding site rather than in the center of the Vt 5-2 site. This is consistent with helix 1 being the predominant occupant of this site even in the presence of an excess of His-LD1/LD2. The data presented here does not rule out binding to this face, but there is not enough information to propose binding residues specific to LD motifs. Further information is needed on this interface before a binding site can be proposed or ruled out completely and it may be necessary to create a construct without helix 1 to analyse specific perturbations caused by LD motifs.

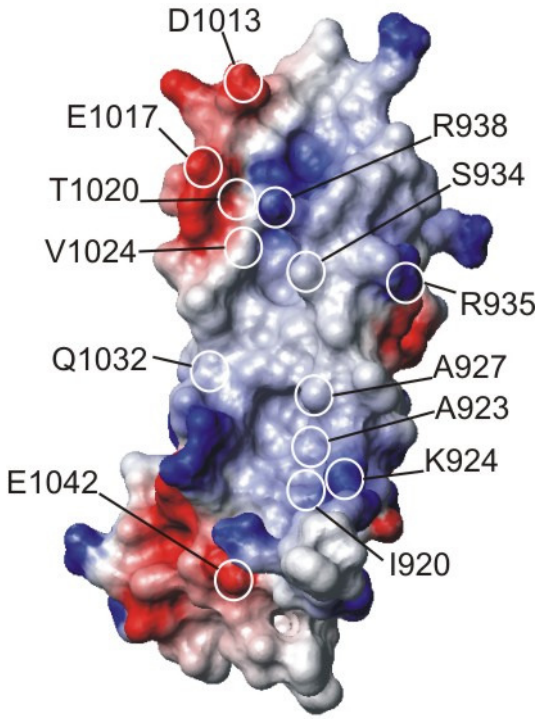


Figure 4.3.5 Chemical shift perturbations mapped onto Vt face 5-2 with helix 1 removed.

The surface potential of Vt face 5-2 is shown, blue regions are positively charged and red regions are negatively charged. Residues E880-D917 have been removed. To indicate the approximate position of the binding sites the most exposed atoms of residues that show significant perturbations in the NMR spectra are circled and labelled with their respective amino acid name. Figure prepared using MOLMOL (269)

4.4 Comparison of Vt-LD interactions with other LD binding proteins

The FAT domain of FAK interacting with LD2 and LD4 provided the first structural insight into the mechanism of LD motif binding (6-9). The knowledge that the structure of Vt was similar to FAT lead to the hypothesis that the LD motif might recognise FAT-like folds as part of the motif recognition mechanism (6). Recently, the structure of the Git1 PBD has been solved (10) which is a four-helix bundle that adopts a similar fold to FAT and Vt, and chemical shift perturbation showed that LD motifs LD2 and LD4 would bind to a single site on one face of the PBD.

Since the initial characterisation, the field has advanced in several ways. FAT is known to have two binding sites, each on opposite sides of the molecule, which can bind simultaneously to LD2 and LD4 (6;9). The affinity appears to be higher on FAT face 1-4 for LD2 and LD4 is suggested to prefer face 2-3 (7;207). It has also been shown that other LD motifs can bind to the FAT domain of FAK (personal communication), but determining the orientation of LD motifs still remains problematic. The structural homologue, Git1 PBD binds LD2 and LD4 but there is only one site on Git1 face 1-4 and the affinity for LD4 is slightly higher than for LD2 but still comparable (10). Finally, the structure of actopaxin (α-parvin) has recently been solved in complex with LD1 and provides an alternative mechanism of binding distinct from the LD-FAT-like fold recognition (11).

The data presented here between Vt and LD1 and LD2 would be consistent with two binding sites on face 3-4 of Vt if the LD peptides bind in a helical fashion. This is the same as face 2-3 of FAT and Git1. Alternatively there may be one site on Vt face 3-4. The affinity and the orientation of LD motifs with Vt still needs to be determined.

A feature that persists throughout the binding sites of FAT, Git1 and Vt is a solvent exposed hydrophobic patch or hydrophobic residue surrounded by charged or polar residues. Figure 4.4.1 shows a comparison of the residues that are involved in the interaction for FAT, Git1 and Vt based on a structural sequence alignment. From this it can

be seen that some of the residue positions are similar, but there is no obvious pattern that arises between the three proteins.

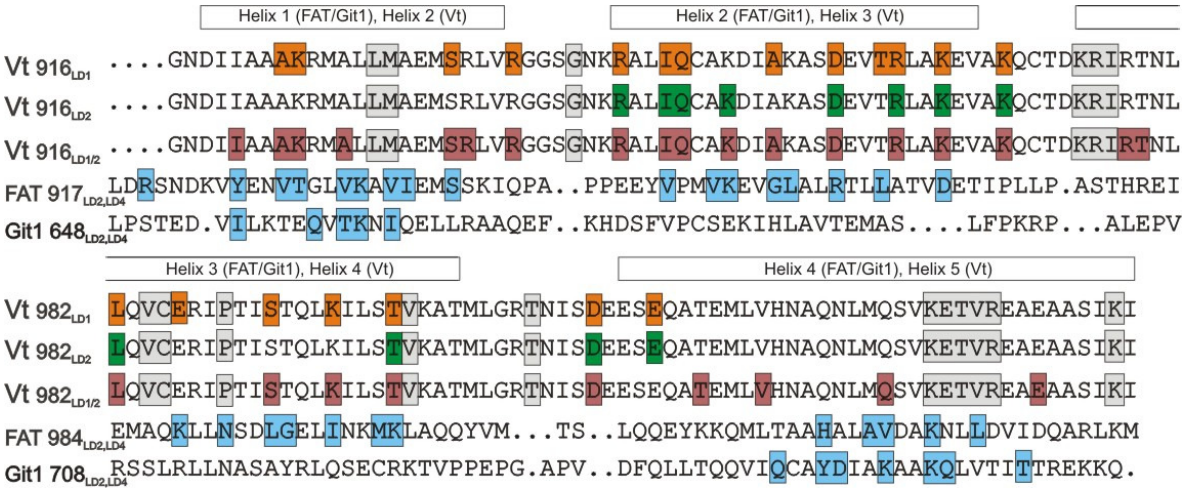


Figure 4.4.1 Sequence alignment based on helical structures of Vt, Git1 PBD and FAT domain of FAK and residues affected by LD motifs

Residues involved in binding LD2 and LD4 for FAT and Git1 shown as blue boxes (6;10), residues in Vt perturbed by LD1, LD2 and His-LD1/LD2 (LD1/2) are shown in orange, green and brown respectively. Grey boxes are residues that are either unassigned or could not be transferred to the wild-type Vt spectrum.

The binding regions for the LD motifs on FAT, Git1 and Vt are shown in Figure 4.4.2 a, c & d. So far the evidence suggests that LD motifs may recognise the FAT fold but the specifics of what determines specificity of motif or binding partner is not clear. An additional insight into LD binding and recognition comes from the recent publication of the structure of actopaxin (α-parvin) in complex with LD1 (11). Actopaxin does not adopt a FAT-like four helix bundle but is a CH domain. It is a known interaction partner of paxillin LD1 and LD4, and the predicted PBS sequence was predicted to reside in residues K273-A290 (77). The new data shows that LD1 does not bind to this PBS but to a different region consisting of helices αL, αA and αF of the CH2 domain (Figure 4.4.2b). This is a new mode of binding distinct from FAT, and it would be interesting to determine whether LD4 binds to actopaxin in the same manner as LD1.

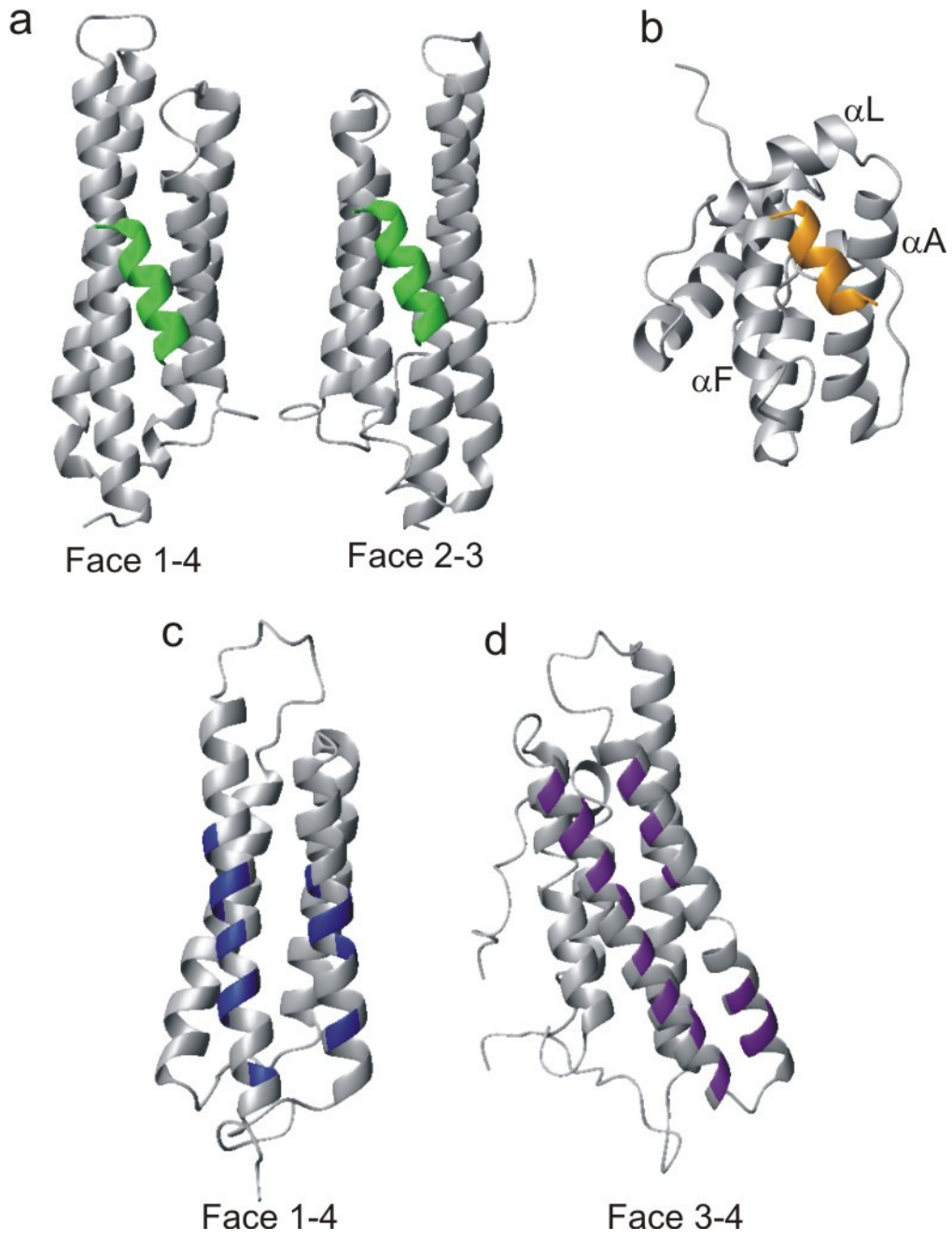


Figure 4.4.2 Structures of the FAT domain (a), actopaxin CH2 domain (b), Git1 PBD (c) and Vt (d) showing LD interaction sites

The FAT domain is shown in complex with LD2 on both faces, actopaxin is shown in complex with LD1, Git1 has the perturbed residues by both LD2 and LD4 in blue, and Vt has the proposed resides perturbed by His-LD1/LD2 mapped in purple.

Searching for similarities that all of the target sequences share shows that the binding interface formed between actopaxin and LD1 contains a hydrophobic surface that interacts with the leucine rich face of LD1, and there are three charged/polar residues. As already mentioned, FAT, Git1 and Vt also share these features. Within the structural families of FAT-like folds the binding interface is formed across two helices, for actopaxin the binding also occurs across helices, despite the actopaxin CH domain being structurally different to the others.

The major difference between all target sequences is the specificity of LD binding and number of binding sites. FAT and Git1 bind to LD2 and LD4 (5-10;80;82;207), actopaxin binds to LD1 and LD4 (11;77), and the data presented here suggests that Vt interacts with LD1 and LD2. Currently there is one known interaction site on one helical interface for actopaxin and Git1; FAT is different as it has two sites on opposite sides of the molecule that can be occupied simultaneously. On the basis of the large area over which chemical shifts were observed, the data presented here suggests that there may be two sites for LD motifs on one face of Vt; alternatively there is one site but LD motifs would need to bind in extended conformation. As stated previously, this would constitute a further mode of interaction between target sequences and LD motifs.

The question still remains as to how an LD motif recognises its prospective binding site, especially now given that the target sequences appear so different? It could be that LD motifs are not specifically seeking their recognition domains but when they are brought close in space by other domain interactions they themselves then interact. In addition to LD motifs, paxillin also contains LIM domains and a polyproline motif that may serve as a docking site for SH3 domains (68). FAK also contains polyproline sequences, so one could imagine a situation where molecules containing SH3 domains bind to paxillin and FAK and bring them close together in space. This would then enable the FAT-LD interaction to ensue and would be consistent with differences between the target sequences whilst LD motifs adopt the same conformation.

This type of mechanism has been suggested to be like 'molecular Velcro'; for example the binding motif of SH3 domains has a structure that will allow it to be 'plugged in' to other target proteins in a relatively fast yet non-specific manner and may be involved in recruiting other protein molecules (279). Remembering that the binary LD-target interaction is potentially part of a much larger signalling complex within FAs, it could be envisaged that other interactions occur that bring the LD motif in close proximity to the target sequence so that they transiently interact to elicit a cellular response. Many of the proteins involved in FA signalling consist of multiple interaction domains, and it is likely that there are multiple interactions occurring that are in dynamic equilibrium with each other (280). Rather than binary interactions occurring one after another, FAs contain non-linear networks of interacting proteins (20). Protein interactions should thus occur when molecules encounter one another in space and time, when local concentration is high and the physiochemical environment correct for the interaction (281)

4.5 Vt-LD interaction: consequences for the cell

4.5.1 FAK competition in apoptosis

Compelling evidence for a cellular role of the paxillin-Vt interaction comes from Subauste *et al.* (4). Here they found that in vinculin null cells, apoptosis could be restored by the addition of Vt (811-1066), and this was shown to be due to modulation of the paxillin-FAK interaction.

This modulation requires that FAK and Vt can compete for the binding of paxillin, so is likely due to the LD motifs in paxillin (4). The data presented in this thesis suggests that LD2 is the motif common to both the FAK and Vt interaction so the competition between FAK and Vt may depend on LD2 in particular. A situation could be imagined where paxillin will bind to either FAK or Vt via LD2, and if LD1 finds a target sequence then Vt signalling continues, and if LD4 finds a target then FAK signalling ensues.

LD4 has previously been shown to be important in the competition between FAK and Git1, expression of paxillin lacking LD4 in paxillin-null cells results in an 11-fold decrease in FA disassembly (51) and mutations that disrupt the LD4 binding site in FAK reduce paxillin binding more than mutations that disrupt the LD2 binding site (9). One could therefore imagine a situation where LD2 is the important motif between FAK and Vt.

Phosphorylation may play a role in competition between FAK and Vt for LD motifs. Phosphorylation of LD4 at serine S272 reduces the binding of FAT by destabilising the helicity of LD4 (87). It has also been shown to affect Git1 binding by increasing the association with LD4 (10;282). The switch is therefore controlled by phosphorylation, when serine S272 is dephosphorylated paxillin prefers to bind to FAK, and once phosphorylated the binding is shifted to more Git1 bound (10). The phosphorylation state of LD2 may affect the binding to FAK and Vt in a similar manner.

For competition to occur, binding affinities need to be comparable. The FAT domain binds two LD motifs on opposite sides. If Vt binds to two LD motifs on one face then the affinity may be similar, and the phosphorylation state could also change the affinities.

4.5.2 Two-site model for Vt-LD interaction: one or two paxillin molecules?

If the prediction proves correct that there are two binding sites in Vt for paxillin, a second question that arises is how many paxillin molecules will bind to Vt. Figure 4.5.1 shows two possible models for paxillin binding, one where one molecule binds simultaneously to both sites and a second where two molecules of paxillin bind via their LD motifs. It will be important to understand the specificity of these two sites to discover whether one is more likely to bind LD1 than LD2 or if they are promiscuous.

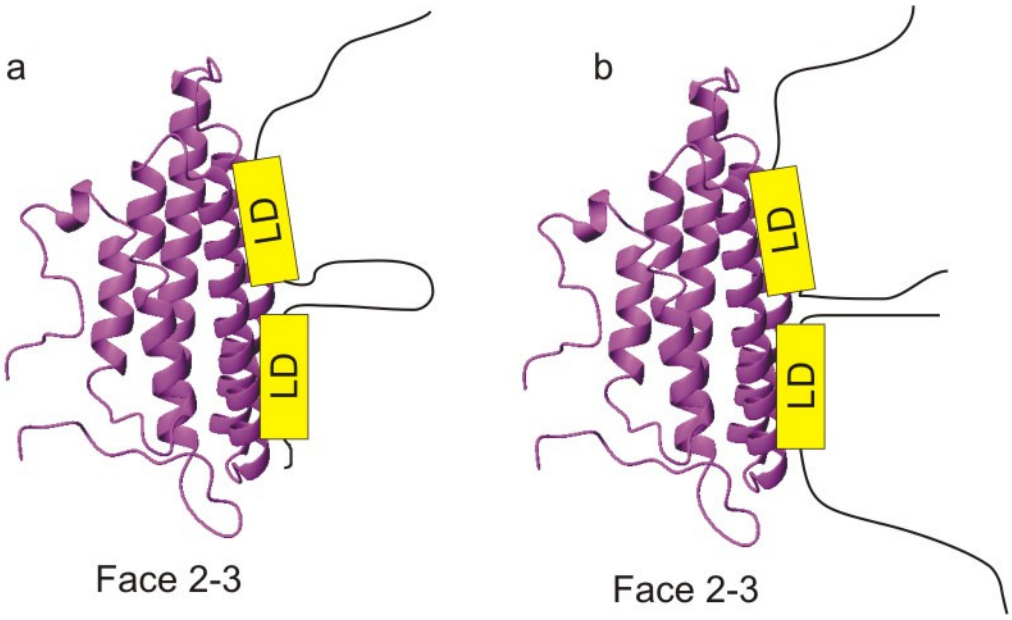


Figure 4.5.1 The two site binding model for Vt-LD motifs showing one paxillin molecule binding (a) or two paxillin molecules binding (b) on Vt face 3-4

Vt is shown from face 2-3 with the binding sites residing on face 3-4. Yellow boxes represent LD motifs in a helical conformation. (a) one paxillin molecule interacts via two LD motifs-here the LD motifs would be different e.g. LD1 and LD2. (b) two separate paxillin molecules bind-in this case the LD motifs could be the same, e.g. LD1 and LD1, or they could be different, e.g. LD1 and LD2.

4.5.3 Full length vinculin and paxillin binding

It is important to remember that Vt is part of a much larger 116kDa protein (the crystal structure can be seen in Figure 1.3.6). Vinculin is an important regulatory protein that exists in a closed, autoinhibited form (152;153) and an open, active form (156;283). It is subject of much debate which binding partners can activate the full length molecule and the question remains as to whether there is a singular molecule that can activate or whether a combinatorial input is required (152;155;157;158;160-162). The major contacts that hold this autoinhibited conformation are between Vt and D1 and Vt and D4 (152-155).

If the proposed binding interface in Vt is true, it provides an interesting idea that paxillin could be involved in activating vinculin. Figure 4.5.2 shows full length vinculin with the

proposed LD1 and LD2 binding sites mapped on. Many of the affected residues are exposed to solvent and not buried in the molecule so it is possible that paxillin could bind to a closed form of vinculin. If the two site model is true then site 1 (L982) is accessible while some of site 2 surrounding I948 is occluded. An attractive idea is that one LD could bind to site 1 which would destabilise the contact between Vt and D4. This may cause site 2 to become available which would weaken the Vt-D1 interaction. This would then allow binding of other ligands such as talin1.

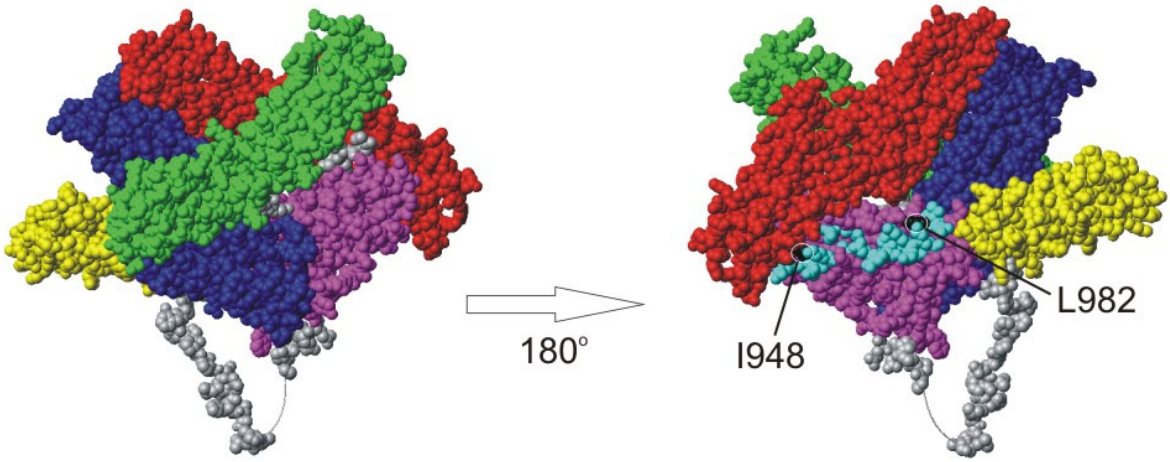


Figure 4.5.2 Full length vinculin showing proposed binding site for LD1 and LD2
The five domains of vinculin are shown , D1 (red), D2 (green), D3 (blue), D4 (yellow) and D5(Vt) (magenta). Paxillin binding site 1 and 2 is shown in cyan on Vt.

4.5.4 The F-actin binding sites overlap with the proposed LD binding sites on Vt

An important interaction within FAs is between vinculin and F-actin (188). Vt binding provides a link between F-actin and the ECM and may play a role in adhesion strengthening and actin crosslinking (41;284). There are two binding sites for F-actin on Vt, and Vt contacts two different F-actin monomers (41). F-actin binding is thought to induce dimerisation of Vt by exposing a cryptic dimerisation site which is thought to be a mechanism for crosslinking actin (204).

The F-actin binding site is deactivated in the autoinhibited form of the full-length molecule and F-actin binds with very low affinity (151). The structure of Vt bound to F-actin reveals that one F-actin binding site is available in the closed vinculin conformation and the second is partially occluded by vinculin D1. Steric hindrance between D1 and F-actin prevents the second binding site from making full contact with actin. Once bound to Vt, crosslinking of actin filaments occurs via activation of the cryptic dimerisation site in Vt, which is believed to occur by changes in the N-terminal strap and C-terminal loop (41).

Figure 4.5.3 shows the binding sites of F-actin in relation to the full length vinculin molecule and the proposed paxillin LD motif binding sites. The F-actin sites partially overlap with the proposed LD binding sites, suggesting that paxillin and F-actin binding to vinculin may be mutually exclusive. This has implications when searching for the vinculin-paxillin interaction, if F-actin and LD binding is mutually exclusive it may not be appropriate to look at mature FAs where there is already a vinculin-F-actin interaction.

It is interesting to note the similarities between F-actin binding to vinculin and the proposed LD binding sites on Vt. Both proteins bind to a similar region on Vt. F-actin consists of two sites, and there is a possibility that there are two sites for LD binding, one of which is partially occluded in the closed conformation. Janssen *et al.* (41) suggest that the partial accessibility of the F-actin binding site in the autoinhibited full-length is evidence towards the combinatorial input model of vinculin activation, where other binding partners act in concert to activate vinculin. Paxillin could work similarly to activate vinculin, and it is possible to imagine two different scenarios, one when F-actin is bound with a set of binding partners such as talin1 which creates one type of FA signalling event such as FA maturation; and a second where paxillin may be bound with other partners to signal a different event such as apoptosis.

Another interesting observation by Janssen *et al.* (41) is the dimerisation interface formed by Vt. The affected regions include the N-terminal strap and C-terminal loop of Vt, and the authors found that this dimerisation only occurred after F-actin binding. The data presented

in this thesis showed that when an excess of Vt was added to His-LD1/LD2, large aggregates occurred-perhaps a similar dimerisation is occurring when paxillin binds to Vt?

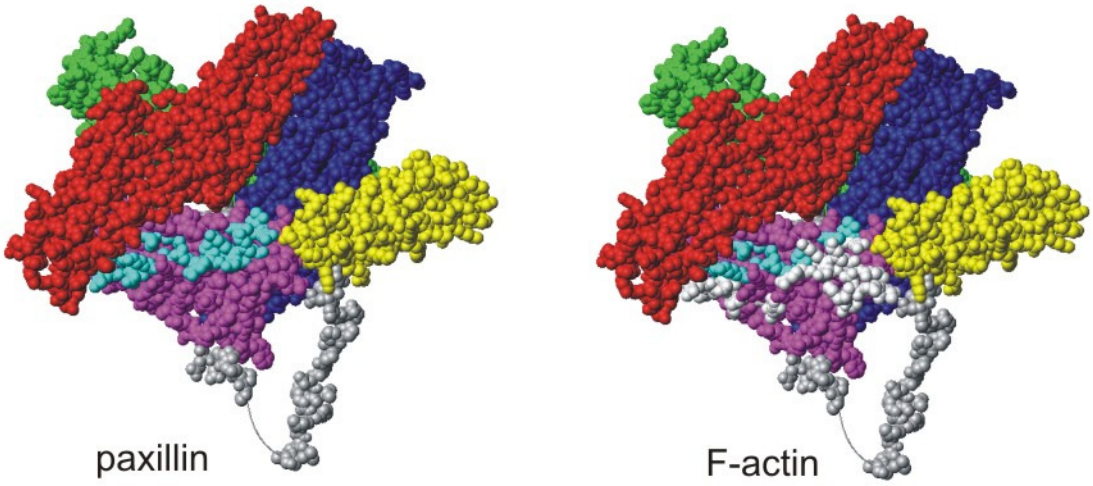


Figure 4.5.3 F-actin binding site overlaps partially with proposed LD binding site
The five domains of vinculin are shown , D1 (red), D2 (green), D3 (blue), D4 (yellow) and D5(Vt) (magenta). Paxillin binding site 1 and 2 is shown in cyan on Vt, and the F-actin binding sites are shown in white in the right hand picture

4.5.5 Could the Vt interaction with LD peptides be non-specific?

In a study by Humphries *et al.* (188) the interaction between Vt and F-actin was shown to be crucial, but it was also noted that co-localisation of Vt and paxillin was not observed. Paxillin co-localised with Vh but not Vt in vinculin enlarged FAs and the authors suggest that there may be very little interaction in cells between Vt and paxillin. As stated in the Introduction (1.4.2) much of the Vt in the cells was associated with actin, and this may explain why the paxillin-vinculin interaction could not be identified as the two binding sites may be mutually exclusive. Similarly, it may be that the cellular situation was not appropriate for detecting this interaction. As other studies have provided compelling evidence for a Vt-LD interaction (4), this is a new insight into FA complexity. As we learn more about FA assembly, turnover and maturation it may become possible to understand the observations in these heterogeneous and dynamic multidomain signalling complexes.

The data presented in this thesis points to a specific interaction *in vitro* between Vt and paxillin, but the nature of that interaction appears complex. It will be of vital importance to test the validity of these binding sites, and then search for them specifically in cells in the appropriate context. The observations of Subauste *et al.* (4) suggest an interaction between Vt and paxillin *in vivo*, but as the full complexity of FA assembly, turnover and maturation is emerging more evidence will be required. The current study has provided hypotheses to investigate based on binary interactions which ultimately will need testing in the complex cellular environment. The following sections explore potential experiments that can now be embarked upon.

4.6 Future experiments to further the knowledge of the Vt-paxillin interaction

4.6.1 *In vitro* experiments

Two potential interaction sites for LD motifs on Vt face 3-4 have been proposed based on chemical shift perturbation data that is less than ideal. It is desirable to establish that this interaction site is specific to LD peptides. If it is shown to be specific then it remains to be determined whether there is one site or two separate sites and whether there is preference for different LD motifs.

An achievable way to study the binding sites would be using paramagnetic labels attached to LD peptides. Similar experiments have been carried out when looking at FAT (7;8) and would involve mutating one residue in the LD peptide to a cysteine and reacting this with MTSSL ((1-oxy-2,2,5,5-tetramethylpyroline-3-methyl)methanethiosulfonate). This reaction produces an oxidized nitroxide which becomes the paramagnetic centre. A previous study (7) made two forms of LD4 where cysteines were either added at the beginning of LD4 (A263Cys) or at the end of the peptide (Ser273Cys) and then were spin labelled. For LD1, Met1Cys and Ser13Cys mutants could be created and for LD2, Asn141Cys and Asn153Cys could be made.

Addition of a paramagnetic spin label increases the relaxation rate of nuclei that are close to the label in a distance dependant manner. Paramagnetic relaxation enhancement (PRE) arises due to the dipolar interaction between a nucleus and unpaired electrons of a paramagnetic centre (285). If a paramagnetic group was added to the LD peptides, residues that are close in space to the interaction site would be broadened, perhaps completely lost from the ^{1}H-^{15}N-HSQC spectrum. These broadened resonances can be assumed to be close to the interaction site and directly affected by the ligand. Any non-specific effects in the spectrum should not be affected by the spin label.

It might be possible to test the two sites hypothesis by introducing a paramagnetic label into one LD peptide and then carrying out competitive chemical shift perturbation experiments by titrating Vt with mixtures of two different LD peptides. The paramagnetic label would reveal the binding site(s) and then the non-spin labelled peptide could be used to out-compete the spin label.

An additional way to look at the interaction sites is to mutate some of the residues predicted to reside in site 1 and site 2 and see whether the ^{1}H-^{15}N-HSQC spectrum is affected in the same manner when titrated by LD peptides. This would rely on the spectrum not being adversely affected by a point mutation. If this was effective, mutagenesis could be used to make one binding site deficient whilst leaving the second site available which could also be useful in determining specificity of LD motifs.

If the Vt face 3-4 interaction interface proves to be true, then it remains to be determined whether the LD peptides interact with Vt in an extended conformation or as a helix. Crystallisation has not been attempted in this study for fears that the affinity between Vt and LD peptides was too low, but perhaps it should be investigated. Filtered 2D-NOESY experiments can be used to filter out proton signals from a doubly labelled protein (Vt) and NOEs can be observed from the unlabelled molecule (LD peptide). Helical peptides will produce NH-NH NOEs and extended peptides will produce strong αH to N (i+1) NOEs (286). NOEs can also be observed between labelled and unlabelled molecules by using filtered/edited 2D-NOESY experiments. These can show NOEs occurring between

residues that interact which could shed light on the specific interaction residues. In practice, as there is only a low population of bound LD peptide compared to the free peptide, searching for NOEs from a small fraction of species would be difficult. It may also be possible to measure the Cα secondary shifts of labelled LD peptides and see if they adopt α-helices but again this may be plagued by the same problem. Determining how these LD motifs interact will be a very difficult problem because of the apparent weak affinity.

Vt face 5-2 was predicted to harbour a binding site for LD peptides (6), but the data does not currently support this. Helix 1 occludes this interface in Vt, so removing helix 1 may aid this investigation if a stable protein can be produced. Based on the F-actin data (41), it would be interesting to investigate whether face 5-2 or the C-terminal loop and N-terminal strap are forming a dimerisation interface for Vt in the presence of paxillin.

To study whether phosphorylation affects binding of LD peptides to Vt, phosphorylated LD peptides can be synthesised. LD4 contains a known phosphorylation site at Ser272 (88), but there are serines in the other peptide mimics that although not predicted could be phosphorylated. These are serine S13 in the LD1 peptide and serine S143 in the LD2 peptide. Chemical shift perturbation experiments can then be repeated for the phosphorylated peptides and effects on the spectrum analysed.

4.6.2 *In vivo* experiments

Ultimately, the interaction between paxillin and Vt needs to be directly observed *in vivo* and it is difficult to construct detailed experiments until the cellular context is understood. If the predicted binding sites in Vt are true and mutants can be made that disrupt binding, then experiments can be repeated as in Subauste *et al.* (4). They found that vinculin null cells could be rescued from apoptosis resistance by Vt, and this was due to the modulation of the FAK-paxillin interaction. A Vt mutant deficient in LD binding should not rescue these cells if it is dependant on the LD-Vt interaction. If the LD2 motif is the important

motif for the competition between FAK and Vt then expression of Vt LD2 mutants should reduce paxillin binding and affect the cells' response to apoptotic stimuli.

The data presented in this thesis is consistent with an overlapping binding site for paxillin and F-actin, which means that the interaction between Vt and paxillin may need to be hunted in cells where Vt is not directly attached to actin filaments. This will require precise analysis of specific types of FAs, and may need to involve looking at the open and autoinhibited forms of vinculin. FRET between Vt and paxillin may prove valuable in deciphering the cellular role of the interaction.

4.7 Conclusions

It is beginning to emerge that LD motifs adopt an α-helical structure when bound to their target and interact via their leucine rich face. However, the target sequences that they bind to seem to be different, and may not be due to structural similarities as previously thought. A recurring feature between all of these interfaces is a hydrophobic element surrounded by polar and charged residues, but how the LD motif specifically determines where it will bind is difficult to determine. It could be that while LD motifs adopt this conformation, the target sequences may be very diverse and the interaction occurs due to the binding partners being brought into close proximity by other domain interactions that are more dominant, for example an SH3 domain binding to polyproline regions or SH2 domains binding to phosphorylated tyrosines. Maybe LD motifs will pave the way in understanding further complexities in molecular recognition and uncover mechanisms of recognition that have not previously been appreciated.

This thesis has opened doors to further understanding of the vinculin-paxillin interaction and has provided new avenues to explore. A binding interface has been proposed and if this is validated will shed new light on the Vt-LD motif interaction. Ultimately, the cellular context needs to be understood; in the meantime studies that strive to understand the molecular details of individual interaction partners are vital for progressing our understanding of focal adhesion signalling.

5 Appendix

5.1 Vt Sequences

Wild-type Vt in pET-15b with amino acids and numbering. Restriction sites are shown.

```
  Bgl II                                  T7 Promoter
 AG ATC TCG ATC CCG CGA AAT TAA TAC GAC TCA CTA TAG GGG

     lac operator                          Xba I
AAT TGT GAG CGG ATA ACA ATT CCC CTC TAG AAA TAA TTT TGT

                     rbs           Nco I
TTA ACT TTA AGA AGG AGA TAT ACC ATG GGC AGC AGC CAT CAT
                                  Met Gly Ser Ser His His

His·Tag                          Thrombin Cleavage Site NdeI
CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG CGC GGC AGC CAT
His His His His Ser Ser Gly Leu Val Pro Arg|Gly Ser His

    879 880 881 882 883 884 885 886 887 888 889 890 891
ATG GAA GAA AAA GAT GAG GAG TTC CCA GAG CAG AAA GCA GGA
Met Glu Glu Lys Asp Glu Glu Phe Pro Glu Gln Lys Ala Gly

892 893 894 895 896 897 898 899 900 901 902 903 904 905
GAA GCT ATT AAT CAG CCC ATG ATG ATG GCT GCT AGG CAG TTG
Glu Ala Ile Asn Gln Pro Met Met Met Ala Ala Arg Gln Leu

906 907 908 909 910 911 912 913 914 915 916 917 918 919
CAT GAC GAG GCT CGG AAA TGG TCT AGC AAG GGT AAC GAC ATC
His Asp Glu Ala Arg Lys Trp Ser Ser Lys Gly Asn Asp Ile

920 921 922 923 924 925 926 927 928 929 930 931 932 933
ATT GCT GCT GCT AAA CGA ATG GCG CTG CTC ATG GCT GAG ATG
Ile Ala Ala Ala Lys Arg Met Ala Leu Leu Met Ala Glu Met

934 935 936 937 938 939 940 941 942 943 944 945 946 947
TCA CGC CTG GTG CGA GGA GGC AGC GGA AAC AAG CGT GCC CTC
Ser Arg Leu Val Arg Gly Gly Ser Gly Asn Lys Arg Ala Leu

948 949 950 951 952 953 954 955 956 957 958 959 960 961
ATC CAG TGT GCA AAA GAT ATT GCT AAG GCA TCG GAT GAA GTC
Ile Gln Cys Ala Lys Asp Ile Ala Lys Ala Ser Asp Glu Val
```

962 963 964 965 966 967 968 969 970 971 972 973 974 975
ACT CGA TTG GCC AAA GAG GTG GCC AAG CAA TGT ACT GAT AAA
Thr Arg Leu Ala Lys Glu Val Ala Lys Gln Cys Thr Asp Lys

976 977 978 979 980 981 982 983 984 985 986 987 988 989
CGC ATT AGA ACA AAC CTC TTA CAG GTC TGT GAG CGA ATC CCA
Arg Ile Arg Thr Asn Leu Leu Gln Val Cys Glu Arg Ile Pro

990 991 992 993 994 995 996 997 998 9991000100110021003
ACC ATC AGC ACG CAG CTC AAA ATT CTT TCC ACA GTG AAA GCT
Thr Ile Ser Thr Gln Leu Lys Ile Leu Ser Thr Val Lys Ala

10041005100610071008100910101011101210131014101510161017
ACC ATG CTG GGC AGG ACT AAC ATC AGC GAT GAA GAA TCA GAA
Thr Met Leu Gly Arg Thr Asn Ile Ser Asp Glu Glu Ser Glu

101810191020102110221023102410251026102710281029103010 31
CAG GCA ACT GAG ATG TTG GTT CAT AAC GCC CAG AAC CTC ATG
Gln Ala Thr Glu Met Leu Val His Asn Ala Gln Asn Leu Met

10321033103410351036103710381039104010411042104310441045
CAG TCT GTG AAG GAA ACT GTG AGA GAA GCT GAA GCA GCA TCC
Gln Ser Val Lys Glu Thr Val Arg Glu Ala Glu Ala Ala Ser

10461047104810491050105110521053105410551056105710581059
ATT AAG ATA AGA ACA GAT GCC GGA TTC ACT CTG CGC TGG GTC
Ile Lys Ile Arg Thr Asp Ala Gly Phe Thr Leu Arg Trp Val

106010611062106310641065 1066 *Bam*H I
AGA AAG ACC CCA TGG TAT CAG TAA CTC GAG GAT CCG GCT GCT
Arg Lys Thr Pro Trp Tyr Gln End Leu Glu Asp Pro Ala Ala

AAC AAA GCC CGA AAG GAA GCT GAG TTG GCT GCT GCC ACC GCT
Asn Lys Ala Arg Lys Glu Ala Glu Leu Ala Ala Ala Thr Ala

*Bpu*1102 I
GAG CAA TAA CTA GCA TAA CCC CTT GGG GCC TCT AAA CGG GTC
Glu Gln End

TTG AGG GGT TTT TTG

Vt Mutants DNA changes
I997S: ATT to AGC
M1022S: ATG to AGC
ΔC: truncated at 1046 ATT to TAG

5.2 Wild-type and mutant amino acid sequences

Wild-type Vt (M879-Q1066) with extra GSH from thrombin cleavage

```
 879 GSHMEEKDEE FPEQKAGEAI NQPMMMAARQ LHDEARKWSS KGNDIIAAAK RMALLMAEMS
 935 RLVRGGSGNK RALIQCAKDI AKASDEVTRL AKEVAKQCTD KRIRTNLLQV CERIPTISTQ
 995 LKILSTVKAT MLGRTNISDE ESEQATEMLV HNAQNLMQSV KETVREAEAA SIKIRTDAGF
1055 TLRWVRKTPW YQ
```

Vt/I997S mutant

```
 879 GSHMEEKDEE FPEQKAGEAI NQPMMMAARQ LHDEARKWSS KGNDIIAAAK RMALLMAEMS
 935 RLVRGGSGNK RALIQCAKDI AKASDEVTRL AKEVAKQCTD KRIRTNLLQV CERIPTISTQ
 995 LKSLSTVKAT MLGRTNISDE ESEQATEMLV HNAQNLMQSV KETVREAEAA SIKIRTDAGF
1055 TLRWVRKTPW YQ
```

Vt/M1022S mutant

```
 879 GSHMEEKDEE FPEQKAGEAI NQPMMMAARQ LHDEARKWSS KGNDIIAAAK RMALLMAEMS
 935 RLVRGGSGNK RALIQCAKDI AKASDEVTRL AKEVAKQCTD KRIRTNLLQV CERIPTISTQ
 995 LKILSTVKAT MLGRTNISDE ESEQATESLV HNAQNLMQSV KETVREAEAA SIKIRTDAGF
1055 TLRWVRKTPW YQ
```

Vt/ΔC (879-1045)

```
 879 GSHMEEKDEE FPEQKAGEAI NQPMMMAARQ LHDEARKWSS KGNDIIAAAK RMALLMAEMS
 935 RLVRGGSGNK RALIQCAKDI AKASDEVTRL AKEVAKQCTD KRIRTNLLQV CERIPTISTQ
 995 LKILSTVKAT MLGRTNISDE ESEQATEMLV HNAQNLMQSV KETVREAEAA S
```

5.3 Primers for site directed mutagenesis of Vt.

```
M1022S
For:CAG AAC AGG CAA CTG AGA GCT TGG TTC ATA ACG CCC AGA ACC  (42 bp)
Rev:GGT TCT GGG CGT TAT GAA CCA AGC TCT CAG TTG CCT GTT CTG  (42 bp)

I997S
For:CCA TCA GCA CGC AGC TCA AAA GCC TTT CCA CAG TGA AAG CTA CC(44 bp)
Rev:GGT AGC TTT CAC TGT GGA AAG GCT TTT GAG CTG CGT GCT GAT GG(44 bp)

C-term
For:GAG AGA AGC TGA AGC AGC ATC CTA GAA GAT AAG AAC AGA TGC CGG(45bp)
Rev:CCG GCA TCT GTT CTT ATC TTC TAG GAT GCT GCT TCA GCT TCT CTC(45bp)

All oligos are presented in the 5' to 3' direction.

M1022SF: 72.7°C
M1022SR: 72.7°C
I997SF:  72.9°C
I997SR:  72.9°C
Cterm:   70.7°C
```

5.4 Paxillin Sequences

His-LD1/LD2 in pET-15b vector. His tag is not removed so protein has potential to start at the Met at the *NcoI* site.

DNA sequence

atg ggc agc agc cat cat cat cat cat cac agc agc ggc ctg gtg
ccg cgc ggc agc cat **atg** gac gac ctc gac gcc ctg ctg gcg gac
ttg gag tct acc acc tcc cac atc tcc aaa cgg cct gtg ttc ttg
tcg gag gag acc ccc tac tca tac cca act gga aac cac aca tac
cag gag att gcc gtg cca ccc ccc gtc ccc cca ccc ccg tcc agc
gag gcc ctc aat ggc aca atc ctt gac ccc tta gac cag tgg cag
ccc agc agc tcc cga ttc atc cac cag cag cct cag tcc tca tca
cct gtg tac ggc tcc agt gcc aaa act tcc agt gtc tcc aac cct
cag gac agt gtt ggc tct ccg tgc tcc cga gtg ggt gag gag gag
cac gtc tac agc ttc ccc aac aag cag aaa tca gct gag tct tca
ccc acc gta atg agc acg tcc ctg ggc agc aac ctt tct gaa ctc
gac cgc ctg ctg ctg gaa ctg aac gct gta cag cat aac ccg cca
gga tcc ggc tgc taa

Protein sequence

M G S S H H H H H H S S G L V P R G S H **M** D D **L D A L L A D L** E S T T S H I S K
R P V F L S E E T P Y S Y P T G N H T Y Q E I A V P P P V P P P P S S E A L N G T I
L D P L D Q W Q P S S S R F I H Q Q P Q S S S P V Y G S S A K T S S V S N P Q D S
V G S P C S R V G E E E H V Y S F P N K Q K S A E S S P T V M S T S L G S N L S E
L D R L L L E L N A V Q H N P P G S G C **Stop**

5.5 Chemical shift table for Vt/I997S assignment

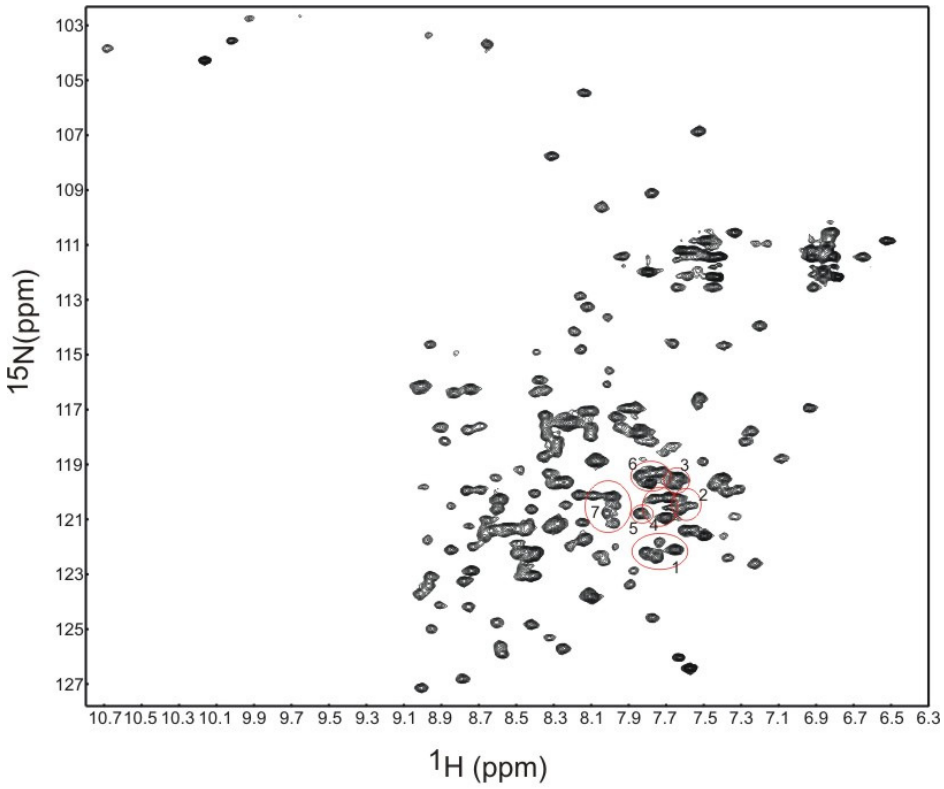
Residue type	Residue number	δHN (ppm)	δCα (ppm)	δCβ (ppm)	δNH (ppm)
Phe	885	8.71	55.86	38.93	127.02
Gln	888	8.67	55.60	28.30	123.15
Lys	889	8.94	54.62	32.71	127.38
Ala	890	8.51	53.28	17.08	126.14
Met	898	8.69	60.01	33.67	118.00
Met	899	7.90	57.80	30.77	120.36
Met	900	8.80	57.76	-	118.34
Ala	901	7.73	54.36	16.97	122.37
Ala	902	8.06	54.90	17.15	121.94
Arg	903	8.61	58.06	28.58	120.23
Gln	904	8.02	58.42	27.54	117.27
Leu	905	7.57	57.77	-	119.96
His	906	8.68	62.31	-	121.44
Asp	907	8.94	56.78	39.70	116.45
Glu	908	7.20	57.79	29.28	118.47
Ala	909	8.28	54.42	17.95	119.01
Ser	914	9.54	58.94	63.92	122.79
Lys	915	8.26	57.35	29.86	125.72
Gly	916	8.88	45.89	-	114.91
Asn	917	7.78	51.56	36.42	118.32
Asp	918	8.86	56.17	40.43	124.27
Ile	919	7.49	66.19	36.15	120.90
Ile	920	7.32	60.33	34.20	121.90
Ala	921	7.94	55.37	17.96	121.03
Ala	922	7.76	54.68	17.40	121.04
Ala	923	8.56	55.07	17.37	121.63
Lys	924	8.32	60.39	32.31	116.63
Arg	925	7.58	58.95	29.67	118.47
Met	926	8.83	59.32	33.04	117.89
Ala	927	8.19	55.30	17.68	121.40
Leu	928	7.69	57.89	39.61	118.20
Ala	931	8.01	54.63	17.18	123.92
Glu	932	7.52	58.64	27.46	121.03
Met	933	8.91	60.18	31.70	123.73
Ser	934	8.05	61.75	-	113.48

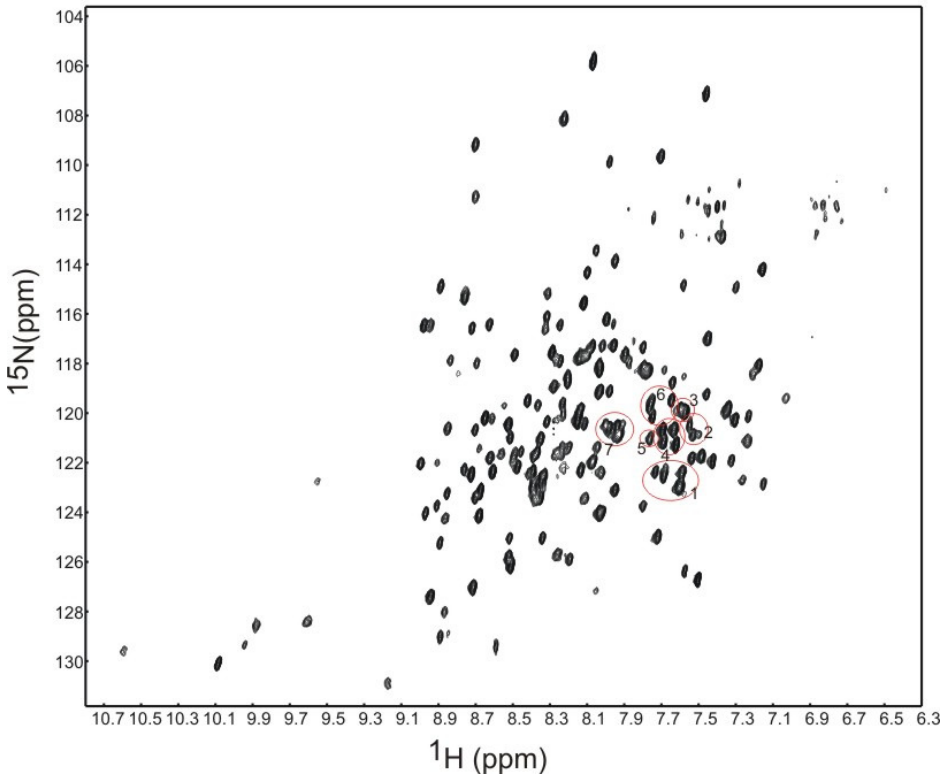
Residue type	Residue number	δHN (ppm)	δCα (ppm)	δCβ (ppm)	δNH (ppm)
Arg	935	7.23	57.63	30.16	120.17
Leu	936	8.45	56.92	43.30	121.55
Val	937	7.98	63.59	30.95	109.84
Arg	938	7.17	53.71	28.31	118.08
Gly	939	7.46	44.18	-	107.17
Gly	940	8.06	43.70	-	105.88
Ser	941	8.76	60.62	62.23	115.35
Asn	943	8.03	52.46	38.56	118.27
Lys	944	7.48	60.85	31.76	121.73
Arg	945	8.29	59.02	28.29	117.54
Ala	946	8.39	54.14	17.88	122.37
Leu	947	8.37	58.26	40.35	121.58
Ile	948	7.45	64.98	37.34	119.24
Gln	949	8.23	57.88	27.48	119.62
Cys	950	8.32	62.38	26.23	120.35
Ala	951	7.75	55.10	17.64	119.42
Lys	952	7.98	60.01	31.29	119.12
Asp	953	8.70	57.43	39.82	123.43
Ile	954	8.89	66.08	36.90	125.22
Ala	955	8.34	54.70	16.75	125.03
Lys	956	8.11	59.17	31.34	120.36
Ala	957	7.69	54.67	16.62	122.55
Ser	958	8.63	62.48	-	116.45
Asp	959	8.19	57.02	39.85	125.90
Glu	960	7.53	58.85	28.57	121.82
Val	961	8.03	67.27	30.73	118.03
Thr	962	8.10	67.33	-	114.29
Arg	963	8.04	59.48	29.30	121.41
Leu	964	8.04	57.10	41.97	119.19
Ala	965	8.76	55.12	18.13	122.26
Lys	966	8.14	59.96	31.29	117.83
Glu	967	7.36	58.49	27.97	119.94
Val	968	8.00	65.69	30.66	120.43
Ala	969	8.39	55.38	16.99	121.73
Lys	970	7.58	58.56	31.66	114.83
Gln	971	7.30	54.44	28.33	114.97
Cys	972	7.55	58.52	27.48	120.63
Thr	973	8.85	62.59	68.47	120.63
Asp	974	7.72	53.06	43.05	124.97

Residue type	Residue number	δHN (ppm)	δCα (ppm)	δCβ (ppm)	δNH (ppm)
Arg	978	8.42	60.23	30.14	119.48
Thr	979	7.95	66.48	68.42	113.89
Asn	980	8.07	55.82	37.10	121.99
Leu	981	8.51	57.27	40.98	121.01
Leu	982	8.28	57.69	40.05	118.92
Gln	983	8.12	58.72	27.59	117.75
Glu	986	7.76	58.03	28.68	119.82
Arg	987	7.45	57.30	30.25	116.99
Ile	988	7.64	67.92	34.38	118.77
Thr	990	7.15	65.61	68.22	114.21
Ile	991	7.95	64.65	37.53	123.08
Ser	992	8.98	61.91	-	116.47
Thr	993	8.15	67.01	68.05	120.13
Gln	994	7.61	58.28	27.19	122.97
Leu	995	8.62	58.45	40.20	121.85
Lys	996	7.98	60.48	30.87	120.64
Ser	997	7.95	61.68	-	117.28
Leu	998	8.99	57.03	41.72	122.05
Ser	999	8.49	62.95	62.09	117.67
Thr	1000	7.60	66.70	67.86	119.80
Lys	1002	8.13	57.20	30.51	119.82
Ala	1003	8.52	54.75	18.07	120.46
Thr	1004	7.70	64.53	69.09	109.69
Met	1005	7.64	56.57	32.70	119.42
Leu	1006	7.30	56.13	40.56	120.28
Gly	1007	8.70	45.15	-	109.20
Arg	1008	7.69	55.63	29.87	120.69
Asn	1010	8.36	52.88	37.14	119.69
Ile	1011	7.69	59.57	38.01	118.84
Ser	1012	8.72	57.59	63.66	122.48
Asp	1013	8.68	56.20	39.91	124.12
Glu	1014	8.65	59.39	28.47	120.20
Glu	1015	7.93	58.81	28.81	120.57
Ser	1016	8.12	60.29	62.53	115.57
Glu	1017	8.61	58.62	28.41	122.40
Gln	1018	8.15	58.54	27.49	120.35
Ala	1019	7.69	54.76	17.58	121.16
Thr	1020	7.99	66.89	67.78	116.20
Glu	1021	8.47	59.33	28.41	122.17

Residue type	Residue number	δHN (ppm)	δCα (ppm)	δCβ (ppm)	δNH (ppm)
Met	1022	8.07	58.16	32.19	117.30
Leu	1023	7.79	58.61	41.19	123.74
Val	1024	8.70	66.73	30.64	120.68
His	1025	8.24	58.54	28.98	117.84
Asn	1026	7.80	57.40	39.72	118.18
Ala	1027	8.85	55.79	18.42	123.22
Gln	1028	8.72	59.12	27.81	116.59
Asn	1029	8.03	55.31	37.12	119.01
Leu	1030	8.97	58.02	40.22	124.04
Met	1031	8.25	58.30	33.96	116.49
Gln	1032	8.20	58.55	27.11	118.74
Ser	1033	8.32	62.56	-	116.11
Val	1034	8.52	67.04	30.12	125.02
Glu	1040	8.54	57.36	28.58	119.69
Ala	1041	8.90	54.49	16.22	121.98
Glu	1042	7.26	59.28	25.62	122.65
Ala	1043	7.15	54.41	16.85	122.82
Ala	1044	8.20	53.28	18.90	118.50
Ser	1045	7.37	60.78	62.92	112.96
Ile	1046	7.03	62.00	37.07	119.42
Ile	1048	7.34	59.81	39.14	119.82
Arg	1049	8.52	56.32	30.32	125.80
Thr	1050	8.31	63.00	68.63	115.24
Asp	1051	8.35	53.35	39.34	121.05
Ala	1052	7.59	52.68	18.68	122.38
Gly	1053	8.22	44.88	-	108.17
Phe	1054	7.63	57.65	39.77	120.66
Thr	1055	7.77	60.05	70.78	118.31
Leu	1056	8.02	53.57	43.01	122.38
Arg	1057	7.23	54.84	30.50	121.15
Val	1059	9.60	61.31	34.45	128.45
Arg	1060	8.89	54.94	29.06	128.99
Lys	1061	8.11	55.81	31.90	123.45
Thr	1062	8.09	59.18	69.01	117.66

5.6 Regions that differ between the wild-type ^{1}H-^{15}N HSQC spectrum and the I997S spectrum





5.7 All assigned wild-type Vt residues affected by chemical shift perturbation experiments.

Criteria set: Percentage intensity decreases ≥50% OR chemical shift differences in the proton and nitrogen dimension of 0.03ppm OR 0.2ppm. Red: specific to His-LD1/LD2 (compared to peptides LD1 and LD2); orange specific to LD1 peptide (compared to LD2 peptide); green specific to LD2 peptide (compared to LD1 peptide)

LD1	LD2	His-LD1/LD2
885	890	889
889	898	898
898	903	900
902	905	902
904	907	903
905	908	904
907	909	905
908	919	907
909	933	908
922	937	909
923	939	918
924	940	919
926	941	920
931	945	921
934	946	922
938	947	923
940	948	924
941	949	926
943	950	927
945	952	932
946	958	933
947	959	934
948	961	935
949	963	937
950	965	938
951	966	939
953	967	943
955	968	944
957	969	945

LD1	LD2	His-LD1/LD2
959	970	947
960	973	948
962	974	949
963	982	950
964	990	952
965	994	953
966	1000	954
967	1003	955
968	1004	957
969	1005	958
970	1006	959
971	1012	960
974	1013	961
981	1014	962
982	1015	963
986	1017	965
987	1023	966
990	1025	967
992	1026	968
994	1027	969
996	1033	970
1000	1049	971
1003	1050	972
1004	1051	974
1005	1056	978
1006	1057	979
1008	1059	981
1012		982
1013		987

LD1	LD2	His-LD1/LD2
1014		990
1015		991
1017		992
1023		994
1025		996
1029		1000
1030		1003
1033		1004
1040		1005
1050		1006
1057		1012
1059		1013
		1014
		1015
		1020
		1023
		1024
		1027
		1030
		1032
		1033
		1034
		1042
		1043
		1044
		1045
		1049

6 Glossary

A_{260}	absorbance at 260nm	for measuring DNA absorbance
A_{280}	absorbance at 280nm	for measuring protein absorbance from aromatic amino acids
Akt1	protein kinase B	serine/threonine protein kinase involved in promoting cell survival
ANK	ankyrin	protein recognition domain-recognition mechanism unclear
Arf	ADP-ribosylation factor	small GTPase involved in actin remodelling
Arp2/3	actin-related protein 2/3	part of the Arp2/3 complex involved in nucleation of actin at the leading edge of cells
ASAP1	Arf GAP containing SH3, ANK repeat and PH domains	PIP_2 dependant Arf GTPase-activating protein involved in actin cytoskeleton regulation
CH	calponin homology	recognition domain that binds to actin (CH1 and CH2) CH3 domains may be regulatory and may not bind actin
Crk	C10-regulator of kinase	adaptor protein within FAs
CrkL	Crk-Like	adaptor protein within FAs
CSA	chemical shift anisotropy	a local magnetic field caused by electron currents in the nucleus in response to the external magnetic field that causes relaxation
Csk	C-terminal Src kinase	tyrosine kinase that inhibits Src family PTKs by phosphorylating near the C-termini
CV	column volume	
δ	chemical shift	The frequency of absorption for a nucleus of interest relative to the frequency of absorption of a molecular standard
DOCK180	180-kDa protein downstream of Crk	guanine nucleotide exchange factor for small GTPases Rac1 and Cdc42
E6	E6 oncogene of human papilloma virus	adapters that change the function of cellular regulatory proteins
ECM	extracellular matrix	extracellular part of animal tissue containing proteins, proteoglycans and polysaccharides
EDTA	ethylenediaminetetraacetic acid	chelating agent
ELMO	engulfment and cell motility	regulator of DOCK180, the Dock180/ELMO complex acts as a bipartite GEF for Rac
ERK	extracellular signal-regulated kinases (MAPK)	serine/threonine protein kinase

ERK1	extracellular signal-regulated kinase 1 (MAPK3)	serine/threonine kinase (classical MAP kinases)
ERK2	extracellular signal-regulated kinase 2 (MAPK1)	serine/threonine kinase (classical MAP kinases)
FA	focal adhesion	protein assemblies at the cell membrane linking cells to the ECM
FAK	focal adhesion kinase	non-receptor tyrosine kinase and adaptor protein
FAT	focal adhesion targeting	domain of FAK necessary for localisation of FAK to FAs
FB	fibrillar adhesion	elongated mature form of adhesion between the cell and the ECM
FERM	band 4.1 and ezrin, radixin, and moesin homology	domain found in many FA proteins, can bind to phospholipids and cytoskeletal proteins
FRAP	fluorescence recovery after photobleaching	optical technique that allows study of diffusion of fluorescently tagged molecules
FRET	förster (fluorescence) resonance energy transfer	energy transfer between two chromophores
FX	focal complex	early adhesion between the cell and the ECM
GAP	GTPase-activating protein	can bind to activated G-proteins and stimulate their GTPase activity. GEFs reverse this
GDP	guanosine diphosphate	product of GTP dephosphorylation in signal transduction
GEF	guanine nucleotide exchange factors	catalyses the removal of GDP and replacement with GTP
Git1	G protein-coupled receptor kinase interacting protein 1	Arf-GTPase-activating proteins and adaptor proteins
GRAF	GAP for Rho associated with focal adhesion kinase	GTPase-activating protein for Rho
Grb	growth factor receptor-bound protein	adaptor protein within FAs
GST	glutathione-S-transferase	affinity tag for protein purification (GST is an enzyme that catalyses conjugation of reduced glutathione to electrophillic centres)
GTP	guanosine-5'-triphosphate	purine nucleotide required in signal transduction where it is converted to GDP
HIS	histidine tag	affinity tag consisting of histidine. Binds to nickel
HSQC	heteronuclear single quantum correlation	a two dimensional NMR experiment

HTI	head-tail interaction	interaction that holds vinculin in its autoinhibited conformation
ILK	Integrin-linked kinase	serine/threonine protein kinase that can bind to B-integrins
INEPT	insensitive nuclei enhanced by polarisation transfer	method for transferring magnetisation to a sensitive nucleus such as protons to an insensitive nuclei such as nitrogen
IpaA	invasion plasmid antigen A	critical for *Shigella* entry and can bind to vinculin
IPTG	Isopropyl ß-D-1-thiogalactopyranoside	molecular mimic of allolactose that induces transcription from the lac operon
JNK	c-Jun N-terminal kinase	MAP kinase responsive to stress stimuli
K_d	dissociation constant	[P][L]/[PL]
kDa	kilodaltons	unit of molecular mass
K_{ex}	exchange rate	in NMR chemical exchange regimes
kinase	phosphorylation enzyme	transfers phosphate groups from donors such as ATP and GTP to target molecules
L	ligand	substance that can form a complex with another molecule
LB	luria broth	nutritionally rich medium used to grow bacteria
LD	leucine/aspartate	motif with consensus LDxLLxxL that can bind to signalling molecules
LIM	lin11, isl-1 & mec-3	double zinc finger motif for protein interactions
MAPK	Mitogen-activated protein kinase	serine/threonine protein kinases
MEK	MAPK/ERK kinase	phosphorylates MAPK (ERK)
MEKK1	MEK kinase 1	phosphorylates MEK
MKK4/7	MAPK kinase 4/7	phosphorylates MAPK (JNK)
Ni-NTA	nickel-nitrilotriacetic acid	agarose resin by Qiagen for purification of 6xhis tagged proteins
NMR	nuclear magnetic resonance	spectroscopic method exploiting the magnetic properties of nuclei
NOE	nuclear Overhauser effect	transfer of magnetisation between nuclei via dipolar relaxation
NOESY	nuclear Overhauser enhancement spectroscopy	NMR exploiting the NOE
N-WASP	N-Wiskott Aldrich syndrome protein	regulates actin cytoskeleton remodeling
OD_{600}	optical density at 600nm	for measuring turbidity of bacterial cultures
p120RasGap	p120 Ras GTPase activating protein	remodelling the actin cytoskeleton

p130Cas	Crk-associated substrate	adaptor protein within FAs
PABA1	poly(A)-binding protein 1	mRNA-binding protein that also binds to paxillin
PAK	p21-activated serine-threonine kinase	disassembly of actin stress fibers at focal adhesions
PBD	paxillin binding domain	domains based on sequence alignment proposed to bind to paxillin
PBS	Paxillin Binding Subdomain	domains based on sequence alignment proposed to bind to paxillin
PCR	polymerase chain reaction	method of in vitro amplification of DNA using thermostable DNA polymerases
PDB	protein data bank	.pdb file containing coordinates for protein structure
PH	pleckstrin homology	domain that binds to phosphatidlinositol lipids
PI3K	phosphoinositide 3-kinase	enzymes that phosphorylate the 3 position hydroxyl group of the inositol ring of phosphatidylinositol
PIP2	phosphatidylinositol bisphosphate	product of cleavage of PIP3
PIX	Pak-interactive exchange factors	guanine nucleotide exchange factors that mediate activation of members of the Rho GTPase family
PKCa	protein kinase C alpha	serine/threonine kinase activated by diacylglycerol and Ca2+
PKL	paxillin kinase linker	ADP-ribosylation-factor–GTPase activating proteins (also called Git2)
PTK	protein tyrosine kinase	enzymes that phosphorylate proteins at tyrosine residues
PTP-PEST	protein tyrosine phosphatase-PEST	tyrosine phosphatase
Rac	rac small GTPase	regulatory e.g. induces actin polymerization at the cell periphery to produce lamellipodia
Rac1	rac small GTPase	member of the rac subfamily
Raf	raf non receptor protein kinase (ser/thr)	serine/threonine kinase that can phosphorylate MEK
Ras	ras small GTPase	responsible for cell proliferation, G-protein, active when bound to guanosine triphosphate
RF	radiofrequency pulse	
Rho	rho small GTPase	responsible for actin dynamics, G-protein, active when bound to guanosine triphosphate
SEC	size exclusion chromatography	also known as gel filtration
SH2	src homology 2 domain	bind phosphorylated tyrosine residues
SH3	src homology 3 domain	bind proline rich sequences

SOC	super optimal broth plus glucose	SOB with added glucose (with glucose for catabolite repression)
Sos	Son of sevenless	Guanine nucleotide exchange factor (GEF), acts on Ras-GTPases
Src	src tyrosine kinase	tyrosine kinase
TALOS	torsion angle likelihood obtained from shifts and sequence similarity	program for prediction ϕ and φ angles in proteins
TEMED	Tetramethylethylenediamine	used in SDS-PAGE
Tm	melting temperature	
TOCSY	total correlation spectroscopy	NMR exploiting J-couplings
TROSY	transverse relaxation optimised spectroscopy	NMR exploiting cross-relaxation to produce higher resolution spectra
UV	ultraviolet	
VASP	Vasodilator-stimulated phosphoprotein	associated with filamentous actin formation, plays a role in cell adhesion and motility

7 References

(1) Vuori K. Integrin signaling: tyrosine phosphorylation events in focal adhesions. *J Membr Biol* 1998; 165(3):191-199.

(2) Lele TP, Thodeti CK, Pendse J, Ingber DE. Investigating complexity of protein-protein interactions in focal adhesions. *Biochem Biophys Res Commun* 2008.

(3) Hynes RO. Integrins: bidirectional, allosteric signaling machines. *Cell* 2002; 110(6):673-687.

(4) Subauste MC, Pertz O, Adamson ED, Turner CE, Junger S, Hahn KM. Vinculin modulation of paxillin-FAK interactions regulates ERK to control survival and motility. *J Cell Biol* 2004; 165(3):371-381.

(5) Brown MC, Perrotta JA, Turner CE. Identification of LIM3 as the principal determinant of paxillin focal adhesion localization and characterization of a novel motif on paxillin directing vinculin and focal adhesion kinase binding. *J Cell Biol* 1996; 135(4):1109-1123.

(6) Hoellerer MK, Noble ME, Labesse G, Campbell ID, Werner JM, Arold ST. Molecular recognition of paxillin LD motifs by the focal adhesion targeting domain. *Structure (Camb)* 2003; 11(10):1207-1217.

(7) Bertolucci CM, Guibao CD, Zheng J. Structural features of the focal adhesion kinase-paxillin complex give insight into the dynamics of focal adhesion assembly. *Protein Sci* 2005; 14(3):644-652.

(8) Liu G, Guibao CD, Zheng J. Structural insight into the mechanisms of targeting and signaling of focal adhesion kinase. *Mol Cell Biol* 2002; 22(8):2751-2760.

(9) Gao G, Prutzman KC, King ML, Scheswohl DM, DeRose EF, London RE et al. NMR solution structure of the focal adhesion targeting domain of focal adhesion kinase in complex with a paxillin LD peptide: evidence for a two-site binding model. *J Biol Chem* 2004; 279(9):8441-8451.

(10) Zhang ZM, Simmerman JA, Guibao CD, Zheng JJ. GIT1 paxillin-binding domain is a four-helix bundle and it binds to both paxillin LD2 and LD4 motifs. *J Biol Chem* 2008.

(11) Wang X, Fukuda K, Byeon IJ, Velyvis A, Wu C, Gronenborn A et al. The structure of alpha -parvin CH2/paxillin LD1 complex reveals a novel modular recognition for focal adhesion assembly. *J Biol Chem* 2008.

(12) Adams JC, Watt FM. Regulation of development and differentiation by the extracellular matrix. *Development* 1993; 117(4):1183-1198.

(13) Curtis AS. The mechanism of adhesion of cells to glass. A study by interference reflection microscopy. *J Cell Biol* 1964; 20:199-215.

(14) Abercrombie M, Heaysman JE, Pegrum SM. The locomotion of fibroblasts in culture. IV. Electron microscopy of the leading lamella. *Exp Cell Res* 1971; 67(2):359-367.

(15) Smilenov LB, Mikhailov A, Pelham RJ, Marcantonio EE, Gundersen GG. Focal adhesion motility revealed in stationary fibroblasts. *Science* 1999; 286(5442):1172-1174.

(16) Zamir E, Katz BZ, Aota S, Yamada KM, Geiger B, Kam Z. Molecular diversity of cell-matrix adhesions. *J Cell Sci* 1999; 112 (Pt 11):1655-1669.

(17) Geiger B, Bershadsky A. Assembly and mechanosensory function of focal contacts. *Curr Opin Cell Biol* 2001; 13(5):584-592.

(18) Zaidel-Bar R, Ballestrem C, Kam Z, Geiger B. Early molecular events in the assembly of matrix adhesions at the leading edge of migrating cells. *J Cell Sci* 2003; 116(Pt 22):4605-4613.

(19) Ballestrem C, Erez N, Kirchner J, Kam Z, Bershadsky A, Geiger B. Molecular mapping of tyrosine-phosphorylated proteins in focal adhesions using fluorescence resonance energy transfer. *J Cell Sci* 2006; 119(Pt 5):866-875.

(20) Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, Geiger B. Functional atlas of the integrin adhesome. *Nat Cell Biol* 2007; 9(8):858-867.

(21) Webb DJ, Parsons JT, Horwitz AF. Adhesion assembly, disassembly and turnover in migrating cells -- over and over and over again. *Nat Cell Biol* 2002; 4(4):E97-100.

(22) Ballestrem C, Hinz B, Imhof BA, Wehrle-Haller B. Marching at the front and dragging behind: differential alphaVbeta3-integrin turnover regulates focal adhesion behavior. *J Cell Biol* 2001; 155(7):1319-1332.

(23) Kaverina I, Krylyshkina O, Small JV. Regulation of substrate adhesion dynamics during cell motility. *Int J Biochem Cell Biol* 2002; 34(7):746-761.

(24) Nobes CD, Hall A. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* 1995; 81(1):53-62.

(25) Zaidel-Bar R, Milo R, Kam Z, Geiger B. A paxillin tyrosine phosphorylation switch regulates the assembly and form of cell-matrix adhesions. *J Cell Sci* 2007; 120(Pt 1):137-148.

(26) Wozniak MA, Modzelewska K, Kwong L, Keely PJ. Focal adhesion regulation of cell behavior. *Biochim Biophys Acta* 2004; 1692(2-3):103-119.

(27) Lock JG, Wehrle-Haller B, Stromblad S. Cell-matrix adhesion complexes: master control machinery of cell migration. *Semin Cancer Biol* 2008; 18(1):65-76.

(28) Kirchner J, Kam Z, Tzur G, Bershadsky AD, Geiger B. Live-cell monitoring of tyrosine phosphorylation in focal adhesions following microtubule disruption. *J Cell Sci* 2003; 116(Pt 6):975-986.

(29) Sastry SK, Burridge K. Focal adhesions: a nexus for intracellular signaling and cytoskeletal dynamics. *Exp Cell Res* 2000; 261(1):25-36.

(30) Burridge K, Fath K, Kelly T, Nuckolls G, Turner C. Focal adhesions: transmembrane junctions between the extracellular matrix and the cytoskeleton. *Annu Rev Cell Biol* 1988; 4:487-525.

(31) Carragher NO, Frame MC. Focal adhesion and actin dynamics: a place where kinases and proteases meet to promote invasion. *Trends Cell Biol* 2004; 14(5):241-249.

(32) Ginsberg MH, Partridge A, Shattil SJ. Integrin regulation. *Curr Opin Cell Biol* 2005; 17(5):509-516.

(33) Gallant ND, Garcia AJ. Model of integrin-mediated cell adhesion strengthening. *J Biomech* 2007; 40(6):1301-1309.

(34) Jiang G, Giannone G, Critchley DR, Fukumoto E, Sheetz MP. Two-piconewton slip bond between fibronectin and the cytoskeleton depends on talin. *Nature* 2003; 424(6946):334-337.

(35) Izzard CS. A precursor of the focal contact in cultured fibroblasts. *Cell Motil Cytoskeleton* 1988; 10(1-2):137-142.

(36) Lo SH. Focal adhesions: what's new inside. *Dev Biol* 2006; 294(2):280-291.

(37) McCulloch CA, Downey GP, El Gabalawy H. Signalling platforms that modulate the inflammatory response: new targets for drug development. *Nat Rev Drug Discov* 2006; 5(10):864-876.

(38) Critchley DR. Cytoskeletal proteins talin and vinculin in integrin-mediated adhesion. *Biochem Soc Trans* 2004; 32(Pt 5):831-836.

(39) Borowsky ML, Hynes RO. Layilin, a novel talin-binding transmembrane protein homologous with C-type lectins, is localized in membrane ruffles. *J Cell Biol* 1998; 143(2):429-442.

(40) Johnson RP, Craig SW. The carboxy-terminal tail domain of vinculin contains a cryptic binding site for acidic phospholipids. *Biochem Biophys Res Commun* 1995; 210(1):159-164.

(41) Janssen ME, Kim E, Liu H, Fujimoto LM, Bobkov A, Volkmann N et al. Three-dimensional structure of vinculin bound to actin filaments. *Mol Cell* 2006; 21(2):271-281.

(42) Schaller MD. Biochemical signals and biological responses elicited by the focal adhesion kinase. *Biochim Biophys Acta* 2001; 1540(1):1-21.

(43) Hildebrand JD, Schaller MD, Parsons JT. Paxillin, a tyrosine phosphorylated focal adhesion-associated protein binds to the carboxyl terminal domain of focal adhesion kinase. *Mol Biol Cell* 1995; 6(6):637-647.

(44) Lee HS, Bellin RM, Walker DL, Patel B, Powers P, Liu H et al. Characterization of an actin-binding site within the talin FERM domain. *J Mol Biol* 2004; 343(3):771-784.

(45) Arnaout MA, Goodman SL, Xiong JP. Structure and mechanics of integrin-based cell adhesion. *Curr Opin Cell Biol* 2007; 19(5):495-507.

(46) Huveneers S, Truong H, Fassler R, Sonnenberg A, Danen EH. Binding of soluble fibronectin to integrin {alpha}5{beta}1 - link to focal adhesion redistribution and contractile shape. *J Cell Sci* 2008; 121(Pt 15):2452-2462.

(47) Leiss M, Beckmann K, Giros A, Costell M, Fassler R. The role of integrin binding sites in fibronectin matrix assembly in vivo. *Curr Opin Cell Biol* 2008.

(48) Chen HC, Appeddu PA, Parsons JT, Hildebrand JD, Schaller MD, Guan JL. Interaction of focal adhesion kinase with cytoskeletal protein talin. *J Biol Chem* 1995; 270(28):16995-16999.

(49) McGough A, Way M, DeRosier D. Determination of the alpha-actinin-binding site on actin filaments by cryoelectron microscopy and image analysis. *J Cell Biol* 1994; 126(2):433-443.

(50) Broussard JA, Webb DJ, Kaverina I. Asymmetric focal adhesion disassembly in motile cells. *Curr Opin Cell Biol* 2008; 20(1):85-90.

(51) Webb DJ, Donais K, Whitmore LA, Thomas SM, Turner CE, Parsons JT et al. FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. *Nat Cell Biol* 2004; 6(2):154-161.

(52) Li S, Guan JL, Chien S. Biochemistry and biomechanics of cell motility. *Annu Rev Biomed Eng* 2005; 7:105-150.

(53) Garton AJ, Tonks NK. Regulation of fibroblast motility by the protein tyrosine phosphatase PTP-PEST. *J Biol Chem* 1999; 274(6):3811-3818.

(54) Bershadsky AD, Balaban NQ, Geiger B. Adhesion-dependent cell mechanosensitivity. *Annu Rev Cell Dev Biol* 2003; 19:677-695.

(55) Chen CS, Tan J, Tien J. Mechanotransduction at cell-matrix and cell-cell contacts. *Annu Rev Biomed Eng* 2004; 6:275-302.

(56) Balaban NQ, Schwarz US, Riveline D, Goichberg P, Tzur G, Sabanay I et al. Force and focal adhesion assembly: a close relationship studied using elastic micropatterned substrates. *Nat Cell Biol* 2001; 3(5):466-472.

(57) Lauffenburger DA, Horwitz AF. Cell migration: a physically integrated molecular process. *Cell* 1996; 84(3):359-369.

(58) Cary LA, Chang JF, Guan JL. Stimulation of cell migration by overexpression of focal adhesion kinase and its association with Src and Fyn. *J Cell Sci* 1996; 109 (Pt 7):1787-1794.

(59) Reiske HR, Kao SC, Cary LA, Guan JL, Lai JF, Chen HC. Requirement of phosphatidylinositol 3-kinase in focal adhesion kinase-promoted cell migration. *J Biol Chem* 1999; 274(18):12361-12366.

(60) Han DC, Guan JL. Association of focal adhesion kinase with Grb7 and its role in cell migration. *J Biol Chem* 1999; 274(34):24425-24430.

(61) Han DC, Shen TL, Guan JL. Role of Grb7 targeting to focal contacts and its phosphorylation by focal adhesion kinase in regulation of cell migration. *J Biol Chem* 2000; 275(37):28911-28917.

(62) Cary LA, Han DC, Polte TR, Hanks SK, Guan JL. Identification of p130Cas as a mediator of focal adhesion kinase-promoted cell migration. *J Cell Biol* 1998; 140(1):211-221.

(63) Wu X, Suetsugu S, Cooper LA, Takenawa T, Guan JL. Focal adhesion kinase regulation of N-WASP subcellular localization and function. *J Biol Chem* 2004; 279(10):9565-9576.

(64) Suetsugu S, Hattori M, Miki H, Tezuka T, Yamamoto T, Mikoshiba K et al. Sustained activation of N-WASP through phosphorylation is essential for neurite extension. *Dev Cell* 2002; 3(5):645-658.

(65) Campbell ID. Modular proteins at the cell surface. *Biochem Soc Trans* 2003; 31(Pt 6):1107-1114.

(66) Glenney JR, Jr., Zokas L. Novel tyrosine kinase substrates from Rous sarcoma virus-transformed cells are present in the membrane skeleton. *J Cell Biol* 1989; 108(6):2401-2408.

(67) Turner CE, Glenney JR, Jr., Burridge K. Paxillin: a new vinculin-binding protein present in focal adhesions. *J Cell Biol* 1990; 111(3):1059-1068.

(68) Brown MC, Turner CE. Paxillin: adapting to change. *Physiol Rev* 2004; 84(4):1315-1339.

(69) Laukaitis CM, Webb DJ, Donais K, Horwitz AF. Differential dynamics of alpha 5 integrin, paxillin, and alpha-actinin during formation and disassembly of adhesions in migrating cells. *J Cell Biol* 2001; 153(7):1427-1440.

(70) Hagel M, George EL, Kim A, Tamimi R, Opitz SL, Turner CE et al. The adaptor protein paxillin is essential for normal development in the mouse and is a critical transducer of fibronectin signaling. *Mol Cell Biol* 2002; 22(3):901-915.

(71) Dawid IB, Breen JJ, Toyama R. LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends Genet* 1998; 14(4):156-162.

(72) Brown MC, Perrotta JA, Turner CE. Serine and threonine phosphorylation of the paxillin LIM domains regulates paxillin focal adhesion localization and cell adhesion to fibronectin. *Mol Biol Cell* 1998; 9(7):1803-1816.

(73) Shen Y, Schneider G, Cloutier JF, Veillette A, Schaller MD. Direct association of protein-tyrosine phosphatase PTP-PEST with paxillin. *J Biol Chem* 1998; 273(11):6474-6481.

(74) Cote JF, Turner CE, Tremblay ML. Intact LIM 3 and LIM 4 domains of paxillin are required for the association to a novel polyproline region (Pro 2) of protein-tyrosine phosphatase-PEST. *J Biol Chem* 1999; 274(29):20550-20560.

(75) Herreros L, Rodriguez-Fernandez JL, Brown MC, Alonso-Lebrero JL, Cabanas C, Sanchez-Madrid F et al. Paxillin localizes to the lymphocyte microtubule organizing center and associates with the microtubule cytoskeleton. *J Biol Chem* 2000; 275(34):26436-26440.

(76) Brown MC, Turner CE. Roles for the tubulin- and PTP-PEST-binding paxillin LIM domains in cell adhesion and motility. *Int J Biochem Cell Biol* 2002; 34(7):855-863.

(77) Nikolopoulos SN, Turner CE. Actopaxin, a new focal adhesion protein that binds paxillin LD motifs and actin and regulates cell adhesion. *J Cell Biol* 2000; 151(7):1435-1448.

(78) Nikolopoulos SN, Turner CE. Integrin-linked kinase (ILK) binding to paxillin LD1 motif regulates ILK localization to focal adhesions. *J Biol Chem* 2001; 276(26):23499-23505.

(79) Turner CE, Miller JT. Primary sequence of paxillin contains putative SH2 and SH3 domain binding motifs and multiple LIM domains: identification of a vinculin and pp125Fak-binding region. *J Cell Sci* 1994; 107 (Pt 6):1583-1591.

(80) Turner CE, Brown MC, Perrotta JA, Riedy MC, Nikolopoulos SN, McDonald AR et al. Paxillin LD4 motif binds PAK and PIX through a novel 95-kD ankyrin repeat, ARF-GAP protein: A role in cytoskeletal remodeling. *J Cell Biol* 1999; 145(4):851-863.

(81) Tong X, Salgia R, Li JL, Griffin JD, Howley PM. The bovine papillomavirus E6 protein binds to the LD motif repeats of paxillin and blocks its interaction with vinculin and the focal adhesion kinase. *J Biol Chem* 1997; 272(52):33373-33376.

(82) West KA, Zhang H, Brown MC, Nikolopoulos SN, Riedy MC, Horwitz AF et al. The LD4 motif of paxillin regulates cell spreading and motility through an interaction with paxillin kinase linker (PKL). *J Cell Biol* 2001; 154(1):161-176.

(83) Sheibani N, Tang Y, Sorenson CM. Paxillin's LD4 motif interacts with bcl-2. *J Cell Physiol* 2008; 214(3):655-661.

(84) Weng Z, Taylor JA, Turner CE, Brugge JS, Seidel-Dugan C. Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells. *J Biol Chem* 1993; 268(20):14956-14963.

(85) Brown MC, Curtis MS, Turner CE. Paxillin LD motifs may define a new family of protein recognition domains. *Nat Struct Biol* 1998; 5(8):677-678.

(86) Wood CK, Turner CE, Jackson P, Critchley DR. Characterisation of the paxillin-binding site and the C-terminal focal adhesion targeting sequence in vinculin. *J Cell Sci* 1994; 107 (Pt 2):709-717.

(87) Bertolucci CM, Guibao CD, Zheng JJ. Phosphorylation of paxillin LD4 destabilizes helix formation and inhibits binding to focal adhesion kinase. *Biochemistry* 2008; 47(2):548-554.

(88) Webb DJ, Schroeder MJ, Brame CJ, Whitmore L, Shabanowitz J, Hunt DF et al. Paxillin phosphorylation sites mapped by mass spectrometry. *J Cell Sci* 2005; 118(Pt 21):4925-4929.

(89) Bellis SL, Miller JT, Turner CE. Characterization of tyrosine phosphorylation of paxillin in vitro by focal adhesion kinase. *J Biol Chem* 1995; 270(29):17437-17441.

(90) Schaller MD, Parsons JT. pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk. *Mol Cell Biol* 1995; 15(5):2635-2645.

(91) Richardson A, Malik RK, Hildebrand JD, Parsons JT. Inhibition of cell spreading by expression of the C-terminal domain of focal adhesion kinase (FAK) is rescued by coexpression of Src or catalytically inactive FAK: a role for paxillin tyrosine phosphorylation. *Mol Cell Biol* 1997; 17(12):6906-6914.

(92) Huang C, Borchers CH, Schaller MD, Jacobson K. Phosphorylation of paxillin by p38MAPK is involved in the neurite extension of PC-12 cells. *J Cell Biol* 2004; 164(4):593-602.

(93) Petit V, Boyer B, Lentz D, Turner CE, Thiery JP, Valles AM. Phosphorylation of tyrosine residues 31 and 118 on paxillin regulates cell migration through an association with CRK in NBT-II cells. *J Cell Biol* 2000; 148(5):957-970.

(94) Nakamura K, Yano H, Uchida H, Hashimoto S, Schaefer E, Sabe H. Tyrosine phosphorylation of paxillin alpha is involved in temporospatial regulation of paxillin-containing focal adhesion formation and F-actin organization in motile cells. *J Biol Chem* 2000; 275(35):27155-27164.

(95) Salgia R, Li JL, Lo SH, Brunkhorst B, Kansas GS, Sobhany ES et al. Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by P210BCR/ABL. *J Biol Chem* 1995; 270(10):5039-5047.

(96) Liu S, Thomas SM, Woodside DG, Rose DM, Kiosses WB, Pfaff M et al. Binding of paxillin to alpha4 integrins modifies integrin-dependent biological responses. *Nature* 1999; 402(6762):676-681.

(97) Schaller MD, Otey CA, Hildebrand JD, Parsons JT. Focal adhesion kinase and paxillin bind to peptides mimicking beta integrin cytoplasmic domains. *J Cell Biol* 1995; 130(5):1181-1187.

(98) Chen LM, Bailey D, Fernandez-Valle C. Association of beta 1 integrin with focal adhesion kinase and paxillin in differentiating Schwann cells. *J Neurosci* 2000; 20(10):3776-3784.

(99) Rose DM, Liu S, Woodside DG, Han J, Schlaepfer DD, Ginsberg MH. Paxillin binding to the alpha 4 integrin subunit stimulates LFA-1 (integrin alpha L beta 2)-dependent T cell migration by augmenting the activation of focal adhesion kinase/proline-rich tyrosine kinase-2. *J Immunol* 2003; 170(12):5912-5918.

(100) Woods AJ, Kantidakis T, Sabe H, Critchley DR, Norman JC. Interaction of paxillin with poly(A)-binding protein 1 and its role in focal adhesion turnover and cell migration. *Mol Cell Biol* 2005; 25(9):3763-3773.

(101) Woods AJ, Roberts MS, Choudhary J, Barry ST, Mazaki Y, Sabe H et al. Paxillin associates with poly(A)-binding protein 1 at the dense endoplasmic reticulum and the leading edge of migrating cells. *J Biol Chem* 2002; 277(8):6428-6437.

(102) Deakin NO, Turner CE. Paxillin comes of age. *J Cell Sci* 2008; 121(Pt 15):2435-2444.

(103) Schlaepfer DD, Hauck CR, Sieg DJ. Signaling through focal adhesion kinase. *Prog Biophys Mol Biol* 1999; 71(3-4):435-478.

(104) Mitra SK, Hanson DA, Schlaepfer DD. Focal adhesion kinase: in command and control of cell motility. *Nat Rev Mol Cell Biol* 2005; 6(1):56-68.

(105) Parsons JT. Focal adhesion kinase: the first ten years. *J Cell Sci* 2003; 116(Pt 8):1409-1416.

(106) Dunty JM, Schaller MD. The N termini of focal adhesion kinase family members regulate substrate phosphorylation, localization, and cell morphology. *J Biol Chem* 2002; 277(47):45644-45654.

(107) Stewart A, Ham C, Zachary I. The focal adhesion kinase amino-terminal domain localises to nuclei and intercellular junctions in HEK 293 and MDCK cells independently of tyrosine 397 and the carboxy-terminal domain. *Biochem Biophys Res Commun* 2002; 299(1):62-73.

(108) Hildebrand JD, Schaller MD, Parsons JT. Identification of sequences required for the efficient localization of the focal adhesion kinase, pp125FAK, to cellular focal adhesions. *J Cell Biol* 1993; 123(4):993-1005.

(109) Schlaepfer DD, Mitra SK, Ilic D. Control of motile and invasive cell phenotypes by focal adhesion kinase. *Biochim Biophys Acta* 2004; 1692(2-3):77-102.

(110) Tachibana K, Sato T, D'Avirro N, Morimoto C. Direct association of pp125FAK with paxillin, the focal adhesion-targeting mechanism of pp125FAK. *J Exp Med* 1995; 182(4):1089-1099.

(111) Harte MT, Hildebrand JD, Burnham MR, Bouton AH, Parsons JT. p130Cas, a substrate associated with v-Src and v-Crk, localizes to focal adhesions and binds to focal adhesion kinase. *J Biol Chem* 1996; 271(23):13649-13655.

(112) Liu Y, Loijens JC, Martin KH, Karginov AV, Parsons JT. The association of ASAP1, an ADP ribosylation factor-GTPase activating protein, with focal adhesion kinase contributes to the process of focal adhesion assembly. *Mol Biol Cell* 2002; 13(6):2147-2156.

(113) Randazzo PA, Andrade J, Miura K, Brown MT, Long YQ, Stauffer S et al. The Arf GTPase-activating protein ASAP1 regulates the actin cytoskeleton. *Proc Natl Acad Sci U S A* 2000; 97(8):4011-4016.

(114) Schaller MD, Hildebrand JD, Parsons JT. Complex formation with focal adhesion kinase: A mechanism to regulate activity and subcellular localization of Src kinases. *Mol Biol Cell* 1999; 10(10):3489-3505.

(115) Tilghman RW, Parsons JT. Focal adhesion kinase as a regulator of cell tension in the progression of cancer. *Semin Cancer Biol* 2008; 18(1):45-52.

(116) Ilic D, Furuta Y, Kanazawa S, Takeda N, Sobue K, Nakatsuji N et al. Reduced cell motility and enhanced focal adhesion contact formation in cells from FAK-deficient mice. *Nature* 1995; 377(6549):539-544.

(117) Hungerford JE, Compton MT, Matter ML, Hoffstrom BG, Otey CA. Inhibition of pp125FAK in cultured fibroblasts results in apoptosis. *J Cell Biol* 1996; 135(5):1383-1390.

(118) Frisch SM, Vuori K, Ruoslahti E, Chan-Hui PY. Control of adhesion-dependent cell survival by focal adhesion kinase. *J Cell Biol* 1996; 134(3):793-799.

(119) Schlaepfer DD, Hunter T. Integrin signalling and tyrosine phosphorylation: just the FAKs? *Trends Cell Biol* 1998; 8(4):151-157.

(120) Oktay M, Wary KK, Dans M, Birge RB, Giancotti FG. Integrin-mediated activation of focal adhesion kinase is required for signaling to Jun NH2-terminal kinase and progression through the G1 phase of the cell cycle. *J Cell Biol* 1999; 145(7):1461-1469.

(121) Igishi T, Fukuhara S, Patel V, Katz BZ, Yamada KM, Gutkind JS. Divergent signaling pathways link focal adhesion kinase to mitogen-activated protein kinase cascades. Evidence for a role of paxillin in c-Jun NH(2)-terminal kinase activation. *J Biol Chem* 1999; 274(43):30738-30746.

(122) Schlaepfer DD, Hanks SK, Hunter T, van der GP. Integrin-mediated signal transduction linked to Ras pathway by GRB2 binding to focal adhesion kinase. *Nature* 1994; 372(6508):786-791.

(123) Gabarra-Niecko V, Schaller MD, Dunty JM. FAK regulates biological processes important for the pathogenesis of cancer. *Cancer Metastasis Rev* 2003; 22(4):359-374.

(124) Cho SY, Klemke RL. Extracellular-regulated kinase activation and CAS/Crk coupling regulate cell migration and suppress apoptosis during invasion of the extracellular matrix. *J Cell Biol* 2000; 149(1):223-236.

(125) Schwartz MA, Ginsberg MH. Networks and crosstalk: integrin signalling spreads. *Nat Cell Biol* 2002; 4(4):E65-E68.

(126) Siesser PM, Hanks SK. The signaling and biological implications of FAK overexpression in cancer. *Clin Cancer Res* 2006; 12(11 Pt 1):3233-3237.

(127) Huang C, Jacobson K, Schaller MD. MAP kinases and cell migration. *J Cell Sci* 2004; 117(Pt 20):4619-4628.

(128) Xu W, Baribault H, Adamson ED. Vinculin knockout results in heart and brain defects during embryonic development. *Development* 1998; 125(2):327-337.

(129) Geiger B. A 130K protein from chicken gizzard: its localization at the termini of microfilament bundles in cultured chicken cells. *Cell* 1979; 18(1):193-205.

(130) Driska S, Hartshorne DJ. The contractile proteins of smooth muscle. Properties and components of a Ca2+-sensitive actomyosin from chicken gizzard. *Arch Biochem Biophys* 1975; 167(1):203-212.

(131) Sobieszek A, Bremel RD. Preparation and properties of vertebrate smooth-muscle myofibrils and actomyosin. *Eur J Biochem* 1975; 55(1):49-60.

(132) Feramisco JR, Burridge K. A rapid purification of alpha-actinin, filamin, and a 130,000-dalton protein from smooth muscle. *J Biol Chem* 1980; 255(3):1194-1199.

(133) Burridge K, Feramisco JR. Microinjection and localization of a 130K protein in living fibroblasts: a relationship to actin and fibronectin. *Cell* 1980; 19(3):587-595.

(134) Geiger B, Tokuyasu KT, Dutton AH, Singer SJ. Vinculin, an intracellular protein localized at specialized sites where microfilament bundles terminate at cell membranes. *Proc Natl Acad Sci U S A* 1980; 77(7):4127-4131.

(135) Rodriguez Fernandez JL, Geiger B, Salomon D, Sabanay I, Zoller M, Ben Ze'Ev A. Suppression of tumorigenicity in transformed cells after transfection with vinculin cDNA. *J Cell Biol* 1992; 119(2):427-438.

(136) Coll JL, Ben Ze'Ev A, Ezzell RM, Rodriguez Fernandez JL, Baribault H, Oshima RG et al. Targeted disruption of vinculin genes in F9 and embryonic stem cells changes cell morphology, adhesion, and locomotion. *Proc Natl Acad Sci U S A* 1995; 92(20):9161-9165.

(137) Saunders RM, Holt MR, Jennings L, Sutton DH, Barsukov IL, Bobkov A et al. Role of vinculin in regulating focal adhesion turnover. *Eur J Cell Biol* 2006; 85(6):487-500.

(138) Xu W, Coll JL, Adamson ED. Rescue of the mutant phenotype by reexpression of full-length vinculin in null F9 cells; effects on cell locomotion by domain deleted vinculin. *J Cell Sci* 1998; 111 (Pt 11):1535-1544.

(139) Rodriguez Fernandez JL, Geiger B, Salomon D, Ben Ze'Ev A. Overexpression of vinculin suppresses cell motility in BALB/c 3T3 cells. *Cell Motil Cytoskeleton* 1992; 22(2):127-134.

(140) Gallant ND, Michael KE, Garcia AJ. Cell adhesion strengthening: contributions of adhesive area, integrin binding, and focal adhesion assembly. *Mol Biol Cell* 2005; 16(9):4329-4340.

(141) Galbraith CG, Yamada KM, Sheetz MP. The relationship between force and focal complex development. *J Cell Biol* 2002; 159(4):695-705.

(142) Chandrasekar I, Stradal TE, Holt MR, Entschladen F, Jockusch BM, Ziegler WH. Vinculin acts as a sensor in lipid regulation of adhesion-site turnover. *J Cell Sci* 2005.

(143) Isenberg G, Leonard K, Jockusch BM. Structural aspects of vinculin-actin interactions. *J Mol Biol* 1982; 158(2):231-249.

(144) Milam LM. Electron microscopy of rotary shadowed vinculin and vinculin complexes. *J Mol Biol* 1985; 184(3):543-545.

(145) Molony L, Burridge K. Molecular shape and self-association of vinculin and metavinculin. *J Cell Biochem* 1985; 29(1):31-36.

(146) Winkler J, Lunsdorf H, Jockusch BM. The ultrastructure of chicken gizzard vinculin as visualized by high-resolution electron microscopy. *J Struct Biol* 1996; 116(2):270-277.

(147) Coutu MD, Craig SW. cDNA-derived sequence of chicken embryo vinculin. *Proc Natl Acad Sci U S A* 1988; 85(22):8535-8539.

(148) Price GJ, Jones P, Davison MD, Patel B, Bendori R, Geiger B et al. Primary sequence and domain structure of chicken vinculin. *Biochem J* 1989; 259(2):453-461.

(149) Johnson RP, Craig SW. An intramolecular association between the head and tail domains of vinculin modulates talin binding. *J Biol Chem* 1994; 269(17):12611-12619.

(150) Kroemker M, Rudiger AH, Jockusch BM, Rudiger M. Intramolecular interactions in vinculin control alpha-actinin binding to the vinculin head. *FEBS Lett* 1994; 355(3):259-262.

(151) Johnson RP, Craig SW. F-actin binding site masked by the intramolecular association of vinculin head and tail domains. *Nature* 1995; 373(6511):261-264.

(152) Bakolitsa C, Cohen DM, Bankston LA, Bobkov AA, Cadwell GW, Jennings L et al. Structural basis for vinculin activation at sites of cell adhesion. *Nature* 2004; 430(6999):583-586.

(153) Borgon RA, Vonrhein C, Bricogne G, Bois PR, Izard T. Crystal structure of human vinculin. *Structure (Camb)* 2004; 12(7):1189-1197.

(154) Izard T, Evans G, Borgon RA, Rush CL, Bricogne G, Bois PR. Vinculin activation by talin through helical bundle conversion. *Nature* 2004; 427(6970):171-175.

(155) Cohen DM, Chen H, Johnson RP, Choudhury B, Craig SW. Two distinct head-tail interfaces cooperate to suppress activation of Vinculin by Talin. *J Biol Chem* 2005.

(156) Chen H, Cohen DM, Choudhury DM, Kioka N, Craig SW. Spatial distribution and functional significance of activated vinculin in living cells. *J Cell Biol* 2005; 169(3):459-470.

(157) Cohen DM, Kutscher B, Chen H, Murphy DB, Craig SW. A conformational switch in vinculin drives formation and dynamics of a talin-vinculin complex at focal adhesions. *J Biol Chem* 2006; 281(23):16006-16015.

(158) Izard T, Vonrhein C. Structural basis for amplifying vinculin activation by talin. *J Biol Chem* 2004; 279(26):27667-27678.

(159) Gilmore AP, Burridge K. Regulation of vinculin binding to talin and actin by phosphatidyl-inositol-4-5-bisphosphate. *Nature* 1996; 381(6582):531-535.

(160) Chen H, Choudhury DM, Craig SW. Coincidence of actin filaments and talin is required to activate vinculin. *J Biol Chem* 2006; 281(52):40389-40398.

(161) Bois PR, O'Hara BP, Nietlispach D, Kirkpatrick J, Izard T. The vinculin binding sites of talin and alpha-actinin are sufficient to activate vinculin. *J Biol Chem* 2006; 281(11):7228-7236.

(162) Bois PR, Borgon RA, Vonrhein C, Izard T. Structural dynamics of alpha-actinin-vinculin interactions. *Mol Cell Biol* 2005; 25(14):6112-6122.

(163) Papagrigoriou E, Gingras AR, Barsukov IL, Bate N, Fillingham IJ, Patel B et al. Activation of a vinculin-binding site in the talin rod involves rearrangement of a five-helix bundle. *EMBO J* 2004; 23(15):2942-2951.

(164) Fillingham I, Gingras AR, Papagrigoriou E, Patel B, Emsley J, Critchley DR et al. A vinculin binding domain from the talin rod unfolds to form a complex with the vinculin head. *Structure (Camb)* 2005; 13(1):65-74.

(165) Chen Y, Dokholyan NV. Insights into allosteric control of vinculin function from its large scale conformational dynamics. *J Biol Chem* 2006; 281(39):29148-29154.

(166) Menkel AR, Kroemker M, Bubeck P, Ronsiek M, Nikolai G, Jockusch BM. Characterization of an F-actin-binding domain in the cytoskeletal protein vinculin. *J Cell Biol* 1994; 126(5):1231-1240.

(167) Huttelmaier S, Bubeck P, Rudiger M, Jockusch BM. Characterization of two F-actin-binding and oligomerization sites in the cell-contact protein vinculin. *Eur J Biochem* 1997; 247(3):1136-1142.

(168) Burridge K, Mangeat P. An interaction between vinculin and talin. *Nature* 1984; 308(5961):744-746.

(169) Gingras AR, Ziegler WH, Frank R, Barsukov IL, Roberts GC, Critchley DR et al. Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. *J Biol Chem* 2005; 280(44):37217-37224.

(170) Nhieu GT, Izard T. Vinculin binding in its closed conformation by a helix addition mechanism. *EMBO J* 2007; 26(21):4588-4596.

(171) Johnson RP, Niggli V, Durrer P, Craig SW. A conserved motif in the tail domain of vinculin mediates association with and insertion into acidic phospholipid bilayers. *Biochemistry* 1998; 37(28):10211-10222.

(172) Bakolitsa C, de Pereda JM, Bagshaw CR, Critchley DR, Liddington RC. Crystal structure of the vinculin tail suggests a pathway for activation. *Cell* 1999; 99(6):603-613.

(173) Fukami K, Endo T, Imamura M, Takenawa T. alpha-Actinin and vinculin are PIP2-binding proteins involved in signaling by tyrosine kinase. *J Biol Chem* 1994; 269(2):1518-1522.

(174) Huttelmaier S, Mayboroda O, Harbeck B, Jarchau T, Jockusch BM, Rudiger M. The interaction of the cell-contact proteins VASP and vinculin is regulated by phosphatidylinositol-4,5-bisphosphate. *Curr Biol* 1998; 8(9):479-488.

(175) Kelly DF, Taylor DW, Bakolitsa C, Bobkov AA, Bankston L, Liddington RC et al. Structure of the alpha-actinin-vinculin head domain complex determined by cryo-electron microscopy. *J Mol Biol* 2006; 357(2):562-573.

(176) Reinhard M, Rudiger M, Jockusch BM, Walter U. VASP interaction with vinculin: a recurring theme of interactions with proline-rich motifs. *FEBS Lett* 1996; 399(1-2):103-107.

(177) Brindle NP, Holt MR, Davies JE, Price CJ, Critchley DR. The focal-adhesion vasodilator-stimulated phosphoprotein (VASP) binds to the proline-rich domain in vinculin. *Biochem J* 1996; 318 (Pt 3):753-757.

(178) Takahashi H, Mitsushima M, Okada N, Ito T, Aizawa S, Akahane R et al. Role of interaction with vinculin in recruitment of vinexins to focal adhesions. *Biochem Biophys Res Commun* 2005; 336(1):239-246.

(179) Kioka N, Sakata S, Kawauchi T, Amachi T, Akiyama SK, Okazaki K et al. Vinexin: a novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organization. *J Cell Biol* 1999; 144(1):59-69.

(180) Demali KA, Barlow CA, Burridge K. Recruitment of the Arp2/3 complex to vinculin: coupling membrane protrusion to matrix adhesion. *J Cell Biol* 2002; 159(5):881-891.

(181) Bourdet-Sicard R, Rudiger M, Jockusch BM, Gounon P, Sansonetti PJ, Nhieu GT. Binding of the Shigella protein IpaA to vinculin induces F-actin depolymerization. *EMBO J* 1999; 18(21):5853-5862.

(182) Hamiaux C, van Eerde A, Parsot C, Broos J, Dijkstra BW. Structural mimicry for vinculin activation by IpaA, a virulence factor of Shigella flexneri. *EMBO Rep* 2006; 7(8):794-799.

(183) Mandai K, Nakanishi H, Satoh A, Takahashi K, Satoh K, Nishioka H et al. Ponsin/SH3P12: an l-afadin- and vinculin-binding protein localized at cell-cell and cell-matrix adherens junctions. *J Cell Biol* 1999; 144(5):1001-1017.

(184) Weekes J, Barry ST, Critchley DR. Acidic phospholipids inhibit the intramolecular association between the N- and C-terminal regions of vinculin, exposing actin-binding and protein kinase C phosphorylation sites. *Biochem J* 1996; 314 (Pt 3):827-832.

(185) Ziegler WH, Tigges U, Zieseniss A, Jockusch BM. A lipid-regulated docking site on vinculin for protein kinase C. *J Biol Chem* 2002; 277(9):7396-7404.

(186) Weiss EE, Kroemker M, Rudiger AH, Jockusch BM, Rudiger M. Vinculin is part of the cadherin-catenin junctional complex: complex formation between alpha-catenin and vinculin. *J Cell Biol* 1998; 141(3):755-764.

(187) Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the cadherin-catenin-actin complex. *Cell* 2005; 123(5):889-901.

(188) Humphries JD, Wang P, Streuli C, Geiger B, Humphries MJ, Ballestrem C. Vinculin controls focal adhesion formation by direct interactions with talin and actin. *J Cell Biol* 2007; 179(5):1043-1057.

(189) Richardson A, Parsons T. A mechanism for regulation of the adhesion-associated proteintyrosine kinase pp125FAK. *Nature* 1996; 380(6574):538-540.

(190) Wade R, Bohl J, Vande PS. Paxillin null embryonic stem cells are impaired in cell spreading and tyrosine phosphorylation of focal adhesion kinase. *Oncogene* 2002; 21(1):96-107.

(191) Thomas JW, Cooley MA, Broome JM, Salgia R, Griffin JD, Lombardo CR et al. The role of focal adhesion kinase binding in the regulation of tyrosine phosphorylation of paxillin. *J Biol Chem* 1999; 274(51):36684-36692.

(192) Turner CE. Paxillin and focal adhesion signalling. *Nat Cell Biol* 2000; 2:E231-E236.

(193) Klemke RL, Leng J, Molander R, Brooks PC, Vuori K, Cheresh DA. CAS/Crk coupling serves as a "molecular switch" for induction of cell migration. *J Cell Biol* 1998; 140(4):961-972.

(194) Wu WS, Wu JR, Hu CT. Signal cross talks for sustained MAPK activation and cell migration: the potential role of reactive oxygen species. *Cancer Metastasis Rev* 2008.

(195) Danen EH, Yamada KM. Fibronectin, integrins, and growth control. *J Cell Physiol* 2001; 189(1):1-13.

(196) Ishibe S, Joly D, Liu ZX, Cantley LG. Paxillin serves as an ERK-regulated scaffold for coordinating FAK and Rac activation in epithelial morphogenesis. *Mol Cell* 2004; 16(2):257-267.

(197) Howell BW, Cooper JA. Csk suppression of Src involves movement of Csk to sites of Src activity. *Mol Cell Biol* 1994; 14(8):5402-5411.

(198) Luxenburg C, Addadi L, Geiger B. The molecular dynamics of osteoclast adhesions. *Eur J Cell Biol* 2006; 85(3-4):203-211.

(199) Bubeck P, Pistor S, Wehland J, Jockusch BM. Ligand recruitment by vinculin domains in transfected cells. *J Cell Sci* 1997; 110 (Pt 12):1361-1371.

(200) Schwartz MA. Integrin signaling revisited. *Trends Cell Biol* 2001; 11(12):466-470.

(201) Bass MD, Smith BJ, Prigent SA, Critchley DR. Talin contains three similar vinculin-binding sites predicted to form an amphipathic helix. *Biochem J* 1999; 341 (Pt 2):257-263.

(202) Gilmore AP, Wood C, Ohanian V, Jackson P, Patel B, Rees DJ et al. The cytoskeletal protein talin contains at least two distinct vinculin binding domains. *J Cell Biol* 1993; 122(2):337-347.

(203) Patel B, Gingras AR, Bobkov AA, Fujimoto LM, Zhang M, Liddington RC et al. The activity of the vinculin binding sites in talin is influenced by the stability of the helical bundles that make up the talin rod. *J Biol Chem* 2006; 281(11):7458-7467.

(204) Johnson RP, Craig SW. Actin activates a cryptic dimerization potential of the vinculin tail domain. *J Biol Chem* 2000; 275(1):95-105.

(205) Hayashi I, Vuori K, Liddington RC. The focal adhesion targeting (FAT) region of focal adhesion kinase is a four-helix bundle that binds paxillin. *Nat Struct Biol* 2002; 9(2):101-106.

(206) Tumbarello DA, Brown MC, Turner CE. The paxillin LD motifs. *FEBS Lett* 2002; 513(1):114-118.

(207) Scheswohl DM, Harrell JR, Rajfur Z, Gao G, Campbell SL, Schaller MD. Multiple paxillin binding sites regulate FAK function. *J Mol Signal* 2008; 3:1.

(208) Glasoe PK, Long FA. Use of glass electrodes to measure acidities in deuterium oxide. *Journal of Physical Chemistry* 1960; 64:188-189.

(209) Campbell ID, Dworzak F. *Biological Spectroscopy*. 1984.

(210) Greenfield N, Fasman GD. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 1969; 8(10):4108-4116.

(211) Rodger A, Norden B. *Circular dichroism and linear dichroism*. Oxford ; New York : Oxford University Press, 1997.

(212) Unneberg P, Merelo JJ, Chacon P, Moran F. SOMCD: method for evaluating protein secondary structure from UV circular dichroism spectra. *Proteins* 2001; 42(4):460-470.

(213) JACOBSON BERT, ANDERSON WA, ARNOLD JT. A Proton Magnetic Resonance Study of the Hydration of Deoxyribonucleic Acid. *Nature* 1954; 173(4408):772-773.

(214) Keeler J. *Understanding NMR spectroscopy*. Chichester : John Wiley, 2005.

(215) Wuthrich K. *NMR of proteins and nucleic acids*. New York ; Chichester : Wiley, 1986.

(216) Freeman R. *A handbook of nuclear magnetic resonance*. Harlow : Longman Scientific & Technical, 1987.

(217) Levitt MH. *Spin dynamics : basics of nuclear magnetic resonance*. Chichester : John Wiley & Sons, 2001.

(218) Cavanagh J, Fairbrother WJ, Palmer AGI, Skelton NJ, Rance M. *Protein NMR spectroscopy : principles and practice*. 2nd ed. ed. Amsterdam ; London : Elsevier, 2007.

(219) Bodenhausen G, Ruben DJ. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 1980; 69(1):185-189.

(220) Bax A, Ikura M, Kay LE, Torchia DA, Tschudin R. Comparison of different modes of 2-dimentional reverse-correlation NMR for the study of proteins. *Journal of Magnetic Resonance* 1990; 86:304-318.

(221) Morris GA, Freeman R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* 1979; 101(3):760-762.

(222) Palmer AG, Cavanagh J, Wright PE, Rance M. Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation Nmr-Spectroscopy. *Journal of Magnetic Resonance* 1991; 93(1):151-170.

(223) Braunschweiler L, Ernst RR. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *Journal of Magnetic Resonance (1969)* 1983; 53(3):521-528.

(224) Davis DG, Bax A. Assignment of Complex H-1-Nmr Spectra Via Two-Dimensional Homonuclear Hartmann-Hahn Spectroscopy. *Journal of the American Chemical Society* 1985; 107(9):2820-2821.

(225) Griesinger C, Sorensen OW, Ernst RR. Three-dimensional Fourier spectroscopy. Application to high-resolution NMR. *Journal of Magnetic Resonance (1969)* 1989; 84(1):14-63.

(226) Overhauser AW. Paramagnetic Relaxation in Metals. *Phys Rev* 1953; 89(4):689.

(227) Solomon I. Relaxation Processes in a System of Two Spins. *Phys Rev* 1955; 99(2):559.

(228) Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. 3-dimensional heteronuclear NMR of 15N-labeled proteins. *Journal of the American Chemical Society* 1989; 111:1515-1517.

(229) Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM et al. Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry* 1989; 28(15):6150-6156.

(230) Zuiderweg ER, Fesik SW. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 1989; 28(6):2387-2391.

(231) Kay LE, Ikura M, Tschudin R, Bax A. 3-dimensional triple resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance* 1990; 89:496-514.

(232) Farmer BT, Venters RA, Spicer LD, Wittekind MG, Muller L. A refocused and optimized HNCA: increased sensitivity and resolution in large macromolecules. *J Biomol NMR* 1992; 2(2):195-202.

(233) Grzesiek S, Bax A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969)* 1992; 96(2):432-440.

(234) Wittekind M, Mueller L. HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- carbon and beta-carbon resonances in proteins. *Journal of Magnetic Resonance, series B* 1993; 101:201-205.

(235) Bax A, Ikura M. An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preeceding residue in uniformly 15N/13C enriched proteins. *Journal of Biomolecular NMR* 1991; 1:99-104.

(236) Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE. A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. *Journal of the American Chemical Society* 1994; 116(26):11655-11666.

(237) Muhandiram DR, Kay LE. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance, series B* 1994; 103(3):203-216.

(238) Pervushin K, Riek R, Wider G, Wuthrich K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* 1997; 94(23):12366-12371.

(239) Krishnan VV, Cosman M. An empirical relationship between rotational correlation time and solvent accessible surface area. *Journal of Biomolecular NMR* 1998; 12(1):177-182.

(240) Garcia dlT, Huertas ML, Carrasco B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* 2000; 147(1):138-146.

(241) Lee D, Hilty C, Wider G, Wuthrich K. Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* 2006; 178(1):72-76.

(242) Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 1989; 28(23):8972-8979.

(243) Kroenke CD, Loria JP, Lee LK, Rance M, Palmer AG. Longitudinal and transverse H-1-N-15 dipolar N-15 chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *Journal of the American Chemical Society* 1998; 120(31):7905-7915.

(244) Garcia dlT, Navarro S, Lopez Martinez MC, Diaz FG, Lopez Cascales JJ. HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys J* 1994; 67(2):530-531.

(245) LeMaster DM. Uniform and selective deuteration in two-dimensional NMR of proteins. *Annu Rev Biophys Biophys Chem* 1990; 19:243-266.

(246) Gardner KH, Kay LE. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* 1998; 27:357-406.

(247) Clore GM, Gronenborn AM. NMR structure determination of proteins and protein complexes larger than 20 kDa. *Curr Opin Chem Biol* 1998; 2(5):564-570.

(248) Venters RA, Farmer BT, Fierke CA, Spicer LD. Characterizing the use of perdeuteration in NMR studies of large proteins: 13C, 15N and 1H assignments of human carbonic anhydrase II. *J Mol Biol* 1996; 264(5):1101-1116.

(249) Wuthrich K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *J Biomol NMR* 2003; 27(1):13-39.

(250) Spera S, Bax A. Empirical correlation between protein backbone conformation and C.alpha. and C.beta. 13C nuclear magnetic resonance chemical shifts. *Journal of the American Chemical Society* 1991; 113(14):5490-5492.

(251) Wishart DS, Sykes BD. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J Biomol NMR* 1994; 4(2):171-180.

(252) Schwarzinger S, Kroon GJ, Foss TR, Wright PE, Dyson HJ. Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J Biomol NMR* 2000; 18(1):43-48.

(253) Clarkson J, Campbell ID. Studies of protein-ligand interactions by NMR. *Biochem Soc Trans* 2003; 31(Pt 5):1006-1009.

(254) Zuiderweg ER. Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* 2002; 41(1):1-7.

(255) Lian L, Roberts GCK. Effects of chemical exchange on NMR spectra. In: Roberts GCK, editor. NMR of Macromolecules. Oxford University Press, 1993: 153-182.

(256) Zuiderweg ER, Hamers LF, Rollema HS, de Bruin SH, Hilbers CW. 31P NMR study of the kinetics of binding of myo-inositol hexakisphosphate to human hemoglobin. Observation of fast-exchange kinetics in high-affinity systems. *Eur J Biochem* 1981; 118(1):95-104.

(257) Englander SW, Mayne L, Bai Y, Sosnick TR. Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci* 1997; 6(5):1101-1109.

(258) Englander SW, Sosnick TR, Englander JJ, Mayne L. Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol* 1996; 6(1):18-23.

(259) Miller DW, Dill KA. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci* 1995; 4(9):1860-1873.

(260) HVIDT A, LINDERSTROM-LANG K. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim Biophys Acta* 1954; 14(4):574-575.

(261) HVIDT A, Nielsen SO. Hydrogen exchange in proteins. *Adv Protein Chem* 1966; 21:287-386.

(262) Ferraro DM, Lazo ND, Robertson AD. EX1 hydrogen exchange and protein folding. *Biochemistry* 2004; 43(3):587-594.

(263) Loh SN, Rohl CA, Kiefhaber T, Baldwin RL. A general two-process model describes the hydrogen exchange behavior of RNase A in unfolding conditions. *Proc Natl Acad Sci U S A* 1996; 93(5):1982-1987.

(264) Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 1995; 6(3):277-293.

(265) Johnson BA, Blevins RA. NMR View: A computer program for the visualization and analysis of NMR data. *Journal of Biomolecular NMR* 1994; 4(5):603-614.

(266) Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J Biomol NMR* 1993; 3(2):185-204.

(267) Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 1999; 13(3):289-302.

(268) Charbonnier S, Zanier K, Masson M, Trave G. Capturing protein-protein complexes at equilibrium: The holdup comparative chromatographic retention assay. *Protein Expr Purif* 2006; 50(1):89-101.

(269) Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 1996; 14(1):51-32.

(270) Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Humana Press, 2005: 571-607.

(271) Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 1985; 230(4732):1350-1354.

(272) Braman J, Papworth C, Greener A. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol Biol* 1996; 57:31-44.

(273) Golovanov AP, Hautbergue GM, Wilson SA, Lian LY. A simple method for improving protein solubility and long-term stability. *J Am Chem Soc* 2004; 126(29):8933-8939.

(274) Tugarinov V, Hwang PM, Kay LE. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu Rev Biochem* 2004; 73:107-146.

(275) Griffey RH, Redfield AG, Loomis RE, Dahlquist FW. Nuclear magnetic resonance observation and dynamics of specific amide protons in T4 lysozyme. *Biochemistry* 1985; 24(4):817-822.

(276) Lopukhov LV, Ponomareva AA, Yagodina LO. Inhibition of [(15)N]Valine Transamination during Selective Labeling of Barstar in a T7 Polymerase System. *Biochemistry (Mosc)* 2004; 69(7):806-808.

(277) Mueller GA, Kirby TW, DeRose EF, London RE. NMR assignment of protein side chains using residue-correlated labeling and NOE spectra. *J Magn Reson* 2003; 165(2):237-247.

(278) John DM, Weeks KM. van't Hoff enthalpies without baselines. *Protein Sci* 2000; 9(7):1416-1419.

(279) Morton CJ, Campbell ID. SH3 domains. Molecular 'Velcro'. *Curr Biol* 1994; 4(7):615-617.

(280) Mayer BJ. SH3 domains: complexity in moderation. *J Cell Sci* 2001; 114(Pt 7):1253-1263.

(281) Nooren IM, Thornton JM. Diversity of protein-protein interactions. *EMBO J* 2003; 22(14):3486-3492.

(282) Nayal A, Webb DJ, Brown CM, Schaefer EM, Vicente-Manzanares M, Horwitz AR. Paxillin phosphorylation at Ser273 localizes a GIT1-PIX-PAK complex and regulates adhesion and protrusion dynamics. *J Cell Biol* 2006; 173(4):587-589.

(283) Jockusch BM, Isenberg G. Vinculin and alpha-actinin: interaction with actin and effect on microfilament network formation. *Cold Spring Harb Symp Quant Biol* 1982; 46 Pt 2:613-623.

(284) Ziegler WH, Liddington RC, Critchley DR. The structure and regulation of vinculin. *Trends Cell Biol* 2006; 16(9):453-460.

(285) Clore GM, Tang C, Iwahara J. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol* 2007; 17(5):603-616.

(286) Peterson RD, Theimer CA, Wu H, Feigon J. New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. *J Biomol NMR* 2004; 28(1):59-67.