Investigation of binding preferences and identification of novel binding partners for the SH3 domains of the multifunctional adaptor protein CD2AP

Evgenia Rouka

University College

University of Oxford

Thesis submitted for the *Doctor of Philosophy* degree

Hilary 2014

Content

Figures and Tables

Investigation of binding preferences and identification of novel binding partners for the SH3 domains of the multifunctional adaptor protein CD2AP

Evgenia Rouka

University College, Submitted for the *DPhil* degree, Hilary 2014

Abstract

CD2AP is a member of the CD2AP/CIN85 family of adaptors and involved in several cellular processes, such as kidney podocyte development, actin-mediated membrane trafficking and T-cell activation. It contains three SH3 domains whose binding properties and interaction partners remain largely unexplored.

The CD2AP SH3 interaction with the novel partner Rab5-activating GEF RIN3 was studied extensively by isothermal titration calorimetry (ITC), peptide scanning arrays, mutagenesis and X-ray crystallography. Mapping of the interaction regions showed that human RIN3 contains two binding sites for the CD2AP SH3 domains. From these studies, the CD2AP SH3 recognition motif P-x-P/A-x-x-R emerged. Two crystal structures (1.65 Å and 1.2 Å) of the SH3-1 and SH3-2 domains in complex with RIN3 epitopes 1 and 2 respectively revealed that these residues serve as anchoring points. With the aid of bioinformatics tools, this motif was used to conduct a peptide array-based screen for additional signalling partner candidates. One of the hits was the Arf-GAP ARAP1. ITC data indicate that the three SH3 domains differentially recognise three ARAP1 epitopes, with the first ARAP1 epitope binding to SH3-2 in the nanomolar range. A crystal structure (1.6 Å) of the SH3-2 domain in complex with the first ARAP1 epitope implicates two additional anchoring residues that extend beyond the PPII helical region of the canonical motif. The CD2AP/ARAP1 interaction was confirmed in podocytes and cancer cells at the endogenous protein level. Even though RIN3 and ARAP1 are involved in membrane trafficking, a direct link to CD2AP had not been reported before. Other candidates from the peptide array analyses were also investigated by ITC.

In conclusion, this study led to the elucidation of the CD2AP SH3-1 and SH3-2 domain binding signatures and the identification of putative novel binding partners for all three SH3 domains. Lastly, insight was gained into the binding preferences of CD2AP SH3-3.

Acknowledgements

I am grateful to my supervisor, Stephan Feller, for giving me the opportunity to study for a DPhil in his lab and for his support. I am also grateful to Philip Simister for his support throughout the project, and time and patience to teach me the experimental methods. I would like to thank both for the helpful discussions and introducing me into science. Moreover, I am indebted to my parents for their support to study for a DPhil. I am also thankful to the MRC, Funds for Women Graduates and Stephan Feller for the additional funding.

Many thanks to previous lab members Tassos Konstantinou and Melanie Janning for their preliminary work for my DPhil project; Kathrin Kirsch and Rudolf Volkmer for vital materials; Stefan Knapp and Frank von Delft (Structural Genomics Consortium Oxford) for allowing me to use the SGC crystallisation facilities, and my transfer viva examiners, Val Macaulay and Panagis Fillipakopoulos, for the helpful discussion on my project. I am also thankful to University College for supporting me to present my work at conferences.

In addition, I am thankful to members of the WIMM and SGC for their help. Special thanks to lab members Jess, Marc, Toshi, Martin and Gang for their encouragement in the good and bad days of the project. Huge thanks to Katha. Without her friendship, lab times would not have been the same.

Huge thanks to Claudia, Jess, Lucy, Teresa, Lyle and Adam for four memorable years!

Lastly, I would like to thank friends, especially Electra, Kiki and Marina, and family in Greece for their encouragement.

Abbreviations

ALIX	ALG-2 interacting protein X
APS	Ammonium persulphate
ARAP1	Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1
BSA	Bovine serum albumin
c-CBL	E3 ubiquitin-protein ligase CBL
CD2AP	CD2-associated protein
CIN85	Cbl-interacting protein of 85 kDa
DLS	Dynamic Light Scattering
DMSO	Dimethyl sulfoxide
DOC	Deoxycholate
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
EGFR	Epidermal growth factor receptor
EVH1	Enabled/VASP (vasodilator-stimulated protein) Homology-1
FBS	Fetal bovine serum
FSGS	Focal segmental glomerulosclerosis
Grb2	Growth factor receptor-bound protein 2

GSH-WB	Glutathione Sepharose -bead wash buffer
HCS	Hampton Core Screen
HIN	Hampton Index screen
HLB	Hypotonic lysis buffer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
JCSG	Joint Centre for Structural Genomics core screen
LB	Luria-Bertani medium
LFS	Ligand Friendly Screen
NP-40	Nonident P-40 buffer
PBS	Phosphate-buffered saline
pDC	Plasmacytoid dendritic cell
PID	Protein interaction domain
PMSF	Phenylmethylsulfonylfluoride
PODs	Immortalised human podocytic cell line
PPII	Poly-L-proline-II helix
PRD	Proline-rich sequence recognition domain
RIN3	Ras and Rab interactor 3
RTKs	Receptor tyrosine kinases

SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SH2	Src homology 2
SH3	Src homology 3
TB	Terrific browth medium
TBS	Tris buffered saline
TCR	T-cell receptor
TEMED	Tetramethyl ethylene diamine
TGFβ	Transforming growth factor β
TXB	1% Triton X-100 buffer
UEV	Ubiquitin E2 variant
β-ME	β-mercaptoethanol
ΔG	Change in free energy
ΔH	Change in enthalpy
ΔS	Change in entropy

Chapter 1: Introduction

1.1 Cell Signalling

Cells need to constantly monitor their environment in order to be able to survive and grow. Nutrients, growth factors, hormones, light, temperature, pH or pathogens are among the environmental parameters that cells sense and respond to. The cell-sensing mechanisms include a large number of receptor families, often located on the cell surface, such as receptor tyrosine kinases (RTKs), G-protein-coupled receptors (GPCRs), T-cell receptors (TCR) and the transforming growth factor β (TGFβ) superfamily (Acuto *et al*. 2008, Lemmon *et al*. 2010, Audet *et al*. 2012, Wrana 2013). Apart from external sensing mechanisms, cells have developed internal sensors as well. For instance, cells sense DNA damage through nuclear protein complexes, such as the Mre11-Rad50-Nbs1 (MRN) complex (Lee *et al*. 2005). Another example is the regulation of oxygen homeostasis by the hypoxia-inducible factor 1 (HIF-1) and HIF-prolyl 4-hydroxylases (HIF-P4Hs) (Semenza 2007).

Different cells have developed a variety of complex signalling networks to integrate and further process these signals in a highly coordinated manner. As an example, key components of the EGF (epidermal growth factor) receptor pathway and their multiple roles are depicted in Figure 1.1. Signalling is mediated by small molecules, such as ions, phospholipids and peptides as well as numerous proteins. Also, various types of smaller or longer non-protein-coding RNA molecules are implicated in gene expression regulation and cell defence mechanisms (Carthew *et al*. 2009, Malone *et al*. 2009, Ponting *et al*. 2009). The outputs resulting from the ‘computational’ processes

within signalling networks are typically appropriate cellular responses, such as changes in cell metabolism, proliferation and differentiation. Membrane trafficking is also instrumental for the balance between the degraded and recycled receptors that control these cellular processes (Grant *et al*. 2009, Goh *et al*. 2013).

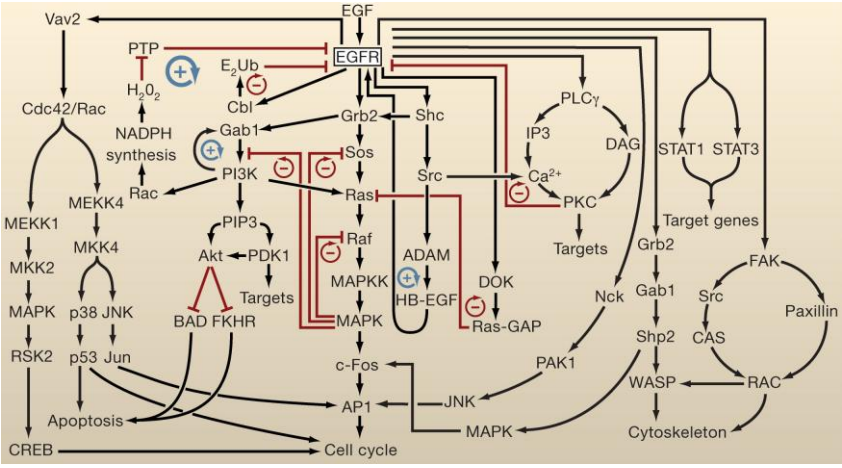


Figure 1.1

Complexity of cell signalling networks.

The epidermal growth factor receptor (EGFR) intracellular pathway is implicated in different cellular processes. A selection of key components and their roles are depicted (Lemmon *et al*. 2010).

1.2 Signal processing in protein complexes

Intracellular signal processing commonly occurs in large multi-protein complexes. It depends to a significant degree on the proper activity of enzymes, which respond to environmental cues and transform the information into appropriate functional outputs (Pawson *et al*. 2003). For example, the protein kinase superfamily consists of roughly 520 members in *Homo sapiens* that phosphorylate their target proteins. Phosphorylation is the addition of a phosphate group to Tyr, Thr and Ser in eukaryotes. Generally,

phosphorylation occurs also on Asp or His in prokaryotes (Fischer *et al*. 1955, Deribe *et al*. 2010). 1.7% of eukaryotic genes are devoted to protein phosphorylation. Out of the 520 human protein kinases, 90 kinases are tyrosine kinases and 50 kinases may lack catalytic activity (Manning *et al*. 2002). Protein phosphatases, another enzyme superfamily, dephosphorylate their targets. There are 140 protein phosphatases in *Homo sapiens* (Tonks 2006, Shi 2009). The perturbation of protein phosphorylation, which is regulated by kinases and phosphatases, may lead to disease, such as cancer. This shows how critical the control of these processes is and the need to have a good understanding of their regulation (Futreal *et al*. 2004, Deribe *et al*. 2010). For a long time, the catalytic domains of these enzymes were thought to be the sole elements critical for enzymatic activity. However, it is now well understood that such enzymes frequently contain other segments, which are also critical for target recognition and multi-protein complex formation (Pawson *et al*. 2003, Good *et al*. 2011). These segments interact with regulatory signalling proteins, which lack catalytic activity and exclusively mediate protein-protein interactions. They are classified into three broad and partially overlapping categories: adaptor, scaffold and large multi-site docking proteins (Figure 1.2). Adaptor proteins bridge two signalling proteins. This is done by establishing direct (Figure 1.2A) or indirect interactions between the signalling proteins and inducing signalling cascades. Scaffold proteins serve as signalling platforms, in which two or more signalling partners of a pathway bind to (Figure 1.2B). Docking proteins also serve as signalling platforms, but in a phosphorylation-dependent manner, and are found in the cell plasma membrane, or they translocate to the membrane upon activation. Once

located in the plasma membrane, they may bind to activated receptors and activate further one or more signalling pathways (Figure 1.2C) (Buday *et al*. 2010, Simister *et al*. 2012).

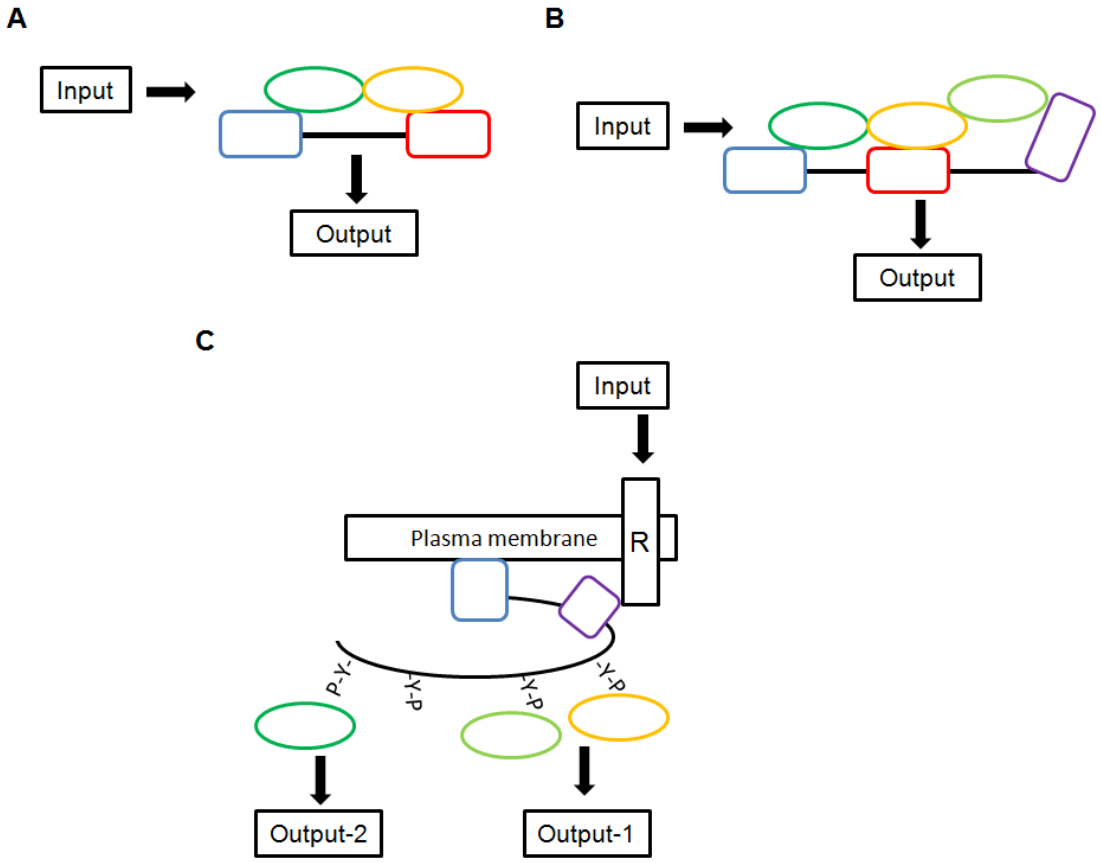


Figure 1.2

Schematic diagrams of non-catalytic mediators of protein-protein interactions.

Adaptors **(*A*)**, scaffolds **(*B*)** and large multi-site docking proteins **(*C*)**: Structured domains are represented by rectangles; unstructured (intrinsically disordered) regions are shown as black lines. Oval shapes represent interaction partners. Please note that some docking proteins lack any folded domains and are stably inserted into the plasma membrane.

The formation of multi-protein complexes is controlled spatially and temporally (Good *et al*. 2011) (Figure 1.3A). Typically these complexes are formed in an environment characterized by macromolecular crowding (Gershenson *et al*. 2011). Protein concentrations in the cytoplasm of cells can apparently reach

up to 300 mg/ml (Figure 1.3B). Therefore, it is challenging for each protein to recognise its targets and to avoid binding non-specifically to other proteins (McGuffee *et al*. 2010).

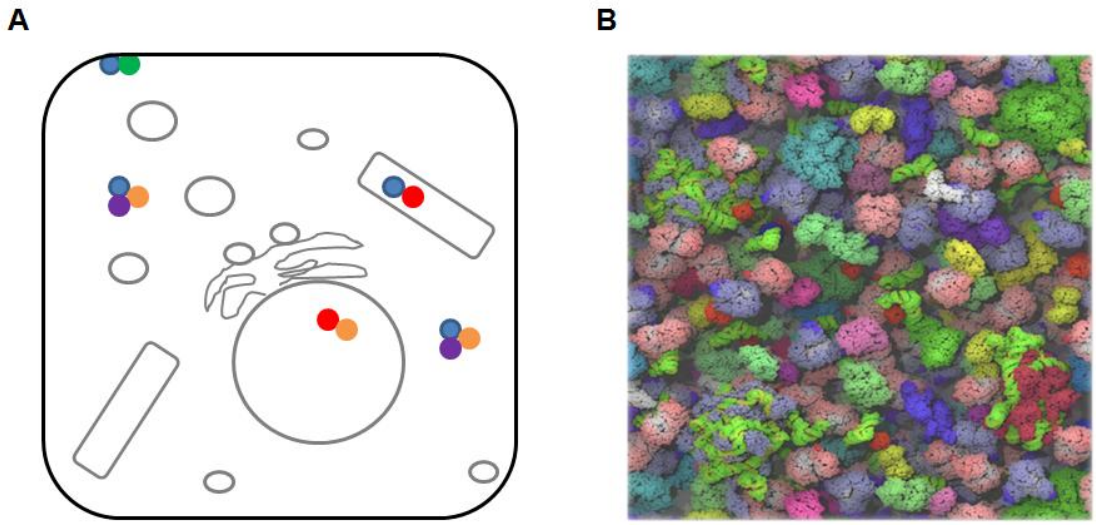


Figure 1.3

Cell compartmentalization and macromolecular crowding are two of the factors that prominently affect cell signalling.

A, Schematic structure of a cell. Black solid line: cell membrane; grey shapes: cell compartments; coloured circles: proteins. ***B***, Model of the *E. coli* cytoplasm and its components (McGuffee *et al*. 2010).

To achieve a high level of specificity, cells might use adaptors and related molecules (Good *et al*. 2011). These interact through their protein interaction domains (PIDs), of which more than 200 have been reported so far (Letunic *et al*. 2012). PIDs are characterised by independent folds, which recognise lipids, post-translationally modified motifs, other target motifs or nucleic acid sequences. A selection of inducible and non-inducible PIDs and their typical recognition motifs are depicted in Figure 1.4.

For instance, SH2 (Src Homology 2) domains recognise specific motifs that contain a phosphorylated Tyr residue, while SH3 (Src homology 3) domains

typically recognise proline-rich motifs (Pawson *et al*. 2003). Domains can bind to other domains, short linear motifs in their target proteins, or, less often, to composite epitopes. Domain-domain interactions are often characterised by large contact surfaces and tight binding. On the other hand, domain-epitope interactions are commonly characterised by smaller contact interfaces and moderate affinity. This allows the quick assembly and disassembly of the interactors, which is a prerequisite for highly dynamic signalling processes (Stein *et al*. 2009).

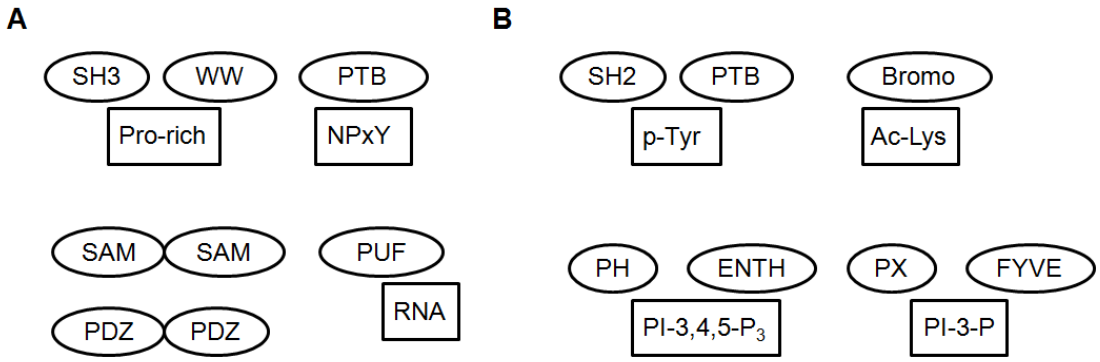


Figure 1.4

Selected protein interaction domains (PIDs), which provide non-inducible (*A*) or inducible (*B*) associations, and their recognition sequences.

Domains: oval shapes; Domain targets: rectangular shapes.

Prior to binding, linear peptides can be structured or disordered. When the peptides are structured, they adopt secondary structures, such as 3_{10} and poly-L-proline-II (PPII) helices (London *et al*. 2010) (Figure 1.5). Such helical structures are not predicted by standard structure prediction programs (Simister *et al*. 2012). The 3_{10} and PPII helices are the third most frequent (2-3%) protein secondary structure elements after the α-helix and β-strand. The left-handed PPII helix is characterised by three residues per turn (Adzhubei *et al*. 2013). The right-handed 3_{10} helix also has three residues per

turn. Hydrogen bonds between residues *i* and *i+3* are critical for its formation leading to a 10-atom ring structure (hence the name 3_{10} helix). 3_{10} helices of increasing length become thinner and irregular (i.e. exhibit varied stereochemical parameters) compared to short, regular 3_{10} helices. This makes them inherently unstable. Due to their instability, 3_{10} helices are usually only up to eight residues long (Pauling *et al*. 1951, Enkhbayar *et al*. 2006, Vieira-Pires *et al*. 2010).

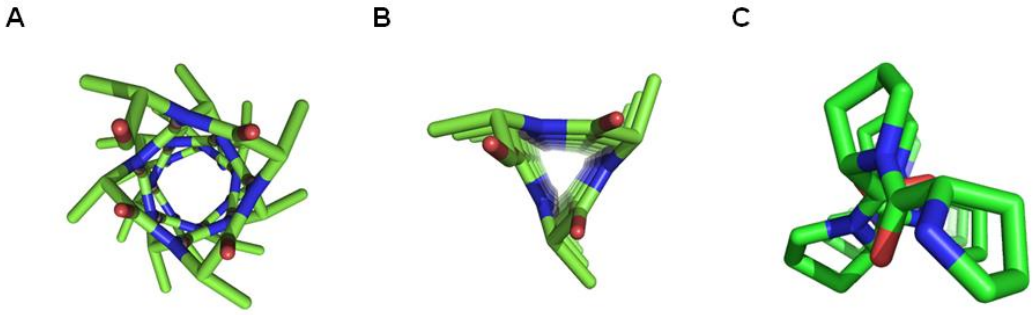


Figure 1.5

Helical secondary structures in proteins.

A, α-helix; ***B***, 3_{10} helix; ***C***, poly-L-proline II (PPII) helix (Vieira-Pires *et al*. 2010).

The high number of PIDs enables the interaction between different proteins leading to different functions. Also, point mutations within the PID recognition motifs may result in the rewiring of existing signalling networks and the formation of new ones (Good *et al*. 2011). For example, p53 mutations lead to mutant p53 proteins with oncogenic activity rather than mutants, which have only lost their wild-type tumour suppressing activity (Muller *et al*. 2013). These characteristics, which enable plasticity in signalling networks, could potentially lead to a lack of specificity in some cases. However, cells appear to be able to exploit the plasticity offered by PIDs without compromising their specificity (Good *et al*. 2011). This is achieved by multiple, complementary mechanisms.

PIDs have different sequences, folds, and recognition motifs. However, it should be noted that different PIDs have been reported to recognise identical recognition motifs (Kaneko *et al*. 2011). Another mechanism is the co-enrichment of regulatory signalling proteins in specific compartments. For example, membrane trafficking leads to the recycling or degradation of cell surface receptors by vesicle transport. Therefore, membrane trafficking regulates signalling by receptors and their biological responses. The formation of complexes of specific scaffold and related molecules is critical for proper vesicle sorting (Palfy *et al*. 2012). Moreover, the affinity of PIDs for their partners contributes to specificity. The low-to-moderate affinities of their interactions require that several interactions have to be combined in a timely fashion to lead to a biological response (Good *et al*. 2011). Moreover, post-translational modifications of proteins, such as phosphorylation and ubiquitinylation, and their recognition by specific PIDs, regulate signal transduction and contribute to many cellular processes (Deribe *et al*. 2010).

Numerous experimental and computational methods are employed to find novel protein-protein interactions, such as yeast two-hybrid screens, affinity purification and co-evolution pattern searches. Each method has advantages and drawbacks. Also, all methods are characterised by significant levels of false positives (30-60%) and false negatives (40-80%). This may result from protein misfolding, degradation or nonselective interactions (von Mering *et al*. 2002, Aloy *et al*. 2006, Liu *et al*. 2012). Therefore, a combination of several different methods increases confidence that the discovered protein-protein interactions are important *in vivo*.

In this thesis, peptide array overlay blots, isothermal titration calorimetry, protein X-ray crystallography and co-immunoprecipitations were used to detect and confirm novel SH3 domain-mediated protein-protein interactions and uncover the SH3 binding modes.

1.3 Proline-rich sequence recognition domains (PRDs)

Proline-rich sequences are abundant in prokaryotes and eukaryotes (Rubin *et al*. 2000, Chandra *et al*. 2004), and participate in the formation of many protein complexes (Freund *et al*. 2008). As mentioned earlier, modular domains are the building blocks of proteins with adaptor, scaffold or docking functions. A number of modular domains that recognise proline-rich sequences have been identified: SH3 (Src homology 3) (Mayer *et al*. 2004), WW (named after two signature tryptophan residues) (Macias *et al*. 2002), UEV (ubiquitin E2 variant) (Pornillos *et al*. 2002), EVH1 (Enabled/VASP [vasodilator-stimulated protein] Homology-1) (Peterson *et al*. 2006), GYF (named after the conserved Gly-Tyr-Phe tripeptide) (Kofler *et al*. 2006), CAP (cytoskeleton-associated protein)-Gly (Saito *et al*. 2004) domains and profilin (Mahoney *et al* 1999).

SH3 and WW domains are the most abundant PRDs in the human proteome (Castagnoli *et al*. 2004). Even though they differ in their three-dimensional fold and amino acid sequence, they both contain crucial aromatic residues, which form hydrophobic grooves that bind to their ligands. These PRDs bind typically to ligands adopting a PPII helical conformation with an affinity of 0.1-500 μM. This conformation was initially identified in proline-rich stretches. However, segments that lack prolines can also adopt this conformation and

bind to PRDs. The propensity of the PPII helix to be formed at the protein surface contributes to its participation in PRD-mediated interactions (Li *et al*. 2005, Freund *et al*. 2008). The 2-fold rotational pseudosymmetry of the PPII helix allows it in principle to bind to PRDs in opposite directions. PPII helices are found in ordered as well as disordered protein regions (Adzhubei *et al*. 1987, Adzubei *et al*. 2013).

Even though PRDs recognise different consensus sequences, there is often a ligand overlap between different PRDs. The flat surface of the PRD docking grooves, their moderate binding affinity and the PPII helix recognition by many of them contribute to their cross-reactivity. Specificity is gained due to the flanking regions of the consensus motifs, the loop variability of the PRDs and cell compartmentalization (Li *et al*. 2005, Freund *et al*. 2008).

In the following section, the characteristics of PRD folds other than the SH3 domain will be briefly mentioned. Examples of complexes of these PRDs with ligands are depicted in Figure 1.6. Also, the subgroups of these PRDs and their recognition motifs are shown briefly in Table 1.1. The SH3 features will be described extensively in a subsequent section.

WW domains typically encompass ca. 40 amino acids, and form a triple stranded β-sheet (Figure 1.6, top right). They are the smallest binding domain known to date (Bork *et al*. 1994, Chen *et al*. 1995). Some WW domains recognise proline-rich sequences that have a phosphorylated Tyr or Thr. According to their recognition preferences, they are divided into four groups (Macias *et al*. 2002). Examples of each group are displayed in Table 1.1. Although the four groups have somewhat different binding mechanisms, in all

cases, the prolines of the peptides dock onto the WW hydrophobic groove (Macias *et al*. 2002). WW appear as single entities or in tandem and can bind independently or synergistically (Sudol *et al*. 2005, Chong *et al* 2010).

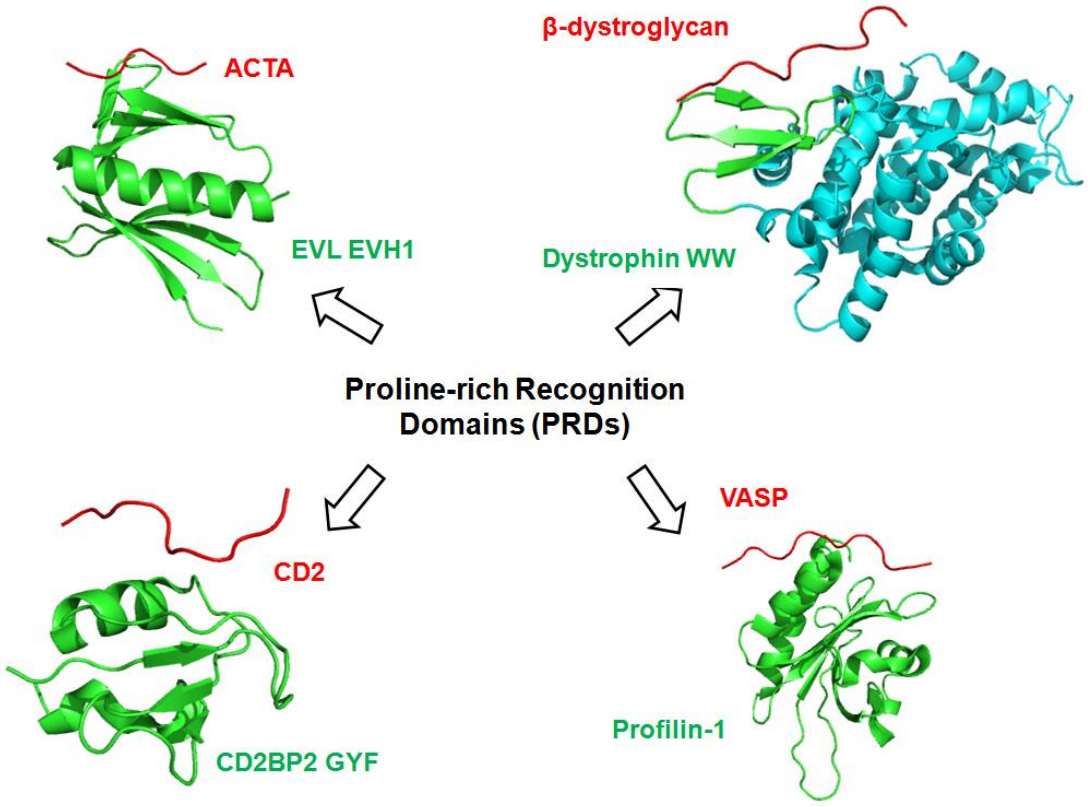


Figure 1.6

Non-SH3 proline-rich recognition domains (PRDs), whose structures have been solved.

EVH1 (top left, PDB ID: 1QC6), WW (top right, PDB ID: 1EG4), GYF (bottom left, PDB ID: 1L2Z) and profilin (bottom right, PDB ID: 1PAV). The PRDs are coloured green. Their ligands are depicted in red. In the case of dystrophin (top right), the protein chain outside of the WW domain is shown in light blue.

UEV domains typically recognise ubiquitin. However, in contrast to other UEVs, the UEV of human TSG101 (Tumor susceptibility gene 101) folds differently at its termini and binds to the proline-rich P-T/S-A-P motif of the HIV-1 L-domain (Garrus *et al*. 2001, VerPlank *et al*. 2001). The proline-rich

and ubiquitin binding sites are located in different regions of the domain (Pornillos *et al* 2002).

Table 1.1 PRDs (GYF, WW, EVH1, CAP-Gly and profilin) and their typical recognition motifs.

PRD	Interaction type	Protein	Binding motif	Binding candidate	Reference
GYF	CD2BP2-type	Hs CD2BP2	PPGW	CD2	Hahn *et al*. 1993, Freund *et al*. 2002
	SMY2-type	Sc SMY2	PPGΦ	MSL5	Rutz *et al*. 2000
WW	Type-I	Hs DMD	PPxY	β-dystroglycan	Huang *et al*. 2000
	Type-II	Sc FBP28	PPLP	Synthetic peptide	Macias *et al*. 2000
	Type-III	Hs NPW38	PGR	NPWBP	Komuro *et al*. 1999
	Type-IV	Hs PIN1	pS-Pro-rich peptide	RNA polymerase II large subunit	Verdecia *et al*. 2000
EVH1	Ena/ VASP-type	Hs VASP	(F/L)PPPP	zyxin, vinculin	Ball *et al*. 2000
	WASP-type	Hs WASP	extended poly-proline epitope	WIP	Volkman *et al*. 2002
	Homer/ Vesl-type	Hs HOMER1	PPxxF	mGluR	Tu *et al*. 1998, Beneken *et al*. 2000
	SPRED-type	Hs SPRED3	Unknown	Unknown	Kato *et al*. 2003
CAP-Gly	-	Hs CYLD	SQRRSPP EEPPDF	NEMO	Saito *et al*. 2004
Profilin	-	Hs HPP	Poly-proline peptide	Synthetic peptide	Mahoney *et al*. 1999

Hs: Homo sapiens; Sc: Saccharomyces cerevisiae.

The EVH1 (Ena VASP Homology-1) domain belongs to the Pleckstrin Homology (PH) domain-like superfamily. It adopts the PH fold, which is characterised by two anti-parallel β-sheets forming a β-sandwich followed by a C-terminal α-helix (Blomberg *et al*. 1999, Peterson *et al*. 2006) (Figure 1.6, top left). EVH1s are ca. 115 amino acids long and usually found at the N-termini of proteins. They can be divided into four categories, whose binding preferences are displayed in Table 1.1. EVH1s bind to proline-rich ligands that adopt a triangular PPII orientation to dock onto the EVH1 hydrophobic groove. Ligands do not adopt this orientation to bind to SH3, WW or GYF (Peterson *et al*. 2006).

The CD2BP2 (CD2 antigen cytoplasmic tail-binding protein 2) GYF domain was initially identified as binding a CD2-derived proline-rich peptide (Nishizawa *et al*. 1998). GYFs consist of an α-helix and a β-sheet that are connected by loops (Figure 1.6, bottom left). A hydrophobic and a positively charged specificity pocket are both critical for the binding of the ligands (Freund *et al*. 2002). According to their proline-rich binding preferences, GYFs can be divided into two categories (Table 1.1). The CD2BP2-type GYFs bind to tryptophan-bearing peptides (Hahn *et al* 1993), and the SMY2-type GYFs bind to an extended epitope that lacks a Trp (Rutz *et al* 2000). In some cases, GYFs bind to the same target peptides as SH3 or WW domains (Kofler *et al*. 2006). For example, a CD2-derived peptide binds to CD2BP2 GYF and Fyn SH3 (Freund *et al* 2002).

The CAP-Gly domain is approximately 80 amino acids long. Originally it was found in four cytoskeleton-associated proteins (CAPs) (Riehemann *et al*. 1993). CAP-Gly consists of five anti-parallel β-sheets and a β-hairpin. Single

or tandem CAP-Glys bind among others to tubulin or microtubules. Their typical recognition motif is E-E-Y/F (Steinmetz *et al*. 2008). However, a proline-rich NEMO (NF-κB essential modulator) peptide binds to the hydrophobic groove of the third CAP-Gly of CYLD (Cylindromatosis gene) (Table 1.1). Other CAP-Gly recognition motifs have also been identified (Steinmetz *et al*. 2008).

Profilins are 15 kDa proteins that usually bind to proline-rich ligands (Tanaka *et al*. 1985, Mahoney *et al*. 1999) and actin monomers (Kang *et al*. 1999). The conserved profilin fold consists of seven-stranded anti-parallel β-sheets and four α-helices (Figure 1.6, bottom right). Hydrophobic interactions and hydrogen bonds contribute to the interactions between different profilin isoforms and poly-proline peptides (Mahoney *et al*. 1999, Haikarainen *et al*. 2009) (Table 1.1).

1.4 SH3 domains

SH3 domains typically consist of roughly 60 residues. They are abundant in the human proteome and more than 300 SH3s have been found (Castagnoli *et al*. 2004). SH3 domains mediate the formation of various complexes in vital intracellular signalling processes that regulate cell proliferation, apoptosis and cell migration, intracellular transport and organisation of the cell architecture (Mayer *et al*. 2004). Gene ontology analysis of the human SH3 domain interactome, which was constructed according to literature mining data, indicated that the SH3 domain-containing proteins and their partners may be enriched in cellular processes, such as cytoskeleton organization, endocytosis and regulation of cell death (Carducci et al. 2012). Also, gene ontology

analysis of the putative SH3 domain interactome of *Saccharomyces cerevisiae* indicated that the SH3 domain-containing proteins and their binding partners may be enriched in cell polarity and endocytic processes. Yeast two-hydrid screen, peptide array and phage display binding data from the 25 of 27 buddy yeast SH3 domains were integrated to construct this SH3 interactome (Tonikian *et al*. 2009). Moreover, SH3 domains are utilised in pathogen-host interactions. For instance, HIV-1 Nef participates in the cell infection by HIV-1. Its mode of action remains unknown. However, it was found to bind to Src family kinases, such as Fyn in T-cells (Arold *et al*. 1997). Furthermore, SH3-mediated interactions can affect the enzymatic activity of the proteins in which they reside. For example, the Abl SH3 domain forms intramolecular contacts with the Abl SH2-kinase linker region, which leads to the inactivation of the enzyme (Nagar *et al*. 2003).

In several cases, SH3 domain amino acid chains have been reported to partially insert into the fold of other SH3 domains. Even though domain swapping can regulate protein function (Rousseau *et al*. 2012), the biological relevance of the documented SH3 swappings has not been understood yet. Four instances of intertwined SH3s have been found so far (Harkiolaki *et al*. 2006, Camara-Artigas *et al*. 2014).

The SH3 domain β-barrel fold is shared by multiple other, sequentially unrelated domains, including nucleotide-binding domains like the RNA-binding Sm-like (LSm) domains, (Theobald *et al*. 2005). SH3-like folds are also found in prokaryotes (Whisstock *et al*. 1999).

The abundance of SH3 domains and their implication in signal transduction led to an extensive analysis of their structures, binding properties and specificities. At present, experimental and computational methods are used to predict and screen for novel SH3 interaction partners. Such approaches should in time result in a much better understanding of the SH3 domain-mediated interaction network (Kay *et al*. 2012, Liu *et al*. 2012). Variants of the WISE (whole interactome scanning experiment) method, such as PATS (peptide array target screening), are used extensively to screen for SH3 domain interactions. These methods take advantage of the ability of peptide arrays to test simultaneously a large number of possible interactors (Landgraf *et al*. 2004, Wu *et al*. 2007). Other experimental methods include the yeast two-hybrid screen, phage display and mass spectrometry. The above can be combined with computational methods, such as trained algorithms, to increase the confidence level of the predicted interactions (Tonikian *et al*. 2009, Hou *et al*. 2012, Zhang *et al*. 2012, Liu *et al*. 2012).

Originally, SH3s were identified in PLC-γ (phosphatidyl-inositol-specific phospholipase C gamma) and the v-CRK oncogene (Stahl *et al* 1988, Mayer *et al*. 1988). A few years later, structural views of the spectrin and Src SH3 folds were obtained. Structurally, SH3s consist of five anti-parallel β-strands forming a β-barrel. The first 3 strands are connected by the RT-, N-Src- and distal loops respectively, while the last two strands are connected by a 3_{10} helix (Musacchio *et al*. 1992, Yu *et al*. 1992) (Figure 1.7).

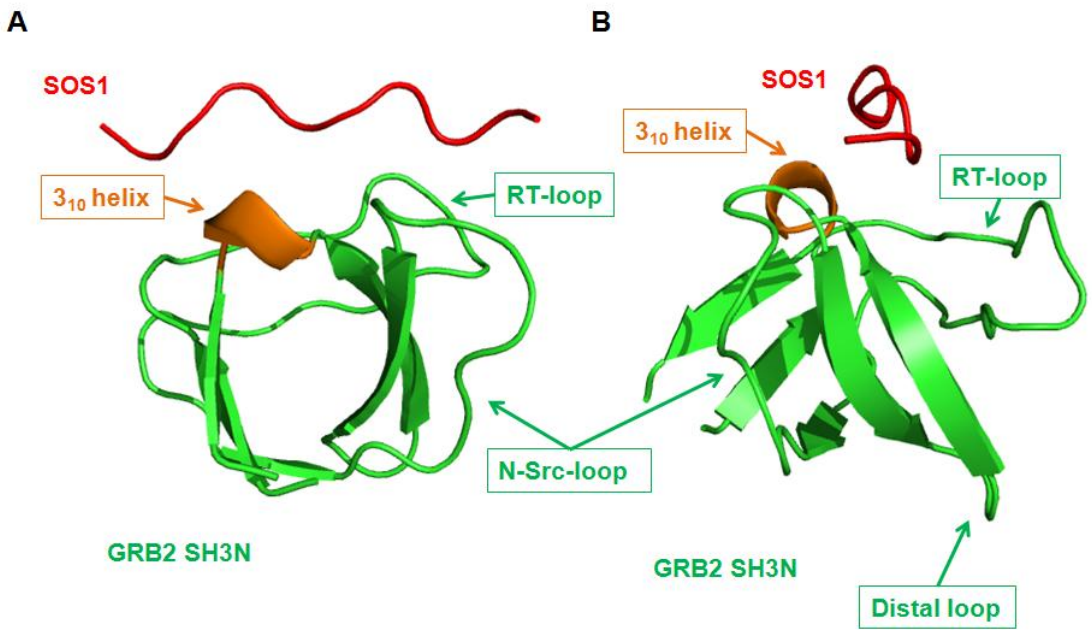


Figure 1.7

Secondary structure of the Grb2 (growth factor receptor-bound protein 2) SH3N domain in complex with its SOS1 peptide ligand.

A, β-barrel structure of the SH3 domain, ***B***, Anti-parallel β-strands of the SH3 domain. The loops and β-strands of the SH3 domain are coloured green, its 3_{10} helix in orange and its ligand in red. (PDB ID: 1GBQ).

The hydrophobic core is conserved among SH3s, while the loops exhibit higher variability (Mayer *et al*. 2004). An additional α-helix is occasionally seen in SH3s (termed as *hSH3*), such as in the ADAP (adhesion- and degranulation-promoting adaptor) protein (Heuer *et al*. 2004). This additional structural feature is involved in the hSH3 recognition of lipids rather than proline-rich regions (Heuer *et al*. 2005).

In the first studies, a typical SH3 domain core binding motif x-P-x-x-P-x was found for a number of SH3 domains, such as those of the Src and Fyn kinases. As later described for other PRDs, the ligands adopt a PPII helical structure (Figure 1.7). This core motif is often flanked by a basic residue, either at its N-terminus (Class I orientation) or C-terminus (Class II

orientation). For instance, the SH3 recognition motif may be R-x-x-P-x-x-P-x or x-P-x-x-P-x-R. The xP dipeptides dock onto two hydrophobic grooves on the SH3 domain, while the basic residue is accommodated in an acidic SH3 pocket, which contributes to the specificity of the interaction. The SH3 binding surface for the ligand is formed by the variable SH3 N-Src and RT-loops, and the 3_{10} helix, while a conserved SH3 tryptophan at the SH3-ligand binding interface is critical for the determination of the ligand's orientation (Ren *et al*. 1993, Feller *et al*. 1994, Mayer *et al*. 2004, Fernandez-Ballester *et al*. 2004, Saksela *et al*. 2012). These two SH3 binding surfaces (hydrophobic grooves and acidic pocket) may be conformationally coupled, and an ensemble of complete and partially engaged SH3-ligand complex structures exist. The mutation of critical residues, which bind to the SH3 surfaces, and other residues within the motif, which do not contact the SH3 surface, results in a decreased ligand population, which forms PPII helices and binds to the SH3 domains (Ferreon et al. 2004, Stollar *et al*. 2012, Krieger *et al*. 2014). Apart from the typical P-x-x-P-x-R and R-x-x-P-x-x-P SH3 recognition motifs, atypical proline-rich motifs, such as the P-x-x-D-Y EPS8 (epidermal growth factor receptor pathway substrate 8) SH3 motif (Aitio *et al*. 2008), have been identified. Even though SH3 recognition motifs, which lack prolines, have been reported, these motifs are flanked by prolines (Carducci *et al*. 2012, Saksela *et al*. 2012). For instance, the C-terminal SH3 domains of the adaptor protein Grb2 and its relative Mona/Gads recognise R-x-x-K core motifs, which are flanked by prolines that contribute to binding (Harkiolaki *et al*. 2009). Also, a subset of SH3 domains binds to ubiquitin (Stamenova *et al*. 2007).

Apart from structural data, thermodynamic data have also been obtained for SH3 domain-mediated interactions. Isothermal titration calorimetry (ITC) is a method that allows the determination of thermodynamic parameters upon complex formation, and compliments the structural data interpretation. As will be explained in section 2.2.5, affinity and hence the change in free energy (ΔG) of an interaction is measured by ITC as well as the change in enthalpy (ΔH). The entropy change (ΔS) can then be determined. ΔH is associated with the number of noncovalent bonds from the uncomplexed to the complexed state. ΔS informs on the changes in the degrees of freedom of a system upon complex formation. Therefore, the quantification of ΔH and ΔS provides additional information on a binding event (Doyle 1997). These data are usually combined with structural data to understand better the SH3 binding signature.

According to the above-mentioned structural data, the interaction between the SH3 hydrophobic grooves and their ligands would lead to a major and favourable change in entropy. The increase of the degrees of freedom of the system would be due to the liberation of bound water molecules from the SH3 surface upon complex formation. Moreover, solvent molecules compete with ligand for binding to the SH3 surface. This would lead to a small favourable change in enthalpy. However, it is well established that the SH3 domain-mediated interactions are characterised by major and favourable changes in enthalpy and small changes in entropy. To our knowledge, unfavourable changes in enthalpy have not been identified for the SH3 domain-mediated interactions. This is in apparent disagreement with the predominantly hydrophobic character of their interactions according to the

structural data. This discrepancy has not been explained fully yet(Ladbury *et al*. 2011).

As mentioned above, there are ensembles of the SH3 domain ligands that adopt partially or completely the PPII conformation in the uncomplexed and states (i.e. pre-formed binding conformation of the ligand prior to the SH3-mediated interaction). For instance, this was shown for the Grb2 SH3C interaction with a Gab2 fragment (Krieger *et al*. 2014). If the ligand adopts the PPII conformation in the uncomplexed state, there would be observed a small change in the flexibility of the system, and hence the ΔS (Ladbury *et al*. 2011). Others suggested that the adoption of a specific conformation by the ligand to enable SH3 binding would lead to loss of flexibility of the system and unfavourable ΔS. This was reported for some SH3 domain interactions, such as the Sem-5 SH3 domain with a SOS (son of sevenless) peptide in *Caenorhabditis elegans* (Ferreon *et al*. 2004, Ladbury *et al*. 2011). Lastly, the lack of water displacement prior to complex formation may lead to small changes in ΔS, and major and favourable ΔH. This was shown for the interactions between the hydrophobic pocket of MUP (major urinary protein) and its ligands (Barratt et al. 2006). However, a similar effect is not well established for the SH3-domain mediated interactions. So far, the contribution of a water-mediated network to the favourable ΔH was shown for the Abl SH3 domain interactionwith a 3BP1 (SH3 domain-binding protein 1)-derived mutant peptide(Palencia *et al*. 2004, Zafra-Ruano *et al*. 2012).

Apart from the recognition of PPII helices by SH3s, tertiary contacts with their ligands have also been reported. For example, the Fyn SH3 binds to the SAP (SLAM-associated protein) SH2 domain in T-cells. This atypical interaction is

mediated mainly by polar contacts and a hydrophobic contact between an SH3 Trp and the side chain of an SH2 Arg (Chan *et al*. 2003). Additionally, PPII helices mediate interactions between SH3 and other folded domains. The Fyn SH3 binds to a PPII helix and also another part of the HIV-1 Nef protein core domain (Arold *et al*. 1997).

SH3 domains participate in stable as well as transient interactions. They commonly bind ligands with modest affinities in the mid to low micromolar range (1 to 100 μM) (Li *et al*. 2005). However, some SH3 domains bind to their ligands with nanomolar affinity, such as the Mona/Gads SH3C in complex with a SLP-76 (SH2 domain-containing leukocyte protein of 76 kDa) epitope peptide (K_d 118 nM). This SLP-76-derived peptide formed a 3_{10} rather than a PPII helix upon binding to the SH3 (Harkiolaki *et al*. 2003). In contrast, the Mona/Gads SH3C recognises a HPK1 (Hematopoietic Progenitor Kinase 1)-derived peptide by a different mechanism and with lower affinity. The affinity of this interaction is in the low micromolar range, and adjacent RxxK (forming a 3_{10} helix) and PxxP (forming a PPII helix) motifs in HPK1 contribute to the interaction (Lewitzky *et al*. 2004). This documents how versatile SH3 domains can be in binding to different ligands.

Moreover, the recognition of tandem PxxP motifs by SH3 domains was reported. The IRTKS (Insulin Receptor Tyrosine Kinase Substrate) SH3 binds to two tandem PxxP motifs in EspFU (*E. coli*-secreted protein F-like encoded on prophage U) with an affinity of 500 nM (Aitio *et al*. 2010).

Proteins encompassing more than one SH3 are quite common. However, their binding mode in the context of full-length proteins is not well

characterised (Li *et al*. 2005). $p47^{phox}$ is one of the regulatory subunits of the NADPH oxidase. It contains two SH3 domains in tandem, which form a single binding surface and bind to an extended region of their linker. This leads to the autoinhibition of $p47^{phox}$. The affinity of the tandem SH3s is 1.5 μM, while it is undetectable for the single SH3s (Groemping *et al*. 2003, Yuzawa *et al*. 2004). A similar surface is used by the tandem SH3 domains to bind to their interaction partner $p22^{phox}$ upon NADPH oxidase assembly. The $p22^{phox}$-derived peptide is capable of binding to $p47^{phox}$ SH3N, but its affinity is lower compared to the tandem $p47^{phox}$ SH3 domains (Groemping *et al*. 2003, Ogura *et al*. 2006).

Despite the wealth of structural and biochemical information available on SH3 domain complexes, the development of small molecule SH3 domain inhibitors with activity *in vivo* has not been successful so far. The affinity of current early stage inhibitors is low, which is in part explained by the relatively shallow binding grooves of the SH3 domains (Vidal *et al*. 2001, Atatreh *et al*. 2008, Simister *et al*. 2013).

The examples mentioned above document the SH3 domains' versatility and promiscuity. RT- and N-Src loop variability, high local domain concentration, surface-surface contacts, cell-type specific effects all further contribute to generating binding specificity *in vivo* (Groemping *et al*. 2003, Chan *et al*. 2003, Li *et al*. 2005). Post-translational modifications of the flanking regions can also modify SH3 binding (Bedford *et al*. 2000). Specificity may be gained by negative selection as well. Negative selection refers to the inability of related domains to bind to a given motif. For example, there could be exclusive recognition of a single binding motif by one SH3 domain, to the

exclusion of all other SH3 domains within that organism. This was shown for the Sho1 SH3/Pbs2 interaction, which is implicated in the osmolarity response pathway in *Saccharomyces cerevisiae*. Of this species' 27 SH3 domains, only Sho1 SH3 can bind to Pbs2. This implies that alongside the recognition determinants in the motif that allow positive selection of Pbs2, there are also features of the Pbs2 binding region that prevent it binding to 26 additional SH3 domains, *i.e.* these are concomitantly negatively selected. However, it should be noted that other mechanisms, such as the ones mentioned above, contribute to SH3-domain mediated interactions in *S. cerevisiae* (Zarrinpar *et al*. 2003, Li 2005).

1.5 CD2AP

1.5.1 CD2AP structure

CD2AP (CD2-associated protein) is a ubiquitously expressed adaptor protein, and part of the small CIN85/CD2AP family (Dikic 2002). The product of the CD2AP gene is also known as CMS or METS-1 (p130Cas ligand with multiple SH3 domains, or mesenchyme-to-epithelium transition protein with SH3 domains). However, the designation CD2AP is most common and will be used throughout this thesis. CD2AP consists of three tandem SH3 domains located in the N-terminal half of the protein (Figure 1.8), which are 40-50% homologous to each other (Figure 1.9). In addition, CD2AP contains one proline-rich region, four actin-binding sites and a coiled-coil region at its C-terminus. The rest of the protein is made up of disordered linker regions.

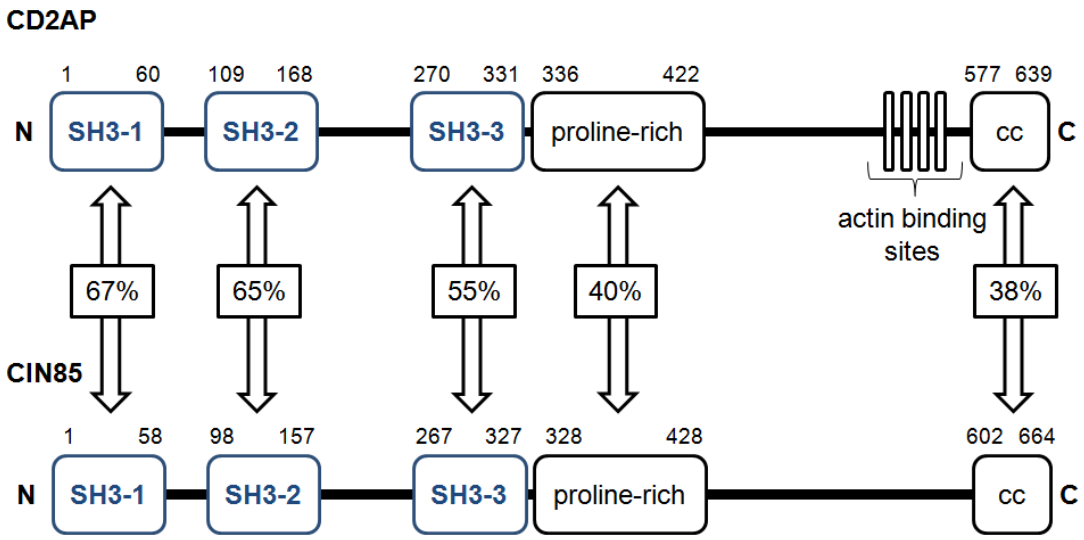


Figure 1.8

Schematic structures of CD2AP and CIN85

The percentages refer to the amino acid identity between two domains or regions, as indicated by the arrows.

Figure 1.9

Sequence alignment of the CD2AP SH3 domains.

Red background, completely conserved residues; blue boxes, residues of similar charge. The alignment was generated using Clustal Omega (Sievers *et al*. 2011) and the image was created with ESPript 3 (http://espript.ibcp.fr, Robert *et al*. 2014).

CIN85 stands for Cbl-interacting protein of 85 kDa, and is also known as SH3KBP1 (SH3 domain-containing kinase-binding protein 1), Ruk (Regulator of ubiquitous kinase) or SETA (SH3 domain-containing gene expressed in tumorigenic astrocytes) (Havrylov *et al*. 2010). As shown in Figure 1.8, CD2AP and CIN85 share a similar architecture. However, CIN85 lacks the four actin-binding sites.

1.5.2 CD2AP functions

The functions of CD2AP are summarised in Figure 1.10, and also described in detail in the following sections. The known interaction partners of CD2AP are shown in Table 1.2.

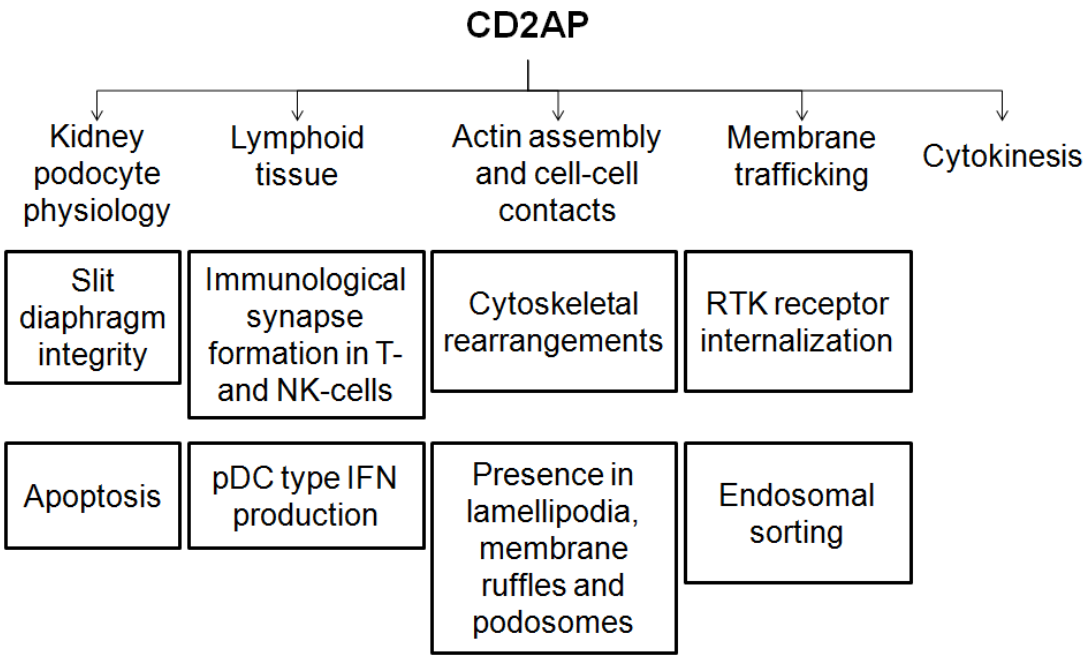


Figure 1.10

CD2AP functions.

Table 1.2 CD2AP interaction partners.

CD2AP partner	Description	CD2AP epitope	Reference
Cytokinesis			
ANLN	Anillin	SH3-1, SH3-2	Monzo *et al*. 2005
Generic binding protein			
ALIX	ALG-2 interacting protein X	N-terminus	Usami *et al*. 2007
CIN85	SH3 domain-containing kinase-binding protein 1	C-terminus	Hutchings *et al*. 2003, Gaidos *et al*. 2007
DAB1	Disables homolog 1	N-terminus	Sato *et al*. 2007
Kinases			
Fyn	Tyrosine-protein kinase Fyn	unclear	Kirsch *et al*. 1999, Huber *et al*. 2006
p85α	Phosphatidylinositol 3-kinase regulatory subunit alpha	N-terminus	Huber *et al*. 2003
Src	Proto-oncogene protein kinase tyrosine Src	unclear	Kirsch *et al*. 1999
Yes	Protein tyrosine kinase Yes	unclear	Kirsch *et al*. 1999
Known podocytic function			
DDN	Dendrin	N-terminus	Asanuma *et al*. 2007, Yaddanapudi *et al*. 2011
NPHS1	Nephrin	C-terminus	Shih *et al* 2001, Huber *et al*. 2003
NPHS2	Podocin	C-terminus	Schwartz *et al*. 2001
PKD2	Polycystin-2	C-terminus	Lehtonen *et al*. 2000
SEPT7	Septin 7	SH3-3, SH3-2	Wasik *et al*. 2012
SHIP2	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2	N-terminus	Hyvonen *et al*. 2010
SYNPO	Synaptopodin	N-terminus	Huber *et al*. 2006
Phosphatases			
PSTPIP1	Proline serine threonine phosphatase-interacting protein 1	Proline-rich region	Badour *et al*. 2003

SHIP1	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1	SH3-1	Bao *et al*. 2012
Cytoskeleton/endocytosis/membrane trafficking			
CAPZ/CP	F-actin-capping protein subunit alpha	C-terminus	Hutchings *et al*. 2003, Zhao *et al*. 2012
CFBP	Multivesicular body subunit 12A	N-terminus	Konishi *et al*. 2007
ASAP1	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 1	SH3-1, SH3-2	Liu *et al*. 2005
CTTN	Src substrate cortactin	Proline-rich region	Lynch *et al*. 2003, Zhao *et al*. 2012
F-actin	F-actin	C-terminus	Gaidos *et al*. 2007
p130Cas	p130Cas/BCAR1	unclear	Kirsch *et al*. 1999
Rab4	Ras-related protein Rab-4A	unclear	Cormont *et al*. 2003
RAC1	Ras-related C3 botulinum toxin substrate 1	N-terminus	van Duijn *et al*. 2010
SH3GL	Endophilin	Proline-rich region	Lynch *et al*. 2003
Receptor			
CD2	CD2 receptor	SH3-1	Dustin *et al*. 1998, Moncalian *et al*. 2006, Ceregido *et al*. 2013
Ubiquitin ligases			
c-CBL	E3 ubiquitin-protein ligase CBL	SH3-1, SH3-2	Kirsch *et al*. 2001, Lynch *et al*. 2003, Cormont *et al*. 2003, Kobayashi *et al*. 2004
CBL-B	E3 ubiquitin-protein ligase CBL-B	SH3-1	Moncalian *et al*. 2006, Ceregido *et al*. 2013
CBL-C	E3 ubiquitin-protein ligase CBL-C	N-terminus	Tsui *et al*. 2008, Calco *et al*. 2014

1.5.2.1 Role of CD2AP in kidney podocyte physiology

The podocyte is a specialised, differentiated epithelial cell type within the glomerular filtration barrier of the kidney. Apart from the podocyte, the glomerular filtration barrier consists of fenestrated endothelial cells and the glomerural filtration membrane. The integrity of the glomerular filtration barrier is critical for urine formation, and podocytes control the plasma permeability (Schell *et al*. 2012, Akchurin *et al*. 2014). Each podocyte consists of a cell body, primary processes and foot processes. Neighbouring podocytes are linked by an intercellular junction (i.e. the podocyte slit diaphragm), which is formed between the podocyte foot processes (Figure 1.11A, B). CD2AP participates in the formation of the podocyte slit diaphragm by anchoring other major components of the diaphragm, such as nephrin and podocin, to the actin cytoskeleton (Figure 1.11C) (Grahammer *et al*. 2013, Akchurin *et al*. 2014).

CD2AP has a similar expression pattern in human and murine kidneys. It is in both cases specifically expressed in podocytes within the glomerulus (Tienari *et al*. 2005). Its *Drosophila* orthologue (Cindr) is a major component of the insect nephrocyte, a cell type similar to the mammalian glomerular podocyte (Weavers *et al*. 2009). Homozygous CD2AP knockout mice exhibit severe kidney pathology after 3 weeks of age, and die due to renal failure within 6 to 7 weeks. The podocyte is the first structure within the murine kidney to be affected by CD2AP absence (Shi *et al*. 1999).

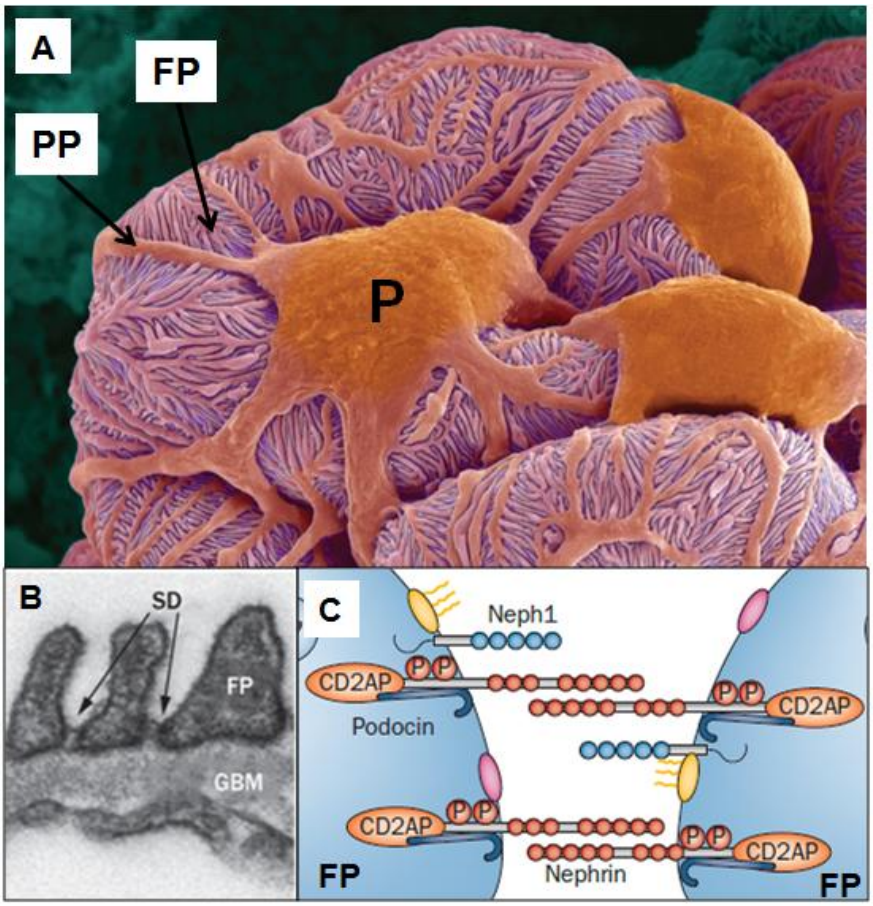


Figure 1.11

Composition of the podocyte.

A, Scanning electron micrograph of the podocyte; ***B***, View of the slit diaphragm and glomerular basement membrane by transmission electron microscopy; ***C***, Schematic representation of a selection of components of the podocyte slit diaphragm. P, cell body; PP, primary processes; FP, foot processes; SD, slit diaphragm; GBM, glomerular basement membrane (Welsh *et al*. 2012, Grahammer *et al*. 2013).

Transgenic mice, which express CD2AP in podocytes and muscle cells, do not display any kidney pathology. In section 1.5.4.2, it will be further described that the testis is the additional tissue, which was affected in this mice (Grunkemeyer *et al*. 2005). This confirms that the CD2AP knockout phenotype of renal failure is due to podocyte dysfunction. Interestingly, even CD2AP heterozygous knockout mice exhibit glomerular abnormalities at 9 months of age, which are similar to the kidney pathology in 3-week-old

homozygous knockout mice. The heterozygous mice are more susceptible to glomerular injury compared to wild-type mice due to defects in late endosome (MVB) formation (Kim *et al* 2003).

Conditionally immortalised podocytes derived from CD2AP knock-out mice exhibit higher levels of apoptosis compared to wild-type podocytic clones. A number of groups tried to explain the anti-apoptotic role of CD2AP.

It could be a consequence of the CD2AP SH3 domain interaction with the p85 subunit of PI3K, which links nephrin to the PI3K/AKT survival pathway (Huber *et al*. 2003, Schiffer *et al*. 2004). However, prolonged AKT activation results in the activation of the pro-apoptotic Smad2/3 pathway as well, which is independent of the anti-apoptotic CD2AP/PI3K/AKT pathway (Xavier *et al*. 2009). Also, CD2AP deficiency results in the activation of the pro-apoptotic p38 MAPK pathway. However, CD2AP-independent pro-apoptotic p38 MAPK pathways are also present (Schiffer *et al*. 2004). Moreover, there exists controversy about CD2AP effects on AKT and ERK1/2 phosphorylation between wild-type and CD2AP knock-out cells. Upon CD2AP deficiency and stimulation with epidermal growth factor (EGF) (Huber *et al*. 2003) or TGFβ (Schiffer *et al* 2004), cells exhibited decreased AKT phosphorylation. The activated ERK1/2 levels did not change. Other groups suggested that the AKT phosphorylation levels did not change after EGF or insulin stimulation between the two clones (Schiffer *et al*. 2004, Tossidou *et al*. 2007), and that ERK1/2 phosphorylation terminated earlier upon application of the above-mentioned stimuli, or platelet-derived growth factor, or hepatocyte growth factor stimulation. FGF (Fibroblast growth factor) or VEGF (Vascular

endothelial growth factor) stimulation resulted in weaker phosphorylation levels of AKT and ERK1/2 (Tossidou *et al*. 2007).

In addition, the CD2AP/dendrin interaction is implicated in podocytic apoptosis. CD2AP absence leads to TGFβ-dependent cathepsin L expression, which degrades components of the slit diaphragm (Yaddanapudi *et al*. 2011). Furthermore, podocytes are insulin-responsive, and transgenic mice lacking the insulin receptor in podocytes exhibit podocytic injury and diabetic nephropathy-like histology features (Welsh *et al*. 2010). The GTPase Septin 7, which may regulate glucose transport, and SHIP2, which is involved in PI3K signalling upon insulin stimulation, are also CD2AP SH3 interaction partners (Hyvonen *et al*. 2010, Wasik *et al*. 2012). Finally, CD2AP mRNA expression is upregulated during mesenchyme-to-epithelium transition in the kidney (Lehtonen *et al*. 2000).

Taken together, these findings indicate that CD2AP is absolutely critical for kidney podocyte stability.

1.5.2.2 CD2AP involvement in lymphoid system

Immunohistochemistry analyses of human tissues suggest that CD2AP has a more restricted expression pattern in lymphoid system. CD2AP is highly expressed in plasmacytoid dendritic cells (pDCs) and is also present in immature (cortical thymocytes) T-cells (Marafioti *et al*. 2008, Rizvi *et al*. 2012). pDCs are antigen-presenting cells that secrete type 1 interferon (Santana-de Anda *et al*. 2013). CD2AP is also highly enriched in murine pDCs. CD2AP-deficient pDCs exhibit migration defects under conditions of inflammation and higher levels of actin polymerization compared to normal

pDCs. After a viral stimulus, defects in leukocyte and B-cell migration to the lymph nodes were also observed (Srivatsan *et al*. 2013). The published results on the contribution of CD2AP to type 1 interferon production are controversial. *Srivatsan et al.* (2013) reported that CD2AP is dispensable for pDC development and type 1 interferon (IFN) production in mice. However, another group suggested that the CD2AP/SHIP1 complex inhibits c-CBL activity in a human pDC-derived cell line. This results in the activation of the BDCA2 (pDC receptor)/FcεR1γ-induced ITAM signalling, which would lead to lower type 1 IFN production (Bao *et al*. 2012).

Even though CD2AP is not found in human mature T-cells (Rizvi et al. 2012), it was found in Jurkat cells to link CD2 to actin polymerization via a CD2/CD2AP-PSTPIP1/WASp cascade (Badour *et al*. 2003) or a CD2/CD2AP/CAPZ cascade (Hutchings *et al*. 2003). Moreover, CD2AP and RAG knock-out mice, which lack functional T and B lymphocytes, exhibit similar renal failure as CD2AP knock-out mice. Furthermore, histological examination of transgenic mice that express CD2AP exclusively in podocytes have normal T-cell activation and unaffected inflammatory responses (Grunkemeyer *et al*. 2005).

In summary, CD2AP is mainly expressed in pDC cells within the lymphoid system. Also, CD2AP is not involved in mature T-cell physiology, but may be involved in acute T cell leukemia.

1.5.2.3 CD2AP involvement in actin assembly, membrane trafficking, cell-cell contacts and cytokinesis

CD2AP is involved in actin polymerization (Welsch *et al*. 2005) and recruited into different actin structures, such as podosomes (Gaidos *et al*. 2007), lamellipodia and membrane ruffles (van Duijn *et al*. 2010, Zhao *et al*. 2012). CD2AP is also found at the leading edge of cells and contributes to the stability of cell-cell contacts (Gaidos *et al*. 2007, van Duijn *et al*. 2010, Tang *et al*. 2013). CD2AP acts by linking proteins to the cell periphery and cytoskeleton. For example, it links cortactin and CAPZ to Rac1 (van Duijn *et al*. 2010) or the Arp2/3 complex (Zhao *et al*. 2012).

Apart from its participation in dynamic actin assembly, CD2AP participates in receptor tyrosine kinase (RTK) internalization and downregulation. It is involved in EGFR internalization by linking cortactin and endophilin to c-CBL (Lynch *et al*. 2003). Without ligand stimulation, the CD2AP/c-CBL complex is recruited to HER2 but not EGFR (Minegishi *et al*. 2013). A c-CBL/CD2AP complex was suggested to be involved in Flt-1 downregulation after VEGF stimulation (Kobayashi *et al*. 2004). CD2AP is also involved in the ubiquitination and degradation of Ret51 by forming a complex with c-CBL via its SH3 domains. This was studied in podocytes and primary neurons (Calco *et al*. 2014). Moreover, CD2AP is found in Rab4-positive vesicles, and may contribute to the sorting from early to late endosomes (MVBs) (McCaffrey *et al*. 2001, Cormont *et al*. 2003, Welsch *et al*. 2005, Gauthier *et al*. 2007).

CD2AP may also be involved in cytokinesis (Monzo *et al*. 2005, Morita *et al*. 2007). *Monzo et al.* (2005) reported that CD2AP knockdown affected the abscission step of daughter cells, while *Morita et al.* (2007) reported the

deficiency defects as moderate. Furthermore, CD2AP participates in the TGF-β3/TGFβR1/ERK signalling cascade in primary Sertoli cells. The additional recruitment of TAB1 (TAK-1 binding protein-1) adaptor protein results in the activation of the proapoptotic p38 MAPK pathway (Xia *et al*. 2006).

1.5.3 CD2AP and CIN85

As already mentioned in section 1.5.1, CD2AP is part of the small CD2AP/CIN85 family (Dikic 2002). CD2AP and CIN85 are both ubiquitously expressed. CD2AP has one predominant isoform, while several isoforms have been reported for CIN85 (Dikic 2002, Havrylov *et al* 2010, Medway *et al*. 2013). Both of the proteins contribute to similar processes, such as cytoskeletal rearrangements and membrane trafficking, by binding to a variety of partners (Dikic 2002, Havrylov *et al*. 2009, Havrylov *et al*. 2010, Medway *et al*. 2013). They have numerous common binding partners, such as CFBP (Konishi *et al*. 2007) and c-CBL (Soubeyran *et al*. 2002), although the binding efficiency between CD2AP and CIN85 might be different for the same partner, such as for F-actin (Gaidos *et al*. 2007). However, there exist proteins that bind to CD2AP and not CIN85, such as CBL-C (Calco *et al*. 2014).

Also, CD2AP and CIN85 could have either overlapping or indeed antagonising functions. Both proteins contribute to the internalization of receptors such as EGFR or HER2 (Haglund *et al*. 2002, Lynch *et al*. 2003, Minegishi *et al*. 2013, Sato *et al*. 2013). The mechanisms through which they contribute to receptor internalization could be independent. Even though the knockdown of CIN85 or CD2AP reduces the internalised HER2 levels compared to wild-type cells, their simultaneous knockdown resulted in unaffected HER2 levels (Minegishi *et al*. 2013).

On the other hand, different effects might occur upon CD2AP or CIN85 binding to the p85 subunit of PI3 kinase. CD2AP/p85 activates the PI3K pathway (Huber *et al*. 2003), while CIN85/p85 inhibits PI3K activity *in vitro* (Gout *et al*. 2000). Another example of antagonism between CD2AP and CIN85 is found in the podocyte. When CD2AP is depleted in podocytes, the CIN85 localization pattern resembles the native CD2AP pattern (Grunkemeyer *et al*. 2005, Tossidou *et al*. 2007). As already mentioned in section 1.5.2.1, CD2AP binds to nephrin and podocin and stabilises the slit diaphragm. In the absence of CD2AP, CIN85 is suggested to internalise nephrin and podocin and destabilise the slit diaphragm (Tossidou *et al*. 2010). The latter might be inhibited due to CD2AP-dependent CIN85 SUMOylation (Tossidou *et al*. 2012). Apart from the kidney glomerulus, CD2AP and CIN85 have distinct non-overlapping localization patterns within testis. These are the two tissues that are clearly most affected by CD2AP deficiency.

Taken together, CD2AP and CIN85 may usually have redundant roles. However, they may antagonise each other in some cases, such as in the podocyte.

1.5.4 CD2AP involvement in disease

1.5.4.1 CD2AP and renal disease

CD2AP mutations are found in a small number of patients with focal segmental glomerulosclerosis (FSGS) (Kim *et al*. 2003, Lowik *et al*. 2007, Gigante *et al*. 2009). These are summarized in Table 1.3.

Table 1.3 CD2AP mutations in patients with focal segmental glomerulosclerosis (FSGS).

CD2AP mutation	CD2AP region	Reference
K301M	SH3-3	Gigante *et al* 2009
T374A	Proline-rich region	
E525 deletion	Linker between proline-rich and actin binding sites	
R612Stop	Coiled-coil region	Lowik *et al* 2007

FSGS is a kidney glomerular disorder, which causes nephrotic syndrome and end-stage renal disease. The pathophysiology of this disease is not fully understood yet. Podocytes are affected by this disorder (Schell *et al*. 2012, Akchurin *et al*. 2014).

CD2AP expression is also downregulated in Chinese children with minimal change nephrotic syndrome (MCNS) and IgA nephropathy (IgAN) (Mao *et al*. 2006).

1.5.4.2 Insights into CD2AP functions from animal studies

As already described in section 1.5.2.1, CD2AP knock-out mice exhibit kidney defects. Apart from this, male mice are infertile when lacking CD2AP in the basal seminiferous tubule (Sertoli cells and immature spermatocytes). The testicular integrity is unaffected by this (Grunkemeyer *et al*. 2005). A potential role of CD2AP in spermatocytes development has not been studied yet.

1.5.4.3 CD2AP and human non-renal diseases

Apart from renal diseases, CD2AP polymorphisms were identified by genome-wide association studies in Alzheimer's and Kashin-Beck (KBD) diseases. The latter is a bone disease (Naj *et al*. 2011, Hollingworth *et al*. 2011, Lambert *et al*. 2013, Yang *et al*. 2014). CD2AP is also expressed strongly in CD4+/CD56+ hematodermic neoplasms and myeloproliferative associated neoplasms, which develop from plasmacytoid dendritic cells (pDCs). Due to the restricted CD2AP expression in pDCs within the lymphoid tissue, CD2AP may be used as a marker for normal and neoplastic pDCs to discriminate pDC from non-pDC cell types (Marafioti *et al*. 2008, Rizvi *et al*. 2012, Kharfan-Dabaja *et al*. 2013). Also, CD2AP was found in a number of solid tumours (Rizvi *et al*. 2012).

1.5.5 CD2AP SH3 domain interactions

Even though CD2AP SH3-mediated interactions contribute to the above-mentioned processes by binding to different partners, the molecular details of their interactions remain largely unexplored. The study of the CD2AP SH3-1 binding mode to CBL-B and CD2 peptides resulted in the identification of the P-x-P/A-x-P-R (x denotes any residue) SH3-1 recognition motif (Moncalian *et al*. 2006). However, controversial results were reported for the peptide positioning on the SH3-1 surface. Moncalian *et al.* (2006) found that two SH3-1 domains bound to CBL-B and CD2 peptides in class I and class II orientation by forming a heterotrimer. However, Ceregido *et al.* (2013) found that SH3-1 forms heterodimers in class I or II orientation with the CBL-B peptide, with class II orientation being preferable. They also found that the same domain forms heterodimers to the CD2 peptide in class II orientation.

Also, the CD2AP SH3 domains bind to ubiquitin (Ortega-Roldan *et al*. 2009, Ortega-Roldan *et al*. 2013).

The CIN85 SH3 domains exhibit the highest similarity for the CD2AP SH3 domains (Figure 1.8). Kowanetz *et al* (2003) reported that CIN85 SH3-1 recognises the P-x-x-x-P-R motif in c-CBL and CBL-B. According to target-assisted iterative screening (TAIS) and mutagenesis studies, the CIN85 SH3 recognition motif is P-x-P/A-x-x-R (Kurakin *et al*. 2003). Similarly to CD2AP SH3-1, controversial results on peptide binding orientations were also reported for the CIN85 SH3-1 domain. CIN85 SH3-1 apparently forms a heterotrimer with a CBL-B peptide in class I and II orientations (Jozic *et al*. 2005, Ceregido *et al*. 2013). However, another group reported that the CBL-B peptide exhibits different binding conformations and forms heterodimers with the CIN85 SH3-1 domain instead. (Ababou *et al*. 2009). Similar to the CD2AP SH3 domains, the CIN85 SH3 domains can bind to ubiquitin (Stamenova *et al*. 2007).

Taken together these findings suggest that the CD2AP and CIN85 SH3-1 domains have similar binding properties. However, it is still not fully understood if and when they recognise their ligands in class I and II orientations. Moreover, the molecular details of the second and third CD2AP SH3 domain-mediated interactions remained unexplored. Studying the binding properties of all three CD2AP SH3 domains in more detail should therefore help to decipher the molecular mechanism by which they selectively interact with their multiple partners in different cellular contexts (Table 1.2).

1.6 Preliminary data for this project

1.6.1 Identification of the recognition motif for the first two CD2AP SH3 domains in c-CBL

Preliminary work prior to this DPhil project was performed by two previous lab members (Dr Tassos Konstantinou, Dr Melanie Janning). Initially, the atypical P-x-P-x-P-R interaction motif in the known partner c-CBL was identified to bind to the first two CD2AP SH3 domains by peptide scanning arrays, subsequent alanine scans and permutation arrays (Dr Tassos Konstantinou). The peptide permutation arrays of the c-CBL epitope, which were probed with the first two SH3 domains, are shown in Figure 1.12. It appears that CD2AP SH3-2 has similar binding preferences to CD2AP SH3-1, because it recognised the same motif in c-CBL. In motif position 3, a Pro_{826} to Ala substitution preserved the interaction. Therefore, ligands with an atypical P-x-P/A-x-P-R motif may be recognised by the first two SH3 domains. This finding is consistent with the published CD2AP SH3-1 recognition motif (Moncalian *et al*. 2006). The arrays were probed with CD2AP SH3-3 as well, but there was not much signal detectable (data not shown). Due to this low affinity no information on the binding preferences of CD2AP SH3-3 were obtained.

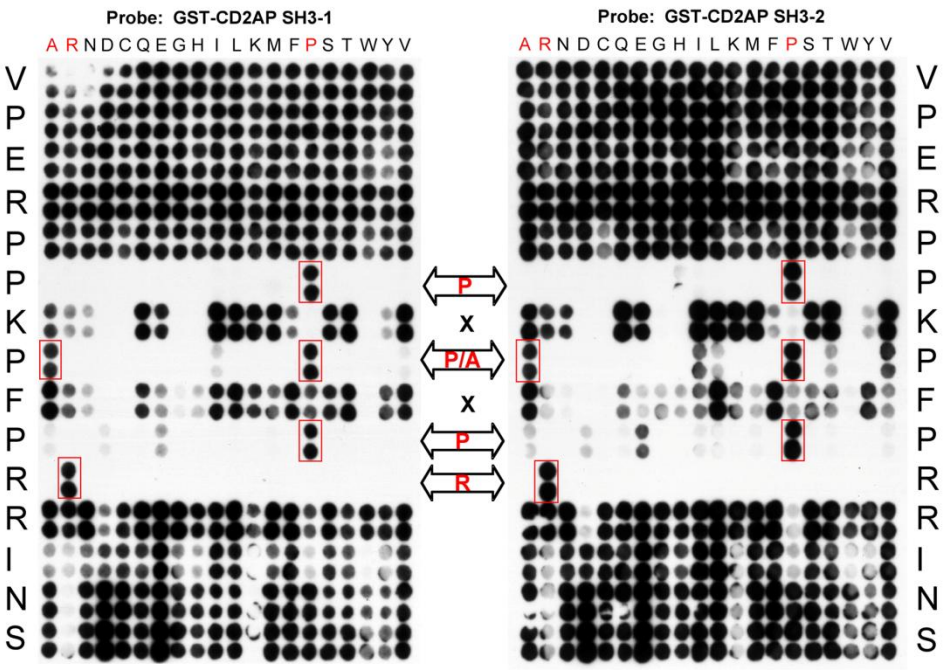


Figure 1.12

The first two CD2AP SH3 domains recognise the P-x-P-x-P-R motif in the c-CBL epitope.

c-Cbl was spot-synthesized as overlapping 21 aa peptides, sliding 3 aa with each step as described in Appendix C.2.1. Two array copies were initially probed with GST and then with GST-CD2AP-SH3-1 and -SH3-2. The positions permutated within the peptides are indicated down the vertical sides of the arrays, and the individual substitutions are listed across the top. The series of spots that gave the strongest signals are shown in red between the two array copies.

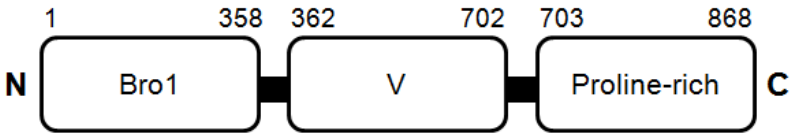
1.6.2 Scan for putative novel CD2AP SH3 interaction partners

The P-x-P-x-P-R motif was used to screen for putative novel CD2AP SH3 interaction partners by peptide scanning array. At that time, this motif was found in over 700 human proteins (ScanProsite bioinformatics tool, http://prosite.expasy.org/scanprosite/). A subset of 70 potential binding sites in proteins with known or potential roles in signalling was further tested using peptide array overlay blots with the three CD2AP SH3 domains (Dr Tassos Konstantinou, data not shown). A brief description is found in Appendix C.2.3. The apparent success rate was 55% for CD2AP SH3-1, 34% for CD2AP

SH3-2 and 8% for CD2AP SH3-3. Most of the CD2AP SH3-1 and SH3-2 hits were similar. The low number of putative SH3-3 hits indicated that CD2AP SH3-3 might exhibit different binding properties. The adaptor protein ALIX (ALG-2 [apoptosis-linked gene 2] interacting protein X) and the GEF (guanine nucleotide exchange factor) RIN3 (Ras and Rab interactor 3) were among the putative hits.

1.6.3 CD2AP/ALIX interaction

ALIX consists of an N-terminal bromo (Bro1) domain, a V domain and a proline-rich region at its C-terminus (Figure 1.13). In the literature, ALIX is also termed AIP1 (ALG-2-interacting protein 1) or PDCD6IP (programmed cell death 6 interacting protein). ALIX is involved in a number of cellular processes through its association with different binding partners (Odorizzi 2006, Bissig *et al*. 2014). These cellular processes are summarized in Figure 1.14, while a selection of interacting proteins and their corresponding ALIX binding regions are depicted in Figure 1.13.



Binding partners:

CMBP4	HIV p6	ALG-2
LBPA	PAR1	CD2AP/CIN85
Src	Syntenin	CEP55
		Endophilin
		Fyn
		Hck
		Src
		TSG101

Figure 1.13

ALIX schematic structure and interaction partners.

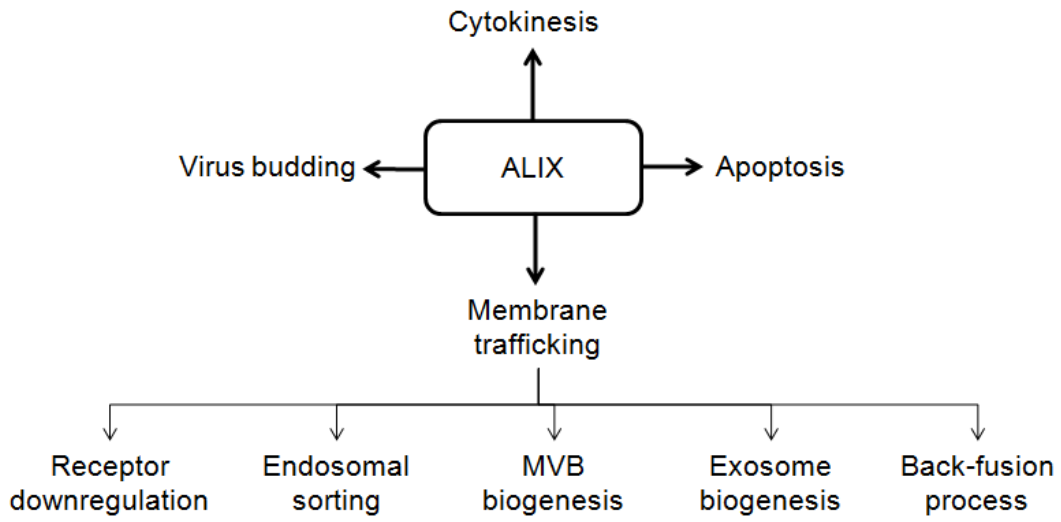


Figure 1.14

Functions of ALIX.

Initially, ALIX was identified as a CIN85 binding partner. The recruitment of ALIX antagonizes, at least in part, EGFR (epidermal growth factor receptor) internalization by the c-CBL/CIN85/endophilin complex and the c-CBL-dependent ubiquitination of EGFR (Schmidt *et al*. 2004, Schmidt *et al*. 2005). CD2AP recognises the same binding site in ALIX as the CIN85 SH3 domains (Figure 1.11, Usami *et al*. 2007). Even though CD2AP and ALIX are involved in similar processes (Figures 1.10 and 1.14), the biological significance of the CD2AP interaction with ALIX has not been determined yet. Their interaction is dispensable for cytokinesis and HIV-1 virus budding (Usami *et al*. 2007, Morita *et al*. 2007). In our research group's initial screen (Dr Tassos Konstantinou), the ALIX epitope was recognised by all three CD2AP SH3 domains (data not shown). In order to define more precisely the recognition motif for the first two SH3 domains as well as for SH3-3, permutation arrays of the ALIX epitope were probed with the three individual SH3 domains (Dr Tassos Konstantinou, Figure 1.15).

According to the permutation arrays, both CD2AP SH3-1 and SH3-2 recognise a P-t-P-x-x-R motif in ALIX (Figure 1.15, left and middle panels). In the case of SH3-1, and similar to the identified SH3-1 recognition motif based on the c-CBL permutation arrays (Figure 1.12), a Pro and Arg are the solely tolerated residues in the first and last position of the motif. However, the arrays indicate that there is tolerance for other residues in positions 2 and 3. In motif position 3, both Ala and Val give similar binding levels with SH3-1 to the native proline in ALIX. In position 2, predominantly aliphatic residues can take the place of Thr without a reduction in signal. On the other hand, motif position 5 is not selective as any residue can be accommodated in this position. Thus, hypothetically, ligands based on some variant of a P-T/V/I-P/A/V-x-x-R motif may also allow SH3-1 binding.

With SH3-2, the array reveals even greater tolerances using this domain. In particular, the substitution of Phe for Pro in motif positions 1 and 3 does not affect signal strength. Again, hypothetically, ligands based on variants of a P/F-T/V/I/L-P/F/A-x-x-R motif could possibly allow SH3-2 binding.

According to the ALIX permutation array, CD2AP SH3-3 recognises an apparent P-t-P-a-P-R-t-x-x-P-x-K motif in ALIX (Figure 1.15, right panel). Similar to the first two CD2AP SH3 domains, Pro and Arg are the single tolerated residues in the first and last position of the motif. In motif position 2, the same tolerance as for the first two SH3 domains is present (Thr, Val, Ile). Similarly in position 3, Ala gives similar binding levels with SH3-3 to the native Pro in ALIX.

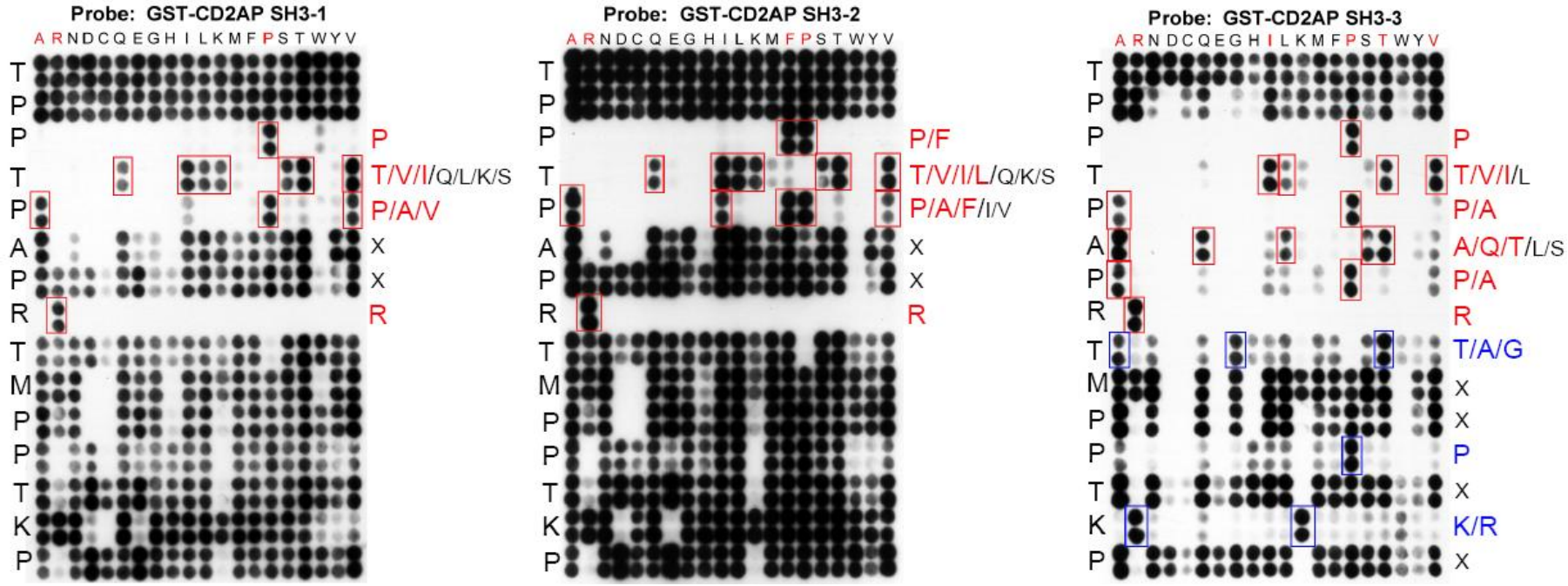


Figure 1.15

The first two CD2AP SH3 domains recognise a similar motif in the ALIX epitope, whereas SH3-3 recognises a more selective and extensive motif.

21 aa peptides of the ALIX epitope AGGHAPTPPTPAPRTMPPTKP were permutated in single positions (underlined) to all other residues in duplicate rows as described in Appendix C.2.2. Three array copies were initially probed with GST and then with GST-CD2AP-SH3-1, -SH3-2 or -SH3-3. The positions permutated within the peptides are indicated down the left sides of the arrays, and the individual substitutions are listed across the top. Peptide spots with highly restricted residues are boxed in red. The amino-acid preferences at these positions are listed down the right side of each array (red font: most strongly binding amino acids; smaller black font: tolerated amino acids). CD2AP-SH3-3 recognises a more extended linear epitope compared to SH3-1 and SH3-2, as shown by the blue boxes. The main amino-acid preferences at these positions are listed on the right-hand side of the array, in blue.

Unlike the first two SH3 domains, the arrays indicate that there is an increased SH3-3 selectivity for other residues in positions 4 and 5. In position 4, Gln and Thr can replace Ala without a signal reduction. In motif position 5, the signal strength is unaffected by the substitution of Pro for Ala. Additionally, SH3-3 selects positively residues in positions 7 (Thr, Ala, Gly), 10 (Pro) and 12 (Lys, Arg). Therefore, hypothetical motif variants based on P-T/V/I-P/A-A/Q/T-P/A-R-T/G-x-x-P-x-K/R, should they exist in proteins, might enable SH3-3 binding. These findings suggest that SH3-3 recognises a more selective and extensive motif in its ligands compared to the first two CD2AP SH3 domains.

1.6.4 CD2AP/RIN3 interaction

RIN3 was detected as a novel putative CD2AP interaction partner in the peptide array overlay blots based upon the CD2AP SH3 interacting motif in c-CBL.

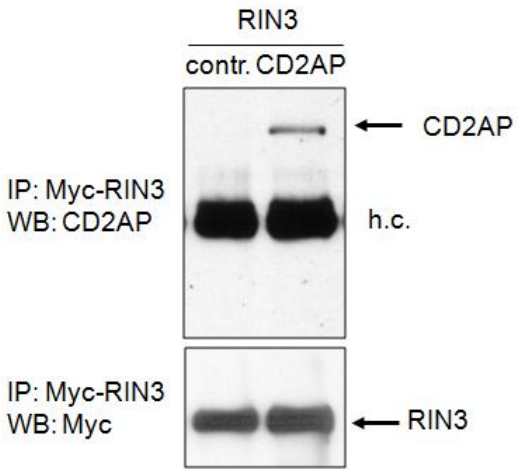


Figure 1.16

Flag-CD2AP and myc-RIN3 co-immunoprecipitate in co-transfected HEK293T cells

The CD2AP/RIN3 interaction was subsequently confirmed by co-immunoprecipitation of co-transfected CD2AP and RIN3 in 293T cells (experiment conducted by our collaborators Dr Kathrin Kirsch *et al*., Boston University, USA; Figure 1.16).

RIN3 is a member of the RIN (Ras and Rab interactors) family, which is involved in membrane trafficking. It consists of an SH2 domain, three proline-rich regions, a Ras-homology domain, a Vps9-domain and a Ras-association domain (Figure 1.17). The proposed functions of RIN3 will be discussed further in Chapter 7.

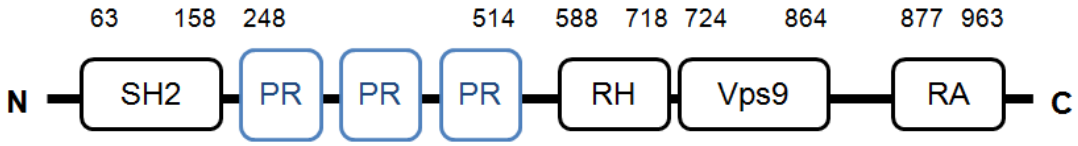


Figure 1.17

Schematic structure of RIN3

SH2, Src Homology 2 domain; RH, Ras Homology domain; RA, Ras Association domain.

The presence of three proline-rich regions in RIN3 prompted us to test whether there are additional CD2AP binding regions in RIN3 using a peptide scanning array covering the entire RIN3 sequence (experiment performed by Dr Melanie Janning, Figure 1.18). The three CD2AP SH3 domains were found to bind to two RIN3 epitopes, and corresponding peptides were synthesised (Nicola O' Reilly, Peptide Synthesis Laboratory, Cancer Research UK, London). Based on the peptide scanning arrays, an epitope is normally identified by the presence of a series of spots (Volkmer *et al*. 2012). In Figure 1.18, this is observed for the first RIN3 binding region using SH3-1 as a

probe. With the other domains, and the second binding region in all cases, there is either one spot or two non-contiguous spots. The lack of a contiguous series of strong spots for each epitope indicates that the RIN3 peptide boundary positions are crucial for their interaction. This binding behaviour is quite peculiar, relative to other SH3 peptide arrays published. Preliminary ITC measurements determined that 16amino acid peptides are sufficient to bind to the SH3 domains (Dr Melanie Janning, data not shown).

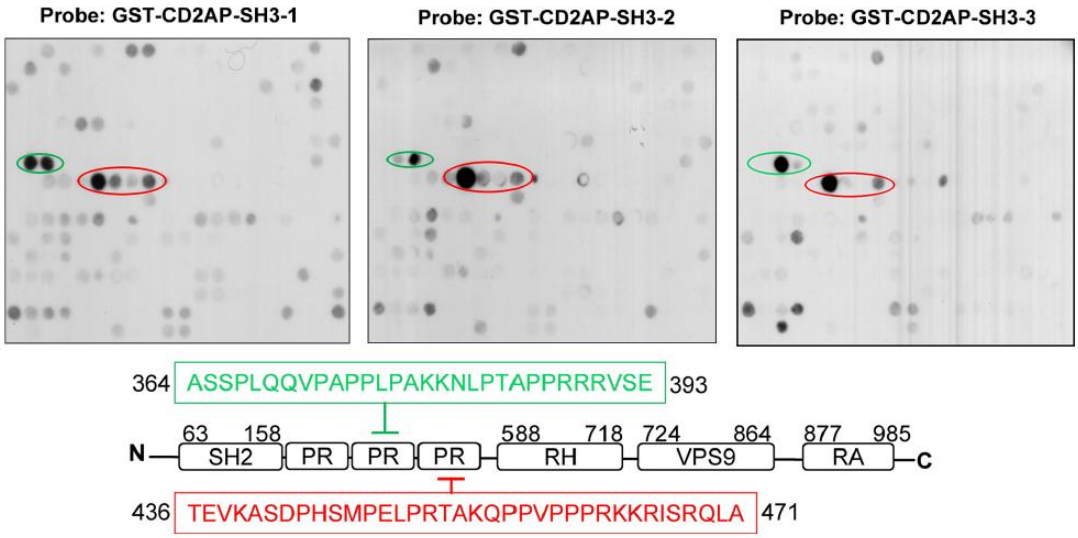


Figure 1.18

The three CD2AP SH3 domains recognise two RIN3 epitopes.

RIN3 was spot-synthesized as overlapping 27 aa peptides, sliding 3 aa with each step as described in Appendix C.2.4. The arrays were initially probed with GST (negative control) and then with GST-CD2AP-SH3-1, -SH3-2 or -SH3-3. Two principal binding regions are evident (ringed in green and red). The boxed sequences below the arrays represent the common segment found in all binding spots for each binding region (same colour-coding). The location of each binding region is depicted in the RIN3 schematic structure. Same abbreviations as for Figure 1.17 are used.

1.7 Aims of this DPhil project

1.7.1 First aim

A major aim of the project was to elucidate the molecular details of the novel and unusual interaction between the CD2AP SH3 domains and RIN3 and ALIX epitopes using biochemical and biophysical methods. This would allow a more detailed understanding of the CD2AP SH3 binding preferences, such as what the differences between them are when binding to the same SH3 domain surfaces.

Chapter 3 focuses on a biophysical analysis of these interactions. Chapter 4 describes the preliminary experiments performed to understand the binding properties of the first two SH3 domains in tandem. A biochemical analysis of the CD2AP/RIN3 interaction is presented in Chapter 5. The biological functions of RIN3 are described in Chapter 7.

1.7.2 Second aim

Having determined the CD2AP SH3 recognition motifs, to search subsequently for putative novel CD2AP SH3 interaction partners harbouring such motifs, and validate their interactions experimentally.

Details on the experimental design of the screen and results are found in Chapter 6. The screen led to the identification of a novel CD2AP interaction partner: the GAP ARAP1 (Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1). The interaction was confirmed by immunoprecipitation at the endogenous level in immortalised human cell lines, which were derived from different tissues. The biological functions of ARAP1 are described in Chapter 7.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Antibodies and applications

Primary antibodies

Antibodies	Source, Origin	Host species	Applications (Dilution)
ARAP1	NB100-68223, Novus Biologicals	Rabbit	WB (1:2000), IP (2 µg/ml)
	NB110-68801, Novus Biologicals	Mouse	WB (1:1000)
CD2AP	sc-8763, Santa Cruz Biotech.	Goat	WB (1:1000)
	sc-25272, Santa Cruz Biotech.	Mouse	IP (2 µg/ml)
	sc-9137, Santa Cruz Biotech.	Rabbit	WB (1:2500), IP (2 µg/ml)
GST	Lab-made	Mouse	Peptide array (1:500)
MLK3	ab51068, Abcam	Rabbit	WB (1:5000)
RIN3	Dr J. Colicelli, UCLA, USA	Rabbit	WB (1:1500), IP (2 µl/ml) (Janson *et al*., 2012)
	12709-1-AP, ProteinTech*	Rabbit	WB (1:300)
	sc-102089, Santa Cruz Biotech.*	Rabbit	WB (1:100)
	AV34618, Sigma-Aldrich*	Rabbit	WB (1:1000)
	SAB4503182, Sigma-Aldrich*	Rabbit	WB (1:1000)

Abbreviations: WB = western blot, IP = immunoprecipitation. * refers to antibodies with nonsatisfactory results (my personal evaluation).

Secondary antibodies

Antibodies	Source, Origin	Host species	Applications (Dilution)
Anti-goat IgG-HRP	sc-2020, Santa Cruz Biotechnology ™	Donkey	WB (1:10000)
Anti-mouse IgG-HRP	715-036-151, Jackson ImmunoResearch™	Donkey	WB (1:10000)
Anti-rabbit IgG-HRP	711-036-152, Jackson ImmunoResearch™	Donkey	WB (1:10000)

Abbreviations: WB = western blot

2.1.2 Plasmids, expression vectors and peptides

Bacterial expression vectors

Plasmids harboring cDNAs coding for the SH3 domains of human CD2AP and for the fragment of human RIN3, which includes the RIN3 epitopes, were kindly provided by Dr Kathrin Kirsch *et al*. (Boston University, USA). In detail:

Protein fragment	Abbreviation	Amino acids	Vector
CD2AP SH3-1 domain	CD2AP SH3-1	1-60	pGEX6-P1
CD2AP SH3-(1+2) domains	CD2AP SH3-(1+2)	1-168	pGEX6-P1
CD2AP SH3-2 domain	CD2AP SH3-2	109-168	pGEX6-P1
CD2AP SH3-3 domain	CD2AP SH3-3-S	270-331	pGEX6-P1
CD2AP SH3-3 domain	CD2AP SH3-3-L	264-334	pGEX-KN
RIN3-(I+II) epitopes	RIN3-(1+2)	378-467	pGEX6-P1

Peptides for isothermal titration calorimetry and protein crystallography

C-terminally amidated ALIX- and RIN3-derived peptides were provided by Dr Nicola O' Reilly (Peptide Synthesis Laboratory, Cancer Research UK, London), and the ARAP1-, DAB1- and MLK3-derived peptides were provided by the WIMM peptide synthesis facility. A description of their synthesis can be found in section 2.2.5.4. Their sequences and the proteins they were derived from are found below:

Protein	Peptide sequence	Residues	Length (a.a.)
Wild-type peptides			
ALIX	AGGHAPTPPTPAPRTMPPTKP-am	732-752	21
ARAP1e1	RPTPRPVPMKRHIFRS-am	76-91	16
ARAP1e2	LPAAPPIPPRRSCLPP-am	112-126	16
ARAP1e3	DPVLPPLPAKRHLAEL-am	139-153	16
DAB1	STNSPPTPAPRQSSPS-am	512-526	16
MLK3e1	EEPKRPVPAERGSSSG-am	622-636	16
MLK3e2	SPLPSPQPAPRRAPWT-am	791-804	16
RIN3e1	AKKNLPTAPPRRRVSE-am	378-393	16
RIN3e2	TAKQPPVPPPRKKRIS-am	452-467	16
RIN3e2	KQPPVPPPRKK-am	454-464	11
RIN3e2-derived mutant peptides			
$P_{457} \rightarrow A$	TAKQPAVPPPRKKRIS-am	-	16
$P_{459} \rightarrow A$	TAKQPPVAPPRKKRIS-am	-	16
$R_{462} \rightarrow A$	TAKQPPVPPPAKKRIS-am	-	16
$R_{462} \rightarrow L$	TAKQPPVPPPLKKRIS-am	-	16
$R_{462} \rightarrow V$	TAKQPPVPPPVKKRIS-am	-	16
$R_{462} \rightarrow I$	TAKQPPVPPPIKKRIS-am	-	16
$I_{466} \rightarrow A$	TAKQPPVPPPRKKRAS-am	-	16
ARAP1e1-derived mutant peptides			
$P_{79} \rightarrow A$	RPTARPVPMKRHIFRS-am	-	16
$R_{87} \rightarrow A$	RPTPRPVPMKAHIFRS-am	-	16
$F_{89} \rightarrow A$	RPTPRPVPMKRHIARS-am	-	16

2.1.3 Cell lines

Cultured cell lines by E. Rouka

Cell line	Description	Culture conditions	Incubator conditions	Provided by
HEK293	Human embryonic kidney	DMEM, 10%FBS, 1% PS	37 °C, 5% CO_2	WIMM, Oxford, UK
HeLa	Human cervical adenocarcinoma	DMEM, 10%FBS, 1% PS		
HepG2	Human hepatoma	DMEM, 10%FBS, 1% PS		A.R. Townsend, WIMM, Oxford, UK
MCF-7	Human breast carcinoma	DMEM, 10%FBS, 1% PS		A. Harris, WIMM, Oxford, UK
POD undifferentiated	Human undifferentiated podocytes	RPMI1640, 10% FBS, 1% ITS, 1% PS	33 °C, 5% CO_2	M. Saleem, University of Bristol, Bristol, UK
POD differentiated	Human differentiated podocytes	RPMI1640, 10% FBS, 1% ITS, 1% PS	37 °C, 5% CO_2	
SCC-9	Oral squamous cell carcinoma	DMEM, 10%FBS, 1% PS		ATCC
UPCI-SCC-154	Oral squamous cell carcinoma	DMEM, 10%FBS, 1% PS		University of Pittsburgh, USA

PS stands for penicillin-streptomycin; ITS stands for insulin-transferrin-selenium.

Head and neck cancer cell line lysate panel

The total cell lysates (RIPA 100, mixed with SDS [sodium dodecyl sulphate]) were prepared by Miss Jessica Doondea (previous lab member).

Cell line	Provided by
BICR16	ECACC, Salisbury, UK
BICR56	
CAL27	ATCC
CAL33	DSMZ, Braunschweig Germany
HSC-3	Health Sciences Research Resources Bank, Japan
HSC-4	
OSC-19	
OSC-20	
SAS	
SCC-4	ATCC
SCC-9	
SCC-15	
SCC-25	
SIHN-005A	S.A. Eccles, University of Surrey, Guildford, UK
SIHN-006	
UT-SCC-10	University of Turku, Finland
UT-SCC-14	
UT-SCC-16A	
UT-SCC-24A	
UT-SCC-30	
UT-SCC-40	
UT-SCC-67	
UT-SCC-73	
UT-SCC-74A	
UT-SCC-76A	
UT-SCC-87	
UPCI-SCC-D56	University of Pittsburgh, USA
UPCI-SCC-016	
UPCI-SCC-21	
UPCI-SCC-40	
UPCI-SCC-75	
UPCI-SCC-103	
UPCI-SCC-122	
UPCI-SCC-154	

Other cell line lysates

The total cell lysates (RIPA 100) were prepared by previous lab members.

Cell line	Description
A431	Human epithelial carcinoma
A549	Human lung carcinoma
Daudi	Human Burkitt's lymphoma
HL60	Human acute myelocytic leukemia
Jurkat	Human acute lymphocytic leukemia
K562	Human acute myelocytic leukemia
MB-MDA-231	Human breast carcinoma
Molt4	Human acute lymphoblastic leukemia
Nalm6	Human B cell precursor acute lymphocytic leukemia
NB4	Human promyelocytic leukemia
SHS5Y5	Human neuroblastoma
SW620	Human colon adenocarcinoma
U937	Human leukemic monocyte lymphoma

2.1.4 Solutions and Buffers

The chemical suppliers were Sigma Aldrich and Fisher Scientific.

2.1.4.1 Bacterial culture and expression of recombinant proteins

GSH-bead wash buffer (GSH-WB)

50 mM Tris pH 7.5

100 mM EDTA (Ethylene diamine tetraacetic acid) pH 8

0.1 % Tween 20

Stored at 4 °C.

TPE lysis buffer

1% Triton X-100

1x PBS (phosphate-buffered saline)

100 mM EDTA pH8

Stored at 4 °C. Appropriate inhibitors (usually 0.7 µg/ml pepstatin A, 0.5 µg/ml leupeptin, 0.1 mM PMSF [Phenylmethylsulfonylfluoride] and 5 µg/ml antipain) added just before use as required for experiment.

Luria-Bertani (LB) (1% [w/v] tryptone,0.5% [w/v] yeast extract, 1% [w/v] NaCl pH 7.0) and Terrific browth (TB) (1.2% [w/v] tryptone, 2.4% [w/v] yeast extract, 0.9% [w/v] K_2HPO_4, 0.2% [w/v] KH_2PO_4, 1% [w/v] glycerol) media, 1 M Tris-HCL pH 7.5, 0.5 M EDTA pH 8, 10x TBS (tris buffered saline) (0.5 M Tris-HCl pH 7.6, 1.5 M NaCl), 10x PBS (0.03 M Na_2HPO_4, 0.01 M KH_2PO_4, 1.5 M NaCl)

Provided by oncology unit at Weatherall Institute of Molecular Medicine (WIMM).

PreScission™ and Thrombin protease cleavage buffers

20 mM Tris pH 7.5

150 mM NaCl

Cooled to 4 °C before use.

Dialysis buffer

5 mM Tris pH 7.5

For the expression of CD2AP SH3-2 and SH3-3, 2 mM β-ME was additionally freshly added. Cooled to 4 °C before use.

4x Bradford protein assay reagent

250 mg Brilliant Blue G are dissolved in 120 ml ethanol. 250 ml concentrated phosphoric acid are carefully added, mixed carefully. H_2O is added to obtain a final volume of 500 ml, stored solution in dark at RT.

1x Bradford protein assay reagent

To 12.5 ml of 4 x Bradford reagent, 37.5 ml of H_2O are added and mixed immediately. Let stand at RT for 15 min then filter through pre-wetted paper filter. Stored dark at RT until use.

BSA (Bovine serum albumin) solution for Bradford assay

Dissolved 10 mg/ml BSA in H_2O (= 100x stock), aliquoted into 0.5 ml aliquots and stored at -20 ^{o}C. To make 1x standards, thawed one aliquot and mixed with 49.5 ml of H_2O (= 0.1 mg/ml), made 1 ml aliquots, stored at -20 ^{o}C.

2.1.4.2 Cell culture, cell lysis and biochemical assays

Cell freeze medium

5% (v/v) DMSO (Dimethyl sulfoxide)

95% (v/v) FBS (Fetal bovine serum, Gibco)

Stored at 4 °C.

Hypotonic lysis buffer (HLB)

10 mM Tris 7.5

10 mM KCl

1 mM EDTA pH 8

1 mM EGTA (ethylene glycol tetraacetic acid, Sigma Aldrich^{TM})

2 mM $MgCl_2$

Stored at 4 °C, protease (Complete^{TM}, Roche) and phosphatase inhibitors inhibitors freshly added as required.

IP buffer

20 mM Tris pH 7.5

1 mM EDTA pH 8

100 mM NaCl

5% (v/v) Glycerol

0.1% Tween 20

Stored at 4°C, fresh protease (Complete™, Roche) and phosphatase inhibitors inhibitors added as required.

1% Triton X-100 buffer (TXB)

20 mM Tris pH 7.5

1 mM EDTA pH 8

100 mM NaCl

5% (v/v) Glycerol

1% (v/v) Triton X-100

Stored at 4 °C, protease (Complete™, Roche) and phosphatase inhibitors freshly added as required.

RIPA 100

20 mM Tris pH 7.5 Added H_2O to 700 ml.

1 mM EDTA pH 8

100 mM NaCl

1% Triton X-100 Stirred 5 min before adding DOC.

0.5% (v/v) Deoxycholate (DOC) Stirred 5 min before adding SDS.

0.1% (w/v) SDS Stirred 5 min, then added H_2O to 1 litre.

Stored at 4 °C and protease (Complete™, Roche) and phosphatase inhibitors

freshly added before use as required for experiment. The number in RIPA 100 indicates a NaCl concentration of 100 mM.

2.1.4.3 Protein separation, protein detection and peptide arrays

Coomassie-Blue Destain (CBD)

7% (v/v) acetic acid

20% (v/v) methanol

Stored at RT.

Coomassie-Blue staining solution

Dissolved 1 g of Serva Blue R = Brilliant Blue R in 1 litre of CBD.

SDS polyacrylamide gel - Stacking gel

For 10 ml:

Acrylamide/Bis-acrylamide solution	1.67 ml
1 M Tris HCl pH 6.8	1.27 ml
10% SDS	0.1 ml
50% Glycerol	0.9 ml
10% APS (Ammonium persulphate)	0.1 ml
TEMED (Tetramethyl ethylene diamine)	10 µl

SDS polyacrylamide gel - 12% Separating gel

For 10 ml:

Acrylamide/bis-acrylamide solution	4 ml
1 M Tris HCl pH 8.8	3.75 ml
10% SDS	0.1 ml
10% APS	0.1 ml
TEMED	10 µl

The components of the SDS gel, except for APS and TEMED, were mixed, filled up with H_2O to the final volume and mixed thoroughly. The solution was again mixed after adding APS and TEMED.

Depending on the desired percentage of separation gel, different amounts of acrylamide/bis-acrylamide solution were added.

10x SDS-PAGE gel running buffer

250 mM Tris base

1.9 M Glycine

1% SDS

Dissolved in H_2O. Stored at RT.

2 - 4x SDS protein gel sample buffer

70 mM Tris pH 6.8

5% (v/v) β-ME (β-mercaptoethanol)

40% Glycerol

3% SDS

0.05% (w/v) Bromophenol blue

Stored aliquots at -20 °C. Before use, warmed up to room temperature and mixed well to redissolve the SDS completely.

Semidry blot-buffer

48 mM Tris base

38.6 mM Glycine

0.037% SDS

Autoclaved and stored at RT.

TBST

TBS with 0.1% (v/v) Tween 20

PBST

PBS with 0.1% (v/v) Tween 20

Blocking buffer for Western blots

TBST or PBST with 5% (w/v) non-fat dry milk or 3% (w/v) BSA (bovine serum albumin) as described in results.

Blot stripping buffer

60 mM Tris pH 6.8

2% SDS

Stored at RT. β-ME added to final concentration of 0.1% (v/v) immediately before use.

2.1.5 Protein Crystallography

The crystallisation screens were the following: HCS (Hampton Core Screen), HIN (Hampton Index Screen), JCSG (Joint Centre for Structural Genomics core screen), LFS (Ligand Friendly Screen). Their conditions are found below:

HCS screen

A1	0.02	M	Calcium chloride dihydrate	0.1	M	Sodium acetate trihydrate	4.6	30	%v/v	2-methyl-2,4-pentanediol
A2								0.4	M	Potassium sodium tartrate tetrahydrate
A3				0.1	M	Tris hydrochloride	8.5	0.4	M	Ammonium dihydrogen phosphate
A4	0.2	M	tri-Sodium citrate dihydrate	0.1	M	Sodium HEPES	7.5	2	M	Ammonium sulfate
A5	0.2	M	magnesium chloride hexahydrate	0.1	M	Tris hydrochloride	8.5	30	%v/v	2-methyl-2,4-pentanediol
A6				0.1	M	sodium cacodylate	6.5	30	%w/v	polyethylene glycol 4000
A7	0.2	M	tri-sodium citrate dihydrate	0.1	M	sodium cacodylate	6.5	1.4	M	sodium acetate trihydrate
A8	0.2	M	ammonium acetate	0.1	M	tri-sodium citrate dihydrate	5.6	30	%v/v	iso-propanol
A9	0.2	M	ammonium acetate	0.1	M	Sodium acetate trihydrate	4.6	30	%w/v	polyethylene glycol 4000
A10				0.1	M	tri-sodium citrate dihydrate	5.6	30	%w/v	polyethylene glycol 4000
A11	0.2	M	magnesium chloride hexahydrate	0.1	M	Sodium HEPES	7.5	1	M	Ammonium dihydrogen phosphate
A12								30	%v/v	iso-propanol
B1	0.2	M	tri-Sodium citrate dihydrate	0.1	M	Tris hydrochloride	8.5	30	%v/v	polyethylene glycol 400
B2	0.2	M	Calcium chloride dihydrate	0.1	M	Sodium HEPES	7.5	28	%v/v	polyethylene glycol 400
B3	0.2	M	ammonium sulfate	0.1	M	sodium cacodylate	6.5	30	%w/v	polyethylene glycol 8000
B4				0.1	M	Sodium HEPES	7.5	1.5	M	lithium sulfate monohydrate
B5	0.2	M	lithium sulfate monohydrate	0.1	M	Tris hydrochloride	8.5	30	%w/v	polyethylene glycol 4000
B6	0.2	M	magnesium acetate tetrahydrate	0.1	M	sodium cacodylate	6.5	20	%w/v	polyethylene glycol 8000
B7	0.2	M	ammonium acetate	0.1	M	Tris hydrochloride	8.5	30	%v/v	iso-propanol
B8	0.2	M	ammonium sulfate	0.1	M	Sodium acetate trihydrate	4.6	25	%w/v	polyethylene glycol 4000
B9	0.2	M	magnesium acetate tetrahydrate	0.1	M	sodium cacodylate	6.5	30	%v/v	2-methyl-2,4-pentanediol
B10	0.2	M	sodium acetate trihydrate	0.1	M	Tris hydrochloride	8.5	30	%w/v	polyethylene glycol 4000
B11	0.2	M	magnesium chloride hexahydrate	0.1	M	Sodium HEPES	7.5	30	%v/v	polyethylene glycol 400
B12	0.2	M	Calcium chloride dihydrate	0.1	M	Sodium acetate trihydrate	4.6	20	%v/v	iso-propanol
C1	0.2	M	ammonium acetate	0.1	M	imidazole	6.5	1	M	sodium acetate trihydrate, 0.8 M
C2	0.2	M	tri-Sodium citrate dihydrate	0.1	M	tri-sodium citrate dihydrate	5.6	potassium dihydrogen phosphate		
C3	0.2	M	sodium acetate trihydrate	0.1	M	Sodium HEPES	7.5	30	%v/v	2-methyl-2,4-pentanediol
C4				0.1	M	sodium cacodylate	6.5	20	%v/v	iso-propanol
C5	0.2	M	ammonium sulfate	0.1	M	Sodium HEPES	7.5	30	%w/v	polyethylene glycol 8000
C6	0.2	M	ammonium sulfate					0.8	M	Potassium sodium tartrate tetrahydrate
C7								30	%w/v	polyethylene glycol 8000
C8				0.1	M	sodium acetate trihydrate	4.6	30	%w/v	polyethylene glycol 4000
C9				0.1	M	Sodium HEPES	7.5	2	M	Ammonium sulfate
C10				0.1	M	Tris hydrochloride	8.5	4	M	sodium formate
C11								2	M	sodium formate
C12								0.8	M	sodium dihydrogen phosphate
								8	%w/v	polyethylene glycol 8000
D1	0.05	M	potassium dihydrogen phosphate	0.1	M	Sodium acetate trihydrate	4.6	8	%w/v	polyethylene glycol 4000, 2 M ammonium sulfate
D2				0.1	M	Sodium HEPES	7.5	1.4	M	tri-sodium citrate dihydrate, 20 %w/v polyethylene glycol 4000
D3	0.2	M	zinc acetate dihydrate	0.1	M	Sodium HEPES	7.5	2	%v/v	polyethylene glycol 400, 20 %w/v polyethylene glycol 4000
D4	0.2	M	calcium acetate hydrate	0.1	M	tri-sodium citrate dihydrate	5.6	20	%v/v	iso-propanol
D5				0.1	M	Sodium HEPES	7.5	10	%v/v	iso-propanol
D6								20	%w/v	polyethylene glycol 8000
D7								30	%w/v	polyethylene glycol 1500
D8				0.1	M	sodium cacodylate	6.5	0.2	M	magnesium formate
D9				0.1	M	sodium cacodylate	6.5	18	%w/v	polyethylene glycol 8000
D10				0.1	M	Sodium acetate trihydrate	4.6	18	%w/v	polyethylene glycol 8000
D11				0.1	M	Tris hydrochloride	8.5	2	M	Ammonium sulfate
D12								1	M	Ammonium dihydrogen phosphate
E1	2.0	M	sodium chloride	7.0				10	%	PEG 6000, 10 %w/v polyethylene glycol 8000
E2	0.5	M	sodium chloride					0.01	M	hexadecyltrimethylammonium bromide
	0.01	M	magnesium chloride hexahydrate	0.1	M	sodium acetate trihydrate	4.6			
E3				0.1	M	sodium acetate trihydrate	4.6	25	%v/v	ethylene glycol
E4	2.0	M	ammonium sulfate	0.1	M	sodium acetate trihydrate	4.6	35	%v/v	dioxane
E5				0.1	M	sodium acetate trihydrate	4.6	5	%v/v	iso-propanol
E6	1.5	M	sodium chloride					1	M	imidazole
E7								10	%w/v	polyethylene glycol 1000
E8	0.2	M	sodium chloride					10	%v/v	ethanol
E9	0.01	M	cobaltous chloride hexahydrate					2	M	sodium chloride
E10	0.1	M	cadmium chloride dihydrate					30	%v/v	MPD
E11								1	M	1,6 hexanediol
E12								30	%v/v	polyethylene glycol 400
F1	0.2	M	ammonium sulfate	0.1	M	sodium acetate trihydrate	4.6	30	%w/v	polyethylene glycol monomethyl ether 200
F2	0.2	M	potassium sodium tartrate tetrahydra	0.1	M	tri-sodium citrate dihydrate	5.6	2	M	ammonium sulfate
F3	0.5	M	ammonium sulfate	0.1	M	tri-sodium citrate dihydrate	5.6	1	M	lithium sulfate monohydrate
F4	0.5	M	sodium chloride	0.1	M	tri-sodium citrate dihydrate	5.6	2	%v/v	ethylene imine polymer
F5				0.1	M	tri-sodium citrate dihydrate	5.6	35	%v/v	tert-butanol
F6	0.01	M	ferric chloride hexahydrate	0.1	M	tri-sodium citrate dihydrate	5.6	10	%v/v	jeffamine M-600
F7				0.1	M	tri-sodium citrate dihydrate	5.6	2.5	M	1,6 hexanediol
F8	0.1	M	sodium dihydrogen phosphate	0.1	M	MES	6.5	1.6	M	magnesium sulfate heptahydrate
F9	0.1	M	potassium dihydrogen phosphate	0.1	M	MES	6.5	2	M	sodium chloride
F10	1.6	M	ammonium sulfate	0.1	M	MES	6.5	12	%w/v	polyethylene glycol 20,000
F11	0.05	M	cesium chloride	0.1	M	MES	6.5	10	%v/v	dioxane
F12				0.1	M	MES	6.5	30	%v/v	jeffamine M-600
G1	0.01	M	cobaltous chloride hexahydrate	0.1	M	MES	6.5	1.8	M	ammonium sulfate
G2	0.2	M	ammonium sulfate	0.1	M	MES	6.5	30	%w/v	polyethylene glycol monomethyl ether 500
G3	0.01	M	zinc sulfate heptahydrate	0.1	M	MES	6.5	25	%v/v	polyethylene glycol monomethyl ether 550
G4				6.5				1.6	M	tri-sodium citrate dihydrate, 5 %v/v MPD
G5	0.5	M	ammonium sulfate	0.1	M	HEPES	7.5	30	%v/v	MPD
G6				0.1	M	HEPES	7.5	10	%w/v	polyethylene glycol 6000
G7	0.1	M	sodium chloride	0.1	M	HEPES	7.5	20	%v/v	jeffamine M-600
G8				0.1	M	HEPES	7.5	1.6	M	ammonium sulfate
G9	0.05	M	cadmium sulfate hydrate	0.1	M	HEPES	7.5	2	M	ammonium formate
G10				0.1	M	HEPES	7.5	1	M	sodium acetate trihydrate
G11				0.1	M	HEPES	7.5	70	%v/v	MPD
G12				0.1	M	HEPES	7.5	4.3	M	sodium chloride
H1	0.2	M	magnesium chloride hexahydrate	0.1	M	HEPES	7.5	10	%w/v	polyethylene glycol 8000, 8 %v/v ethylene glycol
H2				0.1	M	HEPES	7.5	20	%w/v	polyethylene glycol 10,000
H3	0.01	M	nickel (II) chloride hexahydrate	0.1	M	tris	8.5	3.4	M	1,6 hexanediol
H4	1.5	M	ammonium sulfate	0.1	M	tris	8.5	25	%v/v	tert-butanol
H5	0.2	M	ammonium dihydrogen phosphate	0.1	M	tris	8.5	1	M	lithium sulfate monohydrate
H6				0.1	M	tris	8.5	12	%v/v	glycerol anhydrous
H7	0.01	M	nickel (II) chloride hexahydrate	0.1	M	tris	8.5	50	%v/v	MPD
H8	0.1	M	sodium chloride	0.1	M	tris	8.5	20	%v/v	ethanol, 10 %w/v polyethylene glycol 20,000
H9				0.1	M	tris	8.5	20	%w/v	polyethylene glycol monomethyl ether 200
H10				0.1	M	bicine	9.0	20	%v/v	polyethylene glycol monomethyl ether 550
H11				0.1	M	bicine	9.0	2	M	magnesium chloride hexahydrate
H12				0.1	M	bicine	9.0	2	%v/v	dioxane

HIN screen

A1						0.1	M	citric acid	3.5	2.0	M	ammonium sulfate
A2						0.1	M	sodium acetate trihydrate	4.5	2.0	M	ammonium sulfate
A3						0.1	M	bis-tris	5.5	2.0	M	ammonium sulfate
A4						0.1	M	bis-tris	6.5	2.0	M	ammonium sulfate
A5						0.1	M	HEPES	7.5	2.0	M	ammonium sulfate
A6						0.1	M	tris	8.5	2.0	M	ammonium sulfate
A7						0.1	M	citric acid	3.5	3.0	M	sodium chloride
A8						0.1	M	sodium acetate trihydrate	4.5	3.0	M	sodium chloride
A9						0.1	M	bis-tris	5.5	3.0	M	sodium chloride
A10						0.1	M	bis-tris	6.5	3.0	M	sodium chloride
A11						0.1	M	HEPES	7.5	3.0	M	sodium chloride
A12						0.1	M	tris	8.5	3.0	M	sodium chloride
B1						0.1	M	bis-tris	5.5	0.3	M	magnesium formate
										0.14	M	di-potassium hydrogen phosphate
B2						0.1	M	bis-tris	6.5	0.5	M	magnesium formate
										0.91	M	di-potassium hydrogen phosphate
B3						0.1	M	HEPES	7.5	0.5	M	magnesium formate
										1.344	M	di-potassium hydrogen phosphate
B4						0.1	M	tris	8.5	0.3	M	magnesium formate
B5										1.3	M	sodium dihydrogen phosphate
B6										0.5	M	sodium dihydrogen phosphate
B7						0.1	M	HEPES	7.5	0.1	M	sodium dihydrogen phosphate
B8										1.4	M	tri-sodium citrate dihydrate
B9										1.8	M	tri-ammonium citrate pH 7.0
B10										0.8	M	succinic acid pH 7.0
B11										2.1	M	DL-malic acid pH 7.0
B12										2.8	M	sodium acetate trihydrate pH 7.0
C1	0.1	M	sodium chloride			0.1	M	bis-tris	6.5	3.5	M	sodium formate pH 7.0
C2	0.8	M	potassium sodium tartrate			0.1	M	tris	8.5	1.1	M	di-ammonium tartrate pH 7.0
C3	1.0	M	ammonium sulfate			0.1	M	bis-tris	5.5	2.4	M	sodium malonate pH 7.0
C4	1.1	M	sodium malonate pH 7.0			0.1	M	HEPES	7.0	35.0	%v/v	tacsimate pH 7.0
C5	1.0	M	succinic acid pH 7.0			0.1	M	HEPES	7.0	60.0	%v/v	tacsimate pH 7.0
C6	1.0	M	ammonium sulfate			0.1	M	HEPES	7.0	1.5	M	ammonium sulfate
C7	15.00	%w/v	tacsimate pH 7.0			0.1	M	HEPES	7.0	0.5	%w/v	polyethylene glycol MME 5000
C8										1.0	%w/v	polyethylene glycol 3350
C9										0.5	%v/v	jeffamine ED-2001 reagent pH 7.0
C10										1.0	%w/v	polyethylene glycol MME 2000
C11										0.5	%w/v	polyethylene glycol 8000
C12										2.0	%w/v	polyethylene glycol 3350
D1	0.2	M	calcium chloride			0.1	M	HEPES	7.0	25.0	%w/v	polyethylene glycol 1500
D2						0.1	M	HEPES	7.0	30.0	%v/v	jeffamine M-600 reagent pH 7.0
D3						0.1	M	citric acid	3.5	30.0	%v/v	jeffamine ED-2001 reagent pH 7.0
D4						0.1	M	sodium acetate trihydrate	4.5	25.0	%w/v	polyethylene glycol 3350
D5						0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
D6						0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
D7						0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
D8						0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
D9						0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
D10						0.1	M	bis-tris	6.5	20.0	%w/v	polyethylene glycol MME 5000
D11						0.1	M	bis-tris	5.5	28.0	%w/v	polyethylene glycol MME 2000
D12										45.0	%v/v	2-methyl-2,4-pentanediol
E1	0.2	M	calcium chloride			0.1	M	bis-tris	6.5	45.0	%v/v	2-methyl-2,4-pentanediol
E2	0.2	M	ammonium acetate			0.1	M	bis-tris	5.5	45.0	%v/v	2-methyl-2,4-pentanediol
E3	0.2	M	ammonium acetate			0.1	M	bis-tris	6.5	45.0	%v/v	2-methyl-2,4-pentanediol
E4	0.2	M	ammonium acetate			0.1	M	HEPES	7.5	45.0	%v/v	2-methyl-2,4-pentanediol
E5	0.2	M	ammonium acetate			0.1	M	tris	8.5	45.0	%v/v	2-methyl-2,4-pentanediol
E6	0.05	M	calcium chloride			0.1	M	bis-tris	6.5	30.0	%v/v	polyethylene glycol MME 550
E7	0.05	M	magnesium chloride			0.1	M	HEPES	7.5	30.0	%v/v	polyethylene glycol MME 550
E8	0.2	M	potassium chloride			0.05	M	HEPES	7.5	35.0	%v/v	pentaerythritol propoxylate (5/4 PO/OH)
E9	0.05	M	ammonium sulfate			0.05	M	bis-tris	6.5	30.0	%v/v	pentaerythritol ethoxylate (15/4 EO/OH)
E10						0.1	M	bis-tris	6.5	45.0	%v/v	polypropylene glycol P 400
E11	0.02	M	magnesium chloride			0.1	M	HEPES	7.5	22.0	%w/v	polyacrylic acid 5100 sodium salt
E12	0.10	M	cobalt chloride			0.1	M	tris	8.5	20.0	%w/v	polyvinylpyrrolidone K15
F1	0.20	M	proline			0.1	M	HEPES	7.5	10.0	%w/v	polyethylene glycol 3350 0.005 M nickel (II) chloride
										0.005	M	nickel (II) chloride
F2	0.20	M	trimethylamine n-oxide			0.1	M	tris	8.5	20.0	%w/v	polyethylene glycol MME 2000
F3	5.00	%w/v	tacsimate pH 7.0			0.1	M	HEPES	7.0	10.0	%w/v	polyethylene glycol MME 5000
F4						0.1	M	HEPES	7.5	12.0	%w/v	polyethylene glycol 3350 0.005 M cobalt chloride
										0.005	M	magnesium chloride
F5	0.20	M	ammonium sulfate			0.1	M	bis-tris	5.5	17.0	%w/v	polyethylene glycol 10,000
F6	0.20	M	ammonium sulfate			0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
F7	0.20	M	ammonium sulfate			0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
F8	0.20	M	ammonium sulfate			0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
F9	0.20	M	sodium chloride			0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
F10	0.20	M	sodium chloride			0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
F11	0.20	M	sodium chloride			0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
F12						0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
G1	0.20	M	sodium chloride			0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
G2	0.20	M	lithium sulfate			0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
G3	0.20	M	lithium sulfate			0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
G4	0.20	M	lithium sulfate			0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
G5	0.20	M	lithium sulfate			0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
G6	0.20	M	ammonium acetate			0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
G7	0.20	M	ammonium acetate			0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
G8	0.20	M	ammonium acetate			0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
G9	0.20	M	ammonium acetate			0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
G10	0.20	M	magnesium chloride			0.1	M	bis-tris	5.5	25.0	%w/v	polyethylene glycol 3350
G11	0.20	M	magnesium chloride			0.1	M	bis-tris	6.5	25.0	%w/v	polyethylene glycol 3350
G12	0.20	M	magnesium chloride			0.1	M	HEPES	7.5	25.0	%w/v	polyethylene glycol 3350
H1	0.20	M	magnesium chloride			0.1	M	tris	8.5	25.0	%w/v	polyethylene glycol 3350
H2	0.20	M	potassium sodium tartrate							20.0	%w/v	polyethylene glycol 3350
H3	0.20	M	sodium malonate pH 7.0							20.0	%w/v	polyethylene glycol 3350
H4	0.20	M	tri-ammonium citrate pH 7.0							20.0	%w/v	polyethylene glycol 3350
H5	0.10	M	succinic acid pH 7.0							15.0	%w/v	polyethylene glycol 3350
H6	0.20	M	sodium formate							20.0	%w/v	polyethylene glycol 3350
H7	0.15	M	DL-malic acid pH 7.0							20.0	%w/v	polyethylene glycol 3350
H8	0.10	M	magnesium formate							15.0	%w/v	polyethylene glycol 3350
H9	0.05	M	zinc acetate							20.0	%w/v	polyethylene glycol 3350
H10	0.20	M	tri-sodium citrate							20.0	%w/v	polyethylene glycol 3350
H11	0.10	M	potassium thiocyanate							30.0	%w/v	polyethylene glycol MME 2000
H12	0.15	M	potassium bromide							30.0	%w/v	polyethylene glycol MME 2000

LFS screen

A1			0.1 M	SPG	6	30.0 % PEG 1K		
A2			0.1 M	SPG	7			
A3			0.1 M	SPG	8			
A4			0.1 M	SPG	6	60.0 % MPD		
A5			0.1 M	SPG	7			
A6			0.1 M	SPG	8			
A7	0.20 M	NaCl				20.0 % PEG 6K	10.0 %	EtGly
A8	0.20 M	NH_4Cl						
A9	0.20 M	LiCl						
A10	0.10 M	$MgCl_2$						
A11	0.05 M	$CaCl_2$						
A12	0.01 M	$ZnCl_2$						
B1			0.1 M	MIB	6	30.0 % PEG 1K		
B2					7			
B3					8			
B4					6	60.0 % MPD		
B5					7			
B6					8			
B7	0.20 M	NaCl	0.1 M	MES	6	20.0 % PEG 6K	10.0 %	EtGly
B8	0.20 M	NH_4Cl			6			
B9	0.20 M	LiCl			6			
B10	0.10 M	$MgCl_2$			6			
B11	0.05 M	$CaCl_2$			6			
B12	0.01 M	$ZnCl_2$			6			
C1			0.1 M	PCB	6	30.0 % PEG 1K		
C2					7			
C3					8			
C4					6	60.0 % MPD		
C5					7			
C6					8			
C7	0.20 M	NaCl	0.1 M	HEPES	7	20.0 % PEG 6K	10.0 %	EtGly
C8	0.20 M	NH_4Cl						
C9	0.20 M	LiCl						
C10	0.10 M	$MgCl_2$						
C11	0.05 M	$CaCl_2$						
C12	0.01 M	$ZnCl_2$						
D1			0.1 M	MMT	6	30.0 % PEG 1K		
D2					7			
D3					8			
D4					6	60.0 % MPD		
D5					7			
D6					8			
D7	0.20 M	NaCl	0.1 M	TRIS	8	20.0 % PEG 6K	10.0 %	EtGly
D8	0.20 M	NH_4Cl						
D9	0.20 M	LiCl						
D10	0.10 M	$MgCl_2$						
D11	0.05 M	$CaCl_2$						
D12	0.01 M	$ZnCl_2$						
E1	0.20 M	NaF				20.0 % PEG 3350	10.0 %	EtGly
E2		NaBr						
E3		NaI						
E4		KSCN						
E5		$NaNO_3$						
E6		Na(form)						
E7		Na(ac)						
E8		Na_2SO_4						
E9		Na/K(tart)						
E10		Na/KPO_4						
E11		K_3(cit)						
E12		Na(malonate)						
F1		NaF	0.1 M	BTProp	6.5			
F2		NaBr						
F3		NaI						
F4		KSCN						
F5		$NaNO_3$						
F6		Na(form)						
F7		Na(ac)						
F8		Na_2SO_4						
F9		Na/K(tart)						
F10		Na/KPO_4						
F11		K_3(cit)						
F12		Na(malonate)						
G1		NaF	0.1 M	BTProp	7.5			
G2		NaBr						
G3		NaI						
G4		KSCN						
G5		$NaNO_3$						
G6		Na(form)						
G7		Na(ac)						
G8		Na_2SO_4						
G9		Na/K(tart)						
G10		Na/KPO_4						
G11		K_3(cit)						
G12		Na(malonate)						
H1		NaF	0.1 M	BTProp	8.5			
H2		NaBr						
H3		NaI						
H4		KSCN						
H5		$NaNO_3$						
H6		Na(form)						
H7		Na(ac)						
H8		Na_2SO_4						
H9		Na/K(tart)						
H10		Na/KPO_4						
H11		K_3(cit)						
H12		Na(malonate)						

JCSG screen

Well	number	[major]	unit	major	minor(Vol)	minor[stock]	[minor]	unit	minor	minor(pH)	other(Vol)	other(stock)	[other]	unit	other	[buff]	unit	buff	buff(pH)
A1	1	50	%	PEG 300	0.100	2.0	0.20	M	Li_2SO_4							0.1	M	acetate	4.5
A2	2	20	%	PEG 3350												0.1	M	citrate	5.5
A3	3	20	%	PEG 3350	0.080	2.5	0.20	M	$(NH_4)_2H(cit)$	5.0									
A4	4	30	%	MPD	0.020	1.0	0.02	M	$CaCl_2$							0.1	M	acetate	4.6
A5	5	20	%	PEG 3350	0.200	1.0	0.20	M	$Mg(form)_2$	5.9									
A6	6	20	%	PEG 1K	0.100	2.0	0.20	M	Li_2SO_4							0.1	M	cit/phos	4.2
A7	7	20	%	PEG 10K												0.1	M	BICINE	9.3
A8	8	20	%	PEG 3350	0.200	1.0	0.20	M	Na(form)										
A9	9	20	%	PEG 3350	0.040	5.0	0.20	M	$(NH_4)Cl$	6.3									
A10	10	1.00	M	Li_2SO_4	0.250	2.0	0.50	M	TMAO										
A11	11	50	%	MPD	0.029	3.5	0.10	M	$(NH_4)_2HPO_4$							0.1	M	TRIS	8.5
A12	12	20	%	PEG 3350	0.200	1.0	0.20	M	$NaNO_3$										
B1	13				0.229	3.5	0.80	M	$(NH_4)_2SO_4$							0.1	M	citrate	4.0
B2	14	20	%	PEG 3350	0.025	8.0	0.20	M	KSCN										
B3	15	20	%	PEG 6K												0.1	M	BICINE	9.0
B4	16	10	%	PEG 10K	0.080	100	8	%	EtGly							0.1	M	HEPES	7.5
B5	17	40	%	MPD	0.100	50	5	%	PEG 10K							0.1	M	cacodylate	6.5
B6	18	40	%	isopropanol	0.100	50	5	%	PEG 1K							0.1	M	citrate/PO_4	4.2
B7	19	8	%	PEG 3350												0.1	M	acetate	4.6
B8	20	10	%	PEG 10K	0.100	2.0	0.20	M	$MgCl_2$							0.1	M	TRIS	7.0
B9	21	20	%	PEG 6K												0.1	M	citrate	5.0
B10	22	50	%	PEG 300	0.100	2.0	0.20	M	$MgCl_2$							0.1	M	cacodylate	6.5
B11	23	2	M	$K_3(cit)$															
B12	24	20	%	PEG 3350	0.080	2.5	0.20	M	$K_3(cit)$	8.3									
C1	25	20	%	PEG 10K	0.040	5.0	0.20	M	NaCl							0.1	M	citrate/PO_4	4.2
C2	26	20	%	PEG 6K	0.100	10.0	1.00	M	LiCl							0.1	M	citrate	4.0
C3	27	15	%	mPEG 2K	0.200	50	10	%	PVPdne							0.1	M	imidazole	6.5
C4	28	10	%	PEG 6K												0.1	M	HEPES	7.0
C5	29	1.60	M	Na/KPO_4												0.1	M	HEPES	7.5
C6	30	40	%	PEG 300												0.1	M	citrate/PO_4	4.2
C7	31	10	%	PEG 3350	0.200	1.0	0.20	M	$Zn(ac)_2$							0.1	M	acetate	4.5
C8	32	20	%	isopropanol												0.1	M	TRIS	8.5
C9	33	25	%	1,2 propandi	0.100	100	10	%	glycerol							0.1	M	$Na/K-PO_4$	6.2
C10	34	10	%	PEG 10K	0.050	40	2	%	dioxane							0.1	M	BICINE	9.0
C11	35	2.00	M	$(NH_4)_2SO_4$												0.1	M	acetate	4.6
C12	36	10	%	PEG 10K	0.200	50	10	%	PEG 1K										
D1	37	24	%	PEG 1K	0.200	100	20	%	glycerol										
D2	38	30	%	PEG 300	0.100	2.0	0.20	M	$MgCl_2$							0.1	M	HEPES	7.5
D3	39	50	%	PEG 300	0.040	5.0	0.20	M	NaCl							0.1	M	$Na/K-PO_4$	6.2
D4	40	30	%	PEG 10K	0.100	2.0	0.20	M	Li_2SO_4							0.1	M	acetate	4.5
D5	41	70	%	MPD												0.1	M	HEPES	7.5
D6	42	20	%	PEG 10K	0.100	2.0	0.20	M	$MgCl_2$							0.1	M	TRIS	8.5
D7	43	40	%	PEG 300	0.100	2.0	0.20	M	Li_2SO_4							0.1	M	TRIS	8.5
D8	44	40	%	MPD												0.1	M	TRIS	8.0
D9	45	25.5	%	PEG 3350	0.049	3.5	0.17	M	$(NH_4)_2SO_4$		0.150	100	15	%	glyc				
D10	46	40	%	PEG 300	0.200	1.0	0.20	M	$Ca(ac)_2$							0.1	M	cacodylate	6.5
D11	47	14	%	isopropanol	0.140	1.0	0.14	M	$CaCl_2$		0.300	100	30	%	glyc	0.7	M	acetate	4.6
D12	48	16	%	PEG 10K	0.020	2.0	0.04	M	Na/KPO_4		0.200	100	20	%	glyc				
E1	49				0.400	2.5	1.00	M	$K_3(cit)$							0.1	M	cacodylate	6.5
E2	50	2.00	M	$(NH_4)_2SO_4$	0.040	5.0	0.20	M	NaCl							0.1	M	cacodylate	6.5
E3	51	10	%	isopropanol	0.040	5.0	0.20	M	NaCl							0.1	M	HEPES	7.5
E4	52	1.26	M	$(NH_4)_2SO_4$	0.100	2.0	0.20	M	Li_2SO_4							0.1	M	TRIS	8.5
E5	53	40	%	MPD												0.1	M	BICINE	9.3
E6	54	20	%	PEG 3350	0.200	1.0	0.20	M	$Zn(ac)_2$							0.1	M	imidazole	8.0
E7	55	10	%	isopropanol	0.200	1.0	0.20	M	$Zn(ac)_2$							0.1	M	cacodylate	6.5
E8	56				0.500	2.0	1.00	M	Na/KPO_4							0.1	M	acetate	4.5
E9	57				0.640	2.5	1.60	M	$MgSO_4$							0.1	M	MES	6.5
E10	58	10	%	PEG 6K												0.1	M	BICINE	9.0

2.1.6 Isothermal titration calorimetry

ITC buffer

25 mM HEPES-KOH pH 7.5

100 mM potassium acetate

5 mM magnesium acetate

Solution was sterile filtered through a 0.2 µm filter.

2.2 Methods

2.2.1 Bacterial culture and expression of recombinant proteins

The pGEX bacterial expression vectors are used for the expression and purification of GST-fusion proteins in *E. coli*. The protein of interest is expressed as a C-terminal fusion with GST, which is a 26 kDa glutathione S-transferase encoded by *Schistosoma japonicum* (Smith *et al*. 1988). The fusion proteins are expressed upon IPTG (Isopropyl β-D-1-thiogalactopyranoside) induction and can be purified by affinity chromatography on immobilised reduced glutathione (GSH) or by batch method with glutathione agarose (settled resin) under non-denaturing conditions. The bound fusion proteins are released from the affinity matrix with an excess of free reduced glutathione. Protease cleavage sites can be inserted into the vectors, which enables GST-tag removal. The pGEX system is also widely used, because fusion proteins often remain more soluble than the target proteins alone.

The vectors, which were used in this thesis, are pGEX6-P1 and pGEX-KN. A Prescission^{TM} protease cleavage site is present in the pGEX6-P1 vector. After GST-tag removal, five residues derived from GST-tag remain at the N-terminus of the protein of interest (Gly-Pro-Leu-Gly-Ser). A thrombin cleavage site is present in the pGEX-KN vector. If the first two residues of the protein of interest are nonacidic, no GST tag-derived residues remain bound at the N-terminus of protein of interest after cleavage with thrombin protease (Hakes *et al*. 1992).

2.2.1.1 Bacterial transformation

A Luria-Bertani (LB) agar plate with ampicillin was warmed up from the cold to room temperature. Filter tips were used for all bacterial DNA experiments. A 20 µl aliquot of competent *E.coli* strain Rosetta (DE3) cells was thawed on ice. The plasmid DNA was diluted with sterile H_2O or TE buffer to give final concentration of 5 to 10 ng/µl. 1 µl of the diluted DNA was added to the thawed competent cells by stirring gently. The mixture was left on ice for 30 min and then heat-shocked for 90 sec at 42 °C. Bacteria were returned onto the ice for 2 min. 980 µl LB medium at room temperature was added and the bacteria were incubated for 45 min at 37°C with shaking. The cells were sedimented gently by centrifugation at 1000 x *g* for 3 minutes. 850 µl of the supernatant was removed and the cell pellet was resuspended gently in the remaining liquid medium (~ 150 µl). The cells were pipetted onto the LB agar plate and spread evenly over the surface to allow it to soak in. The agar plates were incubated overnight at 37 °C.

2.2.1.2 Bacterial stocks

A few colonies of transformed bacteria were picked from an LB agar plate and transferred into 4 ml LB medium with 75 µg/ml carbenicillin. The culture was grown for several hours at 37 °C in a bacterial shaker until the culture reached an optical density at 600 nm (OD_{600}) of 0.8. The bacterial suspension was then aliquoted and thoroughly mixed with glycerol at a 40% final glycerol concentration. Aliquots were frozen in liquid nitrogen and stored for further use at -80 °C.

2.2.1.3 GST-fusion protein expression

2 ml of LB medium with 75 µg/ml carbenicillin were inoculated by transfer of a small amount of material from a glycerol stock. The culture was grown for several hours at 37°C in a bacterial shaker until the culture reached an OD_{600} of 1-1.5. 200 ml of LB medium with 75 µg/ml carbenicillin were inoculated with 1 ml of bacterial suspension and grown overnight at 37 °C in a bacterial shaker. For each 2 litre Erlenmeyer flask, 20 ml of overnight culture were added to 600 ml TB with 75 µg/ml carbenicillin. The bacterial suspension was grown for several hours in a bacterial shaker (37 °C) until the culture reached an OD_{600} of 1-1.5. Next, expression was induced with 0.1 mM IPTG overnight at 18 °C, and bacteria were then sedimented by centrifugation at 2000 x *g* for 15 minutes at 4 °C. Afterwards, 30 min ice-chilled bacterial pellets were lysed in cold TPE lysis buffer with inhibitors and sonicated on ice for 4 x 30 sec with 1 min intervals. The lysate was clarified by centrifugation at 48000 *g* for 1 h at 4 °C. The supernatant was frozen in liquid nitrogen and stored at -80 °C.

2.2.1.4 GST-fusion protein purification

A glutathione Sepharose (GSH) beads titration was first performed to estimate the optimal ratio of GSH beads (Thermo Scientific Pierce™ or Macherey-Nagel™) to the supernatant. For this, various amounts of supernatant (30 µl, 100 µl, 300 µl, 600 µl, 1 ml) were incubated with 15 µl GSH beads on a nutator overnight at 4 °C. Then, beads were washed with 3 x 1 ml GSH-WB. Each time the GSH beads were sedimented gently by centrifugation at 500 x *g* for 5 minutes. The bound GST-fusion proteins were analyzed by SDS-PAGE and Coomassie Blue staining. Based on the analysis, the optimal ratio GSH beads to supernatant was used for scale-up

purification by co-incubating on a nutator overnight at 4 °C. Then, beads were washed with 3 x 50 ml GSH-WB. Each time the GSH beads were sedimented gently by centrifugation at 500 x *g* for 5 minutes. Bound GST-fusion protein was eluted with elution buffer (100 mM GSH, pH-adjusted to approximately pH 8 with 1 M Tris-HCl pH 8.8) by gentle mixing for 3 hours and the solution was recovered by gravity flow through an Eco column (Biorad). Additional fractions were collected by eluting with 3 x 1 bead volumes of fresh elution buffer. The eluate was dialyzed twice against 5 litres of dialysis buffer. The integrity of the dialyzed protein was analyzed by SDS-PAGE and Coomassie Blue staining, and the protein was assayed by the Bradford method. Purified GST-fusion protein was snap-frozen in aliquots and stored at -80 °C until further use.

2.2.1.5 GST-CD2AP-SH3-1, -2, -3-S purification for ITC and protein crystallography

For crystallization experiments and ITC measurements, the GST-tags of the expressed GST-fusion proteins were removed by Prescission™ protease (GE Healthcare). Proteins were purified as above (section 2.2.1.4). The eluate was dialyzed twice against 5 litres of Prescission™ protease buffer. The GST-tag was removed by incubating the eluate with Prescission™ protease on a nutator overnight at 4 °C (1:250 ratio by weight of protease to GST-fusion protein). The GST-tag, uncleaved and cleaved protein were separated by size exclusion FPLC on a HiLoad 16/60 Superdex S75 column (GE Healthcare™) equilibrated with gel filtration buffer (20 mM Tris pH 7.5, 150 mM NaCl for SH3-1 and additionally 2 mM β-ME for SH3-2 and SH3-3-S). Fractions containing purified protein were pooled and dialyzed against the dialysis

buffer. The integrity of the dialyzed protein was analyzed by SDS-PAGE and Coomassie Blue staining and the protein then concentrated. Final concentrations were determined by UV absorption at 280 nm.

2.2.1.6 GST-CD2AP-SH3-3-L purification for ITC and protein crystallography

For crystallization experiments and ITC measurements, the GST-tag of the GSH-Sepharose purified GST-fusion protein was removed by Thrombin protease as described below. After cleavage, only one residue (Gly) from the vector was left on the N-terminus of SH3-3-L. Before the proteolytic cleavage, the eluate was dialyzed twice against 5 litres of the Thrombin protease buffer. The GST-tag was removed by incubating the eluate with Thrombin protease on a nutator for 60 h at 4 °C (1:100 ratio by weight of protease to GST-fusion protein). The GST-tag, uncleaved and cleaved protein were separated by size exclusion FPLC on a HiLoad 16/60 Superdex S75 column (GE Healthcare) equilibrated with gel filtration buffer (20 mM Tris pH 7.5, 150 mM NaCl, 2 mM β-ME). Fractions containing purified protein were pooled and dialyzed against 5 mM Tris pH 7.5 and 2 mM β-ME. To maximise yield, fractions containing uncleaved protein were pooled and the steps starting from incubation with the Thrombin protease were repeated. The integrity of the dialyzed protein was analyzed by SDS-PAGE and Coomassie Blue staining and the protein was then concentrated. The final protein concentrations were determined by UV absorption at 280 nm.

2.2.1.7 GST-CD2AP-SH3-(1+2) purification for ITC and protein crystallography

For crystallization experiments and ITC measurements, the GST-tag of the expressed GST-fusion protein was removed by Prescission^{TM} protease. Protein was purified as above (section 2.2.1.4). The eluate was dialyzed twice against 5 litres of the Prescission protease buffer. The GST-tag was removed by incubating the eluate with Prescission protease on a nutator overnight at 4 °C (1:150 ratio by weight of protease to GST-fusion protein). The eluate was precipitated with ammonium sulphate to obtain a fraction from 40-50% saturation at 4 °C (see section 2.2.4.5 for more details on this method). The mixture was centrifuged at 10.000 g for 15 minutes at 4 °C . The supernatant was discarded and the pellet was dissolved in gel filtration buffer (20 mM Tris pH 7.5, 150 mM NaCl, 2 mM β-ME) and applied to a HiLoad 16/60 Superdex S75 column equilibrated with the same buffer. Fractions containing purified protein were pooled and dialyzed against the dialysis buffer. Fractions containing protein and GST were pooled and the steps starting from the ammonium sulphate precipitation were repeated. The integrity of the dialyzed protein was analyzed by SDS-PAGE and Coomassie Blue staining and then concentrated. The final concentration was quantified by UV absorption at 280 nm.

2.2.1.8 Protein concentration determination by Bradford assay

On binding of Coomassie brilliant blue G-250 to proteins, the maximum of absorption is shifted from 465 nm to 595 nm (Bradford, 1976). Absorption at 595 nm ($A_{595\ nm}$) was used to determine the concentration of protein samples. Initially, a standard curve was prepared by incubating a range of BSA

concentrations (2-12 µg/ml) with 1 x Bradford solution for 10 min and measuring $A_{595\ nm}$. The same procedure was followed for protein samples of unknown concentration. Based on the observed $A_{595\ nm}$ of the samples and the standard curve, the sample concentration were calculated.

2.2.1.9 Protein concentration determination by UV absorption at 280 nm

For proteins that contain tyrosines or tryptophans, a characteristic peak in the ultraviolet (UV) absorption spectrum around 280 nm is observed. The 280 nm absorbance can be used to calculate the protein concentration in pure protein solutions (Layne 1957, Noble *et al*. 2009). The protein concentration was calculated according to the Beer-Lambert law:

$$A = \mathrm{a}_m cl$$

where A is the absorbance at 280 nm, a_m is the molar extinction coefficient, c is the concentration in mg/ml and l the path length of the cuvette in cm, which is 1.

The protein solution was clarified by centrifugation at ca. 21.000 x *g* for 10 min so that any suspended particles, which may interfere with the measurement, were removed. The molar extinction coefficient of the protein was calculated with the ProtParam bioinformatics tool (Gasteiger *et al*. 2005, http://web.expasy.org/protparam/), based on the amino acid composition of the protein. The protein buffer solution was used to zero the spectrophotometer before measuring A_{280nm}. If the value was above 1.0, the sample was diluted in protein buffer and the measurement was repeated.

2.2.2 Cell culture and lysis

2.2.2.1 Cell culture of HEK293, HeLa, HepG2, MCF-7, SCC-9 and UPCI-SCC-154 cell lines

The growth and incubator conditions for the different cell lines are described in section 2.1.3. The medium of the cells was changed every 2-3 days. Cells were split 1:3 to 1:5 when they reached 80-90% confluency. In order to do so, culture medium was decanted and cells were washed with PBS. Next, cells were incubated for 5 minutes or less with 1-2 ml of 0.25% Trypsin and 0.1% EDTA in PBS solution until the cells detached. The cells were sedimented gently by centrifugation at 200 x *g* for 5 min. The cells were resuspended in fresh medium at the desired density.

2.2.2.2 Cell culture of immortalised podocytic cell line (PODs)

PODs are conditionally immortalised human podocytes, which were kindly provided by Professor Moin Saleem (University of Bristol). This cell line was developed by transfection of primary podocytic cells with the temperature-sensitive SV40-T gene and a telomerase gene. The detained clone, designated 'LY', proliferate at 33 °C and have epithelial morphology (Figure 2.1A). When the cells are transferred to 37 °C, their growth is arrested and they fully differentiate within 10-14 days (Ni *et al*. 2012). Upon differentiation, the cell bodies enlarged in irregular shapes and projections were formed (Figure 2.1B). The growth and incubator conditions are described in section 2.1.3. Cells were cultured as described in section 2.2.2.1 except being split 1:3 when they reached 40-60% confluency.

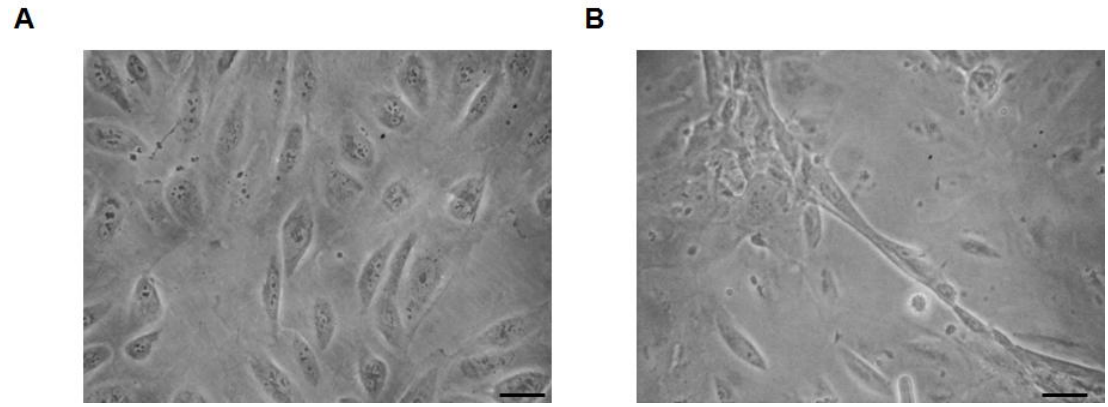


Figure 2.1

Light microscopy images show the morphology of the immortalised human podocytic LY cell line.

A, proliferative and ***B***, differentiated state (day 10). Magnification: x10. Scale bar: 100 µm.

2.2.2.3 Vanadate/ H_2O_2 stimulation of the UPCI-SCC-154 cell line

Vanadate/H_2O_2 mixtures inhibit tyrosine phosphatase activity in cells and animals (Swarup *et al*. 1982, Heffetz *et al*. 1990, Hecht *et al*. 1992, Chen *et al*. 1997). A stock solution with 50 mM vanadate and 500 mM H_2O_2 was made fresh prior to the start of the experiment. This stock solution was diluted 1:500 with cell-free, pre-warmed medium taken from the cells shortly before the experiment, mixed and immediately added onto the cells. The mixture was incubated with the cells for the designated time and then the cells were lysed as described in section 2.2.2.6. A control of untreated cells was set up and lysed as well.

2.2.2.4 Cell stocks

For long term storage of cells, culture medium was decanted and cells were washed with PBS. 1-2 ml of 0.25% Trypsin and 0.1% EDTA in PBS solution was added to cells ensuring the cells were fully covered. Any excess was decanted. Cells were incubated for 5 min or less, until the cells detached.

Cells were sedimented gently by centrifugation at 200 x *g* for 5 min and resuspended in freezing medium (95% FBS, 5% DMSO). Cells were aliquoted into cryovials and frozen slowly in styrofoam boxes at -80 °C. After one day, the vials were transferred into the liquid nitrogen tank.

2.2.2.5 Thawing frozen cells

The frozen vial of cells was transferred to a 37 ^{o}C waterbath for ca. 2 min, until it was fully thawed. The vial was then wiped with a tissue soaked in 70% alcohol prior to opening. Contents of the vial were slowly pipetted into a tube of pre-warmed medium and sedimented gently by centrifugation at 200 x *g* for 5 minutes. Cells were resuspended in fresh medium to achieve the correct flask volume.

2.2.2.6 Cell lysis

Prior to lysis, cells were washed three times with cold PBS. Any excess PBS was sucked off. A small volume (1 ml for a 175 cm^2 flask) of cold lysis buffer with inhibitors was added onto the cells and distributed by tilting the culture flask. Cells were scraped off the plastic surface, collected into prechilled tubes and incubated for 30 min on a nutator at 4 °C. The cell lysate was sedimented gently by centrifugation at 10.000 x *g* for 30 min at 4 °C. The supernatant was collected and the protein concentration was determined by the Bradford assay.

2.2.2.7 Cell fractionation

Cell fractionation allows the enrichment of specific subcellular compartments. This enables the study of intracellular protein distribution and simplifies the identification of low abundance proteins, which are found in specific

compartments (Michelsen *et al*. 2009). Here the method was used to separate cells sequentially into nuclear (P_1), large (P_{10}) and small (P_{100}) vesicle and cytoplasmic fractions (S_{100}) using increasing sedimentation speeds from 1000 to 100.000 x *g*.

Cells were washed twice with cold PBS and twice with cold hypotonic lysis buffer (HLB). Excess HLB was decanted. Cold lysis buffer (HLB buffer with inhibitors) was added onto the cells and distributed by tilting the culture flask. Excess lysis buffer was again decanted. Cells were scraped off the plastic surface, collected into the dounce homogeniser and allowed to swell for 10 minutes on ice. Afterwards, cells were dounce homogenised with 20 pestle strokes and collected into tubes. For the generation of specific subcellular fractions, the following steps were followed:

1. Cell suspension was sedimented by centrifugation at 1000 x *g* for 30 minutes at 4 °C. The supernatant (S_1) and pellet (P_1) were collected. P_1 contains nuclei, as well as unbroken cells and cell fragments, if any (Michelsen *et al*. 2009).
2. The S_1 cell suspension was sedimented by centrifugation at 3000 x *g* for 30 minutes at 4 °C. The supernatant (S_3) and pellet (P_3) were collected. The P_3 fraction is enriched in mitochondria.
3. The S_3 cell suspension was sedimented by centrifugation at 10.000 x *g* for 30 minutes at 4 °C. The supernatant (S_{10}) and pellet (P_{10}) were collected. P_{10} contains the large vesicles, including some remaining mitochondria.
4. The S_{10} cell suspension was sedimented by centrifugation at 100.000 x *g* for 30 minutes at 4 °C. The supernatant (S_{100}) and pellet (P_{100}) were

collected. S_{100} and P_{100} contains the soluble cytoplasmic proteins and small vesicles, respectively.

The pellets (P_1, P_3, P_{10} and P_{100}) were resuspended in cold lysis buffer (RIPA 100 with inhibitors) and incubated for 30 min on a nutator at 4 °C. Then, they were sedimented by centrifugation at 20.000 x *g* for 30 min at 4 °C. The supernatants were collected. The supernatant protein concentrations were determined by the Bradford assay. The workflow is also depicted in Figure 2.2.

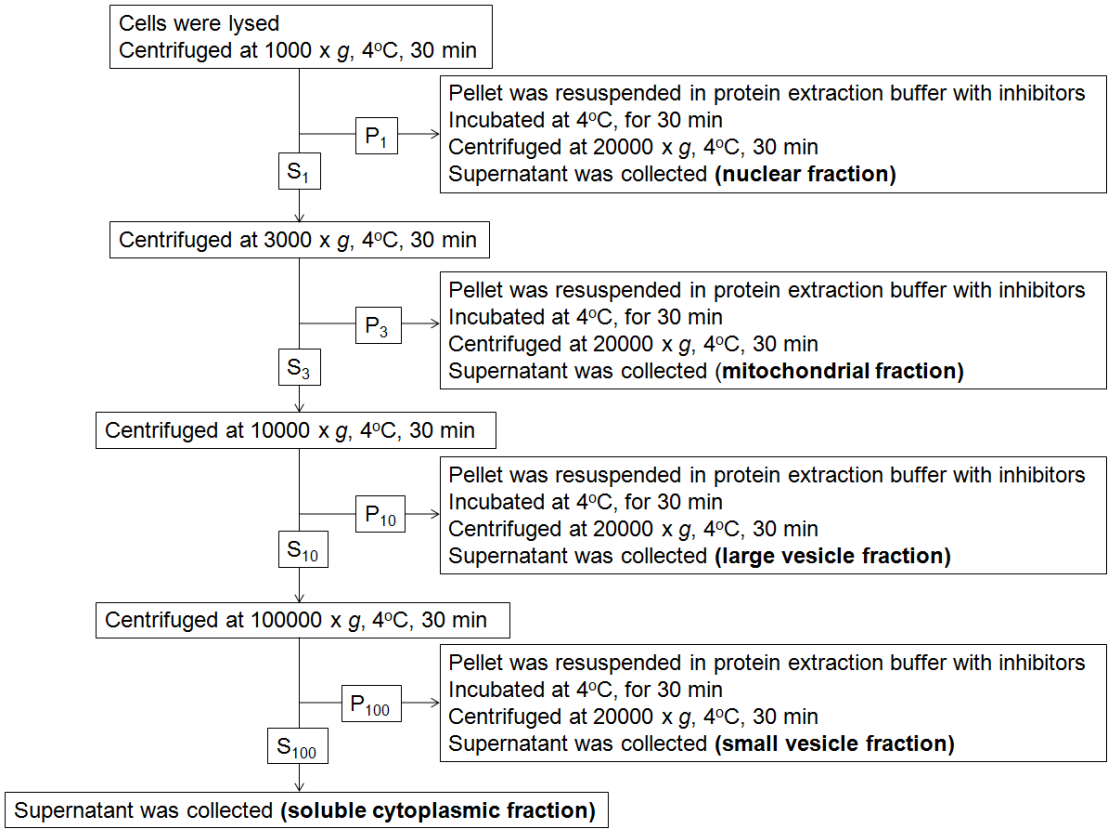


Figure 2.2

Subcellular cell fractionation workflow.

2.2.3 Biochemical assays

2.2.3.1 Precipitation of cellular proteins with GST-fusion proteins

20 µl of GSH beads and equimolar amounts of GST (30 µg) or GST-fusion proteins were mixed with 1 mg of lysate. The suspension was topped up with IP buffer to 500 µl and incubated overnight on a nutator at 4 °C. Then, beads were washed with 3 x 1 ml cold wash buffer (RIPA 100). Each time the GSH beads were sedimented gently by centrifugation at 500 x *g* for 5 min (4 °C). Bound GST-fusion proteins were analyzed by SDS-PAGE and western blot. In some cases, to reduce the amount of non-specifically precipitated proteins, the lysates were pre-cleared by incubation with 20 µg of GSH beads alone for 2 h, then 2 x 20 µl GSH beads + 50 µl GST for 2 h, and a further time with 20 µl GSH beads for 2 h. Beads were pelleted by centrifugation as before and the supernatant was used for precipitation with fresh beads and GST-fusion proteins. For the peptide competition experiments, a 200-fold molar excess of control or competition peptide were added to the GST-fusion proteins 2 h prior addition of the cell extracts. Unless stated otherwise, the experiments were performed at least in duplicates.

2.2.3.2 (Co-) Immunoprecipitations

Primary antibodies were used to precipitate cellular proteins together with Protein A or Protein G Sepharose beads (Sigma Aldrich). 2 µg of antibodies and 20 µl of packed Sepharose beads were mixed with 2 mg of cell lysate in ice-cold Eppendorf tubes. As controls, samples with control antibody (IgG [Santa Cruz Biotechnology], IgG_1 [Cell Signalling Technology], pre-immune rabbit serum [Santa Cruz Biotechnology]) and cell lysate or with antibody but without cell lysate were used. The stringency of cell lysis and wash buffer and

the incubation times were individually determined for optimal detection of specific signals. Proteins were precipitated at 4 °C on a nutator. Sepharose beads were washed with 3 x 1 ml cold IP buffer. Each time the Sepharose beads were sedimented by centrifugation at 20.000 x *g* for 1 min. Bound proteins were analyzed by SDS-PAGE and western blot. Experiments were performed at least in duplicates.

2.2.3.3 siRNA

RNA interference (RNAi) is a method to silence the expression of genes of choice through the interference with and degradation of their mRNA transcripts. This can be achieved by different types of RNA molecules, such as siRNA, miRNA and shRNA (Fire *et al* 1998, Bernstein *et al* 2001, Ketting 2011).

Human RIN3 siRNAs were purchased from QIAGEN™ (FlexiTube GeneSolution, GS79890). This kit included four individual siRNAs. The sequences are the following:

RIN3 siRNA	Catalogue number	Sequence
RIN3-1	SI04319021	AGUCGAUAUUGUACCUGGA
RIN3-2	SI04334036	CGGAGAUGCUUCUCAAUGU
RIN3-3	SI04212131	GCAGCAUGUUCCACGCUUU
RIN3-4	SI04308794	GGAGCUCGAAGCAAUUGUA

The stock of each individual siRNA was prepared according to QIAGEN's guidelines. HEK293 and HepG2 cells were seeded in 6-well plates at a density of $2x10^5$ cells/well and transfected with the RIN3-targeted siRNAs (individually or as a pool) or negative control siRNA (Luciferase GL2 duplex,

Thermo Scientific Dharmacon^{TM}). The negative control siRNA is expected to not affect gene expression, because it has no homology to any known human gene. Moreover, a positive control well with p130Cas customary siRNA (sequence: GCAGCAGCUGAAGCAGUUU), which was validated by a previous lab member, was used to confirm that the transfection protocol was successful. A control well with mock-transfected cells was also used to test if there were any off-target effects due to the transfection reagents. Different transfection reagents, time points, media and procedures were used so that the optimal protocol of siRNA transfection could be determined. Details can be found in section 5.4.3. The incubator conditions for the HEK293 and HepG2 cell lines are described in section 2.1.3.

2.2.4 Protein separation and detection

2.2.4.1 SDS polyacrylamide gel electrophoresis

4x SDS PAGE sample buffer was added to the protein solution and the mix incubated at 100 ^{o}C for 5-10 min to denature proteins. The SDS-PAGE sample buffer contains SDS and β-ME. The former unfolds hydrophobic regions of the proteins, while the latter causes a reduction of disulphide bonds present. The denatured proteins were loaded onto polyacrylamide gels that consist of a separation gel (pH 8.8), which is below a low-percentage stacking gel (pH 6.8). Proteins were separated by size on vertical polyacrylamide gels in a Protean II xi electophoresis chamber (Bio-Rad).

2.2.4.2 Detection of proteins immobilised on membranes with antibodies (Western blot)

After electrophoresis, the gel was cut to an appropriate size and transferred into semi-dry blot buffer (SDBB). Proteins were transferred from the gel onto a PVDF membrane (Fisher Scientific) by semi-dry blotting (TransBlotSD Semi-Dry Transfer Cell, Bio Rad). The PVDF membrane was pre-wetted in methanol for 15 sec and then transferred into SDBB. Six layers of Whatman 3MM chromatography paper were soaked in SDBB and placed onto the transfer cell electrode. Afterwards, the PVDF membrane, the gel and six further layers of Whatman paper were positioned sequentially on top of the Whatman paper, while removing remaining air bubbles. A constant voltage of 20 V for 1 h was chosen for the successful transfer of the proteins. The PVDF membrane was then incubated in blocking buffer appropriate for the primary detection antibody for 1 h at room temperature. This was done to ensure the saturation of non-specific binding sites on the membrane. The membrane was then incubated with the primary detection antibody overnight at 4 °C on a nutator. After washing three times for 15 min in the appropriate wash buffer, the secondary HRP-coupled antibody in the same fresh blocking buffer (1:10.000) was added and incubated for 1 h at room temperature. After washing three times with the same buffer for 15 min, the freshly mixed ECL detection solutions (Thermo Scientific Pierce^{TM}) were applied for 1 min. Excess liquid from the membrane was drained with filter papers. Light produced by oxidation of luminol in the detection solution, catalysed by antibody-coupled horse radish peroxidase, was detected by autoradiography on X-ray films or an imaging system (Fusion FX7^{TM}).

2.2.4.3 Silver Staining of SDS-Gels

Proteins were separated by SDS-PAGE. The gel was fixed by shaking for 20 min in 50% (v/v) methanol with 5% (v/v) acetic acid, then washed for 10 min in 50% methanol, and a further 10 min in H_2O. The gel was then incubated for 1 min in 0.02% (w/v) sodium thiosulphate and afterwards thoroughly washed for 1 min under a running water tap. Silver was bound by incubation of the gel in 0.1% (w/v) silver nitrate solution for 20 min. It was then washed again for 2 x 1 min by shaking rapidly with H_2O. Proteins were detected by developing in 0.04% (v/v) formaldehyde with 2% $NaHCO_3$. To stop the reaction, the gel was immersed for 10 min in 5% (v/v) acetic acid and stored in 1% acetic acid.

2.2.4.4 Gel filtration chromatography separation of S_{100} subcellular fraction

The S_{100} fraction was prepared as described in section 2.2.2.7. Nine 175cm^2 flasks were used, which were ca. 95-100% confluent. The obtained S_{100} lysate was loaded onto a HiLoad^{TM} 16/60 Superdex^{TM} S200 column (GE Healthcare). The column was run at 1 ml/min with a modified buffer (10 mM Tris pH 7.5, 150 mM NaCl, PBS, 50 mM EDTA pH 8, 0.7 µg/ml pepstatin A, 0.5 µg/ml leupeptin, 0.1 mM PMSF, 5 µg/ml antipain, 1 mM sodium orthovanadate, 1 mM sodium molybdate).

2.2.4.5 Ammonium sulphate precipitation

Precipitation or 'salting out' of proteins is caused by the formation of protein aggregates due to neutralization of protein surface charges by addition of salt. A number of salts can be used for this process, but ammonium sulphate is one of the most widely used protein precipitants. Since proteins differ in their size, shape and surface properties, they will precipitate at different salt

concentrations. This allows fractionation of proteins according to their solubility and precipitability (Englard *et al*. 1990).

The protein mixture (GST-tag, GST-SH3-[1+2], SH3-[1+2]) was fractionated by the addition of solid ammonium sulphate at 4^{o}C. Fractions were collected at concentrations of 30, 40, 50, 60 and 80% ammonium sulphate. For the collection of each fraction, the required amount of solid ammonium sulphate was added to the protein solution with stirring for 1 h to allow the equilibration of the protein mixture. The mixture was then centrifuged at 10.000 x *g* for 15 min. The resulting pellet was dissolved in 20 mM Tris pH 7.5, 150 mM NaCl and 2 mM β-ME. More solid ammonium sulphate was added to the supernatant to achieve the next higher concentration, while the stirring, and centrifugation steps were repeated as before. The proteins of each fraction were identified by SDS-PAGE and Coomassie Blue staining.

2.2.5 Isothermal titration calorimetry

2.2.5.1 Overview

Isothermal titration calorimetry (ITC) is a method that enables the direct measurement of heat associated with complex formation at constant temperature. This allows the simultaneous determination of changes in the thermodynamic parameters, such as free energy of binding (ΔG), enthalpy (ΔH) and entropy (ΔS). Therefore, ITC has many applications and is used to study enzyme kinetics and to characterize the interactions between macromolecules, such as proteins, DNA and lipids (Doyle 1997).

2.2.5.2 Instrumentation and ITC data calculation

In the ITC experiment, typically a ligand is injected in aliquots into a solution containing the macromolecule of interest and the heat released (exothermic reaction) or absorbed (endothermic reaction) is detected by measuring the power required to maintain constant temperature between the sample and a reference solution (Figure 2.3A). The syringe that contains the ligand has a stirrer, which is rotated continuously, and ensures that the contents of the sample cell are quickly and evenly mixed. The heat released or absorbed upon each addition of ligand into the sample cell corresponds to the area under the signal-versus-time curve (Figure 2.3B, top panel) and relates to the amount of complex formed. In an ideal case (i.e. macromolecule and ligand are soluble and bind with moderate to high affinity) with excess ligand, the injections of ligand lead to the saturation of the binding sites, which causes the heat of reaction to tend to zero (Figure 2.3B). At the end of the reaction a residual heat may be observed which can be attributed among others to buffer dilution effects and mechanical heat of stirring of the syringe. These contributions must be determined by performing control experiments and substracted from the total heats to derive the true heat of complex formation.

The resulting binding isotherm is a plot of heat absorbed or released per injection versus the molar ratio of the ligand to macromolecule of interest (Figure 2.3B, bottom panel). By assuming the interaction occurs under equilibrium conditions, the binding isotherm enables the determination of the binding enthalpy ΔH, the equilibrium binding affinity K and the stoichiometry N (i.e. number of binding sites per mole of macromolecule) (Figure 2.3B, bottom panel).

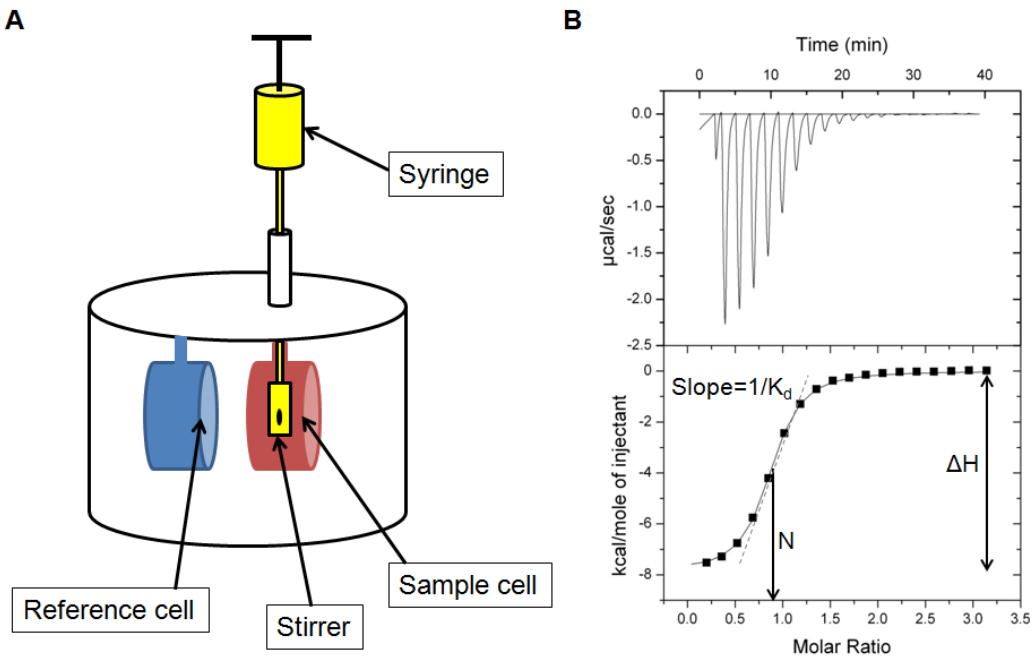


Figure 2.3

A, Illustration of an isothermal titration calorimetry instrument. The sample and reference cells and the syringe are depicted. B, Raw ITC data (upper panel) and the resulting binding isotherm (bottom panel).

The solid line represents the nonlinear best fit to the data assuming a single-site binding model. It is also depicted how the parameters ΔH, N and K_d are derived from the isotherm.

It is possible to calculate the N value because the macromolecule and ligand concentrations are known and thus it is determined by the molar ratio of interacting species at the equivalence point. This leads to the determination of the Gibbs free energy of binding ΔG, the change in entropy ΔS and the dissociation constant K_d from the following equations:

$$\Delta G = -RT \ln K$$

$$\Delta G = \Delta H - T\Delta S$$

$$Kd = \frac{1}{K}$$

As shown above, affinity is quantified by ΔG. ΔH and ΔS provide additional information, which were already described in section 1.4. In general, ITC

measurements are able to calculate disassociation constant values between approximately 10 nM and 100 µM (Perozzo *et al* 2004, Ladbury 2010, Ghai *et al* 2012).

2.2.5.3 ITC experiments

ITC was performed with a VP-ITC MicroCalorimeter (MicroCal). Peptides were dissolved at 0.75 mM in ITC buffer, clarified by centrifugation for 10 min at 20.000 x *g* and degassed prior to use. Clarified and degassed protein solutions (1.43 ml in total) comprising 0.05 mM CD2AP SH3-1, SH3-2 or SH3-3 domains in ITC buffer were placed in the sample chamber. Upon reaching the equilibrium temperature of 25 °C, peptide solutions were titrated into the sample chamber by an initial injection of 4 µl followed by 18 injections of 15 µl. Resulting peaks of measured deviations from the equilibrium temperature were integrated to yield the quantity of heat generated. Data were obtained using χ2 minimization on a model assuming a single set of sites to calculate the binding affinity K_d. All data analysis steps were performed using the manufacturer's ORIGIN (V5.0) software. The reported standard error was computed by the calorimeter.

2.2.5.4 Peptide Synthesis for ITC and Crystallography

A list of the peptides' sequences can be found in section 2.1.2. The ALIX- and RIN3-derived peptides were provided by Dr Nicola O' Reilly (Peptide Synthesis Laboratory, Cancer Research UK, London). Briefly the peptides were synthesized using 9-fluorenylmethyloxylcarbonyl for temporary α-amino group protection, cleaved from the resin, deprotected, freeze-dried, purified on a C8 RP column, and analyzed for purity and mass on an Agilent 1100 LC-

MS. The ARAP1-, DAB1- and MLK3-derived peptides were provided by the peptide synthesis facility at the WIMM (Oxford, UK).

2.2.6 X-ray crystallography - Theory

2.2.6.1 Overview

Molecular models generated by X-ray crystallography are used as tools to study the molecular details of biological processes. A number of steps are followed for the determination of a protein structure. (Rhodes, 2010) has been used as a reference to describe these steps in the following sections.

2.2.6.2 Protein Crystallization

For protein crystal growth, purified protein is dissolved in aqueous buffer solution under non-denaturing conditions. The buffer solution contains a precipitant, such as polyethylene glycol or salts, at a concentration below that necessary to precipitate the protein. Controlled evaporation of water causes an increase in protein and precipitant concentrations, which results in slow precipitation. This may occasionally give rise to the growth of either small crystals or crystals that are large enough for the X-ray diffraction experiments.

Small crystals of good quality might be used for the production of appropriately large crystals. This method is called *microseeding*. The experimental set-up is the same as before, except that the protein solution is seeded with the small crystals, providing a nucleus around which larger crystals may grow.

For crystals to grow, it is important to find the best crystal growth conditions. Several factors are important for crystal growth, including protein and precipitant concentration and purity, pH and temperature.

A useful way of choosing the appropriate starting concentration for the crystallization screens is the *pre-crystallization test (PCT)*. Various protein concentrations are tested which should yield clear drops or varying degrees of precipitation with selected reagents (Hampton Research, US). The protein concentration giving rise to moderate precipitation is considered a suitable starting point.

The crystals described in this thesis were formed by the *sitting-drop vapour diffusion method*. In this method, the protein/precipitant solution is placed on a support and equilibrated with a larger precipitant reservoir in a closed container.

2.2.6.3 Collection of X-ray Diffraction Data

For X-ray data collection, the crystal is mounted onto a *goniometer.* The goniometer allows the rotation of the crystal between an *X-ray source* and an *X-ray detector*, while being centered in the beam. The diffraction by the crystal of the X-ray beam results in the production of a diffraction pattern of X-ray *reflections* (spots in Figure 2.4) on the detector. The intensity of each reflection and their coordinates (or *indices*) in the *reciprocal lattice*, which is the inverse of the *real (or crystalline) lattice*, are directly measurable.

The real lattice is composed of *unit cells*. The unit cell is the minimum repeatable translational unit from which the volume of the crystal may be built in three dimensions. The dimensions of the unit cell can be calculated because the spacing of unit cells in the real lattice is inversely proportional to the spacing of reflections in the reciprocal lattice on the X-ray diffraction pattern (Figure 2.4). The content of the unit cell (i.e. the structure of the protein) is determined by the spot intensity and the phases. As it will be

described in section 2.2.6.4, the phases are not measurable, but are initially estimated by various methods.

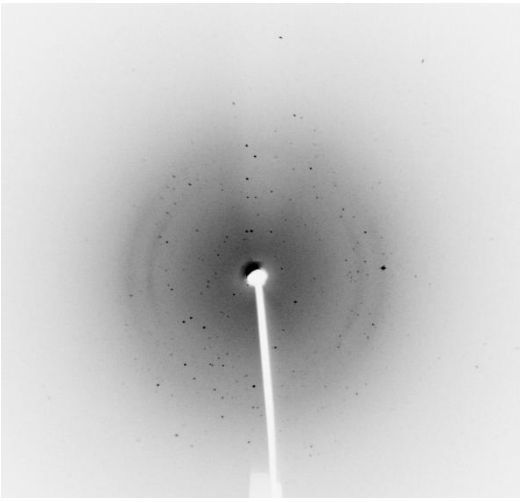


Figure 2.4: X-ray diffraction pattern.

The unit cell may contain a few identical molecules arranged in a symmetric way. The largest group of such molecules that do not possess symmetry elements is called the *asymmetric unit*. The symmetry of the contents of a unit cell is described by its *spacegroup*. The spacegroup is referred to by the lattice type, such as C, and the symmetry directions, such as [121]. There are 230 spacegroups, but due to chirality, only 65 spacegroups are possible for proteins.

2.2.6.4 Electron Density Computation

Software, which is described in section 2.2.7.2, is used to compute the electron density map within the unit cell from the indexed intensities. This map leads to an image of the molecule. Electron density *ρ(x, y, z)* is a periodic function that is represented by the following Fourier sum:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F_{hkl}|e^{-2\pi i(hx+ky+lz-a_{hkl})}$$

where *h*, *k* and *l* are the indices of reflection hkl, V is the volume of the unit cell,$|F_{hkl}|$ is the structure factor amplitude and α the phase of this reflection. The structure factor F_{hkl} is described by frequency, amplitude and phase. The amplitude is derived from the intensity of the reflection, while the frequency is derived from the indices h, k and l. However, the phase of each reflection α_{hkl} is not measurable in the X-ray diffraction experiment and this problem is known as the *phase problem*. A number of methods can be used to calculate initial estimates of α_{hkl}: *isomorphous replacement*, *anomalous scattering* and *molecular replacement*. The calculation of α_{hkl} for the structures solved in this thesis was done by molecular replacement.

2.2.6.5 Molecular Replacement

In molecular replacement, the phases from structure factors of a known protein, which is called a *phasing model*, are used as initial estimates of phases for the target protein by positioning a model of the known protein in the unit cell of the target protein. This is easier when the target protein is identical to a component of the phasing model, as in the case of crystals soaked with a ligand (*isomorphous phasing model*). Another possibility is to use molecular replacement when the target protein and the phasing model exhibit high sequence homology (*nonisomorphous phasing model*). In this case, calculating α_{hkl} might still be feasible, but more challenging.

When the phasing model and target protein are isomorphous, the electron density is described by the following Fourier sum:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F_{hkl}^{target}|e^{-2\pi i(hx+ky+lz-\alpha_{hkl}^{model})}$$

where the structure factor amplitude of each reflection $|F_{hkl}|$ is derived from the intensities of the reflections of the target protein, while the phases α_{hkl} are those of the phasing model.

When the phasing model and target protein are nonisomorphous, the optimal location and orientation of the phasing model have to be found in the unit cell of the target protein so that the phasing model is superimposed on the target model. The unit cell dimensions and symmetry are computed by the data of the target protein as described above. This allows the calculation of α_{hkl} for the appropriately oriented phasing model, which will be used as initial estimates of α_{hkl} for the target protein. The search for the optimal orientation of the phasing model is done by the *Patterson function*. The Patterson function is a Fourier sum. However, it has no phases. If the Patterson maps (i.e. maps of Patterson function) of the phasing model and the target protein are similar, then this optimal orientation of the phasing model is used to search for its optimal location. The search for the optimal location is done by computing the structure factors of the phasing model (F_{calc}) for various model locations and comparing their amplitudes with the calculated amplitudes of the structure factors of our target protein (F_{target}). The optimal location is the one where there is the best match between F_{calc} and F_{target}.

After computation of the initial estimates of α_{hkl}, the phase estimates are improved by refining the initial estimated model. This leads to an improvement of the electron density maps, which results in a structural model that is in closer agreement with the data of our target protein.

2.2.6.6 Assessment of the Molecular Model

The R_{work} and R_{free} factors are used to assess the refined model. Their values are expected to decrease during each refinement cycle. The R_{work} factor is derived when the structure factor amplitude of the resulting model is compared to the calculated structure factor amplitude of the target protein's data:

$$R_{work} = \frac{\sum ||F_{target}| - |F_{model}||}{\sum |F_{target}|}$$

R_{work} values are between 0 and 1. Zero signifies that the structure factors between the model and the data are in complete agreement. In practice, R_{work} values for the final model are around 0.2. The R_{free} factor is computed based on a set of intensities that are randomly chosen prior to refinement and not used. Thus, the R_{free} value is a way of measuring how well the resulting model can predict the values, which were set aside.

An additional way of assessing the model at late stages of refinement is through a *Ramachandran diagram*. The diagram shows the pair of torsional angles (Φ, Ψ) for each residue and whether this pair is within conformationally allowed regions. The torsional angle Φ is the one around the N — Cα bond, while the torsional angle Ψ is around the Cα — C bond. Therefore, the Ramachandran diagram enables the determination of any conformationally unrealistic parts of the model, which have to be further refined.

Finally the *temperature factor* B_j of each atom j or *B-factor* reflects the uncertainty in determining the exact atom positions in our model. It describes whether an atom oscillates more (high B-factor) or less (low B-factor) around its given position. The contribution of thermal motion is included within the B-factor. Therefore, the B-factor is a useful tool to compare the flexibility of

different parts of our model. However, one should keep in mind that other contributions, such as static disorder, is included in this parameter. Therefore, data analysis based on B-factors should be performed with caution.

2.2.7 X-ray Crystallography - Experimental work

2.2.7.1 Pre-crystallisation test protocol

A pre-crystallisation test (PCT) was used to determine the appropriate starting concentrations of CD2AP SH3-1 and SH3-2 domains for the crystallization screens. The test was set up by the hanging-drop vapour diffusion method in 96-well plates. The reagents and tested concentrations are found in Table 2.1.

Table 2.1: Pre-crystallisation test reagents and tested concentrations.

Reagents			
A1	0.1 M Tris HCl pH 8.5, 2.0 M Ammonium Sulphate		
A2	0.1 M Tris HCl pH 8.5, 0.2 M Magnesium Chloride, 30% (w/v) PEG 8000		
A3	0.1 M Tris HCl pH 8.5, 30% isopropanol		
Concentration (mg/ml)			
SH3 domains	**no peptide**	**RIN3e1**	**RIN3e2**
SH3-1	36.7	19.7	18.3
SH3-2	28.2	16.9	15.5

In the crystallisation trials of the SH3 domains in complex with the RIN3 peptides, the complexes were pre-formed just before the PCT test at a ratio of 1:3.5 (SH3 domain:RIN3e1 peptide) and 1:3 (SH3 domain:RIN3e2 peptide) respectively. These ratios were tested based upon the ITC results. Crystallisation trials were set up at 22 °C from a 1:1 ratio of mother liquor (Reagent A1-A3) to protein solution. The total volume of each drop and reservoir solution was 1 and 150 µl, respectively. The drop was placed underneath a cover slip, which was sealed on top of the reservoir well with

vacuum grease silicone. The drops were assessed after 24 h of setting up the experiment.

2.2.7.2 Software

The theoretical background of molecular model generation by protein X-ray crystallography was described in the section 2.2.6. A number of programs are used in practice to achieve this.

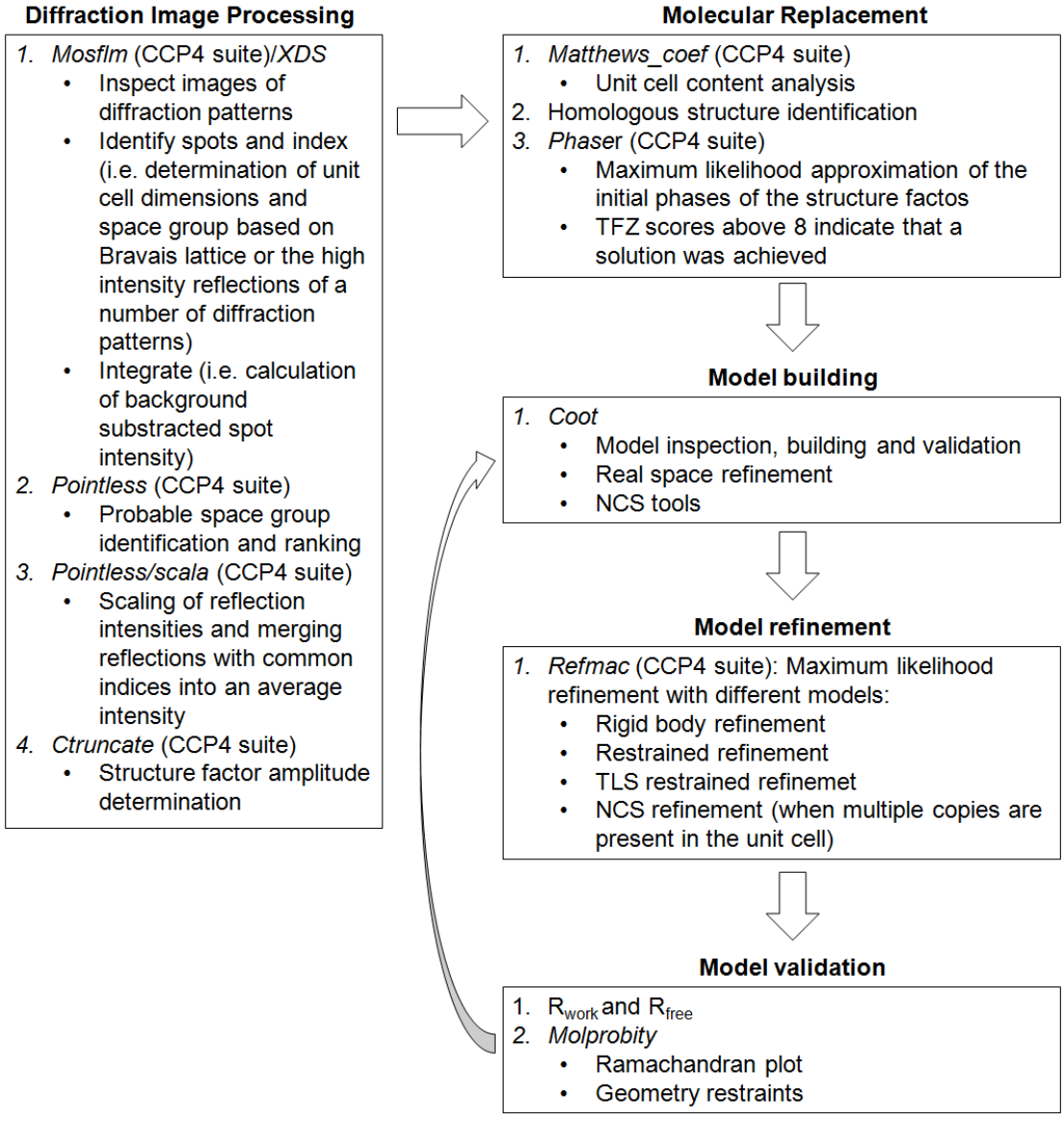


Figure 2.5

Data processing and refinement workflow.

The program name is written in italics and a description of its use is described in the bulleted list.

These programs and the general workflow are summarized in Figure 2.5, and will also be mentioned in sections 3.6.2.1 and 6.8.2.1.

2.2.7.3 Crystallization, data collection and analysis of CD2AP SH3 domains with the RIN3- and ARAP1-derived peptides

Trials were set up by the vapour diffusion method in a sitting nanodrop (total volume of 150 nl) in 96-well plates using the following crystal screens: HCS, HIN, JCSG, LFS. The screen condition are found in section 2.1.5. The starting concentrations for CD2AP SH3-1 and SH3-2 were chosen according to the pre-crystallisation test.

CD2AP SH3-2/RIN3e2 complex: Crystals were grown at 20 °C from a 2:1 ratio of mother liquor (0.1 M HEPES pH 7.5, 1.4 M tri-sodium citrate dehydrate) to CD2AP SH3-2 protein solution (16.7 mg/ml). Dr Philip Simister harvested the crystal and solved the structure. A single, thin square plate (~180 μm in length/width), which grew after 3 weeks, was first transferred to cryoprotectant (mother liquor supplemented with 10% glycerol) before being snap-frozen in liquid nitrogen.

Data were collected to 1.11 Ångstrom on beamline I03 at the Diamond synchrotron (Harwell, UK). Data reduction was carried out with the *XDS* package (Kabsch *et al* 2010). Data was incomplete in most resolution shells (spacegroup C2), so an additional dataset was later collected using the same crystal but reorienting it in the beam (Diamond beamline I04-1). The two datasets were merged with *XSCALE* before being converted to *CCP4* mtz format. Molecular replacement was performed with *Phaser* (McCoy *et al* 2005) using an ensemble of SH3 models homologous to CD2AP SH3-2 (PDB codes: 2G6F, 1OEB, 2AK5, 2FEI, 3IQL). The single, correct solution was

subjected to automatic model building with the *AutoBuild* program in the *PHENIX* suite (Zwart *et al* 2008) and later refinement with *REFMAC 5* (Murshudov *et al* 1997). Additional manual model building was done with *Coot* (Emshley *et al* 2004). Anisotropic B-factors and several alternative conformations were modelled. The final model was deposited in the PDB with accession code 3U23. Data collection, processing and refinement data are depicted in Table 2.2.

Table 2.2: Data collection, processing and refinement data.

Values in parentheses represent the highest resolution shell: 1.14 – 1.11 Å.

Data collection and processing:		
Diamond beamline		I03
Detector		Pilatus 6M
Wavelength (Å)		0.97625
Space group		C2
Unit cell parameters	(Å)	53.13 32.04 39.18
	(°)	90.00 92.56 90.00
Resolution range (Å)		27.43 – 1.11
Highest resolution shell (Å)		1.14 - 1.11
I/σ (I)		14.98 (2.24)
Completeness (%)		92.8 (59.36)
R meas (%)		5.6 (52.1)
Refinement:		
R work (%)		15.1 (33.7)
R free (%)		16.7 (29.1)
R.m.s.d. bond angles	(°)	2.27
bond lengths	(Å)	0.025
Ramachandran plot (%)		
Favoured regions		98.31
Allowed regions		1.69
Outliers		0

CD2AP SH3-1/RIN3e1 and CD2AP SH3-2/ARAP1e1 complexes: Information on the crystallisation, data processing and refinement data are found in sections 3.6.2.1 and 6.8.2.1. The final models were deposited in the PDB with accession codes 4WCI (SH3-1/RIN3e1) and 4X1V (SH3-2/ARAP1e1).

The figures of the crystal structures of this thesis were generated using the PyMOL™ Molecular Graphics System software (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC). The colouring of the crystal structures as a function of B-factors following a gradient of colours was generated using the *color_b* script in PyMOL™ (http://pldserver1.biochem.queensu.ca/~rlc/work/pymol/), which was written by Robert L. Campbell and James Stroud. The structures were coloured with the Blue-White-Red gradient in 30 colours of equal numbers of atoms in each colour.

2.2.8 Peptide arrays

2.2.8.1 Overview

Peptide arrays enable the study of protein-peptide interactions. They are particularly useful for studying the binding of protein interaction domains, which often recognise short linear motifs. As a result, they have a variety of applications such as: mapping of binding epitopes, elucidation of binding determinants and specificities, and identification of interaction partners (Volkmer *et al*. 2012).

2.2.8.2 Synthesis of peptide arrays

The RIN3 peptide permutation arrays were kindly provided by Dr Rudolf Volkmer (Charite, Institute of Medical Immunology, Berlin, Germany). Apart from the RIN3 peptide permutation arrays, all arrays were synthesized by Dr Nicola O' Reilly (Peptide Synthesis Laboratory, Cancer Research UK, London). Details on the synthesis are found in Appendix C.1.

2.2.8.3 Probing of peptide arrays

2.2.8.3.1 RIN3 - Permutation array

Two 17 aa sequence regions from RIN3 (AKKNLPTAPPRRRVSER and TAKQPPVPPPRKKRISR) that bind to CD2AP SH3 domains were permutated in 15 single positions (underlined) to all 20 other residues, and spot-synthesised in single rows. Membranes were briefly pre-wetted in ethanol, then washed with RIPA 100 for 3 x 10 min and blocked in RIPA 100 with 2% (w/v) ovalbumin (Sigma Aldrich) for 4 h at room temperature. The membrane was then incubated with 0.1 µM (2.6 µg/ml) purified GST in blocking buffer overnight at 4 °C to detect any background signals elicited by the GST-tag. After washing 3 times in RIPA 100 for 10 min, the membrane was probed with an anti-GST mAb, washed as before with TBST and incubated with HRP-coupled secondary (anti-Ig) antibody. After washing, the membrane was re-blocked for 4 h and probed with 0.1 µM (3.6 µg/ml) GST-CD2AP SH3 domain overnight at 4 °C. Bound GST-CD2AP SH3 fusion protein was detected as described for GST. No background GST signal could be detected after 2 mins, whereas the signal from GST-SH3 appeared after 1 sec.

The arrays were stripped by incubation for 5 x 20 min at 50 °C in pre-warmed blot stripping buffer, with gentle shaking. Afterwards, the membrane was rinsed 3 x 15 min in RIPA-100, then re-blocked for 2 h. The stripping efficiency was tested by probing the arrays with 0.1 µM GST overnight and following the already described protocol to detect any signals. Next, the arrays were re-probed with the GST-SH3 domains but with blocking buffer supplemented with 2 mM DTT. As will be explained in section 3.4, this was

done to prevent disulphide bond formation between the cysteine-rich GST-tag and the peptides.

2.2.8.3.2 CD2AP proline-rich region scanning array

The CD2AP proline-rich region (aa 330-520) was spot-synthesised as partially overlapping 27 aa peptides, sliding 3 aa with each step. The arrays were probed as described above (section 2.2.8.3.1) but with different blocking buffer (TBST, 5% [w/v] non-fat dry milk) and incubation time (1 h) with the blocking buffer prior to GST or GST-SH3 addition.

2.2.8.3.3 RIN3 scanning array

For RIN3 antibody epitope mapping, the RIN3 scanning array was probed as described in section 2.2.8.3.2 but with the RIN3 antibody (1:100 dilution) rather than GST. Next, the array was stripped as described in section 2.2.8.3.1. The stripping efficiency was tested by overnight incubation with secondary antibody and detection of any signals. Then, the array was re-probed with a different RIN3 antibody.

2.2.8.3.4 Proteome-wide search for novel putative CD2AP SH3 binding partners

The P - {FGHWY} - P/A - {D} - {FGHWY} - R - {P} motif (residues that are not tolerated are shown within curly brackets), which was found to be the dominant linear recognition motif for the three CD2AP SH3 domains, was used as a search sequence for scanning the human proteome with ScanProsite (de Castro *et al*. 2006). This search resulted in 11963 hits. A parallel search with the PrePPI database (Zhang *et al*. 2012) resulted in 1472 potential CD2AP interaction partners. The results from ScanProsite and

PrePPI had 386 proteins in common. From 386 hits, a subset of 300 potential binding sites in proteins was selected. The selection criteria are described in section 6.2. These selected CD2AP SH3 binding sequences were then spot-synthesized on a microarray as 16-aa peptides, in duplicate, for validation of their in-vitro binding ability. The arrays were probed as described in section 2.2.8.3.2 but with different blocking buffer (TBST, 3% ovalbumin, 2 mM DTT).

2.2.8 Dynamic Light Scattering (DLS)

2.2.8.1 Overview

Dynamic light scattering (DLS) is a fast method of determining the size of biomolecules in solution. The time-dependent alteration of the light scattering intensity is measured at a constant specified temperature and solvent composition. This enables the calculation of the hydrodynamic radius due to the Brownian motion of molecules in solution. Therefore, DLS can be used to calculate molecular weights and to distinguish between monodispersed and aggregated samples (Ruf *et al*. 1989).

2.2.8.2 DLS experiment

Dynamic light scattering data were collected with a Viscotek 802 DLS instrument (Malvern Instruments). 15 µl of SH3-1 or SH3-2 domains (2 mg/ml) in 5 mM Tris pH 7.5 and 2 mM β-ME in the presence or absence of peptide were clarified by centrifugation for 10 min at 20.000 x *g* and loaded into the quartz cell. Control measurements with peptide and buffer alone were also performed. OmniSIZE^{TM} software (Malvern Instruments) were used to determine the molecular weights of the protein samples.

Chapter 3: Structural and biochemical investigation of the CD2AP SH3 interaction with RIN3 and ALIX using individual SH3 domains

3.1 Overview

One of the aims of this project was to explore the binding properties and specificity of the three CD2AP SH3 domains by investigating their interactions with different binding partners. The three CD2AP SH3 domains are referred to as SH3-1, -2 and -3 throughout this thesis. The SH3 domains were expressed and purified individually and their interactions with different partners were analysed by peptide scanning arrays, isothermal titration calorimetry (ITC) and X-ray crystallography. These three techniques complement each other, which allows us to have a more detailed and reliable view of the SH3-mediated interactions. The chosen interaction partners were RIN3 and ALIX.

Regarding the CD2AP-RIN3 interaction, the previous analysis by peptide arrays had revealed that the three SH3 domains bind to two RIN3 epitopes (aa 378-393 and aa 452-467) (Dr Melanie Janning, section 1.6.4). The two RIN3 epitopes will be referred to as RIN3e1 and RIN3e2, respectively, throughout this thesis. Corresponding peptides were synthesised (Nicola O'Reilly *et al*., Peptide Synthesis Laboratory, Cancer Research UK, London). Preliminary ITC measurements determined that 16 amino acid peptides are sufficient to bind to the SH3 domains without affinity loss (Dr Melanie Janning, unpublished data). This initial work was followed up by analyzing the CD2AP/RIN3 interaction by the three above mentioned techniques. Two high

resolution structures of the SH3-1 and SH3-2 domains in complex with RIN3e1 and RIN3e2 respectively provided a structural view of this interaction.

Regarding the CD2AP/ALIX interaction, peptide permutation arrays of the ALIX epitope had revealed that the first two SH3 domains have a similar recognition pattern and that the SH3-3 domain recognises a more extensive region in ALIX compared to the first two SH3 domains (Dr Tassos Konstantinou, section 1.6.3). ITC and crystallisation trials between the SH3 domains and the ALIX epitope were performed. Unfortunately, these crystallisation trials were unsuccessful. Therefore, the study focused on a complete structural and biochemical characterization of the CD2AP/RIN3 interaction.

Finally, it was investigated whether the boundaries of the SH3-3 domain are critical for its activity. Recent work concerning the SH3-3 domain of CIN85, which is a CD2AP homologue, has shown that the SH3 domain boundary residues might be critical for domain stability, at least in some cases. Therefore, two SH3-3 domain constructs with differential boundaries (aa 263-334 and aa 270-331) were purified and their properties were compared by performing isothermal titration calorimetry experiments with the two RIN3 epitopes. These constructs will be referred to as SH3-3-S (aa 270-331) and SH3-3-L (aa 263-334) throughout this thesis. The boundaries of the CD2AP SH3-3-L construct were equivalent to the boundaries of the CIN85 SH3-3 domain analysed by Philippe *et al.* (2011).

3.2 Purification of individual CD2AP SH3 domains

Individual SH3 domains were expressed as GST-tagged proteins in *E.coli* and affinity-purified. The GST-tagged proteins were used as probes in peptide scanning array experiments. For ITC experiments and crystallization screens, the GST-tag was removed by proteolysis and proteins were further purified by size exclusion chromatography (SEC) (section 2.2.1). Figure 3.1 represents a summary of the workflow.

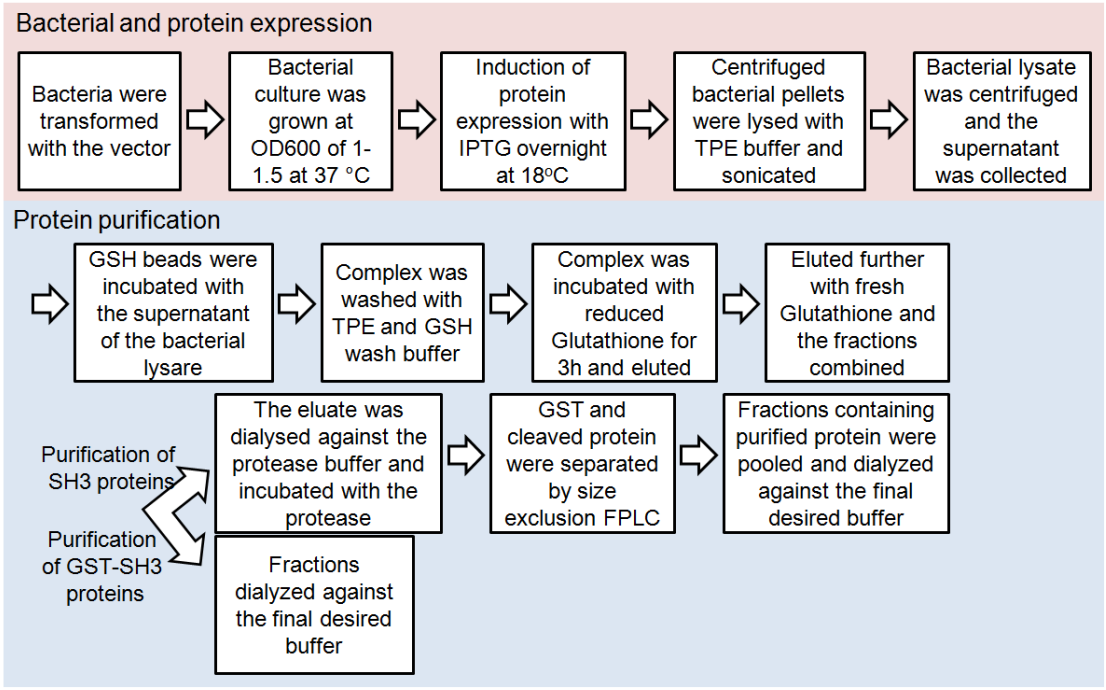


Figure 3.1

General workflow of bacterial expression and purification of the individual CD2AP SH3 domains.

The samples were estimated to be more than 95% pure at the end of the purification process (Figures 3.2-3.5). The final yields of the four SH3 domain construct expressions were 4-5 mg SH3-1/l culture, 1.5-2.5 mg SH3-2/l culture, 0.8-2 mg SH3-3-S/l culture and 3-4 mg SH3-3-L/l culture.

3.2.1 CD2AP SH3-1 and SH3-2

Residues corresponding to SH3-1 (aa 1-60) and SH3-2 (aa 109-168) domains were cloned into pGEX6-P1 vectors (Dr Kathrin Kirsch *et al*., Boston University, USA). The vectors were expressed in *E. coli* and the integrity and purity of the protein were tested at various stages of the purification process (Figures 3.2, 3.3). A Prescission^{TM} protease cleavage test was performed with GST-SH3-1 so that the optimum amount of Prescission^{TM} protease can be calculated for a complete cleavage overnight at 4 °C. This ratio was found to be 1:250 Prescission^{TM} protease to GST-fusion protein and it was adopted for the GST cleavage of the other constructs too.

Figure 3.2 depicts how the purity of GST-SH3-1 and SH3-1 improves through the different purification steps: GSH titration test (Figure 3.2A); elution with reduced glutathione (Figure 3.2B); GST-tag cleavage with Prescission^{TM} protease (Figure 3.2C); separation of the GST-tag and SH3-1 with size exclusion chromatography (Figures 3.2D and E). Figure 3.2E shows the integrity and purity of purified SH3-1. In figures 3.2B - 3.2E, a band of roughly 50 kDa is observed. This band could correspond to a GST dimer and was separated from SH3-1 by SEC. When GST-SH3-1 was purified, the GSH titration test and elution steps (Figures 3.2A, B) were followed. The fractions containing GST-SH3-1 were then pooled and dialysed against the final desired buffer. As shown in Figure 3.2, the bands corresponding to GST-SH3-1 and SH3-1 have the expected molecular weights of 33 kDa and 7.5 kDa respectively.

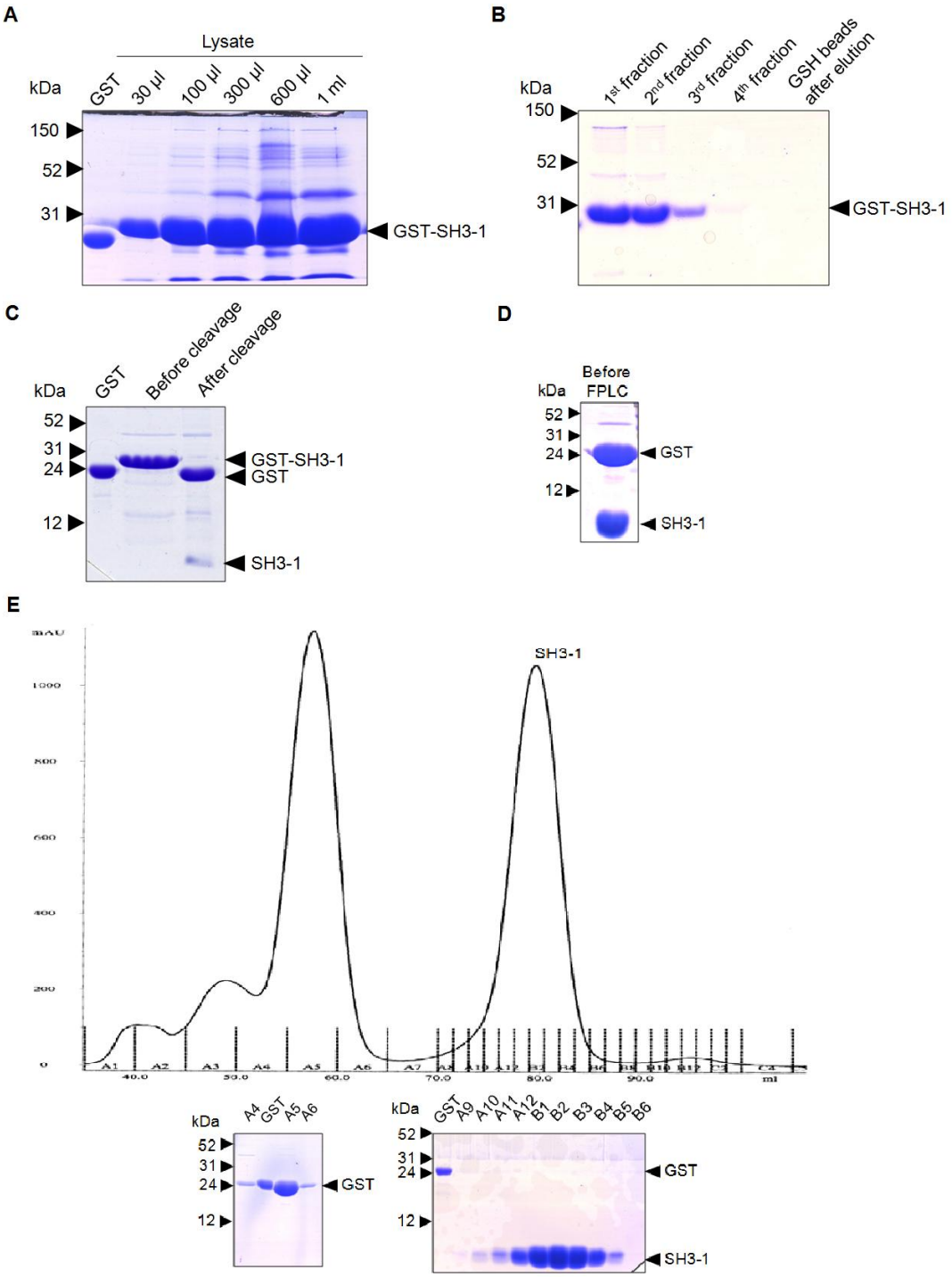


Figure 3.2

CD2AP SH3-1 purification.

All samples were resolved by SDS-PAGE and the gels were stained with Coomassie blue. Arrows on the left of each gel show the position of the markers. ***A*****,** GSH bead titration test was performed by incubating specified amounts of supernatant of bacterial lysate with 15 µl GSH beads (12% resolving gel). ***B*****,** GST-SH3-1 was eluted from the GSH beads and collected into fractions with reduced glutathione (12% resolving gel). ***C*****,** The GST-tag was cleaved off from the SH3-1 domain by overnight incubation with Prescission™ protease (16% resolving gel). ***D and E*****,** The GST-tag and SH3-1 were separated by SEC (16% resolving gel). The integrity and purity of the cleaved protein was verified by SDS-PAGE.

Similarly, Figure 3.3 depicts the purification steps of GST-SH3-2 and the SH3-2: GSH titration test (Figure 3.3A); elution with reduced glutathione (Figure 3.3B); GST-tag cleavage with Prescission^{TM} protease (Figure 3.3C); separation of the GST-tag and SH3-2 with SEC (Figure 3.3D). Figure 3.3D shows also the integrity and purity of final SH3-2. When GST-SH3-2 was purified, the same process was followed as described for GST-SH3-1. As shown in Figure 3.3, the bands corresponding to GST-SH3-2 and SH3-2 have the expected molecular weights (33 kDa and 7.4 kDa respectively).

3.2.2 Two different constructs of CD2AP SH3-3

Two different constructs of the third domain were tested following a recent paper on the SH3-3 domain of CIN85, which is 54% identical to the CD2AP SH3-3 domain (Philippe *et al*. 2011). The authors suggest that a construct with extended N- and C-terminal boundaries of CIN85 SH3-3 increases solubility and thermostability. According to their NMR structure (PDB code: 2K9G), this is due to a 90^{o} bend of the C-terminus over the N-terminus. Residues corresponding to SH3-3-S (aa 270-331) and SH3-3-L (aa 263-334) domains were cloned into pGEX6-P1 and pGEX-KN vectors respectively. The boundaries of the CD2AP SH3-3-L construct were chosen according to the boundaries of the CIN85 SH3-3 domain construct (Philippe *et al*. 2011). The pGEX-KN vector was used to minimise any amino acid overhang.

The vectors were expressed in *E. coli* and the integrity and purity of the protein was tested at various stages of the purification process (Figures 3.4 and 3.5). Both constructs were soluble, which suggests that the boundaries of the CD2AP SH3-3 constructs do not affect its solubility in a major way.

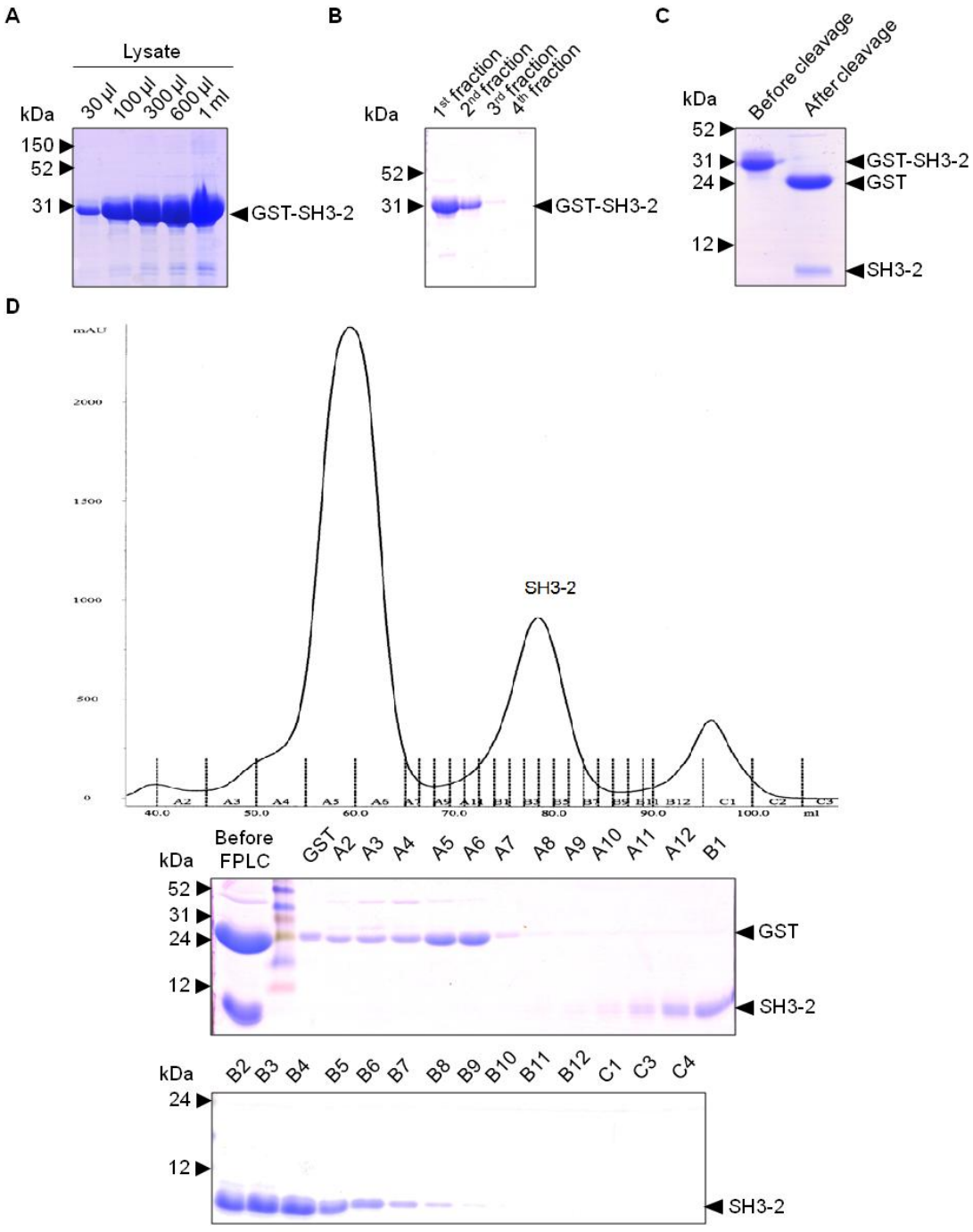


Figure 3.3

Purification of CD2AP SH3-2.

The same steps described in Figure 3.2 were performed. All samples were resolved by SDS-PAGE and the gels were stained with Coomassie blue. Arrows on the left of each gel show the position of the markers. ***A***, GSH bead titration test (12% resolving gel). ***B***, GST-SH3-2 elution with reduced glutathione (12% resolving gel). ***C***, GST-tag cleavage with Prescission™ protease (16% resolving gel). ***D***, The GST-tag and SH3-2 were separated by SEC (16% resolving gel) The integrity and purity of the cleaved protein was verified by SDS-PAGE.

Similar purification steps compared to SH3-1 and SH3-2 domains were followed for SH3-3-S (Figure 3.4) by incubating GST-SH3-3-S with Prescission™ protease to cleave off the GST-tag overnight at 4 °C. The expected molecular weights of GST-SH3-3-S and SH3-3-S are 33 and 7.5 kDa respectively and are in accordance with the migration positions of the visualised bands.

The purification steps of GST-SH3-3-L do not differ from the ones already described (Figures 3.5A-B). Its corresponding protein band migrates at its expected molecular weight (34 kDa). Thrombin protease was used to cleave off the GST-tag from SH3-3-L (domain cloned into pGEX-KN). As shown in Figure 3.5C, the GST-tag cleavage was still incomplete after 48 h of incubation with thrombin protease (1:100 ratio of thrombin to GST-fusion protein in µg). This process was optimised and a complete GST-tag cleavage was achieved after 60 h with the protease at 4 °C at the above-mentioned ratio of thrombin to GST-fusion protein (data not shown). The expected molecular weight of SH3-3-L is 8.2 kDa, which corresponds to the band below the 12 kDa marker (Figure 3.5D).

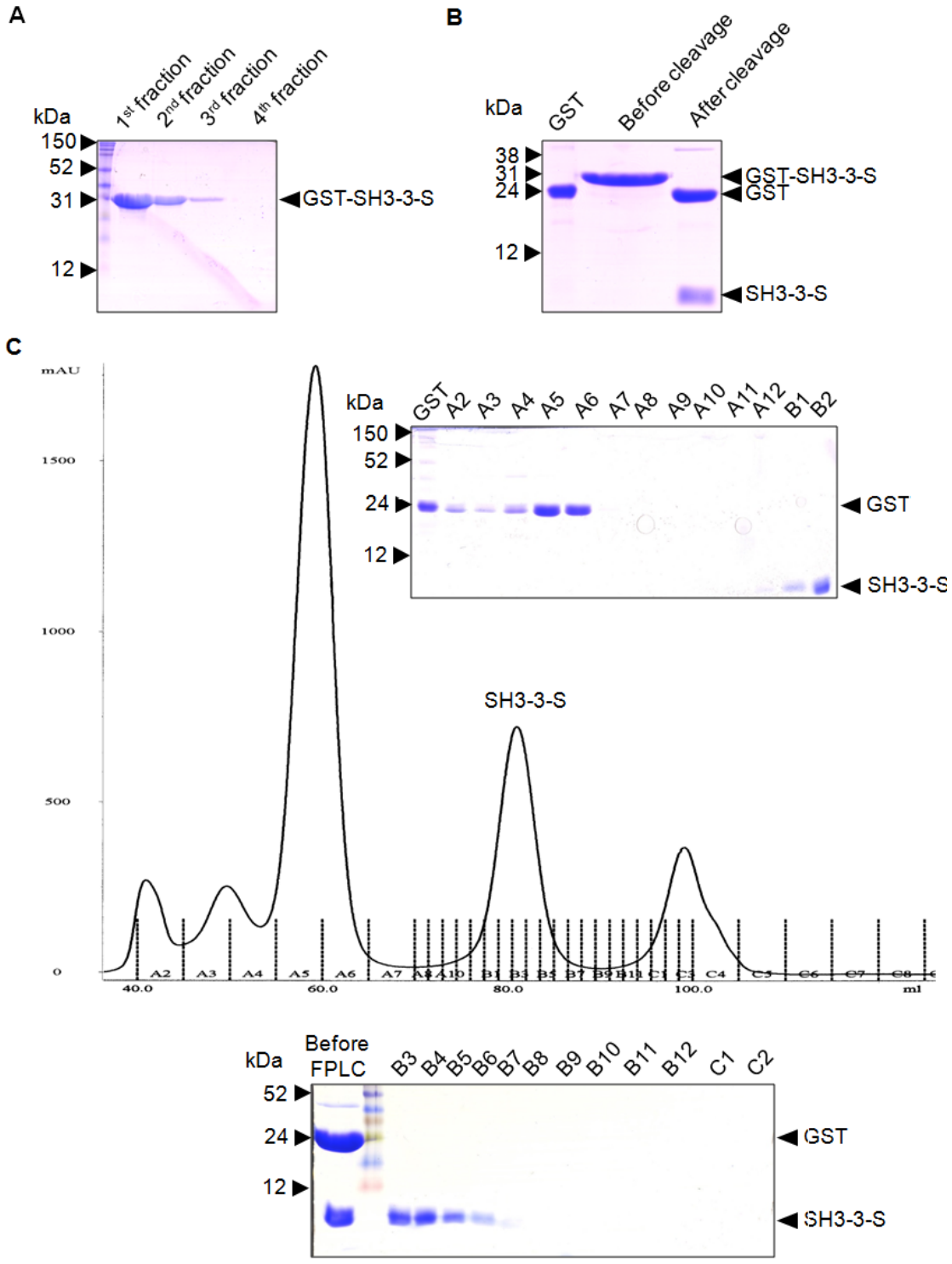


Figure 3.4

Purification of CD2AP SH3-3-S.

The same steps described in Figure 3.2 were performed. All samples were resolved by SDS-PAGE and the gels were stained with Coomassie blue. Arrows on the left of each gel show the position of the markers. ***A***, GST-SH3-3-S elution with reduced glutathione (12% resolving gel). ***B***, GST-tag cleavage with Prescission™ protease (16% resolving gel). ***C***, The GST-tag and SH3-3-S were separated by SEC (16% resolving gel). The integrity and purity of the cleaved protein was verified by SDS-PAGE.

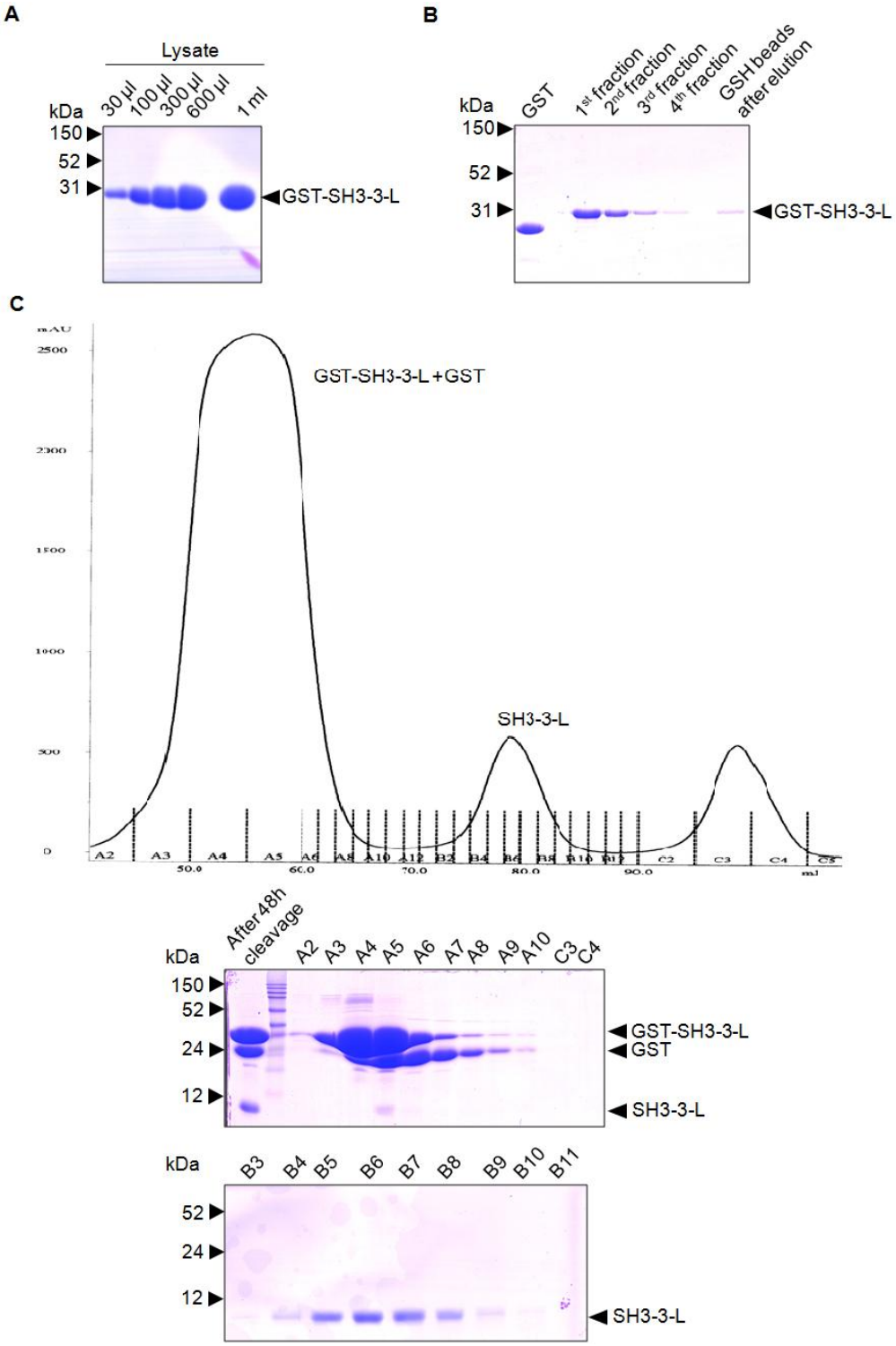


Figure 3.5

Purification of CD2AP SH3-3-L.

All samples were resolved by SDS-PAGE and the gels were stained with Coomassie blue. Arrows on the left of each gel show the position of the markers. ***A***, GSH bead titration test (12% resolving gel). ***B***, GST-SH3-3-L elution with reduced glutathione (12% resolving gel). ***C***, Incomplete GST-tag cleavage with thrombin protease after 48h at 4 ^{o}C (16% resolving gel). The GST-SH3-3-L, GST-tag and SH3-3-L were separated by SEC (16% resolving gel). The integrity and purity of the cleaved protein was verified by SDS-PAGE.

3.3 Mass Spectrometry of CD2AP SH3-1, -2 and -3-3-S

Electrospray Ionization Mass Spectrometry (ESI-MS) was performed to confirm that the purified SH3 domains (SH3-1, SH3-2 and SH3-3-S) had the expected mass (Georgina Berridge, Structural Genomics Consortium Oxford, UK). Figure 3.6 depicts the graph of the counts versus deconvoluted mass of the ESI scans for the SH3 domains. A single peak was observed for each SH3 domain. Additionally, small peaks were present, whose distance from the main peak was similar for each domain. These peaks might represent salt ions. The theoretical molecular weight calculated with the ProtParam tool (Gasteiger *et al*. 2005), and the experimentally obtained molecular weight were almost identical (Table 3.1).

Table 3.1 Accuracy of the molecular weight (MW) data for the CD2AP SH3 domains. MW_{actual}, molecular weight determined by the ESI-MS experiment; $MW_{theoretical}$, molecular weight computed by the ProtParam tool (Gasteiger *et al*. 2005).

CD2AP SH3 domains	MW_{actual}	$MW_{theoretical}$
SH3-1	7584.67	7583.5
SH3-2	7438.28	7437.4
SH3-3-S	7507.24	7506.23

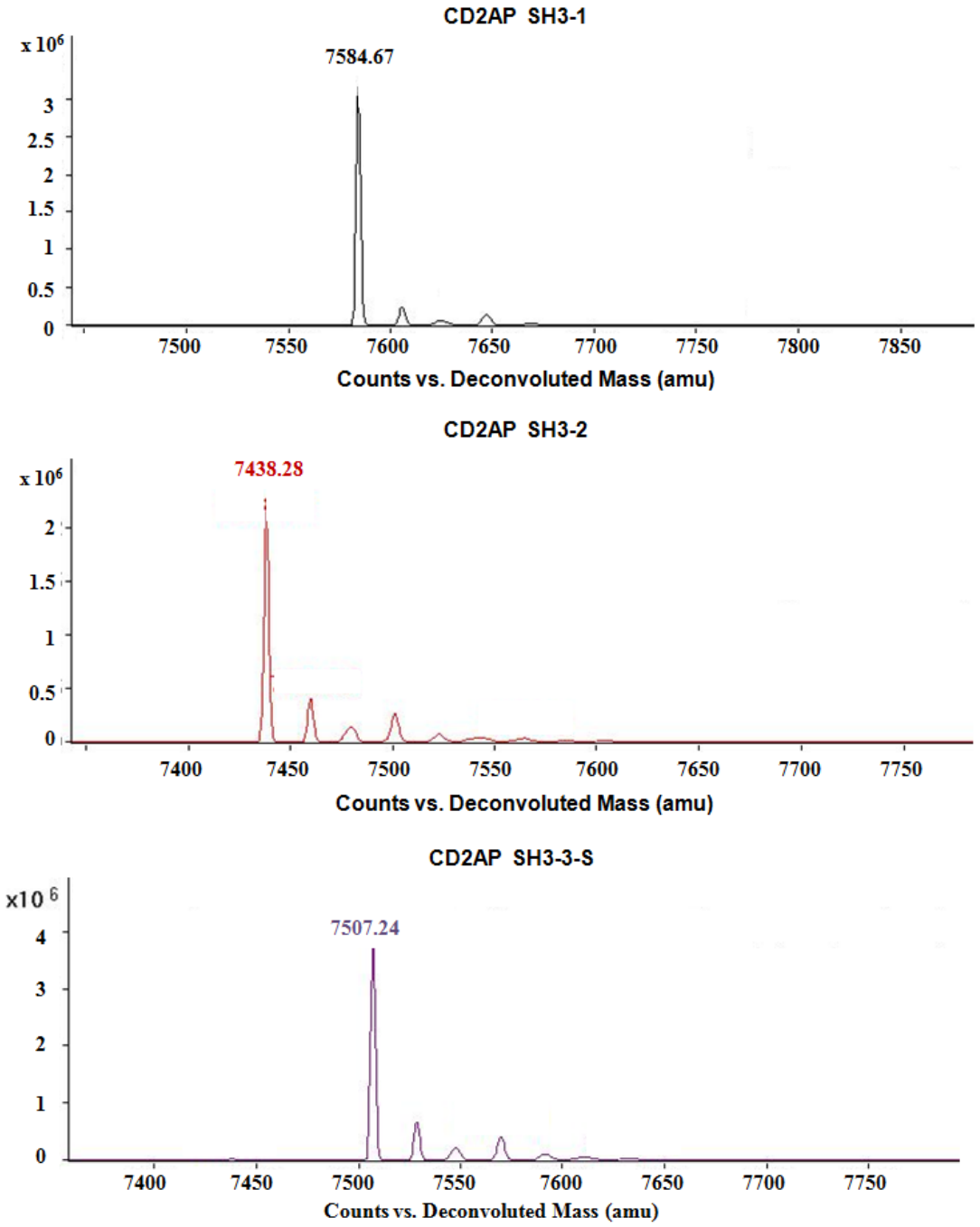


Figure 3.6

ESI-MS data for the CD2AP SH3 domains (SH3-1, SH3-2, SH3-3-S).

3.4 Analysis of the CD2AP SH3 interaction with RIN3 epitopes using peptide arrays

Peptide spot arrays are a useful tool for the characterization of the short linear motifs, recognised by protein interaction domains (Volkmer *et al*. 2012). They allow to understand better the binding mode and specificity of the CD2AP SH3 domains. Therefore, synthesised RIN3 permutation arrays of both RIN3 epitopes were probed with the SH3 domains (SH3-1, SH3-2 and SH3-3-S). These were kindly provided by Dr Rudolf Volkmer (Charite, Institute of Medical Immunology, Berlin, Germany). In the peptide permutation arrays, each residue of each RIN3 epitope was mutated against all other naturally occurring amino acids. Therefore, the arrays inform on the preferred as well as the negatively selected residues in specific positions. We found that the choice of the appropriate blocking buffer is critical to avoid non-specific binding. Different buffer conditions were tested and 2% ovalbumin in RIPA 100 was determined to be the best (Appendix C.3). Initially, the arrays were probed with GST. No background signal was detected (data not shown).

When the RIN3e1 peptide permutation array was probed with SH3-3, the substitutions of P_{382}, A_{384} and R_{387} to Cys appeared to partially preserve the interaction (data not shown). This result could have been an artifact due to disulphide bond formation between the cysteine-rich GST-tag and the peptides. Therefore, the membranes were re-probed with the three SH3 domains in the presence of the reducing agent DTT. In this case, the previously observed spots due to Cys substitutions did not appear. Therefore, the arrays re-probed in the presence of DTT are depicted (Figure 3.7).

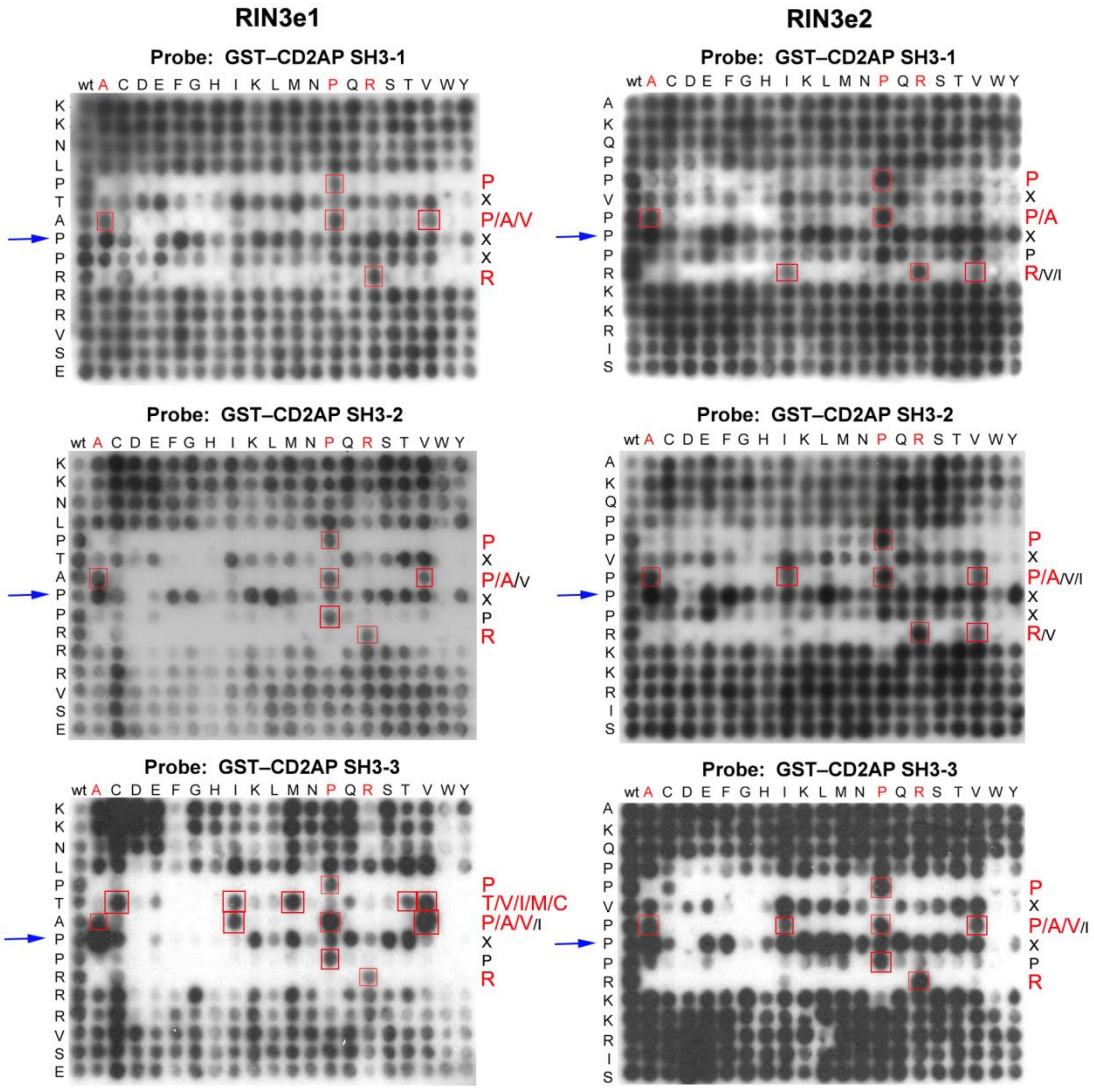


Figure 3.7

All three CD2AP SH3 domains recognise the atypical P-x-P/A-x-p-R motif in RIN3.

The two RIN3 binding regions were spot-synthesized as 17 aa peptides (AKKNLPTAPPRRRVSER and TAKQPPVPPPRKKRISR) and permutated in 15 single positions (underlined) against all 20 natural residues. Three array copies were initially probed with GST and then with GST-SH3-1, -SH3-2 or -SH3-3-S. The positions permutated within the peptides are indicated on the left side of each array, and the individual substitutions are listed across the top. Peptide spots with substantially restricted residues are boxed in red. The amino acid preferences at these positions are listed on the right side of each array (red font: most strongly binding amino acids; smaller black font: tolerated amino acids). Blue arrows indicate the wild-type residues, whose substitution to Ala appears to result in tighter binding.

For all three SH3 domains, Pro and Arg are the single residues in the first and last positions of the motif in both RIN3 epitopes that give strong binding

signals (Figure 3.7); all others prevent or weaken binding. Thus, these two residues are essential for the interaction. SH3-1 recognises the P-x-A-x-x-R motif in RIN3e1 (Figure 3.7, upper left panel). In motif position 3, Pro and Val give similar binding levels with SH3-1 to native Ala in RIN3e1. SH3-1 recognises the P-x-P-x-p-R motif in RIN3e2 (Figure 3.7, upper right panel). In position 3, a Pro_{459} to Ala substitution preserved the interaction.

SH3-2 recognises the P-x-A-x-p-R motif in RIN3e1. In position 3, Pro gives similar binding levels to native Ala in RIN3e1. SH3-2 recognises the P-x-P-x-x-R motif in RIN3e2. In position 3, a Pro_{459} to Ala substitution does not affect signal strength. Motif position 5 is not as selective in RIN3e2 compared to RIN3e1 for SH3-2 binding, as any residue can be accommodated in this position.

According to the RIN3e1 permutation array, SH3-3 recognises a P-t-A-x-p-R motif in RIN3e1 (Figure 3.7, bottom left panel). In motif position 3, the same tolerance as for SH3-1 is present (Ala, Pro, Val). In position 5, the same tolerance as for SH3-2 is present (Pro). Unlike the first two SH3 domains, there is an increased SH3-3 selectivity for other residues in position 2. A T_{383} to Val, Ile, Met or Cys substitution preserved the interaction with SH3-3, but any other residue greatly affected the signal strength. SH3-3 recognises a P-x-P-x-p-R motif in RIN3e2 (Figure 3.7, bottom right panel). In motif position 3, Ala and Val can replace Pro without a reduction in signal. Unlike the SH3-3 selectivity in position 2 of RIN3e1, any residue can replace the native Val in RIN3e2. As a speculative note, hypothetical motif variants based on P-T/V/I/M/C-P/A/V-x-P-R, should they exist in proteins, may also allow SH3-3 binding.

Furthermore, it appears that $P_{385}A$ (RIN3e1) and $P_{460}A$ (RIN3e2) lead to tighter binding between the SH3 domains and the two RIN3 epitopes (Figure 3.7, blue arrow). This is particularly obvious for the SH3-3 domain.

As mentioned above, Arg is the only residue in the last position of the motif that allows an interaction between the three SH3 domains and the two RIN3 epitopes. However, there may be a Arg_{462} to Val or Ile substitution, which reduces the affinity of the interaction with the first two SH3 domains.

3.5 ITC analysis of the three CD2AP SH3 domains with one ALIX and two RIN3 epitopes

The generation of high purity protein material enabled ITC measurements with the SH3 domains to investigate binding of the ALIX- and RIN3-derived peptides. As described in section 2.2.5, ITC informs on the affinity, stoichiometry (N) and thermodynamic signature of an interaction (Doyle 1997).

In section 3.5.1, the ITC data on affinity and N will be discussed. Table 3.2 shows the dissociation constant and stoichiometry values. Examples of ITC curves are found in Figure A.1 (Appendix A). In section 3.5.2, the thermodynamic parameters (changes in free energy ΔG, enthalpy ΔH and entropy ΔS) will be presented.

3.5.1 Affinity and stoichiometry

3.5.1.1 Wild-type ALIX and RIN3 peptides

Initially ITC measurements were performed with the SH3 domains and an ALIX peptide (Table 3.2, first row). The first two SH3 domains exhibited

similar affinities in the low micromolar range while SH3-3 bound 5 to 7-fold more weakly. Next, 16 residue RIN3e1 and RIN3e2 peptides were tested with the SH3 domains. As shown in Table 3.2 (second and third rows), the first two SH3 domains bind ca. 4 times more strongly to RIN3e1 compared to SH3-3-S and SH3-3-L. SH3-1, SH3-2 and SH3-3-S binds approximately 2-fold, 5-fold and 6 to 10-fold stronger to RIN3e2 than to RIN3e1. No major difference in the affinity of SH3-3-S and SH3-3-L for the RIN3 epitopes was observed. These results show that the SH3 domains bind to ALIX- and RIN3-derived peptides in the low micromolar range, and the first two SH3 domains bind tighter to ALIX and RIN3e1 compared to the SH3-3.

Subsequently, the SH3 domain binding to a truncated 11 residue RIN3e2 peptide with the already described P-x-P-x-p-R motif present was tested. The first two SH3 domains bind the truncated peptide 2.5- and 7-fold weaker, respectively, while the peptide binds 9-fold more weakly to the third SH3 domain. This result is not in agreement with the peptide permutation data, which showed that point mutations of the N- and C-terminal RIN3e2 residues against all other amino acids did not affect binding to the SH3 domains (Figure 3.7). However, It could be that flanking residues may be needed to stabilise a conformation, which allows RIN3e2 binding to the SH3 domains.

The stoichiometry values for the SH3s were usually close to 1 with the 16 residue and truncated RIN3e2. The N values for SH3-1 and SH3-3-L reached 1 with RIN3e1, and were close to 1 for the first two SH3 domains and the ALIX epitope.

Table 3.2 Isothermal titration calorimetric measurements of CD2AP SH3 domain interactions with ALIX- and RIN3-derived peptides: Dissociation constant and stoichiometry values. The critical mutations, which lead to at least 8-fold reduction in affinity compared to the wild-type peptide, are coloured red. SE = standard error; ND= not done; TLQ = affinity too low for accurate quantification by ITC; -am indicates C-terminal amidation.

Protein epitope	Peptide sequence	Length (a.a.)	K_d ± SE (μM)				N ± SE			
			SH3-1	SH3-2	SH3-3-S	SH3-3-L	SH3-1	SH3-2	SH3-3-S	SH3-3-L
ALIX	AGGHAPTPPTPAPRTMPPTKP-am	21	6.6 ± 0.002	4.2 ± 0.001	32.2 ± 1.9	ND	1.09 ± 0.01	0.98 ± 0.004	1.42 ±0.03	ND
RIN3e1	AKKNLPTAPPRRRVSE-am	16	8.9 ± 0.2	10 ± 0.2	39.8 ± 8.3	47.4 ± 3.3	0.85 ± 0.03	0.64 ± 0.02	0.69 ± 0.06	1.18 ± 0.05
RIN3e2	TAKQPPVPPPRKKRIS-am	16	4.6 ± 0.4	2.1 ± 0.8	6.7 ± 0.2	4.6 ± 0.6	1.18 ± 0.01	0.94 ± 0.03	0.84 ± 0.02	1.16 ± 0.01
RIN3e2	KQPPVPPPRKK-am	11	10.4 ± 1.3	15.1 ± 0.9	57.5 ± 6.5	ND	1.04 ± 0.03	1.38 ± 0.02	1.27 ± 0.05	ND
RIN3e2 $P_{457}A$	TAKQPAVPPPRKKRIS-am	16	126 ± 1.7	19.9 ± 2.6	68 ± 0.24	35.3 ± 2.5	0.797 ± 0.32	1.46 ± 0.04	0.98 ± 0.3	1.3 ± 0.04
RIN3e2 $P_{459}A$	TAKQPPVAPPRKKRIS-am	16	2.2 ± 0.3	7.8 ± 2	6.9 ± 0.02	7.2 ± 0.7	1.43 ± 0.02	1.5 ± 0.06	0.68 ± 0.04	1.57 ± 0.02
RIN3e2 $P_{460}A$	TAKQPPVPAPRKKRIS-am	16	2 ± 0.3	0.7 ± 0.07	3.9 ± 0.3	2.2 ± 0.2	1.74 ± 0.02	1.4 ±0.01	0.79 ± 0.01	1.16 ± 0.01
RIN3e2 $R_{462}A$	TAKQPPVPPPAKKRIS-am	16	TLQ				TLQ			
RIN3e2 $R_{462}L$	TAKQPPVPPPLKKRIS-am	16								
RIN3e2 $R_{462}V$	TAKQPPVPPPVKKRIS-am	16	TLQ	47.1 ± 0.05	TLQ					
RIN3e2 $R_{462}I$	TAKQPPVPPPIKKRIS-am	16		42.7 ± 0.05						
RIN3e2 $I_{466}A$	TAKQPPVPPPRKKRAS-am	16	1.9 ± 0.1	1.6 ± 0.2	11.8 ± 2	4 ± 0.5	1.35 ± 0.01	1.32 ± 0.01	1.07 ± 0.04	1.08 ± 0.02

These experimentally observed stoichiometry values indicate that a number of issues need to be considered when analyzing the interactions between the SH3 domains and these synthetic RIN3 epitopes. As was described in section 1.4, the binding equilibrium may involve a conformation manifold (i.e. multiple conformations) of the peptides. Therefore, only a subset of the peptides may employ the PPII conformation to enable binding at any given time at the hydrophobic grooves and acidic pocket of the SH3 domains (Stollar *et al*. 2012, Krieger *et al*. 2014). Another possibility would be that the protein and/or ligand concentration were incorrect. Measuring the peptide concentration correctly is important for the calculation of N, ΔH and association constant K_A, while the accuracy of the cell concentration measurement is critical for the calculation of N (Myszka *et al* 2003). As described in sections 2.2.1.5, 2.2.1.6 and 2.2.1.9, the SH3 domain concentration was calculated by measuring the absorbance at 280nm. These measurements were repeated and the results were consistent. Unfortunately, the peptide concentrations could not to be calculated with this method because they lacked aromatic residues.

Stoichiometry can also be ascertained for the crystallography experiments. As described below (sections 3.6.2 and 3.6.3), two crystal structures of SH3-1/RIN3e1 and SH3-2/RIN3e2 were solved. These structures showed 1:1 binding of peptide to the SH3 domains.

Lastly, the N value of the SH3-2 complex with RIN3e1 was further investigated by performing dynamic light scattering (DLS). No robust results were obtained due to high background noise (data not shown). Also, size exclusion chromatography and multi-angle static light scattering (SEC-MALS) experiments were attempted, but complex dissociation was observed during

the course of the SEC run, preventing interpretation (Dr Yin Dong & Anna-Maria Tessitore, Structural Genomics Consortium Oxford, UK; data not shown). Therefore, these experimental methods were not pursued further.

3.5.1.2 RIN3e2-derived mutant peptides

According to the peptide permutation array analysis, the SH3 domains recognise three critical residues in RIN3e1 (P_{382}, A_{384}, R_{387}) and RIN3e2 (P_{457}, A_{459} and R_{462}) (section 3.4). However, further analysis should be performed to verify the importance of these residues. Therefore, ITC experiments using the SH3 domains (SH3-1, SH3-2, SH3-3-S, SH3-3-L) and point mutant 16 residue RIN3e2 peptides were conducted. The above-mentioned key RIN3e2 residues (i.e. P_{457}, P_{459}, R_{462}) were point-mutated to Ala, and R_{462} was mutated additionally to Leu, Val and Ile. An R_{462}L mutant peptide was also included to determine if the hydrophobic part of the R_{462} side chain contributes to binding. Furthermore, R_{462}V and R_{462}I peptides were tested, because the peptide permutation arrays showed that these mutations might weaken binding with the SH3 domains. Lastly, a P_{460}A peptide was tested, because the RIN3 permutation arrays suggested that this mutation may lead to increased affinity to the SH3 domains (section 3.4). Table 3.2 (5^{th} to 11^{th} row) summarizes the dissociation constant and stoichiometry values.

As shown in Table 3.2, the R_{462} point mutations usually abolished the interaction with the three SH3 domains, except for a 20-fold reduction in affinity of the SH3-2 interaction with R_{462}V and R_{462}I mutant peptides. Compared to the wild-type RIN3e2, P_{457}A weakened the interaction with SH3-1 ca. 30-fold and with the SH3-2 and SH3-3 domains 10-fold. Even though P_{459}A did not affect the SH3-1 and SH3-3 binding significantly, it

caused a 3.5-fold reduction in affinity to SH3-2 compared to wild-type. The latter is comparable to the affinity of SH3-2 for RIN3e1. Therefore, P_{457} and R_{462} are the most critical residues for the SH3 interaction with RIN3e2. Also, P_{459}A preserves the interaction between the SH3 domains and RIN3 epitopes. These results confirm that the SH3 domains recognise the P-x-P/A-x-p-R motif in RIN3.

When P_{460} was mutated to Ala, the SH3 domains bound 2-fold more strongly to the peptide. Thus, P_{460}A does not lead to significantly tighter binding to the SH3 domains as the RIN3 permutation arrays suggested (section 3.4).

Lastly, the SH3-1 stoichiometry values in complex with the mutated RIN3e2 peptides reach 1.5, except the N value of 1 with the P_{459}A peptide. The N values of SH3-2, SH3-3-S and SH3-3-L domains with the same peptides range between 1.3-1.5, 0.7-1.1 and 1.1-1.6 respectively. Therefore, the SH3 domains form largely binary complexes with the RIN3e2-derived mutant peptides.

3.5.2 Thermodynamics

As mentioned in sections 1.5 and 2.2.5, the affinity is quantified by the change in free energy (ΔG) upon complex formation. ΔG is associated with changes in enthalpy (ΔH) and entropy (ΔS) ($\Delta G = \Delta H - T\Delta S$). Therefore, ΔH and ΔS provide additional information on binding events. ΔH is relative to the net change in the number of formed and broken noncovalent bonds between the uncomplexed and complexed states. For instance, increased hydrogen bond formation will result in a favourable (i.e. negative) ΔH. ΔS informs on the change in the degrees of freedom of the system. For example, hydrophobic

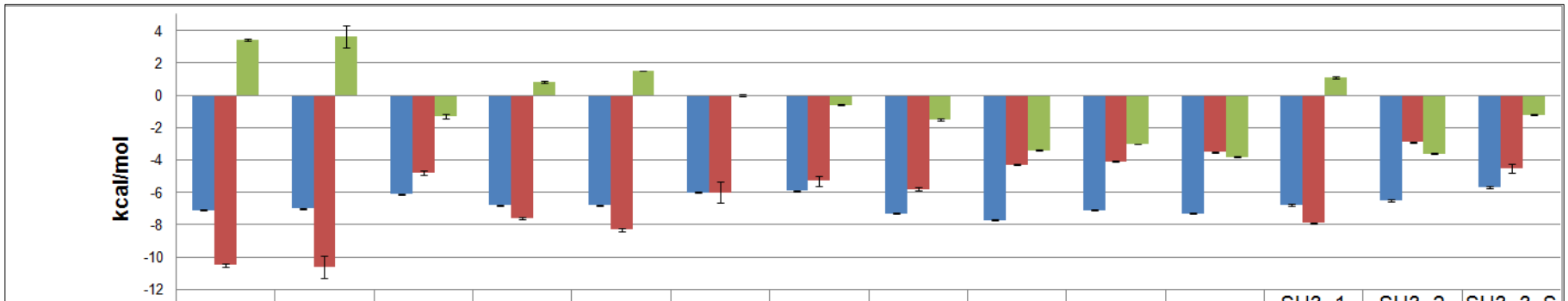
contacts due to complex formation leads to the displacement of water molecules from the binding interface. This would result in favourable (i.e. negative -TΔS) changes in entropy (Doyle 1997, Ladbury 2010).

3.5.2.1 Wild-type peptides

Figure 3.8 depicts the thermodynamic parameters (changes in free energy ΔG, enthalpy ΔH and entropy ΔS) associated with the interaction between the SH3 domains and the ALIX and RIN3 peptides. The values are found in the bottom panel, while their graphic representation is found in the upper panel of Figure 3.8. As described before, negative ΔH and -TΔS values describe favourable changes.

The first two SH3 domains behave similarly upon binding to ALIX and RIN3e1. Dominant favourable ΔH is observed, which suggests that polar contacts drive the interaction. Small unfavourable ΔS is also observed. The SH3-3-S binding to ALIX and RIN3e1 is characterised by major favourable enthalpic changes. Small favourable entropic changes are observed upon SH3-3-S/ALIX complex formation. This implies that even though polar interactions drive the SH3-3 interaction with ALIX and RIN3e1, hydrophobic interactions and/or displacement of solvent molecules from the binding interface might contribute partially upon SH3-3 binding to ALIX.

The SH3-2 and SH3-3-S binding to RIN3e2 is characterised by similar favourable ΔH and ΔS. This suggests that polar and hydrophobic interactions contribute to their interaction with RIN3e2. The favourable entropy could also be attributed to the liberation of solvent molecules from the SH3 surface upon docking of RIN3e2.



	SH3_1 +ALIX	SH3_2 +ALIX	SH3_3_S +ALIX	SH3_1 +RIN3e1	SH3_2 +RIN3e1	SH3_3_S +RIN3e1	SH3_3_L +RIN3e1	SH3_1 +RIN3e2	SH3_2 +RIN3e2	SH3_3_S +RIN3e2	SH3_3_L +RIN3e2	SH3_1 +RIN3e2 truncated	SH3_2 +RIN3e2 truncated	SH3_3_S +RIN3e2 truncated
ΔG (kcal/mol)	-7.1	-7	-6.1	-6.8	-6.8	-6.003	-5.9	-7.3	-7.7	-7.1	-7.3	-6.8	-6.5	-5.7
ΔH (kcal/mol)	-10.5	-10.6	-4.8	-7.6	-8.3	-6	-5.3	-5.8	-4.3	-4.1	-3.5	-7.9	-2.9	-4.5
-TΔS (kcal/mol)	3.4	3.6	-1.3	0.8	1.5	-0.003	-0.6	-1.5	-3.4	-3	-3.8	1.1	-3.6	-1.2

Figure 3.8

Thermodynamic parameters of binding of the ALIX and RIN3 epitopes to the CD2AP SH3 domains.

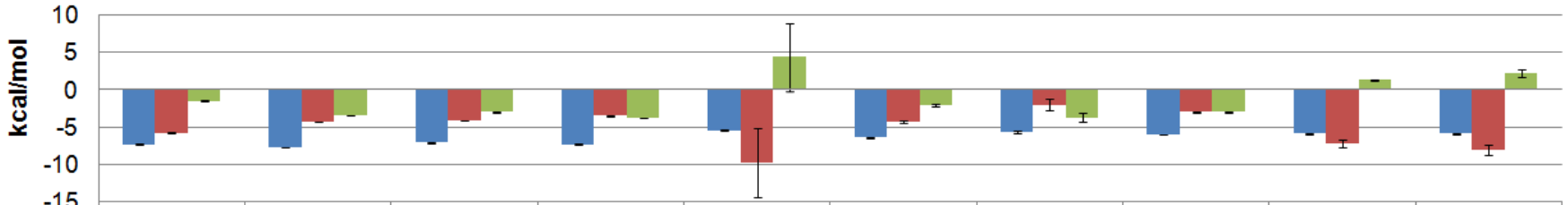
RIN3e2 truncated refers to the 11 aa peptide.

The SH3-1 binding to RIN3e2 leads to major favourable ΔH and small favourable ΔS. Based on the ΔS values, hydrophobic interactions contribute less to the SH3-1/RIN3e2 compared to SH3-2/RIN3e2 and SH3-3/RIN3e2 interactions.

Moreover, SH3-3-S and SH3-3-L behave in a similar way, because they bind to the RIN3 epitopes with similar affinities and exhibit a similar thermodynamic signature.

The truncated form of RIN3e2 includes the SH3 recognition motif, which was found by the peptide permutation arrays (section 3.4). We would expect that the truncated RIN3e2 would exhibit higher ΔH compared to the 16 residue RIN3e2 peptide due to higher solvent accessibility. Also, more unfavourable (i.e. positive) or less favourable (i.e. negative) ΔS compared to wild-type would be observed due to lower degrees of freedom upon binding. As shown in Figure 3.8, this was observed for SH3-1 and SH3-3-S. However, SH3-2 binding in the truncated form of RIN3e2 leads to decreased ΔH and unchanged ΔS. This indicates a reduction in polar interactions and no changes in the flexibility of the system.

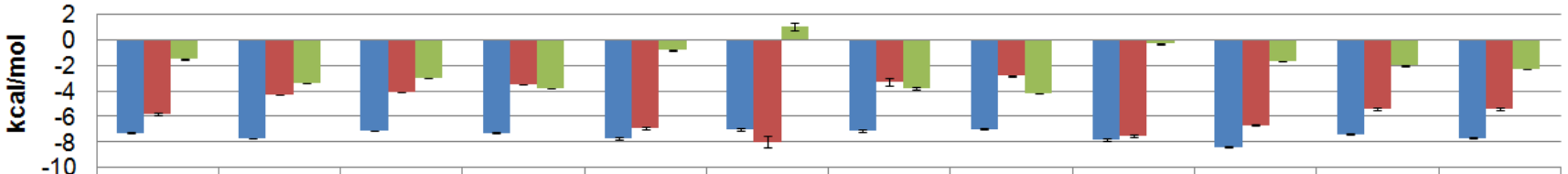
Obtaining a structural view through protein crystallography may help to understand these values better and they will be further discussed in the section 3.6.4.3.



	SH3_1 +RIN3e2	SH3_2 +RIN3e2	SH3_3_S +RIN3e2	SH3_3_L +RIN3e2	SH3_1 + P457A	SH3_2 + P457A	SH3_3_S + P457A	SH3_3_L + P457A	SH3_2 + R462V	SH3_2 + R462I
ΔG (kcal/mol)	-7.3	-7.7	-7.1	-7.3	-5.4	-6.4	-5.7	-6	-5.9	-5.9
ΔH (kcal/mol)	-5.8	-4.3	-4.1	-3.5	-9.8	-4.3	-2	-3	-7.2	-8.1
-TΔS (kcal/mol)	-1.5	-3.4	-3	-3.8	4.4	-2.1	-3.7	-3	1.3	2.2

Figure 3.9

Thermodynamic parameters of binding of the RIN3e2-derived P457A, R462V and R462I peptides to the CD2AP SH3 domains.



	SH3_1 +RIN3e2	SH3_2 +RIN3e2	SH3_3_S +RIN3e2	SH3_3_L +RIN3e2	SH3_1 + P459A	SH3_2 + P459A	SH3_3_S + P459A	SH3_3_L + P459A	SH3_1 + P460A	SH3_2 + P460A	SH3_3_S + P460A	SH3_3_L + P460A
ΔG (kcal/mol)	-7.3	-7.7	-7.1	-7.3	-7.7	-7	-7.1	-7	-7.8	-8.4	-7.4	-7.7
ΔH (kcal/mol)	-5.8	-4.3	-4.1	-3.5	-6.9	-8	-3.3	-2.8	-7.5	-6.7	-5.4	-5.4
-TΔS (kcal/mol)	-1.5	-3.4	-3	-3.8	-0.8	1	-3.8	-4.2	-0.3	-1.7	-2	-2.3

Figure 3.10

Thermodynamic parameters upon binding of the RIN3e2-derived P459A and P460A peptides to the CD2AP SH3 domains.

3.5.2.2 RIN3e2 mutant peptides

The thermodynamic parameters of the $P_{457}A$, $R_{462}V$ and $R_{462}I$ interactions with the SH3 domains are depicted in Figure 3.9, while the parameters for $P_{459}A$ and $P_{460}A$ are shown in Figure 3.10.

For the peptides, whose affinity was reduced compared to wild-type peptide (i.e. SH3 binding to $P_{457}A$, SH3-2 binding to $P_{459}A$, $R_{462}V$ and $R_{462}I$), it would be expected that the lost contact points would cause a reduction in ΔH, and the increase of the degrees of freedom of the system would lead to an increase in ΔS. This signature is only observed when SH3-3-S binds to $P_{457}A$. For SH3-1/$P_{457}A$, SH3-2/$P_{459}A$, SH3-2/$R_{462}V$ and SH3-2/$R_{462}I$, increased favourable ΔH is observed. ΔH might be increased due to an increase of solvent-mediated hydrogen and ionic bonds. For SH3-2/$P_{457}A$, ΔH remained unchanged, possibly due to rearrangement of the bonds. The reduction in affinity and hence ΔG was due to unfavourable ΔS for SH3-1/$P_{457}A$, SH3-2/$R_{462}V$ and SH3-2/$R_{462}I$, and reduced favourable ΔS for SH3-2/$P_{457}A$. As mentioned in section 1.4, less favourable or unfavourable ΔS indicates loss of flexibility of the system. This could be due to the restraining of the peptide in certain conformations, such as PPII. Similarly, the above mentioned ΔS values (i.e. SH3-1/$P_{457}A$, SH3-2/$P_{457}A$, SH3-2/$R_{462}V$ and SH3-2/$R_{462}I$) suggest that the mutant peptides adopted restrained conformations. In this case, the adopted conformations of the mutant peptides did not enable as tight binding compared to the wild-type peptides.

The opposite effects are expected upon an increase in affinity due to the peptide docking onto the SH3 domains i.e. increase in ΔH and decrease in ΔS. This is observed for all domains binding to $P_{460}A$, and for SH3-1 binding

to $P_{459}A$. Even though the affinity and hence the ΔG of SH3-3-S for $P_{459}A$ did not change, ΔH reduced and ΔS increased, which indicates more hydrophobic and less polar interactions between SH3-3-S and $P_{459}A$. A reduction in ΔH and an increase in ΔS was observed for SH3-3-L/$P_{459}A$, which resulted in a slight reduction in the free energy (ΔΔG = 0.3 kcal/mol). This indicates that there are small differences in the recognition of $P_{459}A$ by SH3-3-S and SH3-3-L. However, this does not affect the binding affinity of the interaction. The thermodynamic parameters of SH3-3-L compared to SH3-3-S were similar for the rest peptides.

In summary, the ITC data of the mutant RIN3e2 peptides are consistent with the peptide permutation array results. The three SH3 domains recognise the P-x-P/A-x-p-R motif in the two RIN3 epitopes.

Having determined the binding affinity between the SH3 domains and the ALIX and RIN3 epitopes, crystallisations of the SH3/peptide complexes were attempted. In preparation for this, the saturation point (i.e. ratio of SH3 domain to peptide, which is sufficient for the saturation of SH3 binding sites) was identified by ITC. The ratios of SH3 domains and peptides are found in Table 3.3.

Table 3.3 SH3 domain/RIN3 epitope ratio to ensure that all binding sites are saturated.

Protein epitopes	Ratio		
	SH3-1	SH3-2	SH3-3-S
ALIX	1:3	1:3.5	1:5
RIN3e1	1:3.5	1:3.5	1:5
RIN3e2	1:3	1:3	1:3

3.6 Analysis of the CD2AP SH3/RIN3 interactions by protein X-ray crystallography

3.6.1 Crystallisation trials with CD2AP SH3 domain constructs and RIN3-derived peptides

A pre-crystallization test (PCT) was used to determine the appropriate protein concentration for the crystallization screens. The first two SH3 domains uncomplexed or in complex with the RIN3 epitopes were tested. The reagents and protein concentrations are found in section 2.2.7.1 and the results are summarized in Table 3.4. Colour annotation is used to show whether precipitates occurred. The production of precipitates with reagent A2 indicates that polyethylene glycol and salt (reagent A2) appear to be more appropriate chemicals compared to ammonium sulphate and alcohol (reagents A1 and A3). Also, these findings suggest that the tested concentrations of 15.5 to 36.7 mg/ml are a reasonable starting range for the crystallization screens.

Table 3.4 Pre-crystallization test.

Results for the first two SH3 domains alone or in complex with the two RIN3 epitopes.

Result for SH3-1			
Reagents	**no peptide**	**RIN3e1**	**RIN3e2**
A1	clear	clear	clear
A2	light	light	heavy
A3	clear	clear	phase separation
Result for SH3-2			
Reagents	**no peptide**	**RIN3e1**	**RIN3e2**
A1	light	clear	clear
A2	heavy	medium	medium
A3	clear	clear	clear
Drop annotation			
	clear		
	light granular precipitate		
	medium granular precipitate		
	heavy granular precipitate		

The aim of the crystallography experiments was to obtain, if possible, a high resolution structural view of each RIN3 and ALIX epitope in complex with one or more SH3 domains. Since it was not predictable which complexes would crystallise, most SH3/peptide combinations were tested, apart from the SH3-3/RIN3e1, since only a low affinity interaction was observed by ITC. Crystal formation was attempted by extensive sparse matrix sampling. The crystallisation screens used were: Hampton Research Index (HIN) screen, Hampton Research Crystal (HCS) screen, Joint Centre for Structural Genomics (JCSG) core screen and Ligand Friendly Screen (LFS). The results obtained are summarized in Table 3.5. Examples of crystals that were formed can be found in Figure 3.11. Crystal structures were solved for SH3-1 in complex with RIN3e1 and SH3-2 in complex with RIN3e2. Lastly, the high affinity of the SH3-3 domains with the P_{460}A mutant peptide prompted us to try to crystallise this complex. However, the crystallisation trials were unsuccessful.

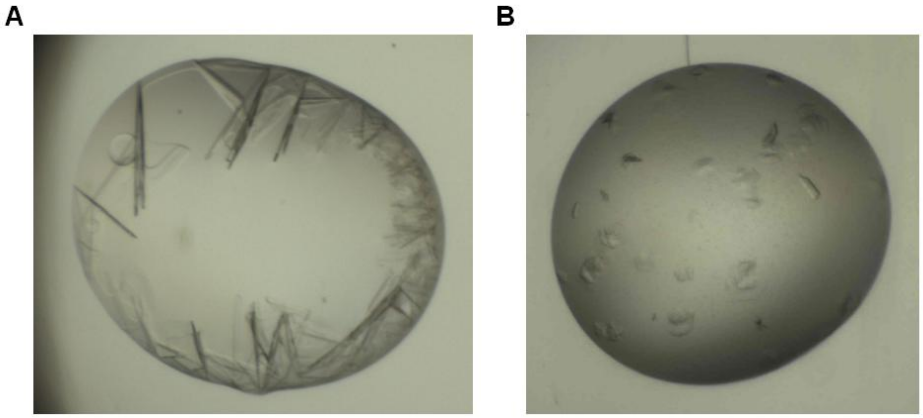


Figure 3.11

Examples of SH3/RIN3 crystals.

A, Crystals of the complex of SH3-1 with RIN3e1 that diffracted to 3.2 Å. ***B***, Crystals of SH3-2 in complex with RIN3e1 that diffracted poorly.

Table 3.5 Crystallization trial results.

SH3 domains alone (blue boxes); SH3 in complex with RIN3e1 (purple boxes), RIN3e2 (red boxes) or ALIX (green boxes) peptides. The boxes with a red outline indicate the successful crystallizations, which led to the generation of a structure.

	PRIMARY SCREENING	NEXT STAGE (including OUTCOMES)
SH3-1	• Structure previously deposited in the PDB by other group (PDB code: 2J6K)	
SH3-2	• No crystal formation in HCS screen • Theoretical pI is 4.45, so tested solubility in pH below the pI • Heavy precipitation in pH 4 (5 mM citric acid) • Crystal formation in 0.1 M cadmium chloride dihydrate, 0.1 M sodium acetate trihydrate pH 4.6 and 30% (v/v) polyethylene glycol 400	• X-ray diffraction pattern was not consistent with protein • Solubilisation of the crystal and mass spectrometry did not detect any protein • No crystal formation after setting up follow-up screen of the condition
SH3-3	• Tested both constructs • Tested concentrations ranging from 28 to 82 mg/ml in buffer without salt • No crystal formation	
SH3-1 + RIN3e1	• Crystals were formed in various conditions • Crystals, which diffracted to 3.2 Å resolution, were formed in 0.2 M Li_2SO_4, 0.1 M bis-tris pH 5.5 and 25% (w/v) PEG 3350 • X-ray diffraction pattern cannot be indexed due to multiple lattices	• Set up a follow-up screen of the best condition • Crystals were formed in Ca(ac) pH4, 0.2 M Li_2SO_4 and 25% (w/v) PEG 10k • X-ray diffraction data were collected to 1.6 Å resolution • Structure was solved and deposited • PDB code: 4WCI
SH3-2 + RIN3e1	• Crystals were formed in various conditions • Cluster of plates in all cases • X-ray diffraction pattern cannot be indexed due to multiple lattices	• Follow-up screens of the conditions did not result in better diffracting crystals • Additive screen and seeding were unsuccessful too
SH3-1 + RIN3e2	• No crystal formation in primary screens • Tried primary screens in low-salt buffer in concentration of 19.6 mg/ml	• Tried primary screens by increasing the concentration in buffer without salt (20-30 mg/ml) • Set up a screen at 4°C instead of 20°C • No crystal formation
SH3-2 + RIN3 e2	• Crystal was formed in 0.1 M HEPES pH 7.5, 1.4 M tri-sodium citrate dihydrate	• X-ray diffraction data were collected to 1.11 Å resolution • Structure was solved • PDB code: 3U23
SH3-3 + RIN3e2	• Crystal was formed in 0.01 M zinc sulphate heptahydrate, 0.1 M MES pH 6.5 and 25% (v/v) polyethylene glycol monomethyl ether 550	• X-ray diffraction data were collected to 4.5 Å resolution • Tried to optimise the condition by setting up a follow-up screen of the condition but no crystal formation
SH3-2 + ALIX epitope	• No crystal formation in primary screens • Tested concentrations ranging from 20 to 30 mg/ml in buffer without salt • No crystal formation	
SH3-3 + ALIX epitope	• No crystal formation in primary screens • Tested concentrations ranging from 20 to 30 mg/ml in buffer without salt • No crystal formation	

3.6.2 Crystal structure of CD2AP SH3-1 domain in complex with RIN3e1

3.6.2.1 Crystallisation, data collection, processing and refinement of the CD2AP SH3-1/RIN3e1 complex

Initial crystals grew as a cluster of needles at 20 °C from a 1:1 ratio of mother liquor (0.1 M acetate pH 4.5, 0.2 M Li_2SO_4 and 30% PEG 10000) to CD2AP SH3-1 protein solution (25 mg/ml) (Figure 3.11A). These crystals diffracted to 3.2 Ångstrom (Å) on beamline I04 at Diamond synchrotron (Harwell, UK; data were collected by Structural Genomics Consortium Oxford members).

An optimization screen was set up around this condition and crystals were grown at 20 °C from a 1:1 ratio of mother liquor (Ca(ac) pH 4.2, 0.2 M Li_2SO_4 and 25% PEG 10000) to SH3-1 protein solution (19 mg/ml). A cluster of needles (~160 μm long), which grew after 2 weeks, was first broken into separate pieces and then transferred to cryoprotectant (mother liquor supplemented with 25% glycerol) before being snap-frozen in liquid nitrogen. Data were collected to 1.6 Å on beamline I04 at Diamond synchrotron (by Structural Genomics Consortium Oxford members). Data reduction was carried out with the *XDS* package (Kabsch *et al*. 2010). The output file was loaded into *Pointless* (Collaborative Computational Project 1994), which is a program that identifies the possible space groups and scores them based on their probability of occurrence. Pointless suggested that the most probable space group was C2 (probability of 0.92), while another possible space group was P221 (probability of 0.06). Next, the dataset was converted to *CCP4* mtz format and scaled with *SCALA* (Evans 2006); data were truncated to 1.65 Å, giving a signal-to-noise ratio (I/sigma[I]) in the last resolution shell of 2.4,

rather than below 2.0. The cell content was analysed with the *Matthews_coef* program (Collaborative Computational Project 1994) and it was found that the complex crystallised as a trimer (*i.e.* three SH3-1 domains [chains A-C] and three RIN3 peptides [chains D-F]). Molecular replacement was performed with *Phaser* (McCoy *et al*. 2005) using as a search model the SH3 domain alone from a reported binary complex of CD2AP SH3-1 / CBL-B peptide (PDB code: 2J6F). The single, correct solution was refined with *REFMAC 5* (Murshudov *et al*. 1997). Additional manual model building was done with *Coot* (Emsley *et al*. 2004). The RIN3e1 peptide was initially modelled by superimposing the SH3-2 domain from the complex of CD2AP SH3-2/RIN3e2 peptide (PDB code: 3U23, section 3.6.3) on the three SH3-1 domains of this model and by exchanging the RIN3e2 residues to those present in RIN3e1.

Manual model building, a rigid body refinement cycle and several cycles of restrained refinement improved the model (*i.e.* reduction of the R_{work} and R_{free} values as shown in Figure 3.12). Non-crystallographic symmetry (NCS) restraints were applied to the three NCS copies of the SH3/peptide complex and the model was further improved (Figure 3.12). One SH3-1 copy (chain A, excluding residues 32-34) was used to restrain tightly the second SH3-1 copy (chain B) and restrain tightly the main chain of the third SH3-1 copy (chain C), with medium restraints applied to its side chains. One RIN3e1 copy (chain E) was used to restrain tightly the other two RIN3e1 copies (chains D and F). TLS (torsion, libration, screw) restraints were included for 6 groups, one for each chain in the asymmetric unit (*i.e*. three SH3-1 domains and three RIN3 peptides). However, the model improved marginally with one cycle of TLS restraint (Figure 3.12).

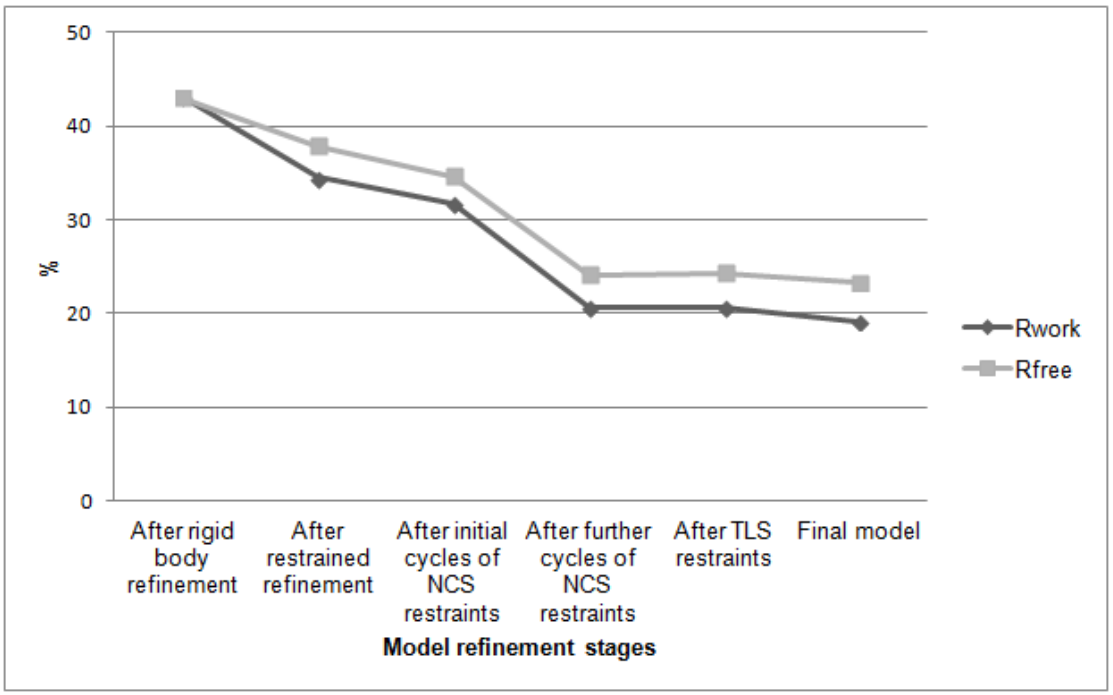


Figure 3.12

Reduction of R_{work} and R_{free} values during the SH3-1/RIN3e1 complex model refinement.

The different stages of model refinement are represented along the x-axis.

The final model was obtained after additional cycles of NCS restraints and manual model building. The data collection, processing and refinement data of the final model are found in Figure 3.13A. The Ramachandran plot (Lovell *et al* 2003) for the final model is found in Figure 3.13B. As depicted in Figure 3.13B, the model has good geometry. The residues are found in either favoured or allowed regions. The final model was deposited in the PDB with accession code 4WCI.

A

Data collection and processing	PDB: 4WCI
Diamond beamline	I04
Detector	Pilatus 2M
Wavelength (Å)	0.9173
Space group	C2
Unit cell parameters (Å)	95.53 55.36 64.56
(°)	90.00 131.7 90.00
Resolution range (Å)	48.19 - 1.65
I/σ(I)	6.6 (2.4)
Completeness (%)	97.7 (93.0)
R meas (%)	15.5 (56.6)
Refinement	
R work (%)	19.2 (27.8)
R free (%)	21.9 (32.1)
R.m.s.d. bond angles (°)	1.01
bond lengths (Å)	0.007
Ramachandran plot (%)	
Favoured regions	98.61
Allowed regions	1.39
Outliers	0
Additional data	
Highest resolution cell	1.70-1.65

B

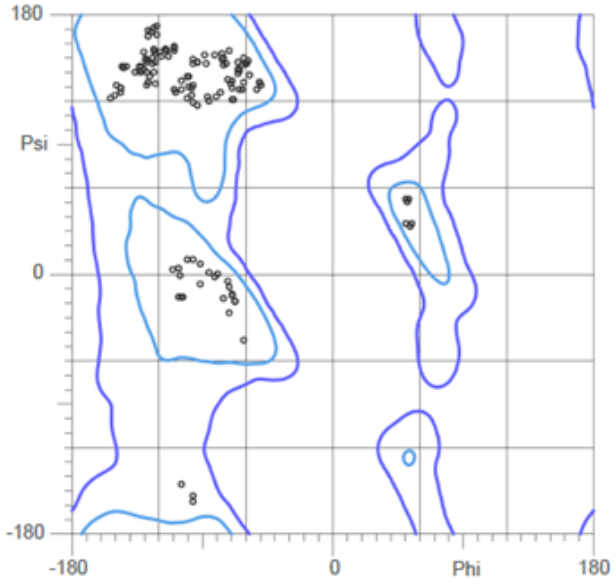


Figure 3.13

Quality of the final model of the SH3-1/RIN3e1 complex.

A, Data collection, processing and refinement parameters. Values in parentheses represent the highest resolution shell: 1.69 – 1.65 Å. ***B***, Ramachandran plot. The x-axis represents the phi value, while the y-axis represents the psi value for each residue. The allowed regions are shown in dark blue and the favoured regions are represented in light blue.

3.6.2.2 CD2AP SH3-1/RIN3e1 complex characteristics

The SH3-1 domain M_1-K_{58} and nine of the 16 RIN3e1 residues (N_{381}-R_{389}) were visible. The main chains of R_{390} and K_{380} were also visible in some peptide chains. GST-tag derived residues, water molecules, acetate ions and sulphate ions were also modelled. The structure suggests that the electrostatic interactions between the ions, water molecules, SH3-1 domains and the GST-tag derived residues assist trimer formation. This did not affect complex formation between the SH3-1 domain and RIN3e1 (cyan and orange ribbons in Figure 3.14, respectively) and the stoichiometry of the complex was one RIN3e1 peptide per SH3-1 domain. This was in accordance with the ITC measurements (Table 3.2).

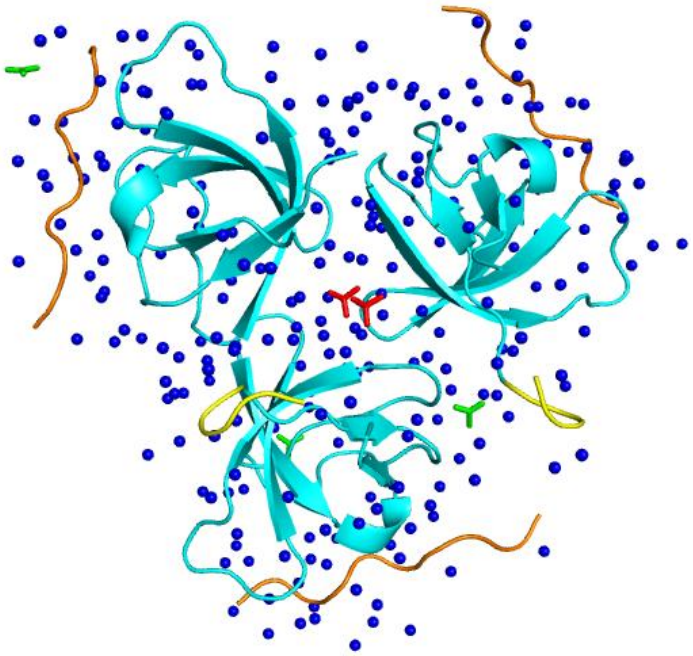


Figure 3.14

Asymmetric unit contents of the SH3-1 domain in complex with RIN3e1.

SH3-1 domain: cyan ribbon; RIN3e1: orange ribbon; GST-tag derived residues: yellow ribbon; $SO4^{2-}$: red sticks; $CH3COO^{-}$: green sticks; waters: blue spheres.

When the three SH3-1 molecules were superimposed according to the orientation of their secondary structure elements, they were almost identical (r.m.s.d ranging from 0.073 to 0.075 Å). In Figure 3.15, the high similarity of the electron density maps of the SH3-1 domain chains is shown for Asp_{15}-Ile_{20}. Therefore, chains A (i.e. one copy of SH3-1 domain) and D (i.e. one copy of RIN3e1 peptide) were used for further analysis.

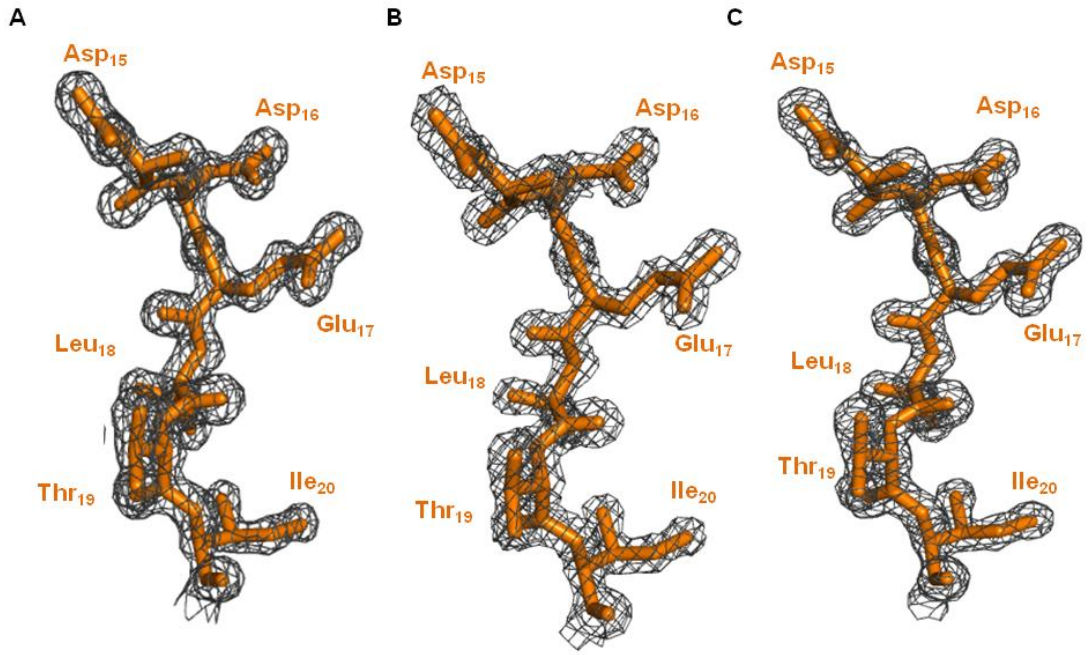


Figure 3.15

The three complexes in the asymmetric unit are almost identical (depicted for Asp_{15}-Ile_{20}).

Refined *2Fo-Fc* electron density maps (grey mesh) of the SH3-1 domain chains of the asymmetric unit (residues Asp_{15}-Ile_{20} modelled as orange sticks). The map is at 1.65 Å resolution contoured at 2.0 electron/$Å^3$: **A**, Chain A; **B**, Chain B; **C**, Chain C.

As depicted in Figure 3.16, the SH3-1 domain had the typical SH3 domain fold that was described in section 1.5. The five anti-parallel β-strands were formed by Tyr_4-Val_6, Ile_{25}-Lys_{30}, Trp_{37}-Leu_{42}, Arg_{45}-Pro_{50} and Val_{54}-Glu_{56}. Asp_{51}-Phe_{53} formed a 3_{10} helix. DSSP (Dictionary of Secondary Structure of Proteins) program (Kabsch *et al*. 1983, Joosten *et al*. 2010) was used for the

secondary structure assignment. This program assigns secondary structure elements based on the proteins' three dimensional structure. RIN3e1 docks onto SH3-1 in class II orientation and the domain's 3_{10} helix, RT and N-Src loops are positioned close to the peptide (Figure 3.16).

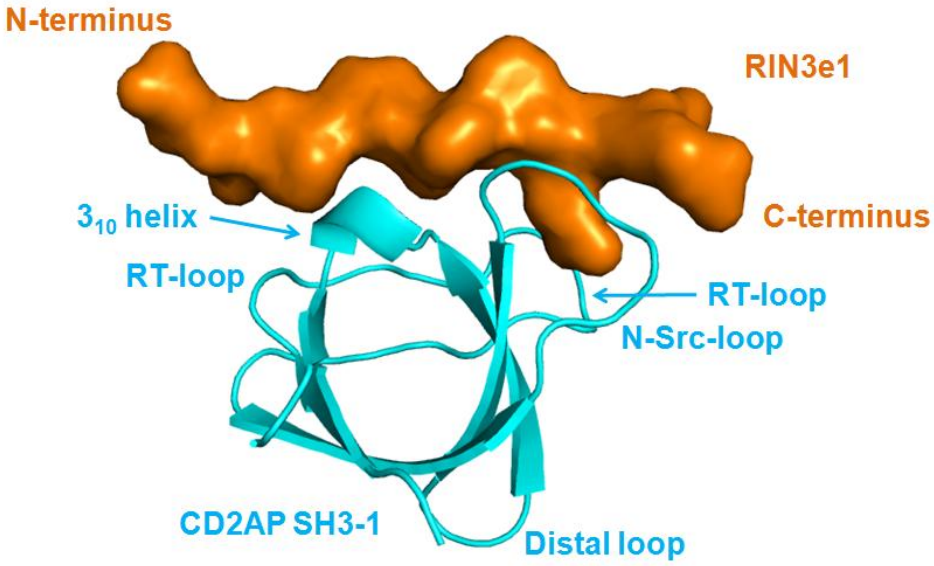


Figure 3.16

Surface representation of RIN3e1 bound to the β-barrel of the CD2AP SH3-1 domain.

SH3-1: cyan ribbon with the loops and 3_{10} helix being labelled; RIN3e1; orange surface.

SH3-1 and RIN3e1 are represented as surface and sticks respectively (Figure 3.17). RIN3e1 adopts a poly-proline type II (PPII) helical conformation and P_{382}, A_{384} and R_{387} were identified as its interaction points with SH3-1. According to the electrostatic potential surface representation, P_{382} and A_{384} slot into hydrophobic grooves (grey areas in Figure 3.17), while R_{387} is accommodated in a negatively charged groove (red area in Figure 3.17). In detail, P_{382} interacts with Y_8 and F_{53}, A_{384} interacts with Y_{10}, P_{50} and F_{53}, while R_{387} interacts with D_{16}, E_{17}, E_{34}, W_{37} and a peptide main chain atom. Also, N_{52}

contacts a peptide main chain atom (Figure 3.18). Therefore, the crystal structure confirms that SH3-1 recognises the P-x-A-x-x-R motif in RIN3e1.

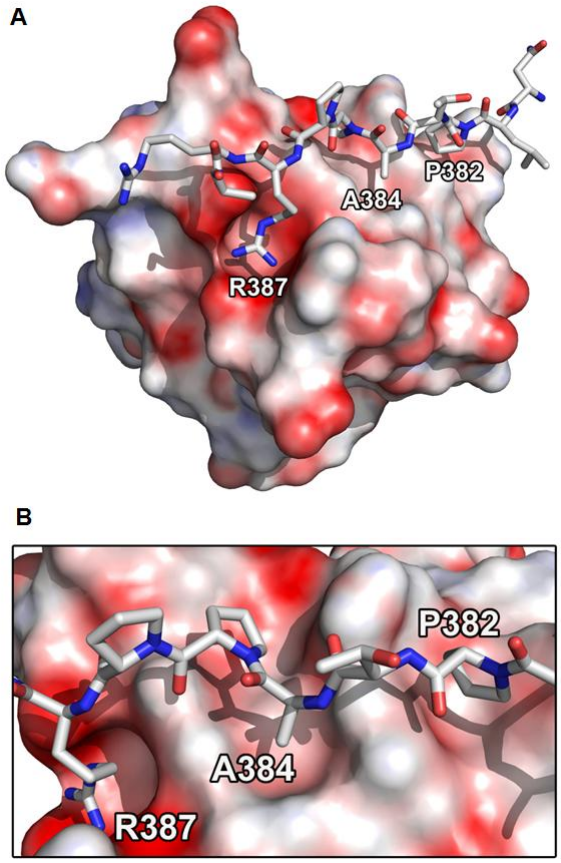


Figure 3.17

Electrostatic potential surface representation of the CD2AP SH3-1 domain in complex with bound RIN3e1.

The RIN3 peptide is shown in stick representation and coloured by element (carbon; white; nitrogen: blue; oxygen: red). The key residues that form contacts with the SH3 surfaces are labelled: ***A***, General view and ***B***, Close-up view of the docking positions of RIN3e1 peptide on the SH3-1 domain. These two views were generated by rotating the positions shown in *A* by -30^{o} around the y axis and 20^{o} around the x axis.

For further insight, it is sometimes informative to analyse the B-factors. In Figure 3.18, SH3-1 and RIN3e1 were coloured according to their B-factors. The blue and red represent low and high B-factor values respectively, and white is the intermediate colour. Low B-factor values indicate that a residue shows less thermal motion and is less flexible. Based on the B-factor values,

the SH3-1 domain core, β-strands and 3_{10} helix show less thermal motion than the loops. Similarly, the RIN3e1 residues contacting the SH3-1 surface show less thermal motion than the N- and C- termini. E_{34} and the side chain of Y_8 in SH3-1 are more flexible compared to the other SH3-1 critical residues as expected due to their position at the edge of the domain. This indicates that they were contributing less to the interaction with RIN3e1.

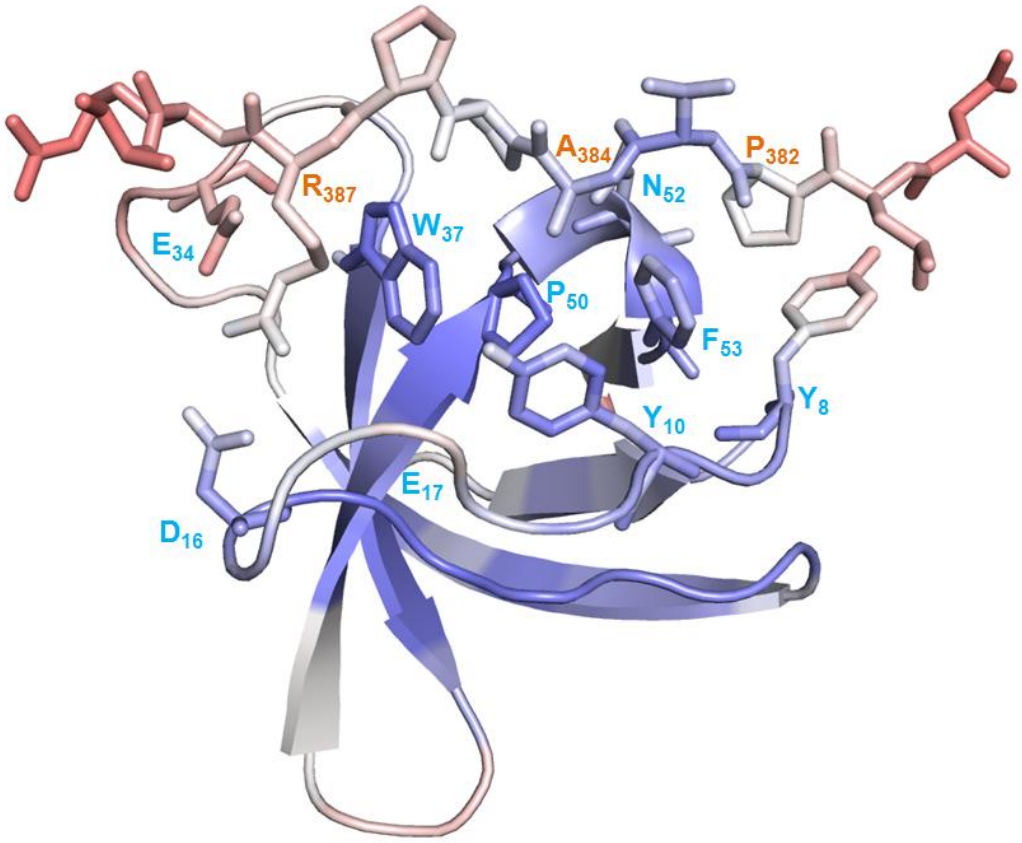


Figure 3.18

B-factors and points of contact between SH3-1 and RIN3e1.

The RIN3 peptide and the SH3-1 residues are shown in stick and secondary structure representation respectively and coloured according to their B-factor (blue to white to red). Blue signifies less thermal motion, and red signifies more thermal motion. Key SH3-1 residues that form contacts with RIN3e1 are depicted as sticks. Key RIN3e1 and SH3-1 residues that form contacts are labelled. The colour coding of individual amino acids is in accordance with previous figures: SH3-1 residues: cyan; RIN3e1 residues: orange.

As mentioned in section 1.5.5, the SH3-1 binding properties were studied by different groups. SH3-1 was studied by crystallography and reported to recognise the P-x-P-x-x-R motif and form trimers, which consist of two domains binding to the same CBL-B- or CD2-derived peptide in class I and II orientations (Moncalian *et al*. 2006). Another group studied the same interaction in solution and found that SH3-1 forms heterodimers in class I or II orientation with a CBL-B-derived peptide, and heterodimers in class II orientation with a CD2-derived peptide (Ceregido *et al*. 2013). When we compared the SH3-1 residues that contributed to the class II binding orientation to CBL-B, CD2 or RIN3e1 peptides, most key residues were identical (Y_8, D_{16}, $E_{17,}$ E_{34}, W_{37}, P_{50}, N_{52}, F_{53}). However, there were additional SH3-1 contacts to RIN3e1 (Y_{10}), CBL-B (E_{35}, G_{36}) and CD2 (E_{35}). Ceregido *et al*. (2013) showed that an Arg in -2 position upstream of the SH3-1 recognition motif in CBL-B enabled binding to class I orientation. The lack of an equivalent Arg in RIN3e1 might explain the SH3-1 binding to RIN3e1 in class II orientation.

3.6.3 Crystal structure of CD2AP SH3-2 with RIN3e2

The structure of the SH3-2 domain in complex with RIN3e2 was solved by Dr Philip Simister at 1.1 Å resolution. The asymmetric unit contains one complex. The electron density of R_{111}-E_{166} of SH3-2 and Q_{455} main chain to S_{467} of RIN3e2 is visible. As illustrated in Figure 3.19, the SH3-2 also has the typical β-barrel SH3 architecture. The β-barrel is formed by five anti-parallel β-strands, which consist of residues Q_{112}-V_{115}, I_{134}-V_{141}, W_{145}-L_{150}, K_{153}-P_{158} and V_{162}-E_{164}, and a 3_{10} helix, which is formed by residues S_{159}-F_{161}. The 3_{10} helix and RT and N-Src loops are positioned close to RIN3e2.

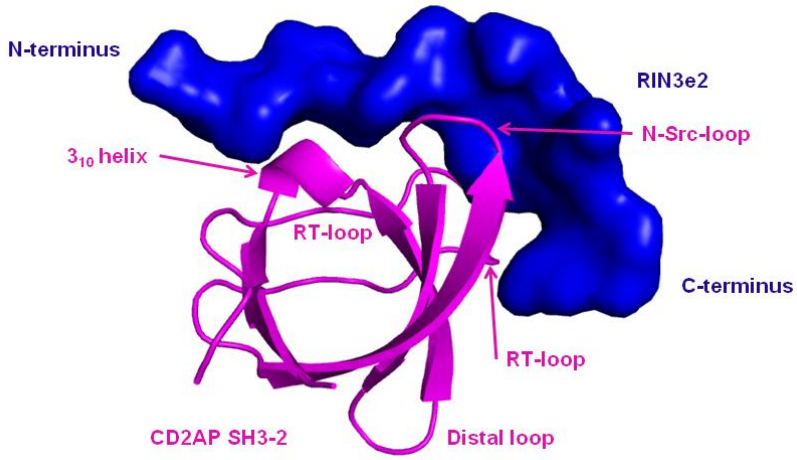


Figure 3.19

Surface representation of RIN3e2 docking onto the β-barrel of the CD2AP SH3-2 domain.

SH3-2: magenta ribbon with the loops and 3_{10} helix are labelled; RIN3e2: blue surface.

The peptide docks in a poly-proline type II (PPII) helical conformation onto the SH3-2 domain. Looking more closely at the structure, the following RIN3e2 residues form bonds with SH3-2: P_{457}, P_{459}, R_{462} and I_{466} (Figure 3.20): P_{457}, P_{459} and P_{460} and I_{466} are accommodated in three hydrophobic pockets and R_{462} in a negatively charged specificity pocket of the SH3-2 domain. The hydrophobic pocket, in which I_{466} sits, is more shallow than the two hydrophobic pockets, onto which P_{457} and P_{459} are docked. Thus, the crystal structure suggests that the SH3-2 domain might recognise an P-x-P-x-x-R-x-x-x-I motif in RIN3e2.

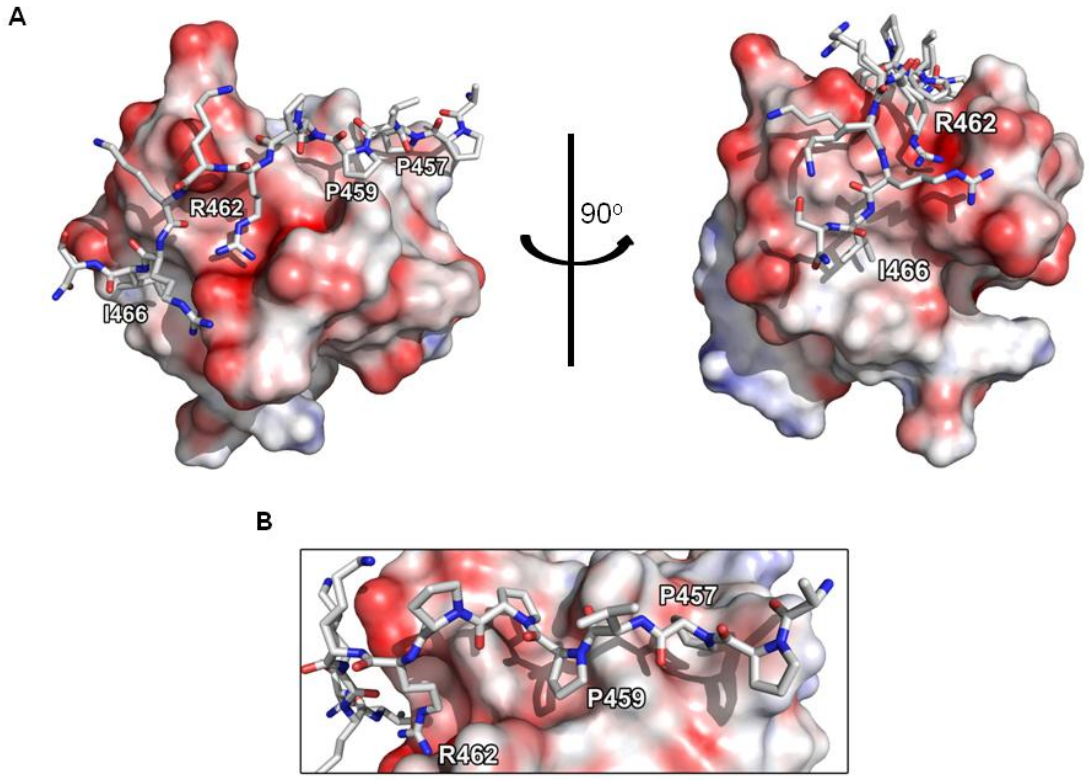


Figure 3.20

Electrostatic potential surface representation of the CD2AP SH3-2 domain in complex with RIN3e2.

RIN3e2 is shown in stick representation and coloured by element (carbon: white; nitrogen: blue; oxygen: red). The key residues that form contacts with the SH3 surfaces are labelled: ***A***, General view (2 views rotated by 90^{o}) and ***B***, Close-up view of the docking positions of RIN3e2 peptide on the SH3-2 domain. This view was generated by rotating the position shown in *A (left panel)* by -30^{o} around the y axis and 20^{o} around the x axis. An alternative conformation of R462 is not shown for clarity.

As illustrated in Figure 3.21, the SH3-2 contact points with RIN3e2 are F_{117}, Y_{119}, D_{125}, E_{126}, $W_{145,}$ V_{141}, L_{154}, L_{156}, N_{160} and F_{161}. F_{117} and F_{161} make hydrophobic contacts with P_{457}, while P_{459} interacts with F_{161} and Y_{119}. $D_{125,}$ E_{126}, W_{145} and K_{464} main chain atom contact R_{462}. R_{462} employs two alternative conformations. I_{466} forms hydrophobic bonds with V_{141}, L_{154} and L_{156}. N_{160} contacts the main chain of V_{458}. SH3-2 and RIN3e2 are coloured according to their B-factors in Figure 3.21.

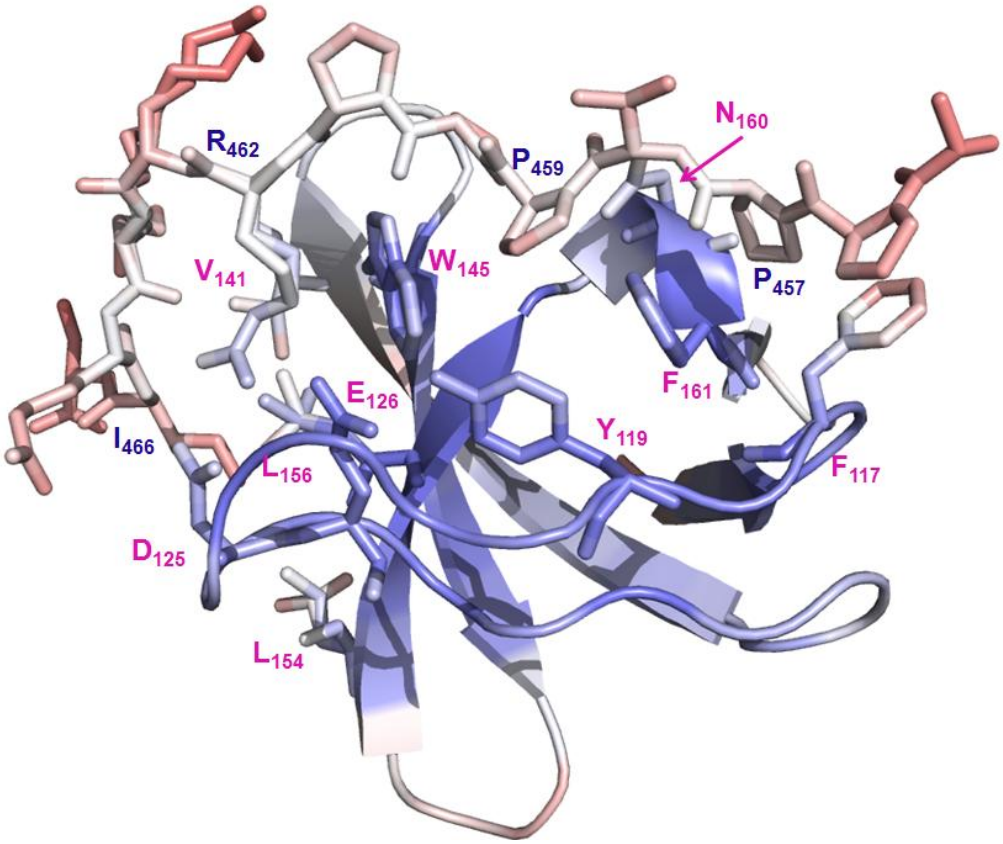


Figure 3.21

B-factors and points of contact between SH3-2 and RIN3e2.

The RIN3e2 peptide and the SH3-2 residues are shown in stick and secondary structure representation, respectively, and coloured according to their B-factor. Key SH3-2 residues that form contacts with RIN3e2 are also depicted as sticks. The key RIN3e2 and SH3-2 residues that form contacts are labelled. The colour coding of individual amino acids is in accordance with previous figures: SH3-2 residues: magenta; RIN3e1 residues: blue.

According to the B-factor values, the SH3-2 domain core, β-strands and 3_{10} helix are less flexible than the loops. As expected, P_{457}, P_{459} and R_{462} are also more thermo stable than the N- and C-termini. I_{466} (RIN3e2) and F_{161} (SH3-2) appear more flexible compared to the other residues critical for binding, which might indicate that they contributed less to their interaction.

According to the peptide permutation arrays, I_{466} is not critical for the SH3-2 binding to RIN32, because its point mutation against the other amino acids

does not affect binding (Figure 3.4). An ITC experiment was done to validate this using a point-mutant 16 residue RIN3e2 peptide ($I_{466}A$). The affinity of the interaction with SH3-2 and SH3-3 domains was similar to wild-type, while a 2-fold increase was observed for SH3-1. This clearly shows that this position is not critical for the SH3 interaction with RIN3e2.

However, RIN3e2 truncation (Table 3.2) decreased the peptide affinity to the three SH3 domains. It may be that the main chain atoms of the peptide contribute to the SH3-2/RIN3e2 interaction. This would explain how the truncation of the peptide rather than the mutation of a side chain affected the affinity of the interaction. When we looked closer at the structure, a water-mediated interaction was formed between a I_{466} main chain atom of the peptide and D_{125} of SH3-2 (Appendix B, Figure B.2).

3.6.4 Discussion

3.6.4.1 Comparison of CD2AP SH3-1 uncomplexed and complexed with RIN3e1

The uncomplexed SH3-1 crystal structure was solved by another group (PDB code: 2J6K). When our complexed and the uncomplexed SH3-1 domains are superimposed, the N-Src and some RT loop residues are the only SH3 regions that do not align well (r.m.s.d=0.35 Å) (Figure 3.22). These regions moved towards the SH3 domain core in the complexed structure so that the peptide remains docked onto the SH3 surface.

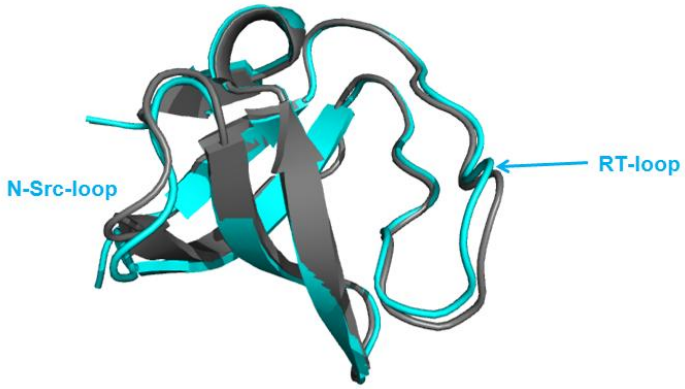


Figure 3.22

Comparison between the uncomplexed and complexed SH3-1 domains.

The uncomplexed SH3-1 is coloured grey (PDB code: 2J6K) and the complexed SH3-1 is coloured cyan. The N-Src and RT-loops are labelled.

The uncomplexed SH3-1 domain was coloured according to its B-factors (Figure 3.23A). When the uncomplexed SH3-1 B-factors were compared to the SH3-1 B-factors upon RIN3e1 binding (Figure 2.23), the SH3-1 core and the key SH3-1 binding residues, apart from $Y_{8,}$ are more thermo stable upon RIN3e1 docking. Y_8 might be more flexible, because it is positioned at the edge of the SH3-1 and RIN3e1 interface.

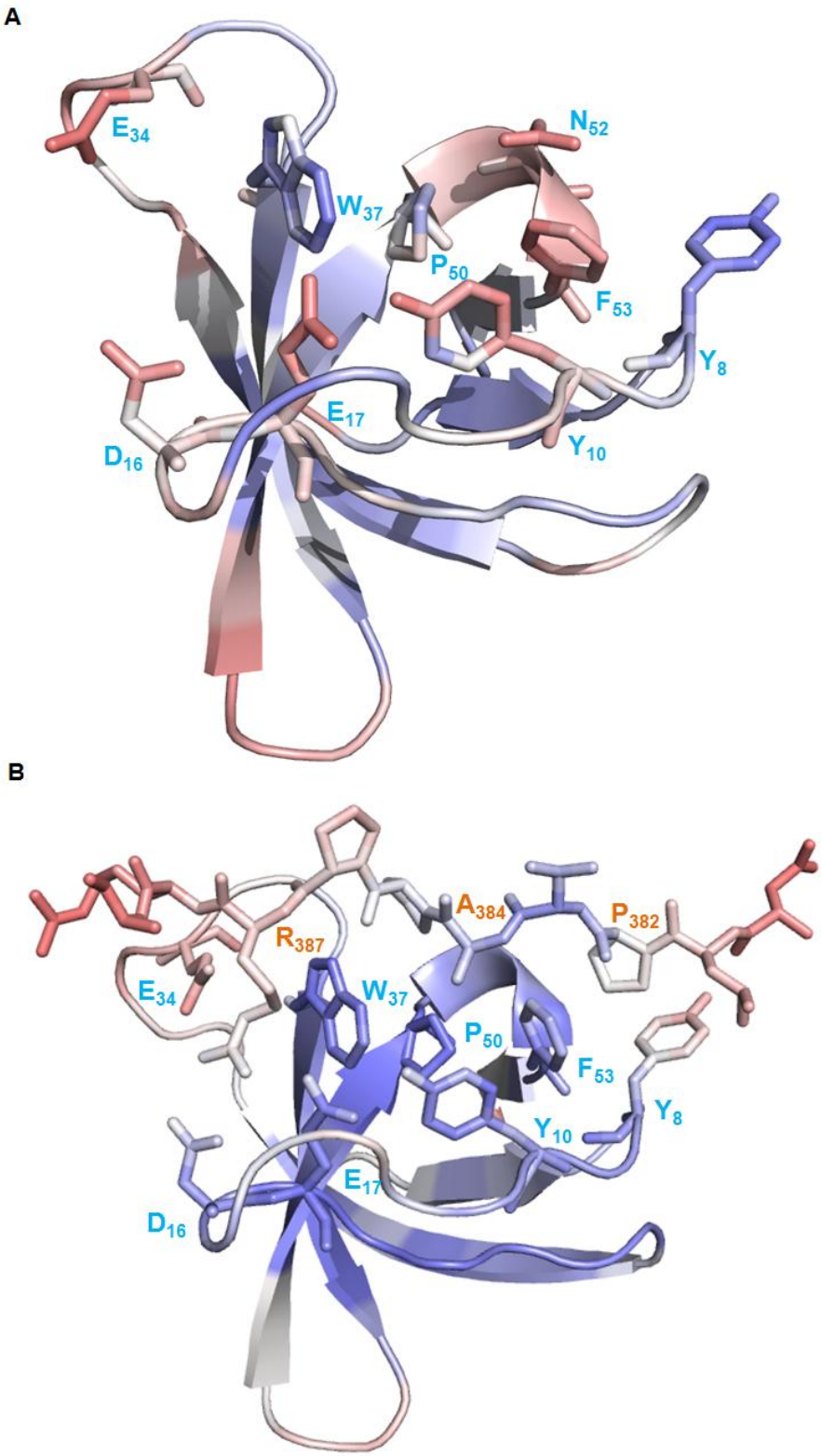


Figure 3.23

B-factors of uncomplexed (*A*) and complexed with RIN3e1 (*B*) SH3-1 domain (PDB: 2J6K).

A, The SH3-1 is represented by its secondary structure elements and coloured according to its B-factor (blue to white to red). Key SH3-1 residues that form contacts with RIN3e1 are depicted in sticks and are labelled cyan. Labelling in ***B*** is the same as for Figure 3.19.

3.6.4.2 Comparison of the CD2AP SH3-2 complexed with RIN3e2 and its unbound solution structure

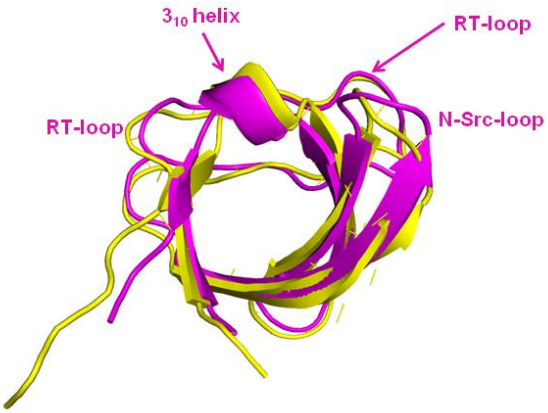


Figure 3.24

Comparison between the uncomplexed and complexed SH3-2 domains.

The uncomplexed SH3-2 is coloured yellow (PDB code: 2FEI) and the complexed SH3-2 is coloured magenta. The N-Src and RT-loops are labelled.

The unbound SH3-2 domain solution structure has been solved by Yao *et al* (2006) (PDB code: 2FEI), but the SH3-2 domain in complex with RIN3e2 presented here is the first crystal structure of this domain. Superimposition shows that the N-Src and RT loops and the 3_{10} helix move to the SH3-2 surface to accommodate RIN3e2 (r.m.s.d=0.98 Å) (Figure 3.24). Differences are found at the N- and C-termini, while the β-strands are almost identical.

3.6.4.3 Comparison of the CD2AP SH3-1 and SH3-2 complexed with the two RIN3-derived peptides

When overlaying the SH3-1/RIN3e1 and SH3-2/RIN3e2 complexes (Figure 3.25), the positioning of the RIN3-derived peptides in the two hydrophobic and the negatively charged groove (Figure 3.25A) are almost identical. The only major difference is the presence of A_{384} in RIN3e1. RIN3e2 has a Pro (P_{459}). Note that the overlay in Figure 3.25A is manually adjusted by moving the side

chain of A_{384} above P_{459}. This was done to show more clearly A_{384}. The unmodified overlay is depicted in Figure 3.25B. The N- and C- termini of both peptides (Figure 3.25B) do not align as well. Unlike I_{466} in RIN3e2, there is no visible density in an equivalent position in RIN3e1. The lack of density suggests that the N- and C-termini of the RIN3e1 peptide are still disordered upon RIN3e1 binding and do not make additional contacts with SH3-1.

As shown in Figure 3.25B, the first two SH3 domains fold similarly. Their loops and secondary structure elements largely overlap. The majority of the SH3 residues, critical for interaction with the RIN3-derived peptides are conserved within the first two SH3 domains, apart from F_{117} and E_{124}. The latter are conserved between SH3-2 and SH3-3, while SH3-1 has residues of similar charge (Y_8 and D_{15}). There are two bonds that seemed to contribute to the SH3-1 interaction with RIN3e1, which were not present equivalently in the SH3-2 interaction with RIN3e2: E_{34}-R_{387} and P_{50}-A_{384}. Conversely, the V_{141}, L_{154} and L_{156} SH3-2 residues interacted with I_{466} of RIN3e2.

According to the RIN3 permutation arrays, position 5 in the P-x-P/A-x-p-R is not critical for the interaction between the SH3 domains and the two RIN3 epitopes (Figure 3.7). The SH3-1/RIN3e1 and SH3-2/RIN3e2 structures revealed that P_{386} in RIN3e1 and P_{461} in RIN3e2 (i.e. motif position 5) do not contact the SH3 surface (Figure 3.25). These results confirm that this position is not critical for the SH3 binding to the two RIN3 epitopes. As described in section 1.4, residues within a motif, which do not contact the SH3 binding interface, may still play a role by assisting the ligand to employ a PPII conformation (Ferreon *et al*. 2004, Krieger *et al*. 2014). It could be that a Pro in motif position 5 favours the PPII conformation.

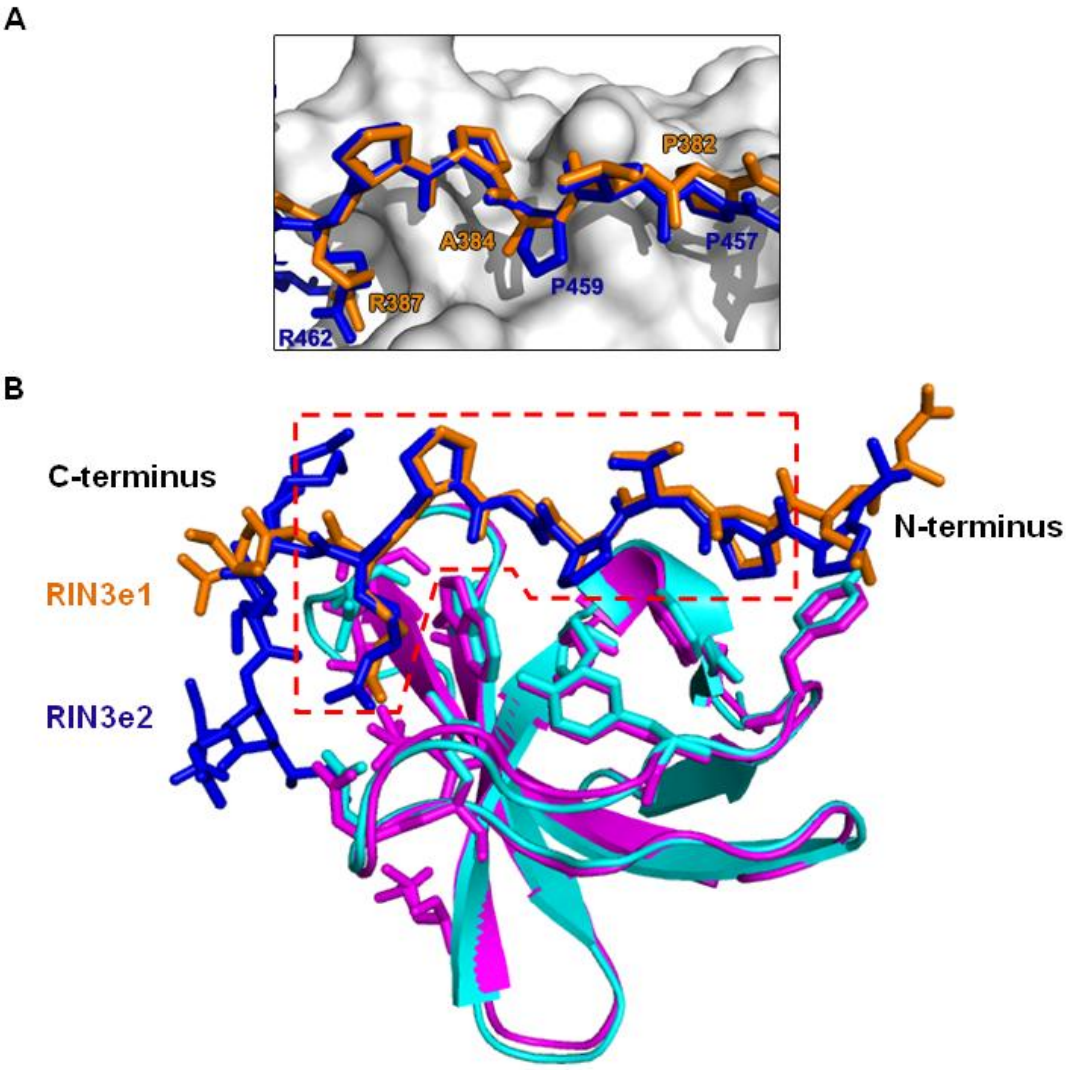


Figure 3.25

Overlay of SH3-1/RIN3e1 and SH3-2/RIN3e2 complexes.

RIN3e1: shown in orange stick representation; RIN3e2: blue sticks ***A***, Superposition of the two RIN3 peptides on the SH3-2 domain (white surface). The residues critical for the interaction are indicated. To note: A384 is slightly distorted for better clarity (the undistorted overlay is found in 3.25*B*.) ***B***, SH3-1 and SH3-2 in complex with RIN3e1 and RIN3e2 respectively. RIN3 peptides and the already described critical SH3-1 and SH3-2 residues for the interaction are shown in stick representation. The secondary structure elements of SH3-1 and SH3-2 are also depicted. Area in *A* represented in *B* is enclosed within red dashed lines. SH3-1: cyan; SH3-2 magenta.

The presence of two hydrophobic grooves, which appear to contribute to the SH3-1 and SH3-2 interactions with RIN3e1 and RIN3e2, respectively, is expected to cause a dominant favourable ΔS due to the liberation of surface

water molecules upon peptide docking and to contribute less to the ΔH of the interaction. Also, the competition between water molecules or ions and the peptides is expected to cause a less favourable ΔH. However, it is well published that the SH3 domain thermodynamic signature is commonly characterised by favourable ΔH and small ΔS (section 1.4). As already described (section 3.5.2), our thermodynamic data support the literature. The SH3-1 interactions were characterized by dominant favourable ΔH and small entropic changes. In the case of the SH3-1 complex with RIN3e1, the entropy was unfavourable. When SH3-2 interacted with RIN3e2, a higher entropic contribution was observed, which might be explained by I_{466} making hydrophobic contacts with SH3-2. However, the enthalpic contribution was bigger than the entropic contribution to the free energy.

A number of groups have tried to explain this discrepancy between the structural and thermodynamic SH3 data based on the observed populations of partially and complete engaged peptides to the SH3 surface. The loss of flexibility of the complete engaged peptide in the PPII conformation and of the SH3 domain's 3_{10} helix, RT- and N-Src loops may explain the unfavourable changes in entropy. Also, the polar contacts between the critical basic residue of the complete engaged peptide with the acidic SH3 pocket would contribute to the favourable ΔH (Ferreon *et al*. 2004, Ladbury *et al*. 2011, Stollar *et al*. 2012). This may explain the thermodynamic signature of the SH3 interactions with ALIX and RIN3e1.

Other groups suggested that the dominant enthalpic contribution might also be explained by water-mediated hydrogen bonds at the interface of the SH3 domains and their ligands (Palencia *et al*. 2004, Palencia *et al*. 2010,

Zafra-Ruano *et al*. 2012). Water networks between the first two SH3 domains and the RIN3 peptides were observed, which could contribute to the favourable enthalpy (see Appendix B, Figures B.1 and B.2). However, more biophysical studies have to be performed to study their importance, which is beyond the scope of this thesis.

3.7 Conclusion

The three CD2AP SH3 domains recognise two RIN3 epitopes and one ALIX epitope. Their binding modes and specificity was investigated by isothermal titration calorimetry, crystallography and peptide arrays using individual SH3 domains.

The ITC data strongly imply that there are differences in the binding affinities and thermodynamic signatures of the three SH3 domains upon interaction with the ALIX and RIN3 peptides. The first two SH3 domains bind with similar affinities in the low micromolar range to these peptides, while SH3-3 binds with lower affinity to the first RIN3 epitope and the ALIX epitope. Interestingly, SH3-3 binds with higher affinity to RIN3e2, which is comparable to the first two SH3 domain binding affinities. This suggests that the SH3-3 domain might contribute to binding similar to the first two SH3 domain in some cases, but that its binding requirements are different from the other CD2AP SH3 domains. The CD2AP SH3-mediated interactions are characterised by dominant favourable enthalpy, which is typical for SH3 domains. Interestingly, favourable entropy is observed upon binding of the SH3-2 and SH3-3 domains to RIN3e2, while small favourable or unfavourable entropic changes are exhibited when the three domains bind to the ALIX and RIN3e1 epitopes.

Based on the protein crystallography experiments, a structural view of the SH3-1/RIN3e1 and SH3-2/RIN3e2 complexes was obtained. Conserved residues or residues of similar charge in the first two SH3 domains, which are located within the 3_{10} helix and N-Src and RT-loops, contributed to the interaction with the RIN3-derived peptides. The latter dock onto three grooves formed by the critical SH3 residues. ITC analysis with mutant peptides and RIN3 permutation arrays revealed that the recognition motif of the three SH3 domains is P-x-P/A-x-p-R.

As described in section 3.4, the peptide permutation arrays inform on the positively and negatively selected residues within a motif. Regarding negative selection, the RIN3 permutation array data suggest that the CD2AP SH3 P-x-P/A-x-p-R recognition motif may be further refined to P-{FGHWY}-P/A-{D}-{FGHWY}-R-{P} (Figure 3.7). The curly brackets of this motif refer to the negatively selected residues.

Finally, the structural analysis revealed that there is a network of water molecules present at the binding interface between the first two SH3 domains and the RIN3 peptides. Extensive biophysical analysis, such as by NMR and computational methods, would have to be performed to verify how much these water molecules contribute to binding. Such studies would also help to delineate the thermodynamic signature upon SH3 domain binding to the ALIX and RIN3 epitopes and bridge the gap between the thermodynamic and structural data. However, such studies are beyond the scope of this thesis and were not pursued.

Chapter 4: ITC analysis of the CD2AP SH3-RIN3 interaction with a tandem SH3 domain

4.1 Overview

Having determined the molecular details of the interaction between the isolated CD2AP SH3 domains and the two RIN3 epitopes by ITC, protein X-ray crystallography and peptide arrays (see Chapter 3), we attempted to investigate whether the individual SH3 binding properties were affected by the presence of the other CD2AP SH3 domains. Moreover, the tandem RIN3 epitopes might be a good model to test whether an increase in affinity and/or selectivity is observable. Aitio *et al.* (2010) showed that the SH3 domain of IRTKS (insulin receptor tyrosine kinase substrate) binds very tightly (K_d=0.5 μM) to tandem PxxP motifs in $EspF_u$ (E.coli-secreted protein F-like protein encoded on prophage U), while the affinity to each motif individually is lower.

The binding mode of tandem SH3 domains remains largely unexplored and the studies focus mainly on the two tandem SH3 domains of $p47^{phox}$(Groemping *et al*. 2003, Yuzawa *et al*. 2004, Ogura *et al*. 2006). Apart from SH3 domains, the WW domain is another proline-rich sequence recognition domain (PRD) that is found in tandem in proteins. In general, the WW domains interact with their partners independently and in tandem (Sudol *et al*. 2005).

To further gain insight into tandem SH3 binding, tandem constructs of the first two CD2AP SH3 domains and of both RIN3 epitopes were generated by our collaborator (Dr Kathrin Kirsch *et al*., University of Boston, USA). These constructs will be referred to as SH3-(1+2) and RIN3-(1+2), respectively,

throughout this thesis. The first step was to express and purify these constructs without the GST-tags for ITC and protein crystallography experiments to obtain a structural and thermodynamic view of their interaction. Such data would have led to a better understanding of the selectivity of the first two SH3 domains.

The purification of the GST-RIN3-(1+2), GST-SH3-(1+2) and SH3-(1+2) constructs was successful. However, the purification of RIN3-(1+2) was not possible. Therefore, ITC analysis was performed with SH3-(1+2) and the individual RIN3e1 and RIN3e2 peptides (Table 3.2). Furthermore, attempts to crystallise SH3-(1+2) were unsuccessful.

Lastly, the presence of a proline-rich region in CD2AP prompted us testing whether the CD2AP SH3 domains might form intra- or intermolecular interactions by binding to this region.

4.2 Purification of the CD2AP SH3-(1+2) construct

Residues corresponding to the first two SH3 domains and their linker region (aa 1-168) were cloned into a pGEX6-P1 vector. A detailed description of the bacterial expression and purification procedure is found in section 2.2.1. Initially, the same general workflow of bacterial and protein expression and protein purification as for the individual SH3 domains (Figure 3.1) was followed. The integrity and purity of the protein was tested at various stages of the purification process.

As shown in Figure 4.1A, bound GST-SH3-(1+2) to GSH beads was successfully eluted with reduced glutathione (elution buffer). A small amount of GST-SH3-(1+2) remained bound to the GSH beads after elution. The

GST-SH3-(1+2) band was found below the 52 kDa marker, which was in agreement with its predicted molecular weight (45.5. kDa). The GST-tag was cleaved off with Prescission^{TM} protease and only a small amount of GST-SH3-(1+2) remained uncleaved (Figure 4.1B). The predicted molecular weights of GST and SH3-(1+2) were 26 and 19.5 kDa respectively. As depicted in Figure 4.2B, the GST and SH3-(1+2) bands ran at a similar position to the 24 kDa marker respectively. Even though the GST and SH3-(1+2) molecular weights were similar, their separation by size exclusion chromatography might be possible due to GST dimerization. The remaining GST-SH3-(1+2) was separated from GST and SH3-(1+2). However, GST and SH3-(1+2) still co-eluted from the SEC column (HiLoad 16/60 Superdex S75 column) (Figure 4.1C).

This elution profile of SH3-(1+2) could be explained in a number of ways. It could have resulted from SH3-(1+2) oligomerization. Also, SH3-(1+2) comprises two globular SH3 domains and a 50-residue linker region. Therefore, its elution volume could correspond to an apparent higher molecular weight rather than its actual molecular weight due to its possibly non globular conformation. To overcome this problem, ammonium sulphate precipitation was attempted.

As illustrated in Figure 4.1D, the majority of SH3-(1+2) could be isolated by ammonium sulphate precipitation (40-50%). These fractions contained only small amounts of GST-SH3-(1+2) and GST. The majority of GST precipitated in the 80% fraction. Next, proteins in the 40-50% ammonium sulphate fractions, were further separated by SEC.

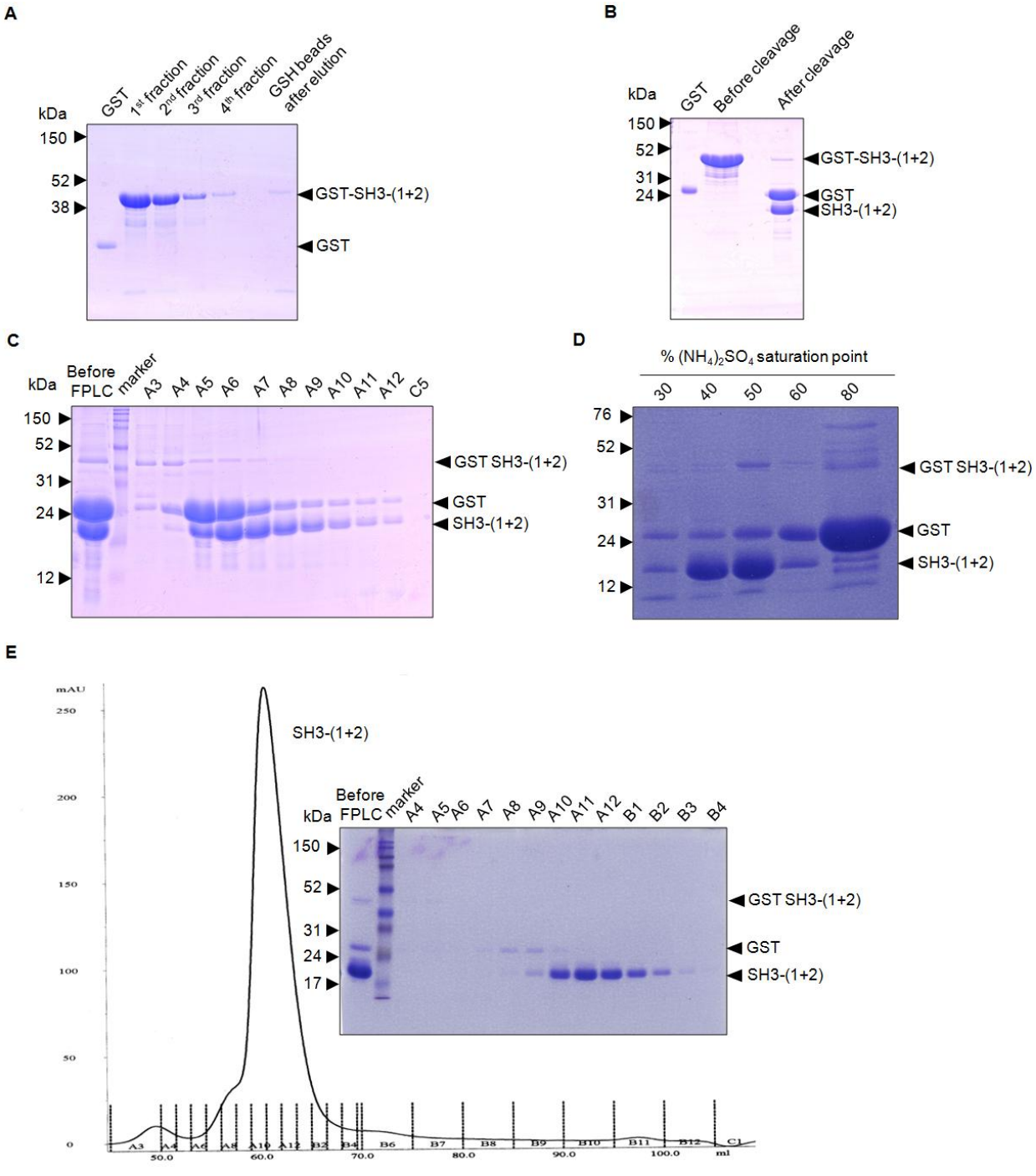


Figure 4.1

CD2AP SH3-(1+2) purification.

All samples were resolved by SDS-PAGE and the gels were stained with Coomassie blue. Arrows on the left of each gel show the position of the markers. All gels were 12% resolving apart from ***B***, which was a 16% resolving gel. ***A***, GST-SH3-(1+2) was eluted from the GSH beads and collected into fractions with reduced glutathione. ***B***, The GST-tag was cleaved off SH3-(1+2) by overnight incubation with Prescission™ protease. ***C***, The GST-tag and SH3-(1+2) co-eluted by size exclusion chromatography. ***D***, GST and SH3-(1+2) were separated by ammonium sulphate precipitation. ***E***, The remaining GST was separated from SH3-(1+2) by SEC.

As shown in Figure 4.1E, the majority of SH3-(1+2) was successfully separated from GST-SH3-(1+2) and GST. SH3-(1+2) eluted at 59 ml, which corresponds to 39 kDa. This elution profile suggests that SH3-(1+2) might dimerize or adopt a more extensive conformation. GST and SH3-(1+2) were present in the A11 fraction. The fractions that contained only SH3-(1+2) according to the SDS-PAGE gel, were combined. The integrity and purity of the final SH3-(1+2) is illustrated in Figure 4.1E. The fractions containing GST and SH3-(1+2), were collected.

4.3 Attempts to purify RIN3-(1+2)

The two RIN3 epitopes and their linker region (aa 378-467), were cloned into a pGEX6-P1 vector. Initially, the same general workflow of bacterial and protein expression and protein purification as for the individual SH3 domains (Figure 3.1) was followed. As shown in Figure 4.2A, the majority of GST-RIN3-(1+2) was successfully eluted with reduced glutathione from the GSH beads. The predicted molecular weight of GST-RIN3-(1+2) is 35.8 kDa and its corresponding band was observed above the 31 kDa marker. Also, a number of bands were present in the first two elution fractions that extended between the GST and GST-RIN3-(1+2) bands. These bands might correspond to incomplete GST-RIN3-(1+2). As depicted in Figure 4.2B, the GST-tag was cleaved with Prescission protease and a small amount of GST-RIN3-(1+2) remained uncleaved. The predicted molecular weights of GST and RIN3-(1+2) were 26 and 9.8 kDa respectively. The GST band was visible around the 24 kDa marker as expected. However, RIN3-(1+2) was not detected. Another band of roughly 76 kDa was visible before and after cleavage, which might be a contaminant. The stain used in Figure 4.2B, was

Coomassie R Brilliant Blue. This dye is more efficient in staining protonated basic rather than acidic residues, but it also binds to non-polar protein regions non-covalently (de St Groth *et al*. 1963). Even though the 90 residue RIN3-(1+2) contains eight Arg, one His and six Lys, the protein did not bind the dye effectively. Also, staining with Ponceau S was also unsuccessful to detect RIN3-(1+2) (Figure 4.2C). Ponceau S recognises similar residues to Coomassie R Brilliant Blue (Salinovich *et al*. 1986). Therefore, a number of different stains were tested so that RIN3-(1+2) might become more visible in acrylamide gels. Roti-Blue is a colloid form of a dimethyl derivative of Coomassie R Brilliant Blue, which is called Coomassie G Brilliant Blue (Diezel *et al*. 1972). With this, only the bands already described were detectable (Figure 4.2D). Finally, silver staining was attempted, but no RIN3-(1+2) could be detected according to its predicted molecular weight (Figure 4.2E). This stain detects individual cysteines and homopolymers of methionine and basic residues (Merril *et al*. 1986). RIN3-(1+2) does not contain cysteines but it has small stretches of basic residues. In all cases, the GST band and some higher molecular weight bands were detectable.

RIN3-(1+2) detection was also attempted by immunoblotting GST-RIN3-(1+2) samples, which were incubated for different times (2 h, 4 h, 6 h, 7 h and 15 h) with Prescission^{TM} protease, with an anti-RIN3 antibody (SAB4503182, section 2.1.1). GST-RIN3-(1+2) was detected, however, no signal was detected in the lower molecular weight band range (data not shown). Unfortunately, RIN3-(1+2) lacks aromatic residues, and it could not be detected by UV spectroscopy. RIN3-(1+2) is predicted to be disordered by the

IUPred server (Dosztanyi *et al*. 2005). Therefore, it might be degraded upon GST-tag cleavage.

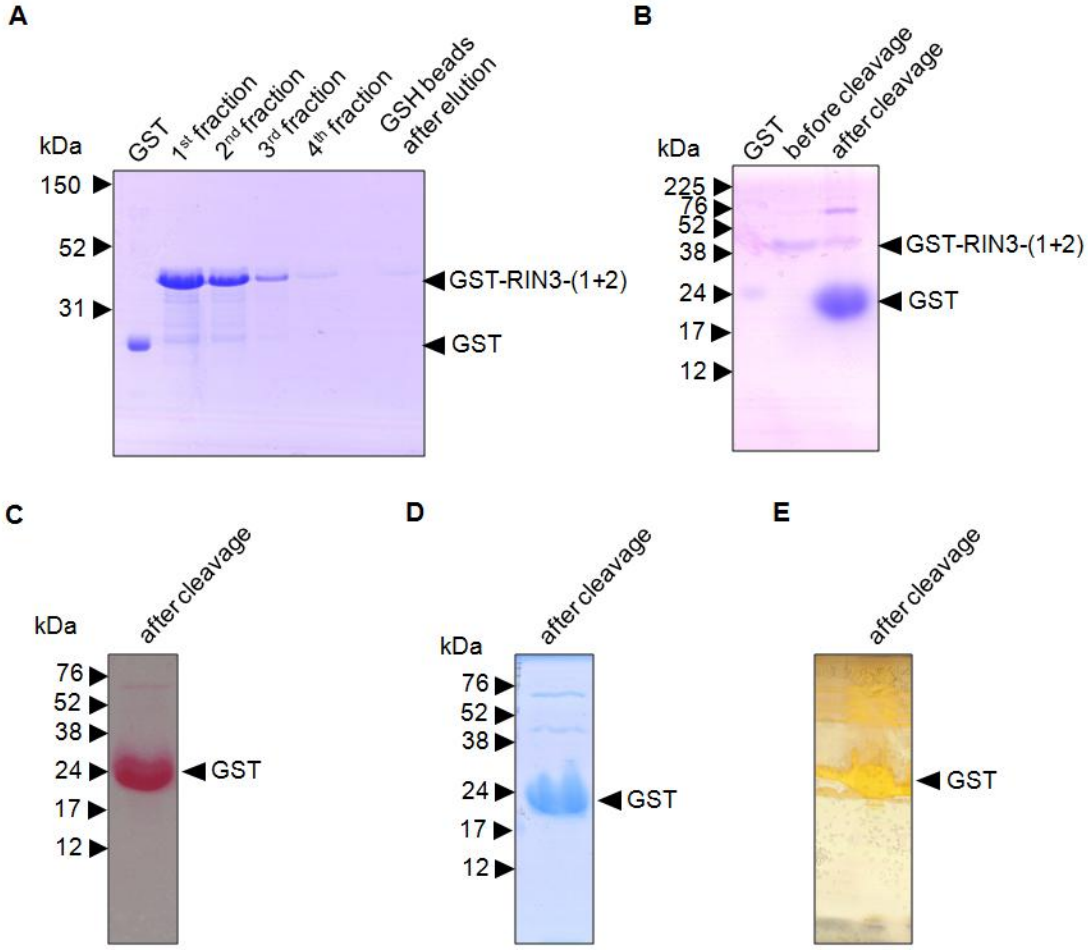


Figure 4.2

RIN3-(1+2) purification.

A, GST-RIN3-(1+2) was eluted from the GSH beads and collected into fractions with reduced glutathione (12% resolving gel). ***B***, The GST-tag was cleaved off RIN3-(1+2) by overnight incubation with Prescission™ protease (16% resolving gel). RIN3-(1+2) was not detected. The gel was stained by Coomassie Brilliant R Blue; Samples from the mixture upon GST cleavage were stained with: ***C***, Ponceau S (PVDF membrane). ***D***, Roti-Blue (16% resolving gel). ***E***, Silver stain (16% resolving gel). All gels were resolved by SDS-PAGE.

In summary, the study of RIN3-(1+2) by ITC and protein crystallography was not possible due to the lack of detection of the RIN3-(1+2) band upon GST cleavage. However, the purification of GST-RIN3-(1+2) was successful and it

could be used for other types of biochemical assays. In Chapter 5, its use for CD2AP precipitation via its GST-tag will be further discussed.

4.4 ITC with CD2AP SH3-(1+2) and the two RIN3-derived epitopes

The successful purification of SH3-(1+2) but not RIN3-(1+2) only allowed ITC experiments between SH3-(1+2) and the RIN3e1 or RIN3e2 peptides. The dissociation constant and stoichiometry values are shown in Table 4.1, while the free energy, enthalpy and entropy data are depicted in Figure 4.3.

The K_d values of SH3-(1+2) upon binding to RIN3e1 and RIN3e2 were 10.5 and 4.2 μM respectively (Table 4.1). Comparing these values with the equivalent values of the individual SH3 domains with the same peptides, similar interaction affinities are observed. Thus, avidity effects are not apparent.

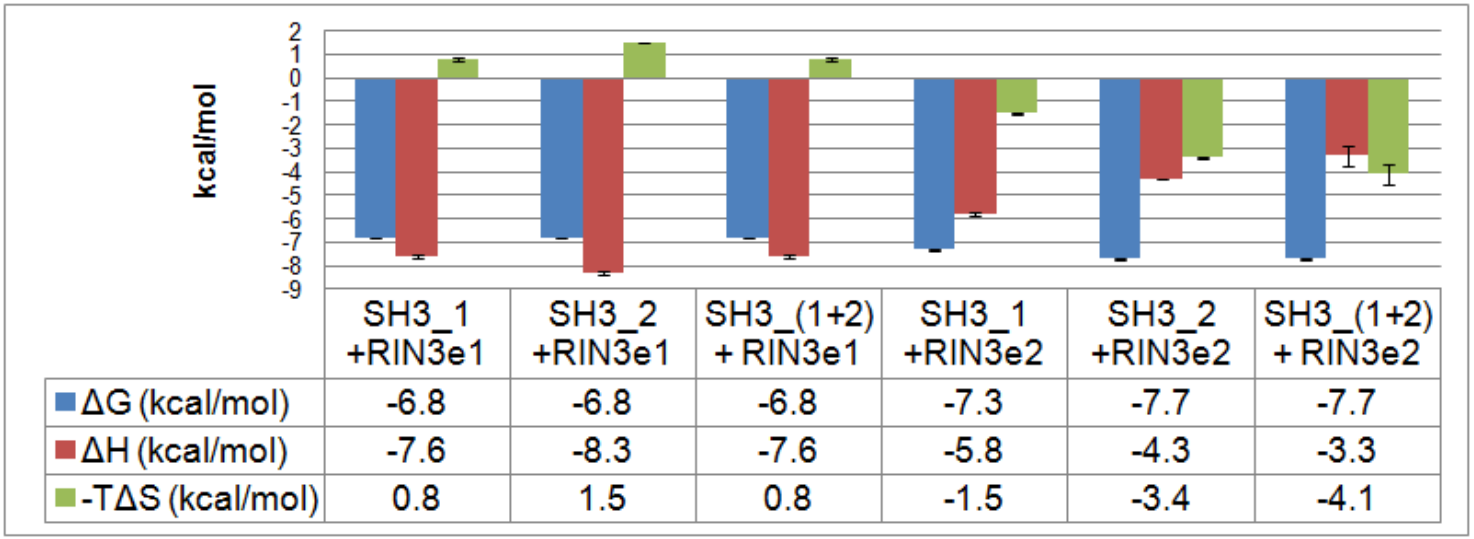
If we look at the N values, SH3-(1+2) binds to one RIN3e1. SH3-(1+2) may bind to one or two RIN3e2. Based on the N value of 1 of the individual SH3 complexes with RIN3e1 or RIN3e2 peptides (Table 4.1), an N value of 2 was expected upon binding of RIN3e1 or RIN3e2 to SH3-(1+2).

The thermodynamic signature of the SH3-(1+2) interaction with RIN3e1, is characterised by a dominant favourable ΔH and small unfavourable ΔS (Figure 4.3). This indicates that polar interactions contribute to binding. The dominant ΔH might be caused due to the docking of one peptide to SH3-(1+2), which would result in higher solvent accessibility to SH3-(1+2). The SH3-(1+2)/RIN3e1 thermodynamic values are identical to the SH3-1/RIN3e1 ones (Figure 4.3).

Table 4.1 Isothermal titration calorimetric measurements of CD2AP SH3-1, SH3-2 and SH3-(1+2) domain interactions with wild-type RIN3 peptides: Dissociation constant and stoichiometry values.

RIN3 epitope	Peptide sequence	Length (a.a.)	K_d ± SE (μM)			N ± SE		
			SH3-1	SH3-2	SH3-(1+2)	SH3-1	SH3-2	SH3-(1+2)
e1	AKKNLPTAPPRRRVSE-am	16	8.9 ± 0.2	10 ± 0.2	10.5 ± 4.2	0.85 ± 0.03	0.64 ± 0.02	1.16 ± 0.01
e2	TAKQPPVPPPRKKRIS-am	16	4.6 ± 0.4	2.1 ± 0.8	4.2 ± 0.3	1.18 ± 0.01	0.94 ± 0.03	1.65 ± 0.02

SE = standard error.



	SH3_1 +RIN3e1	SH3_2 +RIN3e1	SH3_(1+2) + RIN3e1	SH3_1 +RIN3e2	SH3_2 +RIN3e2	SH3_(1+2) + RIN3e2
ΔG (kcal/mol)	-6.8	-6.8	-6.8	-7.3	-7.7	-7.7
ΔH (kcal/mol)	-7.6	-8.3	-7.6	-5.8	-4.3	-3.3
-TΔS (kcal/mol)	0.8	1.5	0.8	-1.5	-3.4	-4.1

Figure 4.3: Thermodynamic parameters of binding of the RIN3 epitopes to the CD2AP SH3-1, SH3-2 and SH3-(1+2) domains.

In the case of the SH3-(1+2)/RIN3e2 interaction, ΔH and ΔS were favourable, and contributed similarly to the ΔG and hence the affinity (Figure 4.3). This suggests that polar and hydrophobic interactions are critical for binding. The thermodynamic signature upon SH3-(1+2) binding to RIN3e2 resembles the thermodynamic parameters of SH3-2/RIN3e2 (Figure 4.3).

In summary, the SH3-(1+2) interaction with RIN3e1 or RIN3e2 does not result in avidity effects compared to the SH3-1 or SH3-2 binding to the same peptides.

4.5 Attempts to study CD2AP SH3-(1+2) by protein X-ray crystallography

The successful purification of SH3-(1+2) might allow the study of the tandem domains by crystallography. A JCSG screen was set up at 15.9 mg/ml and precipitation occurred after 11 h at 50% of the drops, but no crystals were formed. Afterwards, another JCSG primary screen was set up at 4.9 mg/ml, but no crystal formation was observed up to 3 months.

4.6 Investigation of possible interactions between the CD2AP SH3 domains and the CD2AP proline-rich region

The already described CD2AP SH3 P-x-P/A-x-p-R recognition motif in the two RIN3 epitopes (Chapter 3) is not present in the CD2AP proline-rich region. This suggests that the SH3 domains might not recognise the CD2AP proline-rich region. However, the high Pro (52%) and Arg (32%) content in the proline-rich region could indicate that interactions between the SH3 domains and the proline-rich region might still be possible via a different binding mode.

Peptide scanning arrays of the proline-rich region (aa 330-520) were provided by Nicola O' Reilly (Peptide Synthesis Laboratory, Cancer Research UK, London) to further test this hypothesis. The arrays were probed with the three individual SH3 domains (SH3-1, SH3-2, SH3-3-S) as GST-fusions. Protocol is detailed in section 2.2.8.3.2. The overlay blots are shown in Figure 4.4.

As illustrated in Figure 4.4, the arrays were initially probed with GST and no signal was detected. As a positive control, a RIN3e2-derived peptide (aa 435-462) was spot-synthesised in duplicate (Figure 4.4, green box). The proline-rich region-derived peptides were spot synthesised as single spots. As shown in Figure 4.4, SH3-3-S did not bind to the proline-rich region, while the first two SH3 domains recognised three principal binding regions (cyan, yellow and pink boxes in Figure 4.4). However, the blot exposure times indicate that SH3-1 and SH3-2 bound with much lower affinity to these regions compared to RIN3e2. The blot exposure times of Figure 4.4 were 1 min, but the proline-rich region spot signal was obvious after 30 sec. By comparison, the RIN3e2 spot signal was present after 1 sec, which indicates that the SH3 domains is unlikely to make prominent inter- or intra-molecular contacts with the CD2AP proline-rich region.

This finding is further supported by the determination of the structure of the full-length CD2AP by electron microscopy by another group. This study was published after the completion of the above mentioned experiment and it shows that the coiled-coil region, but not the SH3 domains, of CD2AP contributes to tetramer formation (Adair *et al*. 2014).

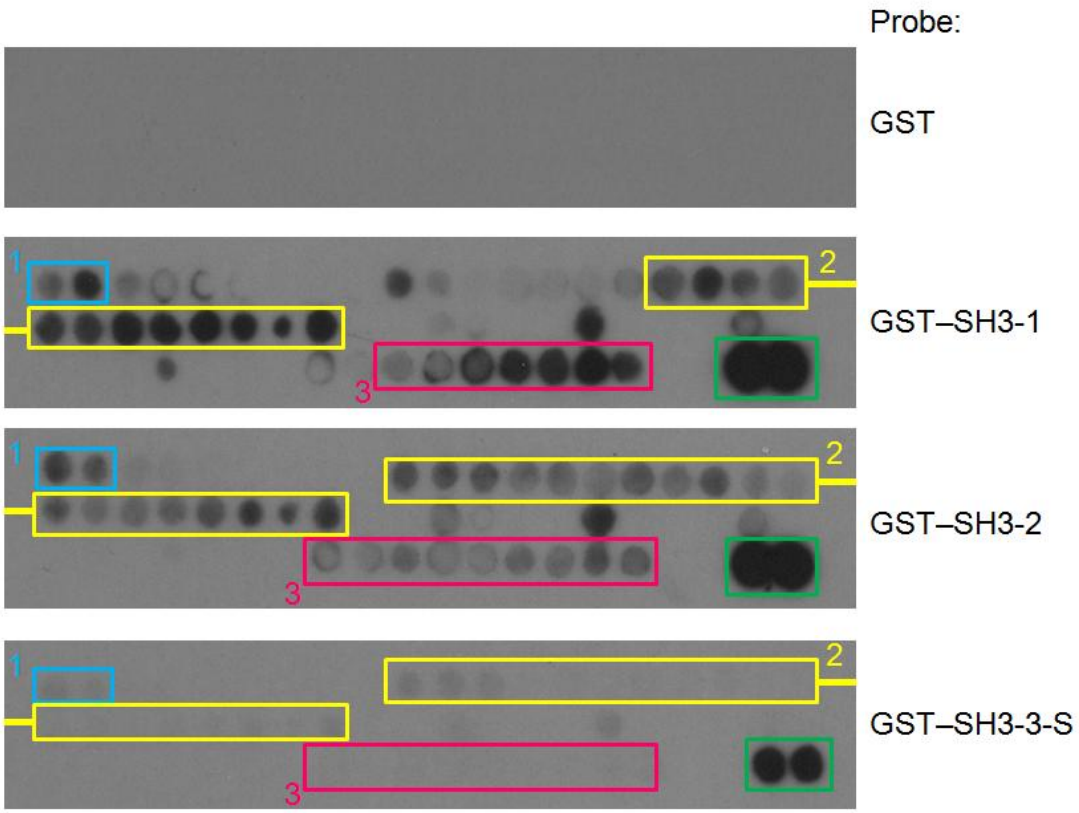


Figure 4.4

The first two SH3 domains recognise the CD2AP proline-rich region with low affinity.

The CD2AP proline-rich region (aa 330-420) was spot-synthesized as overlapping 27 aa peptides, sliding 3 aa with each step as described in section 2.2.8.3.2. The arrays were initially probed with GST (top panel) and then with GST-CD2AP-SH3-1, -SH3-2 or -SH3-3-S as described in the same section. Three potential binding regions are evident (boxed in cyan, yellow and pink). The positive control, which is duplicate spots of RIN3e2-derived peptide, is boxed in green. Exposure time is 1 minute.

4.7 Conclusion

Studying the interaction patterns of individual SH3 domains enables a more confident analysis of their binding properties compared to their investigation within a complex signal transduction network (Liu *et al*. 2012). However, the subsequent analysis of multi-domain systems is equally important for understanding the biological impact of the domain-mediated interactions (Randles *et al*. 2012).

The interactions between the first two SH3 domains and the individual RIN3e1 and RIN3e2 were analysed by ITC. Based on the results obtained, it seemed that there were not any major avidity effects and the thermodynamic signature of the SH3-(1+2) with both peptides was similar to the observed signatures with the individual SH3 domains. SH3-(1+2) was found to bind to either one RIN3e1 or possibly two RIN3e2. Such differences cannot be understood based on the structural and biophysical information of the individual SH3 domains.

The binding mode of the first two SH3 domains of CIN85 (i.e. CIN85-SH3-[1+2]), which is a CD2AP homolog, to c-CBL- and CBL-B-derived peptides was studied in isolation and in tandem. It was found that the first two individual CIN85 SH3 domains bound to one peptide, and CIN85-SH3-(1+2) bound to two peptides (Ababou *et al*. 2009). This might indicate that the first two CD2AP and CIN85 SH3 domains behave differently in tandem upon interaction with their partners.

The generation of stable wild-type and mutant RIN3-(1+2) constructs will allow characterisation of the key determinants of the CD2AP-RIN3 interaction by ITC and crystallography. Choosing the boundaries of the RIN3-(1+2) so that they include aromatic residues might be a good approach for RIN3-(1+2) to be detected by UV spectroscopy. However, this approach is not applicable if RIN3-(1+2) is degraded after cleavage. Moreover, generating RIN3-(1+2) constructs, which include globular domains of RIN3, might enhance its stability. If the crystallography experiments are not successful, other methods, such as NMR could be used to visualise the interaction. Another approach would have been to synthesise RIN3-(1+2) peptides. However, this might be

difficult due to its length (90 residues). Alternatively, linkers of different lengths could be synthesised to bridge the two RIN3 epitopes in an artificial construct.

Due to the successful purification of SH3-(1+2), a study of constructs of all three SH3 domains or the last two SH3 domains could be attempted by the above mentioned methods. Furthermore, it would be interesting to test whether the linker regions between the SH3 domains and RIN3 epitopes are important for the CD2AP interaction with RIN3. Also, generating CD2AP SH3 constructs by switching their positions within their sequence (i.e. for example, switch the positions of SH3-1 and SH3-3) and performing biophysical and structural studies will help to delineate whether this will affect their interaction with RIN3 or whether these SH3 domains complement each other.

This chapter describes work carried out in the final year of my DPhil and lack of time prevented these ideas from being pursued.

Chapter 5: Analysis of the CD2AP/RIN3 interaction by biochemical experiments

5.1 Overview

As mentioned in section 1.6.4, the CD2AP interaction with RIN3 was confirmed by co-immunoprecipitation of co-transfected flag-CD2AP and myc-RIN3 in HEK293T cells (Dr Kathrin Kirsch *et al*., Boston University, USA). Also, Chapter 3 described the molecular details of this interaction, which were determined by ITC, peptide arrays and protein crystallography. We continued by studying the endogenous CD2AP and RIN3 protein levels in human immortalised cell lines, which were derived from different tissues, and investigated their interaction at the endogenous level.

The CD2AP expression levels and pattern are well studied (Tienari *et al*. 2005, Lehtonen *et al*. 2008, Rizvi *et al*. 2012). When our study began, information on RIN3 expression was only available at the transcript level. At this level, human RIN3 was found to be ubiquitously expressed, but at higher levels in blood cells (Saito *et al*. 2002, Kajiho *et al*. 2003, Bliss *et al*. 2006). During our study, the RIN3 protein expression was studied across different haematopoietic cell lines and found to be highly enriched in mast cells. Low or undetectable levels were exhibited in the other tested haematopoietic cell lines (Janson *et al*. 2012).

As expected, CD2AP was ubiquitously expressed and its intracellular distribution was also investigated. Unfortunately, the commercially available anti-RIN3 antibodies were non-specific, which made the study at the endogenous level initially impossible. At later stages of this study, the

laboratory of Dr John Colicelli (University of California Los Angeles, USA) generated their own anti-RIN3 antibody (Janson *et al*. 2012), which they kindly provided. This antibody could be used for precipitation assays (precipitation with GST-fusion proteins, immunoprecipitation), but not for immunoblotting of total protein levels.

5.2 Analysing the CD2AP protein expression

The CD2AP expression in human immortalised cell lines, which were derived from different tissues, was studied by immunoblotting with an anti-CD2AP antibody.

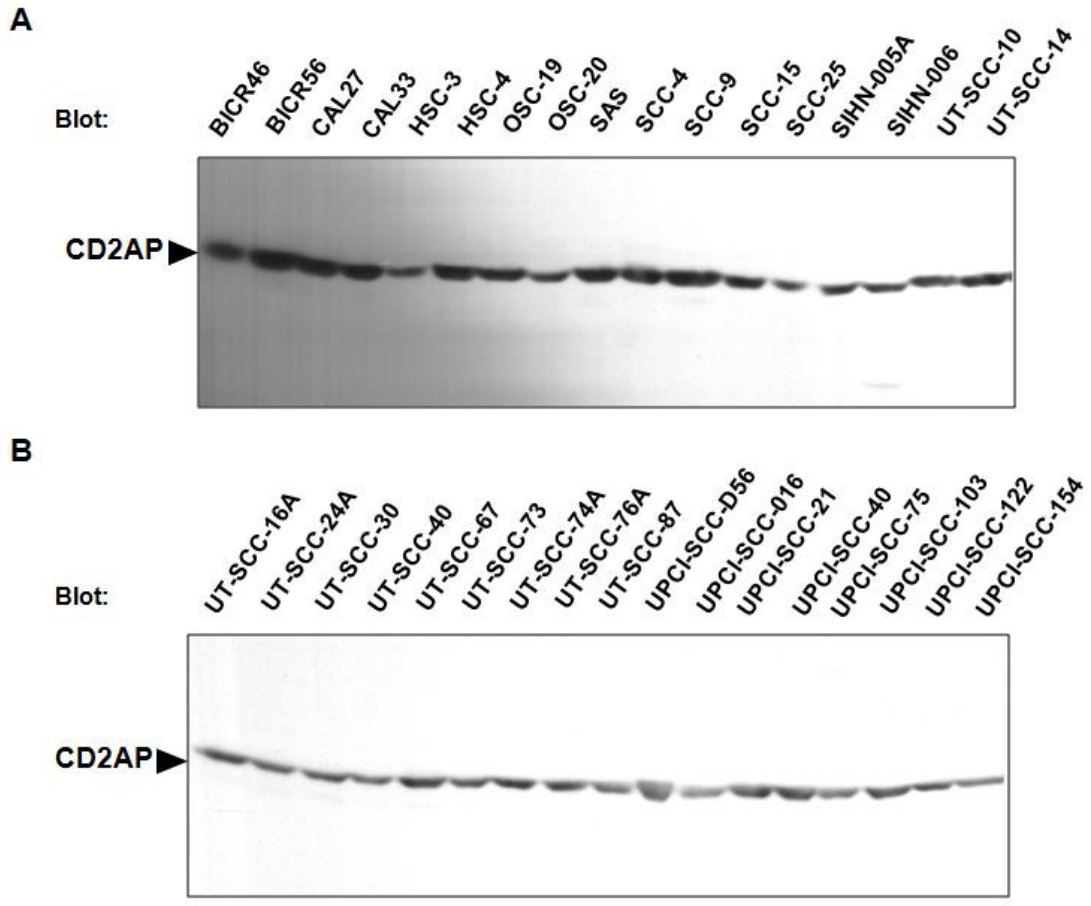


Figure 5.1

CD2AP expression levels in a panel of human head and neck cancer squamous epithelial cell lines.

According to protein quantification by Bradford, 100 µg of cell lysate/cell line was loaded.

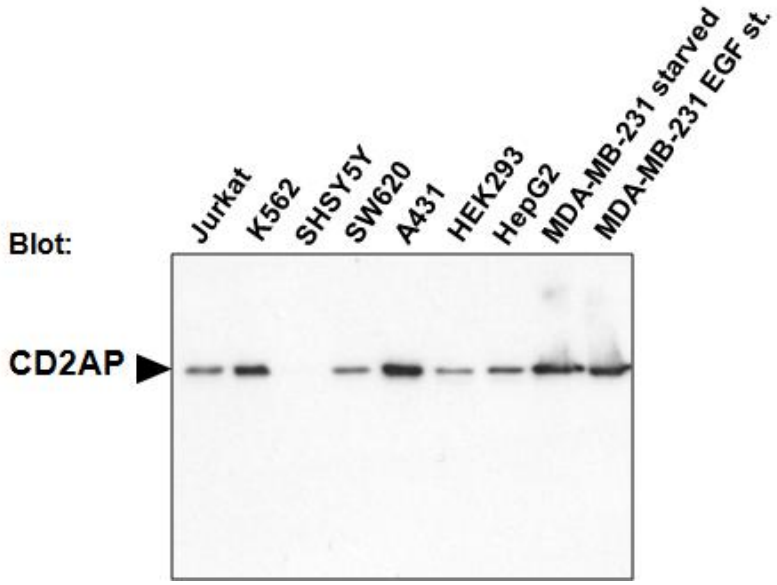


Figure 5.2

CD2AP expression levels in human immortalised cell line extracts.

According to protein quantification by Bradford, 100 µg of cell lysate/cell line was loaded. MDA-MB-231 cells were either starved or stimulated with epidermal growth factor (EGF) for 10 minutes prior to cell lysis.

The chosen cell lines were a panel of 34 head and neck cancer squamous epithelial cell lines and selected T cell (Jurkat), CML (K562), neuron (SHSY5Y), colon (SW620), skin (A431), kidney (HEK293), liver (HepG2) and breast (MDA-MB-231) cancer lines (see section 2.1.3 for more details). Figures 5.1 and 5.2 depict representative blots. Apart from SHSY5Y, CD2AP was expressed in all cell lines, with somewhat higher levels in A431, K562 and MDA-MB-231 cells. Even though CD2AP has been found in sympathetic neurons (Tsui *et al*. 2008, Medway *et al*. 2013), it was not detected in the SHSY5Y neuroblastoma cell line. This cell line was derived from the neural crest, which generates cell types of the sympathetic nervous system as well (Hammerling *et al*. 1987).

Next, the CD2AP intracellular localization was studied by subcellular fractionation in HepG2 and UPCI-SCC-154 (head and neck cancer cell line) cell extracts. The protocol is described in section 2.2.2.7. Equal volumes of the final soluble fraction (S100) and the RIPA 100 extracted pellets of the 1000 x *g* (P1), 3000 x *g* (P3), 10000 x *g* (P10) and 100.000 x *g* (P100) spins were analysed by immunoblotting with an anti-CD2AP antibody (Figure 5.3).

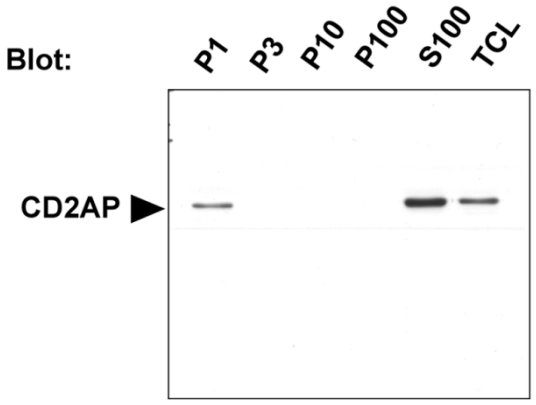


Figure 5.3

Intracellular localization of CD2AP in HepG2 cells.

P1, P3, P10, P100 refer to the extracted pellets of the 1000, 3000, 10.000 and 100.000 x *g* spins. S100 denotes the retrieved supernatant after the 100.000 x *g* spin.

CD2AP was mainly found in the cytoplasm, while a limited amount was present in the P1 fraction. This is in accordance with microscopy and immunohistochemistry studies in other cell lines, which reported that it is found mainly in the cytoplasm. However, these studies found CD2AP at the leading edge of the cell as well (Gaidos *et al*. 2007, van Duijn *et al*. 2010, Tang *et al*. 2013). The mitochondrial and membrane fractions (P3, P100) had negligible CD2AP.

Finally, CD2AP expression was investigated in an immortalised human podocytic cell line. As already described in section 1.5.2.1, CD2AP is important in podocytes for kidney development. Podocytes were kindly provided by Prof. Moin Saleem (University of Bristol, UK). Details on the culture conditions can be found in section 2.2.2.2. CD2AP expression was studied by immunoblotting in the undifferentiated state and at various time points of differentiation (days 1, 3, 6, 10 and 14). Similar CD2AP levels were observed in the proliferative state and throughout different stages of differentiation (data not shown), which is consistent with the literature (Saleem *et al*. 2002).

5.3 CD2AP precipitation with the GST-RIN3-(1+2) fragment

As described in section 4.3, the expression and purification of the GST-RIN3-(1+2), which consists of the two SH3 binding regions in RIN3, was successful.

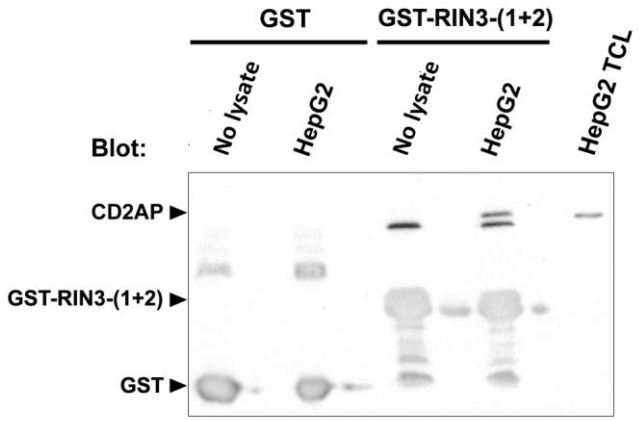


Figure 5.4

GST-RIN3-(1+2) can bind to CD2AP *in vitro*.

1 mg of HepG2 total cell lysate (TCL) was precipitated with 50 µg of bead-immobilised GST or GST-RIN3-(1+2). Precipitates were subjected to immunoblotting with an anti-CD2AP antibody.

This fragment was used to precipitate endogenous CD2AP from HepG2 cell extracts (Figure 5.4). The protocol is detailed in section 2.2.3.1. The precipitation assay indicates that CD2AP and RIN3 form a complex. As shown in Figure 5.4, apart from the CD2AP band, the anti-CD2AP antibody recognised other bands non-specifically, corresponding to GST and GST-RIN3-(1+2).

5.4 Analysing the RIN3 protein expression with commercial antibodies

A panel of anti-RIN3 antibodies was tested to find the appropriate antibody for further studies at the endogenous level. At the beginning of this thesis research, there was not any published literature with anti-RIN3 antibodies. According to the Uniprot database (Uniprot Consortium 2014), there are two putative RIN3 isoforms (108 and 99 kDa) produced by alternative splicing. Both of them contain the two RIN3 epitopes that are recognised by the SH3 domains. Details on the tested RIN3 antibodies are found in section 2.1.1. In summary, the panel consisted of the following anti-RIN3 antibodies: sc-102089 (Santa Cruz Biotechnology), AV34618 (Sigma Aldrich), SAB4503182 (Sigma Aldrich) and 12709-1-AP (Protein Tech).

5.4.1 Testing a panel of anti-RIN3 antibodies by immunoblotting

Initially, we wanted to study the RIN3 expression across different tissues, in which CD2AP is ubiquitously expressed. So, extracts from the above mentioned representative cell line panel (section 5.2, Figure 5.2) were chosen to be tested by immunoblotting with the panel of anti-RIN3 antibodies (AV34618, sc-102089, SAB4503182 and 12709-1-AP) (Figure 5.5). Due to

the limited expression of CD2AP in the SHSY5Y cell line (Figure 5.2), the SHSY5Y cell extract was substituted by the HeLa cell extract, which is representative of cervix tissue. These experiments were done in duplicate for each antibody.

The anti-RIN3 antibodies showed differential non-specific profiles. In the case of AV34618, weak bands of predicted size (100-108 kDa) were found across the cell line panel except Jurkat and HeLa cells, but additional bands of higher and similar intensity were also present. As illustrated in Figure 5.5, a similar immunoblotting profile was obtained with sc-102089 compared to AV34618.

Apart from the Jurkat lysates, circa 102 kDa weak bands and multiple bands of higher or similar intensity were observed. A different immunoblotting profile was observed for SAB4503182. A band corresponding to the predicted molecular weight, was found across the different cell lines. However, bands of similar or lower intensity were also found for all lysates apart from Jurkat. A number of bands of higher or similar intensity except the predicted molecular weight bands were observed for 12709-1-AP (Figure 5.5).

Even when higher primary or secondary antibody amounts were used, the antibody sensitivity did not improve. This approach was not successful (data not shown). The blocking condition was 5% (w/v) non-fat dry milk in TBST. Different blocking conditions (5% [w/v] non-fat dry milk in PBST, 3% [w/v] BSA in TBST or PBST, *E.coli* lysate with 1% Triton X-100 in TBST) were also tested to possibly improve antibody specificity, but there was none (data not shown).

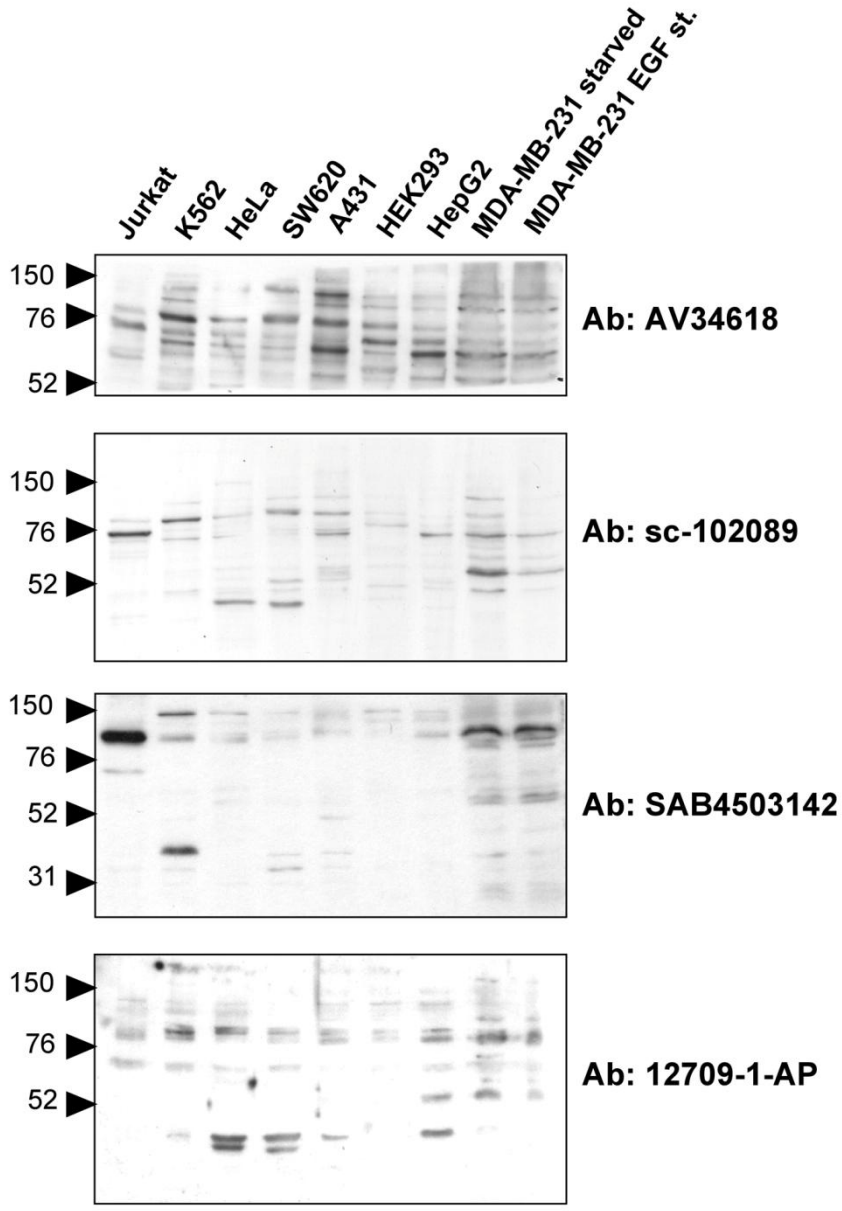


Figure 5.5

The panel of four commercially available RIN3 antibodies was tested by immunoblotting of human immortalised cell extracts.

The tested antibodies are noted on the right of each blot.

In summary, the immunoblotting experiments did not identify a suitable anti-RIN3 antibody. AV34618 and sc-102089 shared a similar non-specific profile. 12709-1-AP appeared to be non-specific too. SAB4503182 seemed promising due to the presence of a band at the predicted molecular weight for

RIN3, but the presence of other bands of similar intensity implied that it was not very specific. The variation of the band sizes across the different cell lines and antibodies was also worrying. Due to the limited literature, we were uncertain whether this variation was due to any post-translational modifications or additional splice variants. Therefore, these antibodies need further validation, for example with siRNA knockdowns.

One alternative approach tested was to enrich subcellular fractions prior to immunoblotting. As a result, the immunoblotting profile might have improved due to potential RIN3 enrichment or reduction of the non-specific bands in certain fractions. The RIN3 localization profile has not been extensively studied. RIN3 was reported to be distributed in the cytoplasm and to translocate to Rab5-positive vesicles in HeLa cells by tyrosine phosphorylation signals (Kajiho *et al* 2003, Yoshikawa *et al* 2008). HepG2 cells were subjected to subcellular fractionation. Immunoblotting was performed on the P1, P3, P10, P100 and S100 fractions with the commercially available RIN3 antibodies. Bands of different sizes were found in the different fractions (data not shown). Therefore, in the absence of protein expression knockdowns, the antibody specificity could not be determined.

5.4.2 Identification of the anti-RIN3 antibody epitopes by RIN3 peptide scanning arrays

The anti-RIN3 antibodies were also validated by studying whether they recognise RIN3. AV34618 and 12709-1-AP were raised against the RIN3 N-terminal (aa 72-121) and C-terminal regions (GST fusion protein containing 350 residues) respectively. However, the sc-102089 and SAB4503182 RIN3 epitopes were unknown. In order to identify them, a RIN3 scanning array was

probed with sc-102089 (Figure 5.6A), stripped successfully (Figure 5.6B) and re-probed with SAB4503182 (Figure 5.6C). The stripping efficiency was tested by probing the array with secondary antibody. The probing and stripping protocols can be found in sections 2.2.8.3.1 and 2.2.8.3.3. This approach enables the identification of linear, but not conformational antibody epitopes (Reineke *et al*. 2009).

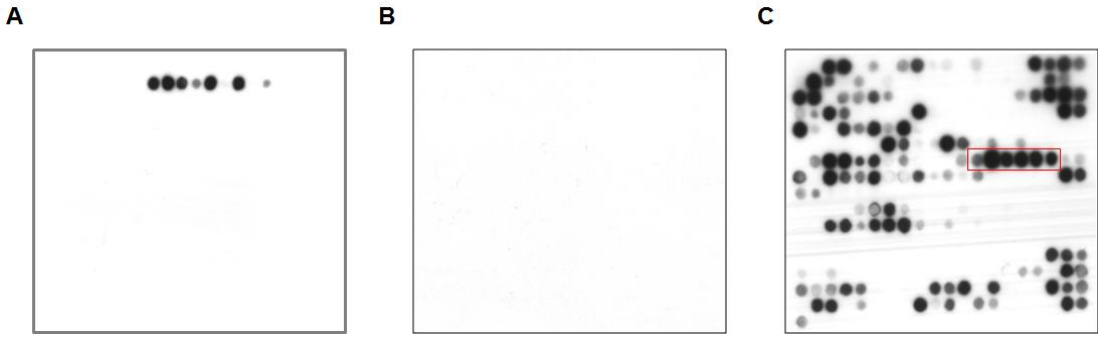


Figure 5.6

Identification of the sc-102089 and SAB4503182 epitopes in RIN3.

A RIN3 peptide scanning array was probed with sc-102089 **(*A*)**, stripped successfully **(*B*)** and re-probed with SAB4503182 **(*C*)**. The extended RIN3 epitope, which was recognised by SAB4503182, is shown within the red box.

As shown in Figure 5.6A, sc-102089 recognised a single epitope in RIN3 (aa 85-122). The minimal RIN3 segment common to all peptide spots (aa 97-111) that was recognised by sc-102089, was also present in the AV34618 epitope. SAB4503182 recognised an extensive RIN3 epitope (aa 412-426). Other small RIN3 epitopes were also identified (Figure 5.6C, extensive epitope within red box).

In summary, apart from AV34618 and sc-102089, there were differences in the linear RIN3 epitopes that the panel of anti-RIN3 antibodies recognised. Even though AV34618 and sc-102089 are rabbit polyclonal antibodies, they

shared a single similar epitope. Due to the similarity of the AV34618 and sc-102089 immunoblotting profiles and the recognition of the similar RIN3 epitopes, AV34618 was not used for the remaining experiments in this thesis. The SAB4503182 antibody recognised more RIN3 epitopes compared to AV34618 and sc-102089.

5.4.3 Testing the panel of anti-RIN3 antibodies by siRNA and immunoblotting

As described in section 5.4.1, the immunoblotting experiments were inconclusive for the quality of the commercial anti-RIN3 antibodies. However, these antibodies recognised linear RIN3 epitopes (section 5.4.2). Therefore, the anti-RIN3 antibodies need to be further validated.

Another approach to validate the sc-102089, SAB4503182 and 12709-1-AP anti-RIN3 antibodies was by siRNA. RIN3 siRNA treatment and immunoblotting with the anti-RIN3 antibodies would lead to the disappearance of the band that corresponds to RIN3. This would indicate the anti-RIN3 antibody that recognises RIN3. At the time of the study, there was no validated RIN3 siRNA available. A RIN3 siRNA kit containing four different siRNAs was purchased. Details of this kit, and the positive and negative controls tested are found in section 2.2.3.3.

Initially, HEK293 and HepG2 cells were transfected with the individual RIN3 siRNAs, a positive control (p130Cas siRNA) or a negative control (GL2 siRNA). Dharmafect 1 and 4 were used as transfection reagents for HEK293 and HepG2 cells respectively. The cells were seeded in wells one day prior to transfection with the pre-formed complexes of the siRNA and the transfection

reagent (starting protocol). Cells were lysed 48 h after transfection. The HepG2 lysates, which were blotted against p130Cas (Figure 5.7A) and RIN3 with sc-102089 (Figure 5.7B) or SAB4503182 (Figure 5.7C) are shown. The positive control was successful (32% reduction, Figure 5.7A).

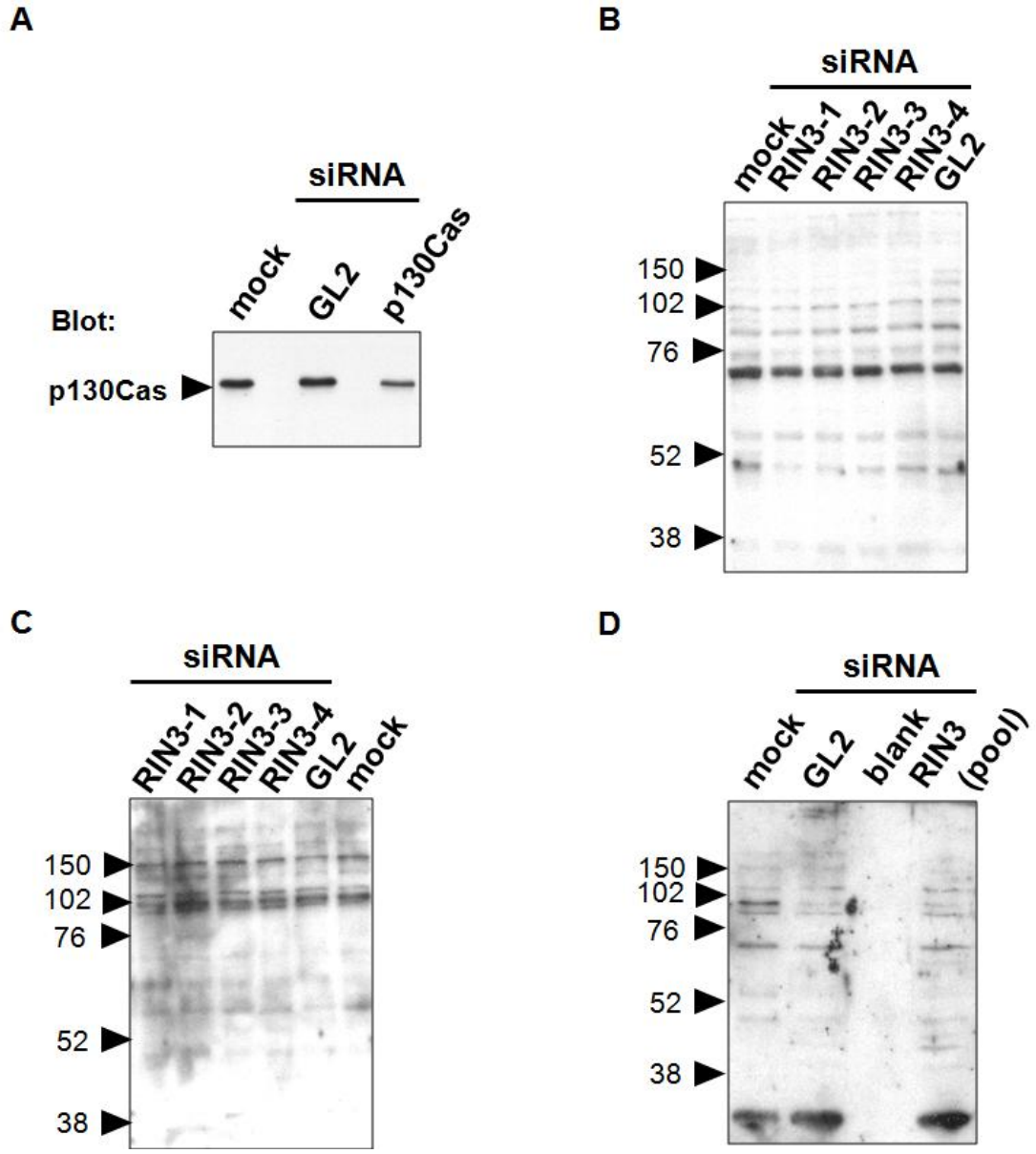


Figure 5.7

The sc-102089, SAB4503182 and 12709-1-AP anti-RIN3 antibodies were tested by siRNA.

A, HepG2 cells were successfully transfected with a p130Cas siRNA; HepG2 cells were transfected with the four individual RIN3 siRNAs (RIN3-1 to -4) and probed with: ***B***, sc-102089 and ***C***, SAB4503182; ***D***, HepG2 cells were transfected with a pool of the four RIN3 siRNAs and probed with 12709-1-AP. GL2 stands for the negative control. Mock denotes cells treated with the transfection reagent.

As shown by immunoblotting with sc-102089 (Figure 5.7B) or SAB4503182 (Figure 5.7C), the band intensities between the individual RIN3 siRNA, mock- and GL2-transfected lanes remained unchanged. Similar results were obtained after immunoblotting with AV34618 (data not shown). In an attempt to improve the HEK293 or HepG2 transfection protocol, the four siRNAs (pool) were used in a simultaneous transfection. Also, the pre-formed complexes of siRNA and transfection reagent were added to the wells prior to cell addition (reverse transfection protocol). The band intensities of the GL2- and pool-transfected HepG2 cells did not change upon immunoblotting with 12709-1-AP (Figure 5.7D). Similar results were obtained with sc-102089 and SAB4503182 (data not shown).

In order to optimise the RIN3 siRNA transfection, six different transfection reagents (Dharmafect 1, 2 and 4, Oligofectamine, HiPerFectamine and Lipofectamine 2000), three different post-transfection cell lysis time-points (24, 48 and 72 h) and two different media (Opti-MEM and DMEM) were attempted in HEK293 and HepG2 cells. Lastly, the reverse transfection protocol was coupled to transfection with equal amounts of pre-formed complexes of siRNA and transfection reagent after 24 h of initial transfection (double transfection reverse protocol). In all cases, immunoblotting of HEK293 or HepG2 RIN3 siRNA transfected cells with sc-102089, SAB4503182 or 12709-1-AP did not affect the intensity of any band. These results indicate that the RIN3 antibodies, siRNAs, or both, do not recognise RIN3.

5.4.4 Testing the panel of anti-RIN3 antibodies by precipitation with GST-SH3 fusion proteins

Another approach taken to evaluate the anti-RIN3 antibodies was increasing the RIN3 amounts prior to immunoblotting by pulldowns with the isolated GST-SH3 or GST-SH3-(1+2) domains from HepG2 extracts. A peptide competition experiment was set up in parallel by pre-incubating GST-SH3 domains with the RIN3e2 $I_{466}A$ or $R_{462}A$ peptides. According to ITC measurements, $I_{466}A$ binds to the SH3 domains with a similar affinity as the wild-type peptide (section 3.6.3), while the SH3 affinity to $R_{462}A$ was too low to be quantified (Table 3.2). Therefore, $I_{466}A$ would potentially compete for and block the RIN3 binding site, while $R_{462}A$ is a negative control. The protocol is described in section 2.2.3.1. The SH3 precipitated proteins were blotted against RIN3 with sc-102089, SAB4503182 and 12709-1-AP.

Unfortunately, SAB4503182 could not be validated by this method due to high background signal. The reduction of the background signal was attempted by trying different blocking conditions and adding fewer GSH beads and/or less cell lysate, but it remained unchanged (data not shown).

Figure 5.8 depicts the sc-102089 (A) and 12709-1-AP (B) blots. Immunoblotting with sc-102089 revealed that SH3-1, SH3-2 and SH3-(1+2) pulled down a number of bands, whose molecular weights extended between circa 102 and 225 kDa (Figure 5.8A). The profiles for SH3-(1+2) and the individual domains were different. Also, circa 40 kDa bands were detected due to the GST-SH3-1 and -2 domains. Lastly, when GST-SH3-1 was pre-incubated with the $I_{466}A$ peptide ('I466A' lane in Figure 5.8A) all detected bands disappeared.

Figure 5.8B depicts a similar experimental set-up to Figure 5.7A. SH3-1 and SH3-(1+2) were used for the pulldowns and $I_{466}A$ and $R_{462}A$ peptide competition experiments. The blot was probed with 12709-1-AP. SH3-1 and SH3-(1+2) pulled down a number of bands, whose molecular weights ranged between 76 and 225 kDa. These bands were not visible when SH3-1 and SH3-(1+2) were incubated with the competitive binding peptide ('I466A' lane in Figure 5.8B).

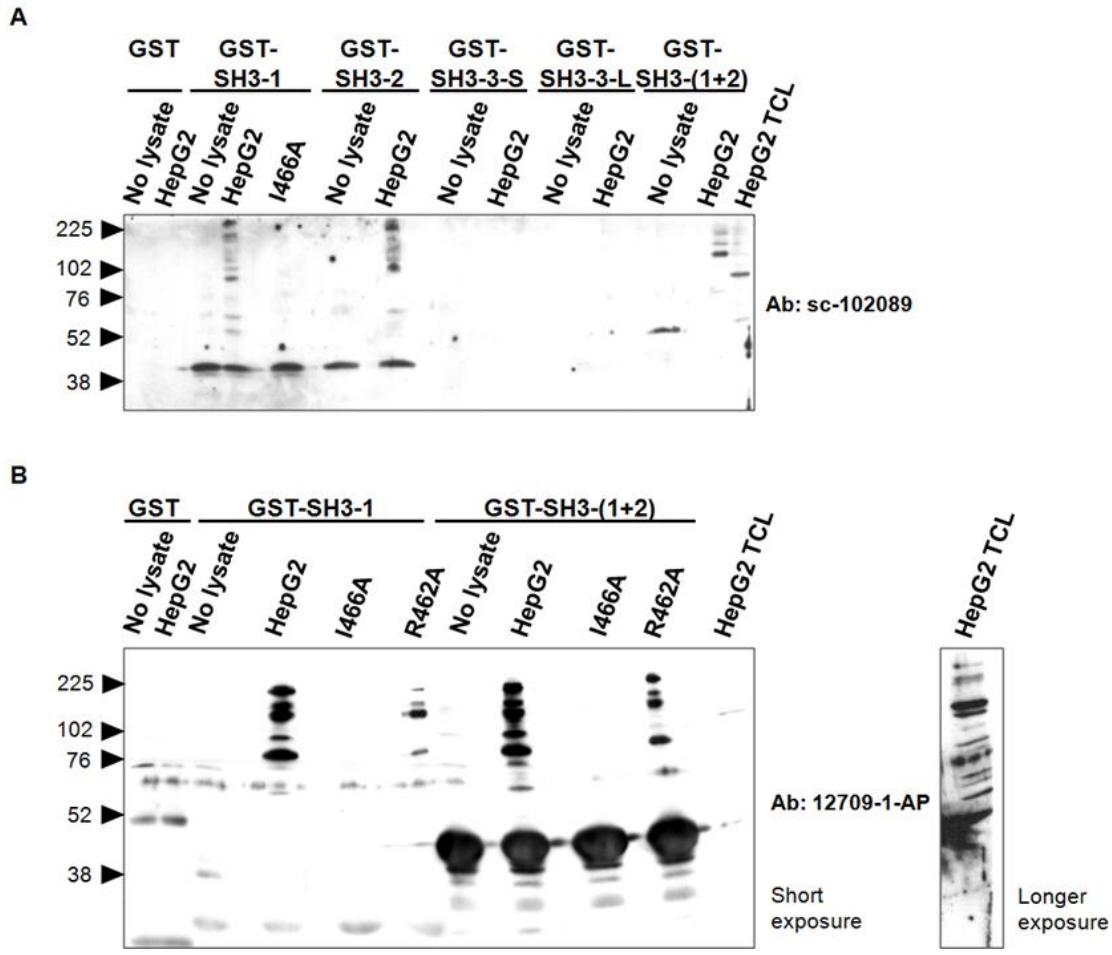


Figure 5.8

Studying the sc-102089 and 12709-1-AP anti-RIN3 antibodies by pulldown with GST-CD2AP-SH3 domains.

A, 1 mg of HepG2 cell lysate was precipitated with 50 µg of bead-immobilised GST, isolated GST-SH3 or GST-SH3-(1+2) proteins, precipitates were subjected to immunoblotting with sc-102089. ***B***, 1 mg of HepG2 cell lysate was precipitated with 50 µg of bead-immobilised GST, GST-SH3-1 or GST-SH3-(1+2) proteins, precipitates were subjected to immunoblotting with 12709-1-AP.

When the same domains were pre-incubated with the negative control peptide, some lower intensity bands were again visible ('R462A' lane in Figure 5.8B). Roughly 40 kDa (burn-through) bands were detected due to the presence of GST-SH3-(1+2).

Finally, some bands might be present due to RIN3 degradation. Therefore, the lysates were incubated for 2 h instead of overnight with the GST-SH3 domains. The immunoblotting profile for sc-102089 was similar to Figure 5.8 (data not shown). In summary, these results indicated that sc-102089 and 12709-1-AP may not be specific for RIN3, but might recognise additional proteins that interact with RIN3, or also bind to the CD2AP SH3 domains. Alternatively, a range of constitutive splice variants, degradation products and/or unexpected modification products of unknown origin might be present, making these results difficult to interpret.

5.5 Studying RIN3 protein expression with a custom antibody

An anti-RIN3 antibody was generated by the laboratory of Dr John Colicelli. This antibody was previously validated with a RIN3 siRNA using the ON-TARGETplus system from Dharmacon by that lab (Janson *et al*. 2012).

5.5.1 Identification of the RIN3 epitope by peptide scanning array

Initially, the RIN3 epitopes, which were recognised by the Colicelli lab antibody, were identified as described before (section 5.4.2). This was done to confirm that this antibody recognised RIN3. The RIN3 scanning array that had been probed with the SAB4503182 was stripped successfully (Figure 5.9A). Then, the array was re-probed with the Colicelli lab anti-RIN3 antibody, which

recognised a number of linear epitopes (Figure 5.9B). The three major epitopes (aa 298-321, 355-369, 856-873) are boxed in red (Figure 5.9B).

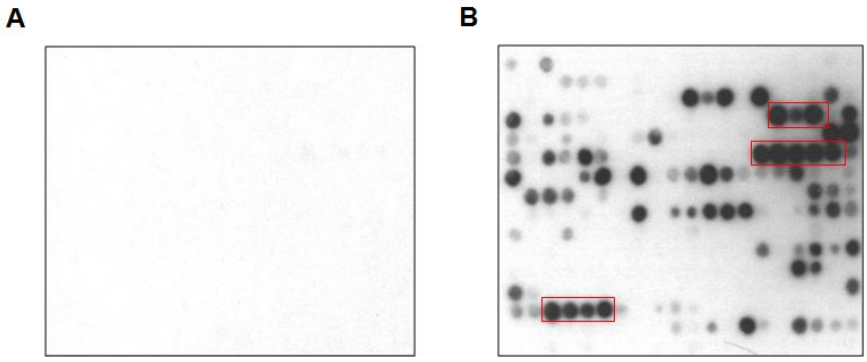


Figure 5.9

Identification of the RIN3 epitopes recognised by the Colicelli lab antibody.

The RIN3 peptide scanning array (Figure 5.5C) was stripped successfully (***A***) and re-probed with the Colicelli lab anti-RIN3 (***B***). The major RIN3 epitopes are enclosed in red boxes.

5.5.2 Antibody testing by immunoblotting

The above mentioned cell panel (section 5.4.1, Figure 5.2) was tested by immunoblotting with the Colicelli lab anti-RIN3 antibody. Unfortunately, the background signal was high (data not shown). The blocking condition was 5% (w/v) non-fat dry milk in TBST. Optimisation of this assay was attempted using different blocking conditions (5% [w/v] non-fat dry milk in PBST, 3% [w/v] BSA in TBST or PBST, *E.coli* lysate with 1% Triton X-100 in TBST) and different amounts of primary or secondary antibodies. 5% (w/v) non-fat dry milk in PBST reduced slightly the background signal and was used as the blocking buffer in the subsequent immunoblotting assays with this antibody. Despite the slight reduction of the background signal, it was still significant (data not shown). Therefore, we sought other cell lines to test this antibody. According

to the Gene Expression Omnibus database (Edgar *et al*. 2002, Barett *et al*. 2013), high RIN3 transcript levels are found in squamous cell carcinomas and the A549 cell line. Therefore, lysates of the panel of 34 head and neck cancer squamous and A549 cell lines were blotted against RIN3 with this antibody. A representative blot of some of them is shown in Figure 5.10. Multiple bands of different sizes and a high background signal were observed. However, a higher intensity band of higher molecular weight than the predicted one (shown with red arrow in Figure 5.10) was observed for some cell lines (UT-SCC-40, UT-SCC-67, UPCI-SCC-103, UPCI-SCC-154-coloured in red). Apart from this blot, a similar band was observed for HSC-4, SCC-9 and UT-SCC-10 and A549 cell lines (data not shown).

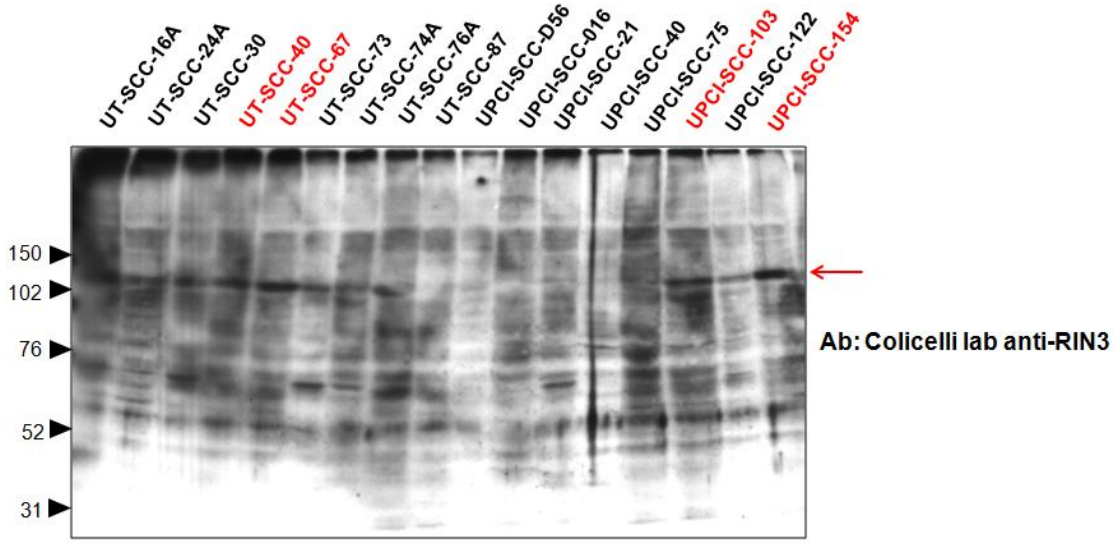


Figure 5.10

Head and neck cancer cell lysates were immunoblotted with the Colicelli lab anti-RIN3 antibody.

The name of each cell line is described at the top of the blot. The suggested RIN3 band is shown by the red arrow on the right of the blot.

The UPCI-SCC-154 cell line was chosen for further experiments. This cell line was chosen, because the band, which might correspond to RIN3, is of high intensity (red arrow, Figure 5.10). Lysates of high (RIPA 100) or medium stringency (1% Triton X-100 buffer) were prepared. When these UPCI-SCC-154 lysates were immunoblotted with the Colicelli lab anti-RIN3 antibody, the background signal was reduced. Apart from the previously mentioned high intensity band, circa 60 kDa bands were also detected (Figure 5.11). The profile differences of Figures 5.10 and 5.11 are concerning. High background signal was also present when recently prepared lysates of other cell lines, such as HepG2 and HEK293, were probed with the Colicelli lab anti-RIN3 antibody (data not shown).

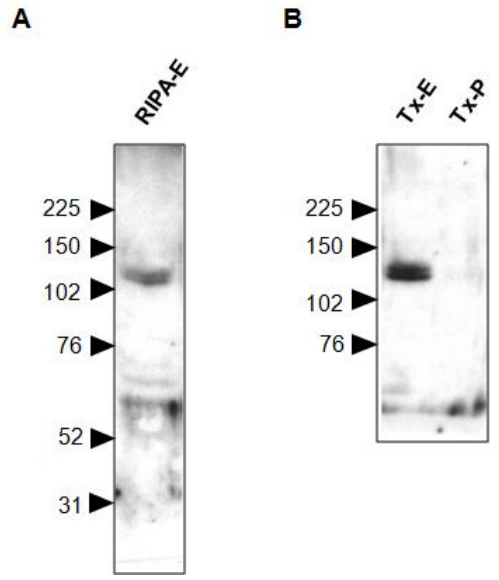


Figure 5.11

UPCI-SCC-154 cell lysates were immunoblotted with the Colicelli lab anti-RIN3 antibody.

Proteins were extracted with ***A***, RIPA 100 (RIPA-E) or ***B***, 1% Triton X-100 buffer (Tx-E). Proteins of the pellet (Tx-P) were extracted with RIPA 100.

Finally, RIN3 expression was investigated in undifferentiated and differentiated podocytic cell extracts. According to immunoblotting with the Colicelli lab anti-RIN3 antibody, RIN3 had very low and low expression levels in the undifferentiated and differentiated states respectively (data not shown).

5.5.3 Antibody testing by precipitation with GST-SH3 domains

In order to investigate whether the above-mentioned band corresponded to RIN3, precipitation assays with GST-SH3 domains were performed on differentiated podocytic (days 6 and 10 of differentiation) cell extracts and immunoblotted with the Colicelli lab anti-RIN3 antibody. As shown in Figure 5.12, a double band was detected, which might represent the two RIN3 isoforms or post-translationally modified forms. This finding implied that the Colicelli lab anti-RIN3 antibody was more specific than the tested commercial antibodies. However, it recognised other non-specific bands when it was used to blot against the total RIN3 levels. This experiment was also repeated on UPCI-SCC-154, UT-SCC-10 and differentiated podocytic (day 14 of differentiation) cell extracts and the results were consistent with Figure 5.12. In the case of UT-SCC-10, a $I_{466}A$ peptide competition experiment was set up as described above (section 5.4.4), which resulted in the disappearance of the double band.

Pulldown experiments with GST-SH3-1 were also successful in MOLT-4, HL-60, UT-SCC-40, UT-SCC-67, UPCI-SCC-103, HSC-4 and SCC-9 lysates (data not shown). A similar experiment was set up with the A549 cell line, but was unsuccessful.

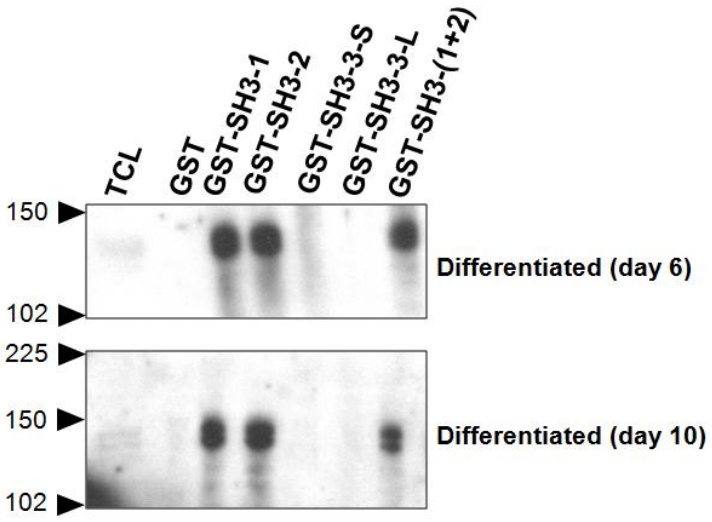


Figure 5.12

The Colicelli lab anti-RIN3 antibody was tested by pulldown assays on day 6 (upper panel) and day 10 (lower panel) differentiated podocytic cell extracts.

1 mg of lysates from differentiated cells were precipitated with 50 µg of bead-immobilised GST, isolated GST-SH3 or GST-SH3-(1+2) proteins; precipitates were subjected to immunoblotting with the Colicelli lab anti-RIN3 antibody.

5.6 Attempts to study the CD2AP/RIN3 interaction by co-immunoprecipitation

After the successful RIN3 precipitation with the GST-SH3 domains, we attempted to confirm the CD2AP/RIN3 interaction by co-immunoprecipitation. Initially, it was sought to precipitate the complex from UPCI-SCC-154 cell extracts (2 mg/assay) by overnight incubation with an anti-CD2AP or the Colicelli lab anti-RIN3 antibody. Even though the antibodies were able to precipitate the corresponding proteins, immunoblotting of the binding partner was unsuccessful (Figure 5.13). Different cell (UPCI-SCC-154, SCC-9, differentiated podocytes [day 6, 10, 14 of differentiation]) extracts, extraction buffers (1% Triton X-100 buffer, Nonident P-40 buffer) and anti-CD2AP antibodies were tried. Different beads to precipitate the antibodies (A or G sepharose) and incubation times (2 h, 4 h or overnight) were also tested.

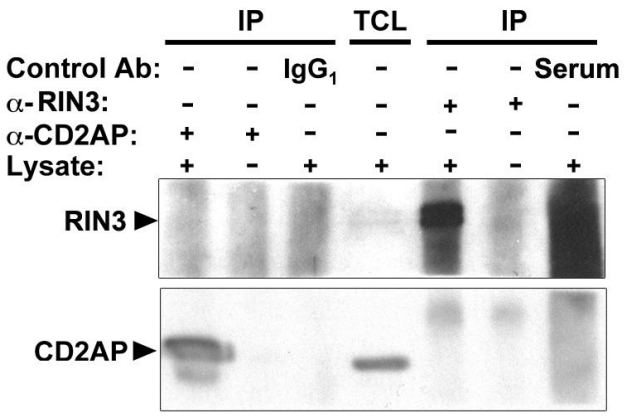


Figure 5.13

Attempt to co-immunoprecipitate the CD2AP/RIN3 complex from UPCI-SCC-154 cell extracts with an anti-CD2AP antibody and the Colicelli lab anti-RIN3 antibody.

Lower panel: Western blot for CD2AP; *upper panel*: Western blot for RIN3. Serum denotes pre-immune rabbit serum.

Also, the tyrosine phosphorylation dependence of the complex formation was tested by stimulating UPCI-SCC-154 cells with a vanadate/H_2O_2 mixture prior to cell lysis, but the co-immunoprecipitation assay under these conditions was not successful either. The detailed method protocol is described in section 2.2.2.3.

Lastly, subcellular fractionation was performed in UPCI-SCC-154 cell extracts and the soluble S100 fraction was resolved by size exclusion chromatography. Immunoprecipitation was attempted from fractions with CD2AP and RIN3 co-elution. Unfortunately, these attempts were also unsuccessful.

5.7 Conclusion

The findings of the CD2AP expression levels and intracellular distribution by immunoblotting of human cell lines across different tissues are in accordance

with the published literature (Saleem *et al*. 2002, Tienari *et al*. 2005, Lehtonen *et al*. 2008, Rizvi *et al*. 2012). Even though peptide array studies indicated that the panel of commercially available anti-RIN3 antibodies recognised linear RIN3 epitopes, it appears from immunoblotting and pulldown studies that these antibodies are not useful reagents. The custom Colicelli lab anti-RIN3 antibody recognised RIN3 in SH3 pulldowns and immunoprecipitation assays. However, a high background signal and the recognition of additional lower intensity bands were observed by immunoblotting against the total RIN3 levels. Also, the differences in the background signal among cell lines is concerning. According to immunoblotting results of the panel of the 34 head and neck cancer squamous epithelial cell lines with this antibody, moderate or high RIN3 levels are found in eight of the 34 cell extracts. The EGFR and c-Met expression levels of this panel have been studied (Wu *et al*. 2011), but there was no correlation between EGFR or c-Met and RIN3 expression levels. The possible involvement of RIN3 in c-KIT receptor endocytosis (Janson *et al*. 2012) led us to study the RIN3 expression levels with the Colicelli lab anti-RIN3 antibody in three colorectal cancer cell lines (HDC114, LS174T, C70), which have high c-KIT levels. However, no RIN3 was detected (data not shown). Therefore, the immunoblotting studies with the Colicelli lab anti-RIN3 antibody suggested that RIN3 had a limited expression pattern across human cancer cell lines derived from different tissues. Testing these cell lines by immunoblotting with other specific anti-RIN3 antibodies is important for the possibly limited endogenous RIN3 expression levels to be validated. Unfortunately, the lack of such tools make this study impossible so far.

The affinity pulldown studies on different cell lysates at the endogenous level with the GST-SH3 domains and GST-RIN3-(1+2) indicated that the CD2AP interaction with RIN3 was possible in principle. Unfortunately, the interaction was not confirmed by immunoprecipitation assays at the endogenous level. These results indicate that this complex might be transient, of low abundance or does not exist in all cells due to spatiotemporal restrictions.

The anti-RIN3 epitope mapping by RIN3 peptide scanning arrays enabled the identification of the dominant linear RIN3 epitopes (aa 97-111, 298-321, 355-369, 412-426, 856-873), which could be used as immunogens for production of an anti-RIN3 antibody. These residue ranges contain the minimal RIN3 segments common to all peptide spots that are within red boxes in Figures 5.6C (aa 412-426) and 5.9B (aa 298-321, 355-369, 856-873), and the single RIN3 epitope in Figure 5.6A (aa 97-111).

Finally, possible ways of studying further the CD2AP/RIN3 interaction will be proposed in section 7.2. Due to time limitations, these studies were not pursued. The next and final chapter of this thesis addresses another aim of the project, namely the identification of novel CD2AP SH3 interaction partners.

Chapter 6: Identification and analysis of novel CD2AP SH3 interacting proteins

6.1 Overview

As described in section 1.4, the SH3 binding specificity is generated by a number of mechanisms, such as positive selection (i.e. recognition of a binding motif within a ligand by an SH3 domain) and negative selection (i.e. prevention of SH3 binding to a given motif). Taking into account the positive selection determinants of the CD2AP/RIN3 interaction, the initially identified CD2AP SH3 recognition motif in c-CBL, P-x-P-x-P-R, was refined to P-x-P/A-x-p-R in the two RIN3 epitopes. This refined motif was then used to search for novel CD2AP SH3 binding candidates harbouring the motif. In order to improve the success rate of our search, we combined our knowledge of both positive and negative selection determinants of the CD2AP SH3/RIN3 interaction identified in the RIN3 peptide permutation arrays (section 3.4). Taking into account negative selection, the RIN3 permutation arrays indicated that the CD2AP SH3 recognition motif, P-x-P/A-x-p-R, may be further refined to P-{FGHWY}-P/A-{D}-{FGHWY}-R-{P} (section 3.7). The curly brackets refer to the negatively selected residues. The latter motif was used in combination with the ScanProsite bioinformatics tool (de Castro et al. 2006) and the PrePPI database (Zhang et al. 2012, Zhang et al. 2012) to search for novel putative CD2AP SH3 interaction partners. Peptide arrays were then used to screen a subset of hits.

The proteins ARAP1 (ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1), MLK3 (mitogen-activated protein kinase 11) and DAB1 (disabled

homolog 1) were among the hits chosen for further validation. SH3 binding to ARAP1, MLK3 and DAB1 peptides was studied by ITC, while a structural view of the SH3-2 domain in complex with an ARAP1 peptide was obtained by protein crystallography. This crystal structure pointed towards additional ARAP1 residues anchoring on the SH3-2 surface, the importance of which was further validated by ITC measurements with ARAP1 point mutant peptides.

Finally, the CD2AP interaction with ARAP1 and MLK3 was studied by pulldown with GST-CD2AP-SH3 domains and immunoprecipitation in a number of immortalised human cell lines derived from different tissues.

6.2 Experimental design for screening for novel CD2AP SH3 partners

Several variants of the whole interactome scanning experiment (WISE) method (Landgraf et al. 2004) have been used to decipher the human SH3 domain interactome (Wu et al. 2007, Carducci et al. 2012, Kay 2012). A variation of these methods was used in our study to screen for novel CD2AP SH3 interaction partners. Our method is summarised in Figure 6.1. Initially, the identified SH3 recognition P-{FGHWY}-P/A-{D}-{FGHWY}-R-{P} motif was used as input to the ScanProsite bioinformatics tool (release 2012_10, http://prosite.expasy.org/scanprosite/) (Obenauer et al. 2003, de Castro et al. 2006) to scan the UniprotKB/TrEMBL database for human proteins matching this motif. To clarify, hits would necessarily contain the residues P-x-P/A-x-x-R (positive selection), but would not contain those residues denoted in positions indicated by curly brackets (negative selection). This search resulted in 10960

potential binding sites within 8178 proteins making a further reduction of candidates for a more detailed analysis desirable. These potential hits were filtered by giving preference to potential binding sites within ordered regions, as only a few examples of SH3 domains binding to poly-proline type II (PPII) helices within ordered regions have been reported (Arold et al. 1997). Clearly, this introduces a substantial bias in the proteins that are further considered, but it could also lead to the preferential detection of unusual interactions.

Recently, the PrePPI database, which combines structure- and non-structure-based approaches to predict protein-protein interactions (Zhang et al. 2012, Zhang et al. 2013), was developed. The search with the PrePPI algorithm retrieved 1472 CD2AP potential binding proteins. When the PrePPI database-based hits were compared to the ScanProsite motif-based hits, 370 proteins overlapped (upper panel of Figure 6.1). Peptides were excluded if they:

i. had identical sequences to other hits (only one hit was then included)
ii. were derived from known interaction partners
iii. were predicted to be within extracellular regions
iv. were derived from proteins with limited expression profiles according to Metacore^{TM} software suite
v. had less than ten PubMed listed papers.

These exclusion criteria are summarised in the middle panel of Figure 6.1. This limited the potential CD2AP SH3 binding regions to 292 binding sites within 210 proteins. For each candidate peptide on the array, the N- and C-termini were extended by five and four residues, respectively, so that the motif was positioned centrally in the final peptide.

Peptide arrays were then used to screen for novel binding partners. Positive and negative controls were included to ensure the quality of the data obtained (Katz et al. 2011, Liu et al. 2012). c-CBL, RIN3e1 and RIN3e2 peptides were used as positive controls, while the RIN3e1($R_{387}A$) and RIN3e2 ($R_{462}A$) mutant peptides were used as negative controls. The SH3 binding affinities for these control peptides were previously quantified by ITC (Table 3.2). The above-mentioned sequences were spot-synthesised as 16-residue peptides in duplicate as described in section 2.2.8.3.4 and probed with purified GST-SH3-1, -2 and -3-S domains (see bottom panels of Figure 6.1).

6.3 Peptide array screen for novel CD2AP SH3 interacting partners

As depicted in Figure 6.1, the SH3 domains bound to a substantial number of peptides. The hit rate was ca. 25% for SH3-1 and SH3-2 but only 5% for SH3-3-S. The low number of SH3-3 binders again implies that this SH3 may prefer a somewhat different recognition sequence compared to the first two CD2AP SH3 domains. A list of the hits is found in Appendix E. There are many strong spots of interest (Figure 6.1), which are particularly interesting for further studies.

Preference was given to hits with medium to high spot intensity and with multiple SH3 binding epitopes within one protein. Based on these filters, ARAP1 (3 epitopes, orange box, Figure 6.1) and MLK3 (2 epitopes, green box, Figure 6.1) were chosen for follow-up experiments.

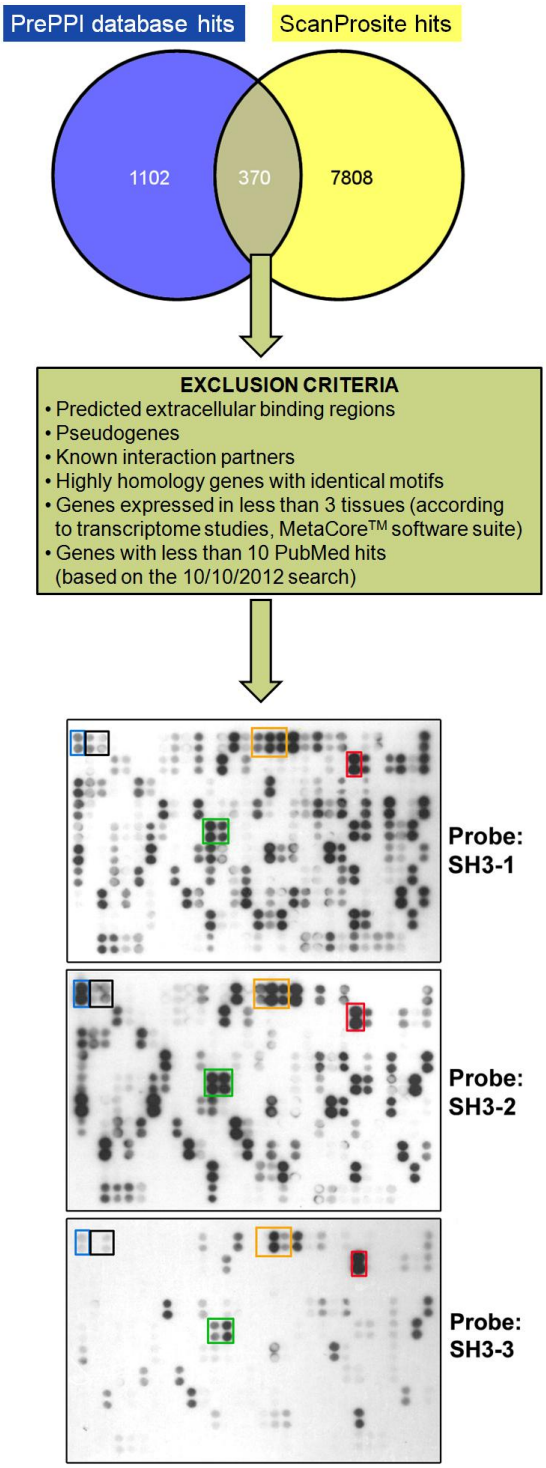


Figure 6.1

Identification of CD2AP SH3 interacting candidates by proteome-wide motif search.

Potential CD2AP SH3 binding proteins, which were selected based on the overlap between ScanProsite and PrePPI hits and further exclusion criteria, were spot-synthesized as 16 aa peptides in duplicates. The arrays were initially probed with GST and then with GST-SH3-1, -2 or -3-S. 210 proteins were tested. The positive controls and hits chosen to be pursued further are boxed in different colours: c-CBL, blue; RIN3, black; ARAP1, orange; DAB1, red; MLK3, green. The Venn diagram was generated by Venny (Oliveros 2007).

Additionally, DAB1 was included in the list for further validation because it was the only hit giving a strong spot intensity with all three SH3 domains (1 epitope, red box, Figure 6.1). It became evident later that the CD2AP/DAB1 interaction has already been reported (Sato et al. 2007). The potential SH3 binding regions within each protein are depicted in Figure 6.2. Despite our usage of the PrePPI database, these epitopes are present within disordered regions. ITC was employed for the validation of the above-mentioned hits.

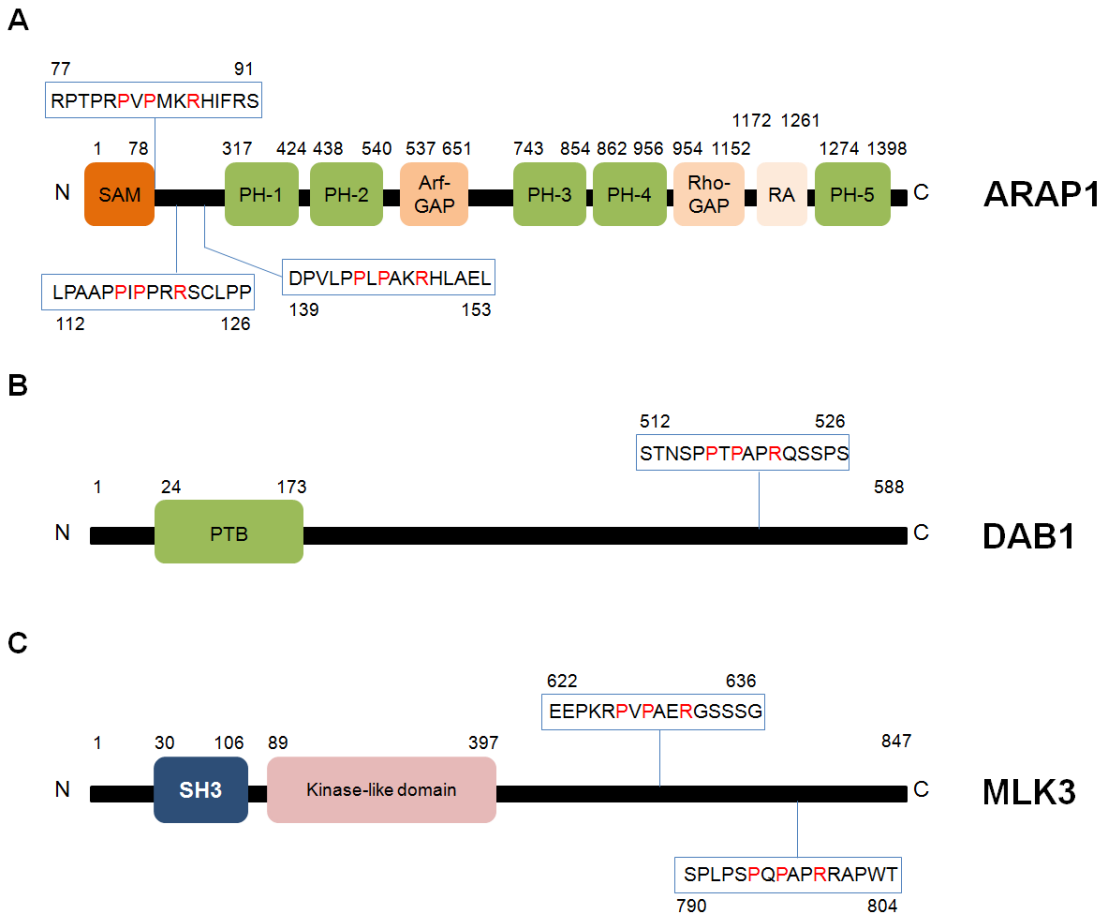


Figure 6.2

Schematic domain composition and position of the potential SH3 binding regions (blue boxes) of the selected hits for further validation.

Red letters within the potential SH3 binding regions show the presence of the P-x-P-x-x-R core motif: ***A***, ARAP1; ***B***, DAB1; ***C***, MLK3.

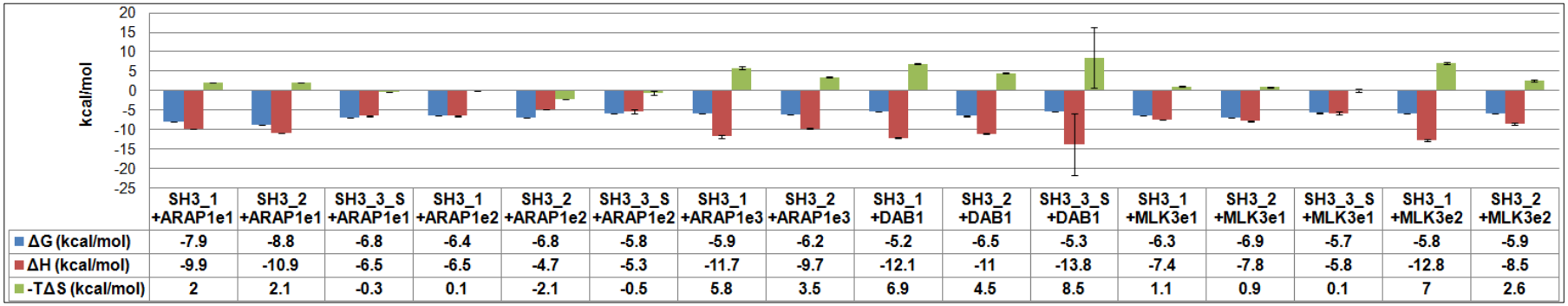
6.4 ITC analysis of the interactions between the SH3 domains and ARAP1, MLK3 and DAB1 peptides

ITC was used to study the binding of the SH3 domain to 16residue peptides corresponding to the three ARAP1 (ARAP1e1-e3), one DAB1 and two MLK3 (MLK3e1-e2) epitopes. The synthetic free peptide boundaries were the same as those for the array peptides. Table 6.1 shows the K_d and N values, while Figure 6.3 depicts the thermodynamic parameters of the interactions.

As shown by the K_d values in Table 6.1, the three SH3 domains bound more tightly to ARAP1e1 by at least one order of magnitude compared to the other peptides. The first two SH3 domains bound with higher affinity to the other peptides compared to SH3-3. Somewhat unusual for SH3 interactions, SH3-2 bound to ARAP1e1 in the nanomolar range. SH3-1 and SH3-3-S bound to ARAP1e1 4-fold and 30-fold more weakly, respectively, compared to SH3-2. There was a 2-fold difference in the binding affinity of SH3-1 and SH3-2 with ARAP1e2, ARAP1e3 and MLK3e1, but these domains have a similar affinity to the DAB1 peptide. SH3-1 bound 10-fold more tightly to MLK3e2 compared to SH3-2. Apart from ARAP1e1, the other peptides bound to SH3-3-S with low affinity. The low affinity of SH3-3-S to DAB1 indicated that this was a false positive in the screen (Figure 6.1). These data show that there are subtle differences in the SH3-1 and SH3-2 recognition patterns, since there are peptides that bind similarly (ARAP1e2, ARAP1e3, DAB1) or differently (ARAP1e1, MLK3e1, MLK3e2) to the first two SH3 domains of CD2AP.

Table 6.1 Isothermal titration calorimetric measurements of CD2AP SH3 domains with ARAP1, DAB1 and MLK3 peptides: Dissociation constant and stoichiometry values. SE = standard error; TLQ= too low to be quantified; -am indicates C-terminal amidation.

Protein	Peptide sequence	Length	K_d ± SE (μM)			N ± SE (μM)		
			SH3-1	SH3-2	SH3-3-S	SH3-1	SH3-2	SH3-3-S
ARAP1e1	RPTPRPVPMKRHIFRS-am	16	1.5 ± 0.00001	0.36 ± 0.0001	11.4 ± 0.001	1.51 ± 0.002	1.37 ± 0.001	1.34 ± 0.01
ARAP1e2	LPAAPPIPPRRSCLPP-am	16	20.0 ± 0.004	9.6 ± 0.35	63.3 ± 0.65	1.27 ± 0.008	1.36 ± 0.009	0.91 ± 0.08
ARAP1e3	DPVLPPLPAKRHLAEL-am	16	52.9 ± 2.5	27.1 ± 0.008	TLQ	1.24 ± 0.03	1.32 ± 0.01	TLQ
DAB1	STNSPPTPAPRQSSPS-am	16	23.2 ± 0.3	15.4 ± 0.004	117.0 ± 0.001	1.10 ± 0.01	1.03 ± 0.01	0.56 ± 0.28
MLK3e1	EEPKRPVPAERGSSSG-am	16	22.7 ± 0.007	8.3 ± 0.1	64.8 ± 0.05	1.66 ± 0.01	1.45 ± 0.004	1.15 ± 0.06
MLK3e2	SPLPSPQPAPRRAPWT-am	16	5.6 ± 0.1	50.0 ± 1.8	TLQ	1.09 ± 0.02	1.25 ± 0.02	TLQ



	SH3_1 +ARAP1e1	SH3_2 +ARAP1e1	SH3_3_S +ARAP1e1	SH3_1 +ARAP1e2	SH3_2 +ARAP1e2	SH3_3_S +ARAP1e2	SH3_1 +ARAP1e3	SH3_2 +ARAP1e3	SH3_1 +DAB1	SH3_2 +DAB1	SH3_3_S +DAB1	SH3_1 +MLK3e1	SH3_2 +MLK3e1	SH3_3_S +MLK3e1	SH3_1 +MLK3e2	SH3_2 +MLK3e2
ΔG (kcal/mol)	-7.9	-8.8	-6.8	-6.4	-6.8	-5.8	-5.9	-6.2	-5.2	-6.5	-5.3	-6.3	-6.9	-5.7	-5.8	-5.9
ΔH (kcal/mol)	-9.9	-10.9	-6.5	-6.5	-4.7	-5.3	-11.7	-9.7	-12.1	-11	-13.8	-7.4	-7.8	-5.8	-12.8	-8.5
-TΔS (kcal/mol)	2	2.1	-0.3	0.1	-2.1	-0.5	5.8	3.5	6.9	4.5	8.5	1.1	0.9	0.1	7	2.6

Figure 6.3

Thermodynamic parameters upon binding of the ARAP1, DAB1 and MLK3 peptides to the SH3 domains. The error bars denote the standard error. Any 'TLQ' interaction (Table 6.1) was excluded from this analysis.

In most cases, the N values were quite close to 1 indicating 1:1 binding. Three values suggested deviation from 1:1 binding (i.e. SH3-1/ARAP1e1, SH3-1/MLK3e1 and SH3-3-S/DAB1), but this is unlikely to be functionally relevant and might be explained by the same factors described in section 3.5.1.1. The absence of aromatic residues in the peptides except in ARAP1e1 precluded measurement of the peptide concentration by UV spectroscopy at 240-300 nm. For ARAP1e1, the absorbance at 257 nm was attempted (data not shown), but since Phe does not absorb as highly as Tyr or Trp, definite conclusions could not be drawn. As will be described later, the crystal structure between SH3-2 and ARAP1e1 was obtained and showed that N is indeed 1. Studies with different methods would have to be performed to verify whether the N values are 1 or 2 in the cases of SH3-1 in complex with ARAP1e1 (N=1.5) and MLK3e1 (N=1.66), but due to a lack of time, such studies were not be pursued as part of this thesis.

If we focus on the thermodynamic parameters of the interaction between the SH3 domains and the ARAP1, DAB1 and MLK3 peptides, Figure 6.3 shows that the interactions are enthalpically driven (i.e. negative ΔH), which indicates that polar contacts mediate the interaction between the two molecules. Small favourable (i.e. negative) or unfavourable (i.e. positive) changes in entropy were observed for ARAP1e1, ARAP1e2 and MLK3e1. In the case of the SH3 interactions with ARAP1e3, DAB1 and MLK3e2 peptides, the high favourable changes in enthalpy were compensated by high unfavourable changes in entropy (i.e. positive $-T\Delta S$), which resulted in a small reduction in the free energy. This suggests that hydrophobic interactions, displacement of solvent

molecules and loss of degrees of freedom of the system might contribute additionally to the interaction.

6.5 Analysis of the CD2AP interaction with ARAP1 and MLK3 by precipitation with GST-fusion proteins

The interaction between the CD2AP and DAB1 proteins has been reported before (Sato et al. 2007). Therefore, the ARAP1 and MLK3 interactions with CD2AP were further investigated.

ARAP1 and MLK3 were found to be expressed in HEK293, HeLa and MCF-7 cell extracts. One anti-MLK3 and two anti-ARAP1 antibodies were found to be of good quality and were thus used for immunoblotting assays (section 2.1.1, Appendix D). To study initially the CD2AP/ARAP1 interaction, pulldowns from HEK293, HeLa and MCF-7 cell extracts were attempted with the individual GST-SH3 domains and GST-SH3-(1+2) (protocol details in section 2.2.3.1). Figure 6.4 depicts the ARAP1 and MLK3 blots with the individual SH3 domains and SH3-(1+2). SH3-1, SH3-2 and SH3-(1+2), but not SH3-3-S, pulled down endogenous ARAP1 and MLK3 from the three tested cell line lysates (Figure 6.4). SH3-3-L did also not pull down ARAP1 or MLK3 (data not shown). This might have been because MLK3 was present in other complexes or formed transient or low abundance complexes.

The inability of GST-SH3-3-S to pull down ARAP1 or MLK3 might be due to the low binding affinity with ARAP1e2, ARAP1e3, MLK3e1 and MLK3e2 (Table 6.1). Surprisingly, the low micromolar range affinity of SH3-3-S with ARAP1e1 (K_d= 11.4 μM, Table 6.1) was apparently not sufficient for this domain to precipitate ARAP1.

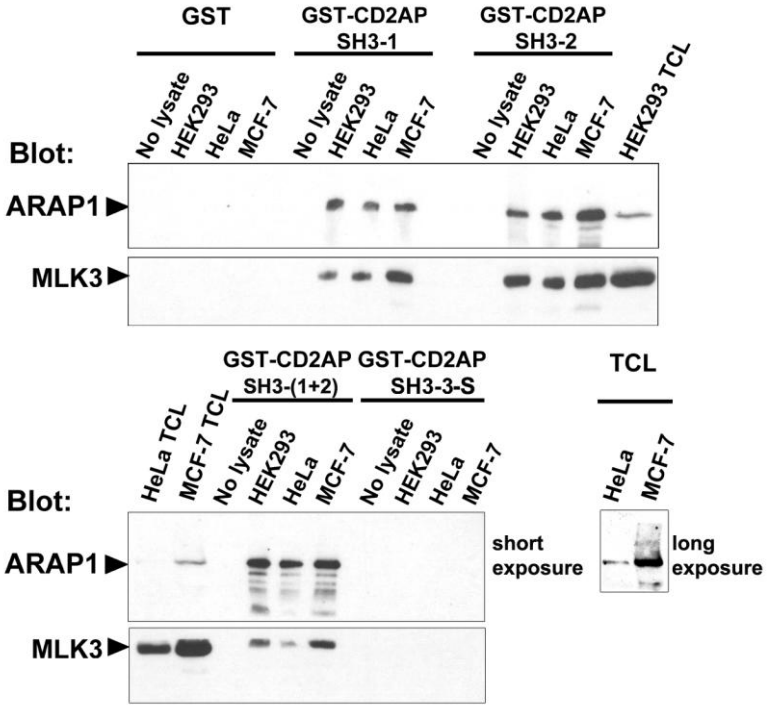


Figure 6.4

SH3-1 and SH3-2 but not SH3-3 can bind to ARAP1 and MLK3 in vitro.

1 mg of HEK293, HeLa or MCF-7 total cell lysates (TCL) were precipitated with 50 µg of bead-immobilised GST, isolatedGST-SH3 or GST-SH3-(1+2) proteins, precipitates were subjected to immunoblotting (WB) with anti-ARAP1 or anti-MLK3 antibodies.

As already mentioned in section 1.5.2.1, CD2AP participates in kidney podocyte physiology. Therefore, we investigated whether ARAP1 is expressed in podocytes. The ARAP1 expression levels were found to be similar in proliferative and differentiated podocytic cell extracts (Figure 6.5). This prompted us to investigate if the SH3 domains could pull down endogenous ARAP1 from undifferentiated and differentiated (day 14) states (Figure 6.5). Consistent with Figure 6.4, the first two SH3 domains but not SH3-3-S pulled down ARAP1. SH3-3-L was also not able to pull down ARAP1

(Figure 6.5). SH3-(1+2) pulled down similar ARAP1 amounts to the individual SH3-1 and SH3-2 domains (Figure 6.5).

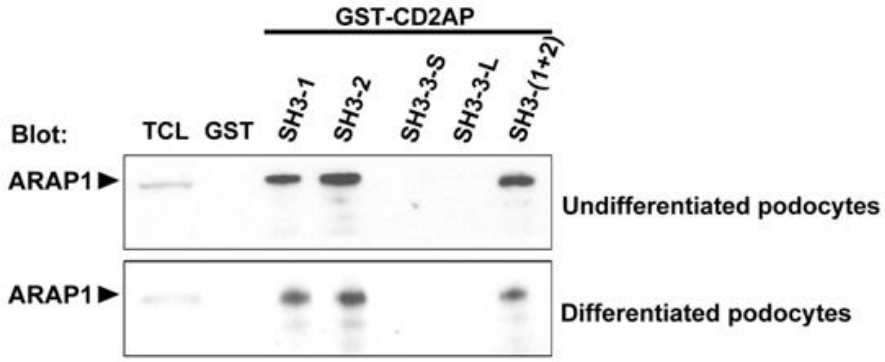


Figure 6.5

SH3-1 and SH3-2 but not SH3-3 can bind to ARAP1 and MLK3 in vitro.

1 mg of lysates from undifferentiated (upper panel) or differentiated (lower panel) cells were precipitated with 50 µg of bead-immobilised GST, isolated GST-SH3 or GST-SH3-(1+2) proteins; precipitates were subjected to immunoblotting with anti-ARAP1 antibody.

6.6 Analysis of the CD2AP/ARAP1 interaction by co-immunoprecipitation

We proceeded by analysing the CD2AP interaction with ARAP1 by co-immunoprecipitation. These experiments were performed in the above-mentioned cell lines (HEK293, HeLa, undifferentiated and differentiated podocytes). The protocol is detailed in section 2.2.3.2. The immunoprecipitations are shown from HEK293, HeLa (Figures 6.6A, 6.6B), undifferentiated and differentiated podocytes (Figures 6.6C, 6.6D) with one set of anti-CD2AP and anti-ARAP1 antibodies (Figure 6.6). Similar CD2AP and ARAP1 amounts were precipitated from the above-mentioned cell line extracts. These findings show that the CD2AP/ARAP1 interaction is present in a number of cell lines derived from different tissues.

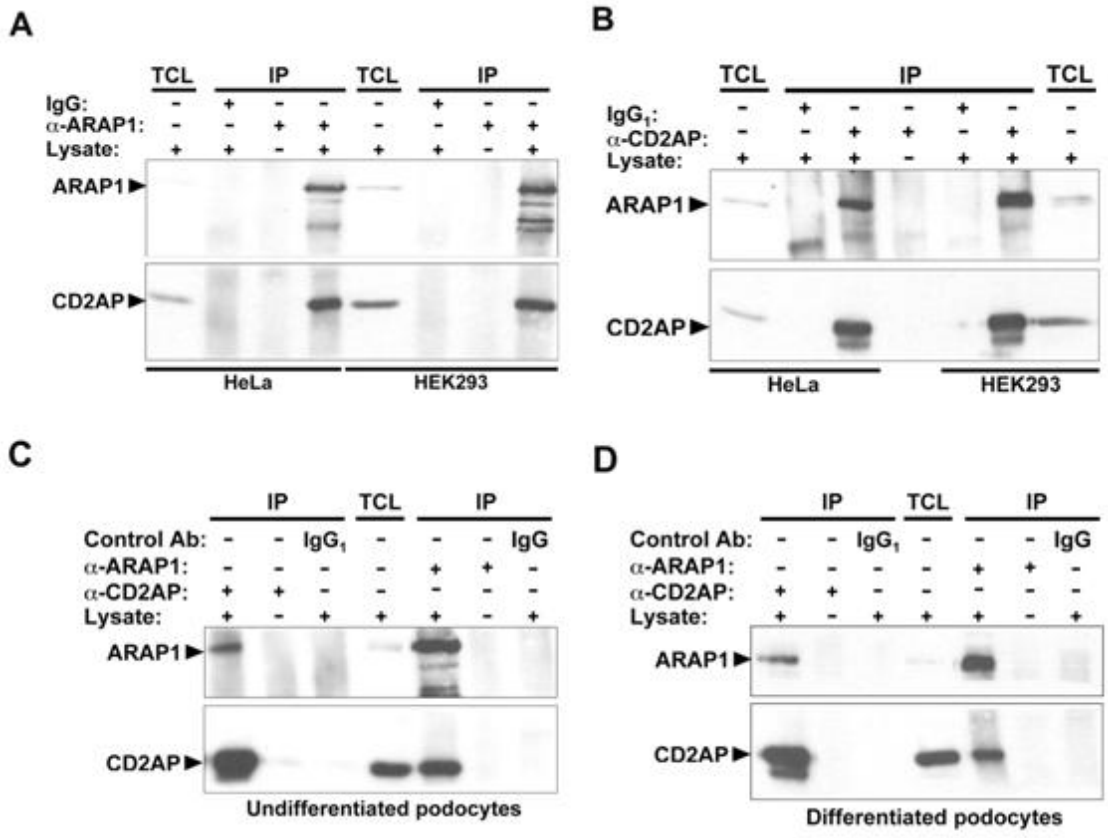


Figure 6.6

Endogenous CD2AP and ARAP1 interact in lysates from HEK293, HeLa and immortalised human podocytes.

CD2AP was co-immunoprecipitated in HEK293, HeLa (A), podocytes in undifferentiated (C) and differentiated states (D) with an anti-ARAP1 antibody. ARAP1 was co-immunoprecipitated in HEK293 and HeLa (B), or podocytes in undifferentiated (C) and differentiated states (D) with an anti-CD2AP antibody.

6.7 Attempts to investigate the CD2AP interaction with MLK3 by co-immunoprecipitation

The study of the CD2AP/MLK3 interaction was attempted by co-immunoprecipitation. HEK293, HeLa and MCF-7 cell extracts were used to detect the complex with anti-CD2AP antibodies. However, the co-immunoprecipitations were unsuccessful. Some non-specific Ig binding was observed in the MLK3 blot. A representative experiment is found in Figure 6.7.

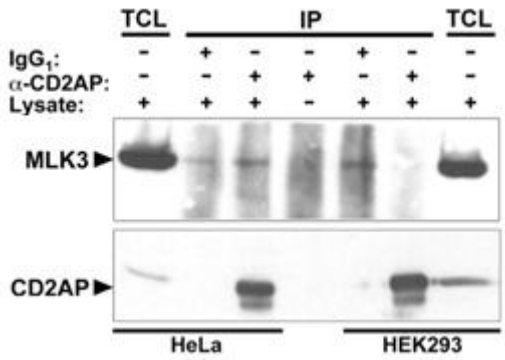


Figure 6.7

Attempt to co-immunoprecipitate the CD2AP/MLK3 complex from HEK293 and HeLa cell extracts with an anti-CD2AP antibody.

Lower panel: Western blot for CD2AP; *upper panel*: Western blot for MLK3.

6.8 Analysis of the SH3-1 and SH3-2 interaction with ARAP1e1 by protein X-ray crystallography

6.8.1 Crystallisation trials with the CD2AP SH3 domains and the ARAP1e1 peptide

Due to the high binding affinities between SH3-1 or SH3-2 and ARAP1e1 and the confirmation by immunoprecipitation of the CD2AP/ARAP1 interaction, we wanted to gain a structural view of these interactions. Crystal formation was attempted by extensive sparse matrix sampling. The attempted crystallisation screens were Hampton Research Index (HIN) and Joint Centre for Structural Genomics core (JCSG). The SH3-1 concentration was 23.6 mg/ml and the SH3-2 concentration was 20 mg/ml. In the case of the SH3-2 in complex with ARAP1e1, crystal formation was observed in 0.1 M HEPES pH 7.5 and 1.4 M tri-sodium citrate dehydrate (Figure 6.8). These are the same conditions in which the SH3-2/RIN3e2 crystal formed. Small crystals, which diffracted

poorly, were also obtained for the SH3-1 complex with ARAP1e1 in 0.1 M citric acid pH 3.5 and 25% w/v polyethylene glycol 3350 (not shown).

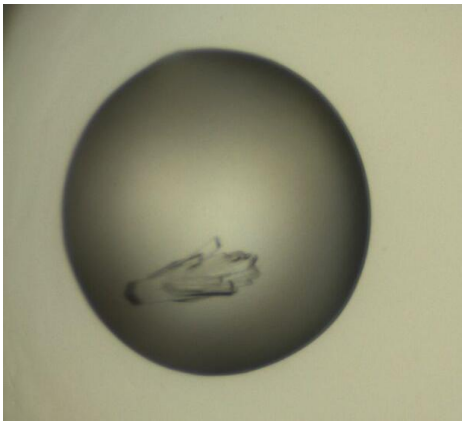


Figure 6.8

Crystals of the SH3-2 complex with ARAP1e1.

6.8.2 Crystal structure of CD2AP SH3-2 domain in complex with ARAP1e1

6.8.2.1 Crystallisation, data collection, processing and refinement of the CD2AP SH3-2/ARAP1e1 complex

Crystals were grown at 20°C from a 1:1 ratio of mother liquor (0.1 M HEPES pH 7.5, 1.4 M tri-sodium citrate dehydrate) to CD2AP SH3-2 protein solution (20 mg/ml). A cluster of plates (~400 μm in length and ~250 μm in width), which grew after 1 week (Figure 6.8), was first transferred to cryoprotectant (mother liquor supplemented with 10% glycerol) before being snap-frozen in liquid nitrogen.

Data were collected to 1.58 Ångstrom on beamline I04 at Diamond synchrotron (by Structural Genomics Consortium Oxford members) by finding a region in the crystal cluster giving favourable diffraction. Data reduction was

carried out with *Mosflm* (Battye et al 2011), which suggested that the space group could be C2, C222 or $C222_1$. *Pointless* indicated that C2 was the most probable space group (probability of 0.909). Next, the dataset was converted to *CCP4* mtz format and scaled with *SCALA* (Evans 2006); data to 1.58 Å were kept for refinement, resulting in a signal-to-noise ratio (I/sigma[I]) of 4.4 in the last resolution shell. The *Matthews_coef* (Collaborative Computational Project 1994) indicated that the complex crystallised as a trimer (*i.e.* three SH3-2 domains and three ARAP1 peptides). Molecular replacement was performed with *Phaser* (McCoy et al. 2005) using as a search model the SH3-2 domain alone from the already solved complex of CD2AP SH3-2/RIN3e2 peptide (PDB code: 3U23). Unfortunately, the program was not able to find a solution. The Translation Function Z-score (TFZ) was around 5, which indicated that the chosen parameters were wrong. The function value did not improve using space group C222 and the above-mentioned steps repeated.

As mentioned in Figure 2.5, the dataset was indexed by choosing a few representative images from two sets of contiguous images, each set being approximately 90 degrees apart. Therefore, we sought to re-index the dataset by choosing other images. Surprisingly, different space groups ($P4_1$, oC, mC and P2) and smaller unit cell dimensions compared to the previous solution were proposed by *Mosflm* (Battye et al. 2011). The dataset was also processed with the *XDS* package (Kabsch et al. 2010) and space group P2 was proposed as the most probable, while $P4_1$ was the second best solution. Due to the different suggestions from the two programs, we chose initially to re-index the dataset by in space group $P4_1$ using *Mosflm* (Battye et al. 2011).

Pointless (Collaborative Computational Project 1994) suggested that $P4_1$ was the most probable choice. The dataset was converted to *CCP4* mtz format and scaled with *SCALA* (Evans 2006). *Matthews_coef* (Collaborative Computational Project 1994) suggested that the complex crystallised as a monomer (*i.e.* one SH3-2 domain and one ARAP1 peptide). Molecular replacement was performed with *Phaser* (McCoy et al. 2005) using as a search model the SH3-2 domain alone from the already solved complex of CD2AP SH3-2 / RIN3e2 peptide (PDB code: 3U23). However, the model was again incorrect, with a low TFZ value of around 5.

Finally, *Phaser* was run by choosing all possible space groups within P4 point group again using the same SH3-2 domain as a search model. This time, the TFZ scores were above 8, revealing a single, correct solution in space group $P4_3$. The difficulty the programs had to index the dataset could be due to over-estimation of the unit cell dimensions by inclusion of additional intensities contributed by secondary lattice(s), this as a result of the crystal being multiple.

The model output from *Phaser* was then refined with *REFMAC 5* (Murshudov et al. 1997). Additional manual model building was done with *Coot* (Emsley et al. 2004). The ARAP1e1 peptide was modelled with reference to the RIN3e2 peptide in the CD2AP SH3-2/RIN3e2 structure. A cycle of rigid body refinement, several cycles of restrained refinement and manual model building were performed to improve the model (Figure 6.9). Next, TLS (torsion, libration, screw) restraints were included for 2 groups, one for each chain in the asymmetric unit (*i.e.* one SH3-2 domain and one ARAP1 peptide).

Additional model building and cycles of TLS and restrained refinement decreased the R_{free} and R_{work} values further (Figure 6.9).

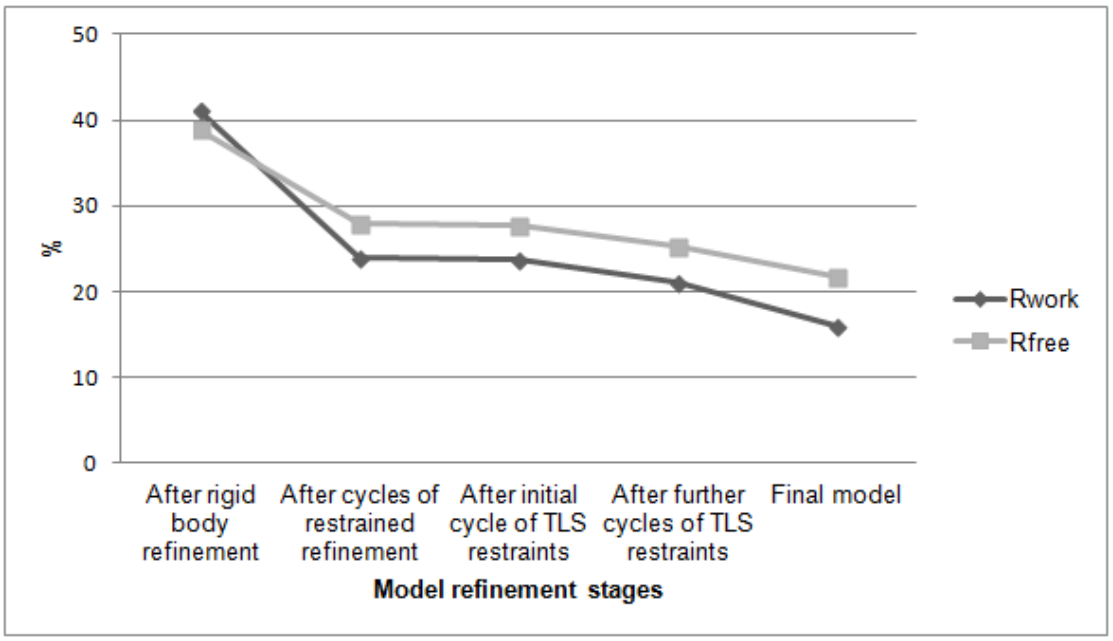


Figure 6.9

Reduction of R_{work} and R_{free} values during the SH3-2/ARAP1e1 complex model refinement.

The different stages of model refinement are represented along the x-axis.

The data collection, processing and refinement data of the final model are found in Figure 6.10A. The Ramachandran plot (Lovell et al. 2003) for the final model, which shows good residue geometry, is found in Figure 6.10B. The final model was deposited in the PDB with accession code 4X1V.

A

Data collection and processing	PDB: 4X1V
Diamond beamline	I04
Detector	Pilatus 6M
Wavelength (Å)	0.97625
Space group	$P4_3$
Unit cell parameters (Å)	44.59 44.59 32.94
(°)	90.00 90.00 90.00
Resolution range (Å)	44.59 - 1.58
I/σ(I)	9.4 (4.2)
Completeness (%)	99.9 (100.0)
R meas (%)	27.0 (46.1)
Refinement	
R work (%)	16.0 (18.2)
R free (%)	21.7 (22.2)
R.m.s.d. bond angles (°)	2.15
bond lengths (Å)	0.02
Ramachandran plot (%)	
Favoured regions	98.8
Allowed regions	1.2
Outliers	0
Additional data	
Highest resolution cell	1.62-1.58

B

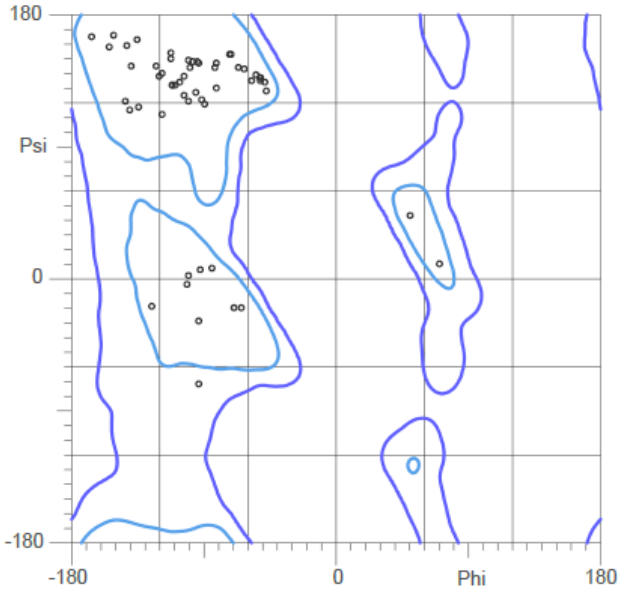


Figure 6.10

Quality of the final model of the SH3-2/ARAP1e1 complex.

A, Data collection, processing and refinement parameters. Values in parentheses represent the highest resolution shell: 1.62 – 1.58 Å, ***B***, Ramachandran plot. The x-axis represents the phi value, while the y-axis represents the psi value for each residue. The allowed regions are shown in dark blue and the favoured regions are represented in light blue.

6.8.2.2 CD2AP SH3-2/ARAP1e1 complex characteristics

The main chain of K_{110}, R_{111}-V_{167} of SH3-2 and P_{79}-F_{89} of ARAP1e1 were visible in the electron density maps.

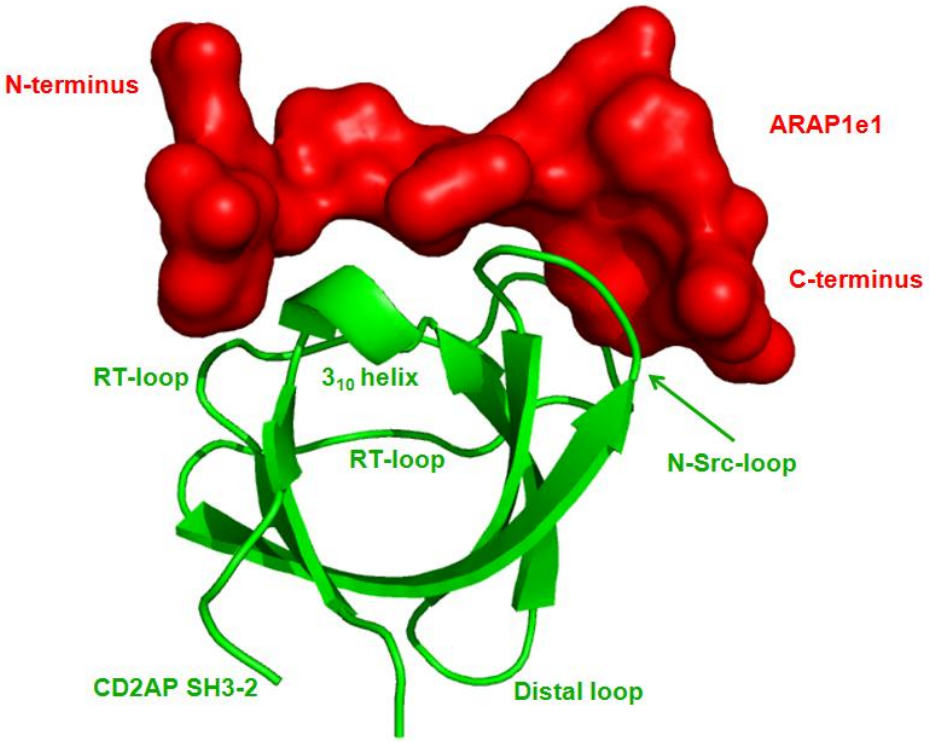


Figure 6.11

Surface representation of ARAP1e1 docking onto the β-barrel CD2AP SH3-2 domain.

SH3-2: green ribbon; the loops and 3_{10} helix are designated; ARAP1e1: red surface.

As illustrated in Figure 6.11, SH3-2 has the typical SH3 β-barrel structure. The β-barrel is formed by the same residues as described for the SH3-2/RIN3e2 complex (section 3.6.3). The 3_{10} helix and the RT and N-Src loops are positioned close to ARAP1e1. As shown in Figure 6.12, the peptide adopts a PPII helical conformation to bind to SH3-2 in class II orientation. P_{79}, P_{81} and P_{83} are docked into three hydrophobic pockets, and R_{86} into the negatively charged specificity pocket of SH3-2. F_{89} makes hydrophobic contacts with R_{86}

of ARAP1e1. Therefore, the crystal structure suggests that SH3-2 recognises an extended P-x-P-x-P-x-x-R-x-x-F motif in ARAP1e1.

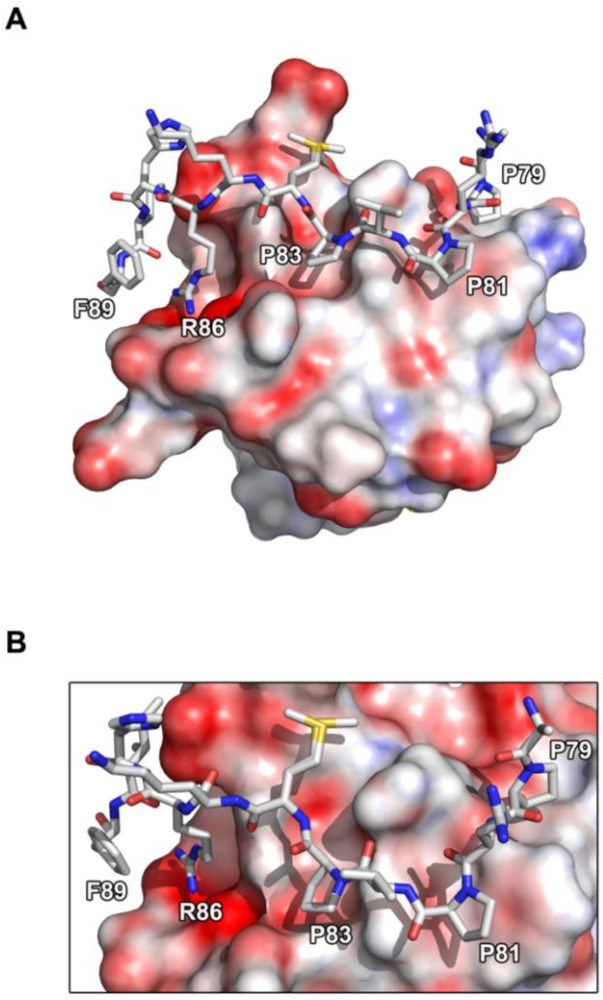


Figure 6.12

Electrostatic surface potential representation of the CD2AP SH3-2 domain in complex with ARAP1e1.

The ARAP1e1 peptide is shown in stick representation and coloured by element (carbon: white; nitrogen: blue; oxygen: red). The key residues that form contacts with the SH3 surfaces are labelled: ***A***, General view and ***B***, Close-up view of the docking positions of ARAP1e1 peptide on the SH3-2 domain. These two views were generated by rotating the positions shown in *A* by -20^{o} around the y axis and -10^{o} around the z axis.

Figure 6.13 depicts the following SH3-2 contact points with ARAP1e1: L_{116}, F_{117}, Y_{119}, N_{123}, D_{125}, E_{126}, $W_{145,}$ V_{141}, P_{158}, N_{160} and F_{161}. L_{116} and F_{117} make hydrophobic contacts with P_{79}, while P_{81} interacts with F_{161} and F_{117}. P_{83} binds

to Y_{119}, F_{161} and P_{158}. R_{86} interacts with N_{123}, D_{125}, E_{126} and W_{145} of SH3-2, and F_{89} and a peptide main chain atom of ARAP1e1. Also, E_{142} might contribute to the interaction with ARAP1e1 via water-mediated interactions (Appendix B, Figure B.3). However, more biophysical studies, which are beyond the scope of this thesis, would have to be performed to investigate the contribution of the water molecules.

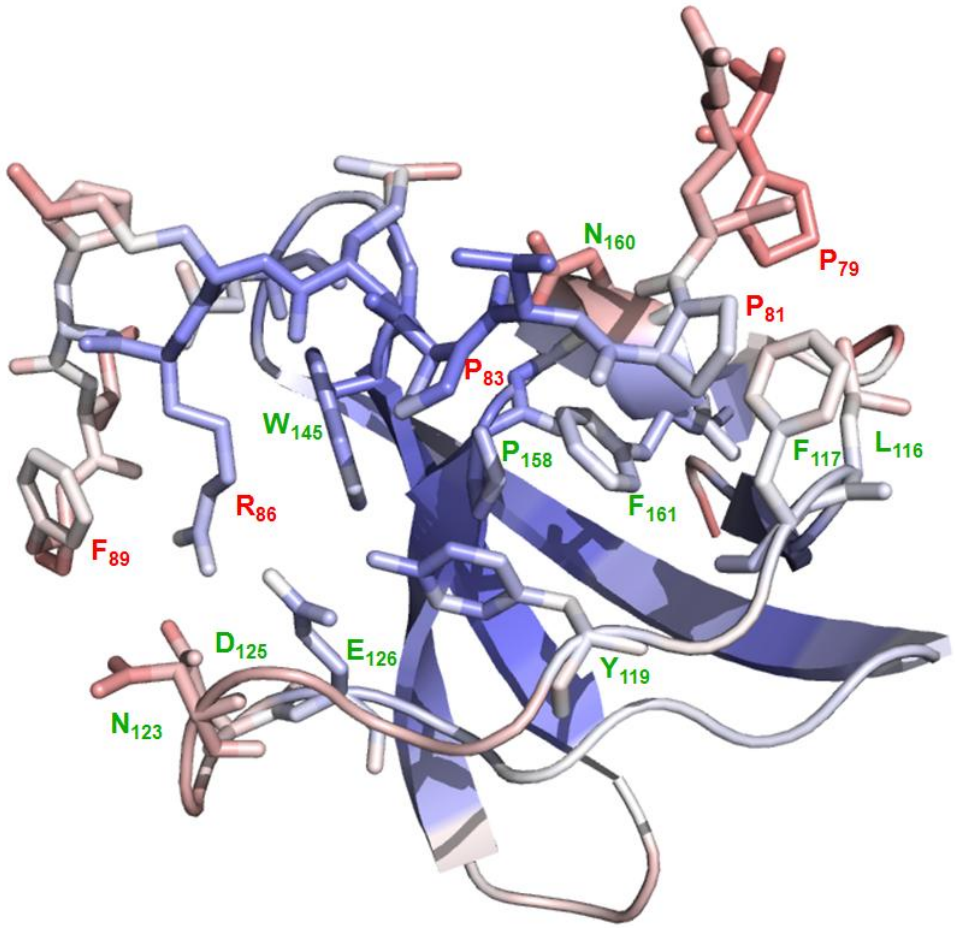


Figure 6.13

B-factors and points of contact between SH3-2 and ARAP1e1.

The ARAP1e1 peptide and the SH3-2 residues are shown in stick and secondary structure representation respectively and coloured by b-factor (blue to white to red). The key SH3-1 residues that form contacts with RIN3e1 are depicted in sticks. The key RIN3e1 and SH3-1 residues that form contacts are labelled. The colour labelling is in accordance with previous figures. SH3-2 residues: green; ARAP1e1 residues: red.

In Figure 6.13, SH3-2 and ARAP1e1 are coloured according to their B-factors with a blue-white-red gradient. According to the B-factor values, the SH3 domain core and β-strands are less flexible (blue color) than the 3_{10} helix and loops (white to red color). In the 3_{10} helix, N_{160} is less thermo stable than F_{161}. Also, N_{123} and D_{125} of the RT loop of SH3-2, which interacts with R_{86} of ARAP1e1, were more flexible than other residues in that loop. The N- and C-termini of ARAP1e1 are more flexible than the central region of the peptide. P_{79} is less thermo stable than $F_{89.}$ So, it might contribute less to the interaction with SH3-2 compared to F_{89}.

As described in section 3.6.4.3, the SH3-1/RIN3e1 and SH3-2/RIN3e2 complex crystal structure data do not correlate with the expected SH3 thermodynamic signatures. This is also observed for the SH3-2/ARAP1e1 complex. The high number of hydrophobic contacts would lead to high favourable entropy and a small enthalpic contribution. As depicted in Figure 6.3, the SH3-2 binding to ARAP1e1 leads to high favourable changes in enthalpy and small unfavourable changes in entropy.

6.8.3 Comparison of the CD2AP SH3-2 structures complexed with RIN3e2 and ARAP1e1 peptides

The structures of SH3-2 in complex with two different peptides (ARAP1e1 and RIN3e2) enabled comparison between the binding modes of each different ligand. Figure 6.14 shows a structural alignment of these two complexes. It is clear that the SH3-2 domains display similar conformations upon binding to the ARAP1e1 and RIN3e2 peptides (r.m.s.d.= 0.302 Å). The secondary structure elements and loops align well with each other (Figure 6.14A).

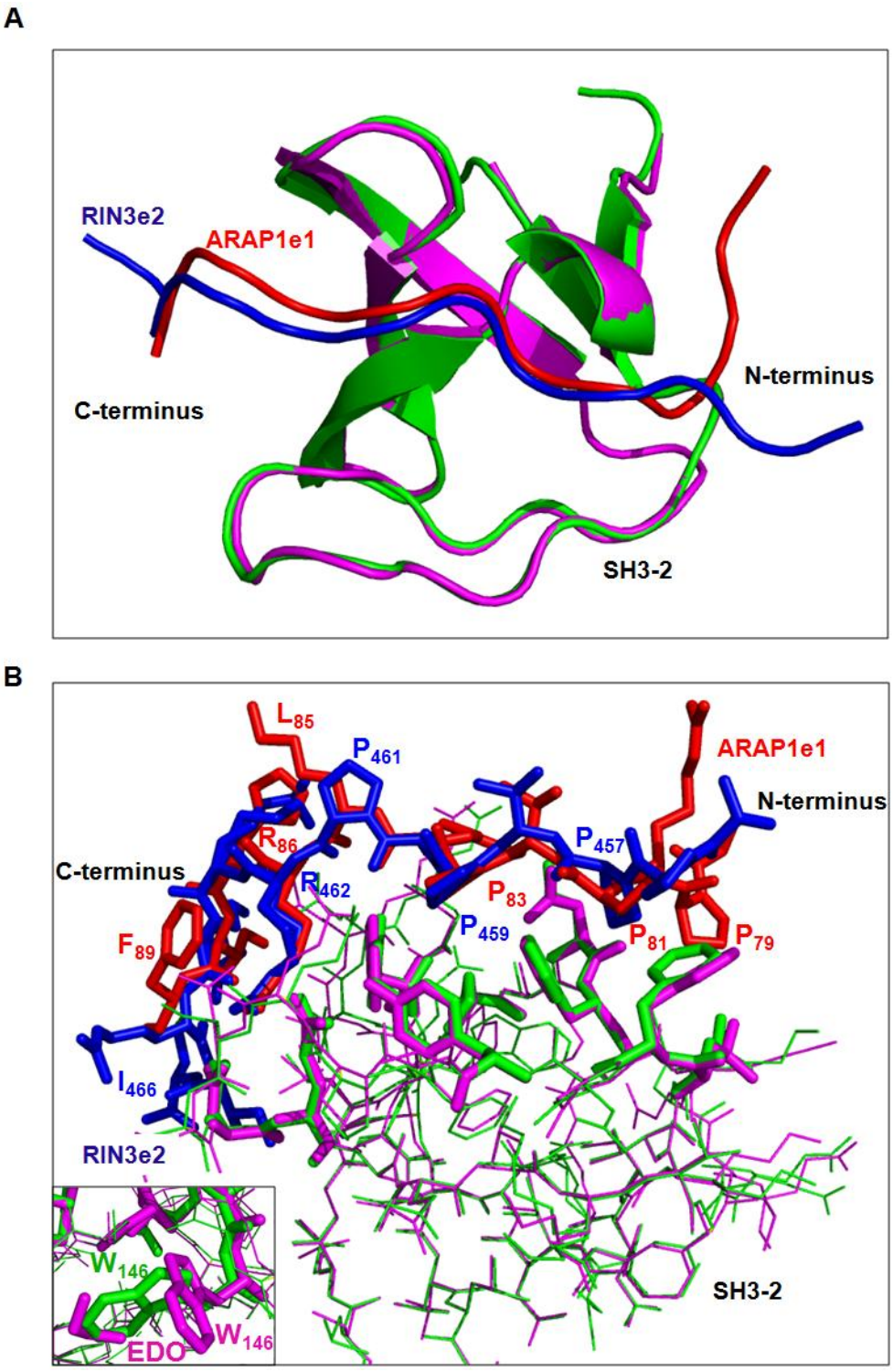


Figure 6.14

Overlay of SH3-2 complexes with ARAP1e1 and RIN3e2.

SH3-2 (green colour) in complex with ARAP1e1 (red colour); SH3-2 (magenta colour) in complex with RIN3e2 (blue colour): ***A***, The SH3-2 domains and the peptides are shown in ribbon representation. The secondary structure elements of the SH3-2 domains are also indicated. ***B***, The ARAP1e1 and RIN3e2 and the SH3-2 residues critical for the interaction are shown in stick representation. The rest of the SH3-2 domains are shown in lines. *Bottom insert*: The difference in the positioning of the W146 residue between the two complexed SH3-2 domains. EDO: ethylene glycol.

If we examine the SH3-2 residues, there are no major differences in their positioning within the SH3 fold, apart from W_{146}, whose position in both SH3-2 domains is featured in the inset of Figure 6.14B. In the SH3-2/ARAP1e1 structure, the position of W_{146} is likely a crystal packing artefact, because it makes contacts with a nearby molecule of the unit cell. In the SH3-2/RIN3e2 structure, this position is occupied by an ethylene glycol molecule and there are no contacts between W_{146} and nearby molecules of the unit cell. This suggests that the position of W_{146} in the SH3-2/RIN3e2 complex is most likely to be the physiologically relevant one.

According to the SH3-2/ARAP1e1 structure, SH3-2 recognises an extended P-x-P-x-P-x-x-R-x-x-F motif in ARAP1e1. As described in Chapter 3, SH3-2 recognises a P-x-P-x-p-R motif in RIN3e2. Motif positions 1, 3 and 6 are critical for binding, while position 5 may play a role in binding (section 3.6.4.3). If we compare the SH3-2 recognition motifs in ARAP1e1 and RIN3e2, the core P-x-P-x-x-R motif is present in both ligands. In motif position 5, a Pro is found in RIN3e2, and a Lys in ARAP1e1. As shown by the overlay of the SH3-2/ARAP1e1 and SH3-2/RIN3e2 complexes, the ARAP1e1 and RIN3e2 residues within the core motif are positioned similarly onto the SH3-2 surface. Pro461 (RIN3e2) and Lys85 (ARAP1e1) (i.e. residues in motif position 5) do not contact the SH3 surface. This suggests that motif position 5 is not critical for SH3-2 binding to ARAP1e1 or RIN3e2.

However, major differences occur at the positioning of the ARAP1e1 and RIN3e2 N- and C-termini. The SH3-2/ARAP1e1 structure indicates two additional anchoring ARAP1e1 residues (P_{79} [motif position -1] and F_{89} [motif position 9]), which extend beyond the core P-x-P-x-x-R motif. The differences

in the orientation of the N- and C-termini of the RIN3e2 and ARAP1e1 peptides are clearly shown in Figure 6.14. Figures 6.14A and 6.14B depict in different representations how the N-terminus of RIN3e2 moves away from the SH3-2 surface, while the N-terminus of ARAP1e1 makes additional contacts with SH3-2 via P_{79}. It is more difficult to conclude on the direction of the C-terminus of ARAP1e1, because F_{89} is the last visible residue and is positioned towards R_{86}. The electron density of RIN3e2 at its C-terminus is more extended and I_{466} is the last residue that contacts SH3-2. As already described in section 3.6.3, I_{466} does not contribute to the affinity of the SH3-2/RIN3e2 interaction. Since SH3-2 binds to ARAP1e1 with an affinity in the nanomolar range (Table 6.1), the additional ARAP1e1 contacts made by P_{79} and F_{89} with SH3-2 might be responsible for this tight interaction. This was further investigated by ITC and is described in the next section.

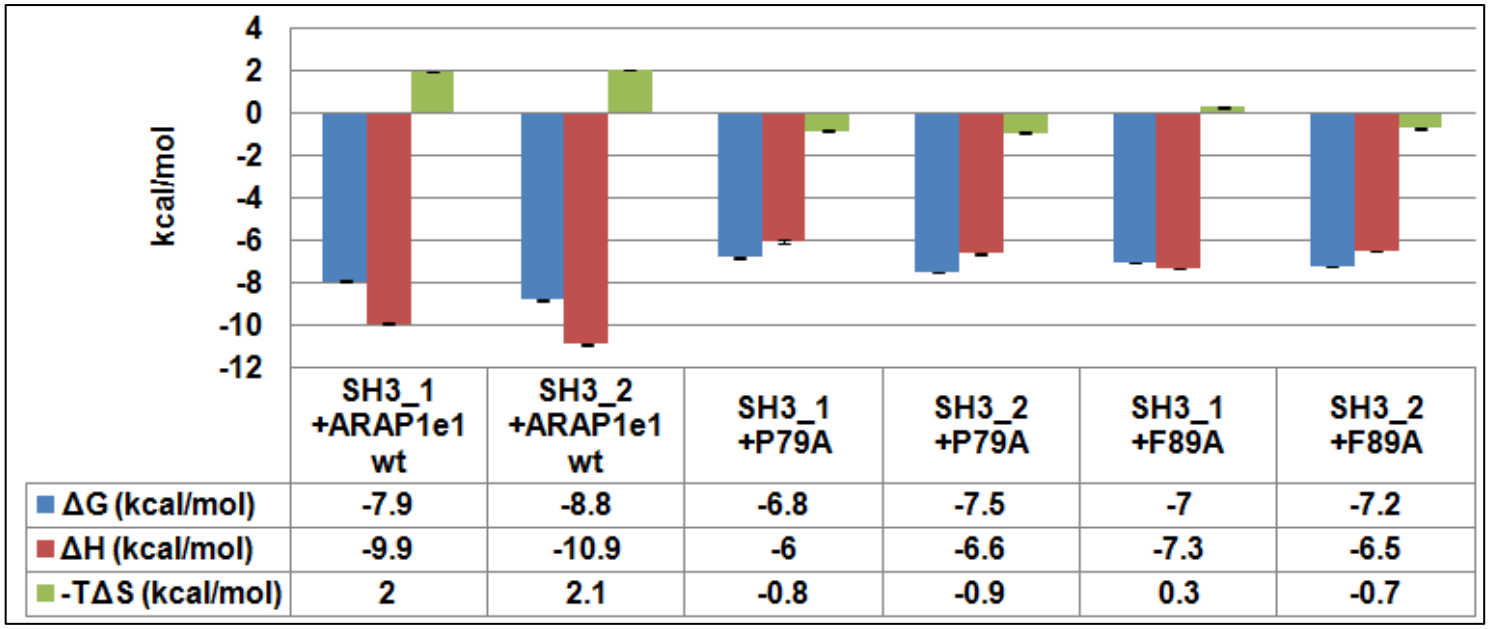
Finally, when comparing the water molecule network at the interface of both structures, the majority of them are conserved between the two structures (Appendix B, Figures B.2-B.3, data not shown).

6.9 ITC analysis of SH3-1 and SH3-2 domains with ARAP1e1 mutant peptides

In order to examine whether P_{79} and F_{89} contributes to the ARAP1e1 interaction with SH3-2, alanine point mutant (P_{79}A and F_{89}A) peptides were tested by ITC.

Table 6.2 Isothermal titration calorimetric measurements of SH3-1 and -2 domains with ARAP1e1 wild-type and mutant peptides: Dissociation constant and stoichiometry values. The data were fitted with a one-site binding model. SE = standard error; TLQ= too low to be quantified; -am indicates C-terminal amidation.

Mutant	Peptide sequence	Length	K_d ± SE (μM)		N ± SE	
			SH3-1	SH3-2	SH3-1	SH3-2
wt	RPTPRPVPMKRHIFRS-am	16	1.50 ± 0.00001	0.36 ± 0.0001	1.51 ± 0.002	1.37 ± 0.001
R_{86}→A	RPTPRPVPMKAHIFRS-am	16	TLQ	TLQ	TLQ	TLQ
P_{79}→A	RPTARPVPMKRHIFRS-am	16	10.80 ± 0.9	3.30 ± 0.002	2.92 ± 0.03	2.56 ± 0.011
F_{89}→A	RPTPRPVPMKRHIARS-am	16	7.80 ± 0.2	5.80 ± 0.16	2.20 ± 0.006	2.05 ± 0.006



	SH3_1 +ARAP1e1 wt	SH3_2 +ARAP1e1 wt	SH3_1 +P79A	SH3_2 +P79A	SH3_1 +F89A	SH3_2 +F89A
ΔG (kcal/mol)	-7.9	-8.8	-6.8	-7.5	-7	-7.2
ΔH (kcal/mol)	-9.9	-10.9	-6	-6.6	-7.3	-6.5
-TΔS (kcal/mol)	2	2.1	-0.8	-0.9	0.3	-0.7

Figure 6.15

Thermodynamic parameters upon binding of the ARAP1e1 wild-type and mutant peptides to the first two SH3 domains.

The error bars denote the standard error.

Also, the importance of the R_{86} of the P-x-P-x-x-R motif in ARAPe1 was analysed by testing a $R_{86}A$ point mutant peptide by ITC. This peptide was chosen because it was found that mutating the critical Arg to any other residue abolished or weakened significantly the interaction with RIN3e2 (Table 3.2). These peptides were tested with SH3-1 and SH3-2 only, because they bound tightly to the wild-type peptide (Table 6.1). Table 6.2 shows the K_d and N values for $R_{86}A$, $F_{89}A$ and $P_{79}A$ titrations. The thermodynamic parameters of the SH3 interactions with these peptides are found in Figure 6.15.

The affinities of the interactions between the first two SH3 domains and $R_{86}A$ were too low to be quantified (Table 6.2). This confirmed that R_{86} is critical for the interactions between the first two SH3 domains and ARAP1e1. As shown in Table 6.2, $P_{79}A$ causes a 7-fold reduction in affinity to SH3-1 and a 10-fold reduction to SH3-2 compared to wild-type. Also, the affinity of $F_{89}A$ is reduced 5-fold for SH3-1 and 16-fold for SH3-2 compared to the wild-type peptide. Surprisingly, the N value was two for the SH3-1/$F_{89}A$ and SH3-2/$F_{89}A$ and three for the SH3-1/$P_{79}A$ complexes. The N value for the SH3-2/$P_{79}A$ was two or three.

If we look at the thermodynamic parameters of the SH3-1 or SH3-2 binding to $P_{79}A$ (Figure 6.15), ΔH is similar and dominant to ΔG compared to a small favourable (i.e. negative) ΔS. As depicted in Figure 6.15, the SH3/$F_{89}A$ interactions are driven by favourable changes in enthalpy, while ΔS contributes little to ΔG, and hence the affinity.

If we compare these values with the thermodynamic parameters of the wild-type peptide with SH3-1 or SH3-2, a reduction in ΔH was observed

(ΔΔH=2.6-4.4 kcal/mol), while ΔS became more favourable in all cases. The reduction in affinity might explain the reduction in ΔH due to a loss of number of contacts and an increase (i.e. more favourable) in ΔS due to the higher flexibility of the peptide. However, this thermodynamic signature is contradictory to the observed N values for the SH3-1 and SH3-2 complexes with P_{79}A or F_{89}A. Based on the N values, the most likely scenario is that one SH3 domain binds to two or three peptides. However, this should have led to a ΔH increase, while ΔS should have become more unfavourable due to additional peptide and SH3 constraints.

In summary, these measurements imply that the point mutations of P_{79} and F_{89} lead to a reduction in the binding affinity between SH3-1 or SH3-2 and ARAP1e1. However, the resulting stoichiometries of these mutants cannot be explained yet. These results provide a basis for future work, which will be described in Chapter 7.

6.10 Conclusion

Bioinformatics tools and peptide arrays were combined to screen for novel putative CD2AP SH3 interaction partners containing the refined CD2AP SH3 recognition motif (Chapter 3). As described in section 1.5.2, the most important function of CD2AP is in podocyte stability. Our screen (section 6.2) tested five proteins with podocyte-specific expression and/or known podocytic functions (MYOF, NFASC, PLCE1, TJP1, UTRN). Also, ROBO1, which is predicted by an algorithm to have podocyte-specific expression was also tested (Ju et al. 2013). The first two SH3 domains bound to PLCE1 and ROBO1. Also, UTRN bound to SH3-2.

Three proteins (ARAP1, DAB1 and MLK3) were selected for further evaluation. The CD2AP/ARAP1 interaction was confirmed in a number of cell lines that were derived from different tissues. Moreover, the pulldowns of ARAP1 and MLK3 with the first two GST-SH3 domains show a trend that one domain (SH3-1 or SH3-2) drives the interaction with ARAP1 and MLK3.

SH3-2 binds to ARAP1e1 in the nanomolar range. So far, only a small number of tight SH3 binding peptides derived from wild-type proteins, have been reported (Lee et al 1995, Harkiolaki et al 2003, Mayer et al 2004). The structural view of the interaction indicates two additional anchoring residues in ARAP1 (P_{79}, F_{89}), which extend beyond the SH3 core recognition motif (P-x-P-x-x-R). The importance of the Arg was confirmed by testing the SH3 binding with an ARAP1e1 $R_{86}A$ peptide by ITC. Also, point mutations of P_{79} and F_{89} to Ala reduce the binding affinity to the first two SH3 domains. However, these ITC results should be confirmed with other methods, which will be discussed further in Chapter 7.

Lastly, the ARAP1e1 and DAB1 spot intensities of the peptide arrays did not correlate with the binding affinities measured by ITC. Peptide arrays are a semi-quantitative method and the peptides are immobilised on a solid support, while ITC is highly quantitative, with the reaction components free in solution. Therefore, it was used to validate the peptide array hits (Liu et al. 2012).

Chapter 7: General remarks and future strategy

7.1 CD2AP SH3 binding mode and specificity

SH3 domains function as components of dynamic multi-protein complexes. Due to their abundance and cross-reactivity, it is still not fully understood how their specificity is generated in cells. The elucidation of the SH3 binding modes should therefore help us to understand better complex formation and how specificity is achieved. This study addresses this question for the three CD2AP SH3 domains.

The three CD2AP SH3 domains recognise the atypical core P-x-P/A-x-x-R motif in their ligands. The crystal structures of SH3-1/RIN3e1, SH3-2/RIN3e2 and SH3-2/ARAP1e1 show that these residues serve as anchoring points. Depending on the ligands, motif positions 2 and 5 may play a role in the ligand recognition by the CD2AP SH3 domains. In motif position 5, a Pro may play a role for the interaction of the three SH3 domains with the two RIN3 epitopes, the binding of the first two SH3 domains to c-CBL and the SH3-3 interaction with ALIX. However, ARAP1e1, which lacks Pro in this position, binds with high affinity to the three SH3 domains. In motif position 2, a Thr may contribute to the SH3-3 interaction with ALIX and RIN3e1.

The SH3 domains bind their ligands in the class II orientation. The first two SH3 domains bind to the same partners more tightly than SH3-3. In general, SH3-1 binds to its interaction partner peptides with a K_d of 1.5-23 μM, SH3-2 with 0.36-27 μM and SH3-3 with 32-120 μM. However, there are exceptions. SH3-3 binds to RIN3e2 and ARAP1e1 peptides with a comparable affinity to the first two SH3 domains (Tables 3.2 and 6.1, 6.7 and 11.4 μM respectively).

Moreover, SH3-1 binds to ARAP1e3 with lower affinity (Table 6.1, 53 μM), and SH3-2 binds to MLK3e2 with similar affinity (Table 6.1, 50 μM). These results, together with the pulldown of endogenous ARAP1 with the first two SH3 domains from cell lysates (section 6.5), show that SH3-1 and SH3-2 drive mainly the interactions with the binding partners analysed so far. Furthermore, the comparable SH3-1 and SH3-2 binding affinities and the above-mentioned crystal structures indicate that the first two SH3 domains have similar binding modes.

As was shown by our crystal structures, the SH3-1 and SH3-2 ligands dock onto two hydrophobic pockets, which are formed by the SH3 domain's RT loop and 3_{10} helix, and one acidic pocket, which is formed by the RT and N-Src loops (Figures 3.17, 3.20 and 6.11). The SH3-1/ubiquitin structure indicated that the above-mentioned regions (i.e. RT and N-Src loops, and 3_{10} helix) in SH3-1 formed the ubiquitin binding region (Ortega-Roldan *et al*. 2013). The structure of mouse SH3-3 in complex with the low affinity binder, ubiquitin (K_d=132 μM), has also been solved (Ortega-Roldan *et al*. 2009) (Figure 7.1). Mouse SH3-3 is 92% identical to human. Even though the RT and N-Src loops of SH3-3 bind to ubiquitin, the 3_{10} helix does not make contacts. This is due to an extra residue in the N-Src loop of SH3-3 compared to SH3-1, which resulted in the formation of a shallow region with the RT loop, and the rearrangement of the 3_{10} helix. Furthermore, the SH3-3 surface is more negative than the SH3-1 and SH3-2 surfaces. It may therefore be that SH3-3 preferentially recognises more positively charged ligands (Roldan *et al*. 2011, Ortega-Roldan *et al*. 2013). In addition, the superimposition of the SH3-2/RIN3e2 complex onto the SH3-3/ubiquitin complex shows that the

critical residues in ubiquitin and RIN3e2 do not overlap (Figure 7.1). The differential positioning of the critical ubiquitin and RIN3e2 residues relative to the SH3-3 surface may explain the low affinity of the SH3-3/ubiquitin interaction. Taken together, the above-mentioned studies and our results indicate that SH3-3 may have different binding preferences compared to the first two SH3 domains.

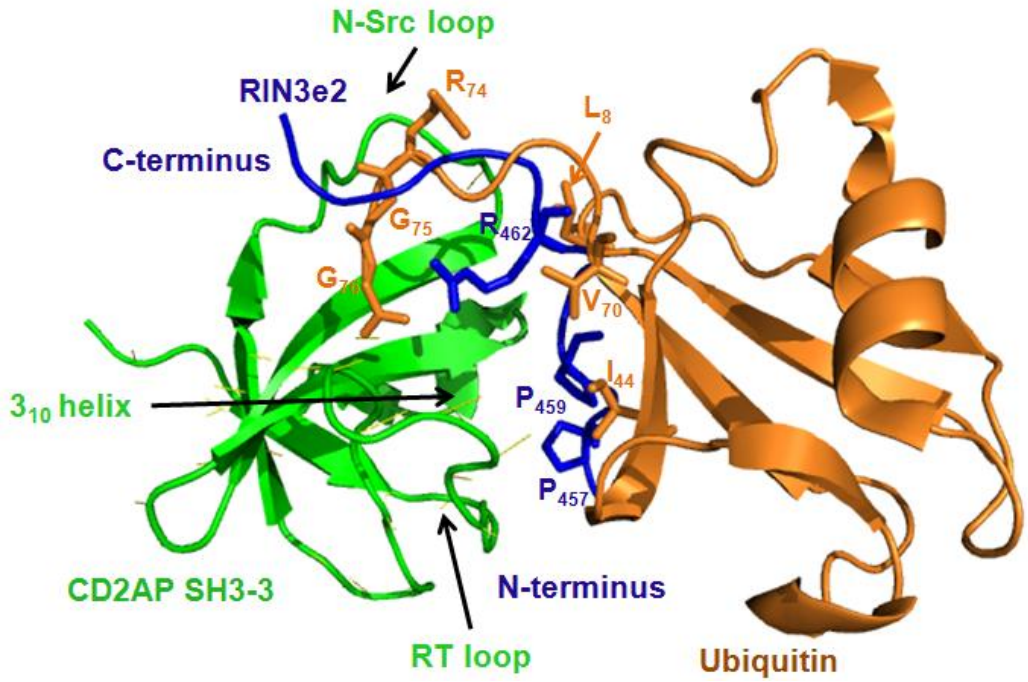


Figure 7.1

Ribbon representation of ubiquitin docking onto CD2AP SH3-3 domain and superposition of RIN3e2 on this complex.

SH3-3: green ribbon; Ubiquitin: orange ribbon (PDB code: 2LZ6); RIN3e2: blue ribbon (PDB code: 3U23). For the superimposition of RIN3e2 to the SH3-3/ubiquitin complex, a structural alignment of the SH3-2/RIN3e2 to the SH3-3/ubiquitin complex was generated. For clarity, the SH3-2 domain is not depicted. The key RIN3e2 and ubiquitin residues that form contacts with the SH3-2 and SH3-3 surfaces are labelled.

Furthermore, Ortega-Roldan *et al*. (2013) showed that a C-terminal His-tag in ubiquitin leads to an order of magnitude tighter binding to the SH3-3 domain. This suggests that the length of the C-terminus of the ligand affects SH3-3

binding. The superimposition of the SH3-2/RIN3e2 complex onto the SH3-3/ubiquitin complex shows that the C-termini of RIN3e2 and ubiquitin are positioned similarly (Figure 7.1). In addition, the truncated form of RIN3e2 binds to SH3-3 10-fold more weakly (Table 3.2). Thus, the extended C-terminus might contribute to the tighter binding between SH3-3 and RIN3e2 (Table 3.2, 6.7 μM) compared to the other peptides. Even though the ALIX permutation arrays indicate that SH3-3 may recognise an extensive motif at the C-terminus of ALIX (Figure 1.13), the SH3-3 affinity for the 21 amino acid ALIX peptide was in the mid-range (Table 3.2, 32 μM). Therefore, the SH3-3 preference for an extended C-terminus of the P-x-P/A-x-x-R motif may only apply to a subset of proteins.

Preliminary binding experiments with SH3-(1+2) and RIN3e1 or RIN3e2 peptides indicate that there are not major avidity effects of the first two SH3 domains with the RIN3 peptides. Also, the SH3-(1+2) binding to one RIN3e1 or possibly two RIN3e2 suggests that there may be additional binding preferences between the first two SH3 domains in the context of full-length CD2AP.

The SH3-2 domain binary crystal structure presented herein is the first reported X-ray structure of this domain. It was recently used to aid reconstruction of a low-resolution model of full-length CD2AP obtained by electron microscopy. This structure suggests that CD2AP forms tetramers through its coil-coiled region, and that the SH3 domains in all molecules are accessible for binding (Adair *et al*. 2014). Our study indicates several ways by which CD2AP SH3 specificity can be achieved. As described above, SH3-3 may have different binding preferences compared to the first two SH3

domains. Moreover, the recognition of multiple epitopes within a protein by the three SH3 domains implies that more than one contact contributes to CD2AP binding to its partners. This would result in a high local concentration, which would regulate complex formation. One method to investigate further SH3 domain cooperativity is by *in vivo* chemical photo-cross-linking (Suchanek *et al*. 2005, Okada *et al*. 2011). Cells could be transfected with constructs expressing the CD2AP SH3 domains individually or in tandem. Next, their covalently trapped interactions in cells could be identified. As a result, it could be tested if differences in the CD2AP SH3-mediated interactions by individual or tandem CD2AP SH3 domains exist. Intriguingly, the combination of positive and negative selection determinants to screen for novel CD2AP SH3 binding candidates did not result in a high success rate (section 6.2). This suggests that additional mechanisms, except positive and negative selection, contribute to the recognition of a ligand by an SH3 domain.

Another mechanism for generating specificity is compartmentalisation. Cell type-specific interactions or spatiotemporal restrictions may contribute to CD2AP SH3 complex formation with different partners, which could result in several complexes of distinct functions. Our failure to detect endogenous CD2AP/RIN3 and CD2AP/MLK3 complexes might be due to compartmentalisation issues in the absence of appropriate physiological stimuli.

7.2 Possible functions of the CD2AP complexes with RIN3, ARAP1, and MLK3, and future work

7.2.1 CD2AP/RIN3

RIN3 is a member of the RIN (Ras and Rab interactors) family (section 1.6.4). It is suggested to act as a Rab5 guanine nucleotide exchange factor (GEF) in a tyrosine-phosphorylation dependent manner (Kajiho *et al*. 2003, Yoshikawa *et al*. 2008). Additionally, RIN3 is proposed to act as a Rab31 (or Rab22B) GEF (Kajiho *et al*. 2011). Rab5 regulates endosome biogenesis. Thus, Rab5 is critical for cargo sorting in endocytic processes (Zeigerer *et al*. 2012). The mechanism of Rab31 action remains unclear. Some scientists support that Rab31 is involved in membrane trafficking at the trans-Golgi network (Ng *et al*. 2007) and in nerve growth factor (NGF) signalling (Wang *et al*. 2011).

Moreover, RIN3 was reported to be enriched in mast cells and participates in the downregulation of c-KIT receptor (Janson *et al*. 2012). Mast cells are suggested to be involved in immune responses, angiogenesis and tumour growth (Gilfillan *et al*. 2011). The c-KIT receptor is critical for mast cell survival and differentiation (Ribatti *et al*. 2014). Furthermore, the locus near RIN3 was found through genome-wide studies to be involved in chronic obstructive pulmonary disease (Cho *et al*. 2014) and Alzheimer's disease (Lambert *et al*. 2013). As mentioned in section 1.5.4.3, CD2AP has also been implicated by genome-wide studies with Alzheimer's disease (Naj *et al*. 2011, Hollingworth *et al*. 2011, Lambert *et al*. 2013).

Regarding future work for the CD2AP/RIN3 complex, the CD2AP and RIN3 expression could be studied in immortalised human cell lines derived from

brain and mast cells. If both proteins are present, immunoprecipitation assays at the endogenous level could be performed to confirm their interaction. Next, localization studies, such as the proximity ligation assay (Fredriksson *et al*. 2002, Soderberg *et al*. 2006), should be performed to study this complex further. Another possible way of studying the CD2AP/RIN3 interaction would be by measuring RIN3 enzymatic activity. However, this approach would be successful only if the RIN3 interaction with CD2AP affects its GEF activity. This could be done by cell-free or *in cell* GEF activity assays (Saito *et al.* 2005, Bliss *et al.* 2006).

Apart from the two CD2AP SH3 P-x-P/A-x-x-R recognition motifs in RIN3, there is one potential binding motif present in both RIN2 (aa 335-361) and RINL (aa 192-198). The screen by Dr Tassos Konstantinou (section 1.6.2) showed that RIN2 binds to CD2AP SH3-1 and SH3-2 domains. Therefore, the above-mentioned methods could be applied to study any possible interactions between CD2AP and different RIN family members.

7.2.2 CD2AP/ARAP1

ARAP1 (Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1) is a member of the ARAP subfamily, which belongs to ArfGAPs (ADP-ribosylation factor GTPase activating proteins). ArfGAPs are scaffold proteins that are involved in signalling processes, such as membrane trafficking (Miura *et al*. 2002, Inoue *et al*. 2007). ARAP1 is an Arf1 and Arf5 GAP (Hasagawa *et al*. 2012).

The CIN85/ARAP1 complex might participate in EGFR recycling and ARAP1e1 is the CIN85 SH3 binding epitope (Yoon *et al*. 2011). CD2AP is

also involved in EGFR internalization (Lynch *et al*. 2003). The CD2AP/ARAP1 and CIN85/ARAP1 complexes might have similar functions. The ARAP1e1 residues P_{81}, P_{83}, R_{86} and R_{90} are critical for the CIN85/ARAP1 interaction (Yoon *et al*. 2011). We confirmed the importance of R_{86} for the CD2AP SH3/ARAP1 interaction. $R_{86}A$ led to a significant decrease in the SH3 affinity for ARAP1e1 (Table 6.2). The loss of ARAP1e2 and ARAP1e3 did not affect the CIN85/ARAP1 interaction (Yoon *et al*. 2011). The 10- to 100-fold difference in CD2AP SH3 affinity for ARAP1e1 compared to ARAP1e2 and e3 implies that ARAP1e1 drives the CD2AP/ARAP1 interaction. Finally, there are no data published on whether P_{79} and F_{89} might contribute to the CIN85/ARAP1 interaction as we found for the CD2AP/ARAP1 interaction (section 6.4.3).

Future experiments to study the CD2AP/ARAP1 complex could be to:

(i) generate point mutations ($P_{79}A$, $F_{89}A$ and $P_{79}A/F_{89}A$) in the ARAP1 protein, and test if there is an effect on CD2AP binding in cells.

(ii) perform localization studies to confirm the CD2AP/ARAP1 interaction and test whether the complex is involved in EGFR internalization or recycling.

(iii) knock down ARAP1 in podocytes, and test the cell viability and morphology.

(iv) test the importance of the CD2AP/ARAP1 interaction in podocytes. This could be tested by knocking down ARAP1 and knocking in ARAP1-$P_{81}A/P_{83}A/R_{86}A$. Then, cell viability and morphology should be analysed.

7.2.3 CD2AP/MLK3

MLK3 (mixed lineage 3) or MAP3K11 (Mitogen-activated protein kinase kinase kinase 11) is a serine-threonine protein kinase, which activates mitogen-activated kinases, which activates further the JNK, ERK and p38 pathways (Gallo *et al*. 2002). MLK3 is involved in a number of cellular processes, such as neuronal cell death, cell proliferation and migration (Chadee 2013). It might also have roles in inflammation and immunity (Handley *et al*. 2007).

MLK3 activates the JNK pathway in neuronal cells among others through its association with c-CBL and CBL-B (Sproul *et al*. 2009), which are CD2AP SH3 interaction partners (Table 1.3). Moreover, an MLK3 $R_{799}C$ mutation (critical Arg in MLK3e2) was found to have tumourigenic capacity in gastrointestinal tumours (Velho *et al* 2010). Therefore, in future work, it might be worth investigating the CD2AP/MLK3 interaction in gastrointestinal cancer or neuronal cell lines with immunoprecipitation assays at the endogenous level. Since CD2AP SH3 domains interact with c-CBL and CBL-B (Table 1.2), it could be further tested in neuronal cells whether a CD2AP/MLK3/c-CBL/CBL-B cascade is present.

7.3 Concluding remarks

In conclusion, this thesis resulted in the elucidation of the CD2AP SH3 binding mode to RIN3, a new interaction partner, through detailed biophysical, structural and biochemical studies. Next, with refinement of the CD2AP SH3 recognition motif, taking into account both positive and negative selection determinants, two additional, novel CD2AP SH3 interaction partners, namely

ARAP1 and MLK3, were identified. Biophysical analyses of the CD2AP/ARAP1 interaction indicated that there may be additional anchoring residues in ARAP1 that extend beyond the known SH3 recognition motif. These residues confer higher affinity binding to CD2AP than previously reported with known interactors, in the nanomolar range. Lastly, the CD2AP interactions with ARAP1, MLK3 and RIN3 were investigated in a number of cell lines derived from different tissues.

Therefore, this DPhil work expands our knowledge of the CD2AP SH3 interactome and provides a greater insight into CD2AP function.

Appendices

Appendix A: Isothermal titration calorimetry measurements

A.1. Isothermal titration calorimetry curves

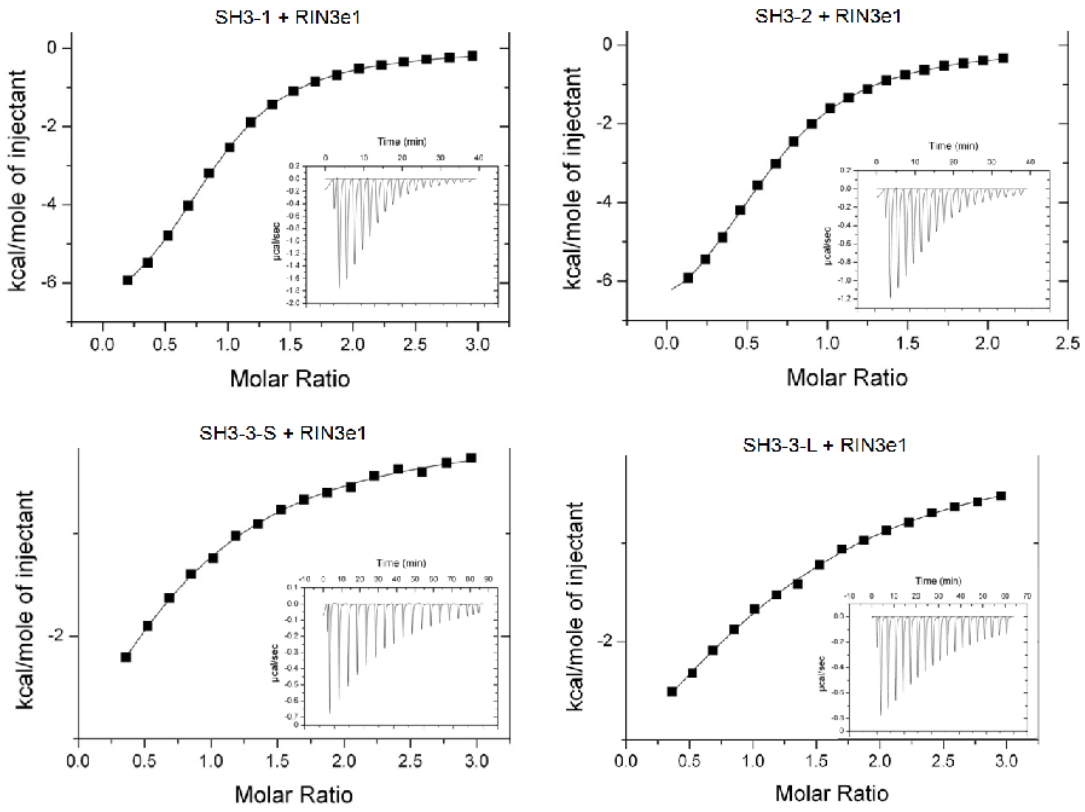


Figure A.1

ITC curves for binding of the RIN3e1 peptides to CD2AP SH3 domains.

Isotherms calculated from raw data. Insets show the integrated heats for binding.

Appendix B: Protein X-ray crystallography

B.1 Structure of CD2AP SH3-1 domain in complex with RIN3e1

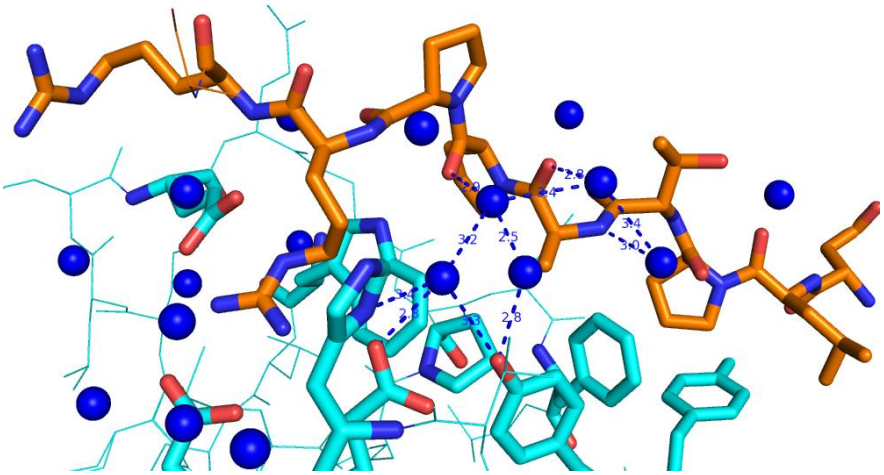


Figure B.1

Network of water molecules at the interface of SH3-1 and RIN3e1.

The RIN3e1 peptide and the critical SH3-1 residues are shown in stick representation and coloured by element (carbon; orange for RIN3e1 and cyan for SH3-1; nitrogen: blue; oxygen: red). Water molecules are shown as blue spheres and their possible hydrogen bonds to the SH3 domain or RIN3e1 peptide are shown as dashed lines for a selection of them, with the distances in Ångstrom displayed.

B.2 Structure of CD2AP SH3-2 domain in complex with RIN3e2

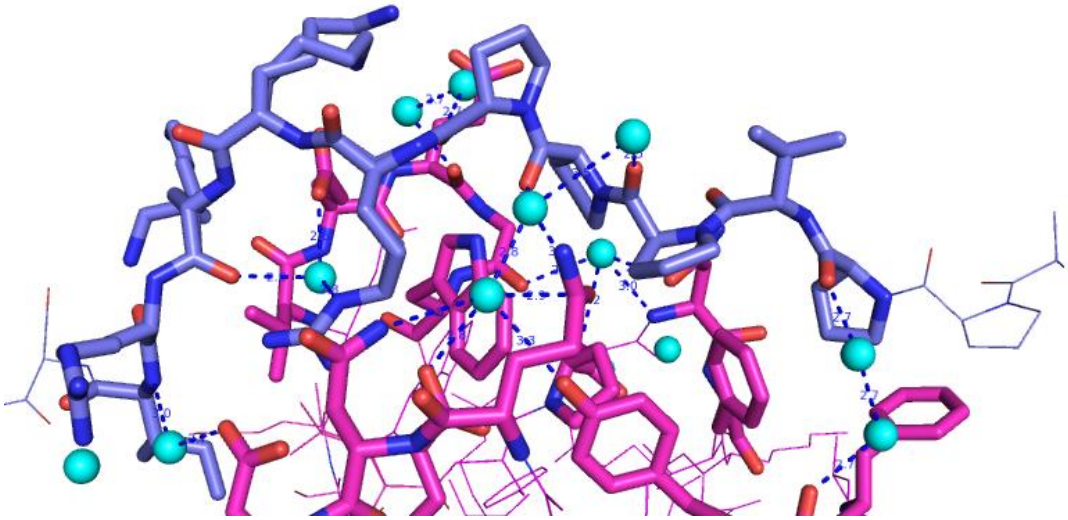


Figure B.2

Network of water molecules at the interface of SH3-2 and RIN3e2.

The RIN3e2 peptide and the critical SH3-2 residues are shown in stick representation and coloured by element (carbon; light blue for RIN3e1 and magenta for SH3-1; nitrogen: blue; oxygen: red). Water molecules are shown as blue spheres and their possible hydrogen bonds to the SH3 domain or RIN3e2 peptide are shown as dashed lines for a selection of them, with the distances in Ångstrom displayed.

B.3 Structure of CD2AP SH3-2 in complex with ARAP1e1

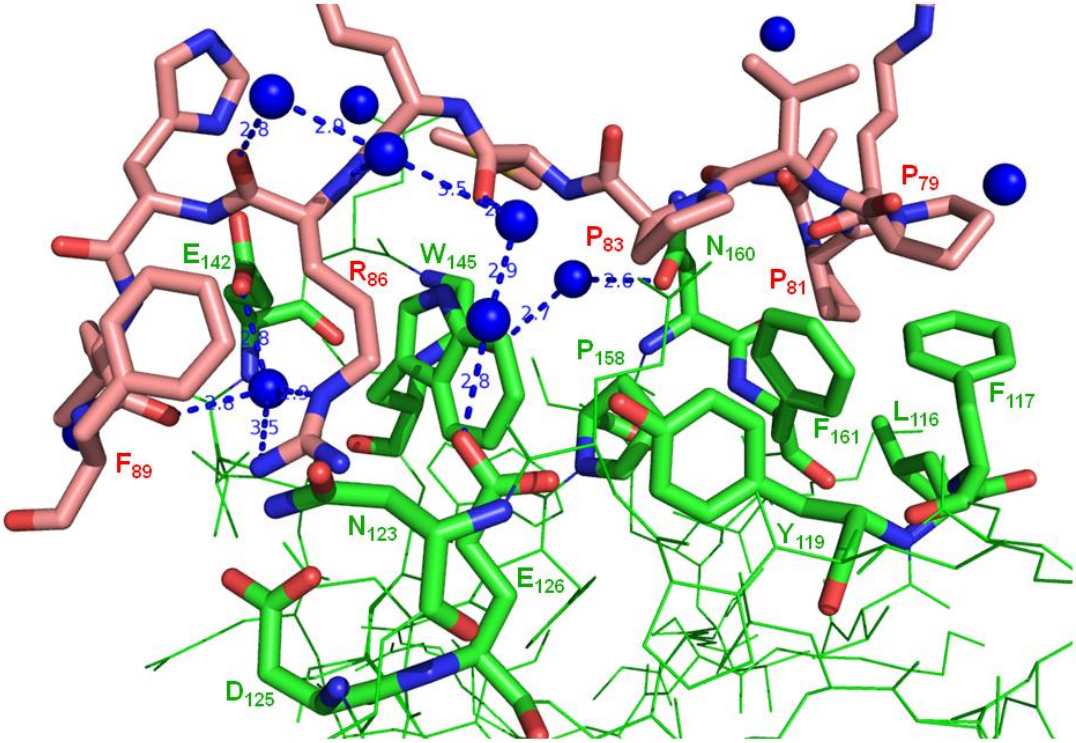


Figure B.3

Network of water molecules at the interface of SH3-2 and ARAP1e1.

The ARAP1e1 peptide and the critical SH3-2 residues are shown in stick representation and coloured by element (carbon; light blue for RIN3e1 and magenta for SH3-1; nitrogen: blue; oxygen: red). Water molecules are shown as blue spheres and their possible hydrogen bonds to the SH3 domain or ARAP1e1 peptide are shown as dashed lines for a selection of them, with the distances (in Å) displayed.

Appendix C: Peptide arrays

C.1 Synthesis of peptide arrays

The RIN3 peptide permutation arrays were prepared on cellulose-(3-amino-2-hydroxy-propyl)-ether membranes (Landgraf *et al*. 2004) (CAPE membranes) using a SPOT robot (Intavis AG, Köln, Germany) according to the company's standard spot synthesis protocol. The array design was generated using the software LISA. The CAPE membranes were used because of the better signal to noise ratio in incubation experiments compared to a standard β-alanine cellulose membrane (Landgraf *et al.* 2004, Tonikian *et al*. 2009).

Apart from the RIN3 peptide permutation arrays, all arrays were synthesized on a Multipep Synthesizer (Intavis Bioanalytical Instruments, Cologne, Germany) on derivatized cellulose (amino–Peg500 UC540, acid-hardened, loading 400 nmol/cm^2; Intavis). After peptides were spot-synthesised and deprotected, membranes were washed extensively with dichlormethane, N-methylpyrolidone, and ethanol.

C.2 Probing of peptide arrays by previous lab members

C.2.1 c-CBL permutation arrays

These arrays were probed by Dr Tassos Konstantinou (previous lab member). 21 aa peptides of the CBL epitope GSQVPERPPKPFPRRINSERK were permutated in single positions (underlined) to all other residues, in duplicate rows. Membranes were probed as described in section 2.2.8.3.2 but with different GST and GST-SH3 concentration (0.56 µM).

C.2.2 ALIX permutation arrays

These arrays were probed by Dr Tassos Konstantinou. Spot-synthesised 21 aa peptides of the ALIX epitope AGGHAPTPPTPAPRTMPPTKP were permutated in single positions (underlined) to all 20 other residues, in duplicate rows. Membranes were probed as described in section C.2.1.

C.2.3 Proteome-wide search for novel putative CD2AP SH3 binding partners

These arrays were probed by Dr Tassos Konstantinou. The P-x-P-x-P-R motif in c-CBL, found to be the dominant linear recognition motif for CD2AP SH3-1 and -2, was used as a search sequence for scanning the human proteome with ScanProsite (de Castro *et al.* 2006). From 700 hits, a subset of 70 potential binding sites in proteins with known or possible roles in signalling was selected. Preference was given to proteins also involved in membrane trafficking. These selected CD2AP SH3 binding sequences were then spot-synthesized on a microarray as 21 aa peptides, in duplicate, for validation of their in-vitro binding ability. The arrays were probed as described in section C.2.1.

C.2.4 RIN3 scanning array

The RIN3 scanning arrays were probed by Dr Melanie Janning (previous lab member). The entire human RIN3 protein sequence was spot-synthesised as partially overlapping 27 aa peptides, sliding 3 aa with each step. Three array copies were probed as described in section C.2.1, except that 0.1 µM (2.6 µg/ml) GST or 0.1 µM (3.6 µg/ml) GST-SH3 were used, in the same blocking buffer (TBST, 5% [w/v] non-fat dry milk).

C.3 Search for optimal blocking buffer for the peptide permutation arrays of both RIN3 epitopes

The conditions described in Figure C.1 were tested to determine a good blocking buffer for probing the RIN3 peptide permutation arrays. As illustrated in Figure C.1, the background was low in the case of using 5% (w/v) non-fat dry milk or 2% (w/v) ovalbumin in RIPA 100 as a blocking buffer (Figures C.1D, C.1E). The latter was chosen as the appropriate blocking buffer because it gave the best signal-to-background ratio.

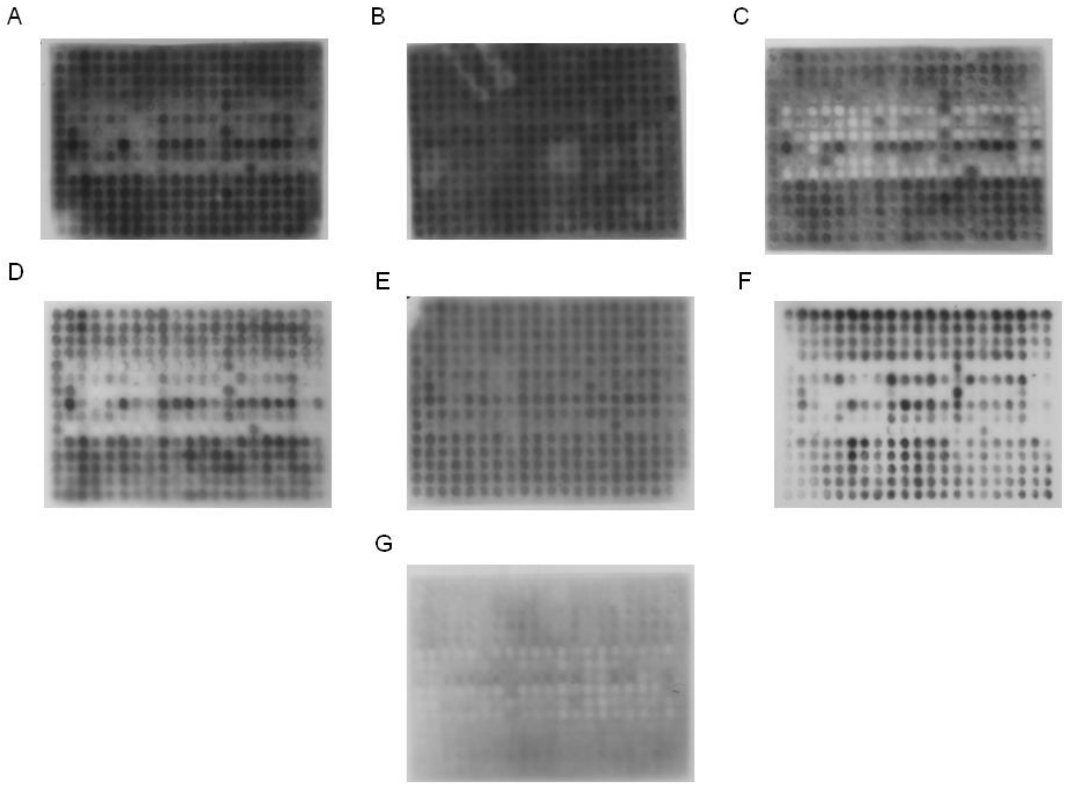


Figure C.1

Optimisation of the blocking buffer conditions for the peptide permutation arrays of RIN3e1 (A, C, E) and RIN3e2 (B, D, F).

The arrays were probed with 0.1 μM GST-SH3-1. Different blocking buffer was applied to each array: ***A***, E.coli lysate with 1% Triton X-100 in TBST. ***B***, 2% (w/v) BSA (Bovine Serum Albumine) in TBST. ***C***, 2% (w/v) ovalbumin in TBST. ***D***, 5% (w/v) non-fat dry milk in RIPA 100. ***E***, 10x block buffer (Sigma) with 5% (w/v) sucrose in TBST. ***F***, 2% (w/v) ovalbumin in RIPA 100. ***G***, 5% (w/v) non-fat dry milk in TBST. All the exposures were 10 seconds.

Appendix D: Immunoblotting

D.1 Analysis of ARAP1 and MLK3 protein expression with commercial antibodies

The ARAP1 and MLK3 expression was studied in HEK293, HeLa and MCF-7 cell extracts. NP-40 or TX-100-buffer extracts from the above mentioned cell lines were tested by immunoblotting with two anti-ARAP1 (NB100-68223, NB110-68801) and one anti-MLK3 (ab51068) antibodies. The antibodies are described in detail in section 2.1.1. Figure D.1 depicts representative blots.

According to the Uniprot database, ARAP1 has seven potential splice variants (77, 127, 135, 136, 137, 160, 162 kDa). The 136, 137, 160 and 162 kDa splice variants have already been experimentally confirmed. In the case of NB100-68223 (Figure D.1A), bands of predicted size, whose molecular weights ranged between 76 and 160 kDa, were found. A 38 kDa band of high intensity was also observed. In the case of NB110-68801 (Figure D.1C), bands of predicted size were visualised, which may correspond to the 77, 160 and 162 kDa splice variants. Bands of similar intensity, whose molecular weights range between 30 and 50 kDa, were also found. Both buffers (NP-40 and TX-100) extracted the same immunoreactive proteins. The three ARAP1 epitopes, which are recognised by the CD2AP SH3 domains, are found in the 160 and 162 kDa splice variants. A band of similar molecular weight was detected in sections 6.5-6.6.

MLK3 has one isoform of 92 kDa (Uniprot database). When immunoblotting was performed with ab51068 (Figure D.1B), one major band between the 76 and 102 kDa markers was present. Some MLK3 was not extracted from the

lysates with both extraction buffers. A 130 kDa band, which is potentially not specific, was found in the RIPA 100 extracts of the pellets. In summary, anti-ARAP1 NB100-68223 antibody is more specific than NB110-68801. The anti-MLK3 antibody is specific.

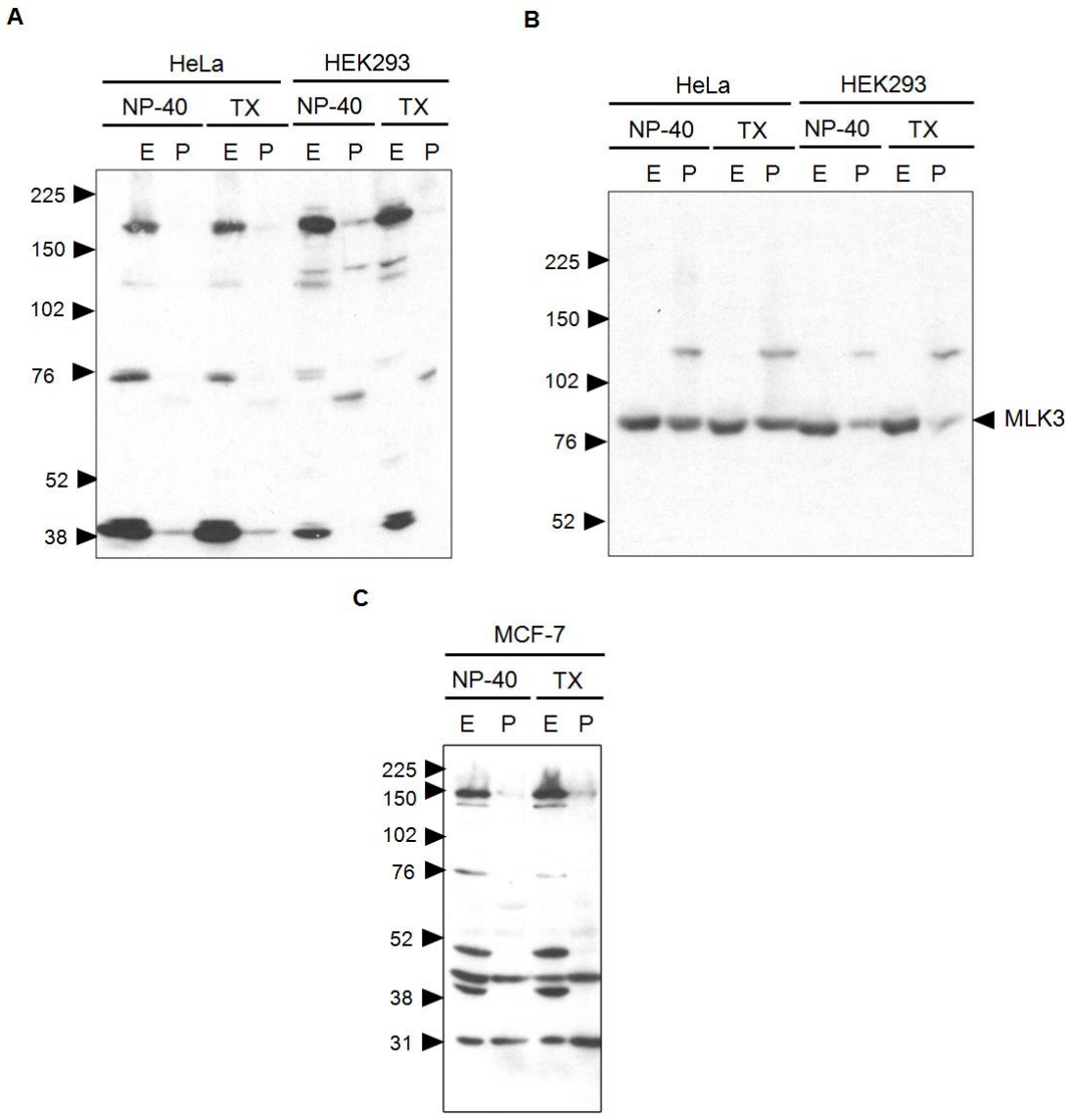


Figure D.1

Antibody testing by immunoblotting in HEK293, HeLa and MCF-7 cell lines.

A, NB100-68223 (anti-ARAP1 antibody) was tested in HEK293 and HeLa cell extracts. ***B***, ab51068 (anti-MLK3 antibody) was tested in the HEK293 and HeLa cell extracts. ***C***, NB110-68801 (anti-ARAP1 antibody) was tested in MCF-7 cell extracts. In all cases, the proteins were extracted from the lysates by either NP-40 or TX-100-buffer and both extracts (denoted as 'E') or the RIPA 100 extract of the pellet (denoted as 'P') were analysed by immunoblotting.

Appendix E: List of hits of the peptide array screen for novel CD2AP SH3 interacting partners

The peptide arrays, which were probed with the CD2AP SH3 domains, are found in Figure 6.1. The list of hits is found below:

Protein	Description	CD2AP SH3 hit
Generic binding protein		
BRD1	bromodomain containing 1	SH3-1, SH3-2
DAB1	disabled homolog 1	SH3-1, SH3-2, SH3-3
DLG5	discs, large homolog 5	SH3-1, SH3-2
DSCAML1	Down syndrome cell adhesion molecule like 1	SH3-1, SH3-2
ICAM5	intercellular adhesion molecule 5, telencephalin	SH3-1, SH3-2, SH3-3
LDB3	LIM domain binding 3	SH3-1, SH3-2, SH3-3
JPH3	junctophilin 3	SH3-1, SH3-2
MSH6	mutS homolog 6	SH3-1, SH3-2
MYO15A	myosin XVA	SH3-1, SH3-2, SH3-3
NFATC2IP	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 interacting protein	SH3-1, SH3-2
PBRM1	polybromo 1	SH3-1, SH3-2, SH3-3
REC8	REC8 homolog	SH3-1, SH3-2, SH3-3
RPGRIP1L	RPGRIP1-like	SH3-1, SH3-2
SCRIB	scribbled homolog	SH3-1, SH3-2
SMN1	survival of motor neuron 1, telomeric	SH3-1, SH3-2
SORBS1	sorbin and SH3 domain containing 1	SH3-1, SH3-2
TNXB	tenascin XB	SH3-1, SH3-2
Generic channel		
PKD1	polycystic kidney disease 1	SH3-1, SH3-2, SH3-3
Generic enzyme		
PJA1	praja ring finger 1, E3 ubiquitin protein ligase	SH3-1, SH3-2
UBE2H	ubiquitin-conjugating enzyme E2H	SH3-1, SH3-2
Kinases		
ABL2	v-abl Abelson murine leukemia viral oncogene homolog 2	SH3-1, SH3-2
LATS2	LATS, large tumor suppressor, homolog 2	SH3-1, SH3-2

MAP3K1	mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase	SH3-1, SH3-2
MAP3K3	mitogen-activated protein kinase kinase kinase 3	SH3-1, SH3-2
MLK3	mitogen-activated protein kinase kinase kinase 11	SH3-1, SH3-2, SH3-3
RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	SH3-1, SH3-2
Known/predicted podocytic function		
PLCE1	phospholipase C, epsilon 1	SH3-1, SH3-2
ROBO1	roundabout, axon guidance receptor, homolog 1	SH3-1, SH3-2
UTRN	utrophin	SH3-1, SH3-2
Lipid phosphatase		
SYNJ1	synaptojanin 1	SH3-1, SH3-2, SH3-3
Protease		
SOLH	small optic lobes homolog (Drosophila)	SH3-1, SH3-2
Receptors		
BMPR1B	bone morphogenetic protein receptor, type IB	SH3-1, SH3-2
CELSR2	cadherin, EGF LAG seven-pass G-type receptor 2	SH3-1, SH3-2, SH3-3
CELSR3	cadherin, EGF LAG seven-pass G-type receptor 3	SH3-1, SH3-2
PDGFRB	platelet-derived growth factor receptor, beta polypeptide	SH3-1, SH3-2
Regulators (GTPase, GDI, GAP, GEF)		
ALS2	amyotrophic lateral sclerosis 2	SH3-1, SH3-2, SH3-3
ARAP1	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	SH3-1, SH3-2, SH3-3
ARAP3	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 3	SH3-1, SH3-2, SH3-3
ARHGAP4	Rho GTPase activating protein 4	SH3-1, SH3-2
ARHGAP10	Rho GTPase activating protein 10	SH3-1, SH3-2
DLC1	deleted in liver cancer 1	SH3-1, SH3-2
DNM1	dynamin 1	SH3-1, SH3-2
DNMBP	dynamin binding protein	SH3-1, SH3-2
KALRN	kalirin, RhoGEF kinase	SH3-1, SH3-2
OBSCN	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	SH3-1, SH3-2
OPHN1	oligophrenin 1	SH3-1, SH3-2
RALGDS	ral guanine nucleotide dissociation stimulator	SH3-1, SH3-2
TRIO	trio Rho guanine nucleotide exchange factor	SH3-1, SH3-2

TSC2	tuberous sclerosis 2	SH3-1, SH3-2, SH3-3
Transcription factor		
MSL3	male-specific lethal 3 homolog	SH3-1, SH3-2

References

UniProt database (2014). Activities at the Universal Protein Resource (UniProt). Nucleic acids research *42*, D191-198.

Ababou, A., Pfuhl, M., and Ladbury, J.E. (2009). Novel insights into the mechanisms of CIN85 SH3 domains binding to Cbl proteins: solution-based investigations and in vivo implications. Journal of molecular biology *387*, 1120-1136.

Adair, B.D., Altintas, M.M., Moller, C.C., Arnaout, M.A., and Reiser, J. (2014). Structure of the Kidney Slit Diaphragm Adapter Protein CD2-Associated Protein as Determined with Electron Microscopy. Journal of the American Society of Nephrology : JASN *25*, 1465-1473.

Adzhubei, A., Eisenmenger, F., Tumanyan, V., Zinke, M., Brodzinski, S., and Esipova, N. (1987). Third type of secondary structure: Noncooperative mobile conformation. Protein Data Bank analysis. Biochemical and Biophysical Research Communications *146*, 934-938.

Adzhubei, A.A., Sternberg, M.J., and Makarov, A.A. (2013). Polyproline-II helix in proteins: structure and function. Journal of molecular biology *425*, 2100-2132.

Aitio, O., Hellman, M., Kazlauskas, A., Vingadassalom, D.F., Leong, J.M., Saksela, K., and Permi, P. (2010). Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. Proceedings of the National Academy of Sciences of the United States of America *107*, 21743-21748.

Aitio, O., Hellman, M., Kesti, T., Kleino, I., Samuilova, O., Paakkonen, K., Tossavainen, H., Saksela, K., and Permi, P. (2008). Structural basis of PxxDY motif recognition in SH3 binding. Journal of molecular biology *382*, 167-178.

Akchurin, O., and Reidy, K.J. (2014). Genetic causes of proteinuria and nephrotic syndrome: Impact on podocyte pathobiology. Pediatric nephrology (Berlin, Germany), 1-13.

Aloy, P., and Russell, R.B. (2006). Structural systems biology: modelling protein interactions. Nature Reviews Molecular Cell Biology *7*, 188-197.

Arold, S., Franken, P., Strub, M.P., Hoh, F., Benichou, S., Benarous, R., and Dumas, C. (1997). The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling. Structure (London, England : 1993) *5*, 1361-1372.

Asanuma, K., Campbell, K.N., Kim, K., Faul, C., and Mundel, P. (2007). Nuclear relocation of the nephrin and CD2AP-binding protein dendrin promotes apoptosis of podocytes. Proceedings of the National Academy of Sciences of the United States of America *104*, 10134-10139.

Atatreh, N., Stojkoski, C., Smith, P., Booker, G.W., Dive, C., Frenkel, A.D., Freeman, S., and Bryce, R.A. (2008). In silico screening and biological evaluation of inhibitors of Src-SH3 domain interaction with a proline-rich ligand. Bioorganic & medicinal chemistry letters *18*, 1217-1222.

Audet, M., and Bouvier, M. (2012). Restructuring G-protein- coupled receptor activation. Cell *151*, 14-23.

Badour, K., Zhang, J., Shi, F., McGavin, M.K., Rampersad, V., Hardy, L.A., Field, D., and Siminovitch, K.A. (2003). The Wiskott-Aldrich syndrome protein acts downstream of CD2 and the CD2AP and PSTPIP1 adaptors to promote formation of the immunological synapse. Immunity *18*, 141-154.

Ball, L.J., Kuhne, R., Hoffmann, B., Hafner, A., Schmieder, P., Volkmer-Engert, R., Hof, M., Wahl, M., Schneider-Mergener, J., Walter, U.*, et al.* (2000). Dual epitope recognition by the VASP EVH1 domain modulates polyproline ligand specificity and binding affinity. The EMBO journal *19*, 4903-4914.

Bao, M., Hanabuchi, S., Facchinetti, V., Du, Q., Bover, L., Plumas, J., Chaperot, L., Cao, W., Qin, J., Sun, S.C.*, et al.* (2012). CD2AP/SHIP1 complex positively regulates plasmacytoid dendritic cell receptor signaling by inhibiting the E3 ubiquitin ligase Cbl. Journal of immunology (Baltimore, Md. : 1950) *189*, 786-792.

Barratt, E., Bronowska, A., Vondrášek, J., Černý, J., Bingham, R., Phillips, S., and Homans, S.W. (2006). Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor. Journal of molecular biology *362*, 994-1003.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M.*, et al.* (2013). NCBI GEO: archive for functional genomics data sets-update. Nucleic acids research *41*, D991-995.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image

processing with MOSFLM. Acta crystallographica. Section D, Biological crystallography *67*, 271-281.

Bedford, M.T., Frankel, A., Yaffe, M.B., Clarke, S., Leder, P., and Richard, S. (2000). Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. The Journal of biological chemistry *275*, 16030-16036.

Beneken, J., Tu, J.C., Xiao, B., Nuriya, M., Yuan, J.P., Worley, P.F., and Leahy, D.J. (2000). Structure of the Homer EVH1 Domain-Peptide Complex Reveals a New Twist in Polyproline Recognition. Neuron *26*, 143-154.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bissig, C., and Gruenberg, J. (2014). ALIX and the multivesicular endosome: ALIX in Wonderland. Trends in cell biology *24*, 19-25.

Bliss, J.M., Venkatesh, B., and Colicelli, J. (2006). The RIN family of Ras effectors. Methods Enzymol *407*, 335-344.

Blomberg, N., Baraldi, E., Nilges, M., and Saraste, M. (1999). The PH superfold: a structural scaffold for multiple functions. Trends in biochemical sciences *24*, 441-445.

Bork, P., and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? Trends in biochemical sciences *19*, 531-533.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry *72*, 248-254.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S.*, et al.* (1998a). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica. Section D, Biological crystallography *54*, 905-921.

Brunger, A.T., Adams, P.D., and Rice, L.M. (1998b). Recent developments for the efficient crystallographic refinement of macromolecular structures. Current opinion in structural biology *8*, 606-611.

Buday, L., and Tompa, P. (2010). Functional classification of scaffold proteins and related molecules. FEBS journal *277*, 4348-4355.

Calabia-Linares, C., Robles-Valero, J., de la Fuente, H., Perez-Martinez, M., Martin-Cofreces, N., Alfonso-Perez, M., Gutierrez-Vazquez, C., Mittelbrunn, M., Ibiza, S., Urbano-Olmos, F.R.*, et al.* (2011). Endosomal clathrin drives actin accumulation at the immunological synapse. Journal of cell science *124*, 820-830.

Calco, G.N., Stephens, O.R., Donahue, L.M., Tsui, C.C., and Pierchala, B.A. (2014). CD2-associated protein (CD2AP) enhances casitas B lineage lymphoma-3/c (Cbl-3/c)-mediated Ret isoform-specific ubiquitination and degradation via its amino-terminal Src homology 3 domains. The Journal of biological chemistry *289*, 7307-7319.

Camara-Artigas, A., Martinez-Rodriguez, S., Ortiz-Salmeron, E., and Martin-Garcia, J.M. (2014). 3D domain swapping in a chimeric c-Src SH3 domain takes place through two hinge loops. Journal of structural biology *186*, 195-203.

Carducci, M., Perfetto, L., Briganti, L., Paoluzi, S., Costa, S., Zerweck, J., Schutkowski, M., Castagnoli, L., and Cesareni, G. (2012). The protein interaction network mediated by human SH3 domains. Biotechnology advances *30*, 4-15.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Castagnoli, L., Costantini, A., Dall'Armi, C., Gonfloni, S., Montecchi-Palazzi, L., Panni, S., Paoluzi, S., Santonico, E., and Cesareni, G. (2004). Selectivity and promiscuity in the interaction network mediated by protein recognition modules. FEBS letters *567*, 74-79.

Ceregido, M.A., Garcia-Pino, A., Ortega-Roldan, J.L., Casares, S., Lopez Mayorga, O., Bravo, J., van Nuland, N.A., and Azuaga, A.I. (2013). Multimeric and differential binding of CIN85/CD2AP with two atypical proline-rich sequences from CD2 and Cbl-b*. The FEBS journal *280*, 3399-3415.

Chadee, D.N. (2013). Involvement of mixed lineage kinase 3 in cancer. Canadian journal of physiology and pharmacology *91*, 268-274.

Chan, B., Lanyi, A., Song, H.K., Griesbach, J., Simarro-Grande, M., Poy, F., Howie, D., Sumegi, J., Terhorst, C., and Eck, M.J. (2003). SAP couples Fyn to SLAM immune receptors. Nature cell biology *5*, 155-160.

Chandra, B.R., Gowthaman, R., Akhouri, R.R., Gupta, D., and Sharma, A. (2004). Distribution of proline-rich (PxxP) motifs in distinct proteomes: functional and therapeutic implications for malaria and tuberculosis. Protein Engineering Design and Selection *17*, 175-182.

Chen, H.I., and Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. Proceedings of the National Academy of Sciences of the United States of America *92*, 7819-7823.

Chen, K., Ruff, S.J., and Cohen, S. (1997). Peroxovanadate Induces Tyrosine Phosphorylation of Multiple Signaling Proteins in Mouse Liver and Kidney. Journal of Biological Chemistry *272*, 1263-1267.

Cho, M.H., McDonald, M.L., Zhou, X., Mattheisen, M., Castaldi, P.J., Hersh, C.P., Demeo, D.L., Sylvia, J.S., Ziniti, J., Laird, N.M.*, et al.* (2014). Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. The Lancet. Respiratory medicine *2*, 214-225.

Chong, P.A., Lin, H., Wrana, J.L., and Forman-Kay, J.D. (2010). Coupling of tandem Smad ubiquitination regulatory factor (Smurf) WW domains modulates target specificity. Proceedings of the National Academy of Sciences of the United States of America *107*, 18404-18409.

Collaborative, C.P. (1994). The CCP4 suite: programs for protein crystallography. Acta crystallographica. Section D, Biological crystallography *50*, 760.

Cormont, M., Meton, I., Mari, M., Monzo, P., Keslair, F., Gaskin, C., McGraw, T.E., and Le Marchand-Brustel, Y. (2003). CD2AP/CMS regulates endosome morphology and traffic to the degradative pathway through its interaction with Rab4 and c-Cbl. Traffic *4*, 97-112.

De Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., and Hulo, N. (2006). ScanProsite:

detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic acids research *34*, W362-365.

De St Groth, S.F., Webster, R., and Datyner, A. (1963). Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. Biochimica et Biophysica Acta *71*, 377-391.

Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. Nature structural & molecular biology *17*, 666-672.

Diezel, W., Kopperschläger, G., and Hofmann, E. (1972). An improved procedure for protein staining in polyacrylamide gels with a new type of Coomassie Brilliant Blue. Analytical biochemistry *48*, 617-620.

Dikic, I. (2002). CIN85/CMS family of adaptor molecules. FEBS letters *529*, 110-115.

Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics (Oxford, England) *21*, 3433-3434.

Doyle, M.L. (1997). Characterization of binding interactions by isothermal titration calorimetry. Current opinion in biotechnology *8*, 31-35.

Dustin, M.L., Olszowy, M.W., Holdorf, A.D., Li, J., Bromley, S., Desai, N., Widder, P., Rosenberger, F., van der Merwe, P.A., Allen, P.M.*, et al.* (1998). A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. Cell *94*, 667-677.

Echols, N., Grosse-Kunstleve, R.W., Afonine, P.V., Bunkoczi, G., Chen, V.B., Headd, J.J., McCoy, A.J., Moriarty, N.W., Read, R.J., Richardson,

D.C., *et al.* (2012). Graphical tools for macromolecular crystallography in PHENIX. Journal of applied crystallography *45*, 581-586.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research *30*, 207-210.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography *60*, 2126-2132.

Englard, S., and Seifter, S. (1990a). Precipitation Techniques. Methods in Enzymology *182*, 285-300.

Enkhbayar, P., Hikichi, K., Osaki, M., Kretsinger, R.H., and Matsushima, N. (2006). 3(10)-helices in proteins are parahelices. Proteins *64*, 691-699.

Evans, P. (2006). Scaling and assessment of data quality. Acta crystallographica. Section D, Biological crystallography *62*, 72-82.

Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta crystallographica. Section D, Biological crystallography *67*, 282-292.

Feller, S.M., Knudsen, B., and Hanafusa, H. (1994a). c-Abl kinase regulates the protein binding activity of c-Crk. The EMBO journal *13*, 2341-2351.

Feller, S.M., Ren, R., Hanafusa, H., and Baltimore, D. (1994b). SH2 and SH3 domains as molecular adhesives: the interactions of Crk and Abl. Trends in biochemical sciences *19*, 453-458.

Fernandez-Ballester, G., Blanes-Mira, C., and Serrano, L. (2004). The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains. Journal of molecular biology *335*, 619-629.

Ferreon, J.C., and Hilser, V.J. (2004). Thermodynamics of binding to SH3 domains: the energetic impact of polyproline II (PII) helix formation. Biochemistry *43*, 7787-7797.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fischer, E.H., and Krebs, E.G. (1955). Conversion of phosphorylase b to phosphorylase a in muscle extracts. The Journal of biological chemistry *216*, 121-132.

Frank, R. (1992). Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. Tetrahedron *48*, 9217-9232.

Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gustafsdottir, S.M., Ostman, A., and Landegren, U. (2002). Protein detection using proximity-dependent DNA ligation assays. Nature biotechnology *20*, 473-477.

Freund, C., Kuhne, R., Yang, H., Park, S., Reinherz, E.L., and Wagner, G. (2002). Dynamic interaction of CD2 with the GYF and the SH3 domain of compartmentalized effector molecules. The EMBO journal *21*, 5985-5995.

Freund, C., Schmalz, H.G., Sticht, J., and Kuhne, R. (2008). Proline-rich sequence recognition domains (PRD): ligands, function and inhibition. Handbook of experimental pharmacology, 407-429.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., and Stratton, M.R. (2004). A census of human cancer genes. Nature Reviews Cancer *4*, 177-183.

Gaidos, G., Soni, S., Oswald, D.J., Toselli, P.A., and Kirsch, K.H. (2007). Structure and function analysis of the CMS/CIN85 protein family identifies actin-bundling properties and heterotypic-complex formation. Journal of cell science *120*, 2366-2377.

Gallo, K.A., and Johnson, G.L. (2002). Mixed-lineage kinase control of JNK and p38 MAPK pathways. Nature reviews. Molecular cell biology *3*, 663-672.

Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Côté, M., and Rich, R.L. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell *107*, 55-65.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook (Springer), pp. 571-607.

Gauthier, N.C., Monzo, P., Gonzalez, T., Doye, A., Oldani, A., Gounon, P., Ricci, V., Cormont, M., and Boquet, P. (2007). Early endosomes associated with dynamic F-actin structures are required for late trafficking of H. pylori VacA toxin. The Journal of cell biology *177*, 343-354.

Gershenson, A., and Gierasch, L.M. (2011). Protein folding in the cell: challenges and progress. Current opinion in structural biology *21*, 32-41.

Ghai, R., Falconer, R.J., and Collins, B.M. (2012). Applications of isothermal titration calorimetry in pure and applied research-survey of the literature from 2010. Journal of molecular recognition *25*, 32-52.

Gigante, M., Pontrelli, P., Montemurno, E., Roca, L., Aucella, F., Penza, R., Caridi, G., Ranieri, E., Ghiggeri, G.M., and Gesualdo, L. (2009). CD2AP mutations are associated with sporadic nephrotic syndrome and focal segmental glomerulosclerosis (FSGS). Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association *24*, 1858-1864.

Gilfillan, A.M., and Beaven, M.A. (2011). Regulation of mast cell responses in health and disease. Critical reviews in immunology *31*, 475-529.

Goh, L.K., and Sorkin, A. (2013). Endocytosis of receptor tyrosine kinases. Cold Spring Harbor perspectives in biology *5*, a017459.

Good, M.C., Zalatan, J.G., and Lim, W.A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. Science (New York, N.Y.) *332*, 680-686.

Gout, I., Middleton, G., Adu, J., Ninkina, N.N., Drobot, L.B., Filonenko, V., Matsuka, G., Davies, A.M., Waterfield, M., and Buchman, V.L. (2000). Negative regulation of PI 3-kinase by Ruk, a novel adaptor protein. The EMBO journal *19*, 4015-4025.

Grahammer, F., Schell, C., and Huber, T.B. (2013). The podocyte slit diaphragm-from a thin grey line to a complex signalling hub. Nature reviews. Nephrology *9*, 587-598.

Grant, B.D., and Donaldson, J.G. (2009). Pathways and mechanisms of endocytic recycling. Nature reviews. Molecular cell biology *10*, 597-608.

Green, P.J., Pines, O., and Inouye, M. (1986). The role of antisense RNA in gene regulation. Annual review of biochemistry *55*, 569-597.

Groemping, Y., Lapouge, K., Smerdon, S.J., and Rittinger, K. (2003). Molecular basis of phosphorylation-induced activation of the NADPH oxidase. Cell *113*, 343-355.

Grunkemeyer, J.A., Kwoh, C., Huber, T.B., and Shaw, A.S. (2005). CD2-associated protein (CD2AP) expression in podocytes rescues lethality of CD2AP deficiency. The Journal of biological chemistry *280*, 29677-29681.

Haglund, K., Shimokawa, N., Szymkiewicz, I., and Dikic, I. (2002). Cbl-directed monoubiquitination of CIN85 is involved in regulation of ligand-induced degradation of EGF receptors. Proceedings of the National Academy of Sciences of the United States of America *99*, 12191-12196.

Hahn, W.C., and Bierer, B.E. (1993). Separable portions of the CD2 cytoplasmic domain involved in signaling and ligand avidity regulation. The Journal of experimental medicine *178*, 1831-1836.

Haikarainen, T., Chen, W.Q., Lubec, G., and Kursula, P. (2009). Structure, modifications and ligand-binding properties of rat profilin 2a. Acta crystallographica. Section D, Biological crystallography *65*, 303-311.

Hakes, D.J., and Dixon, J.E. (1992). New vectors for high level expression of recombinant proteins in bacteria. Analytical biochemistry *202*, 293-298.

Hammerling, U., Bjelfman, C., and Pahlman, S. (1987). Different regulation of N- and c-myc expression during phorbol ester-induced maturation of human SH-SY5Y neuroblastoma cells. Oncogene *2*, 73-77.

Handley, M.E., Rasaiyaah, J., Chain, B.M., and Katz, D.R. (2007). Mixed lineage kinases (MLKs): a role in dendritic cells, inflammation and immunity? International journal of experimental pathology *88*, 111-126.

Harkiolaki, M., Gilbert, R.J., Jones, E.Y., and Feller, S.M. (2006). The C-terminal SH3 domain of CRKL as a dynamic dimerization module transiently exposing a nuclear export signal. Structure (London, England : 1993) *14*, 1741-1753.

Harkiolaki, M., Lewitzky, M., Gilbert, R.J., Jones, E.Y., Bourette, R.P., Mouchiroud, G., Sondermann, H., Moarefi, I., and Feller, S.M. (2003). Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76. The EMBO journal *22*, 2571-2582.

Harkiolaki, M., Tsirka, T., Lewitzky, M., Simister, P.C., Joshi, D., Bird, L.E., Jones, E.Y., O'Reilly, N., and Feller, S.M. (2009). Distinct binding modes of two epitopes in Gab2 that interact with the SH3C domain of Grb2. Structure (London, England : 1993) *17*, 809-822.

Hasegawa, J., Tsujita, K., Takenawa, T., and Itoh, T. (2012). ARAP1 regulates the ring size of circular dorsal ruffles through Arf1 and Arf5. Molecular biology of the cell *23*, 2481-2489.

Havrylov, S., Redowicz, M.J., and Buchman, V.L. (2010). Emerging roles of Ruk/CIN85 in vesicle-mediated transport, adhesion, migration and malignancy. Traffic *11*, 721-731.

Havrylov, S., Rzhepetskyy, Y., Malinowska, A., Drobot, L., and Redowicz, M.J. (2009). Proteins recruited by SH3 domains of Ruk/CIN85 adaptor identified by LC-MS/MS. Proteome science *7*, 21.

Hecht, D., and Zick, Y. (1992). Selective inhibition of protein tyrosine phosphatase activities by H_2O_2 and vanadate *in vitro*. Biochemical and Biophysical Research Communications *188*, 773-779.

Heffetz, D., Bushkin, I., Dror, R., and Zick, Y. (1990). The Insulinomimetic Agents H_2O_2 and Vanadate Stimulate Protein Tyrosine Phosphorylation in Intact Cells. Journal of Biological Chemistry *265*, 2896-2902.

Heuer, K., Arbuzova, A., Strauss, H., Kofler, M., and Freund, C. (2005). The helically extended SH3 domain of the T cell adaptor protein ADAP is a novel lipid interaction domain. Journal of molecular biology *348*, 1025-1035.

Heuer, K., Kofler, M., Langdon, G., Thiemke, K., and Freund, C. (2004). Structure of a helically extended SH3 domain of the T cell adapter protein ADAP. Structure (London, England : 1993) *12*, 603-610.

Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.-C., Carrasquillo, M.M., Abraham, R., Hamshere, M.L., Pahwa, J.S., and Moskvina, V. (2011). Common variants at ABCA7, MS4A6A/MS4A4E,

EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nature genetics *43*, 429-435.

Hou, T., Li, N., Li, Y., and Wang, W. (2012). Characterization of domain-peptide interaction interface: prediction of SH3 domain-mediated protein-protein interaction network in yeast by generic structure-based models. Journal of proteome research *11*, 2982-2995.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M.J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. Nature structural biology *7*, 634-638.

Huber, T.B., Hartleben, B., Kim, J., Schmidts, M., Schermer, B., Keil, A., Egger, L., Lecha, R.L., Borner, C., Pavenstadt, H.*, et al.* (2003). Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling. Molecular and cellular biology *23*, 4917-4928.

Huber, T.B., Kwoh, C., Wu, H., Asanuma, K., Godel, M., Hartleben, B., Blumer, K.J., Miner, J.H., Mundel, P., and Shaw, A.S. (2006). Bigenic mouse models of focal segmental glomerulosclerosis involving pairwise interaction of CD2AP, Fyn, and synaptopodin. The Journal of clinical investigation *116*, 1337-1345.

Hutchings, N.J., Clarkson, N., Chalkley, R., Barclay, A.N., and Brown, M.H. (2003). Linking the T cell surface protein CD2 to the actin-capping protein CAPZ via CMS and CIN85. The Journal of biological chemistry *278*, 22396-22403.

Hyvonen, M.E., Saurus, P., Wasik, A., Heikkila, E., Havana, M., Trokovic, R., Saleem, M., Holthofer, H., and Lehtonen, S. (2010). Lipid phosphatase

SHIP2 downregulates insulin signalling in podocytes. Molecular and cellular endocrinology *328*, 70-79.

Inoue, H., and Randazzo, P.A. (2007). Arf GAPs and their interacting proteins. Traffic *8*, 1465-1475.

Janson, C., Kasahara, N., Prendergast, G.C., and Colicelli, J. (2012). RIN3 is a negative regulator of mast cell responses to SCF. PloS one *7*, e49615.

Joosten, R.P., Te Beek, T.A., Krieger, E., Hekkelman, M.L., Hooft, R.W., Schneider, R., Sander, C., and Vriend, G. (2011). A series of PDB related databases for everyday needs. Nucleic acids research *39*, D411-D419.

Jozic, D., Cárdenes, N., Deribe, Y.L., Moncalián, G., Hoeller, D., Groemping, Y., Dikic, I., Rittinger, K., and Bravo, J. (2005). Cbl promotes clustering of endocytic adaptor proteins. Nature structural & molecular biology *12*, 972-979.

Ju, W., Greene, C.S., Eichinger, F., Nair, V., Hodgin, J.B., Bitzer, M., Lee, Y.S., Zhu, Q., Kehata, M., Li, M.*, et al.* (2013). Defining cell-type specificity at the transcriptional level in human disease. Genome research *23*, 1862-1873.

Kabsch, W. (2010). XDS. Acta crystallographica. Section D, Biological crystallography *66*, 125-132.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577-2637.

Kajiho, H., Saito, K., Tsujita, K., Kontani, K., Araki, Y., Kurosu, H., and Katada, T. (2003). RIN3: a novel Rab5 GEF interacting with amphiphysin II involved in the early endocytic pathway. Journal of cell science *116*, 4159-4168.

Kajiho, H., Sakurai, K., Minoda, T., Yoshikawa, M., Nakagawa, S., Fukushima, S., Kontani, K., and Katada, T. (2011). Characterization of RIN3 as a guanine nucleotide exchange factor for the Rab5 subfamily GTPase Rab31. The Journal of biological chemistry *286*, 24364-24373.

Kaneko, T., Sidhu, S.S., and Li, S.S. (2011). Evolving specificity from variability for protein interaction domains. Trends in biochemical sciences *36*, 183-190.

Kang, F., Purich, D.L., and Southwick, F.S. (1999). Profilin promotes barbed-end actin filament assembly without lowering the critical concentration. Journal of Biological Chemistry *274*, 36963-36972.

Kato, R., Nonami, A., Taketomi, T., Wakioka, T., Kuroiwa, A., Matsuda, Y., and Yoshimura, A. (2003). Molecular cloning of mammalian Spred-3 which suppresses tyrosine kinase-mediated Erk activation. Biochem Biophys Res Commun *302*, 767-772.

Katz, C., Levy-Beladev, L., Rotem-Bamberger, S., Rito, T., Rudiger, S.G., and Friedler, A. (2011). Studying protein-protein interactions using peptide arrays. Chemical Society reviews *40*, 2131-2145.

Kay, B.K. (2012). SH3 domains come of age. FEBS letters *586*, 2606-2608.

Ketting, R.F. (2011). The many faces of RNAi. Developmental cell *20*, 148-161.

Kharfan-Dabaja, M.A., Lazarus, H.M., Nishihori, T., Mahfouz, R.A., and Hamadani, M. (2013). Diagnostic and therapeutic advances in blastic plasmacytoid dendritic cell neoplasm: a focus on hematopoietic cell transplantation. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation *19*, 1006-1012.

Kim, J.M., Wu, H., Green, G., Winkler, C.A., Kopp, J.B., Miner, J.H., Unanue, E.R., and Shaw, A.S. (2003). CD2-associated protein haploinsufficiency is linked to glomerular disease susceptibility. Science (New York, N.Y.) *300*, 1298-1300.

Kirsch, K.H., Georgescu, M.M., Ishimaru, S., and Hanafusa, H. (1999). CMS: an adapter molecule involved in cytoskeletal rearrangements. Proceedings of the National Academy of Sciences of the United States of America *96*, 6211-6216.

Kirsch, K.H., Georgescu, M.M., Shishido, T., Langdon, W.Y., Birge, R.B., and Hanafusa, H. (2001). The adapter type protein CMS/CD2AP binds to the proto-oncogenic protein c-Cbl through a tyrosine phosphorylation-regulated Src homology 3 domain interaction. The Journal of biological chemistry *276*, 4957-4963.

Kobayashi, S., Sawano, A., Nojima, Y., Shibuya, M., and Maru, Y. (2004). The c-Cbl/CD2AP complex regulates VEGF-induced endocytosis and degradation of Flt-1 (VEGFR-1). FASEB journal : official publication of the Federation of American Societies for Experimental Biology *18*, 929-931.

Kofler, M.M., and Freund, C. (2006). The GYF domain. The FEBS journal *273*, 245-256.

Komuro, A., Saeki, M., and Kato, S. (1999). Association of two nuclear proteins, Npw38 and NpwBP, via the interaction between the WW domain and a novel proline-rich motif containing glycine and arginine. The Journal of biological chemistry *274*, 36513-36519.

Konishi, H., Tashiro, K., Murata, Y., Nabeshi, H., Yamauchi, E., and Taniguchi, H. (2006). CFBP is a novel tyrosine-phosphorylated protein that might function as a regulator of CIN85/CD2AP. The Journal of biological chemistry *281*, 28919-28931.

Kowanetz, K., Szymkiewicz, I., Haglund, K., Kowanetz, M., Husnjak, K., Taylor, J.D., Soubeyran, P., Engstrom, U., Ladbury, J.E., and Dikic, I. (2003). Identification of a novel proline-arginine motif involved in CIN85-dependent clustering of Cbl and down-regulation of epidermal growth factor receptors. The Journal of biological chemistry *278*, 39735-39746.

Krieger, J.M., Fusco, G., Lewitzky, M., Simister, P.C., Marchant, J., Camilloni, C., Feller, S.M., and De Simone, A. (2014). Conformational recognition of an intrinsically disordered protein. Biophysical journal *106*, 1771-1779.

Kurakin, A.V., Wu, S., and Bredesen, D.E. (2003). Atypical recognition consensus of CIN85/SETA/Ruk SH3 domains revealed by target-assisted iterative screening. Journal of Biological Chemistry *278*, 34102-34109.

Ladbury, J.E. (2010). Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. Biochemical Society transactions *38*, 888-893.

Ladbury, J.E., and Arold, S.T. (2011). Energetics of Src homology domain interactions in receptor tyrosine kinase-mediated signaling. Methods Enzymol *488*, 147-183.

Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B.*, et al.* (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature genetics *45*, 1452-1458.

Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R., and Cesareni, G. (2004). Protein interaction networks by proteome peptide scanning. PLoS biology *2*, E14.

Layne, E. (1957). Spectrophotometric and turbidimetric methods for measuring proteins. Methods in Enzymology *3*, 447-454.

Lee, J.H., and Paull, T.T. (2005). ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. Science (New York, N.Y.) *308*, 551-554.

Lehtonen, S., Ora, A., Olkkonen, V.M., Geng, L., Zerial, M., Somlo, S., and Lehtonen, E. (2000). In vivo interaction of the adapter protein CD2-associated protein with the type 2 polycystic kidney disease protein, polycystin-2. The Journal of biological chemistry *275*, 32888-32893.

Lemmon, M.A., and Schlessinger, J. (2010). Cell Signaling by Receptor Tyrosine Kinases. Cell *141*, 1117-1134.

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. Nucleic acids research *40*, D302-305.

Lewitzky, M., Harkiolaki, M., Domart, M.C., Jones, E.Y., and Feller, S.M. (2004). Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. The Journal of biological chemistry *279*, 28724-28732.

Li, S.S. (2005). Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. The Biochemical journal *390*, 641-653.

Liu, B.A., Engelmann, B.W., and Nash, P.D. (2012). High-throughput analysis of peptide-binding modules. Proteomics *12*, 1527-1546.

London, N., Movshovitz-Attias, D., and Schueler-Furman, O. (2010). The structural basis of peptide-protein binding strategies. Structure (London, England : 1993) *18*, 188-199.

Lowik, M.M., Groenen, P.J., Pronk, I., Lilien, M.R., Goldschmeding, R., Dijkman, H.B., Levtchenko, E.N., Monnens, L.A., and van den Heuvel, L.P. (2007). Focal segmental glomerulosclerosis in a patient homozygous for a CD2AP mutation. Kidney international *72*, 1198-1203.

Lynch, D.K., Winata, S.C., Lyons, R.J., Hughes, W.E., Lehrbach, G.M., Wasinger, V., Corthals, G., Cordwell, S., and Daly, R.J. (2003). A

Cortactin-CD2-associated protein (CD2AP) complex provides a novel link between epidermal growth factor receptor endocytosis and the actin cytoskeleton. The Journal of biological chemistry *278*, 21805-21813.

Macias, M.J., Gervais, V., Civera, C., and Oschkinat, H. (2000). Structural analysis of WW domains and design of a WW prototype. Nature structural biology *7*, 375-379.

Macias, M.J., Wiesner, S., and Sudol, M. (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. FEBS letters *513*, 30-37.

Mahoney, N.M., Rozwarski, D.A., Fedorov, E., Fedorov, A.A., and Almo, S.C. (1999). Profilin binds proline-rich ligands in two distinct amide backbone orientations. Nature structural biology *6*, 666-671.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science (New York, N.Y.) *298*, 1912-1934.

Mao, J., Zhang, Y., Du, L., Dai, Y., Yang, C., and Liang, L. (2006). Expression profile of nephrin, podocin, and CD2AP in Chinese children with MCNS and IgA nephropathy. Pediatric nephrology (Berlin, Germany) *21*, 1666-1675.

Marafioti, T., Paterson, J.C., Ballabio, E., Reichard, K.K., Tedoldi, S., Hollowood, K., Dictor, M., Hansmann, M.L., Pileri, S.A., Dyer, M.J.*, et al.* (2008). Novel markers of normal and neoplastic human plasmacytoid dendritic cells. Blood *111*, 3778-3792.

Matalon, O., Reicher, B., and Barda-Saad, M. (2013). Wiskott-Aldrich syndrome protein-dynamic regulation of actin homeostasis: from activation through function and signal termination in T lymphocytes. Immunological reviews *256*, 10-29.

Mayer, B.J., Hamaguchi, M., and Hanafusa, H. (1988). A novel viral oncogene with structural similarity to phospholipase C. Nature *332*, 272-275.

Mayer, B.J., and Saksela, K. (2005). SH3 domains. Protein Science Encyclopedia.

McCaffrey, M.W., Bielli, A., Cantalupo, G., Mora, S., Roberti, V., Santillo, M., Drummond, F., and Bucci, C. (2001). Rab4 affects both recycling and degradative endosomal trafficking. FEBS letters *495*, 21-30.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta crystallographica. Section D, Biological crystallography *61*, 458-464.

McGuffee, S.R., and Elcock, A.H. (2010). Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS computational biology *6*, e1000694.

Medway, C., and Morgan, K. (2013). CD2-Associated Protein (CD2AP). In Genetic Variants in Alzheimer's Disease (Springer), pp. 201-208.

Merril, C.R., and Pratt, M.E. (1986). A silver stain for the rapid quantitative detection of proteins or nucleic acids on membranes or thin layer plates. Analytical biochemistry *156*, 96-110.

Michelsen, U., and von Hagen, J. (2009). Isolation of subcellular organelles and structures. Methods Enzymol *463*, 305-328.

Minegishi, Y., Shibagaki, Y., Mizutani, A., Fujita, K., Tezuka, T., Kinoshita, M., Kuroda, M., Hattori, S., and Gotoh, N. (2013). Adaptor protein complex of FRS2beta and CIN85/CD2AP provides a novel mechanism for ErbB2/HER2 protein downregulation. Cancer science *104*, 345-352.

Miura, K., Jacques, K.M., Stauffer, S., Kubosaki, A., Zhu, K., Hirsch, D.S., Resau, J., Zheng, Y., and Randazzo, P.A. (2002). ARAP1: a point of convergence for Arf and Rho signaling. Molecular cell *9*, 109-119.

Moncalian, G., Cardenes, N., Deribe, Y.L., Spinola-Amilibia, M., Dikic, I., and Bravo, J. (2006). Atypical polyproline recognition by the CMS N-terminal Src homology 3 domain. The Journal of biological chemistry *281*, 38845-38853.

Monzo, P., Gauthier, N.C., Keslair, F., Loubat, A., Field, C.M., Le Marchand-Brustel, Y., and Cormont, M. (2005). Clues to CD2-associated protein involvement in cytokinesis. Molecular biology of the cell *16*, 2891-2902.

Morita, E., Sandrin, V., Chung, H.Y., Morham, S.G., Gygi, S.P., Rodesch, C.K., and Sundquist, W.I. (2007). Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. The EMBO journal *26*, 4215-4227.

Muller, P.A.J. and Vousden, K.H. (2013). p53 mutations in cancer. Nature cell biology *15*, 2-8

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica. Section D, Biological crystallography *53*, 240-255.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. Nature *359*, 851-855.

Myszka, D.G., Abdiche, Y.N., Arisaka, F., Byron, O., Eisenstein, E., Hensley, P., Thomson, J.A., Lombardo, C.R., Schwarz, F., Stafford, W.*, et al.* (2003). The ABRF-MIRG'02 study: assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction. Journal of biomolecular techniques : JBT *14*, 247-269.

Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., and Kuriyan, J. (2003). Structural basis for the autoinhibition of c-Abl tyrosine kinase. Cell *112*, 859-871.

Naj, A.C., Jun, G., Beecham, G.W., Wang, L.S., Vardarajan, B.N., Buros, J., Gallins, P.J., Buxbaum, J.D., Jarvik, G.P., Crane, P.K.*, et al.* (2011). Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nature genetics *43*, 436-441.

Ng, E.L., Wang, Y., and Tang, B.L. (2007). Rab22B's role in trans-Golgi network membrane dynamics. Biochem Biophys Res Commun *361*, 751-757.

Ni, L., Saleem, M., and Mathieson, P.W. (2012). Podocyte culture: tricks of the trade. Nephrology *17*, 525-531.

Noble, J.E., and Bailey, M.J.A. (2009). Quantitation of Protein. Methods in Enzymology *463*, 73-95.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic acids research *31*, 3635-3641.

Odorizzi, G. (2006). The multiple personalities of Alix. Journal of cell science *119*, 3025-3032.

Ogura, K., Nobuhisa, I., Yuzawa, S., Takeya, R., Torikai, S., Saikawa, K., Sumimoto, H., and Inagaki, F. (2006). NMR solution structure of the tandem Src homology 3 domains of p47phox complexed with a p22phox-derived proline-rich peptide. The Journal of biological chemistry *281*, 3660-3668.

Okada, H., Uezu, A., Mason, F.M., Soderblom, E.J., Moseley, M.A., 3rd, and Soderling, S.H. (2011). SH3 domain-based phototrapping in living cells reveals Rho family GAP signaling complexes. Science signaling *4*, rs13.

Oliveros, J.C. (2007). VENNY. An interactive tool for comparing lists with Venn Diagrams.

Ortega-Roldan, J.L., Jensen, M.R., Brutscher, B., Azuaga, A.I., Blackledge, M., and van Nuland, N.A. (2009). Accurate characterization of weak macromolecular interactions by titration of NMR residual dipolar couplings: application to the CD2AP SH3-C:ubiquitin complex. Nucleic acids research *37*, e70.

Ortega Roldan, J.L., Casares, S., Ringkjobing Jensen, M., Cardenes, N., Bravo, J., Blackledge, M., Azuaga, A.I., and van Nuland, N.A. (2013). Distinct ubiquitin binding modes exhibited by SH3 domains: molecular determinants and functional implications. PloS one *8*, e73018.

Ortega Roldan, J.L., Romero Romero, M.L., Ora, A., Ab, E., Lopez Mayorga, O., Azuaga, A.I., and van Nuland, N.A. (2007). The high resolution NMR structure of the third SH3 domain of CD2AP. Journal of biomolecular NMR *39*, 331-336.

Palencia, A., Cobos, E.S., Mateo, P.L., Martinez, J.C., and Luque, I. (2004). Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: implications for rational ligand design. Journal of molecular biology *336*, 527-537.

Palencia, A., Camara-Artigas, A., Pisabarro, M.T., Martinez, J.C., and Luque, I. (2010). Role of interfacial water molecules in proline-rich ligand recognition by the Src homology 3 domain of Abl. The Journal of biological chemistry *285*, 2823-2833.

Pálfy, M., Reményi, A., and Korcsmáros, T. (2012). Endosomal crosstalk: meeting points for signaling pathways. Trends in cell biology *22*, 447-456.

Pauling, L., Corey, R.B., and Branson, H.R. (1951). The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proceedings of the National Academy of Sciences of the United States of America *37*, 205-211.

Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. Science (New York, N.Y.) *300*, 445-452.

Perozzo, R., Folkers, G., and Scapozza, L. (2004). Thermodynamics of Protein–Ligand Interactions: History, Presence, and Future Aspects. Journal of Receptors and Signal Transduction *24*, 1-52.

Peterson, F.C., and Volkman, B.F. (2009). Diversity of polyproline recognition by EVH1 domains. Frontiers in bioscience (Landmark edition) *14*, 833-846.

Philippe, D., Ababou, A., Yang, X., Ghosh, R., Daviter, T., Ladbury, J.E., and Pfuhl, M. (2011). Making ends meet: the importance of the N- and C-termini for the structure, stability, and function of the third SH3 domain of CIN85. Biochemistry *50*, 3649-3659.

Pornillos, O., Alam, S.L., Rich, R.L., Myszka, D.G., Davis, D.R., and Sundquist, W.I. (2002). Structure and functional interactions of the Tsg101 UEV domain. The EMBO journal *21*, 2397-2406.

Powell, H.R., Johnson, O., and Leslie, A.G. (2013). Autoindexing diffraction images with iMosflm. Acta crystallographica. Section D, Biological crystallography *69*, 1195-1203.

Randles, L.G., Dawes, G.J., Wensley, B.G., Steward, A., Nickson, A.A., and Clarke, J. (2013). Understanding pathogenic single-nucleotide polymorphisms in multidomain proteins–studies of isolated domains are not enough. FEBS journal *280*, 1018-1027.

Reebye, V., Frilling, A., Hajitou, A., Nicholls, J.P., Habib, N.A., and Mintz, P.J. (2012). A perspective on non-catalytic Src homology (SH) adaptor signalling proteins. Cellular signalling *24*, 388-392.

Reineke, U., and Sabat, R. (2009). Antibody epitope mapping using SPOT peptide arrays. Methods in molecular biology (Clifton, N.J.) *524*, 145-167.

Ren, R., Mayer, B.J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science (New York, N.Y.) *259*, 1157-1161.

Rhodes, G. (2010). Crystallography made crystal clear: a guide for users of macromolecular models.

Ribatti, D., and Crivellato, E. (2014). Mast cell ontogeny: an historical overview. Immunology letters *159*, 11-14.

Riehemann, K., and Sorg, C. (1993). Sequence homologies between four cytoskeleton-associated proteins. Trends in biochemical sciences *18*, 82-83.

Ritter, A.T., Angus, K.L., and Griffiths, G.M. (2013). The role of the cytoskeleton at the immunological synapse. Immunological reviews *256*, 107-117.

Rizvi, H., Paterson, J.C., Tedoldi, S., Ramsay, A., Calaminici, M., Natkunam, Y., Lonardi, S., Tan, S.Y., Campbell, L., Hansmann, M.L.*, et al.* (2012). Expression of the CD2AP adaptor molecule in normal, reactive and neoplastic human tissue. Pathologica *104*, 56-64.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic acids research *42*, W320-324.

Roldan, J.L., Blackledge, M., van Nuland, N.A., and Azuaga, A.I. (2011). Solution structure, dynamics and thermodynamics of the three SH3 domains of CD2AP. Journal of biomolecular NMR *50*, 103-117.

Rousseau, F., Schymkowitz, J., and Itzhaki, L.S. (2012). Implications of 3D domain swapping for protein folding, misfolding and function. Advances in experimental medicine and biology *747*, 137-152.

Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., *et al.* (2000). Comparative genomics of the eukaryotes. Science (New York, N.Y.) *287*, 2204-2215.

Ruf, H., Georgalis, Y., and Grell, E. (1989). Dynamic laser light scattering to determine size distributions of vesicles. Methods in Enzymology *172*, 364-390.

Rutz, B., and Seraphin, B. (2000). A dual role for BBP/ScSF1 in nuclear pre-mRNA retention and splicing. The EMBO journal *19*, 1873-1886.

Saito, K., Kajiho, H., Araki, Y., Kurosu, H., Kontani, K., Nishina, H., and Katada, T. (2005). Purification and Analysis of RIN Family-Novel Rab5 GEFs. Methods in Enzymology *403*, 276-283.

Saito, K., Kigawa, T., Koshiba, S., Sato, K., Matsuo, Y., Sakamoto, A., Takagi, T., Shirouzu, M., Yabuki, T., Nunokawa, E., *et al.* (2004). The CAP-Gly domain of CYLD associates with the proline-rich sequence in NEMO/IKKgamma. Structure (London, England : 1993) *12*, 1719-1728.

Saito, K., Murai, J., Kajiho, H., Kontani, K., Kurosu, H., and Katada, T. (2002). A novel binding protein composed of homophilic tetramer exhibits unique properties for the small GTPase Rab5. The Journal of biological chemistry *277*, 3412-3418.

Saksela, K., and Permi, P. (2012). SH3 domain ligand binding: What's the consensus and where's the specificity? FEBS letters *586*, 2609-2614.

Saleem, M.A., O'Hare, M.J., Reiser, J., Coward, R.J., Inward, C.D., Farren, T., Xing, C.Y., Ni, L., Mathieson, P.W., and Mundel, P. (2002). A conditionally immortalized human podocyte cell line demonstrating nephrin and podocin expression. Journal of the American Society of Nephrology *13*, 630-638.

Salinovich, O., and Montelaro, R.C. (1986). Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by sodium dodecylsulfate-polyacrylamide gel electrophoresis. Analytical biochemistry *156*, 341-347.

Santana-de Anda, K., Gomez-Martin, D., Soto-Solis, R., and Alcocer-Varela, J. (2013). Plasmacytoid dendritic cells: key players in viral infections and autoimmune diseases. Seminars in arthritis and rheumatism *43*, 131-136.

Sato, S., Zhao, Y., Imai, M., Simister, P.C., Feller, S.M., Trackman, P.C., Kirsch, K.H., and Sonenshein, G.E. (2013). Inhibition of CIN85-mediated invasion by a novel SH3 domain binding motif in the lysyl oxidase propeptide. PloS one *8*, e77288.

Sato, Y., Taoka, M., Sugiyama, N., Kubo, K., Fuchigami, T., Asada, A., Saito, T., Nakajima, K., Isobe, T., and Hisanaga, S. (2007). Regulation of the interaction of Disabled-1 with CIN85 by phosphorylation with Cyclin-dependent kinase 5. Genes to cells : devoted to molecular & cellular mechanisms *12*, 1315-1327.

Schell, C., and Huber, T.B. (2012). New players in the pathogenesis of focal segmental glomerulosclerosis. Nephrology Dialysis Transplantation, gfs273.

Schiffer, M., Mundel, P., Shaw, A.S., and Bottinger, E.P. (2004). A novel role for the adaptor molecule CD2-associated protein in transforming growth factor-beta-induced apoptosis. The Journal of biological chemistry *279*, 37004-37012.

Schmidt, M.H., Dikic, I., and Bogler, O. (2005). Src phosphorylation of Alix/AIP1 modulates its interaction with binding partners and antagonizes its activities. The Journal of biological chemistry *280*, 3414-3425.

Schmidt, M.H., Hoeller, D., Yu, J., Furnari, F.B., Cavenee, W.K., Dikic, I., and Bogler, O. (2004). Alix/AIP1 antagonizes epidermal growth factor receptor downregulation by the Cbl-SETA/CIN85 complex. Molecular and cellular biology *24*, 8981-8993.

Schwarz, K., Simons, M., Reiser, J., Saleem, M.A., Faul, C., Kriz, W., Shaw, A.S., Holzman, L.B., and Mundel, P. (2001). Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin. The Journal of clinical investigation *108*, 1621-1629.

Semenza, G.L. (2007). Life with oxygen. Science (New York, N.Y.) *318*, 62-64.

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. Cell *139*, 468-484.

Shih, N.-Y., Li, J., Karpitskii, V., Nguyen, A., Dustin, M.L., Kanagawa, O., Miner, J.H., and Shaw, A.S. (1999). Congenital nephrotic syndrome in mice lacking CD2-associated protein. Science (New York, N.Y.) *286*, 312-315.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology *7*, 539.

Simister, P.C., and Feller, S.M. (2012). Order and disorder in large multi-site docking proteins of the Gab family-implications for signalling complex formation and inhibitor design strategies. Molecular bioSystems *8*, 33-46.

Simister, P.C., Luccarelli, J., Thompson, S., Appella, D.H., Feller, S.M., and Hamilton, A.D. (2013). Novel inhibitors of a Grb2 SH3C domain interaction identified by a virtual screen. Bioorganic & medicinal chemistry *21*, 4027-4033.

Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31-40.

Soderberg, O., Gullberg, M., Jarvius, M., Ridderstrale, K., Leuchowius, K.J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.G.*, et al.* (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nature methods *3*, 995-1000.

Soubeyran, P., Kowanetz, K., Szymkiewicz, I., Langdon, W.Y., and Dikic, I. (2002). Cbl-CIN85-endophilin complex mediates ligand-induced downregulation of EGF receptors. Nature *416*, 183-187.

Sproul, A.A., Xu, Z., Wilhelm, M., Gire, S., and Greene, L.A. (2009). Cbl negatively regulates JNK activation and cell death. Cell research *19*, 950-961.

Srivatsan, S., Swiecki, M., Otero, K., Cella, M., and Shaw, A.S. (2013). CD2-associated protein regulates plasmacytoid dendritic cell migration, but is dispensable for their development and cytokine production. Journal of immunology (Baltimore, Md. : 1950) *191*, 5933-5940.

Stahl, M.L., Ferenz, C.R., Kelleher, K.L., Kriz, R.W., and Knopf, J.L. (1988). Sequence similarity of phospholipase C with the non-catalytic region of src. Nature *332*, 269-272.

Stamenova, S.D., French, M.E., He, Y., Francis, S.A., Kramer, Z.B., and Hicke, L. (2007). Ubiquitin binds to and regulates a subset of SH3 domains. Molecular cell *25*, 273-284.

Stein, A., Pache, R.A., Bernado, P., Pons, M., and Aloy, P. (2009). Dynamic interactions of proteins in complex networks: a more structured view. The FEBS journal *276*, 5390-5405.

Steinmetz, M.O., and Akhmanova, A. (2008). Capturing protein tails by CAP-Gly domains. Trends in biochemical sciences *33*, 535-545.

Stollar, E.J., Lin, H., Davidson, A.R., and Forman-Kay, J.D. (2012). Differential dynamic engagement within 24 SH3 domain: peptide complexes revealed by co-linear chemical shift perturbation analysis. PloS one *7*, e51282.

Suchanek, M., Radzikowska, A., and Thiele, C. (2005). Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nature methods *2*, 261-267.

Sudol, M., Recinos, C.C., Abraczinskas, J., Humbert, J., and Farooq, A. (2005). WW or WoW: the WW domains in a union of bliss. IUBMB life *57*, 773-778.

Swarup, G., Cohen, S., and Garbers, D.L. (1982). Inhibition of membrane phosphotyrosyl-protein phosphatase activity by vanadate. Biochemical and Biophysical Research Communications *107*, 1104-1109.

Tanaka, M., and Shibata, H. (1985). Poly (l-proline)-binding proteins from chick embryos are a profilin and a profilactin. European journal of biochemistry *151*, 291-297.

Tang, V.W., and Brieher, W.M. (2013). FSGS3/CD2AP is a barbed-end capping protein that stabilizes actin and strengthens adherens junctions. The Journal of cell biology *203*, 815-833.

Theobald, D.L., and Wuttke, D.S. (2005). Divergent evolution within protein superfolds inferred from profile-based phylogenetics. Journal of molecular biology *354*, 722-737.

Tienari, J., Lehtonen, S., and Lehtonen, E. (2005). CD2-associated protein in human urogenital system and in adult kidney tumours. Virchows Archiv : an international journal of pathology *446*, 394-401.

Tonikian, R., Xin, X., Toret, C.P., Gfeller, D., Landgraf, C., Panni, S., Paoluzi, S., Castagnoli, L., Currell, B., Seshagiri, S*., et al.* (2009). Bayesian modeling of the yeast SH3 domain interactome predicts spatiotemporal dynamics of endocytosis proteins. PLoS biology *7*, e1000218.

Tonks, N.K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. Nature Reviews Molecular Cell Biology *7*, 833-846.

Tossidou, I., Kardinal, C., Peters, I., Kriz, W., Shaw, A., Dikic, I., Tkachuk, S., Dumler, I., Haller, H., and Schiffer, M. (2007). CD2AP/CIN85 balance determines receptor tyrosine kinase signaling response in podocytes. The Journal of biological chemistry *282*, 7457-7464.

Tossidou, I., Teng, B., Drobot, L., Meyer-Schwesinger, C., Worthmann, K., Haller, H., and Schiffer, M. (2010). CIN85/RukL is a novel binding partner of nephrin and podocin and mediates slit diaphragm turnover in podocytes. The Journal of biological chemistry *285*, 25285-25295.

Tossidou, I., Niedenthal, R., Klaus, M., Teng, B., Worthmann, K., King, B.L., Peterson, K.J., Haller, H., and Schiffer, M. (2012). CD2AP regulates SUMOylation of CIN85 in podocytes. Molecular and cellular biology *32*, 1068-1079.

Tsui, C.C., and Pierchala, B.A. (2008). CD2AP and Cbl-3/Cbl-c constitute a critical checkpoint in the regulation of ret signal transduction. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 8789-8800.

Tu, J.C., Xiao, B., Yuan, J.P., Lanahan, A.A., Leoffert, K., Li, M., Linden, D.J., and Worley, P.F. (1998). Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. Neuron *21*, 717-726.

Usami, Y., Popov, S., and Gottlinger, H.G. (2007). Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. Journal of virology *81*, 6614-6622.

van Duijn, T.J., Anthony, E.C., Hensbergen, P.J., Deelder, A.M., and Hordijk, P.L. (2010). Rac1 recruits the adapter protein CMS/CD2AP to cell-cell contacts. The Journal of biological chemistry *285*, 20137-20146.

Velho, S., Oliveira, C., Paredes, J., Sousa, S., Leite, M., Matos, P., Milanezi, F., Ribeiro, A.S., Mendes, N., Licastro, D.*, et al.* (2010). Mixed lineage

kinase 3 gene mutations in mismatch repair deficient gastrointestinal tumours. Human molecular genetics *19*, 697-706.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. Nature structural biology *7*, 639-643.

VerPlank, L., Bouamr, F., LaGrassa, T.J., Agresta, B., Kikonyogo, A., Leis, J., and Carter, C.A. (2001). Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55Gag. Proceedings of the National Academy of Sciences *98*, 7724-7729.

Vidal, M., Gigoux, V., and Garbay, C. (2001). SH2 and SH3 domains as targets for anti-proliferative agents. Critical reviews in oncology/hematology *40*, 175-186.

Vieira-Pires, R.S., and Morais-Cabral, J.H. (2010). 310 helices in channels and other membrane proteins. The Journal of General Physiology *136*, 585-592.

Volkman, B.F., Prehoda, K.E., Scott, J.A., Peterson, F.C., and Lim, W.A. (2002). Structure of the N-WASP EVH1 domain-WIP complex: insight into the molecular basis of Wiskott-Aldrich Syndrome. Cell *111*, 565-576.

Volkmer, R., Tapia, V., and Landgraf, C. (2012). Synthetic peptide arrays for investigating protein interaction domains. FEBS letters *586*, 2780-2786.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields,

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S. and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature *417*, 399-403.

Wang, L., Liang, Z., and Li, G. (2011). Rab22 controls NGF signaling and neurite outgrowth in PC12 cells. Molecular biology of the cell *22*, 3853-3860.

Wasik, A.A., Polianskyte-Prause, Z., Dong, M.Q., Shaw, A.S., Yates, J.R., 3rd, Farquhar, M.G., and Lehtonen, S. (2012). Septin 7 forms a complex with CD2AP and nephrin and regulates glucose transporter trafficking. Molecular biology of the cell *23*, 3370-3379.

Weavers, H., Prieto-Sanchez, S., Grawe, F., Garcia-Lopez, A., Artero, R., Wilsch-Brauninger, M., Ruiz-Gomez, M., Skaer, H., and Denholm, B. (2009). The insect nephrocyte is a podocyte-like cell with a filtration slit diaphragm. Nature *457*, 322-326.

Welsch, T., Endlich, N., Gokce, G., Doroshenko, E., Simpson, J.C., Kriz, W., Shaw, A.S., and Endlich, K. (2005). Association of CD2AP with dynamic actin on vesicles in podocytes. American journal of physiology. Renal physiology *289*, F1134-1143.

Welsh, G.I., Hale, L.J., Eremina, V., Jeansson, M., Maezawa, Y., Lennon, R., Pons, D.A., Owen, R.J., Satchell, S.C., Miles, M.J.*, et al.* (2010). Insulin signaling to the glomerular podocyte is critical for normal kidney function. Cell metabolism *12*, 329-340.

Welsh, G.I., and Saleem, M.A. (2012). The podocyte cytoskeleton-key to a functioning glomerulus in health and disease. Nature reviews. Nephorology *8*, 14-21.

Whisstock, J.C., and Lesk, A.M. (1999). SH3 domains in prokaryotes. Trends in biochemical sciences *24*, 132-133.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Analytical biochemistry *179*, 131-137.

Wrana, J.L. (2013). Signaling by the TGFβ Superfamily. Cold Spring Harbor perspectives in biology *5*, a011197.

Wu, C., Ma, M.H., Brown, K.R., Geisler, M., Li, L., Tzeng, E., Jia, C.Y., Jurisica, I., and Li, S.S. (2007). Systematic identification of SH3 domain-mediated human protein-protein interactions by peptide array target screening. Proteomics *7*, 1775-1785.

Wu, Z., Doondeea, J.B., Gholami, A.M., Janning, M.C., Lemeer, S., Kramer, K., Eccles, S.A., Gollin, S.M., Grenman, R., Walch, A.*, et al.* (2011). Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer. Molecular & cellular proteomics : MCP *10*, M111 011635.

Xavier, S., Niranjan, T., Krick, S., Zhang, T., Ju, W., Shaw, A.S., Schiffer, M., and Bottinger, E.P. (2009). TbetaRI independently activates Smad- and CD2AP-dependent pathways in podocytes. Journal of the American Society of Nephrology : JASN *20*, 2127-2137.

Xia, W., Mruk, D.D., Lee, W.M., and Cheng, C.Y. (2006). Differential interactions between transforming growth factor-beta3/TbetaR1, TAB1, and CD2AP disrupt blood-testis barrier and Sertoli-germ cell adhesion. The Journal of biological chemistry *281*, 16799-16813.

Yaddanapudi, S., Altintas, M.M., Kistler, A.D., Fernandez, I., Moller, C.C., Wei, C., Peev, V., Flesche, J.B., Forst, A.L., Li, J.*, et al.* (2011). CD2AP in mouse and human podocytes controls a proteolytic program that

regulates cytoskeletal structure and cellular survival. The Journal of clinical investigation *121*, 3965-3980.

Yang, Z., Xu, Y., Luo, H., Ma, X., Wang, Q., Wang, Y., Deng, W., Jiang, T., Sun, G., He, T.*, et al.* (2014). Whole-exome sequencing for the identification of susceptibility genes of Kashin-Beck disease. PloS one *9*, e92298.

Yao, B., Zhang, J., Dai, H., Sun, J., Jiao, Y., Tang, Y., Wu, J., and Shi, Y. (2007). Solution structure of the second SH3 domain of human CMS and a newly identified binding site at the C-terminus of c-Cbl. Biochim Biophys Acta *1774*, 35-43.

Yoon, H.Y., Kales, S.C., Luo, R., Lipkowitz, S., and Randazzo, P.A. (2011). ARAP1 association with CIN85 affects epidermal growth factor receptor endocytic trafficking. Biology of the cell / under the auspices of the European Cell Biology Organization *103*, 171-184.

Yoshikawa, M., Kajiho, H., Sakurai, K., Minoda, T., Nakagawa, S., Kontani, K., and Katada, T. (2008). Tyr-phosphorylation signals translocate RIN3, the small GTPase Rab5-GEF, to early endocytic vesicles. Biochem Biophys Res Commun *372*, 168-172.

Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brugge, J.S., and Schreiber, S.L. (1992). Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science (New York, N.Y.) *258*, 1665-1668.

Zafra-Ruano, A., and Luque, I. (2012). Interfacial water molecules in SH3 interactions: Getting the full picture on polyproline recognition by protein-protein interaction domains. FEBS letters *586*, 2619-2630.

Zarrinpar, A., Park, S.H., and Lim, W.A. (2003). Optimization of specificity in a cellular protein interaction network by negative selection. Nature *426*, 676-680.

Zeigerer, A., Gilleron, J., Bogorad, R.L., Marsico, G., Nonaka, H., Seifert, S., Epstein-Barash, H., Kuchimanchi, S., Peng, C.G., Ruda, V.M.*, et al.* (2012). Rab5 is necessary for the biogenesis of the endolysosomal system in vivo. Nature *485*, 465-470.

Zhang, Q.C., Petrey, D., Deng, L., Qiang, L., Shi, Y., Thu, C.A., Bisikirska, B., Lefebvre, C., Accili, D., Hunter, T.*, et al.* (2012). Structure-based prediction of protein-protein interactions on a genome-wide scale. Nature *490*, 556-560.

Zhang, Q.C., Petrey, D., Garzon, J.I., Deng, L., and Honig, B. (2013). PrePPI: a structure-informed database of protein-protein interactions. Nucleic acids research *41*, D828-833.

Zhao, J., Bruck, S., Cemerski, S., Zhang, L., Butler, B., Dani, A., Cooper, J.A., and Shaw, A.S. (2013). CD2AP links cortactin and capping protein at the cell periphery to facilitate formation of lamellipodia. Molecular and cellular biology *33*, 38-47.

Zwart, P.H., Afonine, P.V., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., McKee, E., Moriarty, N.W., Read, R.J., Sacchettini, J.C.*, et al.* (2008). Automated structure solution with the PHENIX suite. Methods in molecular biology (Clifton, N.J.) *426*, 419-435.