Structure and Function of PH and WW domains

Marko Hyvönen

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Science
of the University of Helsinki for public critisism
in the auditorium 1041 of the Viikki Biocenter 2, Viikinkaari 5, Helsinki,
on September 26th, 1997, 3 o'clock in the afternoon

Structure and Function of PH and WW domains

Marko Hyvönen

Structural Biology Programme
European Molecular Biology Laboratory
Heidelberg
GERMANY

and

Division of Biochemistry
Department of Biosciences
Faculty of Science
University of Helsinki
FINLAND

Helsinki 1997

Supervisor:

Matti Saraste, PhD
European Molecular Biology Laboratory
Heidelberg
Germany

Referees:

Kari Alitalo, MD PhD
Haartman Institute
University of Helsinki
Finland

Juha Rouvinen, PhD
Department of Chemistry
University of Joensuu
Finland

ISBN 951-45-9158-5 (PDF version)

Helsingin yliopiston verkkojulkaisut
Helsinki 2000

Mater artium necessitas

To Maija

Contents

List of original publications 9
List of abbreviations 10

1 Introduction **11**
1.1 Signal transduction in eukaryotic cells 11
1.2 Small domains in signal transduction 12
1.2.1 Src homology domains 12
1.2.2 Pleckstrin homology domains 16
1.2.3 Phosphotyrosine binding domains 23
1.2.4 WW domains 25

2 Aims of this study **27**

3 Materials and methods **28**
3.1 Cloning, mutagenesis and expression 28
3.1.1 β-spectrin PH domain 29
3.1.2 Btk PH domain 29
3.1.3 WW domain 30
3.2 Panning of a phage display library 31
3.3 CD spectroscopy 32
3.4 Crystallography 32
3.4.1 Crystallisation of PH domains 32
3.4.2 Data collection and processing 33
3.4.3 Structure determination and refinement 34
3.5 NMR and modelling 35

4 Results **36**
4.1 The β-spectrin PH domain (I and unpublished) 36
4.1.1 The ligands of the β-spectrin PH domain 36
4.1.2 The structure 38
4.2 The Btk PH domain (III) 39
4.2.1 The structure of the Btk PH domain 39
4.2.2 XLA mutations in the Btk PH domain 40
4.2.3 The structure of the Btk motif 41

4.3 The WW domain of YAP65 (II) 42
4.3.1 The structure 42
4.3.2 Ligand binding 44

5 Discussion **45**
5.1 Inositol phosphate binding site in PH domains 45
5.2 The ligands of the PH domains 46
5.3 Structure and function of the Btk motif 49
5.4 WW domain 50
5.5 Conclusions 50

Acknowledgements **52**

Bibliography **53**

Appendix: Original articles **67**

List of original publications

This thesis is based on following original publications, which will later be referred to by their Roman numerals. In addition some unpublished results will be presented.

I Hyvönen, M., Macias, M. J., Nilges, M., Oschkinat, H., Saraste, M. and Wilmanns, M. (1995) Structure of the binding site for inositol phosphates in a PH domain. *EMBO J*, **14**:4676–4685.

II Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M. and Oschkinat, H. (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature*, **382**:646–649.

III Hyvönen, M. and Saraste, M. (1997) Structure of PH domain and Btk motif from Bruton's tyrosine kinase: Molecular explanations for the X-linked agammaglobulinemia. *EMBO J*, **16**:3396–3404.

List of abbreviations

The standard one-letter abbreviations are used for amino acids. Point mutations are labelled like R28C, in which arginine 28 is mutated to cysteine.

3BP-1	SH3 binding protein 1
β-ARK	β-adrenergic receptor kinase
Btk	Bruton's tyrosine kinase
DAG	diacyl glycerol
DH	Dbl homology (domain)
ENaC	amiloride-sensitive epithelial sodium channel
FBP	formin binding protein
GRB2	growth factor receptor-bound protein 2
g-PIns(4,5)P_2	L-α-glycero-phospho-D-*myo*-inositol (4,5)-bisphosphate
GST	glutathione-S-transferase
Ins(1,3,4)P_3	D-*myo*-inositol (1,3,4)-trisphosphate
Ins(1,4,5)P_3	D-*myo*-inositol (1,4,5)-trisphosphate
Ins(1,3,4,5)P_4	D-*myo*-inositol (1,3,4,5)-tetrakisphosphate
IRS-1	insulin receptor substrate 1
K_D	dissociation constant
NMR	nuclear magnetic resonance spectroscopy
NOE	nuclear Overhauser effect
PCR	polymerase chain reaction
PDB	Protein Data Bank
PH	pleckstrin homology (domain)
PI3K	phosphoinositide-3-OH kinase
PPII	polyproline type II (helix/conformation)
PTB	phospho-tyrosine binding (domain)
PtdIns(3,4)P_2	phosphatidylinositol (3,4)-bisphosphate
PtdIns(4)P	phosphatidylinositol (4)-phosphate
PtdIns(4,5)P_2	phosphatidylinositol (4,5)-bisphosphate
PtdIns(3,4,5)P_3	phosphatidylinositol (3,4,5)-trisphosphate
PLC	phospholipase C
PTK	protein tyrosine kinase
pY	phospho-tyrosine
RK	rhodopsin kinase
SH2	Src homology 2 (domain)
SH3	Src homology 3 (domain)
YAP65	Yes-kinase associated protein
xid	x-linked immunodeficiency
XLA	X-linked agammaglobulinemia
Å	Ångström = 10^{-10} m

1 INTRODUCTION

1.1 Signal transduction in eukaryotic cells

Correct transmission of extracellular signals is vital both for unicellular and multicellular organisms. Stimulation of cells from outside can trigger complicated cascades of signal transduction which result in cellular responses such as growth, differentiation and movement. These signals can be transduced both by small molecules, like cyclic AMP, Ca^{2+} or Ins(1,4,5)P_3 and by networks of interacting proteins (Berridge, 1993; Hunter, 1997). Small molecules allow rapid and often strong amplification of signals, whereas the protein networks offer better controlled and more precise signalling routes.

In signal transduction the correct sub-cellular localisation of participating molecules is very important and re-localisation is an essential part of the activation process. As the signals are coming from outside of the cell and transmitted by transmembrane receptors, the plasma membrane plays an important role as a meeting point for signalling proteins. Association with the membrane can either be direct or indirect. Direct membrane binding of a cytoplasmic protein is for example achieved by covalent lipid modification of the proteins, which itself can be dynamically regulated (Mumby, 1997). Indirect membrane localisation can occur through specific protein-protein interactions, which can as well be regulated by covalent protein modification such as phosphorylation.

Phosphorylation of proteins is one of the hallmarks of intracellular signalling. In higher eukaryotes phosphorylation of tyrosines in particular has a major role in early events of signal transduction. Phosphorylation can serve several purposes: enzymes can be regulated by modification of the active site (and other parts) and phospho-tyrosines serve as docking sites for specialised interaction domains, as will be discussed later. Modification of phosphatidylinositol lipids has gained wide attention in recent years and kinases and phosphatases responsible for this are intensively studied (Carpenter and Cantley, 1996). The role of differentially phosphorylated phosphoinositides is

not clearly understood, but their role as specific membrane targeting signals is slowly emerging (Toker and Cantley, 1997).

1.2 Small domains in signal transduction

Signals inside the cells are often transduced by specific protein-protein interactions, which are mediated by a number of small, non-catalytic protein modules. Networks of interactions bring various proteins together and allow efficient transduction of extracellular stimuli.

Use of small, independent protein domains has allowed nature to use similar activation mechanisms by shuffling these domains between different proteins. The use of similar domains in different proteins allows also cross-talk between signalling pathways and creates the possibility for a protein to participate in several different networks. In addition this can create functional redundancy to signalling pathways. Homologous domains can also be used in cytoskeletal proteins, providing a connection between signal tranduction and structural regulation of the cell.

A large number of signalling domains have been described in the literature and in the following discussion the emphasis is given to domains which are widely distributed in proteins and which share either functional or structural similarities with PH or WW domains (figure 1.1).

1.2.1 Src homology domains

The Src homology domains 2 (SH2) and 3 (SH3) are among the first signalling modules described. They were identified as regions outside the catalytic kinase domain (the SH1 domain) of cytoplasmic protein tyrosine kinases (PTKs) by sequence homology with other signalling proteins (Mayer *et al.*, 1988; Sadowski *et al.*, 1986). These domains are also the hallmark of the well characterised Src-superfamily of PTKs (Superti-Furga and Courtneidge, 1995).

SH2 domains

The SH2 domains are typically found in proteins involved in growth factor signalling such as PTKs, phospholipase C (PLC) isoforms and adaptor proteins like GRB2 (Russell *et al.*, 1992). No SH2 domains have been found in yeasts or plants, which is in correspondence with the lack of tyrosine phosphorylation in these organisms.

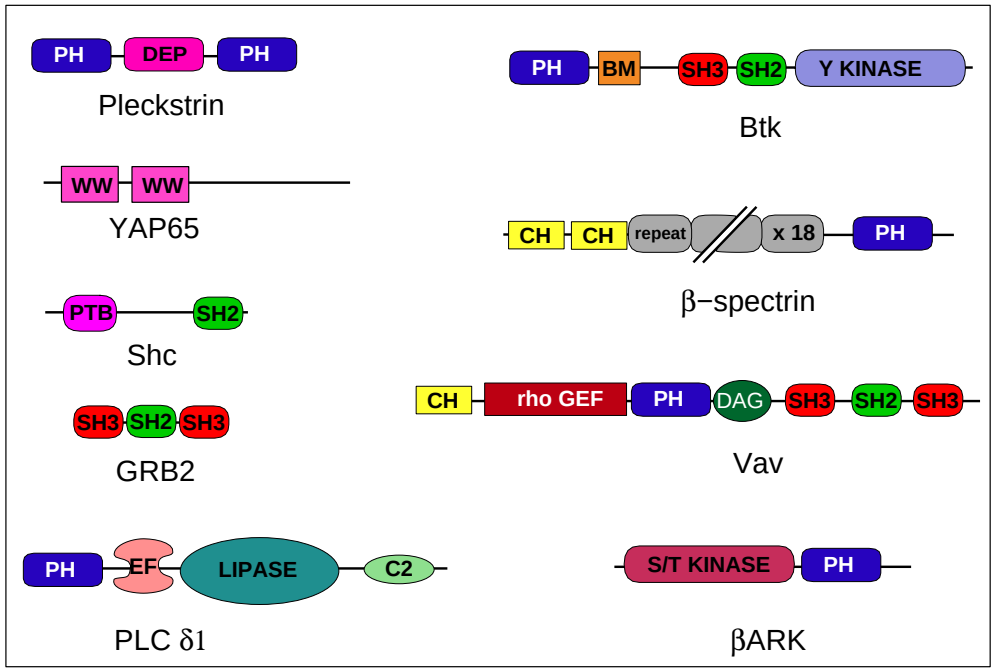


Figure 1.1 Modular proteins and their domains. Picture was drawn following the suggestions of Bork *et al.* (1997) with some modifications and additions.

The SH2 domains interact with phosphorylated tyrosine residues (pY) in their target proteins. The different SH2 domains distinguish between different ligand peptides by the sequences surrounding the phospho-tyrosine residue (Piccione *et al.*, 1993). In particular the sequence following the phospho-tyrosine is important for specificity (Songyang *et al.*, 1993, 1994). Typically the SH2 domains show greatest selectivity for residues +1 and +3 [1]. Only GRB2 SH2 domain has shown clear selectivity in position +2, namely for asparagine (Gay *et al.*, 1997; Songyang *et al.*, 1994). Dissociation constants for SH2-peptide interactions are low, in nanomolar range.

The structure of the SH2 domain consists of a central β-sheet sandwiched between two α-helices (Waksman *et al.*, 1992). The N- and C-termini are very close to each other, a feature later found to be common for many other small domains as well. This feature is believed to facilitate the introduction of these domains to new proteins without disturbing the existing fold of the host protein.

All the SH2 domains bind the phospho-tyrosyl residue in a specific pocket with a strictly conserved arginine βB5 interacting with the phosphate group of

[1] The phospho-tyrosine position is assigned as 0 and other residues are numbered in relation to it, ie. +3 refers to the 3rd residue C-terminal to pY.

the phospho-tyrosine and the loop between β-strands B and C closing over the phosphate (Eck *et al.*, 1993; Lee *et al.*, 1994; Rahuel *et al.*, 1996; Waksman *et al.*, 1992, 1993). In the crystal structures of Src and Lck SH2 domains, the peptide lies in extended conformation perpendicular to the central β-sheet of the SH2 domain and a deep pocket serves as a docking site for a hydrophobic residue in the position +3 of the peptide (Eck *et al.*, 1993; Waksman *et al.*, 1993). The Syp SH2 domain binds the phosphopeptides in same orientation as Src and Lck, but rather than having a specific pocket for the +3 residue, it binds the peptide in a continuous groove which allows a hydrophobic residue in position +1 (Lee *et al.*, 1994). These results are in accordance with the different specificities of these domains for the +1 position; the Src family members prefer glutamate in this position whereas Syp chooses a hydrophobic residue (Songyang *et al.*, 1993).

A recent structure of the C-terminal SH2 domain of GRB2 in complex with a phospho-peptide has revealed an alternative binding mode, where the peptide adopts a β-hairpin conformation and +2 residue interacts specifically with the domain (Rahuel *et al.*, 1996). GRB2 has the only SH2 domain which discriminates residues in this position. The different specificity and mode of binding can be correlated with features of GRB2's primary structure.

Structure of the tandem SH2 domains from ZAP-70 protein tyrosine kinase in complex with doubly phosphorylated peptide from the ζ subunit of the T-cell receptor has unveiled an interesting mechanism by which multiple domains can co-operate to achieve high specificity and affinity interactions with their ligands (Hatada *et al.*, 1995). The two SH2 domains are connected by a 65 residue linker which by forming a coiled-coil brings the domains side-by-side creating a long, continuous binding site for the peptide. Remarkably, the binding pocket for the second phospho-tyrosine is formed by the interface of the two SH2 domains. The relative orientation of the two SH2 domains enables highly selective interaction with the natural ligand. The sequence surrounding the two phospho-tyrosines does not affect the specificity of the binding, whereas correct distance between them is crucial for interaction. Structure of the tandem SH2 domains of protein tyrosine phosphatase SH-PTP2 shows different inter-domain arrangement and two phosphopeptides bind individual domains in roughly anti-parallel orientations (Eck *et al.*, 1996b). The SH2 domains are fixed in their relative positions by a disulfide linkage and are predicted to restrict both the length and conformation of a natural biphosphorylated ligand peptide.

SH3 domains

The SH3 domains contain ca. 60 amino acids and are characterised by several highly conserved aromatic residues. They are more universal than SH2 domains and are found in yeasts, invertebrates and vertebrates (Musacchio *et al.*, 1992a, 1994b). In addition to typical signalling proteins, the SH3 domains are also present in cytoskeletal proteins providing a possible link between the signal transduction pathways and the morphological changes of the cell (Lehto *et al.*, 1988; Musacchio *et al.*, 1992a).

In 1992 Cicchetti and co-workers identified the first SH3 binding protein, 3BP1. A deletion analysis and site-directed mutagenesis of 3BP1 (and of a second SH3 binding molecule molecule 3BP2) pinned down a ten amino-acid proline rich sequence as the binding site for SH3 domains (Ren *et al.*, 1993). A similar recognition motif has since been confirmed for several different SH3 domains. These peptides can adopt left-handed poly-proline type II (PPII) helical conformations in solution and they bind to SH3 domains typically with micromolar affinities (Viguera *et al.*, 1994).

The structure of the SH3 domain is a five-stranded anti-parallel β-barrel (Musacchio *et al.*, 1992b; Noble *et al.*, 1993; Yu *et al.*, 1992). The most distinctive feature of this domain is the clustering of conserved aromatic residues on one of its sides. This hydrophobic surface has proved to be the interaction site for the ligands of SH3 domain.

As predicted before, the proline-rich peptides bind SH3 domain in PPII conformation (Musacchio *et al.*, 1994a). The bound peptide interacts with the SH3 domain by intercalating its sidechains into the ladder of aromatic residues on the surface of the domain. The PPII conformation allows the side-chains of the peptide to be positioned correctly for this interaction. The interactions between the proline-rich peptide and SH3 domain are not, however, very specific - a feature that partly explains the ability of an SH3 domain to interact with several different peptides, and *vice versa*, with similar affinities.

The PPII helix is a pseudo-symmetrical structure which allows binding of peptide ligands to SH3 domains in two opposite orientations. The SH3 domain of SEM-5, a *Drosophila melanogaster* orthologue of GRB2, binds the proline-rich peptide of Sos in an opposite orientation to that of the Abl and Fyn SH3 domains (Lim *et al.*, 1994; Musacchio *et al.*, 1994a). The peptide-domain interactions are very similar in both peptide orientations and the residues outside the proline-rich core of the peptide determine the orientation of the binding (Lim *et al.*, 1994).

The structure of Nef-protein from human immunodeficiency virus (HIV) in complex with R96L mutant of Fyn SH3 domain demonstrates how interactions outside the PPII helix contribute to the specificity and affinity of the interaction

(Lee *et al.*, 1996). The SH3 domain binds the proline-rich tail of Nef in a very similar manner as it binds an isolated peptide. The specificity and higher affinity of the interaction (over 300 fold difference in binding to full-length Nef vs. proline-rich peptide of Nef) is governed by interactions between other parts of the molecules. The so-called RT-loop of the SH3 domain in particular interacts very specifically with the Nef protein.

The latest twist to the SH3 domain - peptide interactions was revealed by the crystal structures of Src and Hck tyrosine kinases (Sicheri *et al.*, 1997; Xu *et al.*, 1997). These structures describe the kinases in the inactive (or closed) state, which is caused by intramolecular interactions of SH2 and SH3 domains. The SH2 domain interacts with the inhibitory phospho-tyrosine in the very C-terminus of the kinase (Y527 in Src) as was known earlier. The SH3 domain binds the linker between the SH2 and kinase domains and sandwiches it in PPII conformation against the upper lobe of the kinase. This linker is variable in sequence between the Src-like kinases and does not carry the consensus elements for classical SH3 ligands. But given the functional similarities between the Src kinases, this inhibitory mechanism is believed to be their common feature. The SH3 domain was known before to be essential for inhibition of Src, but the mechanism by which it acts is somewhat surprising (Superti-Furga *et al.*, 1993). Time will show if more ligands will be found, which, despite the lack of prolines, bind to SH3 domains in the PPII conformation.

1.2.2 Pleckstrin homology domains

The pleckstrin homology (PH) domains were originally identified as an internal repeat in pleckstrin (also called as p47), a phospho-protein from blood platelets (Haslam *et al.*, 1993; Mayer *et al.*, 1993; Tyers *et al.*, 1988). The initial discovery of the domain soon led to more extensive database searches and to date more than 100 PH domains in different proteins have been identified (Gibson *et al.*, 1994; Musacchio *et al.*, 1993; Saraste and Hyvönen, 1995).

The PH domains have ca. 100 amino acids, but some domains can also contain large insertions including several other domains. The PH domains are very poorly conserved and typical pairwise identities between different domains are in the range of 10–20 %. The only highly conserved residue is the nearly invariant tryptophan in the C-terminus of the domain. A multiple sequence alignment of representative PH domains is shown in the figure 1.2.

The low sequence conservation between PH domains makes the database searches very difficult and sophisticated profile searches are needed in order to identify more distant members of the family and still some might be missed (for an example on phospholipase C βs, see Parker *et al.*, 1994).

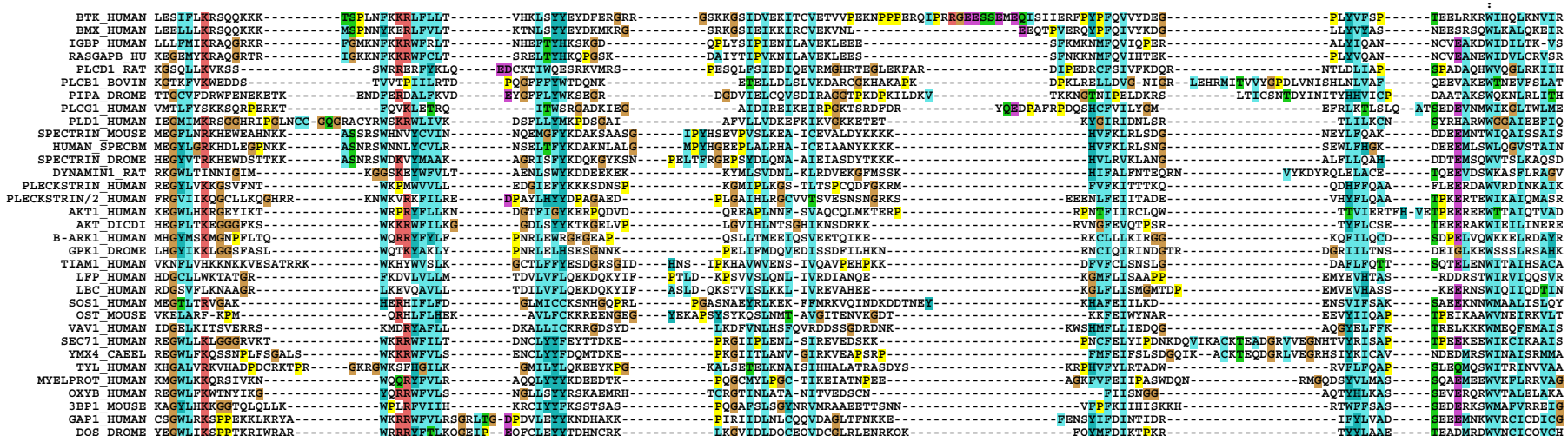


Figure 1.2 Multiple sequence alignment of PH domains. The sequences are coloured according to conservation. Hydrophobic residues are on blue background, negatively charged on purple, prolines on yellow, positively charged on red, glycines on brown and serines and threonines on green.

Distribution of PH domains

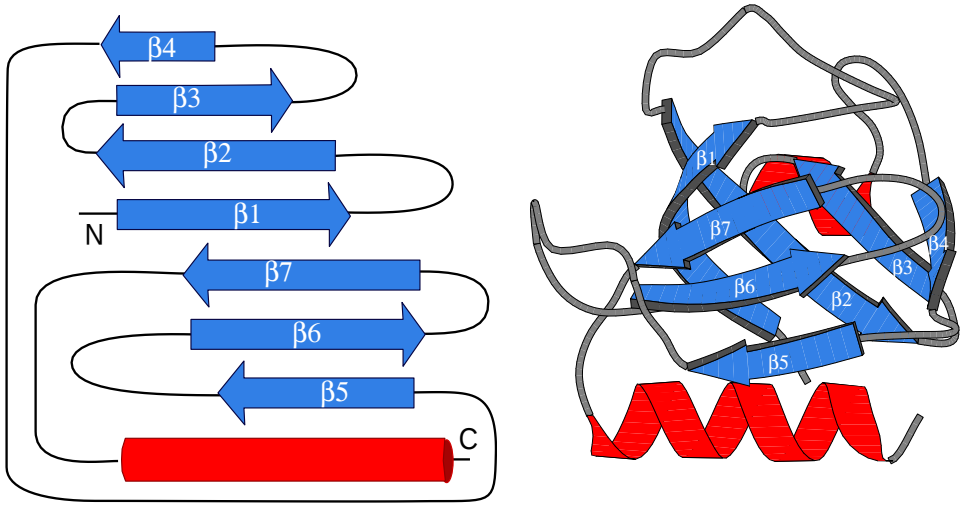
The PH domains are found throughout the animal kingdom and in yeasts—no PH domain has been detected in plants or in bacteria (Musacchio *et al.*, 1993; Gibson *et al.*, 1994). The proteins carrying PH domains are either involved in signal transduction or they are part of the cytoskeleton. The PH domains occupy variable positions in proteins, as demonstrated by pleckstrin, which has PH domains both in the extreme N- and C-termini separated by a so-called DEP domain (This domain is found in proteins like *dishevelled*, egl-10 and pleckstrin, Ponting and Bork, 1996). The PH domain containing proteins are typically modular, *ie.* they are composed of various independent domains (Saraste and Hyvönen, 1995). The PH domains seem independent of other domains. However some domains are always found next to a PH domain. The Dbl homology domain (DH, also called Cdc24 or RhoGEF domain), a guanine nucleotide exchange factor of small G-proteins of Rho and Rac families, is always followed by a PH domain (Cerione and Zheng, 1996). Another such example is the Btk motif which is in all cases preceded by a PH domain (Vihinen *et al.*, 1994).

Structure of PH domains

The first three-dimensional structures determined of PH domains were the N-terminal PH domain of human pleckstrin and the domain of murine β-spectrin, both by NMR (Macias *et al.*, 1994; Yoon *et al.*, 1994). After these the structures of dynamin, PLCδ1, β-spectrin (both from mouse and fruit fly), Sos-1 and BTK PH domains have been solved either by X-ray crystallography or NMR (I;III; Koshiba *et al.*, 1997; Downing *et al.*, 1994; Ferguson *et al.*, 1994; Fushman *et al.*, 1995; Timm *et al.*, 1994; Ferguson *et al.*, 1995; Zhang *et al.*, 1995b).

The structure of the PH domain consists of a seven-stranded β-sandwich which is capped by a C-terminal α-helix. The conserved tryptophane in the C-terminus of the domain is part of the hydrophobic core. Similar to SH3 and SH2 domains, the N- and C-termini of PH domains are very close in space, facilitating the insertion of this domain into new proteins. The topology of the PH domain and the structure of the β-spectrin PH domain are shown in figure 1.3.

The different PH domains share the same fold, but differ greatly in the variable loops connecting the β-strands. A superpositioning of Btk, PLCδ1, β-spectrin, dynamin and pleckstrin PH domain structures in figure 1.4 illustrates this feature. The variable loops play an important role in ligand interactions of PH domains and introduce functional variability to this family, as will be discussed later. The loops can also carry additional secondary structure elements, such as the short α-helices in the loop 3-4 of β-spectrin and loop 5-6 of PLCδ1



A Topology of the PH domains **B** β-spectrin PH domain

Figure 1.3 Topology and structure of the PH domains. The β-sheets are numbered from 1-7 both in the topology diagram and in the ribbon representation of the β-spectrin PH domain. The additional α-helix in the β-spectrin PH domain is not displayed in the topology diagram. Figure B and all the following protein images are prepared with Molscript (Kraulis, 1991).

and Btk PH domains.

Another common feature shared by the PH domains for which the three-dimensional structures have been determined is their striking electrostatic polarisation. The positively charged end of the domain is the area for inositol phosphate binding as demonstrated by the crystal structures of PLCδ1 and β-spectrin PH domains in complex with Ins(1,4,5)P_3 (I; Ferguson *et al.*, 1995). In addition NMR studies with dynamin, pleckstrin and Sos1 PH domains have localised the binding site for inositol phosphates in this part of the molecules (Koshiba *et al.*, 1997; Harlan *et al.*, 1995; Zheng *et al.*, 1996a; Salim *et al.*, 1996).

Although this polarisation is very pronounced in the known structures, it is not a general property of all the PH domains. In a recent homology modelling study Blomberg and Nilges (1997) have divided PH domains in different categories depending on their electrostatic properties. Groups of PH domains were identified which are either very weakly polarised or have reversed polarity to that described above. These domains are likely to have different function and / or ligands compared to the ones studied in detail until now.

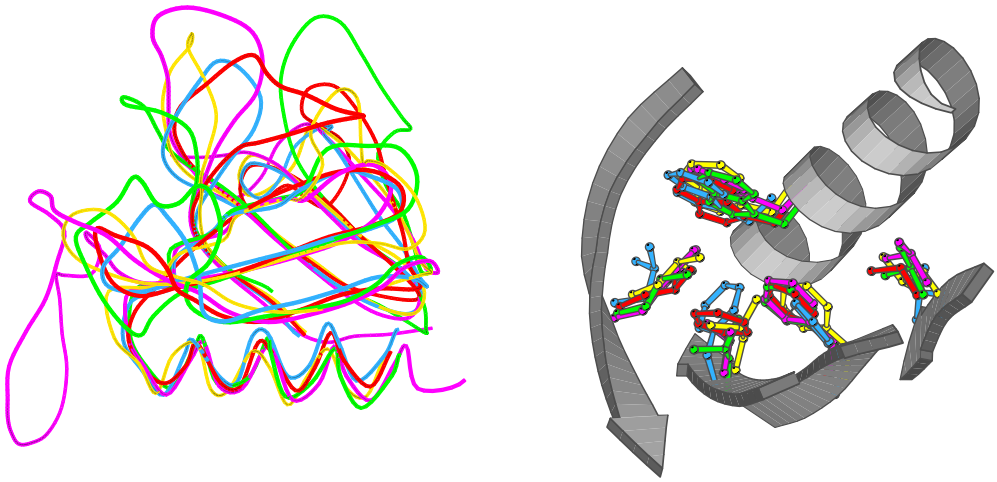
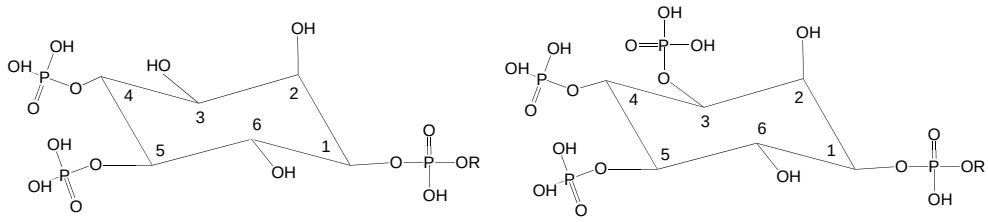


Figure 1.4 Superpositioning of the PH domains. In A the smoothed Cα traces of superimposed Btk (purple), β-spectrin (yellow), dynamin (green), PLCδ1 (red) and pleckstrin (light blue) PH domains are shown. In B the most conserved hydrophobic residues from the same structures are shown as ball-and-stick models. Colouring is the same as in A. The secondary structure elements of the Btk PH domain are shown in grey.

Ligands of PH domains

Unlike other widespread signal transduction domains like SH2 and SH3, the PH domains do not seem to be directly involved in protein-protein interactions. Instead, the ligands for a large fraction of PH domains are membrane-bound inositol phospholipids and / or their soluble head-group derivatives (figure 1.5).



Ins(1,4,5)P_3, R=H

PtdIns(4,5)P_2,R=DAG

Ins(1,3,4,5)P_4, R=H

PtdIns(3,4,5)P_3, R = DAG

Figure 1.5 Chemical formulas of some of the inositol phosphates that bind to PH domains. In the lipid form, the R is a diacyl glycerol (DAG) group

In 1993 an N-terminally truncated form of phospholipase C δ1 was shown to have reduced affinity for PtdIns(4,5)P_2 containing membranes, although it was still catalytically active (Cifuentes *et al.*, 1993; Yagisawa *et al.*, 1994). [2] The truncated enzyme showed impaired processivity, which implied that elements outside to the catalytic core of the enzyme are required to keep PLCδ1 on the membrane. This deletion had occurred in the middle of the yet unidentified PH domain. Later studies have shown that this domain is required both for processive catalysis of the enzyme and for membrane association of PLCδ1 *in vivo* (Paterson *et al.*, 1995).

The first direct evidence for PH domain-lipid interaction came from studies of the N-terminal PH domain of pleckstrin (Harlan *et al.*, 1994). The pleckstrin PH domain bound specifically to vesicles containing PtdIns(4,5)P_2 and this interaction could be inhibited by excess of soluble Ins(1,4,5)P_3. The residues showing chemical shift changes in NMR spectra upon ligand binding were mapped to β-strands 1 and 2 and to loop between β-strands 3 and 4 (loop 3-4). The binding site was later confirmed by site-directed mutagenesis and a dissociation constant of 13.5 μM for PH domain PtdIns(4,5)P_2 interaction was determined (Harlan *et al.*, 1995).

The β-spectrin PH domain binds soluble inositol phosphates with comparable affinity (I). It binds both Ins(1,4,5)P_3 and Ins(1,3,4,5)P_4 and a derivative of PtdIns(4,5)P_2, L-α-glycero-phospho-D-*myo*-inositol (4,5) bisphosphate. The PLCδ1 PH domain has the highest known affinity to Ins(1,4,5)P_3 with an estimated dissociation constant (K_D) of 210 nM (Lemmon *et al.*, 1995). Affinity for corresponding lipids is slightly lower. Several other PH domains have since been shown to bind inositol phosphates; they are listed in table 1.2.2.

The structures of β-spectrin and PLCδ1 PH domains in complex with Ins(1,4,5)P_3 are known and will be discussed later in detail (I and Ferguson *et al.*, 1995).

Increasing amount of *in vivo* data is emerging supporting the role of PH domains as membrane anchors. In cells pleckstrin is associated with membranes and this association requires an intact N-terminal PH domain (Ma *et al.*, 1997). Interestingly, phosphorylated serines just downstream of this domain are regulating this interaction. The C-terminal PH domain does not affect the membrane association and its function is not known. Dbl, a DH domain containing protein, depends on the PH domain for membrane association and requires it to retain oncogenic activity (Zheng *et al.*, 1996b). Related protein Tiam-1 is recruited to membrane by the first of its two PH domains, but in contrast to Dbl, the PH domain next to DH domain can be deleted without noticable effect on cellular localisation of the protein (Michiels *et al.*, 1997). Rapid endocytosis me-

[2]The nomenclature used for the phosphoinositides follows that suggested by Divecha and Irvine (1995). For abbreviations, see page 10.

Protein	Ligand	K_D	Reference
pleckstrin	PtdIns(4,5)P_2	13.4 μM	Harlan *et al.* (1995)
PLCδ1	Ins(1,4,5)P_3	210 nM	Lemmon *et al.* (1995)
PLCδ1	PtdIns(4,5)P_2	1.09 μM	Lemmon *et al.* (1995)
dynamin	PtdIns(4)P	1.8 μM	Zheng *et al.* (1996a)
dynamin	PtdIns(4,5)P_2	1.23 mM	Salim *et al.* (1996)
β-spectrin	g-PIns(4,5)P_2	40 μM	I
Akt/RAC	PtdIns(3,4)P_2	400 nM	Frech *et al.* (1997)
SOS1	PtdIns(4,5)P_2	1.8 μM	Kubiseski *et al.* (1997)
Btk	Ins(1,3,4,5)P_4	40 nM	Fukuda *et al.* (1996)
Btk	PtdIns(3,4,5)P_3	800 nM	Rameh *et al.* (1997)
Gap1^m	Ins(1,3,4,5)P_4	40 nM	Fukuda and Mikoshiba (1996)
p130	Ins(1,4,5)P_3	12.2 μM	Takeuchi *et al.* (1996)

Table 1.1 The table shows inositol phosphate ligands for isolated PH domains and their estimated dissociation constants. (In addition the PH domains of Tsk, pleckstrin C-terminus and Ras-GAP have been reported to bind PtdIns(4,5)P_2 containing vesicles, but little or no data for the specificity or affinity have been presented (Harlan *et al.*, 1994)).

diated by dynamin-1 requires intact PH domain and can be blocked by exogenous dynamin-1 PH domain (Artaljo *et al.*, 1997). This effect is specific for this isoform and is mediated by two amino acids which are different between the PH domains of dynamin 1 and 2. GTPase activity of dynamin is also activated by PtdIns(4,5)P_2 (Salim *et al.*, 1996; Lin and Gilman, 1996). The N-terminal PH domain is required (but not sufficient), for Ca^{2+} mediated membrane targeting of Ras-GRF (Buchsbaum *et al.*, 1996). PH domain is indispensible for effective signalling between insulin receptor and insulin receptor substrate 1 (IRS-1) (Myers Jr. *et al.*, 1995). Its deletion abolishes tyrosine phosphorylation of IRS-1 and decouples it from downstream signalling (Yenush *et al.*, 1996). The proto-oncogene Akt is one of the only proteins that are known to be directly regulated by phosphoinositide 3-OH-kinase (PI3K) (Kaplan and Cantley, 1997). Akt carries a PH domain in its N-terminus and depends on this domain for activation by PtdIns(3,4)P_2 (Franke *et al.*, 1997; Klippel *et al.*, 1997). The Akt is also claimed to form homo-oligomeric complexes through its PH domain and adjacent region (Datta *et al.*, 1995; Franke *et al.*, 1997). The guanine nucleotide exchange factors of the small G-protein ARF1 (containing a so-called Sec7 domain and a PH domain) also bind to membranes and phosphoinositide vesicles with their PH domains (Chardin *et al.*, 1996; Klarland *et al.*, 1997). The Sos PH domain is localized to specific areas of the plasma membrane, namely the leading edge, in serum-induced fibroblasts (Chen *et al.*, 1997). Surprisingly this was independent of the PtdIns(4,5)P_2 binding capasity, pointing to existence of additional

membrane bound ligands or binding partners for PH domains. Microinjected Sos PH domain has dominant-negative effect on Sos dependent signalling, suggesting that it can compete with the endogenous protein for its ligands.

The most often mentioned protein ligands for PH domains are the $\beta\gamma$-subunits of trimeric G-proteins. In β-adrenergic receptor kinase (β-ARK), a serine-threonine protein kinase responsible for the desensitisation of β-adrenergic receptor, a PH domain and its C-terminal extension interact with $\beta\gamma$-subunits. This interaction localises β-ARK to plasma membrane and activates the kinase (Koch *et al.*, 1993). This activation can be further enhanced by the presence of PtdIns(4,5)P_2 in the lipid vesicles carrying the $\beta\gamma$-subunits (Pitcher *et al.*, 1995). Some other PH domains have also been shown to interact with the $\beta\gamma$-subunits, but the interactions described are often weak and require elements outside the PH domain (Touhara *et al.*, 1994; Wang *et al.*, 1994; Tsukada *et al.*, 1994).

PH domains in disease

Bruton's tyrosine kinase is a cytoplasmic protein kinase of the Tec-family, which is affected in hereditary immunodeficiency, X-linked agammaglobulinemia (XLA) (Vetrie *et al.*, 1993; Tsukada *et al.*, 1993). Inactivation of Btk by XLA mutations ceases the maturation of B-cells to pre-B-cell stage. As a result of this the patients have very low levels of circulating B-cells and immunoglobulins in their blood and are very vulnerable to infectious diseases (Smith *et al.*, 1994).

Large number of XLA mutations have been detected throughout the molecule including several point mutations in the PH domain (Vihinen *et al.*, 1996). A Btk PH domain mutation R28C is also the cause of a similar disease, x-linked immunodeficiency (*xid*), in mice (Rawlings *et al.*, 1993; Thomas *et al.*, 1993). Analysis of XLA mutations in Btk can provide information on the function of individual domains in Btk and to the maturation process of B-cells in general.

XLA (and *xid*) is the only inherited disease, which is caused by mutations in the PH domain The Btk PH domain binds specifically Ins(1,3,4,5)P_4 and PtdIns(3,4,5)P_3 (table 1.2.2) and molecular reasons for inactivation of Btk-mediated signal transduction by mutations in this domain will be discussed later.

1.2.3 Phosphotyrosine binding domains

The phosphotyrosine binding (PTB, also called PID or SAIN) domains are a second class of domains in addition to SH2 domains that specifically recog-

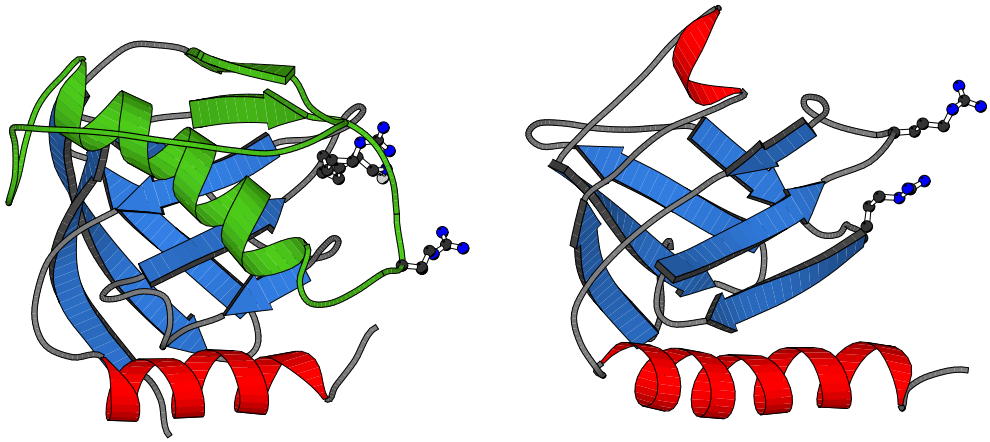
nise phosphorylated tyrosines in their target proteins. The PTB domain was first discovered in the N-terminus of an adaptor molecule Shc for it's capacity to interact with tyrosine-phosphorylated proteins (Blaikie *et al.*, 1994; van der Geer *et al.*, 1995; Gustafson *et al.*, 1995; Kavanaugh and Williams, 1994). Sequence database searches using the experimentally defined PTB domain of Shc revealed the presence of PTB in 13 other proteins (Bork and Margolis, 1995). The PTB domains share very low sequence homology and for example the IRS-1 was never detected in database searches, although it is known from functional and structural studies to carry this domain.

The substrate specificity of PTB domains differs from that of SH2 domains. While the SH2 domains distinguish between different peptides from the sequence C-terminal to the phospho-tyrosine, the PTB domains differentiate the peptides by their N-termini. The Shc PTB domain recognises a consensus motif NP*X*pY in the target proteins c-ErbB2, EGF and IL4 receptors (Gustafson *et al.*, 1995; Kavanaugh *et al.*, 1995). The IRS-1 and Shc PTB domains recognise slightly different phosphotyrosine peptides which vary in the positions -1 , -5 and -6 with respect to phosphotyrosine (Trüb *et al.*, 1995). Dissociation constants of PTB domains for peptides derived from their biological binding partners can be as low as 20 nM (Zhou *et al.*, 1995a).

Structures of both Shc and IRS-1 PTB domains in complex with ligand peptides are known. The structure of Shc PTB domain was solved by NMR in complex with a peptide from nerve growth factor receptor TrkA (Zhou *et al.*, 1995b). The structure of IRS-1 domain has been solved by NMR in complex with IL4 receptor peptide and by X-ray crystallography in complex with insulin receptor peptide (Eck *et al.*, 1996a; Zhou *et al.*, 1996). A surprising feature of the PTB domains is their close structural similarity to PH domains (figure 1.6). The PTB and PH domains share the same core of seven-stranded β-sandwich capped by a C-terminal α-helix. The Shc PTB domain has an insertion in the first loop containing two additional β-strands and an α-helix, but the core of the domain is still the same (Zhou *et al.*, 1995b).

The PTB domains bind their peptide ligands between the C-terminal α-helix and the adjacent β-strand 5. The N-terminus of the peptide continues the β-sheet of the domain by hydrogen bonding to the β-strand 5. The NP*X*pY phosphopeptides have β-turn confirmation in solution and the same confirmation is retained upon binding to PTB domains (Eck *et al.*, 1996a; Trüb *et al.*, 1995; Zhou *et al.*, 1995b, 1996). The phosphotyrosine residue is coordinated both in Shc and IRS-1 by positively charged amino acids, which, however, are not equivalent between the domains. Arginine 67 in Shc PTB comes from the long insertion which is missing from IRS-1 and also the residues in the β-strand 5 and loop 6-7 (numbering corresponds to IRS-1) are different.

The specificity of PTB domains for particular peptides is affected by spacing



A Shc PTB domain **B** IRS-1 PTB domain

Figure 1.6 PTB domains of Shc and IRS-1. The secondary structure elements that correspond to those of PH domains are coloured as in figure 1.3. The insertion between β-strands 1 and 2 in the Shc domain is shown in green. The key residues coordinating the phospho-tyrosine peptide are shown as ball-and-stick models

of hydrophobic residues in the amino terminus of the ligand. An aliphatic leucine in position -5 binds to a hydrophobic pocket in Shc PTB domain, where as IRS-1 has a similar hydrophobic interaction with residues in positions -6 and -8 of the peptide.

The similarities with PH domains are further enhanced by the observation that Shc PTB domain can interact with PtdIns(4,5)P_2 vesicles (Zhou *et al.*, 1995b; Rameh *et al.*, 1997). It will be interesting to see if some of the PH domains could bind phospho(tyrosine)-peptides as well. However, comparison of SH2 and PTB domains reveals more differences than common features between these two domains. In contrast to the SH2 domain, the PTB domain i) recognises the substrates via the N-terminal end of the peptide, ii) binds the peptide parallel to the β-sheet, iii) binds the the peptides in a β-turn conformation and iv) coordinates the phosphotyrosine with variable, non-conserved residues.

1.2.4 WW domains

The WW domain was originally discovered as an additional internal repeat in murine Yes-kinase associated protein (YAP65), when it was compared with

chicken and human isoforms (Sudol *et al.*, 1995). The WW domains are among the smallest signal transduction modules with barely 40 residues. They are characterised by two fully conserved tryptophans (hence the name WW) and a proline. The WW domains are found, often in multiple copies, in proteins with diverse functions such as adaptor protein YAP65, cytoskeletal protein dystrophin, transcription-factor activator FE65 and peptidyl-prolyl isomerase Pin1 (André and Springael, 1994; Bork and Sudol, 1994; Lu *et al.*, 1996). In addition to animals and yeasts, WW domains are also found in plants (Sudol *et al.*, 1995).

Expression library searches with the WW domain of YAP65 have identified two putative protein ligands which share a common sequence motif PPPPY (Chen and Sudol, 1995). This proline-rich sequence is required for interaction with YAP65 WW domain, but does not seem to bind to WW domain of dystrophin (Chen and Sudol, 1995).

The Nedd4 WW domains recognise a similar sequence from the C-terminus of amiloride-sensitive epithelial sodium channel (ENaC) (Staub *et al.*, 1996). This interaction is required for regulation of the ion channel's activity and mutations or deletions of this proline-rich region are found in patients suffering from hereditary form of systemic hypertension, the Liddle's syndrome (Schild *et al.*, 1996; Staub *et al.*, 1996).

Proline-rich region of formins, nuclear phosphoproteins involved in limb and renal development, bind several different SH3 and WW containing proteins from λ-expression libraries (Chan *et al.*, 1996). The WW domains of these protein (called collectively as formin binding proteins or FBPs) are sufficient to bind formin. They can compete for the same binding site with various SH3 domains. The two WW domains of FBP11 recognise proteins carrying a consensus sequence PPLP, but do not bind the YAP WW binding protein with PPxY motif (Bedford *et al.*, 1997).

Both the PPLP motif binding to FBP11 and the proline-rich region of ENaC are recognised by SH3 domains of Abl and α-spectrin, respectively illustrating the overlapping ligand binding specificities between these two domain families (Bedford *et al.*, 1997; Rotin *et al.*, 1994).

The structure of YAP65 WW domain has been solved by heteronucler NMR and will be discussed later (II).

2 AIMS OF THIS STUDY

This study was initiated after the first structures of PH domains had been determined and information on potential ligands for these domains was beginning to emerge. The aims set for this study were:

- To study the structures of small interaction domains in signalling molecules.
- To identify ligands of these domains.
- To study the specificity of ligand interaction.
- To characterize these protein - ligand interactions structurally.

3 MATERIALS AND METHODS

3.1 Cloning, mutagenesis and expression

The domains were amplified from cDNA using polymerase chain reaction (PCR). Cleavage sites for selected restriction enzymes were included in the PCR primers. The amplified PCR fragments were digested and cloned into *E.coli* expression vectors under T7 promoter (Peränen *et al.*, 1996). All DNA constructs were verified by dideoxy sequencing. Clonings were carried out in the *E.coli* strain DH5α. Mutant domains were generated by PCR-directed mutagenesis and cloned into similar *E.coli* expression vectors as the wild-type proteins and sequenced.

The domains and their mutants were expressed in *E.coli* strain BL21(DE3) which carries a chromosomal copy of the T7 RNA polymerase (Studier *et al.*, 1990; Hyvönen and Saraste, 1997). The bacterial cultures were grown at 37°C until the the optical density at 600 nm reached 0.8-1.0 after which the expression was induced by adding 400 μM isopropyl-β-D-thiogalactopyranoside (IPTG). If the expression was done at lowered temperature (as was the case with Btk PH domain -Btk motif), the bacterial cultures were cooled to the desired temperature prior induction. The induction was 3 hours at 37°C and 20 hours at 15°C.

The proteins were released from *E.coli* cells by French press lysis and purified to homogeneity by various chromatographic techniques and analysed by mass-spectroscopy. The results of the mass-analysis of β-spectrin and Btk PH domains and YAP65 WW domain are shown in table 3.1. The protein concentrations were determined spectrophotometrically in 6 M guanidinium chloride at 280 nm using calculated absorption coefficients (Gill and von Hippel, 1989).

Domain	Calculated mass	Calculated mass (-Met)	Measured mass
β-spectrin PH domain	12268.86	12137.67	12266.0
Btk PH domain	20084.05	19952.86	19952.9
YAP65 WW domain	6372.10	–	6372.4

Table 3.1: Results of the mass-analysis of the studied domains.

3.1.1 β-spectrin PH domain

Part of the mouse general form β-spectrin cDNA (genbank:musspna), corresponding to the PH domain (amino acids 2199 - 2304), was cloned to a non-fusion expression vector pBAT4 and to glutathione-S-transferase (GST) fusion-vector pGAT2 to generate a GST fusion protein for panning experiments with phage display libraries.

The β-spectrin PH domain was purified by a combination of cation exchange, gel filtration and anion exchange chromatographies as described (I). The resulting protein was very pure and could be stably stored at +4°C for months in 10 mM sodium phosphate pH 6.5. The GST fusion of β-spectrin PH domain was purified using glutathione-Sepharose affinity chromatography as described before and used immediately (Smith and Johnson, 1988).

3.1.2 Btk PH domain

The human Btk (swissprot:btk_human) PH domain-Btk motif was amplified by PCR from cDNA supplied by Prof. C. I. Edward Smith (Karolinska Institute, Stockholm). Initial attempts to express the canonical PH domain of Btk alone resulted in an insoluble protein which resisted all refolding attempts. Only after the adjacent Btk motif was added to the expression construct, a stable, soluble protein could be produced. The solubility of different Btk expression constructs is shown in figure 3.1.

The amino-terminal 170 amino acids of Btk contain altogether 7 arginines (with two pairs), which are encoded by AGG or AGA, corresponding to 4% of all the amino acids. The Arg-tRNA for these codons is expressed at very low levels in *E.coli* resulting in highly reduced expression of proteins with high AGG / AGA codon content (Emilsson *et al.*, 1993; Brinkmann *et al.*, 1989). In addition to greatly lowering the expression of a protein, these codons can get supressed by lysyl-tRNAs in *E.coli* leading to heterogeneity of the protein sample (Calderone *et al.*, 1996). These problems can be solved by overexpressing this Arg-tRNA from a low-copy-number plasmid (Brinkmann *et al.*, 1989).

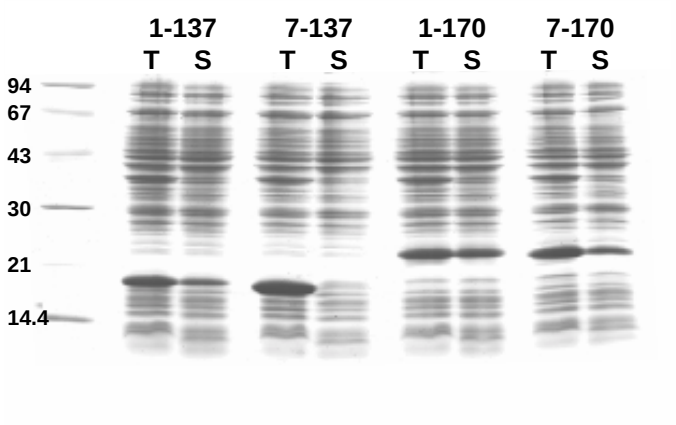


Figure 3.1 Solubility of the Btk PH domain - Btk motif expression constructs. Four different constructs containing either the PH domain alone of both PH domain and Btk motif were created and their solubility in *E.coli* was studied after expression at 15°C. The span of the constructs is shown on top of the gel, the two lanes for each construct represent the total cell lysate (T) and the soluble fractions (S) of the lysate.

This was done in the case of Btk PH domain, as no protein expression could be detected otherwise. The yield of purified protein expressed in this manner varied between 10 and 15 mg per litre of *E.coli* culture.

The protein was purified in three steps using cation exchange and gel filtration chromatographies (III). The concentrated protein was stored frozen at -80°C in 10 mM Tris, 100 mM NaCl, 2 mM DTT pH 8.0 in small aliquots.

3.1.3 WW domain

The WW was cloned from cDNA of human YAP65 (swissprot:ya65_human) supplied by Prof. Marius Sudol (NYU, Mount Sinai School of Medicine, New York). Expression of this small domain was possible only when fused to GST. Several expression constructs were created to define the minimal domain boundaries. The different constructs and their behaviour during the purification and the NMR experiments are shown in table 3.2. The construct 2, corresponding to residues 164-217 of the human YAP65, was used in our studies, as it was the smallest non-degraded and folded protein. It could be expressed in soluble form at 37°C.

The WW domains were purified using a combination of affinity chromatographies and gel filtration. The soluble fraction of the *E.coli* cell lysate was loaded into glutathione-Sepharose column and the column was washed until no more protein was eluting. The GST-WW fusion protein was eluted with 10 mM reduced glutathione in 100 mM Tris pH 8.5, 2 units of bovine thrombin per

1	GSM DDVPLPAGW...WQDPRKAMLSQ	degraded
2	GSM SFEIPDDVPLPAGW...WQDPRKAMLSQMNVTAPTS	folded
3	GS PAGW...WQDPRKAMLSQMNVT	degraded
4	GS VPLPAGW...WQDPRKAMLSQMNVT	degraded
5	GS VPLPAGW...WQDPRKAMLSQMNVTAPTS	unfolded

Table 3.2 WW domain expression constructs. The protein sequences correspond to the final, full-length products after thrombin cleavage from the GST-fusion. Two to three amino acids from N-terminus (GS or GSM) that are separated from the rest of the sequence by a space are derived from the expression vector and are not part of the human YAP65. The sequence between the two conserved tryptophans has been deleted for clarity.

1 mg of fusion protein was added and the sample incubated overnight in room temperature. The digested fusion-protein was loaded to Ni^{2+}-NTA column to which the GST (carrying an N-terminal His_6-tag) bound. The released WW domain was collected from the flow-through, concentrated and loaded into Superdex 30 gel filtration column. Highly purified domain eluted as a single peak at ca. 80 ml. The sample was concentrated for NMR to 1mM and D_2O was added to final concentration of 10 % v / v.

For heteronuclear NMR experiment both C^{13} / N^{15} and C^{13} / N^{15} / H^2 labelled samples were prepared. The *E.coli* were grown in M9 minimal media supplemented with 1 g / l of each sugar-mix, peptone and hydrolysate of labelled *Chaenopodium rubrum* cells and 10 g / l of $N^{15}H_4Cl$ (Ashurst and Oschkinat, unpublished). For the H^2 labelling the growth media was prepared in heavy water (D_2O). Otherwise the expression and purification were as with the unlabelled sample.

3.2 Panning of a phage display library

In an attempt to identify protein ligands for PH domains, the β-spectrin PH domain was used as a receptor for a phagemid library displaying random nonamer peptides on the pVIII coat protein of a filamentous phage M13 (Felici *et al.*, 1991). The phage panning was done as described by Smith and Scott (1993) with some modifications. The main differences were the amplification of the library, for which we needed to use a helper phage, and the immobilisation technique for the receptor.

The PH domain was immobilised as GST fusion on glutathione-Sepharose beads and incubated with the peptide library. Unbound phages were removed by multiple washes with TBST buffer (20 mM Tris, 150 mM NaCl, 0.1 % Tween-

20, pH 7.4). The bound phages were eluted with 20 mM glycine-HCl pH 2.5, neutralised with buffer and used to infect *E.coli* cells (strain JM101). The infected cells were plated on ampicillin plates and grown overnight at 37°C.

The colonies were scraped from the plates the next day, grown for 2 hours in liquid culture and infected with a helper phage M13(K07) to package DNA into phage particles. Infected cultures were grown overnight and produced phages were purified from the media the next morning. This amplified eluate was used as a starting material for the next round of panning. This procedure was repeated three times more lowering the amount of fusion protein in each round. In the end, 54 independent clones from the last round of panning were sequenced by single-stranded dideoxy sequencing.

3.3 CD spectroscopy

The binding of β-spectrin PH domain and inositol phosphates was studied using near-UV circular dichroism (CD) spectroscopy as described earlier (I). In brief, the near UV CD-spectrum of the PH domain was recorded at various concentrations of inositol phosphate ligands and ligand-induced changes in the spectra were used to estimate the binding constants.

3.4 Crystallography

3.4.1 Crystallisation of PH domains

The crystallisation of the PH domain was done using hanging drop vapour diffusion method. The protein solution was mixed with reservoir solution in 1:1 ratio and incubated at 20°C.

The initial screening was done using an incomplete factorial screen of 48 different conditions developed by J. Zeelen at EMBL and using commercial screens from Hampton Research. Promising conditions were then refined until the best possible crystals were obtained.

The β-spectrin PH domain was crystallised from mother liquor containing 10-15 % (w / v) PEG-6000, 30 % ethanol and 100 mM sodium acetate pH 4.8. Initial protein concentration was 20 mg / ml in 10 mM sodium phosphate pH 6.5. Ins(1,4,5)P_3 was added in 3-fold molar excess before the crystallisation drops were prepared. Needle-like crystals, which normally developed in 2-3 days, are shown in figure 3.2A.

The Btk PH domain - Btk motif was crystallised from 32.5 % PEG-3350, 500 mM NaCl, 200 mM $MgCl_2$ and 100 mM Tris-HCl pH 8.5. Protein concentration varied between 6 and 8 mg / ml. The wild-type domain and all the soluble mutants were used in crystallisation trials. The wild-type protein gave only very small "balls" of thin needles, shown in figure 3.2B, which were unsuitable for X-ray diffraction. Mutant R28C, however, gave high quality plate-like crystals which diffracted beyond 1.5 Å in a synchrotron X-ray source. These crystals are shown in figure 3.2C.

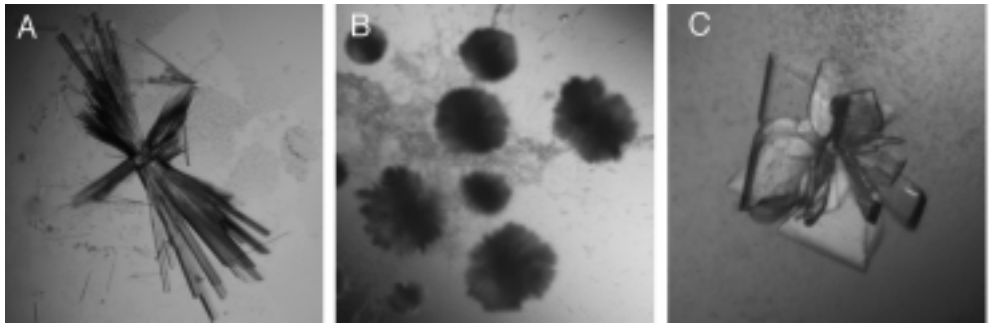


Figure 3.2 PH domain crystals. In A the crystals of β-spectrinPH domain are shown. The needles had maximum thickness of 0.2 mm and length of up to 3 mm. In panel B are the "needle-balls" of the wild-type Btk PH domain- Btk motif. In C are the crystals obtained with mutant R28C. The individual plates were typically ca. 0.5 mm in two dimensions and ca. 50 μm thick.

3.4.2 Data collection and processing

The diffraction data collection and processing was done as described earlier (I and III). All data was collected from frozen crystals at 100–110 K to minimise the radiation damage to the crystals. For the β-spectrin PH domain 10 % glycerol was used for cryoprotection. The Btk crystals did not need a cryoprotectant due to high concentration of PEG-3350 in the mother liquor. The synchrotron dataset of Btk PH domain was collected in two phases. The second phase was needed to secure reliable low-resolution information, as these reflections were saturated during the longer exposures of the high resolution data collection. The high and low resolution data were then merged together omitting all the saturated reflections. The data collection statistics are summarized in table 3.3.

For the heavy metal derivative searches the crystals of Btk PH domain - Btk motif were soaked in original mother liquor into which heavy atom compounds had been dissolved, typically at 1 mM concentration. After variable incubation periods a dataset was collected from a soaked crystal and processed

Dataset	Resol.	λ	R_{symm}	Unique refl.	Completeness
β-spectrin 1	2.8 Å	1.514 Å	4.3 %	3261	98.8 %
β-spectrin 2	2.0 Å	1.514 Å	4.7 %	7830	89.6 %
Btk native 1	2.6 Å	1.514 Å	6.0 %	9817	99.8 %
Btk derivative	2.6 Å	1.514 Å	6.1 %	9319	97.2 %
Btk native 2	1.6 Å	0.882 Å	5.4 %	42840	99.9 %

Table 3.3: The data collection statistics.

as with the native crystals. The different datasets of Btk crystals showed variations in the cell dimensions and were clearly non-isomorphous. To overcome these problems, the datasets used for the phasing were collected from a single crystal. After collecting a native dataset the crystal was thawed and soaked for 24 hours in the heavy-atom solution, after which a derivative dataset was collected. The crystals did not suffer from repeated freezing and thawing and the resulting datasets were isomorphous.

3.4.3 Structure determination and refinement

The phase problem for the β-spectrin PH domain was solved by molecular replacement using the previously determined NMR structure of the same molecule as a search model using the programs AMoRe and X-PLOR (Navaza, 1994; Brünger, 1988). The correct solution was found using artificial temperature (B) factors, which were derived from the root-mean-square deviations of atom positions in the ensemble of the solved NMR structures (for a detailed description of the method, see Wilmanns and Nilges, 1996).

The Btk PH domain - Btk motif structure was solved by so-called heavy-atom method. The screening of potential derivatives and phasing was done using program package Phases-95 (Furey and Swaminathan, 1996). After extensive searches, a single heavy atom derivative, trimethyl lead (II) acetate, was found. This proved to have very good phasing power both for isomorphous and anomalous data and was sufficient to solve the phase ambiguity. Phases were very good to the resolution limit of the in-house datasets, 2.6 Å.

The phases were improved by solvent flattening and two-fold non-crystallographic symmetry averaging in program DM (Cowtan, 1994). Later when high-resolution synchrotron data was available, the experimental phases were extended to 1.8 Å.

The model-building was done using the program O (Jones *et al.*, 1991). The models were refined using programs X-PLOR (Brünger, 1988), TNT (Tronrud *et al.*, 1987) and ARP (Lamzin and Wilson, 1993) as described (I and III). Refine-

ment statistics for β-spectrin and Btk PH domain are shown in table 3.4. Coordinates of the β-spectrin and Btk PH domains have been deposited to Protein Data Bank (PDB) under accession numbers 1BTN and 1BTK, respectively.

Protein	**Resol.**	$\mathbf{R}_{cryst}$	$\mathbf{R}_{free}$
β-spectrin PH domain	8.0-2.0 Å	20.5 %	28.6 %
Btk PH domain	40.0-1.6 Å	23.1 %	28.0 %

Table 3.4: Refinement statistics.

The stereochemical quality of the models was analysed using Procheck and WHAT-Check programs (Laskowski *et al.*, 1993; Vriend, 1990).

3.5 NMR and modelling

The structure of the YAP65 WW domain was determined by heteronuclear NMR as described (II). The model of the WW domain - proline-peptide complex was calculated using a simulated anneling protocol with X-PLOR (Brünger, 1988).

4 RESULTS

4.1 The β-spectrin PH domain (I and unpublished)

4.1.1 The ligands of the β-spectrin PH domain

One of the original aims in our study was to identify protein ligands for PH domains. So far no ligands were known for the β-spectrin (or any other) PH domain and as an attempt to find one, we used this domain as a bait for panning a random peptide library. Similar approach had successfully been used for other signalling modules, *e.g.* for the SH3 domains (Sparks *et al.*, 1994; Cheandle *et al.*, 1994).

After several rounds of binding and amplification of random nona-peptide phage display library, 54 individual clones were sequenced. Of these, 20 either carried no insert, contained in-frame stop codons or were too short. The remaining peptide sequences were analysed for common motifs and divided into four groups shown in table 4.1.

group 1		group 2		group 3	group 4
MRRRQPKNL	5 $\star$	DRRAKHLPL	2 $\star$	RRAANPQAR	QATSTHHAR
RRVQPKALR	2 $\star$	NKKVKGFAW		RKCCSRPRG	ARHHSSDQP
AQPRSWDFP		HNRRLRDSL		ARKAASYPG	RNFGVMAVS
AQPRSSHFP		KRIKDRVGQ		AKRKSSLPR	SICPNKTPP
QPKKVKAPL	2 $\star$	NKKARYGPM		NRRKFSSES	YALKWKPQS
				SRRGVPANP	KESTCNNPP
				ANRKLYYFS	SAKSVSATA
				HRKKVTEIQ	SNFSQIRAL

Table 4.1 The peptides obtained from a nona-peptide library using β-spectrin PH domain as a receptor. The number after the sequence indicates how many times the same peptide was found in the analysed clones.

The peptides in the group 1 share a common motif QPK/R. For the second group a consensus motif KRAK can be derived. Group 3 contains peptides with consecutive positive residues and group 4 all the rest. As can be seen, the sequences are only distantly related even within these groups and no clear consensus pattern can be derived. Peptides marked with a star ($\star$) in table 4.1 were synthesised and binding to β-spectrin PH domain was studied by NMR, CD and fluorescence spectroscopy, but none of the peptides bound to this domain even at millimolar concentrations (data not shown). The reasons for this are unclear. In a control experiment GST was used as a receptor, and the clones sequenced in the end showed similarity neither to each other nor to the peptides obtained with the β-spectrin PH domain. The peptides are displayed in tens or even hundreads of copies on the surface of the phage and it is possible that multiple weak (and unspecific) interaction can lead to amplification of false-positive clones. As this attempt to find a protein ligand for the β-spectrin PH domain proved unsuccessful, we turned to inositol phosphates, which were shown to bind to the pleckstrin PH domain (Harlan *et al.*, 1994).

The ability of β-spectrin PH domain to interact with inositol phosphates was studied using near-UV CD and NMR spectroscopies. Both the CD and NMR spectra showed changes upon mixing the protein with inositol phosphates which carry phosphates in 4- and 5-positions of the inositol ring. These include Ins(1,4,5)P_3, Ins(1,3,4,5)P_4 and g-PIns(4,5)P_2. The other tested compounds are listed in table 4.2. The results suggest that the β-spectrin PH domain interacts with 4- and 5-phosphates, and that the binding is unaffected by the 3-phosphate. The fact that g-PIns(4,5)P_2 bound with similar affinity as Ins(1,4,5)P_3 indicates that the 1-phosphate can be derivatised without affecting the binding. This is noteworthy, as in phosphatidylinositol lipids the phosphatidyl moiety is attached to the 1-hydroxyl of the inositol. The dissociation constant for the β-spectrin PH domain g-PIns(4,5)P_2 complex was estimated to be ca. 40 μM.

Ligand	**Binding**	**No binding**
D-*myo*-inositol		X
Ins(1)P		X
Ins(1,4)P_2		X
Ins(1,4,5)P_3	X	
Ins(1,3,4)P_3		X
g-PIns(4,5)P_2	X	
Ins(1,3,4,5)P_4	X	

Table 4.2 Binding of different phosphoinositides to β-spectrin PH domain. See page 10 for abbreviations.

The experiments with Ins(1,4,5)P_3 and g-PIns(4,5)P_2 induced consistent changes in the NMR spectra of the PH domain and the binding site could be localised to loops 1-2 and 5-6. To verify the role of individual residues in PH domain - ligand interactions, site-directed mutagenesis was used to replace the positively charged amino acids with glutamines in the vicinity of the binding site. The capacity of the mutant domains to bind g-PIns(4,5)P_2 was studied by CD spectroscopy. As expected, the mutants K8Q, R21Q and K71Q were unable to bind the ligand. Mutations R7Q, K17Q or K72Q had no effect on the ligand binding, which is in agreement with the minor shifts these residues showed in the NMR spectra.

4.1.2 The structure

The β-spectrin PH domain was crystallised in complex with Ins(1,4,5)P_3. The crystals diffracted to 2.0 Å in an in-house X-ray source and belonged to space-group $P42_12$ with unit cell parameters a=b=68.9 Å, c=50.8 Å, $\alpha=\beta=\gamma=90°$.

The structure of the β-spectrin PH domain in complex with Ins(1,4,5)P_3 shows no large differences compared to the previously described unliganded NMR structure (Macias *et al.*, 1994; Nilges *et al.*, 1997). The largest differences are seen in the loops involved in ligand interactions (figure 4.1).

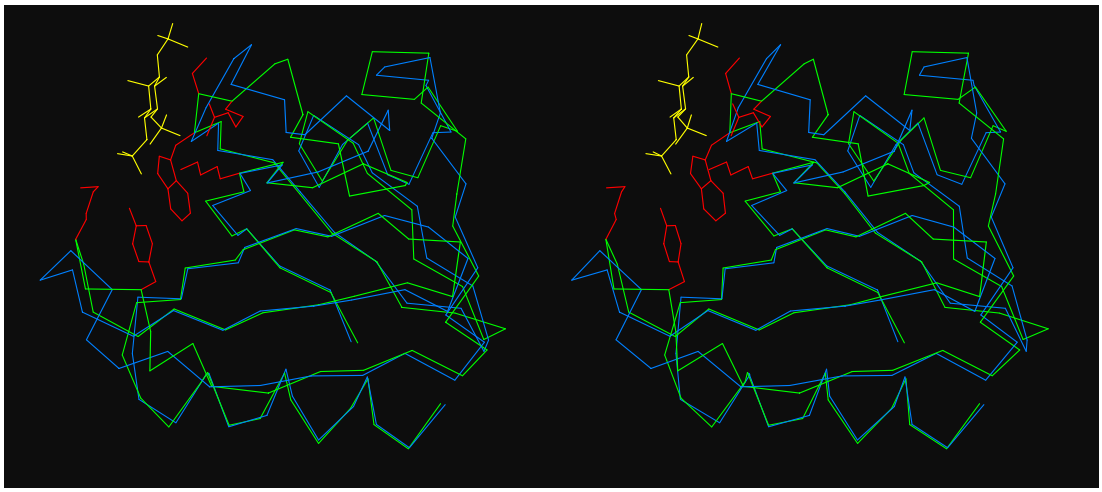


Figure 4.1 Crystal and NMR structures of the β-spectrin PH domain. A stereo view of the superimposed Cα-traces of the crystal structure (green) and the avarage NMR structure (blue) are shown. The ligand binding residues in the crystal structure are shown in red and the ligand in yellow.

The PH domain binds Ins(1,4,5)P_3 between the β-strands 1 and 2 and the loop 5-6. The binding site is a very shallow pocket surrounded by positively

charged residues. The binding site is created by residues K8, R21, S22, W23, Y69 and K71. They interact with Ins(1,4,5)P_3 through the 4- and 5-phosphates. The only contact to 1-phosphate is a hydrogen bond from the γ-hydroxyl group of S22. There are no contacts to the ring or to the free hydroxyl groups of the inositol. The binding site in this crystal structure is in perfect agreement with the ligand specificity studies and localisation of the binding site by NMR spectroscopy and site-directed mutagenesis.

4.2 The Btk PH domain (III)

4.2.1 The structure of the Btk PH domain

The wild-type human Btk PH domain and all the soluble mutants were used for crystallisation attempts.Only the mutant R28C (corresponding to mouse *xid* mutation) yielded crystals suitable for diffraction analysis. Crystals belong to space-group $P2_1$ with two molecules in the asymmetric unit. The unit cell dimensions are a=49.15 Å, b=59.87 Å, c=55.94 Å, α=γ=90°, β=98.2°. The structure was solved using a single heavy metal derivative and is refined at 1.6 Å.

The core of the Btk PH domain is very similar to the previously published PH domains (see figure1.4A for superpositioning of Btk PH domain with the others). Btk contains a long insertion between the β-sheets 5 and 6. The end of this loop forms a one and a half turn α-helix and is part of the interface between the two molecules in the asymmetric unit. The hydrophilic middle part of the loop has weak electron density and is left out of the model in one of the two molecules and is modelled as poly-alanine in the other. In addition to the two protein molecules, 2 Zn^{+2} ions (see next section), 2 Na^{+} ions and 248 waters are included in the model. The two Na^{+} ions are bound between the two protein molecules in the asymmetric unit with octahedral coordination sphere of protein or water oxygen atoms (Nayal and Di Cera, 1996). The non-crystallographic dimer and bound metal ions are shown in figure 4.2.

The structure of the Btk PH domain exposed some errors in the previous sequence alignments. The first β-strand was shifted by one residue which consequently positioned some of the residues in the putative inositol phosphate binding site incorrectly. In analogy to PLCδ1 PH domain the Btk PH domain has a β-bulge in the beginning of the β-strand 5, which is missing in other known structures (β-spectrin, dynamin and pleckstrin). This β-bulge requires a gap to be opened in the multiple sequence alignment in some cases. These corrections, although small, are important and as more structural information accumulates, more accurate database searches can be conducted. A correct sequence alignment is also crucial for successful homology modelling of related domains.

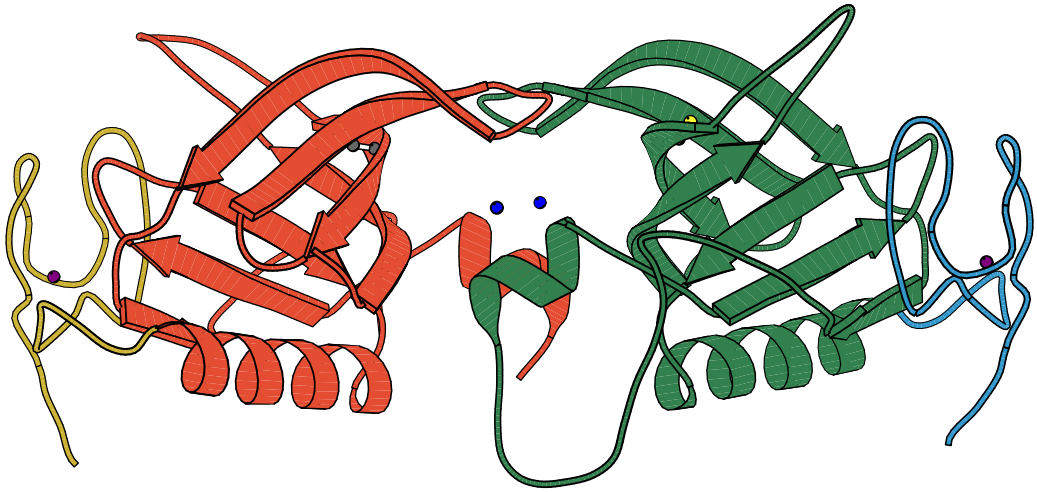


Figure 4.2 The non-crystallographic dimer of PH domain-Btk motif. The PH domains are coloured red and green, whereas the Btk motifs are yellow and blue. The zinc ions are shown in purple and the sodium ions in blue. R28C mutation is shown as ball- and-stick model.

4.2.2 XLA mutations in the Btk PH domain

We wished to study the structural and functional consequences of XLA mutations in the Btk PH domain and created expression constructs to produce many of the mutant domains in *E.coli*. In order to produce the protein in *E.coli*, it was necessary to include the adjacent small domain, so-called Btk motif to the expression constructs.

As predicted, some of the XLA mutations (T33P, V64F and V113D) disrupt the structure of the domain and the corresponding proteins could not be produced in soluble form *E.coli*. The rest of the mutations do not prevent folding and the proteins were produced for structural and functional studies.

The crystal structure allows us to predict the effect of XLA mutations on the Btk PH domain. Mutants T33P, V64F and V113D showed impaired ability to fold in *E.coli* and structural characterisation of these residues supports our biochemical data. V64 and V113 are conserved hydrophobic residues which contribute to the hydrophobic core of the domain and introduction of a bulky phenylalanine or a polar aspartate would be unfavourable. T33 is located in the loop 2-3 and a rigid proline is likely to affect the folding and / or conformation of this part. We have recently expressed also the mutant L11P, which exhibits similar behaviour in *E.coli* and can be classified as a folding mutant (data not shown).

All the other XLA mutations are located in the putative binding site for PtdIns(3,4,5)P_3 and ipppp (for more detailed discussion on the binding sites see section 5.1). Mutations K12R and R28H / C / P affect residues whose analogues in the PLCδ1 PH domain interact directly with Ins(1,4,5)P_3. Mutations S14F and F25S are predicted to have more indirect effect on inositol phosphate binding. Impaired binding of Btk PH domain to an phosphoinositol ligand is anticipated to prevent membrane localisation of the kinase during activation causing a developmental arrest of B-cells and subsequently XLA.

4.2.3 The structure of the Btk motif

In the Tec family of cytoplasmic tyrosine kinases, the PH domain is always followed in sequence by a short motif which is characterised by conserved cysteine and histidine residues (Vihinen *et al.*, 1994). This so-called Btk motif is needed for stability of the Btk PH domain as demonstrated by our failed attempts to express the Btk PH domain alone in *E.coli* (see figure 3.1 on page 30). In the progress of building the initial model of the protein into experimentally derived electron density maps, a metal ion which was bound to the Btk motif was identified . This metal ion was tetrahedrally coordinated by the conserved three cysteines and a histidine, suggesting that it could be a zinc ion. Atomic absorption spectroscopy confirmed that the sample contained 0.94 zinc atoms per protein molecule. The zinc coordination sphere, bond lengths and angles are shown in table 4.3.

	Bond angles in degrees				**Bond lengths to Zn^{+2}**	
	H143 Nδ1	C154 Sγ	C155 Sγ	C165 Sγ	Mol. 1	Mol. 2
C165 Sγ	118.0	100.8	110.9		2.2 Å	2.2 Å
C155 Sγ	95.7	123.0		112.9	2.4 Å	2.4 Å
C154 Sγ	109.6		117.2	97.0	2.4 Å	2.3 Å
H143 Nδ1		114.4	101.4	113.8	2.1 Å	2.1 Å

Table 4.3 Zinc coordination in the Btk motif. The table shows the angles between protein ligands centered on the metal ion. The upper and lower halves of the table show data of the two different molecules in the asymmetric unit. The two last columns list the bond lengths between the protein ligands and the zinc ion.

The Btk motif has a compact, globular core in the middle of which the zinc is bound. The metal ion is completely buried and keeps the structure together. The Btk motif packs against the β-sheets 5-7 of the PH domain and the aromatic side-chain of F146 is inserted into a hydrophobic pocket formed by residues

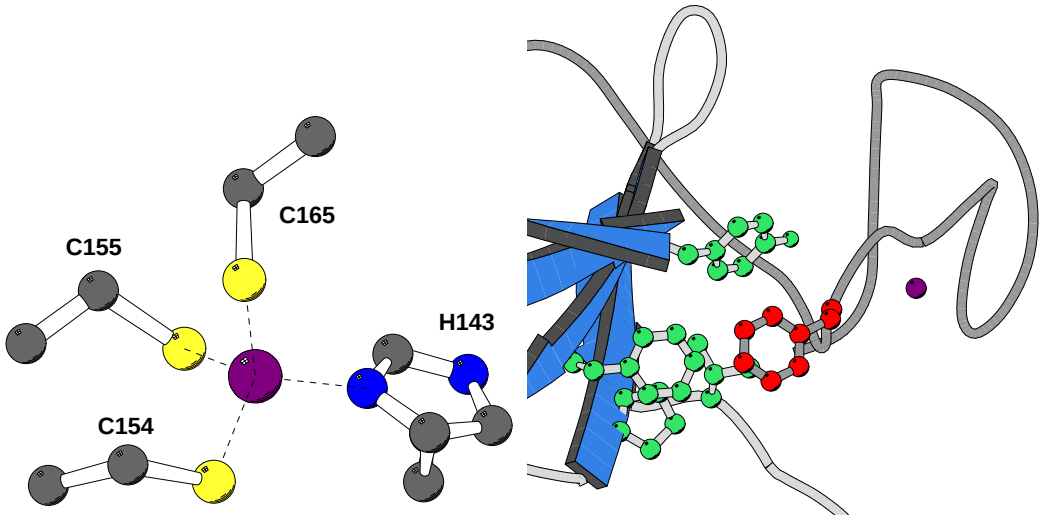


Figure 4.3 The Btk motif. The figure A shows the coordination of the zinc ion by residues H143, C154, C155 and C165. Interface between Btk motif and PH domain is shown B. F146 of the Btk motif (red atoms) and residues forming the hydrophobic pocket in the PH domain (green atoms) are shown as ball-and-stick models.

V67, P101, Y112 and F114 (Figure 4.3B). Additional hydrogen bonding stabilises the domain-domain interaction.

The globular part of the Btk motif is connected to PH domain with a 9-residue linker, which wraps around the edge of the PH domain's β-sheet. The only conserved residue in the linker is N135. It inserts its Nδ atom into a small pocket and donates hydrogen bonds to the main chain carbonyl groups of residues V58 and I61. A multiple sequence alignment of known Btk motifs and the preceding linkers is shown in the figure 4.4.

4.3 The WW domain of YAP65 (II)

4.3.1 The structure

The structure of the WW domain is very simple—a single three-stranded β-sheet (figure 4.5). The structured core of the domain with only 30 residues is smaller than predicted from the sequence alignments. Residues outside this core do not show long- or medium-range nuclear Overhauser effects (NOEs) apart from I7, which binds to a hydrophobic patch on the domain and stabilises the structure. This explains our failed attempts to produce a shorter folded

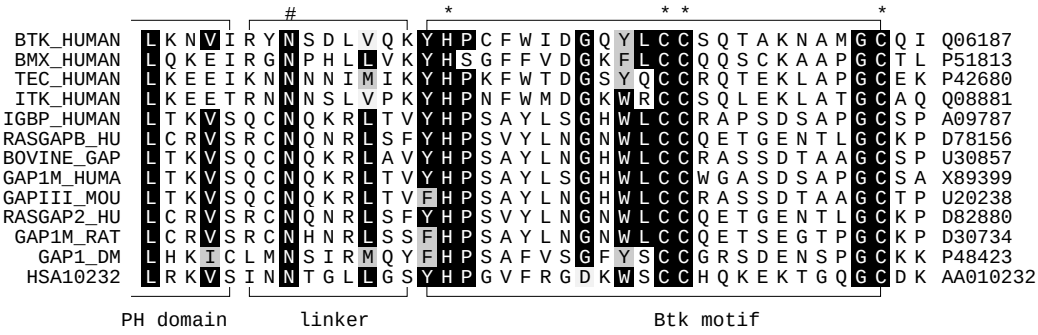


Figure 4.4 Multiple sequence alignment of Btk motifs and and the preceding linker. The database accession numbers of the sequences are shown on the right-hand side of the alignment. Zinc binding residues H143, C154, C155 and C165 are marked on top of the alignment with asterisks.

domain. The position of I7 is not conserved in other WW domains and is considered to be a specific feature of YAP65 WW domain. In analogy to SH3 domains, some of the conserved hydrophobic residues, including one of the two tryptophans, are exposed on the surface of the domain.

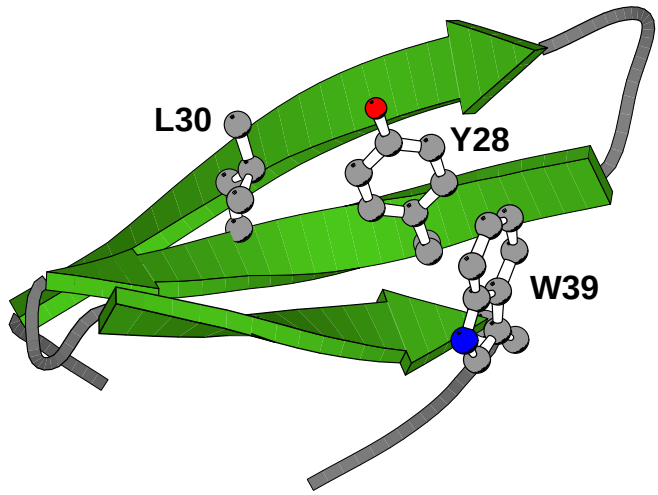


Figure 4.5 Structure of the WW domain from human YAP65. The conserved hydrophobic residues on the surface of the domain are shown as ball-and-stick models

4.3.2 Ligand binding

The structure was solved in complex with a proline-rich peptide GTPPPPYTVG, which is derived from WW domain binding protein 1 (Chen and Sudol, 1995). This peptide binds to YAP65 WW domain with K_D of 47 μM as measured by fluoresence spectroscopy. Binding of the peptide induced chemical shift changes on WW domain (residues Y28, Q35, T36 and W39 showing shifts of > 0.5 p.p.m.) and residues L30, H32 and W39 showed NOEs to the ligand. These NOEs defined the orientation of the peptide and a structure of the WW of the complex was calculated using the limited experimental data available and restraining the peptide into PPII conformation.

In this model the WW binds its ligand on the concave hydrophobic surface. Prolines P4' and P5' (the peptide residues are primed) are in contact with the conserved W38 and Y7' is held in a pocket formed by L30 and H32. The hydroxyl group of Y7' is in close vicinity of H32 and Q35 and could form hydrogen bonds with either of them.

To confirm the location of the binding site, mutations L30K, H32A and Q35A of WW domain were created. H32A and Q35A mutants had completely lost or greatly reduced ability to bind the peptide. Strongest effects were seen with a short peptide SPPPYTV, which bound to the wild-type protein with a similar affinity as the longer peptide. Mutant L30K did not bind the short peptide, but bound the longer peptide with slightly higher affinity than wild-type protein did, for reasons not clearly understood.

Alanine scanning mutagenesis of the PPPPY motif in the shorter peptide was used study the minimal determinants for the WW domain-peptide interaction. The first and the last of the prolines could be mutated without affecting the binding, whereas the two middle prolines and the tyrosine were strictly required for the binding.

5 DISCUSSION

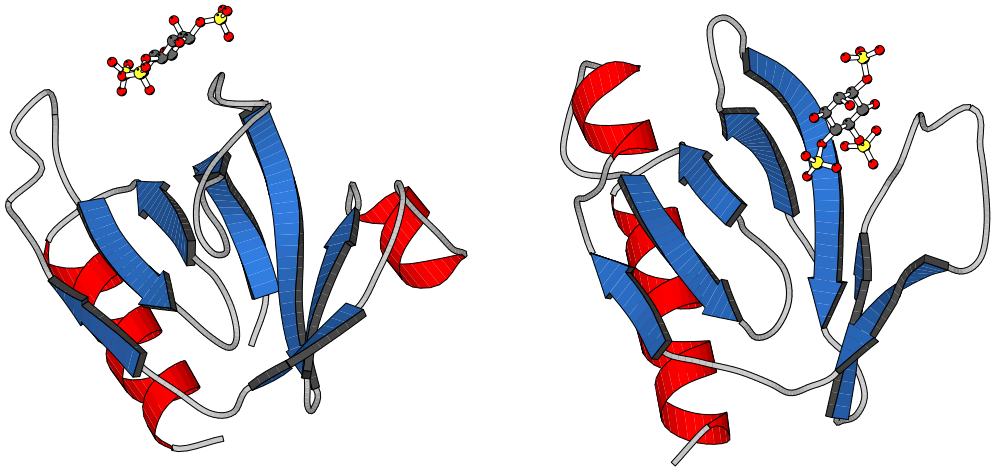
5.1 Inositol phosphate binding site in PH domains

Soon after we had determined the structure of β-spectrin PH domain-Ins(1,4,5)P_3 complex, Ferguson *et al.* (1995) published the structure of PLCδ1 PH domain - Ins(1,4,5)P_3 complex. The PLCδ1 PH domain bound the ligand with polar interactions to 4- and 5-phosphates exposing the 1-phosphate to the solvent. This is very similar mode of binding as was observed for the β-spectrin PH domain. The unexpected finding was the different location of the binding sites in these two domains. The β-spectrin PH domain binds the ligand between the β-strands 1-2 and 5-6 where as in the PLCδ1 domain the binding site is located on the other side of β-strands 1 and 2 and additional residues involved in binding reside in the loop 3-4. Although β-strands 1 and 2 are involved in both binding sites, the key residues for the ligand interactions are different as shown in the alignment in figure 5.1. This part of PH domains is generally rich in positive residues and many PH domains show conservation of the major ligand binding residues.

The residues in loops 3-4 and 5-6 are very variable and it is more difficult to predict how the domains would bind inositol phosphates, if they bind at all. The structures of these loops also differ greatly, further complicating the predictions.

The structure of the Btk PH domain shows strong similarities with the ligand binding residues of the PLCδ1 PH domain in the β-strands 1 and 2. Many of XLA mutations in the Btk PH domain which are located around the putative inositol phosphate binding site are known to impair the domains ability to bind Ins(1,3,4,5)P_4 (Fukuda *et al.*, 1996). The differences in the binding sites of PLCδ1 and Btk PH domains binding sites most likely reflect their preference for differently phosphorylated ligands (PtdIns(4,5)P_2 vs. PtdIns(3,4,5)P_3). The mutation R28C in our structure might have also affected the structure of the binding site in the vicinity of the mutated residue and only structural characterisation of the

```
     PLC-δ1  23  L K G S Q L L K V K S S . . . . . . . . . S W R R E R F Y K L Q
β-spectrin    1  . M E G F L N R K H E W E A H N K K A S S R S W H N V Y C V I N
   dynamin   13  I R K G W L T I N N I G I M . . . . . K G G S K E Y W F V L T
pleckstrin    6  I R E G Y L V K K G S V F . . . . . . . N T W K P M W V V L L
     SOS-1  462  I M E G T L T R V G . . . . . . . . . . . A K H E R H I F L F
                       β1                                    β2
```



B β-spectrin **C** PLCδ1

Figure 5.1 Inositol phosphate binding sites in PH domains from β-spectrin and PLCδ1. The two domains are shown in same the orientation and the ligands are displayed as ball-and-stick models.A multiple sequence alignment of the β-strands 1 and 2 is shown on the top. The residues interacting with the ligands are highlighted. The residues highlighted in pleckstrin, dynamin and Sos1 sequences are those that have shown changes in NMR spectra upon addition of a ligand. The residues on gray background are the most conserved hydrophobic positions in each β-strand.

wild type Btk PH domain can elucidate the details of the ligand binding.

5.2 The ligands of the PH domains

Although many PH domains can bind inositol phosphates, it is still to be seen, whether these are the physiological ligands for all members of this family. Different forms of phosphoinositides are formed in cells as a result of activation by extra-cellular signals. The best studied system is the release of Ins(1,4,5)P_3

by phospholipase C in response to activation of trimeric G-proteins (Berridge, 1993). The soluble Ins(1,4,5)P_3 molecules bind to specific receptors in the endoplasmic reticulum and trigger the release of Ca^{2+} into the cytoplasm. The key enzymes in this cascade, the phospholipases, carry an N-terminal PH domain. In this case it is very easy to understand the role of PH domain, which anchors PLCs to the membrane close to the pool of hydrolysable substrates. Whether these soluble inositol phosphates are also ligands for PH domains of either PLCs themselves allowing feed-back regulation of their activity or other proteins *in vivo*, remains to be seen.

Many other PH domain carrying proteins are not so clearly or directly involved in the biology of phosphoinositides. Most of these proteins are, however, closely associated with membranes in cells. β-spectrin is a major component of the cortical cytoskeleton and is believed to be involved in organisation of the plasma membrane (Bennet and Gilligan, 1993). The C-terminal part of the general form of β-spectrin carrying the PH domain binds to alkaline washed bovine brain membranes (Davis and Bennett, 1994; Wang and Shaw, 1995). The erythrocyte isoform β-spectrin lacks PH domain as result of differential splicing and the corresponding region does not interact with membranes. The β-spectrin PH domain is targeted to plasma membrane also *in vivo* (Wang *et al.*, 1996).

Many of the kinases carrying PH domains (for example β-ARK, Btk, Akt) associate with membranes following activation and are found in the insoluble fraction of the cell lysate, but the molecular mechanism of such translocation is not well established. Assuming that these PH domains do bind phosphatidylinositol lipids *in vivo*, the availability of the lipid in the membrane could affect the localisation of these proteins. Phosphoinositide-3-OH kinase (PI3K) is an enzyme that catalyses the synthesis of various 3-phosphorylated phosphoinositides and it is activated by various extracellular stimuli (Vanhaesebroeck *et al.*, 1996). Btk and Akt PH domains show highest affinity towards 3-phosphorylated phosphoinositides and are belived be become membrane associated in responce to PI3K activity (Franke *et al.*, 1997; Klippel *et al.*, 1997; Rameh *et al.*, 1997; Salim *et al.*, 1996). Regulated synthesis and rapid turnover of these lipids could be used to regulate membrane association of PH domain carrying proteins (figure 5.2. For reviews see Toker and Cantley, 1997; Carpenter and Cantley, 1996; Franke *et al.*, 1997).

In this scenario, the role of PH domains is very much the same as with many other signalling domains - they bring the proteins to the site of action. A membrane bound protein is restricted in its movements to two dimensions, a limitation which would greatly facilitate interactions with other membrane bound proteins. The co-operative activation of β-ARK by $\beta\gamma$-subunits and PtdIns(4,5)P_2 is a good example of such a coordinated action of lipid and pro-

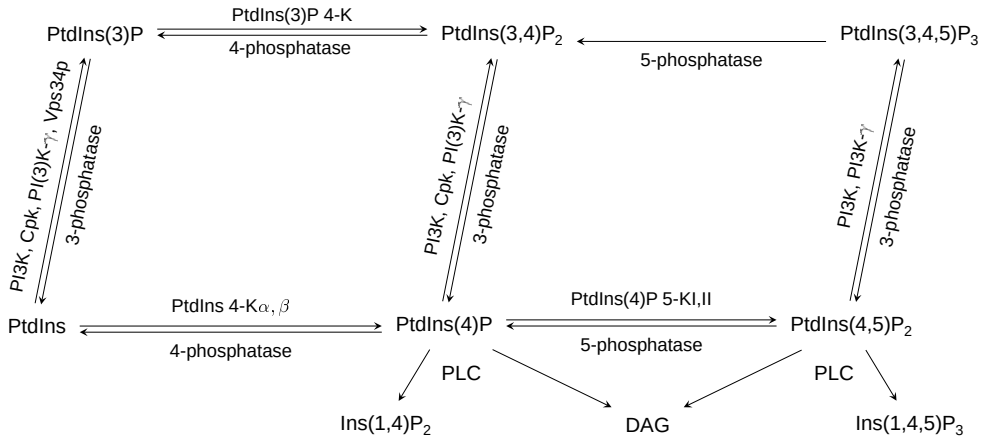


Figure 5.2 Phosphoinositide metabolism in signal transduction. Figure adapted from Toker and Cantley (1997), which should also be referred to for more details.

tein components in the membrane (Pitcher *et al.*, 1995).

Covalent lipid modification is a common way to enhance the membrane localisation of signalling proteins (Casey, 1995; Bhatnagar and Gordon, 1997). The Ras proteins are isoprenylated and often also palmitoylated in their C-terminus and many members of Src PTK family are myristoylated in the N-terminus, to give some examples. In some cases lipid modification and PH domain seem to carry an analogous function. β-ARK is closely related to rhodopsin kinase (RK), and both can desensitise rhodopsin by phosphorylation. The two proteins differ most in their C-termini, as the PH domain is missing and is replaced by an isoprenylation signal in RK. The functional analogy between the PH domain and lipid modification is demonstrated by restoration of function in C-terminally deleted β-ARK by isoprenylation (Koch *et al.*, 1993). Another such example is *lfc*, a member of DH domain containing proteins, in which transformation capacity can be restored by substituting PH domain for isoprenylation signal of Ras, suggesting a membrane anchoring function for this class of PH domains (Whitehead *et al.*, 1995). In Tiam1, however, the PH domain adjacent to DH domain is not required for membrane association, but instead an additional PH domain in the N-terminus localises the protein to membrane and can be substituted with myristoylation signal of c-Src (Michiels *et al.*, 1997).

Src and Btk are another example of homologous proteins which have alternatively either lipid modification or a PH domain, in this case in their very N-termini. Both of these enzymes are translocated to membrane upon activation, but experiments similar to those with β-ARK and RK have not yet been done.

Lipid modifications and PH domains seem to share the same principal func-

tion of membrane targeting, but allow different ways to control this localisation of proteins—both spatially and temporally. Cellular membranes have areas enriched in specific lipids (sphingolipids and cholesterol, in particular), which can serve as localisation sites for various proteins including doubly lipidated PTKs (Simons and Ikonen, 1997). These so-called membrane rafts are also enriched in PtdIns(4,5)P_2 and some of the phosphoinositide-modifing enzymes providing a possible docking site also for PH domains (Hope and Pike, 1996).

The role of PH domain as a regulated membrane targeting module is well exhibited in Btk. As some point mutations in the PH domain can cause XLA by abolishing binding to inositol phosphates and preventing membrane recruitment of Btk, others can enhance the membrane association of the kinase and constitutively activate it (Li *et al.*, 1995).

5.3 Structure and function of the Btk motif

The Btk motif is a new zinc-binding domain with no clear homologies in the Protein Data Bank. The zinc coordination by three cysteines and a histidine is found also in diacyl glycerol/ phorbolester binding domain (DAG/PE), LIM domains and some zinc fingers, but none of them coordinate the metal ion with two adjacent cysteines. These other domains form a tight turn similar to that in the Btk motif at the zinc site, but they have typically two residues between the coordinating cysteines (Pérez-Alvarado *et al.*, 1994; Zhang *et al.*, 1995a).

The function of the Btk motif is unknown. It is always found adjacent to a PH domain and as our structure shows, these two domains are closely associated. The sequence conservation of the Btk motif is restricted to two glycines in addition to the zinc coordinating residues. The role of the glycines is unclear. They do not have phi-psi angles in the disallowed region of the Ramachandran plot, which could structurally explain their conservation. A possibility is that these flexible amino acids are required during the folding of this small motif. Especially the second glycine (G164 in Btk) could have such a role as it is positioned just before the last zinc liganding cysteine. The possibility that these glycines are functionally important residues cannot be ruled out either. The first of these glycines (G150 of Btk) is on the “top” of the Btk motif, when viewed as in figure 4.2, lying next to the side-chain of W147. This residue is partly mobile and has its sidechain partly solvent exposed.

A recently discovered Btk associated protein BAP-135 binds specifically to Btk, more precisely to amino-terminal fragment of Btk containing PH domain and Btk motif (Yang and Desiderio, 1997). BAP-135 is not similar to any other known proteins and it is not known which part of it interacts with Btk.

5.4 WW domain

The function of WW domain as an adaptor module with functional similarity with SH3 domains is getting established. The recognition motifs for WW and SH3 domains are very similar and in many cases these domains can compete for the same ligands (Bedford *et al.*, 1997; Sudol, 1996b). WW domains are very often present in multiple copies in proteins; Nedd4 protein carries four WW domains in a row. Binding of several domains to a long proline-rich sequence in the target protein would increase the affinity of the protein-protein interaction and allow more specific targeting of these proteins.

In Pin1 protein the WW domain is associated with a peptidyl-prolyl isomerase domain, which catalyses *cis-trans* isomerisation of proline residues (Lu *et al.*, 1996). A likely function of WW domain is to attach the enzyme to a proline-rich substrate enhancing the catalytic activity in a very similar manner as the PH domains do in phospholipase Cs.

Involvement of WW domain containing proteins in several human diseases like Alzheimer's disease (FE65), Liddle's syndrome (Nedd4) and muscular dystrophies (dystrophin) has been implied (Sudol, 1996a). Only in the case of Liddle's syndrome evidence for the possible involvement of WW domain in the development of the disease has been presented and further studies are required to establish the role of WW domains in other diseases.

5.5 Conclusions

In this study the structures and ligand interactions of β-spectrin and Btk PH domains as well as the YAP65 WW domain have been characterised. The results bring insight to the function of these domains as molecular adaptors.

The β-spectrin PH domain binds to inositol phosphates with 40 μM affinity recognising specifically the 4- and 5-phosphates of the ligands. The binding site for the Ins(1,4,5)P_3 is located between β-strands 1 and 2 and the loop 5-6. The ligand is coordinated by polar residues in the middle of the positively charged end of the domain. Orientation of Ins(1,4,5)P_3 on the β-spectrin PH domain has lead us to propose a model, where PH domain acts as a membrane targeting module recognising a membrane bound phospholipid, PtdIns(4,5)P_2, specifically. The membrane interaction is strengthened by additional, unspecific interactions between the positively charged surface of the PH domain and the negatively charged lipid headgroups. Spectrin is a large cytoskeletal protein located just under the plasma membrane and the PH domain is believed to serve as a direct membrane anchor of the protein. Further studies are needed to assess the biological role of these interactions.

The structure of the Btk PH domain and Btk motif has allowed us to predict the molecular mechanism of XLA in the cases where the disease causing mutations are located in the PH domain. Some of the XLA mutations affect the fold of the domain preventing the assembly of the native three-dimensional structure. Other mutations affect surface exposed residues which are localised around the predicted binding site for inositol phosphates. Comparison of the Btk PH domain with the Ins(1,4,5)P_3 binding site in the PLCδ1 PH domain has enabled us to predict the role each of the mutated residues in XLA. The structure of the Btk motif has revealed a novel zinc binding fold. A zinc ion is bound inside the Btk motif to three conserved cysteines and a histidine stabilising the structure. The Btk motif is tightly associated with the PH domain, and they are predicted to form a functional entity. This structure provides a rational basis for further studies on the role of PH domain and Btk motif in XLA, in the maturation of B-cells and signal transduction in general.

The structure of the WW domain from human YAP65 has been solved in complex with a proline-rich peptide by NMR. The structure of the WW domain is a single three-stranded β-sheet. Some of the conserved, hydrophobic residues of the WW domain are exposed on the surface of the domain forming a binding site for a proline-rich peptide. The WW domain is predicted to bind its ligand in a PPII helical conformation and model of the complex based on the available experimental data is presented. The binding site has been confirmed by mutagenesis of both the WW domain and the ligand peptide.

Acknowledgements

This work has been carried out at European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. I have enjoyed the international environment of EMBL and am grateful for the excellent facilities that have made my studies possible.

I owe my deepest thanks to Dr. Matti Saraste, my supervisor, for all the support he has given me. I have had great opportunities to learn new techniques and at the same time had very exiting projects to work with.

I grateful to Drs. Kari Alitalo and Juha Rouvinen, for reviewing this thesis and for their constructive comments on it.

I thank Maria Macias, Elena Baraldi, Johan Schultz and Hartmut Oschkinat for the collaboration in the WW and / or PH domain projects. Toby Gibson's PH domain alignments and help are highly appreciated.

I wish to thank the past and present members of our group at EMBL for the good atmosphere in and out of the laboratory: Sonia Bañuelos, Elena Baraldi, Jose Castresana, Kristina Djinović-Carugo, Mathias Gautel, Ulrich Gohlke, Janneke Hendriks, Pekka Lappalainen, Andrea Musacchio, Tor-Björn Olausson, Jaime Pascual, Gilles Travé, Anthony Warne, Matthias Wilmanns, Paul Young and the numerous visitors. Help from Kristina and Matthias in crystallography has been invaluable.

I wish to thank two fellow PhD student, Drs. Morten Grøtli and Jaime Pascual for their friendship and for all those bicycle trips, runs and half-marathons together. Thanks for waiting patiently on the top of the Königstuhl.

My studies at EMBL would have been a lot more difficult without the experience I gained in the Semliki Forest virus group, at the Institute of Biotechnology, Helsinki. In particular I owe my thanks to Dr. Johan Peränen for teaching me practically everything I know of modern molecular cloning techniques. The coffee group from Valimotie is acknowledged for the numerous and fun parties. I want to thank Jaana Suopanki for her friendship and for the financial support.

I would like to thank my mother Tuula Hyvönen for her everlasting interest and support in our lives.

I am ever grateful to Maija for her love and patience during all these years and for making this possible.

This study was supported by an EMBL pre-doctoral fellowship.

Heidelberg, September 1997, 

BIBLIOGRAPHY

André, B. and Springael, J.-Y. (1994) WWP, a new amino acid motif present in single or multiple copies in various proteins including dystrophin and the SH3-binding Yes-associated protein YAP65. *Biochim Biophys Res Comm*, **205**:1201–1205.

Artaljo, C. R., Lemmon, M. A., Schlessinger, J. and Palfrey, H. C. (1997) Specific role for the PH domain of dynamin-1 in the regulation of rapid endocytosis in adrenal chromaffin cells. *EMBO J*, **16**:1565–1574.

Bedford, M. T., Chan, D. C. and Leder, P. (1997) FBP WW domains and Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J*, **16**:2376–2383.

Bennet, V. and Gilligan, D. M. (1993) The spectrin-based membrane skeleton and micron-scale organization of the plasma membrane. *Annu Rev Cell Biol*, **9**:27–66.

Berridge, M. (1993) Inositol trisphosphate and calcium signalling. *Nature*, **361**:315–324.

Bhatnagar, R. and Gordon, J. I. (1997) Understanding covalent modifications of proteins by lipids: where cell biology and biophysics mingle. *Trends Cell Biol*, **7**:14–20.

Blaikie, P., Immanuel, D., Wu, J., li, N. and andB. Margolis, V. Y. (1994) A region in Shc distinct from the SH2 domain can bind tyrosine-phosphorylated growth factor receptors. *J Biol Chem*, **269**:32031–32034.

Blomberg, N. and Nilges, M. (1997) Functional diversity of PH domains: an exhaustive modeling study. Manuscript in preparation.

Bork, P. and Margolis, B. (1995) A phosphotyrosine interaction domain. *Cell*, **80**:693–694.

Bork, P., Schultz, J. and Ponting, C. P. (1997) Cytoplasmic signalling domains: the next generation. *Trends Biochem Sci*, **22**:296–298. The domains are listed in the accompanying poster.

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

Brinkmann, U., Mattes, R. E. and Buckel, P. (1989) High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. *Gene*, **85**:109–114.

Brünger, A. T. (1988) *X-PLOR Version 3.1 A system for X-ray crystallography and NMR* . Yale University Press, New Haven and London.

Buchsbaum, R., Telliez, J. B., Goonesekera, S. and Feig, L. A. (1996) The N-terminal pleckstrin, coiled-coil, and IQ domains of the exchange factor Ras-GRF act cooperatively to facilitate activation by calcium. *Mol Cell Biol*, **16**:4888–4896.

Calderone, T. L., Stevens, R. D. and Oas, T. G. (1996) High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. *J Mol Biol*, **262**:407–412.

Carpenter, C. L. and Cantley, L. C. (1996) Phosphoinositide kinases. *Curr Opin Cell Biol*, **8**:153–158.

Casey, P. J. (1995) Protein lipidation in cell signaling. *Science*, **268**:221–225.

Cerione, R. A. and Zheng, Y. (1996) The Dbl family of oncogenes. *Curr Opin Cell Biol*, **8**:216–222.

Chan, D. C., Bedford, M. T. and Leder, P. (1996) Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J*, **15**:1045–1054.

Chardin, P., Paris, S., Antonny, B., Robineau, S., Bérard-Dufour, S., Kackson, C. L. and Chabre, M. (1996) A human exchange factor for ARF contains Sec7 and pleckstrin homology domains. *Nature*, **384**:481–484.

Cheandle, C., Ivanshchenko, Y., South, V., Searfos, G. H., French, S., Howk, R., Ricca, G. A. and Jaye, M. (1994) Idetification of a Src SH3 domain binding motif by screening a random phage display library. *J Biol Chem*, **269**:24034–24039.

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. *Proc Natl Acad Sci USA*, **92**:7819–7823.

Chen, R.-H., Corbalan-Garcia, S. and Bar-Sagi, D. (1997) The role of the PH domain in the signal-dependent membrane targeting of Sos. *EMBO J*, **16**:1351–1359.

Cicchetti, C., Mayer, B. J., Thiel, G. and Baltimore, D. (1992) Identification of a protein that binds SH3 region of Abl and is similar to Bcr and GAP-Rho. *Science*, **257**:803–806.

Cifuentes, M. E., Honkanen, L. and Rebecchi, M. J. (1993) Proteolytic fragments of phosphoinositide-specific phospholipase C-δ_1. *J Biol Chem*, **268**:11586–11593.

Cowtan, K. (1994) DM. *Joint CCP4 and ESF EACBM newsletter on protein crystallography*, **31**:34–38.

Datta, K., Franke, T. F., Chan, T. O., Makris, A., Yang, S. I., Kaplan, D. R., Morrison, D. K., Golemis, E. A. and Tsichlis, P. N. (1995) AH / PH domain-mediated interaction between Akt molecules and its potential role in Akt regulation. *Mol Cell Biol*, **15**:2304–2310.

Davis, L. H. and Bennett, V. (1994) Identification of two regions of β_G spectrin that bind to distinct sites in brain membranes. *J Biol Chem*, **269**:4409–4416.

Divecha, N. and Irvine, R. F. (1995) Phospholipid signaling. *Cell*, **80**:269–278.

Downing, A. K., Driscoll, P. C., Gout, I., Salim, K., Zvelebil, M. J. and Waterfield, M. D. (1994) Three-dimensional solution structure of the pleckstrin homology domain from dynamin. *Curr Biol*, **4**:884–891.

Eck, M. J., Dhe-Paganon, S., Trub, T., Nolte, R. T. and Shoelson, S. E. (1996a) Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell*, **85**:695–705.

Eck, M. J., Pluskey, S., Trüb, T., Harrison, S. C. and Shoelson, S. E. (1996b) Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. *Nature*, **379**:277–280.

Eck, M. J., Shoelson, S. E. and Harrison, S. C. (1993) Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of $p56^{lck}$. *Nature*, **362**:87–91.

Emilsson, V., Näslund, A. K. and Kurland, C. G. (1993) Growth-rate-dependent accumualtion of twelwe tRNA species in *Eschericia coli*. *J Mol Biol*, **230**:483–491.

Felici, F., Castaglioni, L., Musacchio, A., Jappelli, R. and Cesareni, G. (1991) Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J Mol Biol*, **222**:301–310.

Ferguson, K. M., Lemmon, M. A., Schlessinger, J. and Sigler, P. B. (1994) Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human dynamin. *Cell*, **79**:199–209.

Ferguson, K. M., Lemmon, M. A., Schlessinger, J. and Sigler, P. B. (1995) Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell*, **83**:1037–1046.

Franke, T. F., Kaplan, D. R., Cantley, L. C. and Toker, A. (1997) Direct regulation of the *Akt* proto-oncogene product by phosphatidylinositol-3,4-bisphosphate. *Science*, **275**:665–668.

Frech, M., Ingley, M. A. E., Reddy, K. K., Falck, J. R. and Hemmings, B. A. (1997) High affinity binding of inositol phosphates and phosphoinositides to the pleckstrin homology domain of RAC / protein kinase B and their influence on kinase activity. *J Biol Chem*, **272**:8474–8481.

Fukuda, M., Kojima, T., Kabayama, H. and Mikoshiba, K. (1996) Mutation of pleckstrin homology domain of Bruton's tyrosine kinase in immunodeficiency impaired inositol 1,3,4,5-tetrakisphosphate binding capacity. *J Biol Chem*, **271**:30303–30306.

Fukuda, M. and Mikoshiba, K. (1996) Structure-function relationships of the mouse Gap1m. Determination of the inositol 1,3,4,5-tetrakisphosphate-binding domain. *J Biol Chem*, **271**:18838–18842.

Furey, W. and Swaminathan, S. (1996) *PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules*, vol. 278 of *Methods in Enzymology*. Academic Press, Orlando, Fl. in press.

Fushman, D., Cahill, S., Lemmon, M. A., Schlessinger, J. and Cowburn, D. (1995) Solution structure of pleckstrin homology domain of dynamin by heteronuclear NMR spectroscopy. *Proc Natl Acad Sci U S A*, **92**:816–820.

Gay, B., Furet, P., García-Echeverría, C., Rahuel, J., Chêne, P., Fretz, H., Schoepfer, J. and Caravetti, G. (1997) Dual specificity of Src homology 2 domains for phosphotyrosine peptide ligands. *Biochemistry*, **36**:5712–5718.

Gibson, T., Hyvönen, M., Musacchio, A., Saraste, M. and Birney, E. (1994) PH domain: the first anniversary. *Trends Biochem Sci*, **19**:349–353.

Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, **182**:319–326.

Gustafson, T. A., He, W., Craparo, A., Schaub, C. D. and O'Neill, T. J. (1995) Phosphotyrosine-dependent interaction of SHC and insulin receptor substrate 1 with the NPEY motif of the insulin receptor via novel non-SH2 domain. *Mol Cell Biol*, **15**:2500–2508.

Harlan, J. E., Hajduk, P. J., Yoon, H. S. and Fesik, S. W. (1994) Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature*, **371**:168–170.

Harlan, J. E., Yoon, H. S., Hajduk, P. J. and Fesik, S. W. (1995) Structural characterization of the interaction between a pleckstrin homology domain and phosphatidylinositol 4,5-bisphosphate. *Biochemistry*, **34**:9859–9864.

Haslam, R. J., Koide, H. B. and Hemmings, B. A. (1993) Pleckstrin domain homology. *Nature*, **363**:309–310.

Hatada, M. H., Lu, X., Laird, E. R., Green, J., Morgenstern, J. P., Lou, M., Marr, C. S., Phillips, T. B., Ram, M. K., Theriault, K., Zoller, M. J. and Karas, J. L. (1995) Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature*, **377**:32–38.

Hope, H. R. and Pike, L. J. (1996) Phosphoinositides and phosphoinositide-utilizing enzymes in detergent-insoluble lipid domains. *Mol Biol Cell*, **7**:843–851.

Hunter, T. (1997) Oncoprotein networks. *Cell*, **88**:333–346.

Hyvönen, M., Macias, M. J., Nilges, M., Oschkinat, H., Saraste, M. and Wilmanns, M. (1995) Structure of the binding site for inositol phosphates in a PH domain. *EMBO J*, **14**:4676–4685.

Hyvönen, M. and Saraste, M. (1997) Expression of cDNAs in *Escherichia coli* using T7 promoter. in *Cell Biology - A Laboratory manual*, in press.

Hyvönen, M. and Saraste, M. (1997) Structure of PH domain and Btk motif from Bruton's tyrosine kinase: Molecular explanations for the X-linked agammaglobulinemia. *EMBO J*, **16**:3396–3404.

Jones, T., Zou, J., Cowan, S. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst*, **A47**:110–119.

Kaplan, T. F. F. D. R. and Cantley, L. C. (1997) PI3K: Downstream AKTion blocks apoptosis. *Cell*, **88**:435–437.

Kavanaugh, W. M., Turck, C. W. and Williams, L. T. (1995) PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. *Science*, **268**:1177–1179.

Kavanaugh, W. M. and Williams, L. T. (1994) An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. *Science*, **266**:1862–1865.

Klarland, J. K., Guilherme, A., Holik, J. J., Virbasius, J. V., Chawla, A. and Czech, M. P. (1997) Signaling by phosphoinositide-3,4,5-trisphosphate through proteins containing pleckstrin and Sec7 homology domains. *Science*, **275**:1927–1930.

Klippel, A., Kavanaugh, W. M., Pot, D. and Williams, L. T. (1997) A specific product of phosphatidylinositol 3-kinase directly activates the protein kinase Akt through its pleckstrin homology domain. *Mol Cell Biol*, **17**:338–344.

Koch, W. J., Inglese, J., Stone, W. C. and Lefkowitz, R. J. (1993) The binding site for the $\beta\gamma$ subunits of heterotrimeric G proteins on the β-adrenergic receptor kinase. *J Biol Chem*, **268**:8256–8260.

Koshiba, S., Kigawa, T., Kim, J.-H., Shirouzu, M., Bowtell, D. and Yokoyama, S. (1997) The solution structure of the pleckstrin homology domain of mouse Son-of-sevenless 1 (mSos-1). *J Mol Biol*, **269**:579–591.

Kraulis, P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. *J Appl Cryst*, **24**:946–950.

Kubiseski, T. J., Chook, Y. M., Parris, W. E., Rozakis-Adcock, M. and Pawson, T. (1997) High affinity binding of the pleckstrin homology domain of mSos1 to phosphatidylinositol (4,5)-bisphosphate. *J Biol Chem*, **272**:1799–1804.

Lamzin, V. and Wilson, K. (1993) Automated refinement of protein models. *Acta Cryst*, **D49**:129–147.

Laskowski, R. A., MacArthur, M. W., Hutchinson, E. G. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**:283–291.

Lee, C.-H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E. and Kuriyan, J. (1994) Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure*, **2**:423–438.

Lee, C.-H., Saksela, K., Mirza, U. A., Chait, B. T. and Kuriyan, J. (1996) Crystal structure of the conserved core of HIV Nef complexed with a Src family SH3 domain. *Cell*, **85**:931–942.

Lehto, V.-P., Wasenius, V.-M., Salvén, P. and Saraste, M. (1988) Transforming and membrane proteins. *Nature*, **334**.

Lemmon, M. A., Ferguson, K. M., O'Brien, R., Sigler, P. B. and Schlessinger, J. (1995) Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. *Proc Natl Acad Sci U S A*, **92**:10472–10476.

Li, T., Tsukada, S., Satterthwaite, A., Havlik, M. H., Park, H., Takatsu, K. and Witte, O. N. (1995) Activation of Bruton's tyrosine kinase (BTK) by a point mutation in its pleckstrin homology (PH) domain. *Immunity*, **2**:451–460.

Lim, W. A., Richards, F. M. and Fox, R. O. (1994) Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature*, **372**:375–379.

Lin, H. C. and Gilman, A. G. (1996) Regulation of dynamin I GTPase activity by G protein $\beta\gamma$ subunits and phosphatidylinositol 4,5-phosphate. *J Biol Chem*, **271**:27979–27982.

Lu, K. P., Hanes, S. D. and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature*, **380**:544–547.

Ma, A. D., Brass, L. F. and Abrams, C. S. (1997) Pleckstrin associates with plasma membranes and induces formation of membrane projections: requirements for phosphorylation and the NH_2-terminal PH domain. *J Cell Biol*, **136**:1071–1079.

Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M. and Oschkinat, H. (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature*, **382**:646–649.

Macias, M. J., Musacchio, A., Ponstingl, H., Nilges, M., Saraste, M. and Oschkinat, H. (1994) Structure of the pleckstrin homology domain from β-spectrin. *Nature*, **369**:675–677.

Mayer, B. J., Hamaguchi, M. and Hanafusa, H. (1988) A novel oncogene with structural similarity to phospholipase C. *Nature*, **332**:272–275.

Mayer, B. J., Ren, R., Clark, K. L. and Baltimore, D. (1993) A putative modular domain present in diverse signaling proteins. *Cell*, **73**:629–630.

Michiels, F., Stam, J. C., Hordijk, P. L., van der Kammen, R. A., Ruuls-Van Stalle, L., Feltcamp, C. A. and Collard, J. G. (1997) Regulated membrane localization of Tiam1, regulated by the NH_2-terminal pleckstrin homology domain, is required for Rac-dependent membrane ruffling and C-Jun NH_2-terminal kinase activation. *J Cell Biol*, **137**:387–398.

Mumby, S. M. (1997) Reversible palmitoylation of signaling proteins. *Curr Opin Cell Biol*, **9**:148–154.

Musacchio, A., Gibson, T., Lehto, V.-P. and Saraste, M. (1992a) SH3 - an abundant protein module in search for a function. *FEBS Lett*, **307**:55–61.

Musacchio, A., Gibson, T., Rice, P., Thompson, J. and Saraste, M. (1993) The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem Sci*, **18**:343–348.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. and Saraste, M. (1992b) Crystal structure of a Src-homology 3 (SH3) domain. *Nature*, **359**:851–854.

Musacchio, A., Saraste, M. and Wilmanns, M. (1994a) High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat Struct Biol*, **1**:546–551.

Musacchio, A., Wilmanns, M. and Saraste, M. (1994b) Structure and function of the SH3 domain. *Prog Biophys molec Biol*, **61**:283–297.

Myers Jr., M. G., Grammer, T. C., Brooks, J., Glasheen, E. M., Wang, L. M., Sun, X. J., Blenis, J., Pierce, J. H. and White, M. F. (1995) The pleckstrin homology domain in insulin receptor substrate-1 sensitizes insulin signaling. *J Biol Chem*, **270**:11715–11718.

Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr*, **A50**:157–163.

Nayal, M. and Di Cera, E. (1996) Valence screening of water in protein crystals reveals potential Na^+ binding sites. *J Mol Biol*, **256**:228–234.

Nilges, M., Macias, M. J., O'Donaghue, S. I. and Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J Mol Biol*, **269**:408–422.

Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A. and Wierenga, R. K. (1993) Crystal structure of the SH3 domain of human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J*, **12**:2617–2624.

Parker, P. J., Hemmings, B. A. and Gierschik, P. (1994) PH domains and phospholipases - a meaningful relationship ? *Trends Biochem Sci*, **19**:54–55.

Paterson, H. F., Savopoulos, J. W., Perisic, O., Cheung, R., Ellis, M. V., Williams, R. L. and Katan, M. (1995) Phospholipase C δ_1 requires a pleckstrin homology domain for interaction with the plasma membrane. *Biochem J*, **312**:661–666.

Peränen, J., Rikkonen, M., Hyvönen, M. and Kääriäinen, L. (1996) T7 vectors with modified T7*lac* promoter for expression of proteins in Escherichia coli. *Anal Biochem*, **236**:371–373.

Pérez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Winge, D. R., Beckerle, M. C. and Summers, M. F. (1994) Structure of the carboxy-terminal LIM domain from the cysteine rich protein CRP. *Nat Struct Biol*, **1**:388–398.

Piccione, E., Case, R. D., Domchek, S. M., Hu, P., Chaudhuri, M., Baker, J. M., Sclessinger, J. and Shoelson, S. E. (1993) Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide/SH2 domain binding. *Biochemistry*, **32**:3197–3202.

Pitcher, J. A., Touhara, K., Payne, E. S. and Lefkowitz, R. J. (1995) Pleckstrin homology domain-mediated membrane association and activation of the β-adrenergic receptor kinase requires coordinate interaction with $G_{\beta\gamma}$ subunits and lipid. *J Biol Chem*, **270**:11707–11710.

Ponting, C. R. and Bork, P. (1996) Pleckstrin's repeat performance: a novel domain in G-protein signaling. *Trends Biochem Sci*, **21**:245–246.

Rahuel, J., Gay, B., Erdmann, D., Strauss, A., Garcia-Echeverría, G., Furet, P., Caravatti, G., Fretz, H., Schoepfer, J. and Grütter, M. G. (1996) Structural basis for specificity of GRB2-SH2 revealed by a novel ligand binding mode. *Nat Struct Biol*, **3**:586–589.

Rameh, L. E., Arvidsson, A.-K., Carraway III, K. L., Couvillon, A. D., Rathbun, G., Crompton, A., VanRenterghem, B., Czech, M. P., Ravichandran, K. S., Burakoff, S. J., Wang, D.-S., Chen, C.-S. and Cantley, L. C. (1997) A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains. *J Biol Chem*, **272**:22059–22066.

Rawlings, D. J., Saffran, D. C., Tsukada, S., Largaespada, D. A., Grimaldi, J. C., Cohen, L., Mohr, R. N., Bazan, J. F., Howard, M., Copeland, N. G. and et al. (1993) Mutation of unique region of Bruton's tyrosine kinase in immunodeficient XID mice. *Science*, **261**:358–361.

Ren, R., Mayer, B. J., P.Cicchetti and Baltimore, D. (1993) Identification of a ten-amino acid proline rich SH3 binding site. *Science*, **259**:1157–1161.

Rotin, D., Bar-Sag, D., O'Brodovich, H., Meriläinen, J., Lehto, V. P., Canessa, C. M., Rossier, B. C. and Downey, G. P. (1994) An SH3 binding region in the epithelial Na^+ channel (αrENaC) mediates its localization at the apical membrane. *EMBO J*, **13**:4440–4450.

Russell, R. B., Breed, J. and Barton, G. J. (1992) Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. *FEBS Lett*, **304**:15–20.

Sadowski, I., Stone, J. C. and Pawson, T. (1986) A noncatalytical domain conserved among cutoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus $P130^{gag-fps}$. *Mol Cell Biol*, **6**:4396–4408.

Salim, K., Bottomley, M., Querfurth, E., Zvelebil, M., Gout, I., Scaife, R., Margolis, R., Gigg, R., Smith, C., Driscoll, P., Waterfield, M. and Panayotou, G. (1996) Distinct specificity in the recognition of phosphoinositides by the pleckstrin homology domains of dynamin and Bruton's tyrosine kinase. *EMBO J*, **15**:6241–6250.

Saraste, M. and Hyvönen, M. (1995) Pleckstrin homology domains: a fact file. *Curr Opin Struct Biol*, **5**:403–408.

Schild, L., Lu, Y., Gautschi, I., Schneeberger, E., Lifton, R. P. and Rossier, B. C. (1996) Identification of a PY motif in the epithelial Na channel subunits as a target sequence for mutations causing channel activation found in Liddle syndrome. *EMBO J*, **15**:2381–2387.

Sicheri, F., Moarefi, I. and Kuriyan, J. (1997) Crystal structure of the Src family tyrosine kinase Hck. *Nature*, **385**:602–609.

Simons, K. and Ikonen, E. (1997) Functional rafts in cell membranes. *Nature*, **387**:569–572.

Smith, C. I., Islam, K. B., Vorechovsky, I., Olerup, O., Wallin, E., Rabbani, H., Baskin, B. and Hammarström, L. (1994) X-linked agammaglobulinemia and other immunoglobulin deficiencies. *Immunol Rev*, **138**:159–183.

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.

Smith, G. P. and Scott, J. K. (1993) Libraries of peptides and proteins displayed on filamentous phage. In Wu, R., ed., *Recombinant DNA*, vol. 217 of *Methods in Enzymology*, pp. 228–257. Academic Press, Inc.

Songyang, Z., Shoelson, S. E., Chauduri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B. and Cantley, L. (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell*, **72**:767–778.

Songyang, Z., Shoelson, S. E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X. R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., Ren, R., Baltimore, D., Ratnofsky, S., Feldman, R. A. and Cantley, L. (1994) Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps / fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol Cell Biol*, **14**:2777–2785.

Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J. and Kay, B. K. (1994) Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. *J Biol Chem*, **269**:23853–23856.

Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J. and Rotin, D. (1996) WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na^+ channel deleted in Liddle's syndrome. *EMBO J*, **15**:2371–2380.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*, **185**:60–89.

Sudol, M. (1996a) Structure and function of the WW domain. *Prog Biophys molec Biol*, **65**:113–132.

Sudol, M. (1996b) The WW module competes with the SH3 domain? *Trends Biochem Sci*, **21**:161–163.

Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K. and Lehman, D. (1995) Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain. *J Biol Chem*, **270**:14733–14741.

Superti-Furga, G. and Courtneidge, S. (1995) Structure-function relationships in Src family and related protein tyrosine kinases. *BioEssays*, **17**:321–330.

Superti-Furga, G., Fumagalli, S., Koegl, M., Courtnidge, S. A. and Draetta, G. (1993) Csh inhibition of c-Src activity requires both the SH2 and SH3 domains of c-Src. *EMBO J*, **12**:2625–2634.

Takeuchi, H., Kanematsu, T., Misumi, Y., Yaakob, H. B., Yagisawa, H., Ikehara, Y., Watanabe, Y., Tan, Z., Shears, S. B. and Hirata, M. (1996) Localization of a high-affinity inositol 1,4,5-trisphosphate / inositol 1,4,5,6-tetrakisphosphate binding domain to the pleckstrin homology module of a new 130 kDa protein: characterization of the determinants of structural specificity. *Biochem J*, **318**:561–568.

Thomas, J. D., Sideras, P., Smith, C. I., Vorechovsky, I., Chapman, V. and Paul, W. E. (1993) Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. *Science*, **261**:355–358.

Timm, D., Salim, K., Gout, I., Guruprasad, L., Waterfield, M. and Blundell, T. (1994) Crystal structure of the pleckstrin homology domain from dynamin. *Nat Struct Biol*, **1**:782–788.

Toker, A. and Cantley, L. C. (1997) Signalling through the lipid products of phosphoinositide-3-OH kinase. *Nature*, **387**:673–676.

Touhara, K., Inglese, J., Pitcher, J. A., Shaw, G. and Lefkowitz, R. J. (1994) Binding of G protein $\beta\gamma$-subunits to pleckstrin homology domains. *J Biol Chem*, **269**:10217–10220.

Tronrud, D. E., Eyck, L. F. T. and Matthews, B. W. (1987) An efficient general-purpose least-square refinement program for macromolecular structures. *Acta Cryst*, **A43**:489–501.

Trüb, T., Choi, W. E., Wolf, G., Ottinger, E., Chen, Y., Weiss, M. and Shoelson, S. E. (1995) Specificity of the PTB domain of Shc for β turn-forming pentapeptide motifs amino-terminal to phosphotyrosine. *J Biol Chem*, **270**:18205–18208.

Tsukada, S., Saffran, D. C., Rawlings, D. J., Parolini, O., Allen, R. C., , Kalisak, I., Sparkes, R. S., Kubagawa, H., Mohandas, T., Quan, S., Belmont, J. W., Cooper, M. D., Conley, M. E. and Witte, O. N. (1993) Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell*, **72**:279–190.

Tsukada, S., Simon, M., Witte, O. and Katz, A. (1994) Binding of $\beta\gamma$ subunits of trimeric G proteins to the PH domain of Bruton tyrosine kinase. *Proc Natl Acad Sci USA*, **91**:11256–11260.

Tyers, M., Stewart, R. A. R. M. I., , Varrichio, A. M., Shorr, R. G. L., Haslam, R. J. and Harley, C. B. (1988) Molecular cloning and expression of the major protein kinase C substrate of platelets. *Nature*, **333**:470–473.

van der Geer, P., Wiley, S., Lai, V. K.-M., Olivier, J. P., Gish, G. D., Stephens, R., Kaplan, D., Shoelson, S. and Pawson, T. (1995) A conserved amino-terminal Shc domain binds to phosphotyrosine motifs in activated receptors and phosphopeptides. *Curr Biol*, **5**:404–412.

Vanhaesebroeck, B., Stein, R. C. and Waterfield, M. D. (1996) The study of phosphoinositide 3-kinase function. In *Cell Signalling*, vol. 27 of *Cancer Surveys*, pp. 249–270. Imperial Cancer Research Fund.

Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarström, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C. I. E. and Bentley, D. R. (1993) The gene involved in X-linked agammaglobulinemia is a member of the src family of protein-tyrosine kinases. *Nature*, **361**:226–233.

Viguera, A. R., Arrondo, J. L. R., Musacchio, A., Saraste, M. and Serrano, L. (1994) Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. *Biochemistry*, **33**:10925–10933.

Vihinen, M., Iwata, T., Kinnon, C., Kwan, S. P., Ochs, H. D., Vorechovsky, I. and Smith, C. I. (1996) BTKbase, mutation database for X-linked agammaglobulinemia (XLA). *Nucleic Acids Res*, **24**:160–165.

Vihinen, M., Nilsson, L. and Smith, C. I. (1994) Tec homology (TH) adjacent to the PH domain. *FEBS Lett*, **350**:263–265.

Vriend, G. (1990) WHAT IF - a molecular modelling and drug design program. *J Mol Graph*, **8**:52–56.

Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L. and Kuriyan, J. (1992) Crystal structure of the phosphotyrosine recognition domain SH2 of the v-*src* complexed with tyrosine-phosphorylated peptides. *Nature*, **358**:646–653.

Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. and Kuriyan, J. (1993) Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell*, **72**:779–790.

Wang, D.-S., Miller, R., Shaw, R. and Shaw, G. (1996) The pleckstrin homology domain of human βIΣII spectrin is targeted to the plasma membrane. *Biochim Biophys Res Comm*, **225**:420–426.

Wang, D. S. and Shaw, G. (1995) The association of the C-terminal region of βIΣII spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1,4,5 triphosphate binding site. *Biochem Biophys Res Commun*, **217**:608–615.

Wang, D. S., Shaw, R., Winkelmann, J. C. and Shaw, G. (1994) Binding of PH domains of β-adrenergic receptor kinase and β-spectrin to WD40 / β-transducin repeat containing regions of the β-subunit of trimeric G-proteins. *Biochem Biophys Res Commun*, **203**:29–35.

Whitehead, I., Kirk, H., Tognon, C., Gonzalez, G. T. and Kay, R. (1995) Expression cloning of lfc, a novel oncogene with structural similarities to guanine nucleotide exchange factors and to the regulatory region of protein kinase C. *J Biol Chem*, **270**:18388–18395.

Wilmanns, M. and Nilges, M. (1996) Molecular replacement with NMR models using distance-derived pseudo *B* factors. *Acta Cryst*, **D52**:973–982.

Xu, W., Harrison, S. C. and Eck, M. J. (1997) Three dimensional structure of the tyrosine kinase c-Src. *Nature*, **385**:595–601.

Yagisawa, H., Hirata, M., Kanematsu, T., Watanabe, Y., Ozaki, S., Sukuma, K., Tanaka, H., Yabuta, N., Kamata, H., Hirata, H. and Nojima, H. (1994) Expression and characterization of an inositol 1,4,5-trisphosphate binding domain of phosphatidylinositol-specific phospholipase C-δ1. *J Biol Chem*, **269**:20179–20188.

Yang, W. and Desiderio, S. (1997) BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. *Proc Nat Acad Sci USA*, **94**:604–609.

Yenush, L., Makati, K. J., a. Smith J. ll, H., Ishibashi, O., Myers Jr., M. G. and White, M. F. (1996) The pleckstrin homology domain is the principal link between the insulin receptor and IRS-1. *J Biol Chem*, **271**:24300–24306.

Yoon, H. S., Hajduk, P. J., Petros, A. M., Olejniczak, E. T., Meadows, R. P. and Fesik, S. W. (1994) Solution structure of a pleckstrin-homology domain. *Nature*, **369**:672–675.

Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S. and Shcreiber, S. L. (1992) Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science*, **258**:1665–1668.

Zhang, G., Kazanietz, M. G., Blumberg, P. M. and Hurley, J. H. (1995a) Crystal structure of the Cys2 activator-binding domain of protein kinase Cδ in complex with phorbol ester. *Cell*, **81**:917–924.

Zhang, P., Talluri, S., Deng, H., Branton, D. and Wagner, G. (1995b) Solution structure of the pleckstrin homology domain of Drosophila β-spectrin. *Structure*, **3**:1185–1195.

Zheng, J., Cahill, S. M., Lemmon, M. A., Fushman, D., Schlessinger, J. and Cowburn, D. (1996a) Identification of the binding site for acidic phospholipids on the PH domain of dynamin: implications for stimulation of GTPase activity. *J Mol Biol*, **255**:14–21.

Zheng, Y., Zangrilli, D., Cerione, R. A. and Eva, A. (1996b) The pleckstrin homology domain mediates transformation by oncogenic dbl through specific intracellular targeting. *J Biol Chem*, **271**:19017–19020.

Zhou, M. M., Harlan, J. E., Wade, W. S., Crosby, S., Ravichandran, K. S., Burakoff, S. J. and Fesik, S. W. (1995a) Binding affinities of tyrosine-phosphorylated peptides to the COOH-terminal SH2 and NH2-terminal phosphotyrosine binding domains of Shc. *J Biol Chem*, **270**:31119–31123.

Zhou, M. M., Huang, B., Olejniczak, E. T., Meadows, R. P., Shuker, S. B., Miyazaki, M., Trub, T., Shoelson, S. E. and Fesik, S. W. (1996) Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nat Struct Biol*, **3**:388–393.

Zhou, M. M., Ravichandran, K. S., Olejniczak, E. F., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J. and Fesik, S. W. (1995b) Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature*, **378**:584–592.