The EMBO Journal Vol.16 No.9 pp.2376–2383, 1997

FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands

Mark T.Bedford, David C.Chan[1] and Philip Leder[2]

Department of Genetics, Harvard Medical School, Howard Hughes Medical Institute, 200 Longwood Avenue, Boston, MA 02115, USA

[1]Present address: Department of Biology, Massachusetts Institute of Technology, Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

[2]Corresponding author

WW domains are conserved protein motifs of 38–40 amino acids found in a broad spectrum of proteins. They mediate protein–protein interactions by binding proline-rich modules in ligands. A 10 amino acid proline-rich portion of the morphogenic protein, formin, is bound *in vitro* by both the WW domain of the formin-binding protein 11 (FBP11) and the SH3 domain of Abl. To explore whether the FBP11 WW domain and Abl SH3 domain bind to similar ligands, we screened a mouse limb bud expression library for putative ligands of the FBP11 WW domain. In so doing, we identified eight ligands (WBP3 through WBP10), each of which contains a proline-rich region or regions. Peptide sequence comparisons of the ligands revealed a conserved motif of 10 amino acids that acts as a modular sequence binding the FBP11 WW domain, but not the WW domain of the putative signal transducing factor, hYAP65. Interestingly, the consensus ligand for the FBP11 WW domain contains residues that are also required for binding by the Abl SH3 domain. These findings support the notion that the FBP11 WW domain and the Abl SH3 domain can compete for the same proline-rich ligands and suggest that at least two subclasses of WW domains exist, namely those that bind a PPLP motif, and those that bind a PPXY motif.
Keywords: formin-binding protein/proline-rich/SH3 domain/WW domain

Introduction

Protein–protein interactions form the basis of a wide variety of cellular processes. Often, these interactions depend upon a small number of modular domains that function as specific protein-binding structures. Prominent among these are Src homology region 2 (SH2) domains (Marengere and Pawson, 1994), Src homology region 3 (SH3) domains (Musacchio *et al.*, 1994), pleckstrin homology (PH) domains (Ferguson *et al.*, 1995) and phosphotyrosine-binding (PTB) domains (Bork and Margolis, 1995). These domains have been shuffled extensively during evolution and are now found in a wide array of proteins. SH2 and PTB domains both interact with phosphotyrosine residues. SH3 domains bind proline-rich ligands and recognize the PXXP motif. The PH domain binding motif has yet to be determined, but may be involved in membrane association of the proteins in which they occur through interaction with phospholipids (Zheng *et al.*, 1996).

In the course of a search for proteins that interact with formins, protein products of the mouse *limb deformity* locus (Woychik *et al.*, 1985, 1990), we came upon a group of formin-binding proteins (FBPs) that encode one of two classes of modular elements that interact with the proline-rich motif prominently encoded in all the known formin isoforms (Chan *et al.*, 1996). One class, as expected, contained one or more SH3 domains, known to interact with the formin proline-rich sequence (Ren *et al.*, 1993), while the other contained a novel motif, the WW domain, now known to bind proline-rich peptide stretches (Einbond and Sudol, 1996). This latter motif was identified by computer homology searches (Andre and Springael, 1994; Bork and Sudol, 1994; Hofmann and Bucher, 1995) and is characterized by two highly conserved tryptophan residues and a proline residue. The distance between the first W residue and the P residue is either 26 or 27 amino acids (aa) and demarcates an active binding module (Chen and Sudol, 1995; Chan *et al.*, 1996).

The WW domain has been identified in a variety of proteins. These include a putative transmembrane protein CD45AP (Cahir McFarland and Thomas, 1995), the signal transducing factor YAP65 (YES-associated protein) (Sudol *et al.*, 1995), cytoskeletal components dystrophin and utrophin (Bork and Sudol, 1994), ubiquitin–protein ligases such as NEDD-4 (Staub and Rotin, 1996), nuclear proteins including Pin1/dodo, FE65 and Prp40 as well as the FBPs (Fiore *et al.*, 1995; Chan *et al.*, 1996; Kao and Siliciano, 1996; Lu *et al.*, 1996; Maleszka *et al.*, 1996). The NMR structure of the hYAP WW domain complexed with a proline-rich peptide has been determined. It consists of a three-stranded anti-parallel β-sheet with a hydrophobic pocket which interacts with the ligand (Macias *et al.*, 1996). This structure is quite distinct from that of the SH3 domain (Musacchio *et al.*, 1994).

We have previously shown that the formin-binding protein FBP11 bears two WW domains that are separated by a 15 amino acid spacer and are capable of binding a 10 aa proline-rich motif (ld10) represented in most formin isoforms which contains an essential PPLP sequence (Chan *et al.*, 1996). The FBP11 binding motif (PPLP) differs from the hYAP binding motif (PPPPY) described by Chen and Sudol (1995). The latter motif was identified in two novel proteins, WBP1 and WBP2, and was found to be sufficient for WW domain binding (Chen and Sudol, 1995). Using functional screening of a cDNA expression library, we have identified eight putative ligands for FBP11 (WBP3 through WBP10). Three of these proteins are novel and four of the remaining five are nuclear. Like the

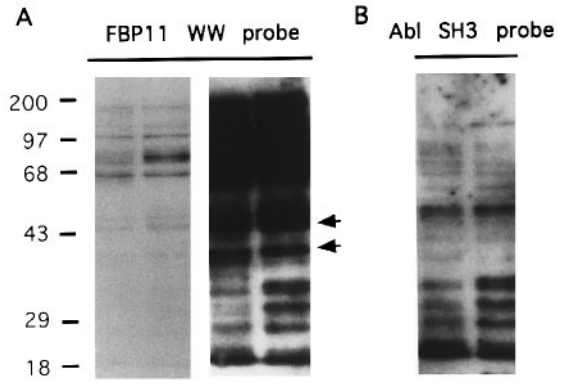

Fig. 1. The WW domain of FBP11 and the Abl SH3 domain bind a common set of proteins as determined by far Western blot analysis of epithelial cell lysates. (**A**) Total cell lysate probed with GST–FBP11 WW domain. Total cell lysates from Neu over-expressing (lane 1) and mutant Neu overexpressing (lane 2) cells were prepared in RIPA buffer. Lysates were run on an SDS–PAGE gel, blotted onto nitrocellulose and incubated with radiolabeled GST–FBP11 WW. After washing, blots were exposed to film (5 h on left; overnight exposure on right). Molecular weight markers (in kDa) are indicated. The arrowheads indicate a 40 kDa and a 44 kDa band which are only bound by the FBP11 WW domain and not the Abl SH3 domain. (**B**) As (A), except that the blot was incubated with radiolabeled GST–Abl SH3 and only the overnight exposure is shown here.

formin proline-rich domain, the majority of these ligands also bind to the Abl SH3 domain, but not SH3 domains found in Fyn or Src. Alignment studies allowed us to identify the consensus sequence noted above this domain (PPLP). We also showed that each of the two individual WW domains of FBP11 are capable of binding the formin proline-rich module (ld10) that contains this consensus and that each WW domain displays a similar binding profile when analyzed by alanine replacement mutagenesis of the ld10 peptide. This binding profile is consistent with the consensus binding motif identified in the putative WBP3–10 ligands, with the most sensitive points in the alanine scan being the most conserved in the alignment. Finally, we demonstrated that the WW domains of FBP11 and YAP are extremely specific in their ligand recognition abilities and that very little, if any, cross-reactivity occurs.

Results

The FBP11 WW domain binds to a number of proteins

The WW domain of FBP11 and the Abl SH3 domain have been shown to bind to the same proline-rich 10 aa sequence encoded in the morphogen protein formin (Chan *et al.*, 1996). Inasmuch as these two domains are quite different from one another, we wished to determine whether the FBP11 WW domain recognizes a distinct subset of proteins. To do this, we carried out blotting analyses of lysates of eight epithelial cell lines using the FBP11 WW domain and the Abl SH3 domain as ^{32}P-labeled GST fusion probes; two representative cell lines are shown in Figure 1. Several common bands were detected by the two probes (Figure 1A and B, see particularly the lower-molecular weight bands), suggesting that these two distinct proline-rich binding domains do recognize some of the same ligands. On the other hand, the FBP11 WW domain bound more strongly to a number of high-molecular weight bands (which may include formin) and specifically to proteins of 40 kDa and 44 kDa, which do not appear to interact with the Abl SH3 probe (Figure 1, arrowheads). These data imply that the ^{32}P-labeled GST–FBP11 WW fusion protein could function as a probe to detect candidate ligands in an expression library screen, and that it might detect two sets of proteins—one that binds both Abl SH3 and FBP11 WW domains, and another that binds only the FBP11 WW domain.

Identification of WW domain binding proteins

As noted above, the protein FBP11 contains two distinct WW domains spaced 15 aa from one another (Chan *et al.*, 1996). Initially, a GST fusion protein containing both of these WW domains was used to screen two expression libraries derived from mouse 10.5-day embryonic limb buds. One library was primed with random oligonucleotides, the other with oligo dT. Of ~10^6 plaques screened, over 100 positives of varying intensity were picked and 31 were plaque-purified. During the plaque purification procedure very weak binders were discarded. Plaque-pure phage were used to infect a bacterial strain harboring the *Cre* gene, allowing recombination at the lox sites and the creation of pET vectors (Novagen) containing the insert of interest. The resulting plasmid DNA was isolated and the inserts were liberated by digestion with the appropriate restriction enzyme. Plasmids were grouped according to their insert size and a member of each group was sequenced. All sequences analyzed encoded proteins with extensive proline-rich tracts. The screen identified eight different proteins (Table I). These proteins have been termed WW domain binding proteins (WBPs) (Chen and Sudol, 1995). Two of the sequences (WBP3 and 5) showed no significant similarity to any known protein in the database. WBP7 was highly homologous to the mouse trithorax protein over two short stretches, and the other five sequences corresponded to known proteins indicated as homologues in Table I. Interestingly, the majority of the known WBPs are found in the nucleus. One is a serine/threonine kinase that is involved in splicing (WBP6), and three contain DNA-binding motifs.

The WBPs vary in their ability to bind the WW domain of FBP28, and the SH3 domains of Abl, Fyn and Src

Apart from the FBP11 protein, there are four other known FBPs that contain WW domains, raising the possibility that each might have a particular affinity for specific proline-rich motifs (Chan *et al.*, 1996). To test this, we determined the ability of a number of SH3 and WW domains to bind the panel of WBPs described above. Immobilized protein binding analyses were carried out by overexpressing the WBP fusion proteins in bacteria, electrophoresing bacterial extracts through SDS–PAGE gels, and immobilizing the proteins on a nylon filter. Several identical filters were prepared in this manner, and they were probed with the ^{32}P-labeled GST–WW domains of FBP11 and FBP28 in addition to the ^{32}P-labeled GST–SH3 domains of Abl, Fyn and Src. All the WBPs were bound well by the FBP11 WW domain under these denaturing conditions (Figure 2). The other WW domain (FBP28) and the SH3 domains bound the panel of WBPs

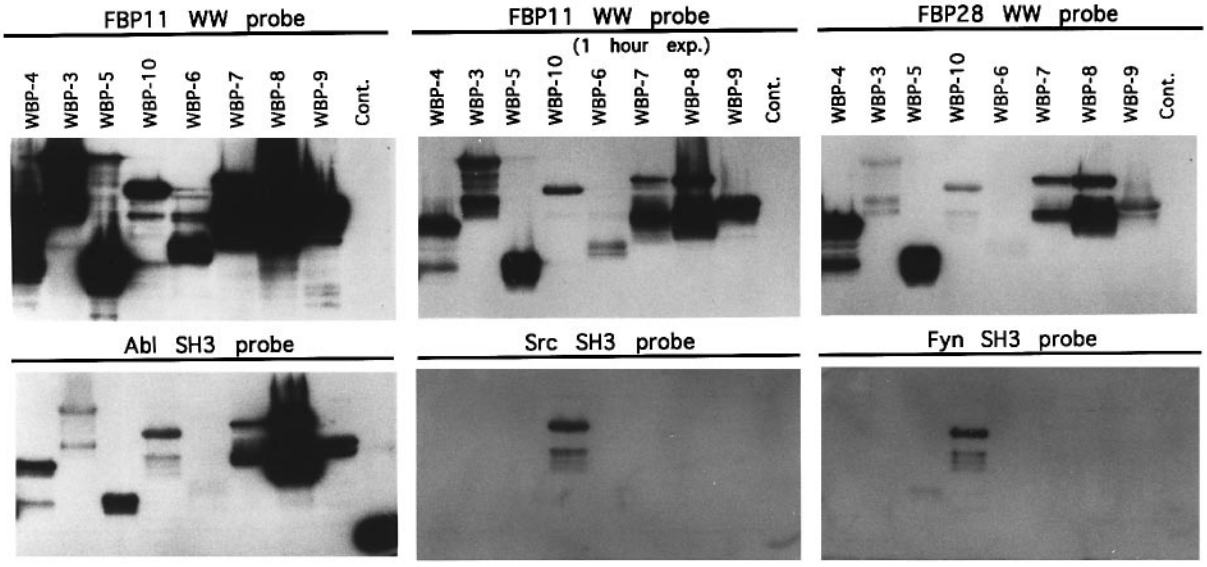

Fig. 2. Determination of binding specificity of the WBPs. Bacteria were transformed with pEXlox expressing WBP fusion proteins. Fusion protein expression was induced with 0.1 mM IPTG (see Materials and methods). Bacterial lysates were run on an SDS–PAGE gel, blotted onto nitrocellulose and incubated with radiolabeled GST probes as indicated in the figure. The control lane on the filters probed with the WW domains is the exlox fusion protein without an insert. The control lane on the filters probed with the SH3 domains is GST–3BP1. All blots were exposed overnight to film, except for the FBP11 WW probed blot, for which we display two exposure times (overnight and 1 h).

Table I. The identity of clones that bind the WW domain of FBP11[a]

Clone name	Homologue	Structural features	Proline-rich tract[b]
WBP3	–	–	10/13 (1 domain)
WBP4	ZFM1	Zinc finger motif	52/175 (multiple domains)
WBP5	–	–	18/19 (1 domain)
WBP6	SRPK-1	S/T kinase	16/20 (1 domain)
WBP7	–	Tri-thorax-like sequence	30/48 (multiple domains)
WBP8	NDPP1	HLH-like domain	21/23 (2 overlapping domains)
WBP9	ATBF1	17 zinc fingers and 4 homeodomains	18/25 (1 domain)
WBP10	MeCP2	AT-rich binding motif	9/13 (1 domain)

[a]The clones are called WBP3–WBP10 (WW domain binding protein). WBP1 and 2 have been previously described (Chen and Sudol, 1995).
[b]The fraction denotes the number of proline residues over the total length of the proline-rich domain.

to varying degrees (Figure 2). The Abl SH3 domain recognizes all of the FBP11 ligands, except for WBP6, whereas the SH3 domains of Src and Fyn only bound WBP10, albeit weakly.

Each WW domain of FBP11 displays a similar binding profile

The expression libraries were screened with a GST–FBP11 fusion protein that contained two WW domains. To determine if these WW domains recognize different binding motifs, they were cloned separately into GST fusion vectors. The individual WW domains, termed FBP11A and B, are both bound by the ld10 ligand (Figure 3A), a proline-rich ligand originally discovered within the morphogenic protein, formin (Chan *et al.*, 1996). However, FBP11A was bound more strongly than FBP11B. When the reciprocal experiment was carried out, both FBP11A and B demonstrated strong binding to ld10 (data not shown), suggesting that in the former experiment, FBP11B may be more sensitive to the denaturing conditions of the assay than FBP11A.

In order to assess the structural binding profiles of these two WW domains, FBP11A and B were used to screen an alanine substitution panel of fusion proteins containing the proline-rich ld10 sequence. With the exception of an alanine substitution in the third position of the ld10 sequence, FBP11A and B both displayed very similar binding profiles which match the profile obtained with the FBP11A+B fusion protein, the protein containing both binding domains (Figure 3B). Since these results indicate similar binding specificities for the two domains, we aligned the nine proline-rich ligands in the hope of finding a motif, or a tandem repeat of a motif, similar to that of ld10.

Alignment of the WBP's proline-rich regions reveals a conserved PPLP motif and an overlapping Abl SH3 ligand

All the WWPs that we isolated contain proline-rich tracts and most have a single proline-rich motif which allows for clear alignment (Figure 4). WBP8 possesses two overlapping regions (which we call WBP8A and WBP8B) that match the predicted consensus sequence. Both of these regions are represented in Figure 4. Like the formin proline-rich domain, two of the ligands, WBP4 and WBP7, have other proline-rich regions (not shown) and may

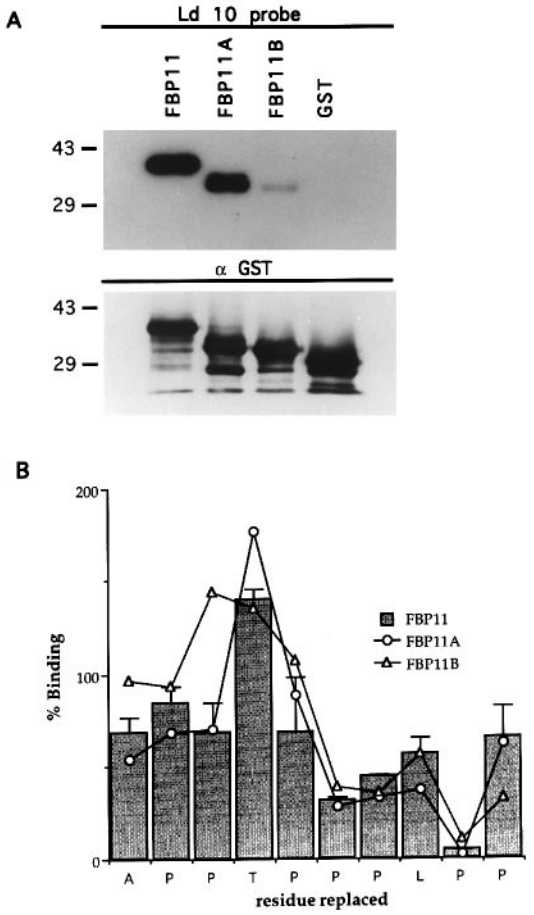

Fig. 3. Alanine-scanning mutagenesis reveals that the two WW domains of FBP11 display similar binding profiles. (**A**) Binding of GST–Ld10 to isolated FBP11 WW domains. GST fusion proteins of the two WW domains in tandem (FBP11) and individual WW domains (FBP11A and FBP11B) were purified (Kaelin *et al.*, 1992), run on a SDS–PAGE gel and blotted onto nitrocellulose. The blot was probed with radiolabeled GST–Ld10, washed and exposed to film. A duplicate blot was probed with antibodies to GST as a loading control. Molecular weight markers (in kDa) are indicated. (**B**) GST fusion proteins containing alanine in place of the indicated residue (or glycine for the first alanine) were tested for binding to radiolabeled GST–FBP11 (both domains), GST–FBP11A (N-terminal domain) and GST–FBP11B (C-terminal domain). Binding was quantitated by a phosphorimager, with wild-type binding (binding to the original APPTPPPLPP fusion protein) defined as 100%.

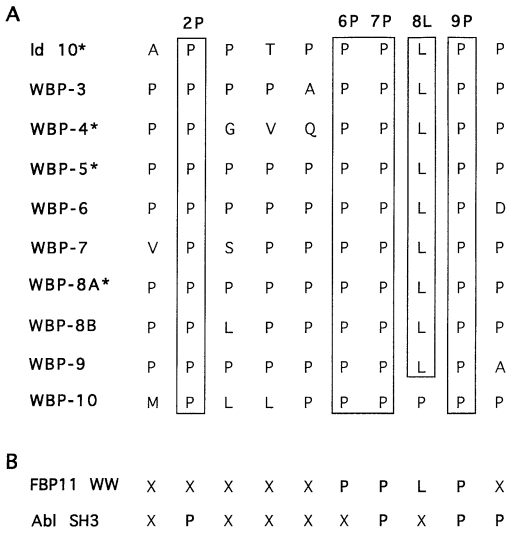

Fig. 4. Alignment of putative FBP11 binding sites in the WBPs. (**A**) The proline-rich regions of the WBPs and Ld10 were aligned using the DNASTAR program MegAlign. Amino acids that are ⩾90% conserved are blocked. The asterisk indicates sequences that were used in GST fusions to ascertain their binding capabilities (Figure 5). (**B**) The four most important residues in ld10 for FBP11 WW binding (Figure 3B) and Abl SH3 binding, as determined by an alanine scan (Chan *et al.*, 1996).

therefore possess additional binding areas. In the interest of clarity, we have depicted only the best fitting sequence from each here.

Using ld10 as a template to which the eight newly isolated ligands were aligned, we were able to identify a common motif. This motif consisted of four conserved proline residues, three of which (6P, 7P and 9P) were shown to be the most critical for WW recognition by sequential replacement of the appropriate amino acid with an alanine residue (Figure 3B). A fourth site sensitive to alanine substitution was the leucine at position eight, which was shown to be conserved in nine of the ten motifs. All other substitutions only mildly affected WW domain recognition of the ld10 ligand. Interestingly, position P2 was conserved in all ten motifs. Binding of the FBP11 WW domain was, surprisingly, unaffected by alanine replacement at this position and its conserved nature may be fortuitous or indicative of an overlapping SH3 domain binding site, as this residue is essential for Abl SH3 binding (Ren *et al.*, 1993; Chan *et al.*, 1996).

FBP11 WW domain binds an identified motif in WBP4 and WBP5

To determine if this common motif was indeed capable of interacting with the WW domain of FBP11, we subcloned the proline-rich 10 aa stretch of WBP4 and WBP5 into a pGex vector, purified fusion proteins and carried out immobilized protein-binding analysis. The WBP4 motif contains the least number of proline residues (6/10) and WBP5 the most (9/10) (see Figure 4). Both of these ligands are bound by the FBP11 WW domain to the same extent as ld10 and the Abl SH3 ligand, 3BP1 (Figure 5). The Abl SH3 domain, however, does not bind WBP4 and WBP5 as well as ld10 or 3BP1. This loss of binding is most striking in the case of WBP5, the ligand with the most proline residues. We have found that the WW domain of FBP11 binds a polyproline stretch of 10 aa (data not shown).

To evaluate the relative affinity of these proline-rich ligands to the FBP11 WW domain and the Abl SH3 domain, we made use of surface plasmon resonance. This analysis was carried out on a BIAcore Biosensor®. GST fusion proteins of three ligands (WBP4, WBP5 and 3BP1) were immobilized to the surface of different flow cells of a sensor chip. To assay for binding, first the GST–FBP11 WW domain, then the GST–Abl SH3 domain, and finally GST, was passed though each of the flow cells. The chip was regenerated with 0.1 M NaOH between each injection. The binding profiles of these interactions (Figure 6) corresponds well with the specificity results obtained using a blot overlay approach (Figure 5). The binding constant was measured for each of these six interactions. Both the SH3 domain and WW domain used in this assay were

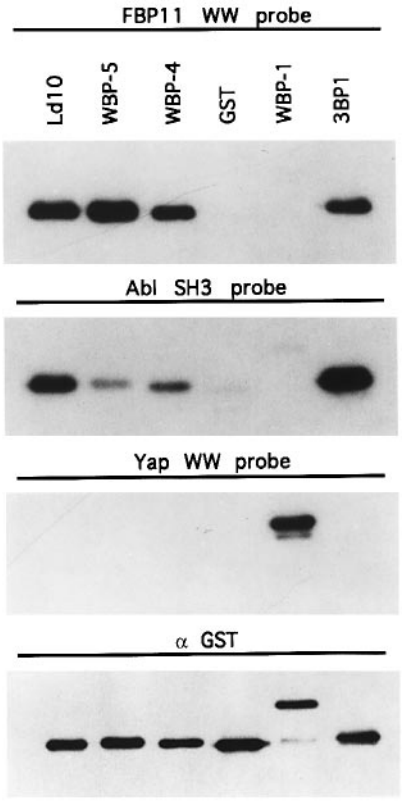

Fig. 5. Binding specificity of the PPLP motif. The indicated purified GST fusion proteins were run on a SDS–PAGE gel and blotted to nitrocellulose. Filters were probed with radiolabeled GST–FBP11 WW, radiolabeled GST–Abl SH3, radiolabeled GST–Yap WW and antibodies to GST.

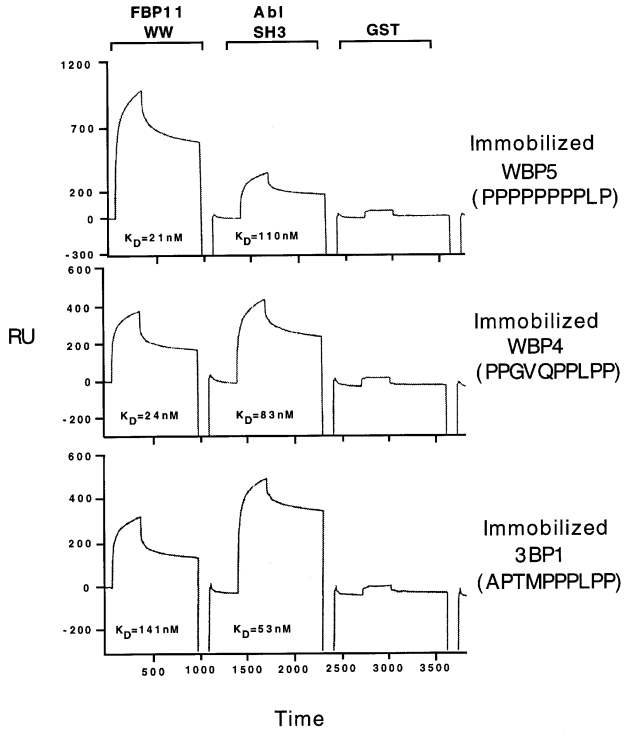

Fig. 6. Comparison of the binding profiles of three proline-rich ligands to the FBP11 WW domain and the Abl SH3 domain. 'BIAcore' technology was used to measure binding of WBP5, WBP4 and 3BP1 to both of these domains. Immobilized ligands were sequentially bound by the GST–FBP11 WW domain, the GST–Abl SH3 domain and finally GST. Between assays, the chip was regenerated with 0.1 M NaOH. Interaction between ligand and analyte is indicated by an increase in RU value. Time is measured in seconds. Amount of analyte injected in each case was: 1.1 μmol of GST–FBP11 WW domain, 1.7 μmol of GST–Abl SH3 domain and 2.1 μmol of GST. Tracings were not corrected for bulk contribution.

fused to GST. GST fusion proteins dimerize (Panayotou *et al.*, 1993) and therefore display an avidity effect. The avidity will result in the binding constant being two or more orders of magnitude lower than the absolute binding constant (Ladbury *et al.*, 1995). However, the relative binding of two different GST-fused domains to the same immobilized ligand can be determined using this approach.

The WW domains of FBP11 and YAP have different binding specificities

The YAP WW domain was shown to bind to a PPXY motif (Chen and Sudol, 1995). To determine whether the YAP WW domain would recognize the FBP11 WW domain ligands, WBP1, WBP4 and WBP5 were expressed as GST fusion proteins and duplicate nitrocellulose blots carrying these fusion proteins were probed with either ^{32}P-labeled GST–FBP11WW or GST–YAPWW. The individual WW domains display extraordinary specificity with regard to ligand recognition (Figure 5). The YAP WW domain only binds WBP1 and does not recognize WBP4 or WBP5, and WBP1 is not bound by the FBP11 WW domain. In addition, the SH3 domain of Abl was capable of recognizing FBP11 WW domain ligands, but not the Yap WW domain ligand.

Discussion

The WW domains of FBP11 and YAP do not recognize the same proline-rich motif

The FBPs were isolated by carrying out a functional screen of a limb cDNA expression library. The probe used in the screen was derived from the proline-rich region of the formin protein, which does not contain a PPXY motif (Chan *et al.*, 1996). This screening method may therefore have selected for a subset of WW domain-containing proteins which selectively bind to a PPLP rather than a PPXY motif. This notion was confirmed by the reciprocal experiment described here, in which eight ligands for FBP11 WW were isolated. None of the eight ligands isolated in this study contained a PPXY motif. We also demonstrated that the WW domains of FBP11 and Yap bind only to their respective ligands and display no cross-reactivity.

All other ligands and putative ligands for WW domains described to date contain the XPPXY motif. The WW domains of NEDD4 bind to the PPXY motifs in the epithelial sodium channel proteins. This interaction was identified by a yeast two-hybrid screen and has been implicated in Liddle's syndrome (Schild *et al.*, 1996; Staub *et al.*, 1996). A functional screen of a cDNA expression library has detected two ligands (WBP1 and WBP2) for YAP, both of which contain a PPXY motif (Chen and Sudol, 1995). The C-terminal 15 amino acids of β-dystroglycan, which contains a PPXY motif, have been shown to interact with a 217 amino acid, WW domain-containing region of dystrophin (Jung *et al.*, 1995). These results suggest that there are at least two distinct classes of WW domain-containing proteins—those that

recognize a PPXY motif, and those that recognize a PPLP motif.

The specificity of the WW domains for different ligands should be reflected in their structure. The NMR structure of the WW domain of human Yap65 in complex with the core motif PPXY has recently been determined. Macias *et al.* (1996) show that the tyrosine residue in the core motif interacts with the L(30) and H(32) residues of the hYap WW domain. All the WW domains known to interact with the PPXY core motif (hYap, NEDD-4 and dystrophin) show conservation of the H(32) residue. However, the FBPs favor no particular residue in this position. To test the role of these two residues in binding the PPXY motif, we replaced Y30 and T32 in the FBP11 WW domain with L and H, respectively. As expected, this chimeric WW domain was unable to bind the ld10 ligand (data not shown); however, nor did it gain the capacity to bind WBP1 (PPXY). Therefore, while L30 and H32 are likely to play important roles in recognizing the PPXY motif, additional residues must be involved. More extensive substitution experiments will be required to define the residues responsible for these two types of binding specificities.

FBP11 binds nuclear proteins

Unlike the docking sites for SH3-containing proteins, which tend to be involved in signal transduction pathways (Bar-Sagi *et al.*, 1993; Gout *et al.*, 1993) or serve as cytoskeletal elements (Freeman *et al.*, 1996), four of the five WBPs of known identity are nuclear proteins. In addition, FBP11 was originally isolated for its ability to bind formin, another protein that displays nuclear localization (Trumpp *et al.*, 1992; Chan and Leder, 1996).

While largely similar in their nuclear localization, the known or presumed functions of the identifiable PPLP-bearing WWB proteins are varied. Two different clones of ZMF1 (WBP4) were isolated, both of which contained the proline-rich domain. This gene harbors both a potential nuclear localization signal and a zinc finger motif (Toda *et al.*, 1994). WBP6 is the mouse homologue of the human serine/threonine protein kinase (SRPK1). SRPK1 is involved in the phosphorylation of the SR family of pre-mRNA splicing factors and in the process of intranuclear distribution of splicing factors during the cell cycle (Gui *et al.*, 1994a,b). It has been detected in nuclear splicing extracts and contains two potential nuclear-targeting signals. The WBP6 clone which we isolated is an alternatively spliced form of SRPK1 which contains a proline-rich cassette of ~27 amino acids that has been introduced adjacent to the C-terminal nuclear localization signal. The introduction of this proline-rich exon may provide a novel way of regulating WW–PPLP interactions by providing alternative isoforms.

Suggesting a potential role in transcription, WBP9 encodes the proline-rich region of the mouse α-fetoprotein enhancer-binding protein (ATBF1) (Morinaga *et al.*, 1991; Yasuda *et al.*, 1994). This protein was originally isolated for its ability to bind an AT-rich DNA sequence present in the enhancer element of the human α-fetoprotein. ATBF1 contains 17 zinc finger domains and four homeodomains; these structural features—and the ability of this protein to bind DNA—suggest that it is localized to the nucleus. Another DNA-binding protein, methyl-CpG-binding protein (MeCP2), was also bound by the FBP11 WW domain. MeCP2 binds DNA that contains a single methyl-CpG pair and was found to be associated with peri-centromeric heterochromatin (Lewis *et al.*, 1992; Nan *et al.*, 1996). The MeCP2 gene has recently been disrupted in ES cells and was shown to be essential for normal embryonic development (Tate *et al.*, 1996).

WBP8 is identical to NDPP1, a protein that displays a proline-rich N-terminus, but no other striking structural features and thus provides no clue as to its function (Sazuka *et al.*, 1992). The subcellular localization of NDPP1 is not known. Curiously, we did not isolate any clones that contained the proline-rich domain of formin (a gene known to be expressed in the embryonic limb bud, the source of the library used in these studies). This is likely due to the fact that the formins are encoded by very low-abundance transcripts and that we did not analyze all positive clones.

The Abl SH3 domain binds a subset of WBPs

Certainly, among the most interesting of our observations is that, of the eight WBPs isolated for their ability to bind the WW domain of FBP11, seven were also bound by the Abl SH3 domain. All of the WBPs bound less strongly to Abl SH3 than to FBP11 WW, except for WBP8. The other two SH3 domains assayed (Fyn and Src) only recognized WBP10. These results support our earlier findings (Chan *et al.*, 1996) that a 10 aa region of formin can be bound by both the Abl SH3 and the FBP11 WW domains, a topic recently addressed in a review by Marius Sudol (Sudol, 1996). Surface plasmon resonance experiments suggest that the relative equilibrium binding constants of FBP11 WW domain and the Abl SH3 domain, for proline-rich ligands, are similar. However, the FBP11 WW domain does bind the very proline-rich ligand (WBP5) with a higher affinity than it does WBP4 and 3BP1 (Figure 6).

Interestingly, when the proline-rich domains of the WBPs are aligned, it is obvious that the resulting consensus sequence is comprised of both an Abl SH3 domain (P2, P7, P10) (Ren *et al.*, 1993; Rickles *et al.*, 1994; Chan *et al.*, 1996; Sparks *et al.*, 1996) and an FBP11 WW domain (P6, P7, P9) (Figures 3 and 4) binding motif. Because most of the WBPs (six of eight) contain relatively short proline-rich domains of <25 amino acids, and two proline-rich binding motifs are superimposed on one another, it is tempting to speculate that these ligands could be the subject of competition by two different protein–protein interaction modules.

Materials and methods

Plasmid constructions

The generation of FBP11 WW, FBP28 WW, GST–Abl SH3, GST–Fyn SH3, GST–Src SH3, GST–Ld10 and GST–3BP1 fusion proteins has been described previously (Chan *et al.*, 1996). To produce Yap WW and WBP1 fusion proteins, 10.5-day post-coitus mouse embryo cDNA was amplified by the polymerase chain reaction (PCR) and subcloned into the pGEX2TK expression vector (Pharmacia). PCR primers for Yap WW were 5′-AGGGATCCCTGCCAGCAGG-CTGGGAGATG-3′ and 5′-TGGGATCCCAGCCTTGGGTCCAGCCAGG-3′. Both primers were flanked by *Bam*HI sites as there is an internal *Eco*RI site in the Yap WW sequence and these primers straddle the WW domains. Primers for WBP1 have been described (Chen and Sudol, 1995) and represent the N-terminal 74 amino acids. PCR primers for the N-terminal FBP11 WW

domain were 5′-CGAGGATCCTGGACAGAACATAAATCACC-3′ and 5′-AGAGAATTCTCAATCATCTGGCTTTTCCCAGG-3′. PCR primers for the C-terminal FBP11 WW domain were 5′-CGAGGATCCTGG-AAAGAGTACAAATCTGA-3′ and 5′-AGAGAATCCTCATTCCTTA-GGTTTGGCCCAGC-3′. GST–WBP5 was constructed by annealing the two oligonucleotides 5′-GATCCCCACCGCCTCCACCGCCTCCTCT-CCCACCATGAG-3′ and 5′-AATTCTCATGGTGGGAGAGGAGGCG-GTGGAGGCGGTGGG-3′ and subcloned in pGEX2TK. GST–WBP4 was constructed by annealing two oligonucleotides 5′-GATCCCCACCG-GGAGTACAACCTCCTCTCCCACCATGAG-3′ and 5′-AATTCTCAT-GGTGGGAGAGGAGGTTGTACTCCCGGTGGG-3′ and subcloned in pGEX2TK. The cloning of the panel of plasmids used in the alanine scanning mutagenesis experiment has been described elsewhere (Chan *et al.*, 1996).

Preparation of cell lysates

Two murine epithelial cell lines (Neu n and NF 639), derived from mammary gland tumors in transgenic animals, were grown to 80% confluency on 10 cm plates. After a PBS wash, the cells were lysed in 1 ml of RIPA buffer (10 mM Tris–HCl, pH 7.4, 5 mM EDTA, 300 mM NaCl) and scraped off the plate. This lysate was then briefly sonicated. For SDS–PAGE gel analysis, 10 μl of this lysate was loaded per lane.

Purification and labeling of fusion proteins

GST fusion proteins were purified as previously described (Kaelin *et al.*, 1992). Purified proteins were labeled with [γ-^{32}P]ATP using heart muscle kinase (Kaelin *et al.*, 1992). Unincorporated nucleotide was removed by purifying the sample over a NICK column (Pharmacia). For most purposes, 10–50 μg of protein were labeled; the typical specific activity was ~1×10^6 c.p.m./μg.

Expression screening

Two mouse limb libraries, one random-primed and one primed with an oligo dT primer, were screened. The construction of these libraries was described previously (Chan *et al.*, 1996). Expression libraries were infected into BL21(DE3)pLysE and plated at a density of 200 000 plaques/22 cm^2 dish. When plaques reached 0.5–1 mm in diameter, the plates were overlaid with nitrocellulose filters previously saturated with 10 mM IPTG and allowed to grow for another 6 h at 37°C. Filters were blocked for 1 h at room temperature in TBST (137 mM NaCl, 2.7 mM KCl, 25 mM Tris, pH 8.0, 0.1% Tween-20) containing 1% non-fat dry milk. Filters were then incubated in binding buffer (TBST containing 1% milk and 0.5×10^6 c.p.m of probe/ml) overnight at 4°C. Filters were washed in TBST four times, 15 min each. Strong signals were obtained after an overnight exposure with an intensifying screen. After plaque purification, positive phages were excised as plasmids by infecting into the *Cre*-carrying strain BM25.8, as recommended by the manufacturer (Novagen). The resulting cloned plasmids were then transformed into BL21(DE3)pLysE for fusion protein analysis.

Expression of λExLox–WBP fusion proteins

A single colony was used to inoculate 2 ml of LB medium and incubated with shaking at 37°C overnight. 1 ml of overnight culture was used to inoculate 10 ml of LB medium and incubated with shaking at 37°C for 1 h. IPTG was added to a final concentration of 1 mM and the incubation was continued for a further 3 h. Induced cells were collected by centrifugation and resuspended in 1 ml of a 10 mM Tris–HCl, pH 8.0, solution. This suspension was then briefly sonicated to denature the DNA. 10 μl of this protein sample was analyzed on a SDS–PAGE gel.

Binding assays

The binding assays were performed using a BIAcore Biosensor®. All GST fusion proteins used in these assays were dialyzed to remove glutathione, then quantitated using the Bradford assay. The ligands were immobilized to the surface of different flow cells of a sensor chip by amine coupling as previously described (Rickles *et al.*, 1994); 480 resonance units (RU) of GST–3BP1, 890 RU of GST–WBP4 and 880 RU of GST–WBP5 were covalently linked to the chip in this manner. The injection and regeneration strategies used here are identical to those used in the SH3 binding study carried out by Rickles *et al.* (1994). The binding constants were determined by measuring the association and dissociation sensorgrams at five different analyte concentrations. The association rate constant (K_a) was determined by plotting the K_s value against the concentration. The slope of the graph equals the K_a. The dissociation kinetics for each of the interactions was obtained from the same sensorgrams, the average dissociation rate constant (K_d) for the five different concentrations was used in the calculation of the equilibrium binding constant (K_D). The K_s and K_d values were calculated using BIAevaluation 2.0® software.

Blot overlay assay

Protein samples were run on SDS–PAGE gels, and transferred onto an Immobilon-P membrane (Millipore) by semi-dry electroblotting (Harlow and Lane, 1988). The blots were blocked for 1 h at room temperature in TBST containing 1% non-fat dry milk. Blots were then incubated in binding buffer (TBST containing 1% milk and 0.5×10^6 c.p.m of probe/ml) overnight at 4°C. Filters were washed in TBST four times, 15 min each, and exposed to film.

Alanine scanning mutagenesis

The panel of alanine-mutated clones has been previously described (Chan *et al.*, 1996). Binding of the fusion proteins using the blot overlay assay was quantitated by a phosphorimager. To correct for protein loading, duplicate blots were subjected to quantitative Western blotting.

Accession numbers

Sequences of WBP3, WBP5, WBP7 and the WBP6 (SRPK-1) splice variant submitted to the DDBJ/EMBL/GenBank Data Library. Accession numbers are: U92453, U92454, U92455 and U92456 respectively.

Acknowledgements

We thank Ella Bedford of assisting in the generation of the BIAcore data. We are grateful to members of the Leder laboratory for advice in the preparation of this manuscript. M.T.B. is a Howard Hughes postdoctoral fellow. D.C.C. was supported in part by a NIH Medical Scientist Training Program fellowship.

References

Andre,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. *Biochem. Biophys. Res. Commun.*, **205**, 1201–1205.

Bar-Sagi,D., Rotin,D., Batzer,A., Mandiyan,V. and Schlessinger,J. (1993) SH3 domains direct cellular localization of signaling molecules. *Cell*, **74**, 83–91.

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

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

Cahir McFarland,E.D. and Thomas,M.L. (1995) CD45 protein-tyrosine phosphatase associates with the WW domain-containing protein, CD45AP, through the transmembrane region. *J. Biol. Chem.*, **270**, 28103–28107.

Chan,D.C. and Leder,P. (1996) Genetic evidence that formins function within the nucleus. *J. Biol. Chem.*, **271**, 23472–23477.

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.

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.

Einbond,A. and Sudol,M. (1996) Towards prediction of cognate complexes between the WW domain and proline-rich ligands. *FEBS Lett.*, **384**, 1–8.

Ferguson,K.M., Lemmon,M.A., Sigler,P.B. and Schlessinger,J. (1995) Scratching the surface with the PH domain. *Nature Struct. Biol.*, **2**, 715–718.

Fiore,F., Zambrano,N., Minopoli,G., Donini,V., Duilio,A. and Russo,T. (1995) The regions of the Fe65 protein homologous to the phosphotyrosine interaction/phosphotyrosine binding domain of Shc bind the intracellular domain of the Alzheimer's amyloid precursor protein. *J. Biol. Chem.*, **270**, 30853–30856.

Freeman,N.L., Lila,T., Mintzer,K.A., Chen,Z., Pahk,A.J., Ren,R., Drubin,D.G. and Field,J. (1996) A conserved proline-rich region of the *Saccharomyces cerevisiae* cyclase-associated protein binds SH3 domains and modulates cytoskeletal localization. *Mol. Cell. Biol.*, **16**, 548–556.

Gout,I. *et al.* (1993) The GTPase dynamin binds to and is activated by a subset of SH3 domains. *Cell*, **75**, 25–36.

Gui,J.F., Tronchere,H., Chandler,S.D. and Fu,X.D. (1994a) Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. *Proc. Natl Acad. Sci. USA*, **91**, 10824–10828.

Gui,J.F., Lane,W.S. and Fu,X.D. (1994b) A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature*, **369**, 678–682.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hofmann,K. and Bucher,P. (1995) The rsp5-domain is shared by proteins of diverse function. *FEBS Lett.*, **358**, 153–157.

Jung,D., Yang,B., Meyer,J., Chamberlain,J.S. and Campbell,K.P. (1995) Identification and characterization of the dystrophin anchoring site on beta-dystroglycan. *J. Biol. Chem.*, **270**, 27305–27310.

Kaelin,W.G. *et al.* (1992) Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. *Cell*, **70**, 351–364.

Kao,H. and Siliciano,P.G. (1996) Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.*, **16**, 960–967.

Ladbury,J.E., Lemmon,M.A., Zhou,M., Green,J., Botfield,M.C. and Schlessinger,J. (1995) Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. *Proc. Natl Acad. Sci. USA*, **92**, 3199–3230.

Lewis,J.D., Meehan,R.R., Henzel,W.J., Maurer-Fogy,I., Jeppesen,P., Klein,F. and Bird,A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905–914.

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

Macias,M.J., Hyvonen,M., Baraldi,E., Schultz,J., Sudol,M., Saraste,M. and Oschkinat,H. (1996) The structure of the WW domain in complex with a proline-rich peptide. *Nature*, **382**, 646–649.

Maleszka,R., Hanes,S.D., Hackett,R.L., de Couet,H.G. and Miklos,G.L. (1996) The *Drosophila melanogaster* dodo (dod) gene, conserved in humans, is functionally interchangeable with the ESS1 cell division gene of *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **93**, 447–451.

Marengere,L.E. and Pawson,T. (1994) Structure and function of SH2 domains. *J. Cell Sci. Suppl.*, **18**, 97–104.

Morinaga,T., Yasuda,H., Hashimoto,T., Higashio,K. and Tamaoki,T. (1991) A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeodomains and seventeen zinc fingers. *Mol. Cell. Biol.*, **11**, 6041–6049.

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

Nan,X., Tate,P., Li,E. and Bird,A. (1996) DNA methylation specifies chromosomal localization of MeCP2. *Mol. Cell. Biol.*, **16**, 414–421.

Panayotou,G., Gish,G., End,P., Truong,O., Gout,I., Dhand,R., Fry,M.J., Hiles,I., Pawson,T. and Waterfield,M.D. (1993) Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. *Mol. Cell. Biol.*, **13**, 3567–3576.

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

Rickles,R.J., Botfield,M.C., Weng,Z., Taylor,J.A., Green,O.M., Brugge,J.S. and Zoller,M.J. (1994) Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *EMBO J.*, **13**, 5598–5604.

Sazuka,T., Tomooka,Y., Kathju,S., Ikawa,Y., Noda,M. and Kumar,S. (1992) Identification of a developmentally regulated gene in the mouse central nervous system which encodes a novel proline rich protein. *Biochim. Biophys. Acta*, **1132**, 240–248.

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.

Sparks,A.B., Rider,J.E., Hoffman,N.G., Fowlkes,D.M., Quilliam,L.A. and Kay,B.K. (1996) Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Contactin, p53bp2, PLCγ, Crk, and Grab2. *Proc. Natl Acad. Sci. USA*, **93**, 1540–1544.

Staub,O. and Rotin,D. (1996) WW domains. *Structure*, **4**, 495–499.

Staub,O., Dho,S., Henry,P.C., 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.

Sudol,M. (1996) 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.

Tate,P., Skarnes,W. and Bird,A. (1996) The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nature Genet.*, **12**, 205–208.

Toda,T., Iida,A., Miwa,T., Nakamura,Y. and Imai,T. (1994) Isolation and characterization of a novel gene encoding nuclear protein at a locus (D11S636) tightly linked to multiple endocrine neoplasia type 1 (MEN1). *Hum. Mol. Genet.*, **3**, 465–470.

Trumpp,A., Blundell,P.A., de la Pompa,J.L. and Zeller,R. (1992) The chicken limb deformity gene encodes nuclear proteins expressed in specific cell types during morphogenesis. *Genes Dev.*, **6**, 14–28.

Woychik,R.P., Stewart,T.A., Davis,L.G., D'Eustachio,P. and Leder,P. (1985) An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature*, **318**, 36–40.

Woychik,R.P., Maas,R.L., Zeller,R., Vogt,T.F. and Leder,P. (1990) 'Formins': proteins deduced from the alternative transcripts of the limb deformity gene. *Nature*, **346**, 850–853.

Yasuda,H., Mizuno,A., Tamaoki,T. and Morinaga,T. (1994) ATBF1, a multiple-homeodomain zinc finger protein, selectively down-regulates AT-rich elements of the human alpha-fetoprotein gene. *Mol. Cell. Biol.*, **14**, 1395–1401.

Zheng,J., Cahill,S.M., Lemmon,M.A., Fushman,D., Schlessinger,J. and Cowburn,D. (1996) 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.

Received on October 4, 1996; revised on December 20, 1996

Note added in proof

WBP4 was recently shown to be the splicing factor SF1 [Arning,S., Grüter,P., Bilbe,G. and Krämer,A. (1996) Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. *RNA*, **2**, 794–810].