THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 14, Issue of April 7, pp. 10359–10369, 2000

Printed in U.S.A.

A Novel Pro-Arg Motif Recognized by WW Domains*

(Received for publication, November 12, 1999, and in revised form, January 7, 2000)

Mark T. Bedford‡§, Dilara Sarbassova¶, Jian Xu¶, Philip Leder‡, and Michael B. Yaffe¶‖****

From the ‡Department of Genetics, Harvard Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115 and the ¶Division of Signal Transduction, Department of Medicine and ‖Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine, Boston, Massachusetts 02215

WW domains mediate protein-protein interactions through binding to short proline-rich sequences. Two distinct sequence motifs, PP*X*Y and PPLP, are recognized by different classes of WW domains, and another class binds to phospho-Ser-Pro sequences. We now describe a novel Pro-Arg sequence motif recognized by a different class of WW domains using data from oriented peptide library screening, expression cloning, and *in vitro* binding experiments. The prototype member of this group is the WW domain of formin-binding protein 30 (FBP30), a p53-regulated molecule whose WW domains bind to Pro-Arg-rich cellular proteins. This new Pro-Arg sequence motif re-classifies the organization of WW domains based on ligand specificity, and the Pro-Arg class now includes the WW domains of FBP21 and FE65. A structural model is presented which rationalizes the distinct motifs selected by the WW domains of YAP, Pin1, and FBP30. The Pro-Arg motif identified for WW domains often overlaps with SH3 domain motifs within protein sequences, suggesting that the same extended proline-rich sequence could form discrete SH3 or WW domain complexes to transduce distinct cellular signals.

Many signaling events in eukaryotic cells involve the assembly of large protein-protein complexes. These diverse associations are mediated through interactions of a limited number of modular signaling domains such as the SH2,[1] SH3, PDZ, and PTB domains and their respective ligands (1, 2). Particular families of domains occasionally share affinity for related ligands. For example, both SH2 and PTB domains can bind phosphotyrosine-containing substrates, while FHA domains, 14-3-3 proteins, and certain WW domains bind phosphoserine/phosphothreonine-containing proteins (3). Recently, several unrelated signaling modules have been identified that specifically recognize proline-rich ligands. These include WW and EVH1 domains in addition to the canonical member of this group, the SH3 domain (4–7). WW domains are encoded as 38–40-amino acid long modules in single copy or tandem repeats in over 25 signaling proteins. WW domain-containing proteins include those implicated directly or indirectly in a variety of human diseases such as Liddle's syndrome, Duchenne muscular dystrophy, Huntington's disease, and Alzheimer's disease (6, 8, 9). Named after two highly conserved tryptophan residues characteristically spaced either 22 or 23 residues apart, WW domains participate in a variety of cellular processes, including ubiquitin-mediated protein degradation (10, 11), viral budding (12, 13), RNA splicing (14, 15), transcriptional co-activation (16–18), and mitotic regulation (19, 20). Recently another WW domain-containing protein, FBP30, was identified as one of the major genes up-regulated by p53 in thymocytes undergoing programmed cell death following γ-irradiation (21). Despite their general importance in normal and disease states, most of the molecular targets recognized by different WW domains *in vivo* remain unidentified.

Sudol and colleagues (22) have proposed a classification of WW domains based on their proline-based ligand specificity. In this scheme, Group I WW domains recognize and bind the motif PP*X*Y, as typified by the prototype WW domain in Yes-associated protein (YAP), whereas Group II WW domains bind a PPLP motif, as observed for the WW domains in formin-binding protein (FBP)-11. A third class of WW domains bind sequences rich in Pro, Gly, and Met residues, although the exact amino acid sequence motif they recognize is not known. A fourth class of WW domain-containing proteins has recently been found that appears to bind Pro residues preceded by phosphoserine (23).

In the present study we use oriented peptide library screening, phage-expression cloning, peptide binding studies, and protein affinity interactions in cell lysates to identify a new motif that defines and re-classifies a novel group of WW domains that specifically recognize and bind to proline-arginine-rich motifs. The prototype member of this group is the first of two WW domains present within the protein FBP30, the sequence of which is reported here. Other members of the Pro-Arg motif-binding class include the WW domains of FBP21 and FE65 and the second WW domain of FBP30. We present a structural model for WW domains that rationalizes Pro-Arg motif selection, and show that many ligands for this novel class of WW domains also bind to the Src- and phosphoinositol 3-kinase-like family of SH3 domains.

MATERIALS AND METHODS

Cloning of FBP30—A partial cDNA for FBP30 containing the first WW domain (FBP30 WW-A) was obtained by screening a mouse random-primed λ cDNA expression library for formin-binding proteins, as described previously (24). To obtain additional FBP30 sequence, a mouse limb bud cDNA expression library (24) and a NIH3T3 cDNA

* This work was supported in part by National Institutes of Health Grants HL03601 (to M. B. Y.) and GM56203 (to Lewis C. Cantley), a Harvard Medical School/Affiliated Hospital Collaborative Seed Grant (to M. B. Y. and P. L.), and a Burroughs-Wellcome Career Development Award (to M. B. Y.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The sequence of FBP30 has been submitted to the GenBank™ data base with the accession number AF237960. Sequences of the WW domain binding proteins identified in this study have been submitted to the GenBank data base with accession numbers AF073932 (WBP12), AF073933 (WBP14), AF073934 (WBP15), AF073935 (WBP16), and AF073936 (WBP17).

§ Supported by the Cancer Research Fund of the Damon Runyan-Walter Winchell Foundation.

** To whom correspondence should be addressed. Tel.: 617-667-0908; Fax: 617-667-0957; E-mail: myaffe@bidmc.harvard.edu.

[1] The abbreviations used are: SH2, Src homology domain 2; FBP, formin-binding protein; YAP, Yes-associated protein; PCR, polymerase chain reaction; GST, gluathione *S*-transferase; PAGE, polyacrylamide gel electrophoresis; WBP, WW domain-binding protein(s).

library cloned into λZap (courtesy of Ari Elson) were screened with the partial cDNA insert of FBP30 obtained above. 5′ Rapid amplification of cDNA ends was employed to obtain a putative start codon using a mouse 11-day embryo library (CLONTECH, Palo Alto, CA) and the primers: 5′-gcagtcgaatctgcctcgggatccgg-3′ (first round of PCR) and 5′-cccggtggagagagctgcaggatagg- 3′ (nested PCR). Two expressed sequence tag clones (GenBank accession numbers AA240645 and AA175408) were sequenced to obtain the stop codon and the poly(A) tail.

Construction, Purification, and Labeling of Fusion Proteins—GST fusion proteins containing the isolated FBP30 WW-A, WW-B, or a fusion of the WW-A and -B domains were generated by PCR and subcloned into pGex-2TK. PCR primers for FBP30 WW-A were 5′-atgggatcccattatgacactcagtgttga-3′ and 5′-gctgaattcagcaaggtattggggta-actc-3′, flanked by *Bam*HI and *Eco*RI sites. PCR primers for FBP30 WW-B were 5′- cttggatccaagcttcaggatgctgcagaa-3′ and 5′-tagggatc-cctacttcacggtcagcttaggcag-3′, flanked by *Bam*HI sites. The resulting PCR fragment also contains an *Eco*RI and a *Hin*dIII site. The fusion of WW-A and -B domains used the PCR primers 5′-atgaagcttcattatgacac-tcagtgttga-3′ and 5′-gctaagcttagcaaggtattggggtaactc-3′ flanked by *Hin*dIII sites for FBP30 WW-A and the resulting PCR fragment was then cloned into the *Hin*dIII site of FBP30 WW-B. The amino acid sequence of FBP30 WW-B fused to GST is: KRKLQDAA-EQLKQYEINATPKGWSCHWDRDHRRYFYVNEQSGESQWEFPDG-EEEEESQTKEVRDESLPKLTVK. The amino acid sequence of FBP30 WW-A&B fused to GST is: KLHYDTQCSLAGVEIEMGDG-DWQEVWDENTGCYYYWNTQTNEVTWELPQYLAKLQDAAEQ-LKQYEINATPKGWSCHWDRDHRRYFYVNEQSGESQWEF*P*DG-EEEEESQTKEVRDESLPKLTVK.

The construction of GST fusions containing the SH3 domains from the p85 subunit of phosphoinositol 3-kinase and Fyn has been described previously (24, 25). Plasmids encoding GST fusion proteins of the YAP and FE65 WW domains were the kind gifts of Dr. Marius Sudol.

GST fusions were expressed in bacteria and purified as described previously (26). For peptide binding experiments, fusion proteins were eluted from the beads by incubation with 20 mM glutathione in 100 mM HEPES, pH 7.4, for 30 min at 4 °C. For blot overlay assays, the purified proteins were labeled with [γ-^{32}P]ATP using heart muscle kinase (26). Unincorporated nucleotide was removed by chromatography on NICK columns (Amersham Pharmacia Biotech). In most experiments, 10–50 μg of protein were labeled, with typical specific activities of 1×10^6 cpm/μg.

Peptide Library Screening—Proline, tyrosine, and proline/arginine-oriented degenerate peptide libraries were synthesized using N^{α}-Fmoc-protected amino acids and standard BOP/HOBt coupling chemistry. The following libraries were used: Pro, Met-Ala-*X*-*X*-*X*-*X*-Pro-*X*-*X*-*X*-*X*-Ala-Lys-Lys, theoretical degeneracy = 1.7×10^{10}; SH3: Met-Ala-*X*-*X*-*X*-*X*-Pro-*X*-*X*-Pro-*X*-*X*-*X*-Ala-Lys-Lys-Lys, theoretical degeneracy = 2.0×10^{11}; Pro-Arg, Met-Ala-*X*-*X*-*X*-*X*-Pro-Pro-Arg-*X*-*X*-*X*-*X*-Ala-Lys-Lys-Lys, theoretical degeneracy = 1.7×10^{10}; and Tyr, Met-Ala-*X*-*X*-*X*-*X*-Tyr-*X*-*X*-*X*-*X*-Ala-Lys-Lys-Lys, theoretical degeneracy = 1.1×10^{10}. *X* represents all amino acids except cysteine. In addition, Trp was omitted from the degenerate positions in the SH3 library, and Tyr was omitted from degenerate positions in the Tyr library. Specific individual peptides for use in fluorescence polarization-based binding assays were synthesized in an identical manner, and the final product purified by reverse phase high performance liquid chromatography following deprotection.

Peptide library screening was performed using 100 μl of glutathione beads containing saturating amounts of each WW domain-GST fusion protein (~1–1.5 mg) as described previously (19, 27). Briefly, beads were packed in small columns and incubated with 0.5 mg of the peptide library mixture for 10 min at room temperature in 50 mM HEPES, pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 1% Triton X-100, 5–10% glycerol, 1 mM Na_3VO_4, and 1 mM dithiothreitol. Unbound peptides were removed by rapidly washing the columns twice with the same buffer followed by two washes with phosphate-buffered saline. Bound peptides were eluted with 30% acetic acid for 10 min at room temperature, dried overnight on a Speed-Vac apparatus, resuspended in H_2O, and sequenced on an Applied Biosystems model 477A or Procise protein sequencer. Preference values for amino acid selection were determined by comparing the relative abundance of each amino acid at a particular cycle of Edman sequencing (*i.e.* mole percentage) in the recovered peptides to that of each amino acid in the original peptide library mixture at the same position.

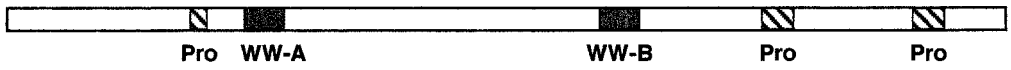
Phage Library Expression Screening—Expression screening was performed using a random-primed mouse limb cDNA library in λ phage, whose construction has been described previously (24). The expression library was infected into BL21(DE3)pLysE and plated at a density of 200,000 plaques/22-cm^2 dish. When plaques reached 0.5–1 mm in diameter, the plates were overlaid with nitrocellulose filters previously saturated with 10 mM isopropyl-1-thio-β-D-galactopyranoside and allowed to grow for another 6 h at 37 °C. Filters were blocked for 1 h at room temperature in TBS-T (137 mM NaCl, 2.7 mM KCl, 25 mM Tris, pH 8.0, 0.1% Tween 20) containing 1% non-fat dry milk. The filters were then incubated in TBS-T containing 1% milk and 0.5×10^6 cpm of ^{32}P-labeled probe/ml overnight at 4 °C, and then washed 4 times for 15 min each. Autoradiography gave strong signals after an overnight exposure with an intensifying screen. After plaque purification, positive phage 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.

Fluorescence Polarization Assays—For measurements of peptide binding affinity, FBP30 WW-A and YAP WW-GST fusion proteins were exchanged into phosphate-buffered saline using NAP-10 Sephadex G-25 columns (Amersham Pharmacia Biotech). Protein concentrations were determined by BCA assay (Pierce) using BSA as standard. Fluorescence polarization anisotropy was measured using a Pan Vera Beacon 2000 Variable Temperature Fluorescence Polarization System. Low fluorescence buffers and reagents (Pan Vera Corp.) were used throughout. WW domain proteins were serially diluted (0–171 μM) to a final volume of 150 μl in phosphate-buffered saline in 6 × 50-mm borosilicate glass tubes. Fluorescein-labeled peptides were added (135 nM final concentration), mixed, and fluorescence polarization measured at 22 °C after a 120-s delay with a 16-s integration. A time course of binding showed peak saturation within 120 s. Background fluorescence was measured for each sample prior to peptide addition. At least three binding curves were measured for each of the WW domains to their optimal peptides. Binding data was analyzed by assuming that fluorescence polarization anisotropy was a linear function of ligand binding, and that each WW domain contained a single peptide-binding site as verified by Scatchard analysis. Curves were fit to the equation, $L_B/L_{tot} = L_f/(k_D + L_f)$, where L_B is bound ligand, L_{tot} is total ligand, L_f is free ligand, and k_D is the dissociation constant in closed form using nonlinear regression analysis (Kaleidograph). Control experiments showed negligible binding of GST alone to either peptide even at 100 μM protein concentration.

Cell Culture and in Vitro Binding Assays—PC12 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% horse serum, 5% fetal calf serum, and antibiotics (50 units/ml penicillin and 50 μg/ml streptomycin). For radiolabeling, 100-cm^2 dishes of 70% confluent cells were switched to 3 ml/plate of 97.5% methionine-free Dulbecco's modified Eagle's medium containing 2% fetal calf serum and 100 μCi of Trans^{35}S-label (NEN Life Science Products) for 6 h. Cells were lysed by the addition of 1 ml of Tris-buffered saline, pH 7.4, containing 1 mM $MgCl_2$, 1 mM $CaCl_2$, 4 μg/ml each of AEBSF, leupeptin, and pepstatin-A, 8 μg/ml aprotinin, 5 mM sodium orthovanadate, 25 mM sodium fluoride, 10 mM β-glycerophosphate, 1 mM dithiothreitol, 1% Nonidet P-40, and 10% glycerol for 15–30 min at 4 °C, followed by centrifugation at 12,000 × *g* for 15 min at 4 °C in a microcentrifuge. Lysates were precleared by incubation with beads containing bound GST for ~3 h at 4 °C, and the supernatants then incubated with beads containing equal amounts of immobilized GST WW domain fusion proteins or GST alone overnight at 4 °C with rocking, in the presence or absence of 100 μM competing peptides. A typical experiment involved 500 μl of lysate incubated with 25 μl of beads containing ~70 μg of fusion protein. The beads were washed three times with phosphate-buffered saline + 0.5% Nonidet P-40, resuspended in sample buffer, and analyzed by SDS-PAGE and autoradiography. 293 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, and antibiotics (50 units/ml penicillin and 50 μg/ml streptomycin) for immunological studies.

Blot Overlay Assays—WW domain-binding proteins isolated by expression cloning were prepared for blot overlay assays by diluting overnight cultures of bacteria containing λExLox-WBP fusion protein plasmids 1:10 into 10 ml of fresh antibiotic-containing medium. Samples were incubated at 37 °C for 1 h, and protein expression induced by addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside for 3 h. Cells were pelleted, resuspended in 1 ml of 10 mM Tris-HCl, pH 8.0, and sonicated. Lysates (10 μl) were analyzed by SDS-PAGE followed by semi-dry transfer onto Immobilon-P membrane (Millipore). Blots were blocked for 1 h at room temperature in TBS-T containing 1% non-fat dry milk, then incubated TBS-T containing 1% milk containing 0.5×10^6 cpm of ^{32}P-labeled probe/ml overnight at 4 °C. Filters were washed in TBS-T 4 times for 15 min each and exposed to film. A similar analysis was performed on proteins isolated from unlabeled PC12 cells by affin-

A



B

```
MRKCTAPVKSGSPREAVPRFMNIDYAPGRCAPAPLICRRSLLALRLMMGKKSRAVPGRRPILQLSPPGPRSS

TPGRDPDPDPDPDPEADSTAAATSQSAPAAATAAAATSPAVPASAAPEDSPSEDEQEVVVEVPNVVQNPPTPVM
                                        -------T--------------------------K---

TTRPTAVKATGGLCLLGAYADSDDDESDVSEKTAQSKEANGNQATDIDSTLANFLAEIDAITAPQPAAPVVA
P----T-------------------------------------------------------------------

SAPPPTPPRPEPKEAATPALSPTASNGSDTAQTPGWHYDTQCSLAGVEIEMGDWQEVWDENTGCYYYWNTQT
------------------------------------------------------------------------

NEVTWELPQYLATQVQGLQHYQPSSVTGTEAAFVVNTDMYTKERTTAASSSKSGPVITKREVKKEVNEGIQA
------------------------------------------------------------------------

LSNSEEERKGVAAALLAPLLPEGVKEEEERWRRKVICKEADPVSETKETSTASEETGPSIKPPEVMMDGTED
------------------------------------------------------------------------

PSQEELCSVVQSGESEEEEEEEEQDTLELELALERKKAELRALEEGDGSVSGSSPRSDISQPASQDGVRRIM
------------------------------------------------------------------------

SKRGKWKMFVRATSPESTSRSSSKTGRDSPENGETAIGAEDSEKIDEISDKETEVEESSEKIKVQLAPKVEE
------------------------------------------------------------------------

EQDLKFQIGELANTLTSKFEFLGINRQSISNFHMLLLQTETRIADWREGALNGNYLKRKLQDAAEQLKQYEI
------------------------------------------------------------------------

NATPKGWSCHWDRDHRRYFYVNEQSGESQWEFPDGEEEEESQTKEVRDESLPKLTVKDKTCTDPNSTESSEN
-------------TSGSNYS

PTGSLCKESFSGQVSSSLMPLTPFWTLLQSNVPVLQPPLPLEMPPPPPPPPESPPPPPPPPPPPPLEDGEI

QEVEMEDEGSEEPPAPGTEEDTPLKPSTQTTAVTSQSLVDSTASSPPSNKAVKRKAPEMSTSVVQRSATIGS

SPVLYSQSAIAAGHQAVGMAHQAVGMAHQAVSASHAAAAGVGHQARGMSLQSNYLGLAAAPALMSYAECSVP

IGVTTPSLQPAQARGTMAAPAVVEPPPPPPPPPTPTPPPPPPAPKVPPPEKTRKGKKDKAKKSKTKMPSLVK

KWQSIQRELDEEDNSSSSEEDRESTAQKRIEEWKQQQLVSGLAERNANFEALPEDWRARLKRRKMAPST.
```

FIG. 1. **The sequence of FBP30.** *A,* schematic diagram of the domain structure of FBP30. The WW domains (*solid shading*) and Pro-rich regions (*hatched shading*) are indicated. *B,* the amino acid sequence of FBP30. Sequences of the WW domains are shown in *bold*, and Pro-rich regions are *underlined*. The partial murine FBP30 sequence reported by Depraetere and Golstein (21) is shown below, where minus (−) indicates amino acid identity and discrepant amino acids are explicitly indicated.

ity purification on beads containing immobilized GST-FBP30 WW-A (above) except that the hybridization solution contained 1×10^6 cpm of ^{32}P-labeled probe/ml.

Molecular Modeling—Atomic coordinates for the NMR structure of the YAP WW domain (kindly provided by H. Oschkinat) and the high resolution x-ray structure of the Pin1 WW domain (PDB accession code 1PIN (43)) were used for modeling the FBP30 WW-A domain. Threading was optimized using SwissPdbViewer 3.1 (28). Modeling was performed using ProModII (28, 29) and energy minimized with Gromos96 using 200 cycles of steepest descent followed by 300 cycles of conjugate gradient minimization. The final structure demonstrated excellent backbone and side chain superpositions with the base model, and yielded nearly identical overall structures regardless of whether YAP or Pin1 was used as the starting model. All ϕ and φ angles fall within acceptable regions of the Ramachandran plot. Molecular surfaces were generated and displayed and electrostatic potentials calculated using GRASP (30). For display, structures were rendered using PovRay.

RESULTS

Cloning of FBP30—A number of WW domain containing proteins have been identified based on their *in vitro* ability to interact with formin, a morphogenic protein involved in limb and kidney development. Collectively referred to as FBPs, this group consists of four members which contain multiple WW domains (FBP11, 21, 23/28, and 30) that were identified by screening a mouse limb expression library for proteins that interacted with the formin homology 1 domain (FH1 or proline-rich domain) (24). The initial FBP30 clone identified in the screen contained a single WW domain near its amino terminus, denoted FBP30 WW-A (Fig. 1A; accession number AAC52478). A fusion protein containing this WW domain linked to the COOH terminus of glutathione *S*-transferase (GST) was seen to interact with a wide variety of proteins in PC12 and 293 cell lysates (*cf.* Fig. 5).

Northern blot analysis of poly(A) selected RNA revealed that the FBP30 transcript is ubiquitously expressed although it is most abundant in spleen and thymus (data not shown and Ref. 21) and that it co-migrates with the 28 S ribosomal RNA. Extension cloning was subsequently performed to obtain additional FBP30 coding sequence based on the initial WW-containing clone that had been identified by screening for formin-binding proteins. The sequence shown in Fig. 1*B* was obtained by combining sequence data from mouse limb bud and NIH3T3 cDNA library screening, 5′-rapid amplification of cDNA ends to obtain a putative start site and two expressed sequence tag clones (GenBank accession numbers AA240645 and AA175408) to obtain the stop codon and the poly(A) tail. The resulting 4748-base cDNA codes for a protein of 1077 amino acids that contain a second WW domain, denoted FBP30 WW-B, 385 amino acids COOH-terminal to the previously identified WW domain as well as 3 proline-rich regions (Fig. 1, *A* and *B*). The size of the cDNA (4748 bases) corresponds well with the size of the FBP30 transcript seen by Northern blot analysis, the mouse 28 S ribosomal RNA is 4700 bases in size (21, 31). It is important to note that we have not yet independently confirmed that our cDNA contains the initiating methionine at the extreme 5′ end, although no additional sequence was obtained despite multiple attempts at 5′-rapid amplification of cDNA ends. Depraetere and Golstein (21) recently published a partial sequence of FBP30 containing the first proline-rich region and WW-A domain based on their finding that FBP30 is a major

p53-regulated protein in lethally irradiated thymocytes as detected by Representational Difference Analysis. A Pfam search (32) of the entire FBP30 sequence failed to reveal any additional protein-interaction domains other than the two WW domains.

The WW Domains of FBP30 Belong to a Class of WW Domains That Recognize the Core Motif Pro-Pro-Arg in Oriented Peptide Library Screening—To investigate the amino acid sequence recognized by the FBP30 WW-A domain and compare this with the sequence motif that binds to the WW domain from YAP (33, 34), we screened GST fusions of each WW domain using a peptide library oriented on a fixed proline residue. As shown in Fig. 2*A*, the FBP30 WW-A domain selected an additional Pro or an Arg residues in the Pro^{-1} position, and showed strong selection for Arg in the Pro^{+1} position. This finding suggests the new motif, (P/R)PR, as optimal for FBP30 WW-A domain recognition. In contrast, the YAP WW domain selected Leu or Pro in the Pro^{-1} position, weakly selected for Pro in the Pro^{+1} position, and showed strong selection for Tyr in the Pro^{+2} position. This (L/P)PPY motif deduced for the YAP WW domain by peptide library screening is in excellent agreement with the PP*X*Y motif previously obtained for the YAP WW domain by alanine scanning mutagenesis and single peptide binding studies (33–35). In contrast, the Pro- and Arg-based motif we obtained for the FBP30 WW domain appears novel.

In order to learn whether this Pro-Arg motif was common to other WW domains, GST fusion proteins containing the WW domains from FBP21, FE65 (36, 37), and the first and third WW domains from NEDD4 (10, 38) were screened using the Pro-oriented peptide library (data not shown). Like the FBP30 WW-A domain, both the FBP21 and FE65 WW domains displayed strong selection for Arg in the Pro^{+1} position (Table I). In contrast, no Arg selection was observed for either of the NEDD WW domains (data not shown).

The second WW domain within FBP30, FBP30 WW-B was also investigated. Like the FBP30 WW-A domain, the WW-B domain also selected for Pro-Arg based peptides, as did a protein containing a GST fusion of both the FBP30 WW-A and WW-B domains (Table I). All of the WW domains that selected for Arg in the Pro^{+1} position also showed moderate selectivity for an additional Pro in the Pro^{-1} position, as well as faint selection for aromatic residues or Pro in the Pro^{-2} position. A crude estimate of binding affinity based on the amounts of total peptide library retained by the GST WW fusion proteins suggested that the FBP30 WW-A, FE65, and FBP21 WW domains bound to peptides at least 4-fold more tightly than the FBP30 WW-B domain.

The other well characterized family of modular signaling domains that bind to Pro-rich peptides containing Arg residues are SH3 domains. In order to contrast the sequence specificity for WW domain binding with the Pro-based peptide specificity selected by SH3 domains, we examined the binding of these WW domains to a Pro-*X*-*X*-Pro oriented peptide library biased for the P*XX*P consensus motif recognized by SH3 domains (39). As shown in the *left panel* of Fig. 2*B*, the FBP30A WW domain appeared to orient on the first fixed Pro residue, selecting for Pro or Ile in the Pro^{-1} position, Arg in the Pro^{+1} position, and Pro or Ala in the Pro^{+2} position, suggesting a PPR or PPRPP-based motif. Similar selection for this core PPR motif was also seen when the WW domains from FBP21 and FE65 were screened with this P*XX*P library (Table I). The FE65 WW domain demonstrated additional selection for Phe and aliphatic side chain containing residues in the Pro^{-2} position. In contrast to the PPR motif we identified for these FBPs and FE65, the YAP WW domain appeared to orient on either of the two fixed proline residues in the library, selecting for Tyr in the Pro^{+2} position and Pro or Leu in the Pro^{-1} position (Table I), consistent with the previous screening results. To independently evaluate the relative importance of this tyrosine residue in YAP WW domain binding in the absence of any orienting Pro residues, the GST-YAP WW fusion protein was screened using

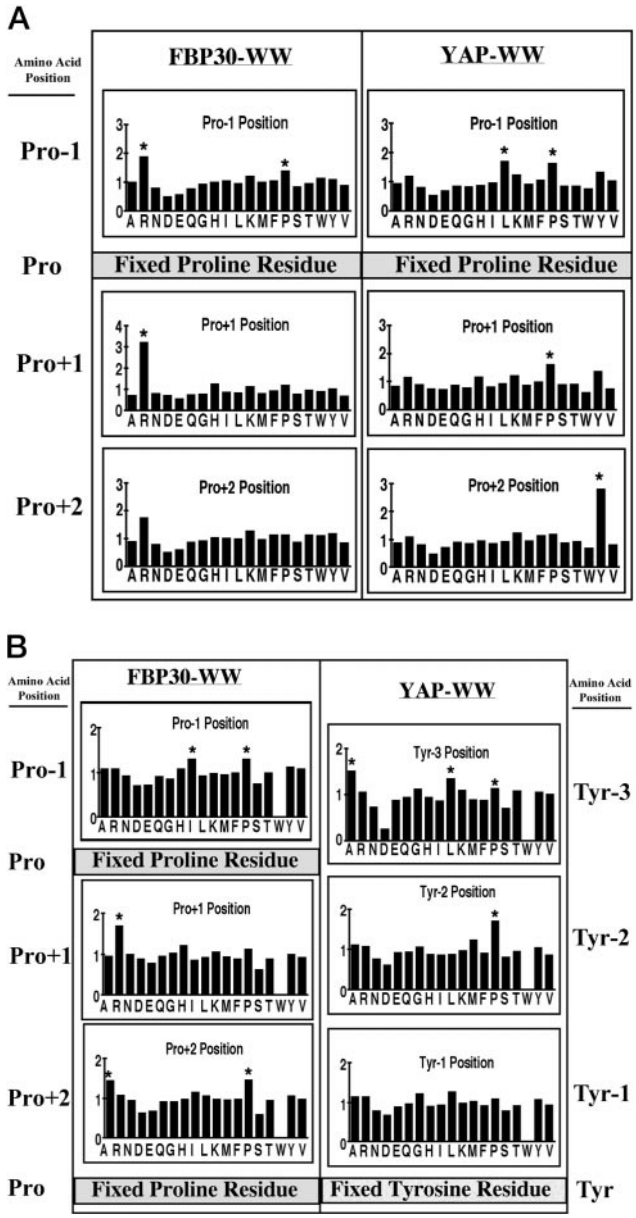


FIG. 2. **Peptide library screening of WW domains.** *A,* motifs for FBP30 WW-A and YAP WW domains using a Pro-oriented degenerate peptide library. Approximately 1–1.5 mg of each WW domain was screened using a library of general sequence MA*XXXX*P*XXXX*AKK, where *X* indicates any amino acid except cysteine. Following extensive washing, bound peptides were eluted and sequenced. Each panel shows the relative abundance of each of the 19 amino acids at a given cycle of sequencing compared with their abundance in the starting library mixture. The one-letter amino acid code listed is used. Strongest selection was observed in positions Pro^{-1}, Pro^{+1}, and Pro^{+2} positions (shown). There was additional weak selection observed in the Pro^{2} position (see Table I). *B,* motif extension and refinement for FBP30 WW-A and YAP WW domains using secondary peptide library screening. The FBP30 WW-A domain was screened with a degenerate peptide library of general sequence MA*XXXX*P*XX*P*XXX*AKKK based on the consensus motif recognized by SH3 domains (*left panels*). The YAP WW domain was screened with a tyrosine-oriented library of general sequence MA*XXXXYXXXX*AKKK. Panels corresponding to amino acid positions showing moderate to strong selection are displayed. Additional weak residue selectivity at other sites for FBP30 WW-A is given in Table I.

TABLE I

Motif selection for WW domains using proline-oriented peptide libraries

Top: WW domain selectivity using a Pro-oriented degenerate peptide library with sequence MAXXXXPXXXXAKK. X indicates all amino acids except Cys. Selectivity values are shown in parentheses. Residues with selectivities ≥1.4 are shown in bold. Bottom: residue selection using a library containing the sequence MAXXXXXPXXPXXXXAKKK optimized for SH domain binding. There was no significant amino acid selection in positions other than those shown.

	−2	−1	**Pro**	+1	+2
YAP WW	F(1.3) M(1.2) P(1.1) L(1.1)	**L**(1.7) **P**(1.6)	P	P(1.0)	**Y**(2.8)
FBP30WW-A	F(1.2) W(1.2) P(1.1)	**P**(1.4)	P	**R**(3.2)	x
FBP30WW-B	x	P(1.3)	P	**R**(3.5) **P**(1.5)	Y(1.3)
FBP30WW A&B	Q(1.5) W(1.3)	**L**(1.4) P(1.3)	P	**R**(3.0)	P(1.2)
FBP21	x	P(1.2)	P	**R**(3.3) P(1.3)	P(1.3)
FE65	F(1.2)	P(1.2)	P	**R**(3.4) **T**(1.6) **H**(1.4)	Y(1.2) P(1.1)

	−2	−1	**Pro**	+1	+2	**Pro**	+4	+5
YAP WW	L(1.3)	**L**(1.7) P(1.3)	P	x	**Y**(2.7) P(1.1)	P	x	**Y**(2.0)
FBP30 WW-A	P(1.1) L(1.1)	P(1.3) I(1.3)	P	**R**(1.7)	P(1.5) A(1.4)	P	R(1.2)	R(1.2)
FBP21	x	P(1.1)	P	**R**(1.8)	x	P	**R**(1.4)	**R**(1.5)
FE65	**K**(2.3) **F**(2.1) **I**(1.9) **P**(1.8) **M**(1.8) **L**(1.8) **V**(1.7)	**Y**(1.4) P(1.3)	P	**R**(2.1)	P(1.2) A(1.2)	P	**R**(1.6) H(1.4)	**R**(1.6)

a peptide library containing only a fixed Tyr residue. As seen in the *left panel* of Fig. 2*B*, the YAP WW domain could bind and orient on these Tyr-containing peptides, selecting for Leu or Pro in the Tyr^3 position and Pro in the Tyr^2 position.

To refine any additional sequence specificity for the other WW domains extending beyond the core PPR-based motif, the WW domains of FBP30, FBP21, and FE65 were screened with a peptide library oriented on the core PPR sequence (Table II). The overall extent of peptide binding we observed to these WW domains was quite large, approximately 10–20-fold larger than that observed using a library in which proline was the sole orienting residue, further supporting the PPR-based motif and suggesting that additional selection beyond this PPR core would likely be modest at best. The FBP30 WW-A, FBP21, and FE65 WW domains demonstrated some additional proline selection in both the PPR-2 and PPR-1 position, suggesting one optimal sequence motif as PPPPR. In addition, all the WW domains showed additional Arg selection flanking both sides of the PPR core, suggesting that sequences such as RPPPR, PRPPR, and PPR*X*R could also bind to this subclass of WW domains.

Identification of FBP30 WW Domain-Binding Proteins by Phage Library Expression Screening—A phage expression strategy was employed to investigate protein targets for the FBP30 class of WW domains (Fig. 3), and determine if the Pro-Arg motifs we had identified by peptide library screening were also present in naturally occurring ligands. In order to screen both WW domains simultaneously, the two WW domains of FBP30 (which are 358 amino acids apart) were engineered to within 24 amino acids of each other and fused to GST. An expression library screen with this GST fusion protein revealed specific interacting plaques. Fifteen of these were selected at random for plaque purification, and the pure phage then used to infect a bacterial strain harboring the *Cre* gene. This allows recombination to occur at loxP sites in the phage, generating pET vectors containing the inserts of interest fused downstream from a T7 gene 10 antigen tag (Novagen). The resulting 15 plasmid DNAs were isolated, sequenced, and found to encode 7 distinct proteins, all with extensive proline-rich tracts (Fig. 3; see below).

Since these proteins were selected using a GST fusion containing both WW domains of FBP30 as a probe, it was possible that each WW domain had a particular affinity for a specific ligand. To test this, we determined the ability of the isolated FBP30 WW-A and WW-B domains to bind to each of the WW domain-binding proteins (WBPs) we had identified by phage expression screening. In addition, we included a control protein (WBP-7) known to bind to an alternative set of WW domains from the formin-binding protein FBP11 through a PPLP motif (40). Bacterial extracts from cells overexpressing the individual T7 gene 10-WBP fusion proteins were analyzed by SDS-PAGE followed by transfer to polyvinylidene difluoride membranes. Identical membranes were then probed using ^{32}P-labeled GST WW domains from FBP30 A&B, FBP30 WW-A alone, FBP30 WW-B alone, and FBP11 A&B (Fig. 3*A*). All the WBPs we had isolated by phage expression bound to both the FBP30 WW-A&B domain probe and the FBP30 WW-A domain probe (*lanes 1–8*). In contrast, the isolated FBP30 WW-B domain probe bound more weakly overall, recognizing predominantly only WBP-15, -17, and -13 (*lanes 3, 6,* and *8*). This suggests that within the FBP30 WW-A&B domain fusion used to screen the expression library, the FBP30 WW-A domain likely dominated over the FBP30 WW-B domain in ligand binding ability.

The WW domains of FBP11 demonstrated extremely strong binding to the WBP-7 protein, Fig. 3, *lane 9*, a proline-rich ligand which completely failed to bind to the FBP30 WW do-

TABLE II

Refinement of Pro-Arg motifs for a class of WW domains

WW domains were screened with a peptide library of sequence MAXXXXPPRXXXXAKK as in Table I. Selectivity values are shown in parentheses. Residues with selectivities ≥1.4 are shown in bold. There was no significant selection beyond the −2 and +2 positions.

	−2	−1	**Pro**	**Pro**	**Arg**	+1	+2
FBP30 WW-A	**R**(1.7) **P**(1.4)	P(1.3) K(1.3)	P	P	R	x	**R**(1.6)
FBP30 WW-B	x	**R**(1.8) K(1.3) H(1.2)	P	P	R	**R**(1.5)	**R**(2.0)
FBP30 WW-A&B	P(1.1)	x	P	P	R	I(1.2)	**R**(1.9)
FBP21	**R**(1.6)	P(1.1)	P	P	R	x	x
FE65	P(1.3)	P(1.2)	P	P	R	x	**R**(1.4) P(1.2)

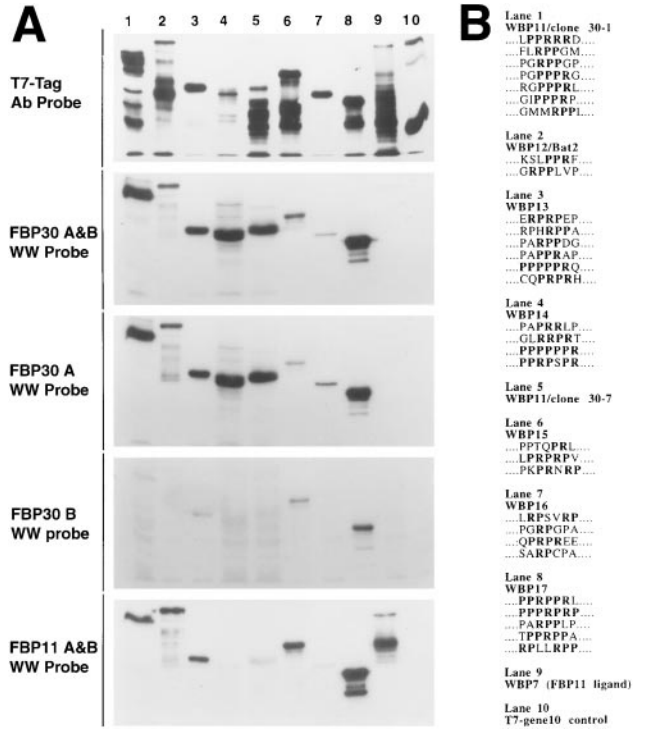


FIG. 3. **Protein binding specificity for the WW domains of FBP30.** *A,* blot overlay analysis of FBP30 WBPs. Bacteria were transformed with pExlox expressing WBP fusion proteins, and induced with isopropyl-1-thio-β-D-galactopyranoside. Crude bacterial lysates were analyzed by SDS-PAGE, transferred to Immobilon-P, and incubated with ^{32}P-radiolabled probes (except for the *top panel*) as indicated in the figure. The *top panel* was probed with an αT7 gene 10 antibody and depicts a loading control of the various fusion proteins. *Lanes 1–8* represent newly identified WBPs that were isolated from a limb expression library using the FBP30 A&B probe (see *panel B* for identity of each lane). Control *lane 9* is WBP7, a previously described ligand for FBP11 (40), which is very proline-rich but does not contain a PPR motif. Control *lane 10* is the exlox fusion protein without an insert. *B,* proline-rich regions within these WBPs contain numerous arginine residues. For each WBP, the sequence that is in-frame with T7 gene 10 was scanned for proline-rich motifs. Proline (*P*) and arginine (*R*) residues that occur within the protein are depicted in *bold*. WBP1/clone 30-1 and clone 30-7 are partial transcripts of the same gene containing overlapping Pro-Arg sequences. The grouping is not intended to represent an alignment.

mains. Conversely, the FBP11 WW domains showed no binding to WBP-11 and -14 (*lanes 4* and *5*), ligands that were the strongest interacters with the FBP30 WW-A&B and WW-A domains, demonstrating selectivity in binding and confirming that the motifs recognized by FBP11 and FBP30 WW domains are distinct. Interestingly, the FBP11 WW domains showed some binding to a subset of ligands that were also recognized by the FBP30 WW domains (*lanes 1–3* and *6*). This suggests either weak overlap between the FBP11 WW and FBP30 WW binding motifs, which seems unlikely based on the peptide library data in Table I and previous characterization of the FBP11 WW motif (40), or more likely, that separate sequence motifs for each of these WW domains lie within the proline-rich portions of these proteins.

FBP30WW-binding Proteins Contain Motifs That Are Arginine- and Proline-rich—Seven different WBPs were identified in 15 cDNA clones identified by phage expression screening. WBP11 was identified 4 times and has been previously isolated as a FBP21 WW domain interacting protein. The sequence of WBP12 corresponded to the proline-rich protein, Bat2 (41). WBP13 represents an out-of-frame fusion with the zinc-binding protein ZnB (42), in which this frame-shifted fusion results in a proline-rich protein. The four remaining sequences (WBP14, WBP15, WBP16, and WBP17) are novel sequences, showing no significant similarity to any known protein in the data base.

All of the WBPs that bound to the FBP30 WW domains were found to contain proline-rich regions that were peppered with arginine residues (Fig. 3*B*). The most common motif in the majority of these proteins contained the core sequence PPR, although there were numerous examples of the reversed motif RPP, as well as the occurrence of PRPR, RP, and RPR sequences within the ligands. All of the ligands contained at least two and as many as seven Pro-Arg based sequences. Five of the seven proteins identified in this screen contained one or more copies of the perfect PPR motif shown in Table II, while the remaining two proteins contained one or more copies of a PXPR sequence that was also selected by peptide library screening (Table I).

The FBP30 WW-A Domain Binds a PPR Peptide with High Affinity—To determine if these Pro-Arg type of motifs were indeed responsible for the interactions of WBPs with the WW domains of FBP30, we proceeded to measure and contrast binding affinities of PPR- and PPXY-based peptides to the FBP30 WW-A and the YAP WW domains (Fig. 4). The Pro-Arg peptide from WBP-11 with the sequence PPGPPPRGPPPR and the Pro-Tyr peptide SGHPGTPPPPYTVG from the YAP WW ligand WBP-1 (34) were synthesized with fluorescein tags attached to their amino termini via an aminohexanoic acid linker, and peptide binding measured by fluorescence polarization anisotropy. The kinetics of binding was rapid, reaching equilibrium within 2 min (Fig. 4, *A* and *B, insets*). The FBP30 WW-A domain demonstrated very strong binding to the Pro-Arg peptide, whereas only weak binding to the Pro-Tyr peptide was observed even at 100 μM FBP30 WW-A concentration (Fig. 4*A*). In contrast, the YAP WW domain strongly bound to the Pro-Tyr peptide and showed minimal binding to the Pro-Arg peptide (Fig. 4*B*). Analysis of fluorescence polarization as a function of protein concentration yielded a K_d for the FBP30 WW-A domain binding to the PPR-based peptide of 14.9 μM, and a K_d for the YAP WW domain binding to the PPPY-based peptide of 5.4 μM, with excellent curve fits (Fig. 4, *C* and *D*). These binding constants compare quite favorably with those of

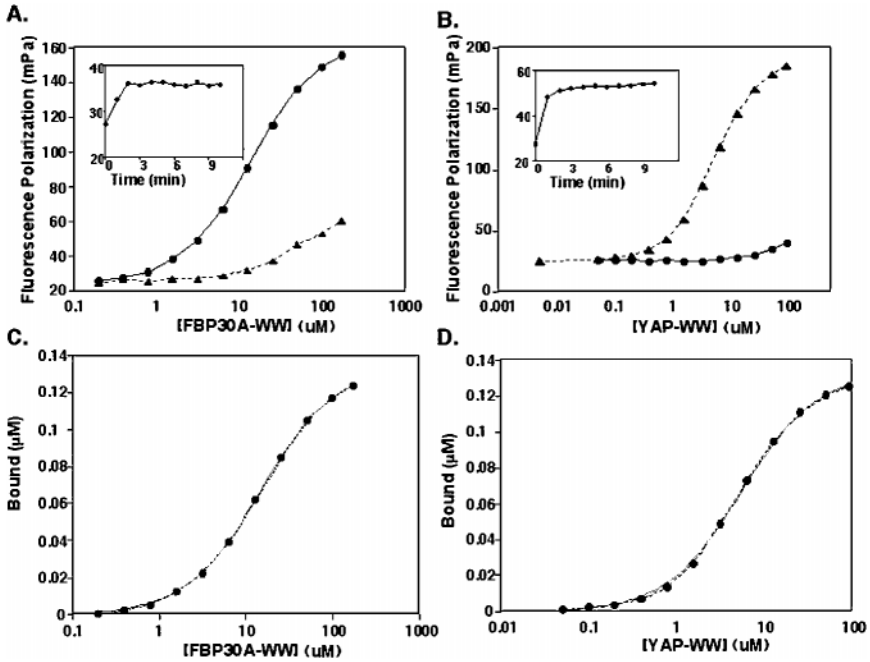


FIG. 4. **Peptide binding affinities for the FBP30 WW-A and YAP WW domains.** *A* and *B,* fluorescence polarization anisotropy measurements demonstrate that the FBP30 WW-A domain binds a PPR-based peptide, but not a PP*X*Y peptide, in contrast to YAP which binds a PP*X*Y peptide, but not a PPR-based peptide. The FBP30 WW-A and YAP WW domains were purified as GST fusion proteins and incubated with a fluorescein-tagged PPR-based peptide from WBP-11 with the sequence PPGPPPRGPPPR (●) or a fluorescein-tagged PP*X*Y-based peptide from the YAP WW ligand WBP-1 with the sequence SGHPGTPPPPYTVG (▲). Binding curves for the FBP30 WW-A domain are shown in *A*, those for the YAP WW domain are shown in *B*. All binding curves for the optimal ligands were measured in triplicate. *Error bars* (±1 S.D.) are present, although they are obscured by the plot symbols for the most part. *Inset,* kinetics of binding for 1 μM of each WW domain to their optimal peptides, showing saturation within 120 s of incubation. *C* and *D,* analysis of binding curves. Binding data were analyzed by fitting the curves using nonlinear regression analyses assuming that each WW domain contained a single peptide-binding site. *Dashed lines* are experimental data, *solid lines* are calculated curves from the regression fit. The k_D for YAP binding to the PPXY peptide (*D*) is 5.4 μM and the k_D for FBP30 WW-A binding to the PPR-based peptide (*C*) is 14.9 μM.

SH3 domains, which typically bind to their proline-rich sequences with K_d values of 10–200 μM. Smaller versions of these peptides containing less than 6 residues were also capable of binding to each of their respective domains, although the binding was weaker, and did not reach saturation even at 100 μM concentrations (data not shown). The most likely explanation for this observation is that extended peptides, particularly those rich in proline residues, reduce the entropic penalty of binding, perhaps by stabilizing a polyproline-II helix conformation.

The PPR-based Peptide Can Disrupt the Association of FBP30 WW-A Domain with Proteins in PC12 Cell Lysates—To verify the selectivity and specificity of PPR and PY-sequence motifs for different WW domains, we investigated the binding of the FBP30-A and YAP WW domains to endogenous proteins within lysates of ^{35}S-labeled PC12 cells in the presence and absence of competing peptides. As shown in Fig. 5, the WW domain of FBP30A bound to numerous proteins within the lysates. The molecular mass of these ranged from ~20 to 120 kDa (*lane 1*). This binding could be competitively inhibited by the presence of 100 μM of the PPR peptide from WBP-11 during the binding reactions, but not by the PY peptide from WBP-1 (*lanes 2* and *3*). In contrast, the YAP WW domain bound most strongly to a protein with molecular mass of ~43 kDa (*lane 4*), whose binding was unaffected by the PPR peptide, but was completely abolished by competition with the PY peptide (*lanes 5* and *6*).

A Structural Model for Pro-Arg Selection—Although the structural basis for Pro-Arg selection by this class of WW domains is unknown, insight comes from molecular modeling studies (Fig. 6). Two WW domain structures have been solved to date, an NMR structure of the YAP WW domain bound to its cognate ligand (54), and a high-resolution x-ray structure of Pin1, a protein containing both a WW domain and a proline isomerase (43). The WW domains in both structures adopt an essentially identical fold, a three-stranded β-sheet with a large shallow hydrophobic binding surface. In the YAP structure, the region that strongly interacts with prolines in the P**P***X*Y ligand motif (particularly the Pro residue shown in bold), is formed from the side chains of Thr^{37}, Tyr^{28}, and Trp^{39} (44), and is shaded *green* in the *left panel* of Fig. 6*A*. In the Pin1 WW domain, the corresponding hydrophobic surface is formed from Ser^{16}, Tyr^{23}, and Trp^{34}, and fortuitously binds a polyethylene glycol molecule from the crystallization solution, effectively demarcating the binding pocket (43) (Fig. 6*A, right panel*). Fig. 6*B* shows the structure of the FBP30 WW-A domain (*blue*) modeled on the high-resolution x-ray structure of the Pin1 WW domain (*yellow*) by threading and energy minimization ("Materials and Methods"). The structures are nearly superimposable, with excellent backbone and side chain fits.

The molecular surfaces for all three WW domains, shaded by electrostatic potential are shown in Fig. 6*C*. The YAP WW domain has an electrostatically neutral surface flanking the shallow peptide binding groove, with the Tyr residue in the PP*X*Y motif stabilized by interactions with Leu^{30} and His^{32} (44). In both the Pin1 and FBP30 WW-A domains, this tyrosine pocket is eliminated by substitution of a bulky aromatic residue (Phe in the case of Pin1 and Tyr in the case of FBP30) for Leu^{32}, and substitution of a Trp residue for His^{32} in FBP30, indicating why these WW domains preserve Pro recognition, but not Tyr selection. In sharp contrast to the neutral/hydrophobic surface of YAP, the peptide binding groove of Pin1 is flanked by a positively charged region at one end, and a negatively charged region at the other, indicated by *blue* and *red arrows*, respectively, in the middle panel of Fig. 6*C*. This charge distribution explains precisely why the optimal ligand for the Pin1 WW

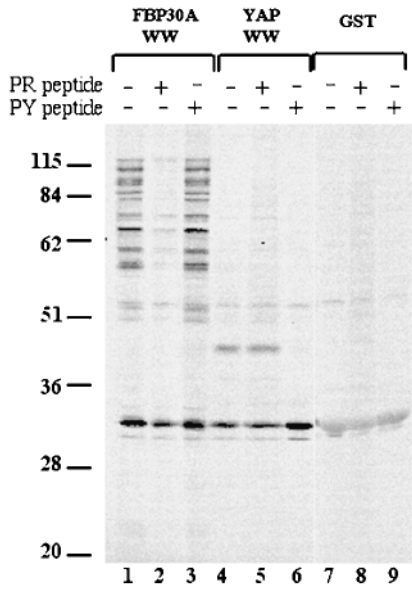


FIG. 5. **Optimal peptides compete for binding of WW Domains to intracellular ligands.** Lysates from ^{35}S-labeled PC12 cells were incubated with beads containing the WW domains of FBP30 WW-A, or YAP as GST fusion proteins, or with GST alone in the presence or absence of the PPR or PY peptides. Bound proteins were analyzed by SDS-PAGE and visualized by autoradiography. The PPR peptide (*lane 2*), but not the PY peptide (*lane 3*) competed the binding of radiolabeled proteins to the FBP30 WW domain (*lane 1*). In contrast, the PY peptide (*lane 6*) but not the PPR peptide (*lane 5*) competed with protein binding to the YAP WW domain (*lane 4*), which predominantly binds to a single major band migrating at ~46 kDa. The positions of molecular mass markers are indicated on the *left* with sizes in kDa.

domain is the sequence *phospho*-Ser-Pro-Arg (19, 23).[2] In the threaded FBP30 WW-A structure, the hydrophobic proline-binding surface is flanked on both ends by negatively charged acidic patches (Fig. 6*C, right panel*), which are likely to be responsible for binding the Arg residues in the PPR motif. Interestingly, the location of these regions at either end of the binding groove provides an excellent rationale for the selection of Arg residues both NH_2- and/or COOH-terminal to the prolines. This explains the occurrence of sequences such as RPP and PRPR that we observed in the FBP30 WW ligands isolated by expression cloning. Furthermore, this analysis suggests that PPR-based ligands might bind in both NH_2- to COOH- and COOH- to NH_2-orientations, as seen with SH3 ligands (45–47).

The Fyn SH3 and p85 SH3 Domains Can Bind Directly to a Subset of the FBP30A Ligand Proteins—The Pro-Arg motif we derived for this novel class of WW domains can be found within a number of proteins, often lying in Pro-Arg sequences recognized by certain SH3 domains (46–49). The Pro-Arg peptide from WBP-11 (Fig. 4), for example, was found to bind to the Fyn SH3 domain with a K_d of ~90 μM by fluoresence polarization anisotropy (data not shown). To investigate whether SH3 domains and this new class of WW domains can, in fact, bind similar ligands, we performed "far Western" blotting experiments using the set of FBP30 WW-A ligands we identified by phage-expression screening (Fig. 3). As shown in Fig. 7*A*, both the Fyn and p85 SH3 domains bound to several of these WBP ligands, interacting strongly with WBP-13, -14, -15, and -17. The exposure shown for the p85 SH3 probe (*bottom panel*) was only one-fifth that for the Fyn SH3 domain probe (*middle panel*), although the probes had identical specific activity. These results demonstrate that proteins containing Pro-Arg-rich stretches can bind to both WW and SH3 domains, and suggests that these WBPs have even higher affinity for the p85 SH3 than for the Fyn SH3 domain.

To further investigate the shared specificity of these WW and SH3 domains, far-Western blotting was performed on endogenous FBP30 WW-A ligands isolated from unlabeled PC12 cell lysates using GST-FBP30 WW-A beads (Fig. 7*B*). Both the radiolabeled FBP30 WW-A and Fyn SH3 probes were found to recognize many of the same protein bands. Similar results were also obtained using a radiolabeled p85 SH3 domain (data not shown). In this experiment far Western blotting identified only a subset of the total FBP30 WW-A-associating ligands shown in Fig. 5. This likely results from the higher sensitivity that affinity purification provides for native ligands compared with far Western blotting which involves refolding of denatured proteins following electrophoresis and transfer to blotting membranes.

DISCUSSION

A Pro-Arg Motif Recognized by the WW Domains of FBP30—Based on data from oriented peptide library screening, phage-expression cloning, affinity measurements of peptide binding, and competition studies for binding of WW domains to cellular ligands, we have identified a new binding motif, the Pro-Arg motif, recognized by a subgroup of WW domains. The prototype members of this class are the WW domains of FBP30, particularly the NH_2-terminal WW domain, WW-A. Depraetere and Golstein (21) demonstrated strong sequence conservation of this portion of the FBP30 molecule between mouse, rat, and humans, suggesting that it performs an essential function. Since FBP30 gene expression is induced upon γ-irradiation in a p53-dependent manner, and p53-driven transcription is essential for γ-irradiation induced cell death (50), these authors hypothesized that FBP30 may play an important role in the apoptotic process. Our elucidation of the Pro-Arg binding motif for the WW domains of FBP30 should assist in the determination of physiological ligands that are involved in this or other aspects of FBP30 function.

The FBP30 sequence presented here extends that reported by Depraetere and Golstein (21) and reveals an additional WW domain and 2 additional proline-rich regions. The largest cDNA that Depraetere and Golstein (21) assembled was 2.5 kilobases, although they noted that the size of the FBP30 transcript detected in T-cells was ~4 kilobases. The additional FBP30 sequence presented here resolves this discrepancy.

Curiously, the first Pro-rich sequence within FBP30 contains a PPRP sequence motif 47-amino acid NH_2-terminal to the WW-A domain, suggesting the possibility of an intramolecular association, a fact that was also noted by Depraetere and Golstein (21) strictly on the basis of its proline-rich content. Preliminary experiments reveal that an FBP30 fragment containing this segment and the WW-A domain of the FBP30 molecule does, in fact, behave as a compact folded structure.[3] Ultimate verification of this intramolecular association, along with confirmation of our structural model for the WW-A domain is being attempted through crystallographic analysis and mutational studies that are currently in progress.

Ligand Targets of FBP21 and FE65 Are Consistent with Pro-Arg Motif Selection—In addition to the WW domains of FBP30, we also identified the PPR motif as the optimal ligand selected by the FBP21 and FE65 WW domains. Previous work has shown that the FBP21 protein is a component of the spliceosomal complex A, where it can associate with the proline-rich regions in the splicing factors U1C, SmB, SmB′, and SF1/mBBP (15). The original motif proposed for FBP21 binding on the basis of ligand sequence alignments was a Pro-Gly-Met-

[2] M. B. Yaffe, X. Z. Zhou, and K. P. Lu, unpublished data.

[3] M. B. Yaffe, unpublished data.

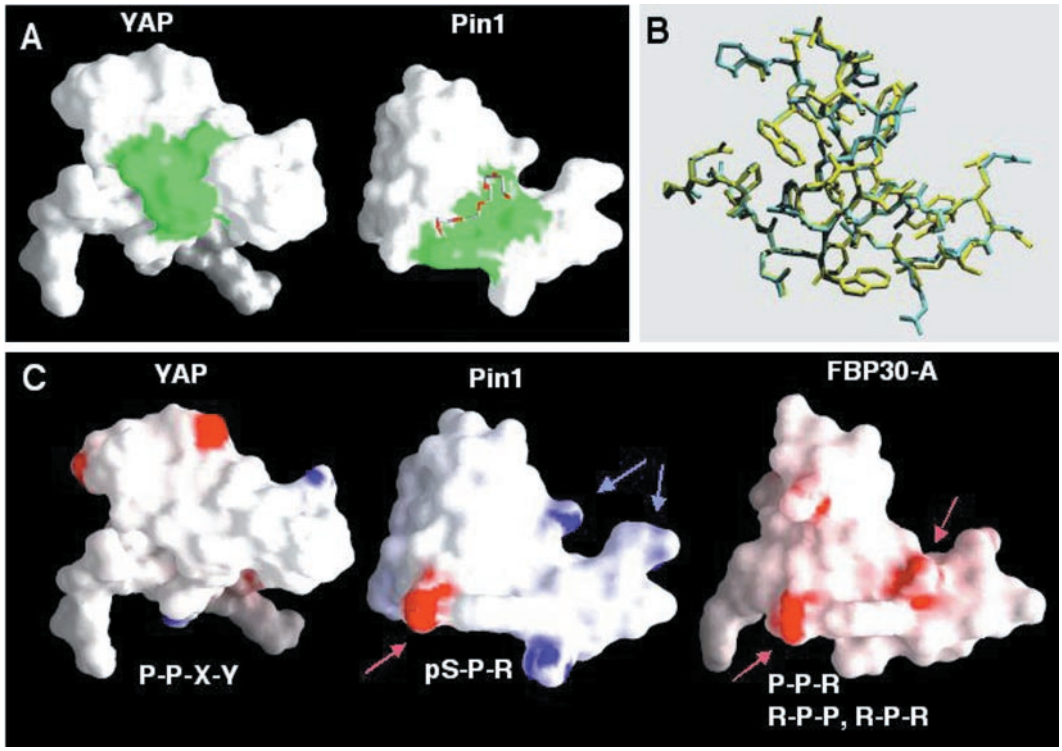


FIG. 6. **Structural basis for proline motif selectivity by different WW domains.** *A,* regions of ligand contact with the YAP and Pin1 WW domains. Molecular surfaces for each domain were constructed using GRASP. The regions of the YAP WW domain that interact with the proline residues in the PP*X*Y are shaded *green* (44), as is the corresponding region in Pin1 (43). In the Pin1 structure, this surface fortuitously bound a molecule of polyethylene glycol from the crystallization solution, shown in stick representation. *B*, modeling the WW domain of FBP30 WW-A based on the Pin1 WW domain structure. The structure of the WW-A domain of FBP30 was modeled by threading the sequence onto that of the high-resolution x-ray structure of Pin1 (36% amino acid identity and 68% homology allowing for conservative substitutions), followed by energy minimization ("Materials and Methods"). The resulting structure for the FBP30 WW domain (*blue*) is nearly superimposable with that of Pin1 (*yellow*). Note that many regions appear as only a single chain colored in a mixture of *yellow* and *blue*. *C,* differences in surface electrostatic potential for different WW domains rationalize ligand specificity. The molecular surfaces for the YAP, Pin1, and FBP30 WW domains are shown shaded by electrostatic potential. *Red* indicates negative charge, *blue* indicates positive charge. The region surrounding the proline binding surface of the YAP WW domain is neutral, whereas in Pin1WW it is flanked at opposite ends by regions of positive and negative charge (*blue* and *red arrows*, respectively). The proline-binding surface of the predicted FBP30 WW-A domain structure is flanked at both ends by negative charge density (*red arrows*). The optimal amino acid motifs recognized by each WW domain are listed beneath each structure. Note that the distribution of neutral, basic, and acidic residues within each motif exactly matches the complimentary charge distribution of the corresponding WW domain molecular surface. All WW domains in *panels A, B*, and *C* are shown in the same orientation.

rich motif. Re-evaluation of the binding experiments described in that paper, in light of the current data, reveals that an RPP and/or a PPR sequence was present in all of the fusion peptides of SmB examined. Furthermore, all of the splicing factors that interact with FBP21 are found to contain PPR motifs within their sequences. To directly investigate the importance of Arg residues within these Pro-Gly-Met-rich sequences, biotinylated peptides were synthesized containing either the wild-type FBP21-interacting peptide from SmB (sequence PPGMRPPP-PGMRRGPPPPGMRPPRP, Ref. 15) or the same peptide with Ala substituted for each Arg, and used in far Western blotting experiments with the WW domains of FBP21 and FBP30. Both peptides contained the identical sequence of Pro, Gly, and Met residues. The WW domains of FBP21 and FBP30 showed strong interaction with the wild-type peptide, but were completely unable to interact with the Arg → Ala mutant peptide despite the persistence of Pro, Gly, and Met residues, verifying the critical role of the Arg residues within Pro-rich sequences in FBP21 and FBP30 WW domain binding (data not shown).

The FE65 protein is an adaptor molecule containing a single WW domain and 2 PTB domains that interacts with the β-amyloid precursor protein and the mammalian homolog of *Drosphila* enabled (Mena) (36). Sudol and colleagues (36) originally proposed a PPLP motif for this WW domain, since it bound to PPLP-containing sequences in proteins such as Mena during cDNA expression library screening. Using a membrane-based peptide binding assay, these authors dutifully noted, however, that substitution of Arg or Lys for Leu in the PPLP motif increased the affinity of the FE65 WW domain for peptides (36). An elegant series of experiments by Espanel and Sudol (22) now makes it clear that the optimal FE65 WW domain ligands contain either a hexa-proline sequence or a PPR-based motif. Thus, the new Pro-Arg motif we have identified can retrospectively rationalize the previous ligand selection observed for FBP21 and FE65 WW domains, and suggests that these WW domains be re-grouped into the Pro-Arg binding class. The fact that the FE65 can also bind to PPLP sequences, albeit more weakly, suggests that this may be an example of a dual specificty WW domain. In summary, our data support a re-categorization of WW domains into 4 classes based on their ligand specificity, as originally proposed by Sudol and colleagues (Table III). In this scheme, class I WW domain molecules recognize the motif PP*X*Y, class II WW domains recognize PPLP, class III WW domains recognize PPR, and class IV WW domains recognize *phospho*-Ser-Pro.

The PPR Motif May Allow Ligand Sharing between WW- and SH3 Domain-containing Proteins—The Pro-Arg motifs we have identified as optimal ligands for these WW domains are often found within, flanking, or partially overlapping consensus Pro/Arg sequences that are recognized by SH3 domains (Fig. 7; *c.f.* Refs. 45, 46, and 49). Many SH3 domains are known to bind to Pro-rich motifs containing Arg residues either NH_2-terminal

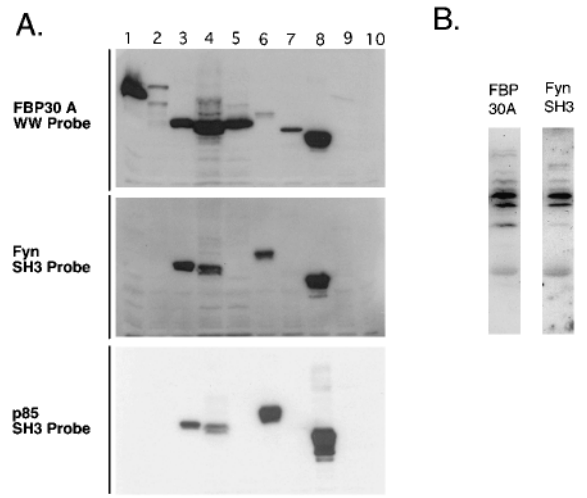


FIG. 7. **FBP30 WW domain ligands can also bind to SH3 domains.** *A,* WBP proteins obtained by expression cloning can bind to the Fyn and p85 SH3 domains. Blot overlay analysis of the WBP fusion proteins obtained by expression cloning was performed using ^{32}P-radiolabeled FBP30 WW-A (*top*), Fyn (*middle*), or p85 (*bottom*) SH3 domain probes as in Fig. 3. Lane identity is the same as that listed in the the legend to Fig. 3. The exposure shown for the p85 SH3 probe is one-fifth of that for the other two probes. *B,* endogenous FBP30 WW-A-binding proteins from PC12 cell lysates are also recognized by the Fyn SH3 domain. PC12 cell lysates were affinity purified using GST-FBP30 WW-A beads as described previously, analyzed by SDS-PAGE, and transferred to Immobilon membranes. Blot overlay analysis was then performed using ^{32}P-radiolabeled FBP30 WW-A or Fyn SH3 domain probes.

TABLE III
Classification of WW domains based on optimal ligand specificity

	Motif	WW domains	Ref.
Class I	PPXY	YAP PDGF-R Dystrophin NEDD4 WW-3	8, 10, 34, 52, 53
Class II	PPLP	FBP11 WW-A FBP11 WW-B	40
Class III	PPR	FBP30 FBP21 FE65	This paper
Class IV	*phospho*-SP	Pin1 NEDD4 WW-2	23

(so-called class I ligands) or COOH-terminal (so-called class II ligands) to a central P*XX*P core. Class I ligands for the SH3 domain of phosphoinositol 3-kinase, for example, have the optimal consensus motif R*XX*PPRP (45), while a variety of class II SH3 ligands, which adopt the consensus sequence P*XX*PZR, contain a proline residue in the Z position (46). We speculate that a tightly shared and/or overlapping Pro-Arg sequence motif recognized by both class III WW domains and SH3 domains may allow individual proteins to form different signaling complexes depending upon their local microenvironment.

The new Pro-Arg rich motif we describe for WW domains lends strong support to a general model in which different classes of WW domains and SH3 domains compete for binding to particular ligand targets based on unique Pro-rich sequence motifs (40, 51). A limited version of this model was first advanced by Chan *et al.* (24) who showed that the PPLP motif recognized by the class II WW domain of FBP11 was also recognized by the SH3 domain of Abl. The consensus motif for the Abl SH3 domain, PP*X*θ*X*PPPφP, however, lacks any Arg or Lys residues (θ represents aromatic amino acids, φ represents aliphatic amino acids) and differs significantly from that of most SH3 domains, which require arginine or lysine residues flanking or within the P*XX*P core (49). The new PPR motif we describe here, which is recognized by the WW domains of FBP30, FBP21, and FE65 and by the SH3 domains of Fyn and p85, significantly expands the range of ligands that could interact with both WW and SH3 domain-containing proteins. Finally, the new Pro-Arg motif we have identified may provide an additional potential means for regulating WW and SH3 interactions, since Arg modification by methyltransferases has recently been found to disrupt SH3, but not WW domain binding.[4] Work in progress should shed additional light on this and other mechanisms through which this PPR class of WW domains and SH3 domain interactions are regulated.

Acknowledgments—We gratefully thank Dr. Lewis C. Cantley for continued encouragement, support, and comments throughout this study, and German G. Leparc for critical reading of the manuscript and assistance with figure construction.

Note Added in Proof—While this manuscript was in press, Komuro *et al.* (Komuro, A., Saeki, M., and Kato, S. (1999) *J. Biol. Chem.* **274,** 36513–36519) reported that the WW domain of Npw38 binds to a Pro-Gly-Arg motif.

REFERENCES

1. Kuriyan, J., and Cowburn, D. (1997) *Annu. Rev. Biophys. Biomol. Struct.* **26,** 259–288
2. Pawson, T. (1995) *Nature* **373,** 573–580
3. Yaffe, M. B., and Cantley, L. C. (1999) *Nature* **402,** 30–31
4. Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., Domann, E., Carl, U. D., Walter, U., Gertler, F. B., Wehland, J., and Chakraborty, T. (1997) *EMBO J.* **16,** 5433–5444
5. Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993) *Science* **259,** 1157–1161
6. Sudol, M. (1996) *Prog. Biophys. Mol. Biol.* **65,** 113–132
7. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *FASEB J.* **14,** 231–241
8. Rentschler, S., Linn, H., Deininger, K., Bedford, M. T., Espanel, X., and Sudol, M. (1999) *Biol. Chem.* **380,** 431–442
9. Faber, P. W., Barnes, G. T., Srinidhi, J., Chen, J., Gusella, J. F., and MacDonald, M. E. (1998) *Hum. Mol. Genet.* **7,** 1463–1474
10. Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J., and Rotin, D. (1996) *EMBO J.* **15,** 2371–2380
11. Wang, G., Yang, J., and Huibregtse, J. M. (1999) *Mol. Cell. Biol.* **19,** 342–352
12. Craven, R. C., Harty, R. N., Paragas, J., Palese, P., and Wills, J. W. (1999) *J. Virol.* **73,** 3359–3365
13. Harty, R. N., Paragas, J., Sudol, M., and Palese, P. (1999) *J. Virol.* **73,** 2921–2929
14. Wang, H. Y., Lin, W., Dyck, J. A., Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X. D. (1998) *J. Cell Biol.* **140,** 737–750
15. Bedford, M. T., Reed, R., and Leder, P. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10602–10697
16. Gavva, N. R., Gavva, R., Ermekova, K., Sudol, M., and Shen, C.-K. J. (1997) *J. Biol. Chem.* **272,** 24105–24108
17. Yagi, R., Chen, L. F., Shigesada, K., Murakami, Y., and Ito, Y. (1999) *EMBO J.* **18,** 2551–2562
18. Waragai, M., Lammers, C. H., Takeuchi, S., Imafuku, I., Udagawa, Y., Kanazawa, I., Kawabata, M., Mouradian, M. M., and Okazawa, H. (1999) *Hum. Mol. Genet.* **8,** 977–987
19. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) *Science* **278,** 1957–1960
20. Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) *Genes Dev.* **12,** 706–720
21. Depraetere, V., and Golstein, P. (1999) *Cell Death Differ.* **6,** 883–889
22. Espanel, X., and Sudol, M. (1999) *J. Biol. Chem.* **274,** 17284–17289
23. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science* **283,** 1325–1328
24. Chan, D. C., Bedford, M. T., and Leder, P. (1996) *EMBO J.* **15,** 1045–1054
25. Kapeller, R., Prasad, K. V., Janssen, O., Hou, W., Schaffhausen, B. S., Rudd, C. E., and Cantley, L. C. (1994) *J. Biol. Chem.* **269,** 1927–1933
26. Kaelin, W. G., Jr., Krek, W., Sellers, W. R., DeCaprio, J. A., Ajchenbaum, F., Fuchs, C. S., Chittenden, T., Li, Y., Farnham, P. J., Blanar, M. A., Livingston, D. M., and Flemington, E. K. (1992) *Cell* **70,** 351–364
27. Yaffe, M. B., and Cantley, L. C. (1999) *Methods Enzymol.*, in press
28. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723
29. Peitsch, M. C. (1996) *Biochem. Soc. Trans.* **24,** 274–279
30. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
31. Kingston, R. E. (1994) in *Current Protocols in Molecular Biology* (Asubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) Vol. 1, p. 4.2.8, John Wiley and Sons, New York
32. Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D., and Sonnhammer, E. L. (1999) *Nucleic Acids Res.* **27**, 260–262
33. Linn, H., Ermekova, K. S., Rentschler, S., Sparks, A. B., Kay, B. K., and Sudol,

[4] M. T. Bedford *et al.*, submitted for publication.

M. (1997) *Biol. Chem.* **378,** 531–537
34. Chen, H. I., and Sudol, M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7819–7823
35. Chen, H. I., Einbond, A., Kwak, S.-J., Linn, H., Koepf, E., Peterson, S., Kelly, J. W., and Sudol, M. (1997) *J. Biol. Chem.* **272,** 17070–17077
36. Ermekova, K. S., Zambrano, N., Linn, H., Minopoli, G., Gertler, F., Russo, T., and Sudol, M. (1997) *J. Biol. Chem.* **272,** 32869–32877
37. Fiore, F., Zambrano, N., Minopoli, G., Donini, V., Duilio, A., and Russo, T. (1995) *J. Biol. Chem.* **270,** 30853–30856
38. Kumar, S., Tomooka, Y., and Noda, M. (1992) *Biochem. Biophys. Res. Commun.* **185,** 1155–1161
39. Grabs, D., Slepnev, V. I., Songyang, Z., David, C., Lynch, M., Cantley, L. C., and De Camilli, P. (1997) *J. Biol. Chem.* **272,** 13419–13425
40. Bedford, M. T., Chan, D. C., and Leder, P. (1997) *EMBO J.* **16,** 2376–2383
41. Banerji, J., Sands, J., Strominger, J. L., and Spies, T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 2374–2378
42. Trompeter, H. I., Brand, I. A., and Soling, H. D. (1989) *FEBS Lett.* **253,** 63–66
43. Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997) *Cell* **89,** 875–886
44. Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) *Nature* **382,** 646–649
45. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) *Cell* **76,** 933–945
46. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) *Science* **266,** 1241–1277
47. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature* **372,** 375–379
48. Alexandropoulos, K., Cheng, G., and Baltimore, D. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 3110–3114
49. Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quillam, L. A., and Kay, B. K. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 1540–1544
50. Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A., and Jacks, T. (1993) *Nature* **362,** 847–849
51. Sudol, M. (1996) *Trends Biochem. Sci.* **21,** 161–163
52. Irusta, P. M., and DiMaio, D. (1998) *EMBO J.* **17,** 6912–6923
53. Kanelis, V., Farrow, N. A., Kay, L. E., Rotin, D., and Forman-Kay, J. D. (1998) *Biochem. Cell Biol.* **76,** 341–350
54. Deleted in proof