Common Mechanism of Ligand Recognition by Group II/III WW Domains

REDEFINING THEIR FUNCTIONAL CLASSIFICATION*

Yusuke Kato‡, Koji Nagata‡, Mihoko Takahashi‡, Lubing Lian§, Juan J. Herrero§,
Marius Sudo¶, and Masaru Tanokura**

From the ‡Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences,
The University of Tokyo, 1-1-1 Yayoi, Bunkyoku, Tokyo 113-8657, Japan, §AxCell Biosciences,
Neuown, Pennsylvania 18940, and the ¶Department of Medicine, Mount Sinai School of Medicine,
New York, New York 10029

WW domain is a well known protein module that mediates protein to protein interactions by binding to proline-containing ligands. Based on the ligand predilections, the WW domains have been classified into four major groups. Group II and III WW domains have been reported to bind the proline-leucine and proline-arginine motifs, respectively. In the present study, using surface plasmon resonance technique we have shown that these WW domains have almost indistinguishable ligand preferences and kinetic properties. Hence, we propose that Group II and III WW domains should be joined together as one group (Group II/III). Unlike Group I and IV WW domains, Group II/III WW domains can bind simple polyprolines as well as the proline-leucine and proline-arginine motifs, and they possess two Xaa-proline (where Xaa is any amino acid) binding grooves similar to SH3 domains. Our work assigns Group II and III WW domains to a larger family of polyproline-binding modules and proteins, which includes SH3 domains and profilin. Because polyprolines belong to the most frequently found peptide motifs in several genomes, our study implies the versatile importance of Group II/III WW domains in signaling.

The WW domain is composed of 30–40 amino acids and named after two tryptophan (W) residues that are highly conserved and spaced 20–22 amino acids apart (1, 2). The domain binds proline-rich or proline-containing ligands (3–5). WW domain-containing proteins have been shown to be involved in a variety of cellular processes including cell cycle control (Pin1/Es41), ubiquitin ligation (Nedd4/Rsp5 and smurf1), and coactivation of transcription (YAP65) (reviewed in Refs. 2 and 6). A few WW domain-containing proteins have been implicated, either directly or indirectly, in a variety of human diseases such as Liddle's syndrome of hypertension, Duchenne and Becker muscular dystrophies, Huntington's disease, and Alzheimer's disease (7–9). FBP11, a mammalian homologue of yeast Prp40, binds the Pro-Leu (PL) motif (4). FBP11 and WAC act as a component of the splicing factor (10, 11). FBP11, HYPB, and HYPC also bind Huntington, a protein responsible for Huntington's disease (7). Fe65 binds ß-amyloid precursor protein, a precursor of ß-amyloid peptide, which constitutes the extracellular neuritic plaques in Alzheimer disease (9, 12). The WW domain of Fe65 binds the proline-rich region of Mena, a mammalian homologue of Dro sophila melanogaster Ena, which has been identified in a specific screen for dominant mutations that alleviate the Aβ/ phenotype (13, 14). Fe65 and ß-amyloid precursor protein are involved in the reconstruction of actin cytoskeleton, which suggests the concerted action of these proteins with Mena and/or Abi (9, 15). The WW domain of FBP30 has been reported to bind the Pro-Arg (PR) motif (10, 16).

WW domains have been classified into four groups according to their ligand specificity: Group I recognizes Pro-Pro-Xaa-Tyr (PY motif); Group II recognizes Pro-Pro-Leu-Pro (PL motif); Group III recognizes proline-rich segments with Arg residues (PR motif); and Group IV recognizes Ser(P)/Thr(P)-Pro- (pS/pT-P motif) (2, 9). According to this classification, the WW domains of FBP11 and Fe65 have been assigned to Group II, whereas that of FBP30 has been assigned to Group III. The crystal structures of the Group I and IV WW domains complexed with their ligands have revealed a common structural basis to recognize the Pro-containing ligands; the XP groove, which is formed by two aromatic rings with nearly parallel alignment, plays a pivotal role in recognizing the Xaa-Pro segment in the ligands (17–19). In addition, Group I and IV WW domains have been shown to possess their respective characteristic ligand recognition sites, the Tyr-binding groove (Tyr groove), and the phosphate-binding patch (“p” patch), which contribute to the specific ligand recognition (20). However, the classification between the Group II and III WW domains has been nebulous. For instance, the WW domain of Fe65 has been classified as Group II in one report (14) but as Group III in another report (16). In addition, there are a few proposed ligand motifs that can also bind some of the WW domains in these classes: the PGR motif and polyprolines (7, 10, 21, 22).

To examine the functional classification of the Group II and III WW domains in detail, we have performed quantitative binding experiments using surface plasmon resonance (SPR).1 The abbreviations used are: SPR, surface plasmon resonance; GST, glutathione S-transferase; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; Fmoc, N-(9-fluorenyl)methoxycarbonyl; PPII, polyproline type II.

1 The abbreviations used are: SPR, surface plasmon resonance; GST, glutathione S-transferase; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; Fmoc, N-(9-fluorenyl)methoxycarbonyl; PPII, polyproline type II.
Ligand Recognition by Group II/III WW Domains

EXPERIMENTAL PROCEDURES

Plasmid Construction—The DNAs encoding the WW domains of human Pin1 and Caenorhabditis elegans Y110 (Y110) were synthesized and amplified by a modified PCR in which several synthetic oligonucleotides were used as the template for the first cycle. The DNA encoding the second WW domain of Saccharomyces cerevisiae Rap1 (Rap1p[WW2]), human F65L2, and the first WW domain of mouse FBP30 (FBP30A) were amplified by PCR using their cDNAs as templates, which are kindly provided by Dr. H. Tanahashi and Dr. M. T. Bedford (16, 24). The PCR products were inserted between the BamHI and EcoRI sites of the expression vector pGEX-4T-1 (Amersham Biosciences). The expression plasmids for GST fusion proteins of the third WW domain of mouse Nedd4 (mNedd4[WW3]) and the first WW domain of human YAP65 (hYAP65[WW1]) were kindly provided by Dr. C.-K. J. Shen (25). The expression plasmid for GST fusion protein of the first WW domain of mouse FBP11 (FBP11A) was kindly provided by Dr. M. T. Bedford (26). The WW domains of HYPB, WAC, and Fe65 were inserted into pGEX-4T-2 vectors as described earlier by Hu et al. (23).

Expression and Purification—All the GST fusion WW domains were expressed in Escherichia coli BL21(DE3) or BL21(DE3)pLysS (Novagen) at 37 °C and purified with glutathione-Sepharose (Amersham Biosciences). The GST tag was removed by thrombin (Sigma) digestion at 4 °C for the SPR analysis. The WW domains were separated from GST and thrombin by reverse phase chromatography with a Resource RPC 1-ml column (Amersham Biosciences) with a linear gradient of 1–40% acetonitrile (1 ml/min, 20 min) in 20 mM ammonium formate, pH 7.0. Their molecular masses were verified by MALDI-TOF mass spectrometry on a Voyager mass spectrometer (Applied Biosystems). The sample used in the proteomic mapping was prepared as described (23).

Preparation—Purified WW domains of human WW domains were plated on 96-well plates. Employing automated equipment, complete cross-affinity matrices were generated between GST-WW fusions and proline-rich peptides derived from the human proteome. Peptide-streptavidin/alkaline phosphatase reaction was used to monitor the relative strength of binding. Quality controls and checks were employed as previously described (23).

Peptide Synthesis for the SPR Analysis—Pro-containing ligand peptides were synthesized by solid phase peptide synthesis on a PSSM8 Peptide Synthesizer (Shimadzu, Kyoto, Japan). Rink amide AM resin (Novabiochem) and Fmoc-amino acids with protected side chains were used (Novabiochem). For phosphotau and phosphothreonine, Fmoc-Ser(PO(OEt)2)OH and Fmoc-Thr(PO(OEt)2)OH (Novabiochem) were used, respectively. Synthesized peptides were cleaved from the resin and deprotected in trifluoroacetic acid in the presence of scavengers (5% water, 5% thioanisole, 3% ethylmethylsulfoxide, 2.5% ethanethiol, 2% thiophenol) and then purified by reverse phase chromatography on a Waters ODS-AQ10S (4.6 mm × 100 mm; YMC, Kyoto) flow cell. Their molecular masses were verified by MALDI-TOF mass spectrometry on a Voyager mass spectrometer (Applied Biosystems). The amino acid sequences of the ligand peptides were GTGPPPpTYPVG (WB1 peptide containing the PY motif), SPPPPPPLPPP (formin peptide containing the PL motif, mouse), PPRPPPPPPPPPPPPPPP (14-mer polyproline called PP in the present paper), CCGGGGPPPPPPGFFPP (WB1 peptide containing the PR motif, green fluorescent protein [GFP] GFP[R126P] (Cdc25c peptide), GMPYVPLPP (prym motif containing the pS/pT-PP motif), and KGPPQAPPY (Myt1 peptide containing the pS/pT-P motif).

SPR Binding Assay—We measured SPR using a BIAcore 2000 (Biacore). Each ligand peptide was immobilized on a flow cell of a sensor chip CM5 (Biacore) by the standard 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride/N-hydroxysuccinimide method with a 1 mg/ml peptide solution dissolved in 25 mM NaHCO3. For reference SPR signals, 25 mM NaHCO3 was used instead of a peptide solution in the immobilization reaction. The remaining activated esters were inactivated with ethanolamine. The binding experiments were performed at 20 °C using the WW domains dissolved as analytes in 20 mM HEPES, pH 7.0, 100 mM NaCl, and 0.5 mM EDTA. The flow cells were regenerated with 100 mM NaOH. The concentrations of the WW domains were determined based on the absorbance at 280 nm of the protein solutions. The dissociation constants (KD) were determined by Scatchard plot analysis (27, 28). First, the raw sensorgrams of a flow cell without an immobilized ligand peptide were fitted to the raw sensorgrams of flow cells. Then, RUeq/C values were plotted against the RUinc values, where RUinc was the SPR response under the equilibrium, and C is the millimolar concentration of the analyte. Plotted data were fit to the equation: RUeq/C = −1/KD × RUinc + A (a constant), to determine KD. All the R2 values for the fittings by linear functions were >0.91. The association and dissociation rate constants (ka and kb) were determined from the direct fitting method using BIAevaluation 2.1.

Molecular Modeling—Homology modeling was carried out using Swiss-Model (29), an automated protein-modeling server combining homology search, model building, and energy minimization. The crystal structures of Pin1 (Protein Data Bank code 1PIN), and dystrophin (Protein Data Bank code 1B53) and the NMR structures of FBP28 (Protein Data Bank code 1EGL) and mNedd4 [WW3] (Protein Data Bank code 1I5H) were used as the templates for modeling the WW domains of FBP11A, F65L2, and FBP30A. Molecular surfaces and electrostatic potentials were calculated and displayed using MOLMOL (30). The Swiss PDB Viewer was used to draw the backbone model (29).

RESULTS

SPR Binding Assay—We performed quantitative binding experiments by SPR to evaluate the ligand specificity of several WW domains belonging to Groups I–IV. Fig. 1, A and B, depict the SPR assay and analysis are shown in Fig. 1, A and B. The polypeptide ligands used were WBP1, Formin, WBP11, and Cdc25c/Myt1 motifs and are the typical ligands for Group I, II, III, and IV WW domains, respectively. In addition, 14-mer polyproline (called PP in the present paper) was used. Eleven different WW domains were used as binding probes: the WW domains of hYAP65, Rsp5p, and mNedd4 (previously classified as Group I); those of FBP11A and HYPB (previously Group II); that of FBP30A (previously Group III); and those of Pin1 and Y110 (previously Group IV) (20). The WW domains of WAC and F65L2 have never been experimentally classified before. None of the WW domains used in the present SPR analysis was in the form of the GST fusion. The GST fusion proteins have the propensity to form dimers because of the nature of the GST component. This activity could cause the problem of avidity in the measurements of SPR. Thus, the proteins we used in the present study were digested and separated from the GST before the measurement.

SPR Binding Assay of WW Domains Previously Classified as Group I—Many proteins have multiple WW domains. In our SPR experiments, we used only individual WW domains from such multi-WW domain proteins. The first WW domain of hYAP65 (hYAP65 [WW1]), the second WW domain of Rsp5 (Rsp5p[WW2]), and the third WW domain of mNedd4 (mNedd4 [WW3]) interacted most strongly with the WB1 peptide that contained a PY motif, the consensus sequence for the Group I ligands, with KD values of 11–71 μM (Fig. 1C). These values are in good agreement with the KD values determined for Group I WW domains using various techniques including isothermal titration microcalorimetry (31, 32). Considering these results, the problem of avidity was well avoided in the present measurement system. The WW domains of hYAP65 (WW1) and rSp5 (WW2) also bound to the formine peptide that contains a PL motif, a consensus sequence for the...
Group II ligands. However, their bindings to the formin peptide were 19 and 50 times weaker in \( K_D \) than those to the WBP1 peptide. These data show that the WW domains of hYAP65(WW1), rsps5p(WW2), and mNed44(WW3) bound highly specifically to the PY motif, the Group I ligand. None of the Group I WW domains bound to PP peptide.

**SPR Binding Assay of WW Domains Previously Classified as Group II**—The binding preference to the PP peptide were shown for the WW domains of FBP11A, HYPB, WAC, Fe65, and Fe65L2 (Fig. 1C). Moreover, the PL and PR motifs exhibited similar binding strength to those WW domains. Thus, a broader range in binding specificity seems common for Group II WW domains. For the WW domain of FBP11A (Group II), it was quantitatively confirmed that the most preferable ligand...
was the formin peptide, the PL motif. The PP and WBP11 peptide (the PR motif) also bound to FBP11A with only 2.1 and 3.3 times larger \( K_D \) values, respectively, than the PL motif. The WW domain of WAC proved to possess the Group II specificity because it bound to the PP peptide most effectively. In addition, its relative binding abilities to the other Group II and III ligands were similar, with at most 2.1 times larger \( K_D \) values. As was the case with the WAC WW domain, the WW domains of HYBP, Fe65, and Fe65L2 bound most strongly to PP in our SPR experiments. The WW domains of WAC, Fe65 and Fe65L2 bound also to the PL and PR motif at most with 8.4 and 2.5 times larger \( K_D \) values than to the PP peptide, respectively. On the other hand, the WW domain of HYBP showed preference to the PP peptide, to which HYBP bound 19 and 20 times more strongly than to the PL and PR motifs, respectively. It is noteworthy that the WW domain of HYBP bound to the PL and PR motifs with similar \( K_D \) values. The similar binding strength to both the PL and PR motifs was also observed in the other Group II WW domains (FBP11A, WAC, Fe65, and Fe65L2) reported here, because the differences of \( K_D \) values were at most only 4.1 times. The controversial group classification of Fe65 was clarified in the present report. Our results showed that Fe65 best preferred the PP peptide, the Group II ligand, but also bound to the PR motif, the Group III ligand, with the small difference of \( K_D \) values (2.5 times). We should consider the PP peptide as one of the general ligand motifs for the WW domains because those of HYBP, Fe65, and Fe65L2 showed significantly stronger binding to PP than to the PL motif.

**SPR Binding Assay of WW Domains Previously Classified as Group III**—The WW domain of FBP30A, a Group III WW domain, expectedly bound most strongly to the PR motif, the typical Group III ligand, and unexpectedly bound to the PL and PP motifs, the Group II ligands, to a considerable extent (Fig. 1C). FBP30A bound to the PP and formin peptide with only 3.3 and 8.0 times larger \( K_D \) values, respectively, than to the WBP11 peptide. It was also found that FBP30A bound to the PL motif more strongly than FBP11A, a Group II WW domain. Our data on the Group II and III WW domains clearly indicate that Group II and III WW domains share their ligands, the PL, PR, and PP motifs, and that in all cases their ligand preferences are very similar or almost indistinguishable.

**SPR Binding Assay of WW Domains Previously Classified as Group IV**—It is confirmed here that the Group IV WW domains bind specifically to peptides containing the pS/pT-P motif. The WW domains of Pin1 and Y110 were previously classified as Group IV (2, 5, 20). In the present study we confirmed that WW domain of Pin1 and Y110 bound most strongly to the cognate peptides derived from the Cdc25c and Myt1 proteins (Fig. 1C). The WW domain of Pin1 interacted weakly with the WBP11 peptide, a Group III ligand, with a \( K_D \) value of 1.13 mM, i.e. 44 times less strongly than with the Cdc25c peptide. Thus, the bindings of the Group IV WW domains were highly specific to the Group IV ligands. It should be noted that none of the Group IV WW domains bound to the PP motif.

**Kinetics of WW Domains Previously Classified as Group II or III**—The kinetic parameters of the Group II and III showed the extremely large \( k_{off} \) values, i.e. the fast dissociation (Fig. 1D). Furthermore, the Groups II and III WW domains share common nature not only in terms of the specificity of ligand recognition but also in the binding kinetics. The kinetic parameters of ligand binding were compared for WW domains from Groups II and III (Fig. 1D). All of these WW domains showed a common tendency in the rate constants for both association and dissociation; those WW domains associated faster with the PP motif than with the PL and the PR motifs and dissociated from the PL motif faster than from the PP and PR motifs. It is noteworthy that the WW domain of HYBP bound to the PL and PR motifs with similar \( k_{off} \) values. The similar binding strength to both the PL and PR motifs was also observed in the other Group II WW domains (FBP11A, WAC, Fe65, and Fe65L2) reported here, because the differences of \( k_{off} \) values were at most only 4.1 times. The controversial group classification of Fe65 was clarified in the present report. Our results showed that Fe65 best preferred the PP peptide, the Group II ligand, but also bound to the PR motif, the Group III ligand, with the small difference of \( k_{off} \) values (2.5 times). We should consider the PP peptide as one of the general ligand motifs for the WW domains because those of HYBP, Fe65, and Fe65L2 showed significantly stronger binding to PP than to the PL motif.
thy that the WW domain of HYPB associated with the PP motif extremely fast, with the rate constant of $2.2 \times 10^7$ M$^{-1}$ s$^{-1}$. This relative speed of association must be one of the underlying features of its high specificity to the PP motif.

Proteomic Mapping—We also performed the proteomic mapping assay for the selected WW domains examined above and confirmed the common ligand predilection of the Group II and III WW domains. From Table I, we can see that all three domains have much lower number of more strongly binding ligands (optical density, $2.5$) from the PY-containing peptides of the Group I. The percentage of more strongly binding ligands in Group I is in the range of $1–2\%$, whereas the percentages of more strongly binding ligands containing the PL, PP, and PR motifs are in the range of $5–8\%$, $10–23\%$, and $5–6\%$, respectively (Tables I and II). Thus, these three WW domains commonly prefer the PP motif best, the PL and PR motifs second best, and the PY motif worst. It is noteworthy that these WW domains prefer the PL (the Group II ligand) and PR (the Group III ligand) motifs to a similar extent. The data indicate that the difference in ligand preference of these three WW domains is almost indistinguishable. These results are consistent with the data from the SPR assay as described above.

The GST fusion proteins of the WW domains were used in this proteomic mapping. However, the problem of avidity was probably significantly minimized because the fused WW domains were immobilized and not presented as free flowing samples in solution.

Comparison of the Ligand-binding Sides of Groups I-IV WW Domains—The results of our survey prompted us to re-examine ligand-binding pockets of WW domains in hope of finding structural denominators common for Group II and III WW domains. Groups I and IV WW domains employ the same side of their bent-sheeted structure to recognize their respective ligands (Fig. 2, A–C). We assumed that the WW domains previously classified as Groups II and III would also use the same side to recognize their ligands. We constructed molecular models of the Group II and III WW domains (FBP11A, Fe65L2, and FBP30A) based on the sequence homology and compared their putative ligand-binding sides with those of Group I and IV WW domains to reveal the structural basis of their ligand specificity.
Ligand Recognition by Group III/III WW Domains

We have defined the common numbering system for residues of WW domains to facilitate the comparison of WW domains (see Fig. 4B). As was the case for Group I and IV WW domains, the XP groove was also found on the putative ligand-binding side of the WW domains of FBP11A, Fe65L2, and FBP30A. This groove is formed by highly conserved aromatic residues, Tyr23 and Trp25; it is likely essential for ligand recognition by Group II and III WW domains, because all of their ligands contain Xaa-Pro segments (hence the name of the groove, XP).

Group I and IV WW domains have respective unique ligand-binding subsites, the Tyr-binding groove (Tyr groove) and the phosphate-binding patch (p patch), which determine their ligand specificity (Fig. 2, A–C) (8, 20, 33). The Tyr groove in Group I WW domains is formed by three residues, Ile/Leu/Val25 (aliphatic), His27, and Lys/Arg/Gln30 (containing C\(^{\alpha}\)H\(_2\)-C\(_{\beta}\)), and is required for recognition of the Tyr side chain in the PY motif of the ligand (see Fig. 2A). None of the WW domains of FBP11A, Fe65L2, or FBP30A can form the Tyr groove because they have neither an aliphatic residue at position 25 nor His at position 27 (Fig. 2, D–F). The p patch in Group IV WW domains is formed by three residues, Ser16, Arg/Lys/Asn/Gln17, and Tyr23, and is required for recognition of the phosphate group in the pS/pT-P motif (Fig. 2C). None of the three WW domains examined: FBP11A, Fe65L2, or FBP30A WW domains could form a p patch because they contain substitutions of the “needed” Ser at position 16 or lack a residue at position 17 (see Fig. 4B).

It is predicted that Group II and III WW domains have a similar ligand recognition surface because their specificity and kinetics in ligand recognition have been shown to be very similar. The modeled structures of Group II and III WW domains actually show common structural features within the putative ligand recognition surface: a groove formed by two aromatic rings at positions 23 and 25 and negative charge(s) in Loop I (Fig. 2, D–F). The two aromatic residues, Tyr23 and Tyr/Trp25, form a groove that resembles the XP groove. To confirm the functional importance of these residues, we carried out site-directed mutagenesis at positions 25 and 16 probing the new groove and the negative charge in the Loop I. We found that the W25V mutant of the Fe65L2 WW domain had a decreased affinity for all of the Group II and III ligands examined. It bound only to the PP motif seven times less strongly than its wild type. On the other hand, the L25W mutant of the WW domain of hYAP65(WW1), which belongs to Group I, acquired the ability to bind to the PP motif (Fig. 3A) (34). The KD value of this binding was 400 \(\mu\)M, which was only 1.5 times larger than that of the WW domain of FBP11A (Figs. 1C and 3A). These data demonstrate the importance of Trp25 for Group II and III WW domains to recognize their ligands, especially the PP motif. Interestingly, the D16A and D16N mutations in the WW domain of Fe65L2 did not affect the ligand binding (data not shown). Thus, the negative charge on Loop I is not important for the ligand binding of Group II and III WW domains. Therefore, Group II and III WW domains recognize their ligands (the PL, PP, and PR motifs) by the XP groove and the adjacent similar groove. Because the ligands of Group II and III WW domains commonly possess two or more Xaa-Pro segments, we propose that the XP groove formed by Tyr23 and Tyr/Trp25 recognizes a second Xaa-Pro segment in a ligand. We thus name this groove between the residues at positions 23 and 25 as the XP2 groove (meaning the second XP binding groove; Figs. 2 and 3).

**DISCUSSION**

After noticing that several WW domains of Group II or III were assigned to either of the two groups of ligand recognition by different investigators, we decided to test the hypothesis that Groups II and III WW domains represent one larger class that also shares similarity with SH3 domains. We elected to use three independent approaches: the SPR binding assay, the analysis of the proteomic mapping of WW domains (23), plus detailed analysis of WW domain structures. Our general conclusion is that the majority (perhaps all) of WW domains that were initially assigned to either Group II or III WW domains belong to one large group of domains with a binding pocket that is generally similar to that of SH3 domains.

It was previously reported that the WW domains of FBP11A and FBP30A bound specifically to the PL motif and the PR motif, and thus they have been classified as Groups II and III, respectively (16, 22, 26). However, we have shown here by SPR binding experiments that the WW domains of FBP11A, WAC, HYBP, Fe65, Fe65L2, and FBP30A are able to bind both the PL and PR motifs as well as simple polyproline with \(K_D\) values so close to each other that their ligand preferences should be considered as similar. Furthermore, all of the above WW domains have similar kinetic properties with regard to both ligand association and dissociation. Our results provide experimental evidence in support of the proposal from the Macias and Sudol laboratories, on the similarity among WW domains of
Proteins contain the sequences that Group II/III WW domains can bind. The protein containing polyprolines have versatile functions such as aromatic rings of Tyr 23 and Phe 25 in Group IV WW domains shown in Fig. 4. This new three-group classification is more appropriate than the four-group classification proposed previously (2, 9). The Group II/III WW domains commonly bind the PL, PP, or PR motifs, all of which have two or more Xaa-Pro segments. Two Xaa-Pro segments are recognized by the two Xaa-Pro binding grooves (the XP and XP2 grooves) formed by three aromatic rings with nearly parallel alignment (Trp, Tyr, and Trp/Tyr).

Interestingly, Group II and III WW domains showed extremely large $k_{on}$ values, the dissociation rate constants, compared with the other domains (28, 35, 36). Thus, the Group II and III WW domains dissociate fast, which in isolated domains and perhaps in the context of full-length proteins may translate into a relatively weak binding of those WW domains compared with the domains of other classes.

The XP2 Groove as a Common Functional Patch of Group II and III WW Domains—Our molecular models show that both the Group II and III WW domains have the XP2 groove as the ligand-binding subsite for their specific ligand recognition. The structural resemblance of the XP and XP2 grooves suggests that the XP2 groove is also involved in the recognition of an Xaa-Pro (including Pro-Pro) segment at positions 1–4 of the PPII helix. If an Xaa-Pro segment at positions 4–5 of the PPII helix stacks onto the XP groove, the second Xaa-Pro segment at positions 1–4 should come onto the XP2 groove and hence the XP2 groove for ligand binding by the Group II and III WW domains. We consider the XP2 groove to be the structural determinant (epsilon determinant; see Ref. 37 for definition) of the Group II and III WW domains for their ligand recognition.

Group IV WW domains also have a groove that might look like the XP2 groove because of the two highly conserved aromatic residues at positions 23 and 25 (Tyr and Phe). But, unlike Group II and III WW domains, they are not able to bind the PL, PP, or PR motifs. In the crystal structure of a Group IV WW domain complexed with its cognate ligand (38), the aromatic ring of Phe is not parallel with but is nearly perpendicular to the aromatic ring of Tyr. Thus, the two aromatic rings of Tyr and Phe in Group IV WW domains do not form a groove suitable for recognition of the Xaa-Pro segment.

More Appropriate Functional Classification of WW Domains: Groups I, II/III, and IV—We have demonstrated the similar features of Group II and III WW domains with respect to their function and structure. We therefore propose that the WW domains previously classified into two distinct groups, Groups II and III, should be classified into a single group called Group II/III. The criteria for the new classification of WW domains are shown in Fig. 4A. This new three-group classification is more appropriate than the four-group classification proposed previously (2, 9). The Group II/III WW domains commonly bind the PL, PP, or PR motifs, all of which have two or more Xaa-Pro segments. Two Xaa-Pro segments are recognized by the two Xaa-Pro binding grooves (the XP and XP2 grooves) formed by three aromatic rings with nearly parallel alignment (Trp, Tyr, and Trp/Tyr).

The ligand specificity of functionally uncharacterized WW domains can be predicted based on the criteria shown in Fig. 4A. We have classified the 200 WW domains listed on the EMBL Web site (www.bork.embl-heidelberg.de/Modules/ww_family_smart.html) and have successfully predicted the ligand specificity for 174 (87%) WW domains; 112 (56%), 54 (27%), and 8 (4%) WW domains satisfy all the criteria to be classified as Groups I, II/III, and IV, respectively, whereas the other 26 (13%) WW domains remain unclassified (Fig. 4B).

Structural Requirements of Ligands for Group II/III WW Domains—Computational simulations suggest that polyproline should form the polyproline type II (PPII) helix, a left-handed three-residue-per-turn helical structure extending the pyrrolidine rings of prolines outside, in aqueous solution (39). In fact, the SH3 domains and profilin bind the proline-rich segments. The two Xaa-Pro segments are recognized by the two Xaa-Pro binding grooves (the XP and XP2 grooves) formed by three aromatic rings with nearly parallel alignment (Trp, Tyr, and Trp/Tyr).

The ligand specificity of functionally uncharacterized WW domains can be predicted based on the criteria shown in Fig. 4A. We have classified the 200 WW domains listed on the EMBL Web site (www.bork.embl-heidelberg.de/Modules/ww_family_smart.html) and have successfully predicted the ligand specificity for 174 (87%) WW domains; 112 (56%), 54 (27%), and 8 (4%) WW domains satisfy all the criteria to be classified as Groups I, II/III, and IV, respectively, whereas the other 26 (13%) WW domains remain unclassified (Fig. 4B).

Structural Requirements of Ligands for Group II/III WW Domains—Computational simulations suggest that polyproline should form the polyproline type II (PPII) helix, a left-handed three-residue-per-turn helical structure extending the pyrrolidine rings of prolines outside, in aqueous solution (39). In fact, the SH3 domains and profilin bind the proline-rich segments forming the PPII helix (40, 41). Thus, the Group II/III WW domains are expected to bind the ligands that form the PPII helix. If an Xaa-Pro segment at positions 1–2 in the PPII helix stacks onto the XP groove, the second Xaa-Pro segment at positions 4–5 should come onto the XP2 groove (Fig. 3, B and C). The two Xaa-Pro segments are recognized by the two Xaa-Pro binding grooves (the XP and XP2 grooves) formed by three aromatic rings with nearly parallel alignment (Trp, Tyr, and Trp/Tyr). This model, a ligand for Group II/III WW domains should form well to the ridges and grooves of the PPII helix. According to this model, a ligand for Group II/III WW domains should be at least 5 residues long, whereas profilin needs at least 10 successive Pro residues to bind (42). In fact, the WW domain of Fe65, a Group II/III WW domain, has been reported to bind 6-residue-long polyproline peptides (34). A PPII helix may be able to bind in the reverse direction according to our model because it has the pseudo-2-fold symmetry of polyprolines. In fact, a PPII helix can bind profilin and SH3 domains in two distinct orientations (35, 44).

Some of the Group II/III WW domains have been reported to bind the PGM and PGR motifs (10, 21). From the alignment of the motifs that can bind Group II/III WW domains, the core
consensus sequence, Xaa$^1$-Pro$^2$-Xaa$^3$-Pro$^4$-Pro$^5$, is deduced (Fig. 3D). The sequence contains Xaa$^1$-Pro$^2$ and Pro$^4$-Pro$^5$ segments that should bind the two XP grooves on the surface of Group II/III WW domains.

It is expected that the XP grooves can recognize only Xaa-Pro sequences. The XP groove of SH3 domains binds only Xaa-Pro sequences because the sequence has a unique hairpin-like structure formed by a general C$^\epsilon$-substituted residue, Xaa, and a unique N-substituted residue, Pro (17, 19, 45). It is therefore predicted that the residues at positions 2' and 5', which are the second residues of the respective XP segments on ligands, should always be Pro. The specific topology of the two XP grooves for recognizing the Xaa-Pro sequences is most likely responsible for the fast and subtle binding property of the Group II/III WW domains. Originally the XP groove was proposed to explain the complexes formed by the SH3 and WW domains and their cognate ligands (17). This special and almost universal topology of XP grooves is thought to enable the formation of delicate yet specific protein complexes that conduct many cellular processes.

Comparison with the SH3 Domains in Terms of Structure and Ligand Specificity—The SH3 domains have two XP binding grooves, a very similar arrangement to the Group II/III WW domains (17, 19), whereas the Group I and IV WW domains have only one XP binding groove. Thus, the ligand recognition by Group II/III WW domains may be more similar to that by the SH3 domains than to that by Group I and IV WW domains (18).

The consensus sequence of the ligands for SH3 domains contains Pro-Xaa-Xaa-Pro, in which the two Pro residues separated by two residues play a key role in the binding. The same pattern of the sequence is also found in the consensus sequence of the ligands for Group II/III WW domains. SH3 domains also require an Arg or Lys residue in their ligand sequences for strong binding; the replacement of the Arg diminishes the affinity by at least 10-fold (40, 46). In contrast, Group II/III WW domains do not need an Arg in the ligand, as revealed by our SPR and protomic mapping described here as well as by methylation experiments of Arg residues (22). The methylation of Arg residues in the Sam68 peptides abrogates binding the SH3 domains of PLC$\gamma$ and Fyn, whereas the binding to WW domain of FBP30A is not affected. The two XP grooves in the Group II/III WW domains are located closer to each other than those in the SH3 domains, which could make the binding of the Group II/III WW domains to the PP motif tighter than that of the SH3 domains. Group II/III WW domains and SH3 domains might compete for the possible ligands in vivo because the consensus sequences of their ligands are common to some extent (18, 22).

Functional Importance of Group II/III WW Domains and the PP Motifs in Intracellular Signaling Networks—The WW domain of Fe65 has been reported to bind to the PL motif of Mena, a cytoskeletal protein involved in cytoskeletal reconstruction (14, 15). But Mena contains not only the PL but also PP motif. Our data from the SPR binding experiments suggest that the interaction between Fe65 and Mena is mediated more by the PP than the PL motif. On the other hand, FBP11, HYPB, and HYPC have been reported to bind Huntingtin (7). Huntingtin contains only the PP motif but neither the PL nor PR motif. Thus, the PP motif should act as an adapter motif in vivo. Simple polyproline is one of the most frequently found motifs in the genomes of fruit fly, worm, and yeast (47). In fact, a large number of polyproline-containing proteins are known to play crucial roles in development, cell movement, cytoskeleton reconstruction, and many other major cellular and physiological processes (Fig. 5).

In the present study, we have shown that the Group II/III WW domains are able to bind simple polyprolines, unlike the Group I and IV WW domains. We have found 54 Group II/III WW domains of 200 WW domains listed in a public data base, suggesting that at least 25% of WW domains play important roles in the recognition of simple polyprolines, the PP motif, as well as the PL and PR motifs. This work elucidates the structural and functional aspects of Group II/III WW domains and their implicit importance in intracellular signaling networks mediated by polyprolines.

Acknowledgments—We thank Dr. H. Tanahashi for the cDNA of Fe65L2, Dr. C.-K. Shen for the expression system of nNedd4(WW3) and hYAP65(WW1), and Dr. M. T. Bedford for the expression plasmid of FBP11A and the cDNA of FBP30A.
44. Mahoney, N. M., Rozwarski, D. A., Fedorov, E., Fedorov, A. A., and Almo, S. C. (1999) Nat. Struct. Biol. 6, 666–671.
45. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N., and Lim, W. A. (1998) Science 282, 2088–2092.
46. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) Science 266, 1241–1247.
47. Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Har Harper, I. R., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brutstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, R. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O’Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H., and Lewis, S. (2000) Science 287, 2204–2215.
Common Mechanism of Ligand Recognition by Group II/III WW Domains: REDEFINING THEIR FUNCTIONAL CLASSIFICATION
Yusuke Kato, Koji Nagata, Mihoko Takahashi, Lubing Lian, Juan J. Herrero, Marius Sudol and Masaru Tanokura

J. Biol. Chem. 2004, 279:31833-31841.
doi: 10.1074/jbc.M404719200 originally published online May 7, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M404719200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 47 references, 19 of which can be accessed free at http://www.jbc.org/content/279/30/31833.full.html#ref-list-1