Determinants of Ligand Specificity in Groups I and IV WW Domains as Studied by Surface Plasmon Resonance and Model Building

Yusuke Kato, Mie Ito, Kunji Kawai, Koji Nagata, and Masaru Tanokura

From the Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, the Center for Analytical Chemistry and Science, Inc., Yokohama Research Center, Mitsubishi Chemical Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama 227-0033, Japan, and the Biotechnology Research Center, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

WW domains are universal protein modules for binding Pro-rich ligands. They are classified into four groups according to their binding specificity. Arg-14 and Arg-17, on the WW domain of Pin1, are thought to be important for the binding of Group IV ligands that have (Ser(P)/Thr(P))-Pro sequences. We have applied surface plasmon resonance to determine the ligand specificity of several WW domains containing Arg-14. Among these WW domains, Rsp5.2 and mNedd4.3 bound only to the Group I ligand containing Pro-Pro-Xaa-Tyr with $K_D$ values of 11 and 55 $\mu$M, respectively. The WW domains of hPin1, Caenorhabditis elegans Pin1 homologue (Y110), PinA, and SspI bound to Group IV ligands with $K_D$ values ranging from 22 to 700 $\mu$M. PinA and SspI do not have Arg-17, unlike Pin1 and Y110. The modeled structures of the WW domains of PinA and SspI revealed that the structure and the network of hydrogen bonds of Loop I, which are also formed in Pin1 and Y110, are conserved. We propose that this configuration of Loop I (referred to as the “p patch”) is necessary for binding Group IV ligands and that it can be used to predict the specificity and functions of other WW domains.

A WW domain is a small structural motif named after a pair of tryptophan residues that are highly conserved (1, 2). It works as an interaction module in proteins with various functions, including cell cycle control (Pin1/Ess1), ubiquitin ligation (Nedd4/Rsp5 and Smurfl), and coactivation of transcription (Nedd4/Rsp5 and Smurf1), and coactivation of transcription (SPR) to examine the ligand specificity of WW domains containing Arg-14. SPR is becoming a widespread method to measure quantitatively biomolecular interactions. WW domains that do not have Arg-17 (such as PinA and SspI) bound to Group IV ligands, as did WW domains that do have Arg-17 (such as Pin1 and the Caenorhabditis elegans Y110A2AL.13 gene product (Y110)). Arg-17 has been thought to be the most important residue for this binding. Thus, we have built structural models of these WW domains to elucidate the mechanism for recognition of ligand peptides with sequences of (Ser(P)/Thr(P))-Pro, and we have proposed the $p$ patch, a new structural patch of Group IV WW domains, which interacts with their ligands. After summarizing these new findings, we have proposed the determinants by which Group IV WW domains function as mitotic arrest (7, 8). hPin1 binds to mitotically phosphorylated proteins such as Cdc25c, Myt1, Plk1, and Cdc27 through its WW domain (7), and it also binds to the phosphorylated RNA pol II C-terminal domain (CTD), implying transcriptional regulation by Pin1 (8).

Several structures have been determined so far for WW domain-ligand complexes, three of Group I and one of Group IV. These structures reveal that WW domains each adopt a triple-stranded antiparallel $\beta$-sheet. For each domain, the same side of the $\beta$-sheet is used to bind all of its respective ligands (2, 6, 9). The crystal structure of the hPin1 WW domain, complexed with a phosphorylated CTD peptide of RNA polymerase II (Tyr-Ser(P)-Pro-Thr-Ser(P)-Pro-Ser), has shown the importance of two arginine residues (Arg-14 and Arg-17) on this side of the sheet in recognizing the ligand. Arg-14 of the WW domain recognizes Pro-3 of the ligand peptide, whereas Arg-17 interacts with the phosphate group of Ser(P)-5' (9). Thus, the arginine residues on the ligand-binding side may determine the ligand specificity of the Group IV WW domains. However, ligand specificity of parvulin-type PPIases other than hPin1 has not been examined so far. The question of whether the arginine residue is essential remains to be addressed because some WW domains of parvulin-type PPIases do not have Arg-17.

To elucidate the roles of these arginine residues on ligand specificity in more detail, we used surface plasmon resonance (SPR) to examine the ligand specificity of WW domains containing Arg-14. SPR is becoming a widespread method to measure quantitatively biomolecular interactions. WW domains that do not have Arg-17 (such as PinA and SspI) bound to Group IV ligands, as did WW domains that do have Arg-17 (such as Pin1 and the Caenorhabditis elegans Y110A2AL.13 gene product (Y110)). Arg-17 has been thought to be the most important residue for this binding. Thus, we have built structural models of these WW domains to elucidate the mechanism for recognition of ligand peptides with sequences of (Ser(P)/Thr(P))-Pro, and we have proposed the $p$ patch, a new structural patch of Group IV WW domains, which interacts with their ligands. After summarizing these new findings, we have proposed the determinants by which Group IV WW domains function as mitotic arrest (7, 8). hPin1 binds to mitotically phosphorylated proteins such as Cdc25c, Myt1, Plk1, and Cdc27 through its WW domain (7), and it also binds to the phosphorylated RNA pol II C-terminal domain (CTD), implying transcriptional regulation by Pin1 (8).

The abbreviations used are: pol II, polymerase II; CTD, C-terminal domain (of RNA polymerase II); ENaC, epithelial sodium channel; h, human; Dmd, Duchenne muscular dystrophy; mNedd4.3, the third WW domain of mouse Nedd4; PPIase, peptidyl-prolyl cis-trans isomerase; Rsp5.2, the second WW domain of S. cerevisiae Rsp5; SPR, surface plasmon resonance; WBPI, WW domain-binding protein 1; GST, glutathione S-transferase; MALDI-TOF-MS, matrix-assisted laser desorption ionization-time of flight-mass spectrometer; Fmoc, N-(9-fluorenylmethoxycarbonyl; Y110, Y110A2AL.13 gene product from C. elegans.
bind to the (Ser(P)/Thr(P))-Pro ligands, and we have compared the determinants of Group IV to those of Group I as deduced from the previously determined structures. By applying these new criteria and alignment analysis, we can predict the specificity of WW domains of Group I or IV.

**Experimental Procedures**

**Plasmid Construction**—The DNAs encoding the WW domains of human Pin1, C. elegans Y110A2AL.13 gene product (Y110), Aspergillus nidulans PinA and Neurospora crassa SspI were synthesized and amplified by a modified PCR method in which several synthetic oligonucleotides were used as the template for the cycle. The DNA encoding the second WW domain of Saccharomyces cerevisiae Rsp5 (Rsp5.2) was amplified by PCR with its cDNA as a template, which was kindly provided by Dr. M. Sudol. The PCR products of each WW domain were inserted between the BamHI and EcoRI sites of the expression vector pGEX-4T1 (Amersham Biosciences). The expression plasmid for GST fusion proteins of the third WW domain of mouse Nedd4 (mNedd4.3) was kindly provided by Dr. C.-K. Shen.

**Expression and Purification**—All the GST fusion WW domains were expressed in *Escherichia coli* BL21(DE3)/pLysS (Novagen) at 37 °C and purified by affinity chromatography with glutathione-Septarose (Amersham Biosciences). The GST tag was removed by thrombin (Sigma) digestion at 25 °C. Then the WW domains were separated from GST and thrombin by reverse-phase chromatography with a Resource R 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). Molecular masses were verified by MALDI-TOF-MS on a Voyager mass spectrometer (PerkinElmer Life Sciences).

**Peptide Synthesis**—Pro-containing ligand peptides were synthesized by solid-phase peptide synthesis on a PSSMS Peptide Synthesizer (Shimadzu, Kyoto, Japan). Rink amide AM resin (Novabiochem) and Fmoc amino acids with protected side chains were used. For phosphoserine and phosphothreonine, Fmoc-Ser(PO(OBzl)OH)-OH and Fmoc-Thr(PO(OBzl)OH)-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, 2% ethylmethylsulfide, 2.5% ethanethiol, 2% thionphenol), and then they were purified by reverse-phase chromatography with a preparative column ODS-ApO303 (4.6 x 250 mm) (YMC, Kyoto, Japan). Molecular masses were verified by MALDI-TOF-MS on a Voyager mass spectrometer (PerkinElmer Life Sciences). The designed sequence of the ligand peptides was GTPPPPTTVG (WW domain-binding protein 1 (WBFI)), mouse PPPPPP PPPPPP (14-mer polyproline (PolyPro)), GVPpTPP (Cdc25c-T48, human), KGPPpA TP (Myt1-T412, human) and KGpYSSpTPS (CTD-S5, human).

**SPR Binding Assay**—We measured SPR using a BiAcore 2000 (Biacore AB). Each ligand peptide was immobilized on a flow cell of a sensor chip CM5 (Biacore AB) by the standard EDC/NHS method recommended by Biacore AB, with a 1 mg/ml peptide solution dissolved in 25 mM NaHCO3. For reference SPR signals, 50 mM NaHCO3 was used instead of a peptide solution in the immobilization reaction. The remaining active carboxylates were inactivated with ethanolamine. The binding experiments were performed at 25 °C with the WW domains 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 from the absorbance at 280 nm of the protein solutions. The dissociation constants (*Kd*) of the WW domains were determined by Scatchard plot analysis (11, 12). A brief explanation of the Scatchard plot method in the present study is as follows. Raw sensorgrams of flow cell without ligand peptide were subtracted from the corresponding raw sensorgrams of ligand-bound flow cell for each sensorgram before Scatchard plot analysis. RU/*C* values were plotted against RU/*A* and the best fit to the data, RU/*C* = -1/1*Kd* × RU/*A* + A, where A is a constant.

**Molecular Modeling**—Homology modeling was carried out using Swiss-Model (13), an automated protein modeling server combining homology search, model building, and energy minimization, by inputting the amino acid sequences and the atomic coordinate of the template structures. The crystal structure of hPin1 in complex with CTD-S2/S5 peptide (PDB code 1F8A) (9) was used as the template for modeling the WW domains of Y110, PinA, and SspI. The NMR structure of mNedd4.3 WW domain in complex with epithelial sodium channel (ENaC) peptide (PDB code 1H5H) (10) was used as the template for modeling the WW domain of Rsp5.2. Molecular surfaces and electrostatic potentials were calculated and displayed using MOLMOL (14). Swiss PDB Viewer was used to draw the backbone and to calculate the hydrogen bonds (13).

**Results and Discussion**

**SPR Binding Assay**—The WW domains of mNedd4.3, Rsp5.2, hPin1, Y110 (15), PinA (16), and SspI (17) conserve Arg-14 (Fig. 1). The WW domains of hPin1 and Y110 also have Arg-17. The amino acid residue at position 14 is shown to be on the binding side of all the WW domains whose structures are revealed (2, 6, 9, 10). Thus, Arg-14 is expected to be on the binding side of the WW domains of Rsp5.2, Y110, PinA, and SspI. The binding specificity of the WW domains containing Arg-14 was examined by SPR (Fig. 2). The binding curves of respective WW domains examined in the present study showed so rapid a transition to steady states that we could not obtain enough data points during transition of binding or dissociation (Fig. 2A). Thus, the determination of *Kd* was carried out by Scatchard plot analysis instead of by calculation of *Kd* from *k* on and *k* off obtained by direct curve fitting (Fig. 2B). The WW domains of Y110, PinA, and SspI as well as hPin1 bound specifically to Group IV ligands such as Cdc25c-T48, Myt1-T412, and CTD-S5 (Table I), whose consensus sequence is (Ser(P)/Thr(P))-Pro (9). The binding of hPin1 to Cdc25c-T48, Myt1-T412, and CTD-S5 has Ser(P) and Thr(P) residues, respectively, which were remarkably stronger than other binding to Group IV ligands. Y110 and hPin1 bind to CTD-S5 with relatively similar *Kd* values. The Thr(P) residues of Cdc25c-T48 and Myt1-T412 might cause large differences in the *Kd* values because only CTD-S5 has Ser(P). The *Kd* values of Y110, PinA, and SspI with the CTD-S5 peptide were 510, 700, and 430 μM, respectively, which were much larger than other *Kd* values. Because the intact CTD has 24 repeats of the Ser-Tyr-Ser-Pro-Thr-Ser-Pro sequence, its *Kd* value should become much smaller than those in Table I. Likewise, multiple (Ser/Thr)-Pro sites can be phosphorylated in each sequence of mitotic proteins including Cdc25c, Myt1, Plk1, Cdc27, or NIMA (7).

On the other hand, the WW domains of mNedd4.3 and Rsp5.2 bound specifically to WBPI (2), a Group I ligand, with *Kd* values of 55 and 11 μM, respectively (Table I). None of the WW domains examined in the present study bound to PolyPro (Table I). Thus, the WW domains with Arg-14 are classified functionally into two groups, I and IV. We have identified the WW domains of PinA and SspI as a new type of Group IV WW domain because both are capable of binding specifically to the Ser(P)- or Thr(P)-containing sequences, although Arg-17, the most important residue for binding in hPin1 (9), is replaced (Fig. 1).

Based on the ligand specificity and the position of Arg resi-
due(s), the Group I and IV WW domains can each be divided into two subgroups, Ia (dystrophin, hAP65, etc.), Ib (mNedd4.3, Rap5.2, etc.), IVA (hPin1, Y110, etc.), and IVB (PinA, Ssp1). Subgroup Ia members have no positively charged residue at positions 14 or 17, whereas Ib members have Arg-14 (Fig. 1). IVA members have both Arg-14 and Arg-17, whereas IVB members, like Ib members, have Arg-14 but not Arg-17 (Fig. 1).

**Comparison of the Ligand-Binding Site**—WW domains use the same side of their bent-sheeted structure to bind respective ligands. We have made homologous models and compared the ligand binding sides of Group I WW domains against that of Group IV domains to analyze the structural basis of the ligand specificity between these two groups. Charge distributions on the binding side of subgroups Ib and IVB are very similar because they have common basic residues at positions 14 and 30 (Fig. 3, A, B, F, and G). On the other hand, charge distributions of subgroups Ia and Ib are totally different (Fig. 3, A, D, and B). Thus, there is little relationship between charge distribution and ligand specificity in Group I. Similarly, Arg-14 and (Arg/Lys)-30 of Group IVB should not be important for the ligand specificity in binding to the (Ser(P)/Thr(P))-Pro motif because WW domains with different specificities share these residues.

The X-P-binding groove, the common patch between Groups I and IV WW domains, recognizes the Pro residues of their ligands (Fig. 3) (18). In addition, the Tyr binding pocket, which is shown in the solved structures of dystrophin and mNedd4.3 (Fig. 3, A and B) (6, 10), should be an important requirement for all Group I WW domains, considering the Tyr residue of the PXFY motif. The Tyr binding pockets of dystrophin and mNedd4.3 are formed commonly by residues at positions 25, 27, and 30 (Fig. 3, A and B). Thus, the Tyr binding pockets of other Group I WW domains should consist of (Ile/Leu/Val)-25 (aliphatic), His-27, and (Lys/Arg/Gln)-30 (containing $-\text{C}_{2}\text{H}_{2}-\text{C}_{2}\text{H}_{1}$), because the model structure of Rsp5.2 formed a pocket similar to those of the solved structures (Figs. 1 and 3, A, B, and D).

**Binding Patch for Binding to the Phosphate Group**—In the case of Group IV WW domains, the hPin1 WW domain, belonging to subgroup IVA, recognizes the phosphate in the (Ser(P)/Thr(P))-Pro motif by the side chains of Ser-16, Arg-17, and Tyr-23 and the backbone amide of Arg-17 (Fig. 4A) (9). The most energetically important interaction seems to be the electrostatic contact between the phosphate of Ser(P)-5' and the side chain ε-amino group of Arg-17 (9). However, the subgroup IVB WW domains replace Arg-17 with Asn/Gln, although PinA and Ssp1 bound to the (Ser(P)/Thr(P))-Pro peptides with $K_D$ values similar to that of Y110, which conserves Arg-17 (Fig. 1 and Table I). Like Arg-17, the side chain of (Asn/Gln)-17 of subgroup IVB can act as a hydrogen bond donor for the phosphate of Ser(P)-5', according to the model structures (Fig. 4B). In addition, the positive charge of Lys-19 in the subgroup IVB WW domains can contribute to this binding through an electrostatic interaction (Figs. 3, F, and G, and 4B). These findings suggest that the combination of the hydrogen bonding by (Asn/Gln)-17 and the positive charge of Lys-19 can compensate for the loss of the electrostatic interaction by Arg-17.

Loop I of the Group IV WW domains is longer by one residue than that of the Group I WW domains because of the insertion of (Arg/Asn/Gln)-17 (Fig. 1). This allows them to adopt a particular conformation of Loop I to recognize the phosphate group by Ser-16 and (Arg/Asn/Gln)-17 (referred to as the p patch) (Fig. 5, A and B). Furthermore, the side chain of Ser-16 forms a network of hydrogen bonds with the backbone of Loop I, thereby stabilizing this conformation (Fig. 4B). Tyr-23, the other residue involved in phosphate binding, is also conserved among all of the Group IV WW domains used in the present study (Fig. 3, E–G). Thus, all Group IV WW domains should

**TABLE I**

Ligand specificity of WW domains containing arginine(s) on the ligand binding side as determined by SPR

| WW | 14/17 | WBP1 | PolyPro | Cdc25c-T48 | Myt1-T412 | CTD-S5 |
|----|-------|------|---------|------------|-----------|-------|
| mNedd4.3 | R/-   | 55 ± 13 | n.b.² | n.b. | n.b. | n.b. |
| Rsp5.2    | R/R   | 11 ± 0.4 | n.b. | n.b. | n.b. | n.b. |
| hPin1     | R/N   | n.b. | n.b. | 22 ± 4 | 29 ± 7 | 170 ± 70 |
| Y110      | R/Q   | n.b. | n.b. | 170 ± 2 | 190 ± 2 | 510 ± 30 |
| PinA      | R/N   | n.b. | n.b. | 190 ± 8 | 210 ± 3 | 700 ± 360 |
| Ssp1      | R/Q   | n.b. | n.b. | 190 ± 40 | 280 ± 60 | 430 ± 270 |

² The column labeled '14/17' indicates kinds of amino acid residues at positions 14 and 17 of each WW domain.

n.b. means that no binding was observed even at an analyte concentration of 0.5 mM. Results of SPR measurements are shown as $K_D$ values (μM).

![Graphs](image-url)
share a p patch as a network of hydrogen bonds and a conformational requirement to bind the (Ser(P)/Thr(P))-Pro motif.

**Prediction of Ligand Specificity and Function**—In addition to the X-P-binding groove, Groups I and IV WW domains have unique binding patches, the Tyr binding pocket and the p patch, respectively. Our results from the binding assay and the molecular modeling support the prediction that these patches are conserved in each group. Thus, we have summarized the criteria for Groups I and IV WW domains (Table II), through which we can predict the ligand specificity of WW domains (Group I, Group IV, or others). All 200 WW domains are aligned automatically by SMART and are listed on Bork’s Web site in.

![EMBL](www.bork.embl-heidelberg.de/Modules/ww_family_smart.html)
binding to the PPXY motif within SMADs because the second domain conserves all the residues listed in Table II, whereas the first domain does not conserve His-27.

Based on these classifications, we predict that all WW domains of parvulin-type PPIases should belong to Group IV (Fig. 6). Inhibition of Pin1 in human tumor cells induces mitotic arrest and apoptosis (20). Both in HeLa cells and Xenopus extracts, the interaction between Pin1 and Cdc25c is cell cycle-regulated (7). This interaction is due to its WW domain (9). Few human parvulin-type PPIases that have WW domains are reported in GenBank™. Thus, these PPIases may compete with each other if they are contained in a single cell, which would explain in part the result that Pin1 knockout mice show no overt phenotype (21). ESS1, a yeast homologue of Pin1, is essential for growth in budding yeast (8) and binds physically to the CTD (19). Thus, it is expected that ESS1 binds to the CTD through its WW domain because that domain belongs to Group IV (Fig. 6).

The CTD of hypothetical protein F36A4.7, a putative RNA pol II of C. elegans, is highly similar to the CTD of human RNA pol II. Thus, nematode RNA pol II may be regulated by Y110 because its WW domain bound to CTD-S5 peptide. We have demonstrated that PinA binds to (Ser(P)/Thr(P))-Pro peptides, which suggests that PinA interacts with NIMA in A. nidulans cells and that PinA may play a role in mitotic regulation.

Acknowledgments—We thank Dr. Marius Sudol for the cDNA of the Rsp5.2 WW domain, Dr. James C.-K. Shen for the expression plasmid of the GST fusion protein of mNedd4.3 WW domain, Dr. Kaori Morimoto (Biacore AB) for technical support, and Dr. Michael E. P. Murphy for discussion.

REFERENCES
1. Sudol, M., and Hunter, T. (2000) Cell 103, 1001–1004
2. Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) Nature 382, 646–649
3. Sudol, M., Siwić, K., and Russo, T. (2001) FEBS Lett. 490, 190–195
4. Zhu, H., Kavek, P., Abdollah, S., Wrana, J. L., and Thomsen, G. H. (1999) Nature 400, 687–693
5. Faber, P. W., Barnes, G. T., Srinidhi, J., Chen, J., Gusella, J. F., and MacDonald, M. E. (1998) Hum. Mol. Genet. 7, 1463–1474
6. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000) Nat. Struct. Biol. 7, 634–638
7. Lu, K. P. (2000) Prog. Cell Cycle Res. 4, 83–96
8. Hanes, S. D., Shank, P. R., and Bottian, K. A. (1989) Yeast 5, 55–72
9. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000) Nat. Struct. Biol. 7, 639–643
10. Kanelis, V., Rotin, D., and Forman-Kay, J. D. (2001) Nat. Struct. Biol. 8, 407–412
11. Lemmon, M. A., Ladbury, J. E., Mandiyan, V., Zhou, M., and Schlessinger, J. (1994) J. Biol. Chem. 269, 31653–31658
12. Oda M., Furukawa K., Sarai A., and Nakamura H. (1999) FEBS Lett. 454, 288–292
13. Guez, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
14. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55
15. Ainscough, R., Bardill, S., Barlow, K., Baslam, V., Baynes, C., Beard, L., Beasley, A., Berks, M., et al. (1998) Science 282, 2012–2018
16. Crenshaw, D. G., Yang, J., Means, A. R., and Kornbluth, S. (1998) EMBO J. 17, 1315–1327
17. Kops, O., Eckerskorn, C., Hottenrott, S., Fischer, G., Mi, H., and Tropschug, M. (1998) J. Biol. Chem. 273, 31971–31976
18. Zarrinpur, A., and Lin, W. A. (2000) Nat. Struct. Biol. 7, 611–613
19. Wu, X., Wilcox, C. B., Devashayam, G., Hackett, R. L., Arevalo-Rodriguez, M., Cardenas, M. E., Heitman, J., and Hanes, S. D. (2000) EMBO J. 19, 3727–3738
20. Lu, K. P., Hanes, S. D., and Hunter, T. (1996) Nature 380, 544–547
21. Fujimori, F., Takahashi, K., Uchida, C., and Uchida, T. (1999) Biochem. Biophys. Res. Commun. 265, 658–663