Cell, Vol. 103, 1001–1004, December 22, 2000, Copyright ©2000 by Cell Press

NeW Wrinkles for an Old Domain

Minireview

Marius Sudol* and Tony Hunter†
*Mount Sinai Medical Center
New York, New York 10029
†The Salk Institute
La Jolla, California 92037

WW domains are protein–protein interaction modules that bind to short proline-rich motifs in proteins that operate in a variety of signaling pathways. Because several of the cognate, proline-rich motifs are terminated by a conserved and potentially phosphorylatable tyrosine, it has been suggested that protein–protein interactions mediated by some WW domains might be modulated by phosphorylation. Moreover, recent data show that some WW domains require phosphorylation of the ligand for binding, specifically, phosphorylation of serine or threonine in Ser-Pro or Thr-Pro containing ligands. The recent high-resolution structures of WW domain complexes have provided molecular details explaining how WW domains recognize diverse and modified sequence motifs. In addition, these structures showed that in certain details of protein ligand binding the WW domain is more similar to the SH3 domain than was originally thought. This minireview focuses on recent data suggesting that in some cases ligand phosphorylation may serve as a switch that regulates WW domain-ligand complexes.

The WW domain is one of the smallest protein modules that folds as a monomer in solution without disulfide bridges or cofactors. Its 40 amino acids form a compact three-stranded, antiparallel β sheet (Macias et al., 1996). The two signature tryptophan (W) residues of the WW domain are spaced 20–22 amino acids apart and play an important role in its structure and function. WW domains bind proteins containing short linear sequence motifs. Based on their ligand predilection, WW domains fall into two major and two minor groups (Table 1). One major group (Group I) binds polypeptides with the minimal core consensus Pro-Pro-X-Tyr, and the other binds ligands with Pro-Pro-Leu-Pro motifs (Group II). The Group III WW domains select polyproline motifs flanked by Arg or Lys, and the Group IV is represented by WW domains with preference for phospho-Ser-Pro or phospho-Thr-Pro containing ligands.

Phosphoserine and Phosphothreonine as Positive Switches

The WW domains of three proteins—Pin1/Ess1p, a *cis-trans* peptidyl prolyl isomerase (PPIase) that is involved in both cell cycle regulation and pre-mRNA 3′ end formation, Nedd4/Rsp5p, a HECT domain E3 ubiquitin ligase that participates in targeted protein degradation, and the splicing factor Prp40p—were all recently shown to bind to proteins phosphorylated on serine or threonine residues (Lu et al., 1999a; Morris et al., 1999; Morris and Greenleaf, 2000; Verdecia et al., 2000). These interactions were convincingly documented as phosphorylation dependent; in the case of Pin1/Ess1p, the WW domain binds short phosphopeptides with a proline C-terminal to the phosphorylated serine or threonine (Table 1). Among the ligands of Pin1/Ess1p are the phosphorylated forms of the Cdc25C phosphatase, the microtubule-associated protein tau and the carboxy-terminal domain (CTD) of RNA polymerase II (Pol II) large subunit (Komuro et al., 1999; Lu et al., 1999a, 1999b; Morris et al., 1999). The fragment of Prp40p spanning its two WW domains bound to phosphorylated CTD of RNA Pol II efficiently, whereas the Rsp5p WW domain region had a capacity to interact with both the phosphorylated and unphosphorylated CTD, showing a clear preference for the latter (Wang et al., 1999; Chang et al., 2000; Morris and Greenleaf, 2000). The WW domain thus joins a group of modules that can bind protein ligands in a phosphorylation-dependent manner, including SH2, PTB, 14-3-3, WD40, FHA, and FF domains.

This unexpected property of the WW domains of Pin1/Ess1p, Prp40p, and Nedd4 family proteins correlates with the proposed function of these proteins. Pin1 can regulate early mitotic events through interaction with several mitosis-specific phosphoproteins including the Cdc25C phosphatase, and the Myt1 and Plk1 protein kinases (Shen et al., 1998; Lu et al., 1999a). The Pin1 PPIase acts on p-Ser/Thr-Pro bonds (Shen et al., 1998). Isomerization of this bond may alter local or global protein conformation and thereby regulate function (e.g., inhibition of Cdc25C phosphatase activity) or facilitate dephosphorylation by PP2A, which requires the bond to be in the *trans* conformation. Interestingly, the Pin1 WW domain can only bind p-Ser-Pro peptides in the *trans* conformation, and it may be used to select targets for isomerization and/or stabilize p-Ser-Pro bonds for subsequent dephosphorylation. Ess1p, the budding yeast homolog of Pin1, interacts with the phosphorylated form of the large subunit of RNA Pol II, and via this interaction it is involved in pre-mRNA 3′ end formation and transcription termination (Morris et al., 1999; Wu et al., 2000). Prp40p is a splicing factor involved in bridging the 5′ and 3′ splice sites during the commitment complex stage of spliceosome formation. The identification of Prp40p as a phospho-CTD RNA Pol II binding protein (Morris and Greenleaf, 2000) provides a molecular clue that may help in understanding numerous reports linking splicing with phosphorylation of RNA Pol II (reviewed by Hirose and Manley, 2000).

Nedd4, and its Pub1 and Rsp5p orthologs in the yeasts, participate in ubiquitination of several phosphoproteins, including the Cdc25 phosphatase and the Fur4p uracil permease, and Fur4p degradation appears to depend on phosphorylation (Marchal et al., 2000). One or more of the three WW domains of Rsp5p is likely to mediate direct interaction with these phosphoproteins in vivo (Lu et al., 1999a; Wang et al., 1999; Chang et al., 2000; Morris and Greenleaf, 2000). In contrast, Nedd4 triggers ubiquitination of the amiloride-sensitive sodium channel, ENaC, in a phosphorylation-independent manner, and the binding of individual WW domains of Rsp5p to unphosphorylated CTD of RNA Pol II is well documented (Chang et al., 2000). Perhaps these data could be reconciled by the finding that the second WW domain of mouse Nedd4 can bind Pro-Pro-X-Tyr and p-Ser-Pro/p-Thr-Pro containing peptides with similar affinities (Lu et al., 1999a). Based on recent reports, the

Table 1. Classification of WW Domains Based on Their Ligand Specificity[a]

Representative Proteins of the Group	Consensus Sequence of the Recognized Peptide	Abbreviation of the Sequence Motif	Representative Ligands (Cognate or Putative)
Group I YAP65, Nedd4/Rsp5p, Dystrophin	PPxY	"PY" or "PPxY"	PEBP2 transcriptional coactivator, ENaC sodium channels, β-Dystroglycan
Group II Formin Binding Proteins, FE65	PPLP usually within long polyproline sequences	"PPLP"	Formin, Mena
Group III Formin Binding Proteins, Npw38/PQBP-1	[R]-R/K/x-PP or PP-R/x-[R]	"PPR"	Splicing factors: SmB, SmB′, U1C, NpwBP
Group IV Pin1/Ess1p, (Nedd4/Rsp5p)[b] (Prp40p?)	(phospho-S/T)P	"(p-S)P or (p-T)P"	RNA Polymerase II Cdc25C phosphatase

[a] See references in a review (Kay et al., 2000) and in Komuro et al., 1999.
[b] Second WW domain of mouse Nedd4 was shown to bind peptides containing p-S/TP or PPxY motifs.

likely explanation for the dual ability of Rsp5p/Nedd4 to associate with both phosphorylated and unphosphorylated CTD is that Rsp5p WW domains can bind to selected imperfect repeats of CTD when they are phosphorylated as well as to all perfect repeats of CTD that are not phosphorylated (Wang et al., 1999; Chang et al., 2000; Morris and Greenleaf, 2000). However, what controls the selection of the phospho-dependent or phospho-independent mode of interaction is not clear.

Generality of Phosphoprotein Binding Function of WW Domains

One of the many questions raised by this new finding is how many of the more than 200 WW domains known so far have the ability to bind to phosphoproteins through recognition of p-Ser or p-Thr? More precisely, can the phosphoprotein binding function of WW domains be predicted? The recent high-resolution structures of the WW domain of Pin1 and dystrophin in complex with their cognate ligands, doubly phosphorylated peptide representing a heptad repeat of CTD, and β-dystroglycan terminal peptide, respectively (Huang et al., 2000; Verdecia et al., 2000), provide an answer to the above question. Moreover, an insightful analysis of these structures by Lim and colleagues allowed them to formulate a unified "protein recognition code" for WW and SH3 domain complexes (Zarrinpar and Lim, 2000). Based on the recent WW domain structures and the results of probing WW and SH3 domains with a collection of N-substituted peptide ligands, they proposed that the WW domain interface recognizes an X-Pro motif (where X is any amino acid), whereas the SH3 domain recognizes two X-Pro motifs (Zarrinpar and Lim, 2000). The authors argue that it is the proximity of a C-substituted and N-substituted X-Pro dipeptide, with only a single intervening carbon atom, which forms the ridge that packs into the aromatic grooves on WW and SH3 domains. Like SH3 domains, the WW domain binds peptide ligands in a polyproline II helix conformation and can bind to an individual peptide in only one particular orientation, which could be in an N- to C-terminal (+) or a C- to N-terminal (−) orientation. In the Pin1-CTD phosphopeptide complex, the ligand binds in the "+" orientation opposite to the orientation of the β-dystroglycan peptide in complex with the dystrophin WW domain ("−" orientation). The new WW domain structures also reveal two variable loops, which lie on the same surface as the aromatic groove that binds X-Pro motif. Loop I is located between β strands 1 and 2, whereas loop II is between β strands 2 and 3. In the "+" orientation of the Pin1 complex, the arginine of loop I makes a contact with the p-Ser5 side chain of the CTD RNA Pol II ligand peptide, and it is predicted that only those WW domains that contain arginine in the loop I will bind p-Ser-Pro or p-Thr-Pro motifs (see also Discussion in Verdecia et al., 2000). Loop II has a constellation of hydrophobic amino acids that accommodate the Tyr of Pro-Pro-X-Tyr peptide motifs. The prediction that WW domains have the ability to bind p-Ser/Thr-Pro peptides based on the presence of arginine in loop I is substantiated by only one additional WW domain, the fourth WW domain of WWP1 protein that was shown to bind phosphopeptides in vitro. However, none of the Nedd4/Rsp5p or the Prp40p WW domains contain arginine or lysine in loop I. Structures are not available for Nedd4/Rsp5p or Prp40p WW domains, but these are necessary to define exactly where β strand 1 ends and loop I begins in these folds. It is conceivable that Arg or Lys located in the Nedd4 and Rsp5p WW domain just two residues before the presumed loop I (Zarrinpar and Lim, 2000) could participate in coordinating the phosphate in Ser/Thr phosphorylated targets. Based on all the available data and less than a half dozen examples of WW domains recognizing phosphorylated proteins, we think that most WW domains probably interact constitutively with their ligands and do not require ligand phosphorylation for binding.

Another new wrinkle in the WW domain complex is revealed by the structure of the dystrophin WW domain bound to the proline-rich peptide of β-dystroglycan. Namely, the stacking of a proline of the ligand with the tryptophan ring of the domain and the formation of a hydrogen bond between a carbonyl oxygen in the peptide and the indole nitrogen in the tryptophan prove to be identical to those found in SH3 domains bound to their ligands (Huang et al., 2000; Verdecia et al., 2000; Zarrinpar and Lim, 2000). This convergence of proline-rich motif recognition by WW and SH3 domains uncovered by the recent structures provides an important protein–protein recognition rule that could be helpful in characterizing new WW domain interactions.

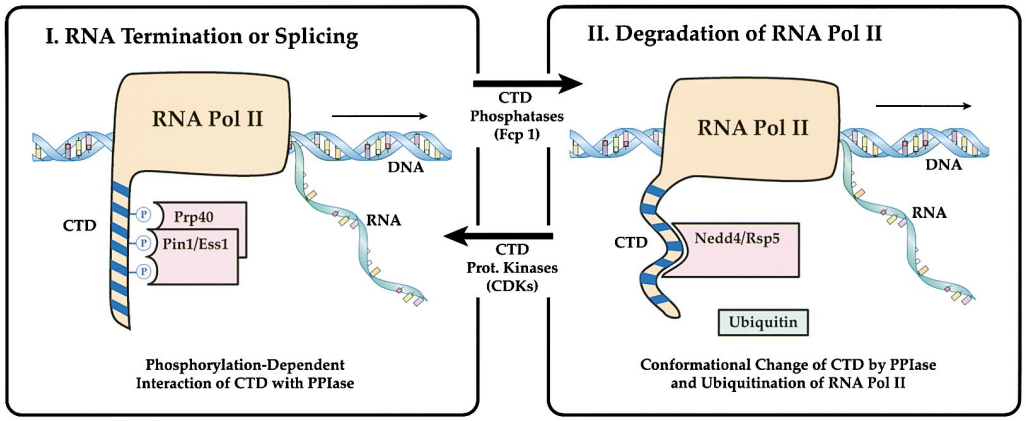

Figure 1. "Binary Switch" Model for Regulation of RNA Polymerase II

"Binary Switch" Model

Two different WW domain–containing proteins bind to the CTD of RNA Pol II as a function of its phosphorylation status. While Pin1/Ess1p WW domains interact with phosphorylated but not nonphosphorylated RNA Pol II CTD (Morris et al., 1999), the second WW domain of mouse Nedd4 or yeast Rsp5p binds preferentially to nonphosphorylated CTD in vivo, but also binds to certain phosphoproteins (Lu et al., 1999a; Wang et al., 1999; Chang et al., 2000; Morris and Greenleaf, 2000). These observations suggest an interesting regulatory scenario, which we present here as a hypothetical model. In eukaryotes, RNA Pol II plays an essential role in transcribing protein-encoding genes. The phosphorylation status of the CTD correlates with different steps in RNA Pol II–mediated transcription, and several factors that associate with the CTD have been shown to function in various aspects of transcription and RNA processing (e.g., Hirose and Manley, 2000). Since WW domains of Nedd4/Rsp5p ubiquitin ligase bind preferentially to perfect repeats of CTD that are not phosphorylated, dephosphorylation, or perhaps partial dephosphorylation of CTD, could be a signal that targets RNA Pol II for ubiquitination and degradation by the 26S proteasome.

Hyperphosphorylation of the CTD is correlated with transcript elongation and splicing. The Pin1/Ess1p PPIase binds the phospho-CTD through its phosphate-sensing WW domain (Morris et al., 1999). We hypothesize that the PPIase activity of bound Pin1/Ess1p might be able to reconfigure the structure of the CTD through isomerization of proline peptide bonds. This conformational change could allow dephosphorylation of the CTD by specific phosphatases, such as Fcp1p, and that could in turn release Pin1/Ess1p and allow other proteins to bind. A possible model for the modulation of RNA Pol II activity and stability by WW domain–mediated interactions is presented in Figure 1. We propose that the phosphorylation of the CTD acts as a "binary switch" for the WW domains of Pin1/Ess1p, Prp40p, and Nedd4/Rsp5p, and perhaps for other WW domain–containing binders of RNA Pol II CTD (Lu et al., 1999a; Morris et al., 1999; Wang et al., 1999; Chang et al., 2000; Morris and Greenleaf, 2000). In this way, the WW domains could contribute to the regulation of RNA Pol II function by bringing to the RNA Pol II holoenzyme additional enzymatic activities, such as is the case for Pin1/Ess1p and Rsp5p/Nedd4. The Prp40p splicing factor could use its WW domains to recognize phospho-CTD, and employ its multiple FF domains to bring additional components, perhaps splicing factors, needed for spliceosome formation.

Phosphotyrosine as a Negative Switch

Examination of the structure of the WW domain-Pro-Pro-X-Tyr ligand complex (Huang et al., 2000), and the negative effect of substituting the tyrosine of the Pro-Pro-X-Tyr core with phosphotyrosine in binding assays supports the idea that tyrosine phosphorylation of the ligand could serve as a negative regulator in vitro and in vivo (Chen et al., 1997). However, there is still no clear cut in vivo example of tyrosine phosphorylation regulating WW ligand complexes. It is also unclear which protein tyrosine kinase might phosphorylate this sequence. If the tyrosine phosphorylation of Pro-Pro-X-Tyr core-containing ligands for Group I WW domains is documented in vivo as a general negative switch, it will add another attractive regulatory feature to this surprisingly versatile module. In addition, it is possible that phosphorylation of the WW domain itself on serine or tyrosine will regulate ligand binding.

WW as a Phosphate-Dependent SH3?

WW domains have the ability to bind proline-rich cores and/or p-Ser-Pro/p-Thr-Pro-containing motifs. Moreover, at least one WW domain, the second WW domain of mouse Nedd4, appears to have dual specificity for proline-rich and phosphoprotein ligands (Lu et al., 1999a; Chang et al., 2000; Morris and Greenleaf, 2000). Based on the available data, one could say that in its functional "design" the WW domain embodies elements of SH2 and SH3 domains, and in the case of Group IV WW domains one could consider them to be phosphorylation-dependent SH3 domains.

Protein Kinases and Phosphatases

Among many questions raised by the recent reports are those about the identity of the protein kinases that are implicated in switching on the ligands for phospho-dependent WW domains. Since the serine and/or threonine phosphorylated peptides that bind Pin1 and the second WW domain of mouse Nedd4 are followed by a single proline residue, the protein kinases that phosphorylate these sites are de facto proline-directed kinases, which include the Cdks, GSK-3, and MAP kinases. It would also be important to identify the protein phospha-

tases that act in concert with protein kinases to regulate the signaling through WW domain–containing proteins.

Signaling Pathways for WW Domains Recognizing Phosphoproteins

What signaling pathways utilize WW domain–containing proteins to sense phosphoprotein ligands? Since HECT family E3 ubiquitin ligases employ WW domains in their regulatory regions as substrate selectors, and since several WD40 repeat–containing proteins represent adaptors that recognize phosphorylated substrates and are part of an E3 complex that recruits an E2 ubiquitin-conjugating activity, it is a reasonable possibility that at least some HECT domain E3s utilize their WW domains to target proteins for proteasome-dependent degradation in a phosphorylation-dependent manner. Further study of how the WW domain–containing Pin1/Ess1p PPIase regulates transcription and other nuclear or cytoplasmic events should reveal details of a concerted action between two molecular switches, phosphorylation/dephosphorylation and conformation. The regulation of the tau and RNA Pol II phosphoproteins by interaction with Pin1/Ess1p provides excellent models for such studies (Lu et al., 1999b; Morris et al., 1999; Verdecia et al., 2000; Wu et al., 2000).

WW Domain Crosstalk with Other Modules

Because of the similar and overlapping sequences of their binding sites in target proteins, there is potential redundancy and crosstalk between WW domains and SH3, WD40, FHA, FF, and PTB domains. There are already several hints that competition at overlapping target sites exists in vitro for SH3 and WW domains, but it would be important to decipher this in an in vivo or at least in a cell culture setting. With the aid of efficient peptidomimetics recently delineated as ligands for SH3 and WW domains (Zarrinpar and Lim, 2000), such studies should progress fast.

It will also be important to understand the cooperation of WW domains with other modular domains located within the same signaling protein. One example is dystrophin, where the WW domain of dystrophin and the two adjacent EF hand-like domains form a composite surface that is required for dystrophin to bind the C-terminal Pro-Pro-X-Tyr motif in β-dystroglycan (Huang et al., 2000). The EF hand region of dystrophin does not bind calcium, but stabilizes the fold of the WW domain and provides additional specificity in recognition of β-dystroglycan (Huang et al., 2000). The WW domain may constitute part of a composite ligand binding domain in other signaling proteins. The recently described FF domain is a good candidate for cooperation with the WW module. The FF domain is frequently present in proteins that contain WW domains and that are involved in RNA splicing or in transcription elongation (Bedford and Leder, 1999). Functional and structural cooperation between WW and FF domains in proteins that build transcriptional and splicing complexes deserves special attention. Two proteins containing both FF and WW domains, CA150, a nuclear factor implicated in transcription elongation, and Prp40p have been shown to interact with CTD of RNA Pol II (Carty et al., 2000; Morris and Greenleaf, 2000).

WW Domain "Returns" to the Cancer Field

The WW domain was originally identified as a repeated module in an oncogene-associated protein. While its function in that complex is still being intensively studied, the WW domain containing *WWOX* gene has recently been shown to encode a potential tumor suppressor (Bednarek et al., 2000), bringing the WW domain back full circle to the cancer arena. The *WWOX* gene encodes an oxidoreductase enzyme, with two WW domains. The gene maps to the human chromosome 16q23, a region affected in breast and ovarian cancers. The investigation of the molecular function of WWOX and the study of YAP expression in transformed cell lines should help in the identification of signaling pathways relevant for certain types of cancer.

Concluding Remarks

The WW domain was first identified only six years ago, but this is relatively ancient by the time standards in the exponentially growing field of signaling domains. Each new wrinkle in its structure and function adds new understanding of its possible roles in cell signaling. Moreover, by utilizing powerful molecular methods, including phage-displayed polypeptide libraries, we are now in a position to introduce additional wrinkles into the domain, potentially giving it the ability to modify or correct a variety of signaling pathways and eventually diseases. Although the WW domain is one of the smallest protein modules, it has a surprisingly large number of functions.

Selected Reading

Bedford, M.T., and Leder, P. (1999). Trends Biochem. Sci. *24*, 264–265.

Bednarek, A.K., Laflin, K.J., Daniel, R.L., Liao, Q., Hawkins, K.A., and Aldaz, C.M. (2000). Cancer Res. *60*, 2140–2145.

Carty, S.M., Goldstrohm, A.C., Sune, C., Garcia-Blanco, M.A., and Greenleaf, A.L. (2000). Proc. Natl. Acad. Sci. USA *97*, 9015–9020.

Chang, A., Cheng, S., Espanel, X., and Sudol, M. (2000). J. Biol. Chem. *275*, 20562–20571.

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.

Hirose, Y., and Manley, J.L. (2000). Genes Dev. *14*, 1415–1429.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M.J. (2000). Nat. Struct. Biol. *7*, 634–638.

Kay, B.K., Williamson, M.P., and Sudol, M. (2000). FASEB J. *14*, 231–241.

Komuro, A., Saeki, M., and Kato, S. (1999). J. Biol. Chem. *274*, 36513–36519.

Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. (1999a). Science *283*, 1325–1328.

Lu, P.J., Wulf, G., Zhou, X.Z., Davies, P., and Lu, K.P. (1999b). Nature *399*, 784–788.

Macias, M.J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996). Nature *382*, 646–649.

Marchal, C., Haguenauer-Tsapis, R., and Urban-Grimal, D. (2000). J. Biol. Chem. *275*, 23608–23614.

Morris, D.P., Phatnani, P.H., and Greenleaf, A.L. (1999). J. Biol. Chem. *274*, 31583–31587.

Morris, D.P., and Greenleaf, A.L. (2000). J. Biol. Chem., in press.

Shen, M., Stukenberg, P.T., Kirschner, M.W., and Lu, K.P. (1998). Genes Dev. *12*, 706–720.

Verdecia, M.A., Huang, H.K., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. (2000). Nat. Struct. Biol. *7*, 639–643.

Wang, G., Yang, J., and Huibregtse, J.M. (1999). Mol. Cell. Biol. *19*, 342–352.

Wu, X., Wilcox, C.B., Devasahayam, G., Hacket, R.L., Arevalo-Rodriquez, M., Cardenas, M.E., Heitman, J., and Hanes, S. (2000). EMBO J. *19*, 3727–3738.

Zarrinpar, A., and Lim, A.W. (2000). Nat. Struct. Biol. *7*, 611–613.