Rsp5 WW Domains Interact Directly with the Carboxyl-terminal Domain of RNA Polymerase II*

Alex Chang, Sonny Cheang, Xavier Espanel, and Marius Sudol‡

From the Department of Biochemistry and Molecular Biology, New York University/Mount Sinai School of Medicine, New York, New York 10029

RSP5 is an essential gene in Saccharomyces cerevisiae and was recently shown to form a physical and functional complex with RNA polymerase II (RNA pol II). The amino-terminal half of Rsp5 consists of four domains: a C2 domain, which binds membrane phospholipids; and three WW domains, which are protein interaction modules that bind proline-rich ligands. The carboxyl-terminal half of Rsp5 contains a HECT (homologous to E6-AP carboxyl terminus) domain that catalytically ligates ubiquitin to proteins and functionally classifies Rsp5 as an E3 ubiquitin-protein ligase. The C2 and WW domains are presumed to act as membrane localization and substrate recognition modules, respectively. We report that the second (and possibly third) Rsp5 WW domain mediates binding to the carboxyl-terminal domain (CTD) of the RNA pol II large subunit. The CTD comprises a heptamer (YSPTSPS) repeated 26 times and a PXY core that is critical for interaction with a specific group of WW domains. Analysis of synthetic peptides revealed a minimal CTD sequence that is sufficient to bind to the second Rsp5 WW domain (Rsp5 WW2) in vitro and in yeast two-hybrid assays. Furthermore, we found that specific “imperfect” CTD repeats can form a complex with Rsp5 WW2. In addition, we have shown that phosphorylation of this minimal CTD sequence on serine, threonine and tyrosine residues acts as a negative regulator of the Rsp5 WW2-CTD interaction. In view of the recent data pertaining to phosphorylation-driven interactions between the RNA pol II CTD and the WW domain of Ess1/Pin1, we suggest that CTD dephosphorylation may be a prerequisite for targeted RNA pol II degradation.

RSP5 (reversion of Spt phenotype) is an essential gene in the yeast Saccharomyces cerevisiae. The gene was originally isolated in a screen for suppressors of SPT3 mutations (1). SPT3 is part of the multicomponent protein named SAGA (Spt-Ada-Gcn5 histone acetyltransferase), which has been implicated in at least two facets of RNA pol II activity. These include the function of the TATA box-binding protein Spt15 and the activity of Gcn5 histone acetyltransferase (1, 2). Moreover, Rsp5 was recently shown to form a physical and functional complex with RNA pol II (3, 4).

In addition to its interaction with RNA pol II, Rsp5 possesses multiple and diverse substrates, suggesting that Rsp5 affects a wide range of cellular processes. For example, RSP5 is known to down-regulate the activity of several plasma membrane-associated proteins, including Fur4 uracil permease, Gap1 amino acid permease, and the integral membrane protein proton ATPase (5–8). Rsp5 was recently proposed to be a component of a complex regulating an early stage of endocytosis (9). Mutations in RSP5 have also been shown to regulate stability of a minichromosome (10) and the cytoplasmic/mitochondrial distribution of Mod5, an enzyme that modifies tRNA (11). In addition, Rsp5-mediated ubiquitination has been implicated in mitochondrial inheritance (12).

The primary structure of Rsp5 reveals a multidomain topology (Fig. 1). The amino-terminal half of Rsp5 consists of four domains: a C2 domain, which may bind membrane phospholipids and inositol phosphates in a calcium-regulated manner; and three WW domains (WW1, WW2, and WW3), which are protein interaction modules that bind proline-rich ligands. The carboxyl-terminal half of Rsp5 contains a HECT (homologous to E6-AP carboxyl terminus) domain (13). The HECT domain encodes an E3 ubiquitin-protein ligase activity, which ligates ubiquitin in the process of targeted protein degradation (14). Presumably, the C2 and WW domains act as membrane localization and substrate recognition modules for Rsp5, respectively (3, 15, 16).

The WW domain is a small protein interaction module composed of 40 amino acids that fold into a three-stranded, antiparallel β-sheet, which constitutes a hydrophobic pocket for binding proline-rich ligands (17–19). The WW domain is named after two conserved tryptophan residues spaced 20–22 amino acids apart (20). Based on their binding specificity, the WW domains can be divided into two major and three minor groups (21). Group I WW domains, exemplified by the YAP and Dystrophiin WW domains, bind proteins with PPXY motifs (18, 22). Although the first proline of the PPXY motif greatly enhances binding to Group I WW domains, our laboratory (24) and others (23, 25) have shown that in a few cases, this position can be substituted by specific amino acids such as leucine and serine. The description for Group I WW domain binding motifs indifferently uses the phrases PPXY core motifs and PPXY motifs. Group II WW domains, exemplified by WW domains of FE65- and Formin-binding proteins, bind to PPLP motifs (26, 27). Among the remaining three groups of WW domains, Groups III and V bind proline-rich sequences (21, 28, 56). However, Group IV WW

* This work was supported by the Human Frontier Science Program Organization, National Institutes of Health Grants CA45757 and CA01605 and the Muscular Dystrophy Association (to M. S.). The costs
† To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, Mount Sinai School of Medicine, Box 1020, One Gustave Levy Place, New York, NY 10029. Tel.: 212-241-9431; Fax: 212-996-7214; E-mail: Marius.Sudol@msm.edu.
‡ The abbreviations used are: RNA pol II, RNA polymerase II; CTD, carboxyl-terminal domain; GST, glutathione S-transferase; PBS, phosphate-buffered saline; YAP, Yes kinase-associated protein; SPOTs, small peptides synthesized on devitiated membranes.
domains were shown to interact with short sequences containing phosphorylated serine or phosphorylated threonine followed by proline. Interestingly, the binding of Group IV WW domains to their ligands was shown to be phosphorylation-dependent (29, 30).

The RNA pol II large subunit is the major component of the transcriptional machinery responsible for regulated expression of >90% of genes (31). The yeast homolog of RNA pol II contains, in its CTD, a heptamer sequence (YSPTSPS) repeated perfectly 22 times and imperfectly four times (Fig. 1). An examination of the CTD sequence revealed a PXYY core motif within the tandemly repeated heptamer peptide YSPTSPSYSPTSPS (32). The CTD has many potential sites for phosphorylation by CTD kinases, most notably serines 2 and 5, and dephosphorylation by CTD phosphatases (31). In addition to these interactions, a well-known function of the CTD is the ability to recruit many transcriptional components (33).

Our strategy for identifying cognate ligands was to utilize phage display combinatorial peptide libraries and to determine the optimal ligand preference for each yeast WW domain, including the three Rsp5 WW domains. After obtaining the consensus sequences from binding phages, the yeast data base was searched for proteins containing the motifs in an effort to predict protein ligands binding to yeast WW domains based on determinations of their consensus binding motifs. Finally, we decided to further investigate a potential interaction between Rsp5 WW domains and the RNA pol II CTD because the PXY motif was found to be a minimal basic core that could interact with Group I WW domains (23–25).

Our strategy for identifying cognate ligands was to utilize phage display combinatorial peptide libraries and to determine the optimal ligand preference for each yeast WW domain, including the three Rsp5 WW domains. After obtaining the consensus sequences from binding phages, the yeast data base was searched for proteins containing the motifs in an effort to identify potential ligands for subsequent biochemical and genetic validation of WW domain-ligand complexes. Data base searches and pull-downs with the Rsp5 WW domains revealed a potential interaction with RNA pol II, and this interaction was extensively characterized. We report that Rsp5 WW2 (and possibly WW3) mediates binding to the CTD of RNA pol II. We were able to uncover a minimal CTD peptide that is sufficient to bind to Rsp5 WW2 in vitro and in the yeast two-hybrid assay. In addition, we have shown that phosphorylation of the minimal CTD sequence on serine, threonine, and tyrosine residues may act as a negative regulator of the Rsp5 WW2-CTD interaction. In view of the recent data pertaining to phosphorylation-driven interactions between the RNA pol II CTD and the Ess1/Pin1 rotamase (30), we suggest that CTD dephosphorylation may be a prerequisite for targeted RNA pol II degradation.

### Table I

**Results of phage-displayed combinatorial peptides using Rsp5 WW domains**

The three Rsp5 WW domains (WW1, WW2, and WW3) were used to screen two biased phage display libraries, X6PPX6 and X6PPXXP6; for more details, see “Experimental Procedures.” The results from screening Rsp5 WW1, WW2, and WW3 are listed in Parts A–C, respectively. The left column indicates the clone number and library used, PP or PXXP. The middle column reveals the peptide sequence from each clone, translated from each phage DNA insert. The right column shows the relative strength of binding of the phages to the respective WW domain, as estimated by enzyme-linked immunosorbent assay and represented using increasing plus symbols. For each WW domain, a consensus sequence of all peptide sequences was organized and is listed in the last row of each section.

| Clone No. | Sequence (X6PPX6 or X6PPXXP6) | Clone binding |
|-----------|---------------------------------|---------------|
| A. 42-PP  | SAREPYYYYGCGWD | +++ |
| 43-PP     | SARIPNAPPPDDVEV | ++ |
| 44-PP     | RLTSPPPPDPGSWE | ++ |
| 45-PP     | RCQQPNPSPPPSELY | ++ |
| 46-PP     | RLEKPNPSPPPIEAA | + |
| 49-PP     | SPPAPGDPSPPPSVG | ++ |
| 510-PP    | SQSNWPPPPYEEALER | + |
| 55-PP     | VLGTVLDPDPSPPPYEIA | + |
| 56-PP     | NFTRILSPPPSYSTI | + |
| 47-PXXP   | LLGYGPPPPPPYEENVGT | + |
| Consensus | DPPYEYEYPS | P |
| B. 51-PP  | RDRGEPSPPPPDYEY | +++ |
| 58-PP     | RLTSISPSPPPDGWE | + |
| 59-PP     | RHYZDEPPPPPTVKKY | + |
| 53-PXXP   | FFGAPPSSPSVQAVGE | + |
| 54-PXXP   | FSPYEPSPPPSQADWS | + |
| 56-PXXP   | FWWPSEPPPPYTQTPK | + |
| 58-PXXP   | QSGMHPYPPPSSRTQ | + |
| 510-PXXP  | LAHEPLPPPPDMVLP | + |
| 51-PXXP   | VVPSSPHPLPPREL | + |
| 52-PXXP   | KNGSPPSPLPPAKAA | + |
| 56-PXXP   | IYVGRPPPPPSGVSQ | + |
| 57-PXXP   | VADGMGPPPPSLWNL | + |
| 58-PXXP   | FTQLSVPPPPSAEEG | + |
| 59-PXXP   | RFPFLDPPPPPEIPSE | + |
| Consensus | PPPPPS | |
| C. 61-PP  | NKEWGPSPPPPZETLF | +++ |
| 62-PP     | DYAEPPPDPGWV | + |
| 64-PP     | ARNWAPPPSPTEMQ | + |
| 69-PP     | AARRPSPPPDGWD | + |
| 610-PP    | SREDWPPPPQCV | + |
| 61-PXXP   | LOVGLFPAPPPPDGW | + |
| 62-PXXP   | YWGEYEPPLPPVYIM | + |
| 63-PXXP   | PHWVLPPPPYNASSS | + |
| 64-PXXP   | DYPWEPPPPPYELAPS | + |
| 65-PXXP   | FFIDNPSPPPPEELDG | + |
| 68-PXXP   | DYPWEPPPPPYELAPS | + |
| 69-PXXP   | ECGAPPPEPEPFDLRF | + |
| 610-PXXP  | MWHRAPPPPPVEVWTVT | |
**Rsp5 WW Domain-RNA Polymerase II CTD Interaction**

All searches were performed at the NCBI web site using the BLAST programs. For the general data base searches, PXY was searched using \\*BLAST searches with the pattern -SPSYS, word size set to 2, and expected number set to 10 million. The other general data base searches, PPPPPP, PPPPYY, PPPPXY, and PPPXYE, were run using Advanced BLAST, with the same settings for word size and expected number. For each general data base search, the total number of matches including open reading frames, an interesting match, and the function of the interesting match are listed in the columns, respectively. For the specific data base searches, each phase display consensus sequence for Rsp5 WW domains was used to search the yeast data base. The resultant searches were analyzed, and the unknown proteins, those proteins from predicted open reading frames, were ignored. For each section (Part A, (D/P)PP(S/P)YE; Part B, PPPPYS; Part C, (A/P)PPPYE; and Part D, LPXY), we included six interesting matches, the actual sequence, and protein function.

### Table II

**BLAST data base searches for yeast proteins containing the optimal peptide sequences**

| Sequence Matches | Interesting match | Function |
|------------------|------------------|----------|
| PXY 613 | Rpb1 | RNA pol II |
| PPPPP 438 | Ms35 | mRNA splicing component |
| PPPFY 365 | Med7 | Mediator of transcription |
| PPPPXY 353 | Prp8 | mRNA splicing and cell cycle |
| PPXYE 74 | NPB1 | Ser/Thr protein kinase |

### II. Specific data base searches

| Match Sequence | Function |
|----------------|----------|
| NPY1 TPDDPYSDES | Ser/Thr protein kinase |
| UBC6 VENPPYLYLA | Ubiquitin-conjugating enzyme |
| TOF1 QIFDPYKLQD | Topoisomerase I-interacting factor |
| RPR7 GGNPVELGGS | RNA pol II, non-essential subunit |
| Bu11, DAG1, RDS1 HDPPSYQEA | Involved in ubiquitination pathway |
| Bu2 HDPPSYQEVE | Bul1 homolog |
| MED7 YPPPPPYVKF | Component of RNA pol holoenzyme |
| RVS167 QGGPPAYSNP | Protein involved in endocytosis |
| NCE4 DITPPAYSAT | Negative regulator of Cts1 Expression |
| THI20 IINTPPYTLT | Kinase needed for thiamine biosynthesis |
| THI21 IINTPPYTLA | Kinase needed for thiamine biosynthesis |
| THI22 IINTPPYTLT | Kinase needed for thiamine biosynthesis |
| Prp8 PPPPPYXIE | mRNA splicing and S-phase regulator |
| SEC9, HSS7 VYPAGYEQV | Component of t-SNARE complex |
| RPL16B or RP23 EGIIPYDHRK | Ribosomal protein |
| RPL16A or RP22 EGIIPYDHRK | Ribosomal protein |
| Kin1 NGPPPSYLPR | Ser/Thr protein kinase |
| TOP1p VIPPPYYQCL | Yeast topoisomerase I |
| TOM1 QLNLPPYEYS | Ubiquitin ligase |
| TAF61 TTKLPYEMO | Component of TFI complex |
| ARC15 ACLYPYPEET | Component of ARP complex |
| MNN10, BED1 KRYPPYSKR | Required for mannann synthesis |
| STO1 YNGPEPYYEM | Large subunit of CAP-binding protein |
| GLY1 HLMQPSYSL | Amino acid metabolism |

---

**EXPERIMENTAL PROCEDURES**

**Clones and Antibodies**—The Rsp5 WW domains and the CTD of RNA polymerase II were amplified and cloned into pGEX-2TK vectors, allowing us to purify GST fusion proteins. The 5’- and 3’-primers of each construct are as follows (dashes indicate restriction enzyme sites): 5’-Rsp5-WW1, ACTCAGGGATCC-AGCGGTACAACAGCTGCT; 3’-Rsp5-WW1, ACTCAGGGATCC-AGGAATTATCTGATGATCCACC; 5’-Rsp5-WW2, ACTCAGCCCGGG-TTCCTCTGTAACAGTTCAAGTG; 3’-Rsp5-WW2, ACTCAGGGATCC-AGGAATTATCTGATGATCCACC; 5’-Rsp5-WW3, TACGACGAAATC-TGGATTTTTGGCTAAGCATAAG, 3’-Rsp5-WW3, TACGACGAAATC-TGGATTTTTGGCTAAGCATAAG, 3’-Rsp5-WW3, TACGACGAAATC-TGGATTTTTGGCTAAGCATAAG, 3’-Rsp5-WW3, TACGACGAAATC-TGGATTTTTGGCTAAGCATAAG, 3’-Rsp5-WW3, TACGACGAAATC-TGGATTTTTGGCTAAGCATAAG.

**Preparation of GST Fusion Proteins**—Expression of the GST fusion proteins was induced after adding 1 mM isopropyl-1-thio-

---

[CONTENT CONTINUES]
amidite labeling of the GST fusion protein; and the dashed arrow identifies GST alone.

Radioactive Labeling of GST Fusion Proteins—The vectors containing the Rsp5 WW domains (pGEX-2TK) have a thrombin-cutting site and protein kinase A phosphorylation site between GST and the WW domain. The fusion proteins were incubated with glutathione-linked agarose beads and washed with PBS and 1% Triton X-100, PBS, and heart muscle kinase buffer (20 mM Tris (pH 7.5), 12 mM MgCl₂). The beads were incubated for 30 min with 60 μl of heart muscle kinase buffer plus 1 mM dithiothreitol, 150 units of protein kinase A (Sigma), and 30 μCi of [γ-32P]ATP (NEN Life Science Products).

After running the gel for ~1 h at 100 V or until the loading buffer ran to the bottom of the gel, the gel was stained with Coomassie Brilliant Blue Dye for 30 min, and the unbound dye was removed.

Phage Display Analysis—Using each WW domain of Rsp5, we searched for preferred binding motifs using various combinatorial peptide libraries. Details concerning the construction of the libraries were discussed previously (38). Briefly, we used two libraries; one library displays 14-mer X₆PPX₆ peptides, and the other displays 16-mer X₆PXXPX₆ peptides, where X denotes any amino acid residue and P represents the position of the fixed proline residue. Each library consists of 1.5 × 10⁹ unique clones. After three rounds of "panning" each Rsp5 WW domain immobilized in wells of microtiter plates, phages were selected for their ability to bind the WW domains and not the GST fusion protein itself. The inserts of selected phages were sequenced and translated for peptide sequences. The consensus sequences identified for each Rsp5 WW domain were used to search the yeast data base for proteins containing the binding motifs. The general and specific data base searches, listed in Table II, utilized the Pam30 advanced BLAST algorithm, filters off, expected number set to 10 million, and word size set to 2. The PXY search employed the hBLAST program using the SPSYS pattern (39).

Pull-down Experiments with Yeast Lysates—Wild-type yeast cells, W303-1A (MATa, ura3-1, leu2-3,112, trp1-1, can1-100, ade2-1, his3-11,15 [psi +]), were grown in YPD (yeast extract, peptone, dextrose) medium overnight. The cells were washed with sterilized water and resuspended in protein extraction buffer (50 mM Tris-HCl (pH 7.4), 1 mM EDTA, and 0.1% Triton X-100) with protease inhibitors. The resuspended cells were divided into 500-μl aliquots using 1.5-ml microcentrifuge tubes. An equivalent amount of 0.2-mm glass beads was added, and the solution was vortexed vigorously for 5 min and placed on ice for 5 min. After vortexing two additional times, the solutions were centrifuged for 4 °C for 30 min at 13,000 rpm. The clarified supernatants were transferred to 1.5-ml Eppendorf tubes, and the pellets were discarded.

Using 800 μg of prepared yeast extracts, beads containing Rsp5 WW1, WW2, or WW3 were incubated in binding buffer (50 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM EDTA, 0.1% Tween, and 0.1% bovine serum albumin) containing 50 μl of [γ-32P]ATP (NEN Life Science Products) as described previously (40, 41). Each preparation was washed twice with binding buffer without bovine serum albumin, and the binding proteins were released in loading buffer containing β-mercaptoethanol and dithiothreitol.

Next, the samples were assayed for proteins binding to respective WW domains. The samples were run on 10.5% SDS-polyacrylamide gel and blotted onto nitrocellulose membranes. The membranes were blocked in Western Wash solution (10 mM Tris-HCl (pH 7.4), 0.9% NaCl, and 10% milk) containing 0.1% Triton X-100) containing 5% milk and probed for 6 h at 4 °C with their respective radiolabeled WW domains used in the binding experiments. Afterward, the blots were washed at least three times with Western Wash solution containing 5% milk and exposed to film.

"Far Western Blotting"—The GST fusion proteins were resolved on 10.5% SDS-polyacrylamide gel. After transferring the gel to nitrocellulose membranes, the blot was blocked for at least 4 h at 4 °C with Western Wash solution containing 5% milk. The blots were probed with radiolabeled Rsp5 WW domains or GST yeast CTD and washed as described above.

SPOTs Membrane Analysis—Small peptides were synthesized on derivatized cellulose membranes (Genosys Biotechnologies, Inc.) as described previously (40, 41). Each membrane was blocked, incubated

FIG. 2. Pull-downs of proteins from yeast cell lysate using each GST-Rsp5 WW domain. A–D, pull-down experiments with yeast extracts. Total yeast lysate, extract, and protein pull-downs from GST alone, GST-Rsp5-WW1, GST-Rsp5-WW2, or GST-Rsp5-WW3 were separated by SDS-polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes. Membranes indicated in A–D were probed with radioactive GST, GST-Rsp5-WW1, GST-Rsp5-WW2, or GST-Rsp5-WW3, respectively. The heavy black arrow indicates a protein migrating around 200 kDa; the double black arrow indicates a protein doublet of unknown identity migrating at ~60 kDa; the thin arrow identifies the GST fusion protein; and the dashed arrow identifies GST alone.

FIG. 3. The 200-kDa band pulled-down by GST-Rsp5-WW2 is recognized by monoclonal antibodies against RNA polymerase II. Each blot from Fig. 1 (A–D) was stripped of labeled probe and incubated with monoclonal antibodies that recognize the phosphorylated CTD of RNA polymerase II. Bound antibodies were identified using ECL. The arrows indicate the 200-kDa band.

FIG. 4A–D. Western Wash solution containing 5% milk and probed for 6 h at 4 °C with their respective radiolabeled WW domains used in the binding experiments. Afterward, the blots were washed at least three times with Western Wash solution containing 5% milk and exposed to film.

"Far Western Blotting"—The GST fusion proteins were resolved on 10.5% SDS-polyacrylamide gel. After transferring the gel to nitrocellulose membranes, the blot was blocked for at least 4 h at 4 °C with Western Wash solution containing 5% milk. The blots were probed with radiolabeled Rsp5 WW domains or GST yeast CTD and washed as described above.

SPOTs Membrane Analysis—Small peptides were synthesized on derivatized cellulose membranes (Genosys Biotechnologies, Inc.) as described previously (40, 41). Each membrane was blocked, incubated
RESULTS

Binding Specificities of the Rsp5 WW Domains Classify Them as Group I WW Domains—We decided to identify potential ligands for each WW domain of Rsp5 and used the following approach. After screening phage-displayed combinatorial peptide libraries, we determined the optimal peptide sequence for ligands to each WW domain and identified potential interacting proteins containing these sequences in the yeast proteome.

The results of “Phage Display Analysis” for each Rsp5 WW domain are presented in Table I. The consensus sequence for each Rsp5 WW domain is as follows: WW1, (D/P)PP(S/P)YE; WW2, PPPPYS; and WW3, (A/P)PPPYE. The preferred binding motif for each WW domain is a PPXY-like motif. Therefore, the three Rsp5 WW domains belong to Group I. Interestingly, two clones, R1-PXXP and 62-PXXP, one from WW2 and one from WW3, respectively, did not conform to the consensus sequence obtained from the other clones. These two clones possess the LPXY sequence, drawing a comparison with the PXYY-type motif suggested previously (23–25).

A search of the yeast data base using the consensus sequences identified the potential ligands listed in Table II (Part 2). The proteins are involved in diverse functions such as transcription, ubiquitination, and amino acid metabolism. The searches also yielded numerous open reading frames. Since the open reading frames represent hypothetical proteins of unknown functions, they were not listed.

We compared the specific data base searches in Part 2 of Table II with a data base search of the yeast proteome using BLAST for PXX and advanced BLAST for PPPPP, PPPPY, and PPXYE, listed in Part 1. The potential ligands identified in Part 2 were recognized in data base searches listed in Part 1. The phage-displayed peptide library screening produced results that are more restrictive than random data base searches.

The Second Rsp5 WW Domain Can Pull-down RNA pol II from Yeast Extracts—To rapidly screen for potential ligand(s) from those listed in Table II (Part 2), we attempted to pull-down binding proteins from total yeast cell lysates. Using GST fusion proteins of Rsp5 WW domains, we identified a protein of 200 kDa in pull-downs with Rsp5 WW2 and WW3 (Fig. 2, C and D). In addition, two faint protein bands were observed migrating around 60 kDa in pull-downs with Rsp5 WW2 (Fig. 2C). Based on the molecular mass of the pull-down proteins and previously published data describing a genetic interaction between Rsp5 and RNA pol II (4), we hypothesized that the 200-kDa protein band was RNA pol II. We stripped the blots and reprobed them with monoclonal antibodies specific for the large subunit of RNA pol II and concluded that the protein, which migrated at 200 kDa, was most likely RNA pol II (Fig. 2, C and D).

All Three Rsp5 WW domains Bind to the CTD of RNA pol II—To investigate the binding specificity driving the protein interaction between WW domains of Rsp5 and the yeast RNA pol II CTD, we used Far Western blot analysis. Bacterial lysates containing GST fusion proteins of yeast RNA pol II CTD, PPXY motifs found in p53BP2 and WBP-1, and a PPLP motif found in Mena were purified, loaded onto SDS gels, transferred to nitrocellulose membranes, and probed with each radioactively labeled Rsp5 WW domain. p53BP2 and WBP-1 contain the PPXY motif that was verified to bind YAP WW domains (18, 38). Mena contains a PPLP motif and was shown to bind to nitrocellulose membranes. The membranes were probed with radioactively labeled GST-Rsp5-WW1 (A), GST-Rsp5-WW2 (B), or GST-Rsp5-WW3 (C). D, the loading of lysates on gels was normalized by Coomassie Blue staining. E, lysates of each GST-Rsp5 WW domain and human YAP were prepared, resolved by SDS-polyacrylamide gel electrophoresis, and transferred to a nitrocellulose membrane. Yeast GST-CTD was radioactively labeled and used to probe the membrane. F, the loading of lysates used in E was normalized by Coomassie Blue staining.
on SPOTs membranes and reduced the length of a repeat from the amino and/or carboxyl terminus by sequentially replacing the flanking sequences with alanine residues.

Based on our analysis, we have chosen -YSPTSPSYSPT (Fig. 5, spot 16) as the “minimal” CTD peptide that binds Rsp5 WW2. Interestingly, a shorter CTD peptide, AAAAAAAPSYSPTSP (Fig. 5, spot 8), was also able to interact with Rsp5 WW2. Parallel SPOTs membranes were prepared with the same repertoires of peptides and probed with radioactive GST-Rsp5-WW1 and GST-Rsp5-WW3 fusion proteins, but no appreciable binding was observed.

Alanine Scanning of the 12-mer CTD Peptide—To identify the amino acid residues involved in the interaction between the minimal CTD peptide and Rsp5 WW2, we scanned the minimal CTD binding motif with alanine residues. Since the peptides are synthesized from the carboxyl terminus, one additional serine residue was added to the 11-mer peptide to provide an anchor and linker for each peptide to the membrane. Replacing tyrosine 1, proline 6, tyrosine 8, and serine 9 with alanine greatly diminished or abolished binding (Fig. 6, spots 2, 7, 9, and 10). Substitutions that had little effect were the following residues: serine 2, proline 3, proline 10, and threonine 11 (Fig. 6, spots 3, 4, 11, and 12). A significant increase in binding was observed when threonine 4 or serine 5 was replaced with alanine (Fig. 6, spot 8). The data suggested that the crucial amino acid residues required for an interaction between the minimal CTD binding motif and Rsp5 WW2 are YXXXXPXYXSPXX. These data confirmed the minimal PXY core shown in phage display screens and revealed the importance of flanking amino acid residues.

Effect of CTD Peptide Phosphorylation on Binding of the Second Rsp5 WW Domain—Next, we investigated whether CTD modification by phosphorylation alters binding to Rsp5 WW2. We synthesized 14-mer peptides of the CTD repeat on SPOTs membranes using phosphorylated serine, threonine, or tyrosine. The Rsp5 WW domains and GST alone were radiolabeled for probing the peptides. In this experiment, we confirmed that only Rsp5 WW2 bound the minimal CTD repeat. Phosphorylation at specific sites of the CTD repeat influenced binding of Rsp5 WW2 in varying degrees. Phosphorylation of an amino-terminal residue flanking the minimal CTD binding motif, PpSYSPYSPTSPYSPX, where pS is phosphorylated serine (Fig. 7, column I, spot 2), showed minimal positive effects on binding. Phosphorylation of serine 2, threonine 4, and serine 7 diminished the interaction with the peptides (Fig. 7, column I, spots 2, 3, and 6; and Fig. 8, spots 3, 4, and 6), but phosphorylation of tyrosine 1, tyrosine 8, serine 9, threonine 11, and serine 12 abolished binding (Fig. 7, column I, spots 7–9; and Fig. 8, spots 7–10).

We repeated the experiments with CTD peptides phosho-

![Fig. 5. Determination of the minimal CTD peptide required for binding to Rsp5 WW2. A, a SPOTs membrane was probed with radioactively labeled GST-Rsp5-WW2. B and C, peptides spanning 15 amino acid residues were synthesized on SPOTs membranes as illustrated. Alanine residues were used to substitute for each CTD amino acid residue to identify the shortest CTD peptide that binds Rsp5 WW2. The substitutions advanced from the amino terminus (spots 2–14), the carboxyl terminus (spots 21–32), and gradually from the amino and carboxyl termini (spots 15–20). Spots 33–39 are imperfect CTD repeats. D, the probed SPOTs membrane was analyzed using a PhosphorImager to compare the relative intensity of signals.](image)
rylated at multiple sites including one phosphorylation in residues flanking the repeat and the other phosphorylation site in the complete repeat. We also varied patterns of multiple CTD phosphorylation. Although not every permutation of the phosphorylated CTD was examined, these experiments have shown that phosphorylation can negatively regulate binding between the CTD and Rsp5 WW2 (Fig. 7, columns II and III). However, although phosphorylation abolishes binding overall, Rsp5 WW2 may have a minor capacity for binding the CTD phosphorylated on serine 7 (Fig. 7, row Rsp5 WW2, compare column I, spots 2 and 6; column II, spot 2; and column III, spots 6, 7, and 9).

Influence of Phosphorylated “Imperfect” CTD Repeats on Binding of Rsp5 WW2—Since the CTD contains imperfect repeats as well, we decided to phosphorylate serine, threonine, and tyrosine residues on imperfect repeats to test the ability of Rsp5 WW2 to bind these peptides. As described under “Alanine Scanning of the 12-mer CTD Peptide,” we observed a much greater binding intensity, using Rsp5 WW2, for a CTD peptide with an alanine substitution at serine 7 than for a minimal CTD peptide (Fig. 6, spot 8). We hypothesize that some imperfect repeats may have greater potential for binding Rsp5 WW2 (see “Discussion”).

We synthesized naturally occurring imperfect CTD peptides, in unphosphorylated and phosphorylated forms, on SPOTs membranes and probed them with Rsp5 WW2. We observed that phosphorylation of imperfect repeats had a negative effect on binding to Rsp5 WW2 compared with the unphosphorylated imperfect repeats (Fig. 8).

For the unphosphorylated imperfect repeats, we observed that not every amino acid can increase binding intensity at position 7 of the minimal CTD peptide, which also corresponds to position X of the PXY core binding motif. For example, in two naturally occurring imperfect repeats in which position 7 harbors an asparagine or glycine residue, there was no observable binding of Rsp5 WW2 for the phosphorylated or unphosphorylated repeat (Fig. 8, spots 26–34 and 41–48). Interestingly, a substitution of serine 7 with alanine increased binding 2–3-fold compared with the perfect repeat (compare spots 1 with spots 8 and 17 in Figs. 6 and 8, respectively). Another 2-fold increase was observed when threonines 4 and 11 were replaced with glycine and lysine, respectively (Fig. 8, spot 35). Finally, the interaction of Rsp5 WW2 with a phosphorylated imperfect CTD repeat at serine 2 or threonine 4 was just as good or greater compared with the binding of Rsp5 WW2 to an unphosphorylated perfect repeat (Fig. 8, compare spot 1 with spots 19, 20, and 37). In summary, selected unphosphorylated imperfect CTD repeats show significantly stronger interactions with Rsp5 WW2 compared with the unphosphorylated perfect repeats.

The Minimal RNA pol II CTD Peptide Interacts with Rsp5 WW2 in a Yeast Two-hybrid Analysis—Rsp5 WW2 and the minimal RNA pol II CTD peptide were cloned into plasmids as fusions with the Gal4 activation domain (AD) and the DNA-binding domain (BD), respectively, to be assayed for protein interaction in vivo using the yeast two-hybrid assay. Clones containing Gal4-AD-Rsp5-WW2 (pGAD-Rsp5-WW2) or Gal4-BD-RNA pol II (pGBT-OneCTD) were plated on agar plates prepared without leucine, tryptophan, and histidine and were unable to grow (Fig. 9). However, a clone containing both plasmids was able to grow (Fig. 9). Thus, we showed that the in

---

**Fig. 6.** Alanine scanning of a 12-mer minimal CTD peptide. A, radioactive GST-Rsp5-WW2 was used to probe the SPOTs membrane. B and C, peptides 12 amino acids long of the CTD repeat were synthesized on SPOTs membranes as illustrated. An alanine residue replaced each position of a perfect repeat. D, the probed SPOTs membrane was analyzed using a PhosphorImager to compare the relative intensity of signals.

**Fig. 7.** Analysis of phosphorylated CTD peptides on SPOTs membranes. As listed to the right, GST alone and each GST-Rep5 WW domain were radioactively labeled and incubated with three strips of spots (columns I-III) containing unphosphorylated and phosphorylated CTD peptides. Grid indicates the order of synthesized peptides, followed by the peptide sequence of each spot. J, phosphoserine; O, phosphothreonine; and Z, phosphotyrosine (in this case).
vitro binding we observed for Rsp5 WW2 and the minimal CTD peptide can also occur in yeast cells.

**DISCUSSION**

We have characterized the binding between Rsp5 WW domains and the CTD of RNA pol II. The second WW domain of Rsp5 was able to pull-down the RNA pol II from yeast cell lysates and bind to CTD-derived peptides. Furthermore, we have determined the minimal length of the CTD peptides that can mediate the Rsp5 WW2-CTD interaction in vitro and in yeast cells. Through the course of our examination of the minimal CTD peptide, the selected imperfect CTD repeats were uncovered as potentially preferential sites for binding Rsp5 WW domains. In addition, we showed that the phosphorylation of the CTD residues has an overall negative effect on binding to Rsp5 WW2.

RNA pol II and Bul1 proteins have been shown to interact with Rsp5 functionally, biochemically, and genetically (3, 4, 32, 43, 44). Using the SPOTs membrane assay, we showed that peptides containing the PPXY motif, derived from Bul1, interacted 30–40 times better than similar length CTD peptides containing a PXK motif. However, when Rsp5 WW2 was used in pull-down assays from yeast cell lysates, RNA pol II was the major protein band that was visualized. We did not observe a
protein migrating between 110 and 140 kDa, the expected molecular mass of Bul1. Assuming that levels of RNA pol II and Bul1 are similar in cells, there could be many explanations for these differences. One favorable explanation is that the numerous (total of 26) CTD PXX-containing repeats may create a scaffolding that assembles Rsp5 WW domains into a large and stable complex with the RNA pol II CTD. The relatively high “local concentration” of the PXX motif combined with the unique structure of the CTD tail could account for such strong complex and/or efficient pull-down results.

The deletion analysis of CTD using alanine substitutions resulted in an interesting observation. After identifying a minimal CTD peptide that binds Rsp5 WW2 (AYSPTSPSYSPTAA), the continued sequential substitution of the CTD with alanine revealed a single peptide (AAAAAAAPSYSPTSPA) that bound to Rsp5 WW2 as well as the minimal CTD peptide (AYSPTSPSYSPTAA). All the intermediate and subsequent peptides in the sequential substitution were negative for binding. This result is difficult to explain. One possibility is that the stretch of alanine residues preceding the PXX core in this peptide (AAAAAAPSYSPTSPA) folds into a conformation similar to the minimal CTD peptide and presents the PXX core for binding. Another possible explanation is that CTD repeats can achieve unique conformations only when a certain length of amino acids and specific amino acid residues of the repeat are present (45). More important, we have chosen an 11-mer CTD sequence as the minimal length peptide for binding Rsp5 WW2 in our study. By NMR analysis, we are investigating the structure of Rsp5 WW2 in complexes using the CTD peptides, as well as other PPXY-containing peptides, to understand these results at the molecular level.

Our complex of Rsp5 WW2 with the 11-mer CTD peptide was confirmed by two methods. The yeast two-hybrid assay confirmed the ability of Rsp5 WW2 to form a complex with the 11-mer peptide of CTD in cells. In addition, we generated an NMR structure of Rsp5 WW2 in complex with the unphosphorylated 11-mer CTD peptide. This structure recapitulates the details previously solved in the complex of the YAP WW domain with the WBP-1 target peptide containing PPXY (19).

A complicated picture has emerged where the status of RNA pol II CTD phosphorylation correlates with transcriptional activity (46). In general, either hypophosphorylated RNA pol II molecules are in the nuclear milieu awaiting targeting by general transcription factors, or RNA pol II rests in quiescence at promoter sites (47). The RNA pol II with a hyperphosphorylated CTD clears the promoter and allows transcription to proceed (31, 48). CTD phosphorylation and dephosphorylation can dynamically change its conformation and drastically alter the proteins that interact with the CTD and therefore the RNA pol II holoenzyme (33). During the preinitiation stage of transcription, the dephosphorylated CTD interacts with proteins involved in transcriptional activation; upon CTD phosphorylation, the proteins that interact with the CTD are involved in co-transcriptional pre-mRNA processing (49, 50).

One of the implications of our data is that the unphosphorylated or partially dephosphorylated CTD may signal an interaction with Rsp5 ubiquitin ligases to target RNA pol II for degradation in the proteasome. Based on our results, we consider two possibilities. CTD phosphatases may mediate a complete dephosphorylation of the CTD, allowing Rsp5 WW domains to form a complex with the CTD at many potential sites; or CTD phosphatases may dephosphorylate specific CTD repeats, possibly selected imperfect repeats, and restrict the ability of Rsp5 WW domains to form a complex with the CTD. The latter possibility seems more likely and is supported by several reports showing that a form of RNA pol II with the phosphorylated CTD is preferentially targeted by ubiquitin ligases (51–53).

In contrast, several reports have shown that Ess1, which contains an isopropyl isomerase and a Group IV WW domain, is able to interact with the phosphorylated RNA pol II CTD and to regulate activity by controlling several processes, which include the formation of the 3′-end mRNA and termination of RNA pol II-mediated transcription (30, 54, 55). Since Ess1 contacts the phosphorylated CTD through its WW domain, an interesting regulatory scenario emerges. The CTD of RNA pol II has an ability to form complexes with two WW domain-containing proteins, Rsp5 and Ess1, which harbor dissimilar enzymatic activities and uniquely affect RNA pol II in terms of transcriptional control mechanisms. A functional test of these regulatory complexes in yeast cells expressing Rsp5 and/or Ess1 with specific point mutations that affect their domain(s) could illuminate novel signaling steps controlling the stability of RNA pol II or maintenance of transcription.

Acknowledgments—We thank David Bregman for monoclonal antibodies to the carboxyl-terminal domain of RNA pol II, Edward Purdue for the W303 cells and assistance in isolation of yeast cell extracts, and Jeanne Hirsch for assistance in isolation of yeast genomic DNA. We also appreciate valuable comments on the manuscript from D. Bregman, A. Greenleaf, S. Hanes, J. Kasanov, B. Kay, A. Korosi, M. Macias, R. Marians, and E. Purdue.

REFERENCES

1. Eisenmann, D. M., Arndt, K. M., Ricupero, S. L., Rooney, J. W., and Winston, F. (1992) Genes Dev. 6, 1319–1331.
2. Roberts, S. M., and Winston, F. (1997) Genetics 147, 451–465.
3. Huijbregts, J. M., Yang, J. C., and Beaudenon, S. L. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3656–3661.
4. Wang, G., Yang, J., and Huijbregts, J. M. (1999) Mol. Cell. Biol. 19, 342–352.
5. Hein, C., Springael, J. Y., Volland, C., Hagenauer-Tsapis, R., and Andre, B. (1995) Mol. Microbiol. 18, 57–77.
6. Galan, J. M., Moreau, V., Andre, B., Volland, C., and Hagenauer-Tsapis, R. (1990) J. Biol. Chem. 271, 10946–10952.
7. de la Fuente, N., Maldonado, A. M., and Portillo, F. (1997) FEBS Lett. 411, 308–312.
8. Springael, J. Y., and Andre, B. (1998) Mol. Cell. Biol. 12, 1253–1263.
9. Wendland, B., Emr, S. D., and Riezman, H. (1998)Curr. Opin. Cell Biol. 10, 513–522.
10. Ussono, Y., Fujita, A., Tohe, A., and Kikuchi, Y. (1994) Gene (Amst.) 143, 135–138.
11. Zoladek, T., Tubiaza, A., Vaduva, G., Boguta, M., Martin, N. C., and Hopper, A. K. (1997) Genetics 145, 595–605.
12. Fisk, H. A., and Yaffe, M. I. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1199–1208.
13. Huijbregts, J. M., Scheffner, M., Beaudenon, S., and Howley, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2563–2567.
14. Ciechanover, A. (1998) EMBO J. 17, 7151–7160.
15. Plant, P. J., Yeger, S., Staub, O., Howard, P., and Rotin, D. (1997) J. Biol. Chem. 272, 32239–32336.
16. Spilka, J. Y., De Craene, J. O., and Andre, B. (1999) Biochem. Biophys. Res. Commun. 257, 561–566.
17. Bork, P., and Sudol, M. (1994) Trends Biochem. Sci. 19, 531–533.
18. Chen, H. I., and Sudol, M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7819–7823.
19. Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) Nature 382, 646–649.
20. Sudol, M. (1996) Prog. Biophys. Mol. Biol. 65, 113–132.
21. Kay, B. K., and Williamson, M. P. (2000) PASEB J. 14, 231–241.
22. Rentschler, S., Linn, H., Deininger, K., Bedford, M. T., Espanel, X., and Sudol, M. (1999) J. Biol. Chem. Hoppe-Seyler 380, 431–442.
23. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N., and Lim, W. A. (1998) Science 282, 2088–2092.
24. Chen, H., Einbond, A., Kwak, S. J., Linn, H., Koepf, E., Petersson, S., Kelly, J. W., and Sudol, M. (1997) J. Biol. Chem. 272, 17070–17077.
25. Pirvouz, G., McConnell, S. M., Uveges, E., Schultz, J., Sudol, M., Sparks, A. B., Kay, B. K., and Fowlkes, D. M. (1997) J. Biol. Chem. 272, 14611–14616.
26. Ermevicka, K. S., Zambrano, N., Linn, H., Minopoli, G., Gertler, F., Russe, T., and Sudol, M. (1997) J. Biol. Chem. 272, 32869–32877.
27. Bedford, M. T., Chan, D. C., and Leder, P. (1997) EMBO J. 16, 2376–2383.
28. Komuro, A., Saecki, M., and Kate, S. (1999) J. Biol. Chem. 274, 36513–36519.

2 C. Civera, A. Chang, M. Sudol, and M. J. Macias, unpublished data.
29. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) Science 283, 1325–1328
30. Morris, D. P., Phatnani, H. P., and Greenleaf, A. L. (1999) J. Biol. Chem. 274, 31583–31587
31. Dahmus, M. E. (1995) Biochim. Biophys. Acta 1261, 171–182
32. Beaudenon, S. L., Huacani, M. R., Wang, G., McDonnell, D. P., and Hubbrecht, J. M. (1999) Mol. Cell. Biol. 19, 6872–6979
33. Myers, L. C., Gustafsson, C. M., Bushnell, D. A., Lui, M., Erdjument-Bromage, H., Temple, P., and Kornberg, R. D. (1998) Genes Dev. 12, 45–54
34. Sudol, M. (1999) Oncogene 17, 1469–1474
35. Espanel, X., and Sudol, M. (1999) J. Biol. Chem. 274, 17284–17289
36. Kim, E., Du, L., Bregman, D. B., and Warren, S. L. (1997) J. Cell Biol. 136, 19–28
37. Patturajan, M., Schultz, R. J., Selton, B. M., Berenney, R., Vincent, M., Bensaude, O., Warren, S. L., and Corden, J. L. (1998) J. Biol. Chem. 273, 4689–4694
38. Linn, H., Ermekova, K. S., Rentschler, S., Sparks, A. B., Kay, B. K., and Sudol, M. (1997) Biol. Chem. Hoppe-Seyler 378, 531–537
39. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
40. Frank, R., and Doring, R. (1988) Tetrahedron 44, 6031–6040
41. Blankenmeyer-Menge, B., Nimitz, B., and Frank, R. (1990) Tetrahedron Lett. 31, 1701–1704
42. Fields, S., and Sung, O. (1989) Nature 340, 245–246
43. Yashiroda, H., Kida, D., Tob, A., and Kikuchi, Y. (1998) Gene (Amst.) 225, 39–46
44. Yashiroda, H., Oguchi, T., Yasuda, Y., Tob, A., and Kikuchi, Y. (1996) Mol. Cell. Biol. 16, 3255–3263
45. Bienkiewicz, E. A., Moon Woody, A., and Woody, R. W. (2000) J. Mol. Biol. 297, 119–133
46. Cho, H., Kim, T. K., Manebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) Genes Dev. 13, 1540–1552
47. Berk, A. J. (1999) Curr. Opin. Cell Biol. 11, 330–335
48. Reines, D., Conaway, R. C., and Conaway, J. W. (1999) Curr. Opin. Cell Biol. 11, 342–346
49. Bensaude, O., Bonnet, F., Casse, C., Dubois, M. F., Nguyen, V. T., and Palancade, B. (1999) Biochem. Cell Biol. 77, 249–255
50. Bentley, D. (1999) Curr. Opin. Cell Biol. 11, 347–351
51. Bregman, D. B., Halaban, R., van Geld, A. J., Henning, K. A., Friedberg, E. C., and Warren, S. L. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 15481–15486