Multivalent Binding of Formin-binding Protein 21 (FBP21)-Tandem-WW Domains Fosters Protein Recognition in the Pre-spliceosome*\textsuperscript{[5]}

Received for publication, May 31, 2011, and in revised form, August 30, 2011 Published, JBC Papers in Press, September 14, 2011, DOI 10.1074/jbc.M111.265710

Stefan Klippel\textsuperscript{1,2,3}, Marek Wieczorek\textsuperscript{1,3}, Michael Schümann\textsuperscript{5}, Eberhard Krause\textsuperscript{5}, Berenice Marg\textsuperscript{6}, Thorsten Seidel\textsuperscript{1}, Tim Meyer\textsuperscript{7}, Ernst-Walter Knapp\textsuperscript{1,3}, and Christian Freund\textsuperscript{1,2}

From the \textsuperscript{4}Protein Engineering Group, Leibniz Institut für Molekulare Pharmakologie and Freie Universität Berlin, Robert-Rösse-Strasse 10, 13125 Berlin, Germany, the \textsuperscript{5}Mass Spectrometry Unit, Leibniz Institut für Molekulare Pharmakologie, Berlin, Robert-Rösse-Strasse 10, 13125 Berlin, Germany, the \textsuperscript{6}Department of Dynamic Cell Imaging, Bielefeld University, Universitätsstrasse 25, 33615 Bielefeld, Germany, and the \textsuperscript{7}Theoretical Chemistry Group, Freie Universität Berlin, Institut für Chemie, Fabeckstrasse 36a, 14195 Berlin, Germany

Background: The role of long proline-rich segments, as they are abundantly present in the spliceosome, is elusive.

Results: Cell biological and biophysical data show the significance of multiple motifs for tandem-WW domain recognition.

Conclusion: The dynamic assembly of the pre-spliceosome is enabled by transient multivalent interactions.

Significance: Our results have general implications for the recognition of proline-rich sequence hubs in modular protein assemblies.

The high abundance of repetitive but nonidentical proline-rich sequences in spliceosomal proteins raises the question of how these known interaction motifs recruit their interacting protein domains. Whereas complex formation of these adaptors with individual motifs has been studied in great detail, little is known about the binding mode of domains arranged in tandem repeats and long proline-rich sequences including multiple motifs. Here we studied the interaction of the two adjacent WW domains of spliceosomal protein FBP21 with several ligands of different lengths and composition to elucidate the hallmarks of multivalent binding for this class of recognition domains. First, we show that many of the proteins that define the cellular protome interacting with FBP21-WW1-WW2 contain multiple proline-rich motifs. Among these is the newly identified binding partner SF3B4. Fluorescence resonance energy transfer (FRET) analysis reveals the tandem-WW domains of FBP21 to interact with splicing factor 3B4 (SF3B4) in nuclear speckles where splicing takes place. Isothermal titration calorimetry and NMR shows that the tandem arrangement of WW domains and the multivalency of the proline-rich ligands both contribute to affinity enhancement. However, ligand exchange remains fast compared with the NMR time scale. Surprisingly, a N-terminal spin label attached to a bivalent ligand induces NMR line broadening of signals corresponding to both WW domains of the FBP21-WW1-WW2 protein. This suggests that distinct orientations of the ligand contribute to a delocalized and semispecific binding mode that should facilitate search processes within the spliceosome.

Protein interactions mediated by proline-rich sequences (PRS)\textsuperscript{3} are characterized by low affinity and promiscuous binding of adaptor domains (1–5). These types of interactions seem to be a hallmark of molecular assemblies that undergo rapid rearrangements. One such example of a highly dynamic molecular machine is the spliceosome where several subunits, also denoted as small nuclear ribonuclear particles (snRNP), are steered into a final complex that is responsible for the excision of introns. Proline-rich sequence hubs have been identified in several snRNP proteins and shown to play a crucial role in the assembly of the early spliceosome (6, 7). WW and GYF domain containing proteins are attracted by these sequences and thereby allow for molecular crowding at the long and repetitive PRS of essential splicing proteins.

Formin-binding protein 21 (FBP21) was initially found as a binding partner of the cytoskeletal protein formin (8). Later investigations, however, clearly put the protein into the context of splicing and transcription (9). It binds to the core splicing protein SmB/B’ and was recently shown to be important for splicing in a cellular context (10). FBP21 contains two WW domains that both have been characterized to interact with PRS ligands of group R\textsubscript{K} (11). In these ligands proline residues are flanked by arginine and are indeed abundantly present in several splicing proteins. The linker connecting the two WW domains is 10–12 amino acids in length and shows enhanced flexibility compared with the individual WW domains. Upon ligand encounter, this linker is thought to adopt a more defined conformation to allow for cooperative binding of both WW domains (10). However, the exact mode of interaction between the tandem-WW domains of FBP21 and elongated ligands is still not well understood and the cellular proteome specifically

\* This work was supported by Deutsche Forschungsgemeinschaft Grants FG806 and SFB765.
\[1\] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S7 and Tables S1 and S2.
\[2\] Both authors contributed equally to this work.
\[3\] To whom correspondence should be addressed. Tel: 49-30-94793-181; Fax: 49-30-94793-181; E-mail: freund@fmp-berlin.de.

\[3\] The abbreviations used are: PRS, proline-rich sequence; snRNP, small nuclear ribonuclear particle; FBP, formin-binding protein; SILAC, stable isotope labeling by amino acids in cell culture; ITC, isothermal titration calorimetry; PRE, paramagnetic relaxation enhancement.
Multivalent Interactions within the Pre-spliceosome

targeted to the tandem-WW domains is also elusive. We therefore set out to profile cellular interaction partners of FBP21-tandem-WW by SILAC/MS analysis. Several candidate proteins were identified that match the requirement of bivalent binding with some targets displaying 3 or more PRS motifs. We show in vitro that the strength of the interaction of FBP21 with SmB and SF3B4 peptides depends on the valency of the ligand. Affinity enhancement due to the tandem arrangement of the WW domains is moderate for bivalent ligands but increases significantly for the tetravalent SmB binding partner. Moreover, the tandem action of both WW domains seems to be more relevant for ligands containing more than two motifs. Variation of the length and composition of the linker between the two WW domains does not compromise the interaction with different SmB ligands. Finally, we show data that are consistent with two perpendicular binding orientations, supporting a model of a dynamic positional and orientational equilibrium of ligands interacting with the FBP21-WW domains.

EXPERIMENTAL PROCEDURES

Constructs—For cloning of the human FBP21-WW constructs, a fragment encoding residues 122–196 (tandem-WW wt) was amplified by PCR and cloned into pET-28a via NdeI and XhoI restriction sites. To disrupt the PRS-binding site, relevant Trp residues 151 and 191 (position 29 and 70 in the tandem-WW construct) were replaced by Ala. Constructs with various linker length between the individual WW domains were designed by deleting residues 156–162 (35–41, “Δ35–41”) or 154–165 (33–44, “Δ33–44”). Additionally we replaced residues 156–163 (35–42, “glycine linker”) by a GGGGSGGGG stretch.

Cloning of the GST fusion construct of FBP21-tandem-WW was achieved by inserting FBP21 residues 122–196 into pGEX4T1 via BamHI/XhoI restriction sites. For the CFP- and YFP-tagged constructs, DNA fragments encoding FBP21 residues 122–196, full-length SmB, and full-length SF3B4 were amplified by PCR and cloned into a pECFP/EYFP-N1 vector (Invitrogen) using EcoRI/AgeI restriction sites for insertion of wt) was amplified by PCR and cloned into pET-28a via NdeI restriction sites. To disrupt the PRS-binding site, relevant Trp residues 151 and 191 (position 29 and 70 in the tandem-WW construct) were replaced by Ala. Constructs with various linker length between the individual WW domains were designed by deleting residues 156–162 (35–41, “Δ35–41”) or 154–165 (33–44, “Δ33–44”). Additionally we replaced residues 156–163 (35–42, “glycine linker”) by a GGGGSGGGG stretch.

Cloning of the GST fusion construct of FBP21-tandem-WW was achieved by inserting FBP21 residues 122–196 into pGEX4T1 via BamHI/XhoI restriction sites. For the CFP- and YFP-tagged constructs, DNA fragments encoding FBP21 residues 122–196, full-length SmB, and full-length SF3B4 were amplified by PCR and cloned into a pECFP/EYFP-N1 vector (Invitrogen) using EcoRI/AgeI restriction sites for insertion of wt) was amplified by PCR and cloned into pET-28a via NdeI restriction sites. To disrupt the PRS-binding site, relevant Trp residues 151 and 191 (position 29 and 70 in the tandem-WW construct) were replaced by Ala. Constructs with various linker length between the individual WW domains were designed by deleting residues 156–162 (35–41, “Δ35–41”) or 154–165 (33–44, “Δ33–44”). Additionally we replaced residues 156–163 (35–42, “glycine linker”) by a GGGGSGGGG stretch.

Cloning of the GST fusion construct of FBP21-tandem-WW was achieved by inserting FBP21 residues 122–196 into pGEX4T1 via BamHI/XhoI restriction sites. For the CFP- and YFP-tagged constructs, DNA fragments encoding FBP21 residues 122–196, full-length SmB, and full-length SF3B4 were amplified by PCR and cloned into a pECFP/EYFP-N1 vector (Invitrogen) using EcoRI/AgeI restriction sites for insertion of wt) was amplified by PCR and cloned into pET-28a via NdeI restriction sites. To disrupt the PRS-binding site, relevant Trp residues 151 and 191 (position 29 and 70 in the tandem-WW construct) were replaced by Ala. Constructs with various linker length between the individual WW domains were designed by deleting residues 156–162 (35–41, “Δ35–41”) or 154–165 (33–44, “Δ33–44”). Additionally we replaced residues 156–163 (35–42, “glycine linker”) by a GGGGSGGGG stretch.
Multivalent Interactions within the Pre-spliceosome

Proteins were dialyzed against PBS, 1 mM EDTA, pH 7.4. Peptides were dissolved in dialysis buffer and solutions were corrected to pH 7.4. Protein concentrations in the measurement cell ranged between 20 and 350 μM and were titrated against increasing amounts of 0.25–4 mM peptide stock solutions. To rule out other unspecified effects, peptide was also titrated against buffer and buffer against protein. To obtain the stoichiometry (N), association constant (K_a), and the change of enthalpy (ΔH) and entropy (ΔS) of the reaction, peak areas were integrated and reaction heats were plotted against the molar ratio by fitting to a “One Set of Sites” model until χ-square reached a minimum.

Protein Preparation—Proteins were expressed as His- or GST-tagged fusion constructs in Escherichia coli BL21(DE3). Proteins were purified by the corresponding affinity matrices, thrombin cleaved in case of the His-tagged proteins, and applied to a gel filtration column (Superdex-75) at 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.5, or 1× PBS, 1 mM EDTA, pH 7.4. Proteins were concentrated by Centrifugal Filter Units (regenerated cellulose, 3 kDa cutoff, Millipore) if necessary. In the case of NMR samples, proteins were expressed in M9 minimal medium supplemented with [15N]NH4Cl and/or [13C]glucose and had a final concentration of 0.1–0.9 mM. Peptide SPOT Experiments—SPOT analyses were performed as described (15). Membranes were incubated with either 7.5 μg/ml of GST-WW1/GST-WW2 or with 10 μg/ml of GST-tandem-WW.

NMR Spectroscopy—Measurements were performed at 300 K on a 750 MHz AV and a 600 MHz DRX spectrometer (Bruker) equipped with triple resonance cryoprobes, respectively. Protein samples were buffered in 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, 10% D_2O at pH 6.5. NMR data were processed using Topspin (Bruker) and spectra evaluation was performed using CCPNMR Analysis software (16).

Assignment of Backbone NH Resonances—Resonance assignment of the unbound tandem-WW WT was performed via two-dimensional 15N-HSQC spectra in combination with standard triple resonance CBCA(CO)NNH/CBCANNH experiments. The resonance assignment of the ligand-bound WT was generated by a 15N-HSQC based titration experiment. Resonances of the tandem-WW domain knock out mutants (W29A/W70A) were assigned by standard triple resonance HNCA/HN(CO)CA-BEST experiments in a constant time mode (17).

Chemical Shift Mapping—The determination of relevant binding epitopes was realized by titrating unlabeled ligand (supplemental Table 1) to 15N-labeled tandem-WW WT (0.2 mM). The protein-ligand ratios were increased stepwise to a final ratio of 1:6 (SmB-4), 1:10 (SmB-2 and SF3B4-2), or 1:100 (SmB-1, SmB-1A, SmB-1B, SmB-1C, SmB-1D, and SF3B4-1). The weighted chemical shift changes observed in 15N-HSQC spectra of every titration end point were calculated based on Equation 2.

$$\Delta \delta^H / 15N = \sqrt{\delta^H + (0.15\delta^{15N})^2}$$  (Eq. 2)

15N-T_1/T_2 Backbone Relaxation Experiments—R_1 and R_2 rates were calculated for 15N-labeled tandem-WW WT (0.2 mM) in the unbound state, in complex with SmB-2 (1:6) and SmB-4 (1:3) at 750 MHz. The extraction of R_1/R_2 rates for NH backbone resonances was based on a set of 15N-T_1/T_2-HSQC spectra (18) with incremented relaxation delays of 6, 10, 14, 18, 22, 26, 34, 50, 82, 162, and 242 ms for T_1 and 12, 52, 102, 152, 202, 402, 802, 1602, and 3202 ms for T_2. Peak heights were plotted against the corresponding relaxation times and fitted to an exponential decay function using CCPNMR Analysis software. Correlation times were calculated based on Equation 3 (19).

$$\tau_c = \frac{1}{2\omega N \sqrt{T_1^2 - 7}}$$  (Eq. 3)

Paramagnetic Relaxation Enhancement (PRE) Experiments—To address the binding orientation of a bivalent ligand, the peptide X-DAPA-MPPPPGMRPPPPGMRGLL, X represents the TEMPO moiety and DAPA the 2,3-diaminopropionic acid (TEMPO-SmB-2), was synthesized (EMC Microcollections). 15N-Labeled samples (0.2 mM) of tandem-WW WT and tandem-WW domain mutants were complexed with TEMPO-SmB-2 added in 6-fold excess. The TEMPO-based PRE effect for the two constructs was analyzed by a comparison of the peak heights of NH resonances in the case of active and inactive spin labels by 15N-HSQC-SOFAST experiments (20). The inactivation of the TEMPO group within the ligand was achieved by the addition of a 2-fold excess of ascorbic acid.

Homology Modeling of FBP21-Tandem-WW in Complex with a Bivalent SmB Ligand (SmB Amino Acid 213–227)—Each FBP21-WW domain bound to a ligand segment was modeled separately. As a starting point the crystal structure of the FE65 WW-domain in complex with a PRS peptide (Protein Data Bank code 2HO2) was used. Missing atoms were added using default internal coordinates of CHARMM22 (21, 22). Furthermore, several possible binding modes were considered. The criteria to select the appropriately bound ligand segments are: (i) maximal contact area between ligand and pocket, (ii) fitting of the prolines into the known pocket of WW-domains, (iii) favorable hydrophobic contacts of the ligand with the WW-domain, and (iv) formation of a conserved hydrogen bond between the pocket tryptophan Nε in both WW domains and the carbonyl group of the residue i+1 to the last proline in the canonical orientation or the carbonyl group of residue i–2 in the inverted orientation. Both ligand segments bound to the individual WW domains were connected to a linear structure followed by the introduction of the interdomain linker. The model was subsequently optimized using the CHARMM22 force field. The procedure requires a stepwise minimization by gradually relaxing the structural strain as described (23, 24). Finally global energy minimization without constraints was applied. Backbone root mean square deviations of WW1/WW2 compared with the first coordinate set of NMR structure PDB code 2JXW is 0.93 Å/1.53 Å for the canonical and 1.01 Å/1.72 Å for the inverted binding mode.

RESULTS

FBP21-Tandem-WW Domains Interact with RNA-processing Factors—Because we were interested in the characteristic features of the FBP21-tandem-WW domains binding to PRS con-
taining proteins we started out by a proteomic experiment based on SILAC/MS. We thereby utilized an epitope inhibition approach as it was previously described for GYF and UEV domains (7, 25) and that capitalizes on the specific isotope enrichment of proteins bound to a binding site blocked by an inhibitor in the second pulldown. We used the SmB-2 peptide (supplemental Table S1) from the known interaction partner SmB as an inhibitor and obtained a large number of highly enriched proteins. To increase the robustness of our results the second experiment was performed with inversely labeled cell lysates (see “Experimental Procedures”). Furthermore, we only considered proteins meeting stringent enrichment criteria as putative binders and the list of the hundred most highly enriched proteins is given under supplemental Table S2. Interestingly, 74% of the proteins in this list are associated either with mRNA processing or RNA binding in agreement with previous results obtained with the second WW domain of FBP21 (6). A prominent fraction of these targets is functionally associated with the spliceosome (supplemental Fig. S1) and only a few expected binding partners of FBP21 could not be detected in our pulldown (6), e.g. the known interactor SmB/B’, which runs close to the GST-WW fusion construct in SDS-PAGE. In our pulldown (6), few expected binding partners of FBP21 could not be detected if only nuclear speckles were considered (Fig. 2B, n = 24).

**Ligand Valency Strongly Affects Binding Affinities**—To assess the ligand binding properties of SmB/B’ in regard to the WW domains of FBP21 we first fragmented the PRS tail into individual peptides. Four peptides each containing a potential binding motif were analyzed for their interaction with the tandem WW domains by NMR and isothermal titration calorimetry (supplemental Fig. S4). All of the peptides induced very similar chemical shift changes in both of WW domains of the tandem construct and therefore indicate a conserved recognition mode. This is in line with the very similar interaction profile of the individual WW domains with respect to longer overlapping PRS sequences (15–19 amino acids in length, see supplemental Fig. S2). The binding affinities for the individual peptides differ, with the best individual peptide still displaying a very high $K_D$ of 220 μM. We then assessed the importance of multivalency within the PRS ligand by measuring ITC curves for SmB ligands containing 1 (SmB-1), 2 (SmB-2), and 4 (SmB-4) binding motifs and for SF3B4 comprising 1 (SF3B4-1) or 2 (SF3B4-2) motifs. A large effect on affinities can be observed when comparing a bivalent and a monovalent PRS ligand. Here, reductions in affinity >10 are observed with $K_D$ values of 20 ± 1 μM for the interaction with a ligand containing two binding sites and 295 ± 73 μM for a monovalent ligand (Fig. 2, A and B). A similar difference in affinity was observed when comparing the SF3B4-derived peptides with either one (SF3B4-1) or two (SF3B4-2) binding sites ($K_D$ values of 570 ± 5 and 38 ± 2 μM, respectively, Table 2). The importance of a second proline-arginine motif in longer ligands is also evident by analyzing a truncation series of peptide spots from bivalent SmB peptides (supplemental Fig. S5). Here the intensity arising from bound tandem-WW domains is significantly reduced when relevant residues of either of the two motifs are deleted. To obtain information about the contribution of each single domain to the binding affinity we determined binding constants for tandem-WW constructs where binding-relevant tryptophans were substituted by alanine. NMR spectroscopy showed that the mutations did not hamper folding of the WW domains (data not shown), whereas binding
to the respective PRS motifs was compromised by the mutation (10). Substitution of the binding relevant Trp-29 or Trp-70 by alanine resulted in $K_D$ values of $\sim 60 \mu M$ for the interaction with SmB-2 (Table 2). A moderate decrease in affinity compared with the WT construct was also observed for the interaction of these W29A and W70A mutants with SF3B4-2 (Table 2). When

---

### TABLE 1

Multiple PRMs found within the 100 most highly enriched proteins

| Gene Name | pppp | pp(x)$_{A,-}$ | pp(x)$_{A,+}$ | Ratio* | Uniprot$^A$ or Bioinformatic Harvester$^B$ comment, shortened |
|-----------|------|---------------|---------------|--------|-------------------------------------------------------------|
| SF3A1     | 1    | 5/6           | 1/8           | 27 ± 2 | Subunit of SF3A required for 'A' complex assembly$^A$. |
| SF3A2     | 2    | 5/6           | 1/6           | 27 ± 3 | As SF3A1$^A$.                                               |
| SF3B2     | 8    | 4/10          | 1/5           | 47 ± 12| Subunit of SF3B required for 'A' complex assembly$^A$.     |
| SF3B4     | -    | 12/12         | -/-           | 18 ± 10| As SF3B2$^A$.                                               |
| SR140     | -    | -/5           | 6/3           | 39 ± 11| U2 snRNP-associated SURP motif-containing protein$^A$.     |
| SFPQ      | 3    | 2/3           | 1/3           | 26 ± 25| Essential pre-mRNA splicing factor required early in spliceosome formation and splicing catalytic step II$^A$. |
| U1C       | -    | 2/2           | -/3           | 30 ± 7 | Associated with U1$^A$.                                    |
| SRPK2     | 9    | -/-           | -/1           | 19 ± 0 | Phosphorylates proteins, such as SFRS1 and SFRS2. Role in spliceosome assembly and in mediating the trafficking of splicing factors.$^A$. |
| CHERP     | -    | 11/11         | 2/6           | 45 ± 10| Participates in pre-mRNA U12-dependent splicing, interacts with PDCD7$^A$. |
| SRRM2     | -    | 42/28         | 13/12         | 17±11  | Involved in pre-mRNA splicing. May function at or prior to the first catalytic step of splicing at the catalytic center of the spliceosome. $^A$. |
| SFRS8     | 1    | 5/-           | 2/3           | 26 ± 9 | Member of the arginine/serine-rich splicing factor family$^B$. |
| SFRS15    | 4    | 8/8           | 9/6           | 31 ± 28| Member of the arginine/serine-rich splicing factor family. A similar protein in Rat appears to bind the large subunit of RNA polymerase II and provide a link between transcription and pre-mRNA splicing$^B$. |
| CA150     | 2    | 10/22         | 2/1           | 25 ± 4 | Regulates transcriptional elongation and pre-mRNA splicing. Interacts with the hyperphosphorylated C-terminal domain of RNA polymerase II via multiple FF domains, and with the pre-mRNA splicing factor SF1 via a WW domain.$^B$. |
| HNRPG     | 1    | 22/22         | 2/2           | 60 ± 61| RNA-binding protein, may be involved in pre-mRNA splicing, Component of complex C$^A$. |
| HNRPU1    | 7    | 3/5           | 1/1           | 19 ± 3 | Acts as a basic transcriptional regulator, also a role in mRNA processing and transport.$^A$. |
| LARP      | -    | 3/12          | 5/6           | 25 ± 10| La ribonucleoprotein domain family, member 1$^B$. |
| CPSF6     | 1    | 8/10          | 1/1           | 30 ± 3 | Component of CFIm, plays key role in pre-mRNA 3’ processing. Binds to cleavage and polyadenylation RNA$^A$. |
| CPSF7     | 2    | 6/6           | 1/1           | 35 ± 15| Component of CFIm, plays key role in pre-mRNA 3’ processing. Binds to cleavage and polyadenylation RNA$^A$. |
| WIPF1     | 11   | 26/24         | 2/8           | 21 ± 1 | Induces actin polymerization/ redistribution. Contributes with NCK1 and GRB2 in the recruitment and activation of WASL.$^A$. |
| WIPF2     | 17   | 32/22         | 3/7           | 28 ± 5 | Interacts with WASL, role in actin-microspike formation through cooperation with WASL$^A$. |
| WASL      | 20   | 28/17         | -/-           | 24 ± 0 | Induces actin polymerization and redistribution$^A$. |

The gray intensity of the gene name column correlates with the protein complex integration/molecular function (spliceosome/splicing) and is listed top-down according to common function or localization with FBP21. Number of motifs in front of the slash represents canonically oriented motifs (N- to C-terminal), whereas those behind the slash are noncanonical (C- to N-terminal). Means of enrichment (ratio) and S.D. results are from two independent experiments, whereby in the second replicate the labeling strategy was inverted (“sample” versus “control,” see “Experimental Procedures”).
both Trp were replaced, no relevant binding between the tandem-WW construct (W29A/W70A) with SmB-2 and SF3B4-2 was observed (data not shown).

To further investigate the effect of multiple motif recognition we studied interaction of the WW domain constructs with a SmB peptide containing four PRS motifs in ITC experiments.

**FIGURE 1.** In vivo interaction of FBP21-tandem-WW domains with splicing factors. A, colocalization of SmB-YFP and FBP21-tandem-WW-YFP (upper panel) and SF3B4-CFP and FBP21-tandem-WW-YFP (lower panel) in nuclear speckles. HEK293FT cells were transfected with the respective constructs and analyzed on a Leica SP2 confocal microscope. The bar is 20 μm. B, FRET efficiencies for the energy transfer from SF3B4-CFP and SmB-CFP to FBP21 tandem-WW-YFP were determined as described under “Experimental Procedures.”

**FIGURE 2.** Increase in affinity due to elevated ligand valency for the interaction of the FBP21-tandem-WW domains with SmB-derived peptides detected by isothermal titration calorimetry and scheme of constructs used in this study. Shown are the interactions detected by ITC of the WT tandem-WW domain of FBP21 with a supposedly monovalent (A), bivalent (B), and tetravalent (C) ligand. Upper panel, shows enthalpy changes upon injection of peptide into the measurement cell containing the FBP21-tandem-WW domains. Bottom panel, shows integrated power peaks fitted with a 1:1 model. Enthalpy and entropy changes are listed in Table 2. S.D. results from at least two experiments. Peptide sequences are listed under supplemental Table S1.

A, scheme of constructs. Top construct, position 122–196 of FBP21-tandem-WW domains WT (WT). Second line left, the tryptophan at position 29 (corresponds to position 150 in the full-length protein) was substituted by alanine resulting in loss of binding of the N-terminal WW domain (W29A). Second line right, the tryptophan at position 70 (corresponds to position 191 in the full-length protein) was substituted by alanine and compromises the binding of the C-terminal WW domain (W70A). Third line left, both tryptophans at positions 29 and 70 were substituted by alanine resulting in a construct with two compromised WW domains (W29A/W70A). Third line right, the linker region regarding positions 35–42 (corresponding to 156–163 in the full-length protein) was replaced by a highly flexible GGGGSGGG stretch (glycine linker), enhancing the flexibility of the linker. Bottom line left, linker residues 35–41 (156–162 in the full-length protein, Δ35–41) were deleted. Bottom line right, the linker region at positions 33–44 (154–165 in the full-length protein, Δ33–44) was deleted.

To further investigate the effect of multiple motif recognition we studied interaction of the WW domain constructs with a SmB peptide containing four PRS motifs in ITC experiments.
An apparent $K_D$ of 4 $\mu$m was measured for the WT construct, indicating that multivalency of the ligand further enhances affinity (Fig. 2c). When the W29A and W70A mutant constructs were used, a $\sim$10-fold higher $K_D$ was observed, indicating that the WW domain tandem arrangement is more efficient in promoting affinity enhancement at higher ligand valencies. Affinity enhancement for the long SmB-4 ligand is driven by increased enthalpy (Table 2), whereas entropy compensation is observed in all cases. Furthermore, the stoichiometry for complex formation of the WW1-WW2 construct with monovalent ligands was 1:2 (Fig. 2), whereas it was 1:1 for all longer peptides, including SmB-4 (Table 2). NMR chemical shift perturbation showed that SmB-1 and SF3B4-1 affect both WW domains of FBP21, suggesting the binding of two peptides to a single tandem-WW fragment (supplemental Fig. S6). In contrast, when more than one motif was available for binding the tandem-WW domains was engaged in a bivalent complex. Interestingly, the binding stoichiometry to the tetravalent ligand, on average, determined by ITC was still 1:1 and this was confirmed by NMR relaxation measurements (supplemental Fig. S7) that showed very similar overall correlation times values for the tandem WW domains in complex with either bivalent ($\tau_c = 8 \pm 0.3$ ns) or tetravalent ligands ($\tau_c = 9 \pm 0.5$ ns), both in agreement with a complex of an approximate molecular mass of 13 kDa.

To investigate whether these bivalent complexes preserve the known binding pockets of WW domains, NMR chemical shift experiments were used to map the respective binding epitopes. The interactions are in fast exchange relative to the NMR time scale, independent of the length of the ligand. Clearly, the epitopes affected by the individual SmB and SF3B4 ligands are very similar and comprise the expected proline-binding pockets within both WW domains (supplemental Fig. S6) similar to those reported by Huang et al. (10). Interestingly, we could observe a clear relationship between the saturation point of chemical shift changes and the number of presented PRS within the ligand. An increase of potential binding sites leads to a decrease in the required concentration based on an increased affinity. Consequently, whereas a 3-fold excess of the tetravalent ligand results in similar shifts as a 6-fold excess of the bivalent and a 60-fold excess of the monovalent peptide (Fig. 3). These findings are in line with our ITC data (Table 2) and show that affinity enhancement is correlated with the multivalency of the ligand.

The Effect of Linking the Two WW Domains—To address the question, whether the binding of a bivalent ligand leads to an altered flexibility within the protein backbone, we compared $R_2$ rates for backbone NH resonances of the unbound tandem-WW and the tandem-WW domains in complex with the bivalent SmB-2 ligand. The comparison was based on averaged and truncated $R_2$ mean values for WW1 (amino acids 7–29), the
linker region (amino acids 37–45), and WW2 (amino acids 48–70). We could observe a significant increase of the $R_2$ value for the linker region (unbound: 7.59 ± 0.16 s$^{-1}$; bound: 9.12 ± 0.1 s$^{-1}$) and for WW2 (unbound, 10.92 ± 0.15 s$^{-1}$; bound, 16.45 ± 0.14 s$^{-1}$). The $R_2$ value of the first WW domain was not affected (unbound, 17.25 ± 0.82 s$^{-1}$; bound, 16.53 ± 0.14 s$^{-1}$) by the binding event. We therefore asked the question of how far changes in linker length would influence coupling of the WW domains and whether correct spacing between the domains is critical for binding to bi- or multivalent ligands. To assess linkage phenomena, we used variants that either contained a more flexible linker of the composition GGGGSGGG or truncation constructs with either 7 or 12 amino acid deletions (Fig. 2D, a scheme of the constructs used here). Dissociation constants for these artificial linker constructs did not differ significantly from WT values (Table 2), indicating that linker composition and flexibility between the domains is not crucial for bivalent ligands with shortly spaced motifs.

Bidirectional Binding of PRS to Tandem-WW Domains—Ligand titration experiments of the tandem-WW domains indicate that bivalent ligands as well as the tetravalent ligand are in fast exchange relative to the NMR time scale, however, the binding mode for these longer ligands is incompletely understood. More specifically, it has been observed that individual motifs can bind in two orientations, due to the pseudosymmetry of the PPII helix. We therefore asked the question of whether such a bidirectional binding mode might be adopted by ligands of FBP21-tandem-WW domains. A TEMPO spin label was attached to the N terminus of the bivalent SmB peptide that should lead to PRE for residues in the vicinity of the unpaired electron. Relaxation enhancement was monitored by comparing signals from measurements with active and inactive TEMPO-labeled peptides. The structure of the TEMPO-labeled peptide is shown as an inlet in panel B. Chemical shift changes of NH backbone resonances are shown as an inlet (dotted lines) for the single domain mutants (C and D).

FIGURE 4. Paramagnetic relaxation enhancement of $^{15}$N-$^1$H resonances of the FBP21-tandem-WW domain upon binding of a bivalent PRS ligand, which comprises an N-terminal TEMPO label. The figure shows the ratio of peak heights for NH backbone resonances of FBP21-tandem-WW WT (A), the binding deficient “W29A/W70A” mutant (B), the single domain mutants W70A (C) and W29A (D) all titrated with a 6-fold excess of TEMPO-labeled bivalent SmB.
Multivalent Interactions within the Pre-spliceosome

FIGURE 5. Model of bidirectional binding of a bivalent SmB-derived ligand to the FBP21-tandem-WW domains confirmed by spin label experiments. The figure shows a model of the tandem-WW domains (gray) in complex with a bivalent peptide (SmB amino acids 213–227, pink). Side chains of hydrophobic residues (Trp, Tyr) involved in ligand binding are displayed. The ligand could bind either in a canonical (right) or inverted (left) orientation as confirmed by spin label experiments. The three most affected residues within each WW domain are highlighted in bright red. Their close spatial location to the spin label attachment side underlines the proposed bidirectional binding mode.

WW2 domain residues cannot explain the effect on the NH groups of the WW1 domain residues in the case of an inverted peptide orientation. We therefore have to postulate a dynamic equilibrium of at least two orientations for the ligand as visualized by a conceivable homology model of FBP21-tandem-WW in complex with a bivalent SmB ligand (Fig. 5).

DISCUSSION

Affinity Enhancement by Multivalency—Proline-rich sequence recognition, with few exceptions, is characterized by recognition rules that allow for peptide degeneracy. Because individual PRS motifs in many cases occur as multiple variants within a single protein, it is of importance to understand the mechanism of multivalent binding at a molecular level. Also, it is not clear how far multiple low-affinity interactions of PRS with protein domains such as WW or GYF contribute to spliceosomal function. Although efficient experimental setups that unambiguously define the functional role of long PRS in proteins are missing we now provide evidence that multivalent binding partners are preferred targets for the tandem-WW domains containing protein FBP21. First, we show that proteins containing long PRS are highly enriched in the soluble cellular fraction when precipitated by FBP21-tandem-WW domains. Furthermore, the cellular interaction of functionally related proteins such as SF3B4 argu es for a role of extended regions of PRS in early spliceosomal assembly. Interestingly, the affinity of the tandem-WW domains construct of FBP21 for single PRS motifs is very low, and this is conceivably true for several of the spliceosomal proteins identified here. For FBP21 binding is already significantly increased for a bivalent ligand (Table 2) and this increase is also observed for the W29A and W70A mutants, indicating the importance of ligand valency. However, whereas only a minor further increase in affinity is observed for the mutant constructs upon addition of the tetravalent ligand a 4–5-fold affinity enhancement is seen for the WT tandem-WW domains. Therefore, the effect of WW domain linkage is more prominent for a multi- versus a bivalent ligand and high-affinity binding in this system is seemingly optimized for efficient binding of the tandem-WW domains to multivalent PRS ligands. It has to be noted, however, that the definition of multivalency is not unambiguous for many of the proteins found in our pulldown, because motifs often overlap and because nonproline residues such as arginine are certainly of importance for the stabilization of the respective complexes.

Influence of the Linker Region—Our results indicate that a shorter linker between the WW domains does not interfere with the observed affinity enhancement upon binding to the SmB and SF3B4 ligands (Table 2). Although the structure of the tandem WW domains does not give any indication of direct domain-domain interactions (10) it has been found that the short ~10 residue linker is dynamically more restricted in the complexed form. Although such motional coupling upon binding could come at an entropic cost, it will be more than counterbalanced by the experimentally observed negative enthalpies (Table 2). Furthermore, the observed overall affinity enhancement, at least in part, will be a consequence of avidity that is operative over a large range of distances between WW domains as well as the PRS motifs. In the case of the YAP and TAZ transcription factors the linker region between WW domains is significantly longer (~28 residues) and a 6-fold higher affinity binding to a bi- versus monovalent ligand is observed, indicating that avidity, whereas still operative, decreases with increasing linker length (30). A somewhat different scenario has recently been described for the Smad ubiquitination regulatory factor, where the presence of an additional 26 residues between its WW2 and WW3 domain in the long isoform of the protein severely uncouples the binding to a composite ligand (31). However, in this case, the two WW domains directly contact each other and the small but common interface probably invokes true cooperativity. An interesting twist of domain-domain coupling has been observed in the Ned44-like protein suppressor of deltex (Su(dx)), where an autoinhibitory interface between the third and fourth WW domain is released upon ligand binding (32). Similarly, the rigid helix connecting the two WW domains of splicing factor Prp40 leads to an overall structure of defined geometry, requiring a distinct topology of proline-rich ligands (33). However, even for Prp40 a set of different ligands can be bound by the tandem-WW domains arguing for promiscuity to be a general property of WW domain recognition (33).

Proline-rich Sequences in the Spliceosome—It is noteworthy that FBP21, as several other nuclear WW domain proteins, also contains an RNA binding zinc finger domain. RNA binding of these types of domains is thought to be of limited specificity and it is conceivable that such promiscuous interaction allows for movement of the domains along longer stretches of nucleotides. This has been investigated in more detail for DNA-binding domains and demonstrated that sparsely populated encounter complexes precede the formation of a specific DNA-binding complex (34). This process allows for inter- as well as intramolecular translocations and is steered by long-range electrostatic interactions. Charge attraction is found to be of importance for the tandem-WW domains binding of FBP21 to a bivalent ligand (supplemental Fig. S3) and we envisage that similar search processes are of importance for spliceosomal maturation (35). It is well conceivable that checkpoint decisions such as splice-site selection or poly(A) tail addition are complemented by highly dynamic protein recognition events. We suggest the PRS within spliceosomal proteins provide a scaffold for “delocalized” binding by protein adaptor domains such as WW
and our data are in line with the existence of multiple complexes along long unstructured PRS. The observation that the long PRS sequences of core splicing proteins, such as SmB/B’ have not been assigned to measurable electron densities by EM (36) could be an indication that these sequences display intrinsic flexibility. Together with the observed degenerate binding of PRS recognition domains, a conformational space within several hundred Å of a certain intracellular locale would be accessible for the formation of encounter complexes within the spliceosome. It will be challenging to prove this hypothesis and to show the importance of semi-specific multivalent binding modes, because proline-rich sequences are generally not highly conserved during eukaryotic evolution. In many cases, certain proteins of the spliceosome have lost PRS recognition sites, whereas others acquired new ones. Despite this shuffling of PRS motifs within larger protein complexes more recent studies suggest that long PRS stretches are of physiological importance.

Motifs within larger protein complexes more recent studies have not been assigned to measurable electron densities by EM (36) could be an indication that these sequences display intrinsic flexibility. Together with the observed degenerate binding of PRS recognition domains, a conformational space within several hundred Å of a certain intracellular locale would be accessible for the formation of encounter complexes within the spliceosome. It will be challenging to prove this hypothesis and to show the importance of semi-specific multivalent binding modes, because proline-rich sequences are generally not highly conserved during eukaryotic evolution. In many cases, certain proteins of the spliceosome have lost PRS recognition sites, whereas others acquired new ones. Despite this shuffling of PRS motifs within larger protein complexes more recent studies suggest that long PRS stretches are of physiological importance.

...