Chapter 6

Ligand recognition of the Prp40 WW domain pair

‡ The WW domain pair of the yeast splicing factor Prp40 is believed to function in cross-intron bridging (Abovich & Rosbach, 1997) and in coupling transcription to splicing. The Prp40 WW domains have been shown to contact the 5' splice-site (ss) and interact with the branch-point binding protein BBP (also referred to as Msl5 and ySF1) bringing the 5' ss and the branch-point in spatial proximity. Furthermore, the Prp40 WW domains interact with the U5 snRNP core component Prp8 (Abovich & Rosbach, 1997). Interestingly, recent studies show that, apart from the U1 snRNP, Prp8 associates with the 5' ss prior to the dissociation of BBP from the commitment complex (Maroney *et al.*, 2000). With both interaction partners present at the same time, the *two* Prp40 WW domains could in principle interact simultaneously with BBP and Prp8. After BBP is displaced from the spliceosome by the U2 snRNP, Prp8 associates with the U4/U6.U5 tri-snRNP at the 3' ss. At this stage, the interaction between the proline-rich N-terminus of Prp8 and the Prp40 WW domains is believed to bridge the 5' ss with the 3' ss (Abovich & Rosbach, 1997). Since BBP is no longer present, the interaction between Prp40 and Prp8 could benefit from cooperative effects by the presence of *two* WW domains and a *tandem* proline-rich stretch at the N-terminus of Prp8. Interestingly, both splicing factors interacting with Prp40 (Prp8 and BBP) contain PPxY motifs (with x being any residue), which are known to be recognised by a number of WW domains (Kay *et al.*, 2000). This suggests that both Prp40 WW domains recognise the same ligand motif and, hence, can both bind to Prp8 and BBP. In addition, there is a growing body of evidence that transcription and splicing are inter-connected processes (Proudfoot *et al.*, 2002; Maniatis & Tasic, 2002). While it is well established that the C-terminal domain (CTD) of the RNA polymerase II largest subunit plays a central role in coordinating transcription and splicing, the participating splicing factors are less well characterised. Several WW domain containing proteins, including the *Sc.* proteins Prp40, Rsp5 and Ess1, have been shown to interact with phosphorylated CTD repeats (Chang *et al.*, 2000; Morris & Greenleaf, 2000), suggesting a direct function of Prp40 in coupling transcription to splicing.

‡This chapter is based on *Wiesner, S., Stier, G., Sattler, M. and Macias, M. J. (2002). J. Mol. Biol. 234: 807–822* with kind permission from Elsevier.

Based on their protein sequences and ligand specificities, WW domains are classified in four groups (Macias *et al.* (2002); Chapter 4). WW domains in group I interact with PPxY motifs and typically recognise the ligand Tyr *via* an aliphatic residue (Ile/Val) and a His (Macias *et al.*, 1996; Huang *et al.*, 2000; Kanelis *et al.*, 2001; Pires *et al.*, 2001). Group II WW domains are rich in aromatic residues and bind ligands containing PPΨΨP motifs (where Ψ is an aliphatic residue (Aasland *et al.*, 2002) or in some cases an arginine) (Bedford *et al.*, 1997; Chan *et al.*, 1996). Group III WW domains encompass an additional residue in the first loop that interacts with the phosphoSer-Pro motifs as contained in phosphorylated CTD repeats (Verdecia *et al.*, 2000; Lu *et al.*, 1999). Group IV WW domains have the most divergent protein sequences of all WW domain groups; they lack the C-terminal Trp and their targets are so far unknown. Based on the protein sequence, the Prp40 WW domains belong to group II WW domains and thus would be expected to bind PPΨΨP motifs. However, the Prp40 interaction partners yet identified genetically and biochemically belong to group I and III WW domain binding motifs, respectively (Abovich & Rosbach, 1997; Morris & Greenleaf, 2000). BBP contains one PPxY motif, Prp8 a tandem stretch with a PPxY and a PPxF motif, while the CTD repeats consist of phosphorylatable motifs with the consensus sequence YSPTSPS.

This raises questions as whether both Prp40 WW domains can recognise all three binding motifs, whether they show individual selectivity for different proline-rich motifs and how their structures enable this binding promiscuity. To gain more information about this intricate scenario, chemical shift mapping experiments were performed to explore the interaction of the tandem Prp40 WW domains with different proline-rich ligands derived from the proteins BBP, Prp8 and CTD. To investigate whether the Prp40 WW domains can also bind group II WW domain ligands two PPΨΨP motif containing peptides were used that have previously been shown to interact with the mammalian Prp40 orthologue FBP11 and the Abl-SH3 domain (Chan *et al.*, 1996).

6.1 Interactions with proline-rich peptides

NMR is an exceptionally powerful technique to directly analyse the binding properties of individual domains in a multi-domain protein, since it allows to identify immediately the residues participating in the binding. To investigate qualitatively whether the tandem WW domains of Prp40 interact with different proline-rich motifs and to map the residues involved in the interactions, ^{1}H and ^{15}N chemical shift changes upon ligand binding were measured for eight selected peptides (Fig. 6.1 and 6.2, Table 6.1). The peptide sequences were chosen according to the results of previous studies (Abovich & Rosbach, 1997; Bedford *et al.*, 1997; Chan *et al.*, 1996; Morris & Greenleaf, 2000).

Three peptides contained PPxY/F motifs corresponding to residues 94–102 of *Sc.* BBP (PSPPPVYDA), residues 24–32 of *Sc.* Prp8 (PPPPGYEIE) and residues 6–14 of *Sc.* Prp8 (PPPPGFEED) (Fig. 6.1(a) and 6.2(a)). Two peptides displayed PPΨΨP motifs cor-

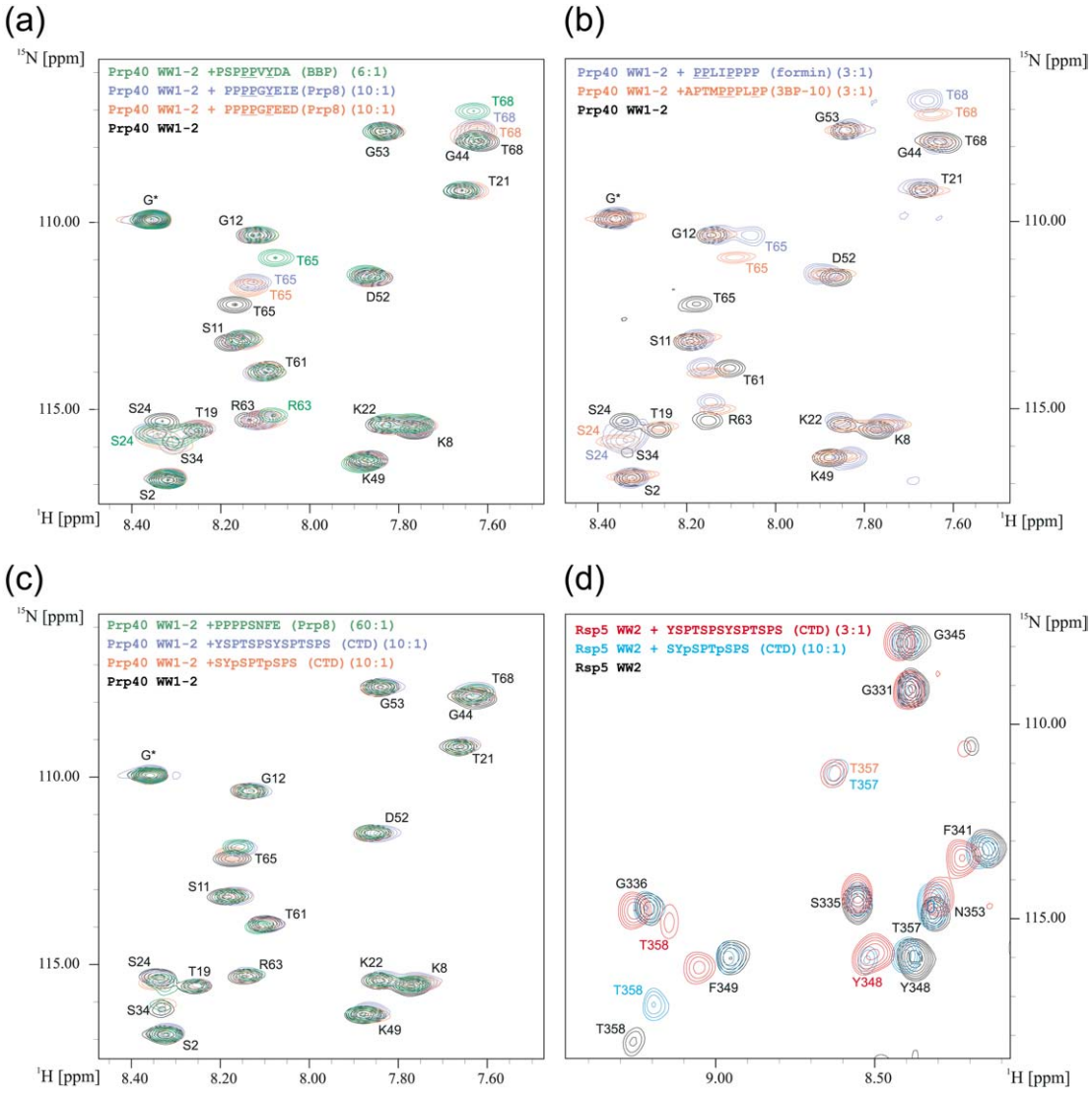

Figure 6.1: Superposition of representative regions of the ^{1}H–^{15}N correlation spectra for the interaction of the Prp40 tandem WW domains and the second Rsp5 WW domain with different proline-rich peptides. The free WW domains are shown in black (reference spectra). In panels (a)–(c), G* corresponds to a glycine residue resulting from the TEV protease cleavage site. (a) Addition of PPxY/F motif containing peptides from BBP (PSPPPVYDA, green) and Prp8 (PPPPGYEIE, blue, and PPPPGFEED, red) to the Prp40 tandem WW domains. (b) Addition of PPΨΨP motif containing peptides from mouse formin (PPLIPPPP, blue) and the Abl-SH3 3BP-10 peptide (APTMPPPLPP, red) to the Prp40 tandem WW domains. (c) Addition of Prp8 peptide (PPPPSNFE, green), the unphosphorylated tandem CTD repeat (YSPTSPSYSPTSPS, blue) and the doubly phosphorylated CTD repeat (SYpSPTpSPS, red) to the Prp40 tandem WW domains. (d) Addition of the unphosphorylated tandem CTD repeat (YSPTSPSYSPTSPS, red) and the doubly phosphorylated CTD repeat (SYpSPTpSPS, cyan) to the second WW domain of Rsp5.

responding to residues 881–888 of the mammalian protein formin (PPLIPPPP) and the Abl-SH3 binding 3BP-10 peptide (APTMPPPLPP) (Fig. 6.1(b) and 6.2(b)). One peptide consisted of residues 53–60 of *Sc.* Prp8 (PPPPSNFE), one of a phosphorylated CTD repeat (SYpSPTpSPS) and another one of an extended, unphosphorylated CTD repeat (YSPTSPSYSPTSPS) (Fig. 6.1(c)). In no case was the binding saturated, since the aim was to detect binding and to characterise the Prp40 WW domain binding sites rather than to provide a quantitative binding analysis. Representative regions of the ^{1}H–^{15}N HSQC spectra of the tandem Prp40 WW domains with all peptides used are shown in Fig. 6.1(a)–(c). The average chemical shift differences in Hz between bound and free forms of the Prp40 WW domains are plotted in Fig. 6.2 (note that the molar protein:peptide ratios refer to the *tandem* WW domains).

In agreement with the relative orientation of the WW domains in the three-dimensional structure of the free protein (see Chapter 5), significant chemical shift changes were in general only observed for residues located in the β-strands. This indicates that no global conformational changes occur upon ligand binding and that each binding site recognises the peptides individually. Among all peptides used only those containing PPΨΨP and PPxY/F motifs interacted with the Prp40 WW domains. Of the PPxY/F containing motifs, the BBP peptide induced larger chemical shift changes than the Prp8 derived peptides, while the PPΨΨP motif caused the largest chemical shift changes of all peptides studied (6.2(a) and (b)). No interaction, however, was detected for the Prp8 peptide PPPPSNFE and for the CTD peptides, which can hence be regarded as negative controls (Fig. 6.1(c)). All chemical shift changes are in fast exchange on the NMR time-scale and are larger for the C-terminal WW domain (WW2) than for the N-terminal WW domain (WW1), despite the fact that both WW domains share 80% sequence similarity in the binding site.

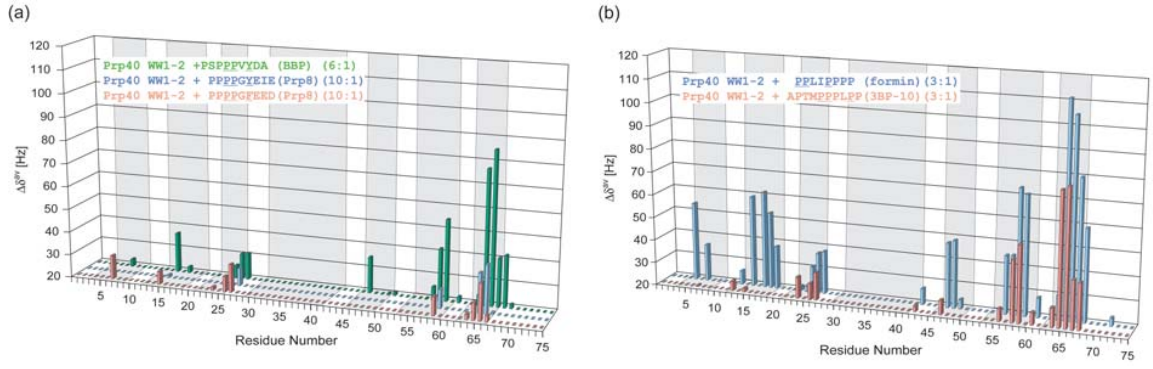

Figure 6.2: Average chemical shift changes in Hz upon ligand binding versus residue number calculated as $\Delta\delta^{av} = [(\Delta\delta_H)^2 + (\Delta\delta_N)^2]^{1/2}$, where $\Delta\delta$ is the difference in the chemical shift observed for the free WW domains and at the respective molar peptide-protein ratio. The peptides are coloured as in Fig. 6.1. Secondary structure elements are represented by grey boxes. (a) PPxY/F motif containing peptides added to the tandem WW domains of Prp40, (b) PPΨΨP motif containing peptides added to the tandem WW domains of Prp40.

Table 6.1: Interaction of the Prp40 tandem WW domains with proline-rich peptides.

Interaction partner	Peptides with PPxY motifs	Peptides with PPΨΨP motifs	Other peptide sequences	Chemical shift changes
BBP (94–102)	PSPPPVYDA			++
Prp8 (6–14)	PPPPGFEED			+
Prp8 (24–32)	PPPPGYEIE			+
Formin (881–888)		PPLIPPPP		+++
3BP-10		APTMPPPLPP		++
Prp8 (53–60)			PPPPSNFE	-
phospho-CTD			SYpSPTpSPS	-
CTD			(Y SPT SPS)$_2$	-

To understand on one hand the differences in the binding between both WW domains and on the other hand, why the induced chemical shift changes are larger for the PPΨΨP motif than for the PPxY/F motifs, the interaction of the single Prp40 WW2 with the BBP peptide was characterised in more detail. To this end, two-dimensional ^{1}H homonuclear NOESY and TOCSY spectra of the Prp40 WW2 in complex with the BBP peptide were recorded at 800 MHz and assigned to derive inter-domain NOEs. The BBP peptide was selected, since it induced the largest chemical shifts amongst all PPxY/F motifs used in this study and because previous studies have only suggested interactions of the Prp40 WW domains with a PPxY motif (Abovich & Rosbach, 1997).

6.2 Interaction with PPxY motifs

The binding site of a WW domain is known to consist of two main cavities (Macias *et al.*, 1996; Huang *et al.*, 2000; Kanelis *et al.*, 2001; Pires *et al.*, 2001). The first cavity accommodates one or two peptide proline-rings and is constituted of residues highly conserved amongst all WW domains (Fig. 6.3), namely the C-terminal Trp (W67 for the Prp40 WW2), a Thr or Ser in the centre of β3 (T65) and an aromatic residue at the beginning of β2 (Y56). Accordingly, in the Prp40 WW2:PSPPPVYDA complex NOEs were observed from W67 and Y56 to two peptide prolines (P4' and P5') (Table 6.2, Fig. 6.3(b)). The second binding pocket (cavity 2), however, is more divergent and hence likely to be responsible for ligand specificity. In the case of the Prp40 WW2, cavity 2 is formed by Y58 (β2), P60 (β2) and R63 (β3) (Fig. 6.3).

In contrast to all known WW:PPxY complexes (Macias *et al.*, 1996; Huang *et al.*, 2000; Kanelis *et al.*, 2001; Pires *et al.*, 2001), the second binding pocket (cavity two) of both Prp40 WW domains and of almost all group II WW domains contains an aromatic residue in the middle of β2 (Y58) (Fig. 6.3), instead of the aliphatic residue characteristic for group I WW

domains. Thus only limited space is left in group II WW domains to allocate another bulky aromatic residue from a PPxY motif. However, an aliphatic residue, as contained in the PPΨΨP peptides used in this study could be accommodated more easily. Hence, a plausible explanation for the recognition of PPΨΨP and PPxY motifs could be that in the PPΨΨP binding (*i.e.* group II) WW domains the aromatic residue in cavity two recognises an aliphatic peptide residue, while in PPxY binding (group I) WW domains an aliphatic residue in cavity two recognises an aromatic peptide residue. This could explain the stronger chemical shift perturbations observed for the PPΨΨP peptides used in this study (Fig. 6.2), since based on their protein sequences the Prp40 WW domains are group II WW domains. Moreover, both Prp40 WW domains lack the His at the end of β2 (Fig. 6.3), which is conserved in PPxY motif binding WW domains and known to form a crucial hydrogen bond with the peptide tyrosine (Macias *et al.*, 2002; Huang *et al.*, 2000).

Despite these disadvantages, which may account for the relatively small chemical shift changes induced by the PPxY/F motif containing peptides, in the WW2:PSPPPVYDA complex the BBP peptide adopts a conformation, in which the peptide tyrosine (Y7') occupies cavity 2 of the Prp40 WW2 (Fig. 6.3). Together with chemical shift changes of 0.32 ppm for the R63 α-proton and of 0.15 ppm for one of the P60 δ-protons, the observed NOEs from Y58, P60 and R63 towards the peptide Y7' (Table 6.2) indicate that the BBP peptide is oriented in the Prp40 WW2 binding site in the same way as in all other known PPxY:WW complexes (Macias *et al.*, 1996; Huang *et al.*, 2000; Kanelis *et al.*, 2001; Pires *et al.*, 2001). The observed

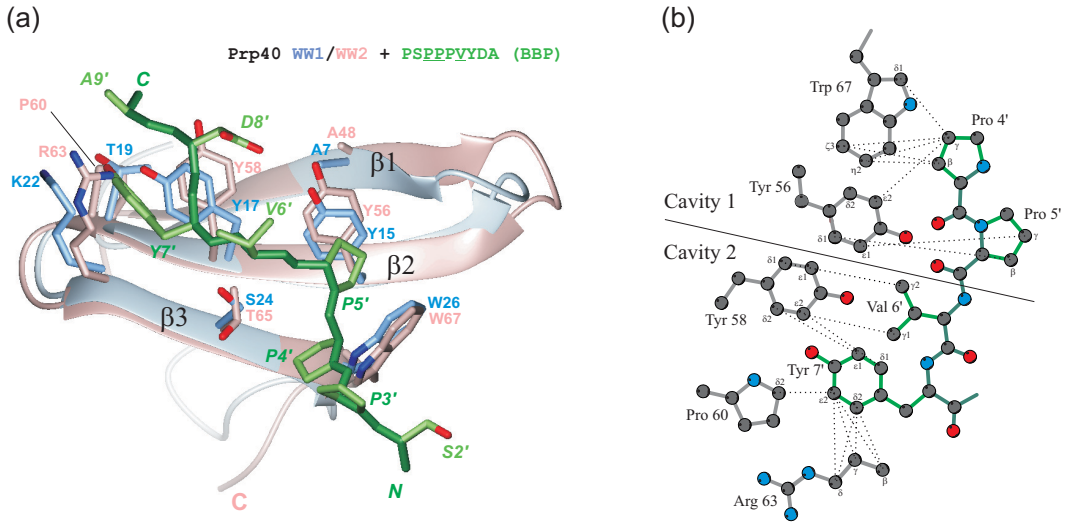

Figure 6.3: The Prp40 WW2:PSPPPVYDA complex. (a) Model of the Prp40 WW2:PSPPPVYDA complex, where the WW2 is depicted in red and the PSPPPVYDA peptide in green. The WW1 is shown in blue for comparison. (b) Schematic representation of the interactions between the Prp40 WW2 (grey) and the BBP peptide (green) produced using the program (Wallace *et al.*, 1995). Intermolecular NOEs are indicated by dashed lines. For reasons of clearness, protons have been removed from the illustration, but proton-proton NOEs are implied. Where NOEs involved diastereotopic protons degenerate in their proton chemical shifts, only one of the possible proton-proton interactions is shown.

Table 6.2: Intermolecular NOEs observed for the Prp40 WW2:PSPPPVYDA complex.

Cavity 1				Cavity2			
Prp40 WW2		PSPPPVYDA		Prp40 WW2		PSPPPVYDA	
Residue	$^1H^{\ddagger}$	Residue	$^1H^{\ddagger}$	Residue	$^1H^{\ddagger}$	Residue	$^1H^{\ddagger}$
Y56	ε_*	P4'	γ_2	Y58	δ_*	V6'	γ_{**}
	δ_*	P5'	β_*		ε_*	V6'	γ_{**}
	ε_*	P5'	γ_*		δ_*	Y7'	ε_*
W67	δ_1	P4'	γ_1		ε_*	Y7'	δ_*
	δ_1	P4'	γ_2	P60	δ_2	Y7'	ε_*
	ζ_3	P4'	β_1	R63	β_1	Y7'	δ_*
	ζ_3	P4'	β_2		β_2	Y7'	δ_*
	ζ_3	P4'	γ_1		γ_1	Y7'	δ_*
	ζ_3	P4'	γ_2		γ_2	Y7'	ε_*
	η_2	P4'	β_1		δ_1	Y7'	δ_*
	η_2	P4'	β_2		δ_2	Y7'	ε_*
	η_2	P4'	γ_1				
	η_2	P4'	γ_2				

‡Protons involved in intermolecular NOEs. H_* and H_{**} denote diastereotopic protons and methyl groups, respectively, with degenerate chemical shifts.

inter-molecular NOEs between the Prp40 WW2 and the BBP peptide are summarised in Table 6.2 and Fig. 6.3(b). To illustrate the interactions in more detail the complex structure was modelled by superimposing the structures of the free Prp40 WW domains (Chapter 5) with the crystal structure of the dystrophin WW:RSPPPYVP complex (Huang *et al.*, 2000).

The reasons why larger chemical shift changes are induced upon ligand binding in Prp40 WW2 as compared to WW1 become apparent from their sequences, assuming that the binding mode is conserved in both domains. The *trans*-configuration of P60 stabilises loop2 in the WW2 contributing to the formation of a β-turn, which is not the case for the threonine-leucine tandem (T19 and L20) in the WW1. In addition, the methyl group of the T19 side-chain occupies more binding surface than the proline ring of P60, which lies mostly in the β-sheet plane (Fig. 6.3(a)).

6.3 Interaction with PPΨΨP motifs

Among all peptides used in this study the PPΨΨP motif containing peptides produced the largest chemical shift perturbations. GST-fusion proteins containing these PPΨΨP motifs have previously been shown to interact with various group II WW domains, namely the mammalian Prp40 orthologue FBP11, the mFBP21 and mFBP28 WW domains, and also to the Abl- and Fyn-SH3 domains (Chan *et al.*, 1996). Together with the data presented in this Chapter, this suggests that the Prp40 WW domains can compete with SH3 domains for the same targets. Moreover, the mammalian Prp40 orthologue FBP11 has been shown to bind not only the Abl-SH3 peptide APTMPPPLPP, but also other PPPLP motif containing peptides (Bedford *et al.*, 1997; Chan *et al.*, 1996). This suggests that the FBP11 WW domains and thus probably also the Prp40 WW domains recognise, besides the polyproline helix, an aliphatic peptide residue (in contrast to the aromatic residue of the PPxY motif).

However, no structural information on how PPΨΨP motifs are recognised by WW domains is available to date. Interestingly, SH3 domains accommodate the proline-rings of PxxP peptide motifs in two aromatic binding pockets in a way that is very similar to the lignad recognition of WW domains (see Chapter 4). Due to the symmetry properties of the polyproline helix these ligands can bind to SH3 domains in two orientations, a C–N orientation with the C-terminal peptide proline in the first binding pocket and the reverse N–C orientation with the N-terminal peptide proline in the first binding pocket (Feng *et al.*, 1994). A similar behaviour has been observed for WW domains, however with different peptide motifs. While the N-terminal proline-rings of PPxY motifs are recognised by the first aromatic binding pocket of WW domains (cavity one in Fig. 6.3(b)), phospho-Ser motifs are bound in the reverse orientation with the N-terminal proline located in the second binding pocket (cavity two) (Fig. 6.5). Therefore, the peptide orientation of PPΨΨP motifs in WW domains cannot be known *a priori* and thus there is considerable interest in a three-dimensional structure of a WW:PPΨΨP complex. To illustrate the interaction between the Prp40 WW2 domain and the SH3 binding peptide APTMPPPLPP, their complex structure was modelled based on the dystrophin WW:RSPPPYVP complex (Huang *et al.*, 2000) to imitate the orientation of PPxY motifs and based on the Pin1 WW:YpSPTpSPS complex (Verdecia *et al.*, 2000) to imitate the reverse peptide orientation (Fig. 6.4).

In the C–N orientation, which is equivalent to the orientation of PPxY peptide motifs, the peptide prolines would occupy the same positions as in the dystrophin complex and the Leu (or Ile in the case of the PPLIPPPP peptide) would point into cavity two and, hence, be responsible for the binding specificity (Fig. 6.4, green ligand). Interestingly, in the reverse peptide orientation, two peptide prolines (P6' and P9', yellow ligand in Fig. 6.4(a)) occupy equivalent positions as compared to the Pin1 WW:YpSPTpSPS complex (P3' and P6') (Fig. 6.5), while the aliphatic peptide residue (L8') points into the first binding pocket instead

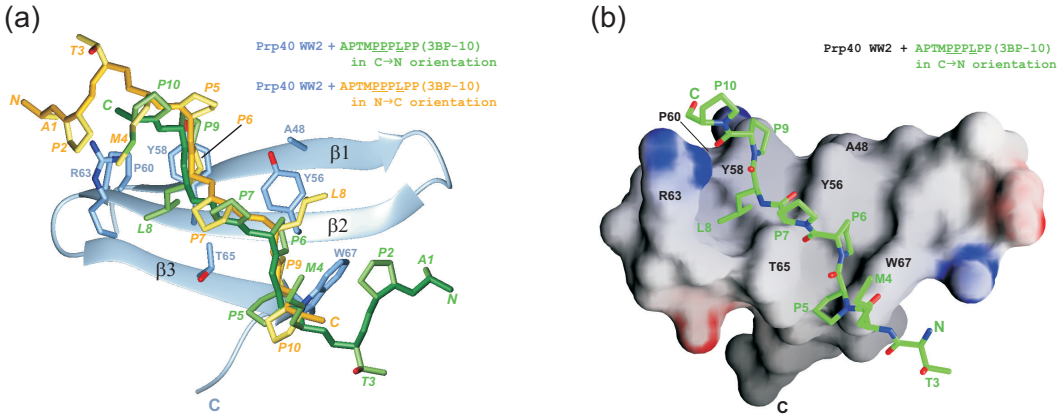

Figure 6.4: Models of the Prp40 WW2:APTMPPPLPP complex. (a) Model of the complexbetween the Prp40 WW2 (shown in blue) with the 3BP-10 peptide APTMPPPLPP (shown in green) orientated as propopsed for PPxY motifs, while the reverse peptide orientation is depicted in yellow. (b) Surface presentation (Nicholls *et al.*, 1993) of the modelled Prp40 WW2:APTMPPPLPP complex. Blue and red surface colours indicate positive and negative electrostatic surface potential, respectively.

of the second binding pocket. However, to fully understand the molecular mechanisms underlying the interactions of WW domains with PPΨΨP motifs three-dimensional complex structures will be necessary.

The novel interactions observed between the PPΨΨP motifs and the Prp40 WW domains pointed us to perform a search in the Saccharomyces Genome Database (http://genome-www.stanford.edu/Saccharomyces) for sequences containing this pattern. Although the exact sequences of the PPLIPPPP and APTMPPPLPP peptides used in this study are not present in *Sc.*, we found 20 potential interaction partners with PPΨΨP motifs. Among these proteins, we found four nuclear proteins (KRH2, CST6, RGM1 and the RNA Polymerase II transcription mediator MED7) and six yeast formins and cytoskeleton organising proteins (BNI1, SRV2, YNL094W, BNR1, PAN1 and VRP1). Intriguingly, WW domains orthologous to those of Prp40 have originally been identified as forming binding proteins (Chan *et al.*, 1996). This suggests that the Prp40 WW domains may have additional interaction partners besides the proteins BBP and Prp8 identified so far.

6.4 Interaction with CTD RNA polymerase II peptides

Already during transcription various pre-mRNA processing reactions are coordinated by the association of elongation factors, capping enzymes and splicing factors to the C-terminal domain (CTD) of the largest subunit of RNA polymerase II (Maniatis & Tasic, 2002). The CTD encompasses a large number of YSPTSPS repeats (Corden, 1990), which are reversibly phosphorylatable at serine positions two and five (Dahmus, 1996). Interestingly, several proteins containing WW domains have been identified as CTD binding partners. The human/*Sc.*

prolyl-isomerase Pin1/Ess1 WW domain as well as the second (WW2) and third (WW3) WW domain of the human/*Sc.* ubiquitin-protein ligase Nedd4/Rsp5 (Chang *et al.*, 2000; Gavva *et al.*, 1997; Wu *et al.*, 2001) interact with both, phosphorylated and unphosphorylated, CTD repeats, respectively (Verdecia *et al.*, 2000; Morris *et al.*, 1999), whereas Prp40 has been shown to interact only with hyperphosphorylated CTD repeats (Morris & Greenleaf, 2000). To test whether the region containing the WW pair in Prp40 is responsible for the interaction with the phosphorylated CTD tail, chemical shift mapping experiments were performed with the Prp40 WW domain pair and a phosphorylated as well as an unphosphorylated CTD peptide (SYpSPTpSPS and YSPTSPSYSPTSPS, respectively).

However, we found that the tandem WW domains of Prp40 interacted neither with the phosphorylated nor with the unphosphorylated CTD repeats (Fig. 6.1(c)). On the contrary and as previously described (Chang *et al.*, 2000; Wu *et al.*, 2001), the control experiment with the ^{15}N-labelled *Sc.* Rsp5 WW2 revealed binding for both peptides, the phosphorylated and unphosphorylated CTD repeats (Fig. 6.1(d)). As observed in previous studies the interaction between the Rsp5 WW2 and the unphosphorylated CTD repeats was stronger than with the phosphorylated CTD peptide as indicated by the larger chemical shift changes produced by the unphosphorylated CTD peptide (Morris & Greenleaf, 2000). In another control experiment, the Ess1 WW domain also interacted, as expected (Morris *et al.*, 1999), with the phosphorylated CTD repeat (data not shown) confirming that the lack of interactions between the Prp40 WW domains and the CTD repeats used in this study is not due to experimental artifacts.

To elucidate the differences in the CTD binding found for the Prp40, Rsp5 and Pin1 WW domains, their domain sequences were compared and the structure of the second Rsp5 WW domain was modelled based on the Pin1 WW:YpSPTpSPS complex (Verdecia *et al.*, 2000) to allow a structural comparison. As can be seen in Fig. 6.5(a) and (b), two key electrostatic interactions occur between the Pin1 WW domain and the phosphorylated serines of the CTD peptide YpSPTpSPS. An Arg (R18) in the middle of β1 interacts with the phosphate group of Ser2', while another Arg in loop1 (R21) forms an extensive hydrogen bond network to the phosphorylated Ser5'. The latter interaction, however, is precluded in the Rsp5 WW2, since it lacks an equivalent Arg in the sequence. On the other hand, the Pin1 Arg in β1 is conserved in the Rsp5 WW2 allowing for the recognition of a phosphorylated and an unphosphorylated Ser2 in the CTD peptides (Fig. 6.5 (b) and (c)). In contrast, the Prp40 WW domains lack both critical arginines. While the Arg in loop1 of the Pin1 WW is deleted in both Prp40 WW domains, an alanine (A7/48) occupies the position of the Pin1/Rsp5 Arg in β1 (Fig. 6.3 (a)).

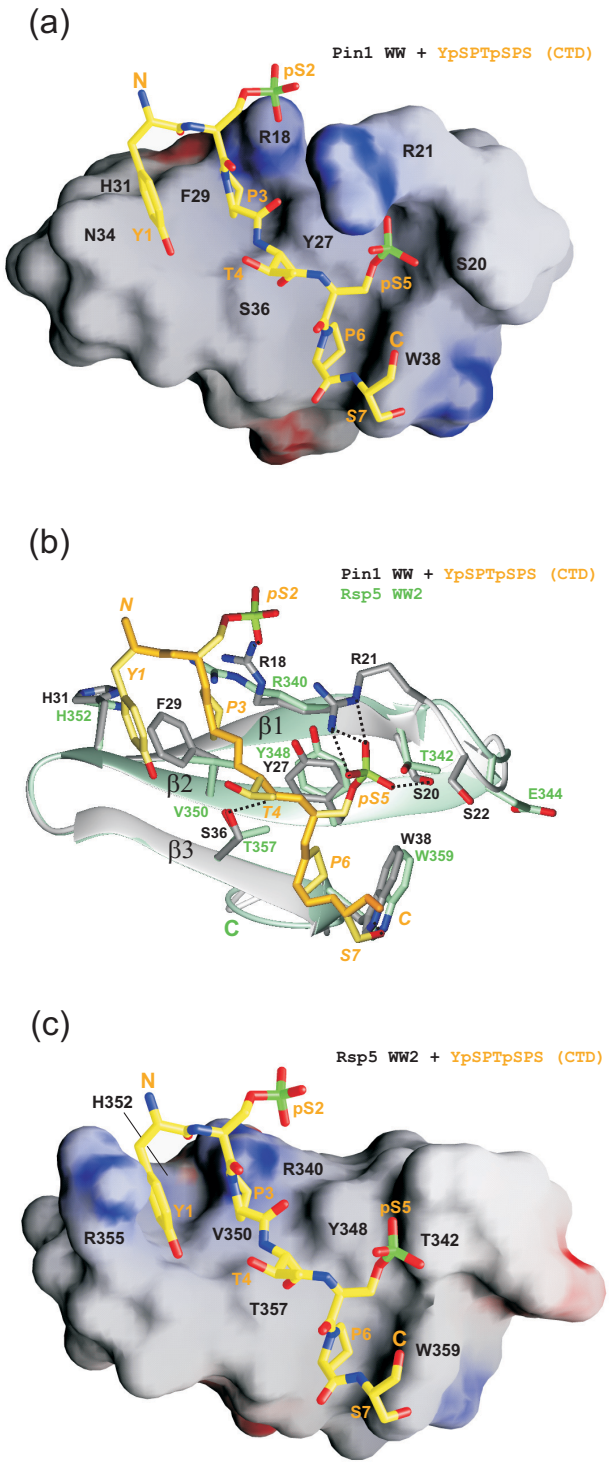

Figure 6.5: Comparison of the interaction between the phosphorylated CTD repeat and the Pin1 WW domain and the second Rsp5 WW domain. Hydrogen bonds are illustrated by dashed lines. In panel (a) and (c), blue and red surface colours indicate positive and negative electrostatic surface potential, respectively. (a) Surface representation of the Pin1 WW:YpSPTpSPS complex (Verdecia *et al.*, 2000). (b) Superposition of the Pin1 WW domain (shiwn in grey) and the modelled second Rsp5 WW domain (shown in green) in complex with the phosphorylated CTD repeat (depicted in yellow). (c) Surface representation of the modelled Rsp5 WW2:YpSPTpSPS complex.

Consequently, the formation of both aforementioned crucial hydrogen bond networks to the CTD is precluded in the Prp40 WW domains. Moreover, the Prp40 WW1 encompasses a negatively charged residue (D9) at the end of β1 pointing towards the phospho-Ser5' binding site in the Pin1 WW:YpSPTpSPS complex making an interaction with the phosphorylated CTD repeat unfavourable. Taken together, based on the residues constituting the two binding sites of the Prp40 WW domains an interaction with the CTD repeats seems unlikely. However, it cannot be excluded that the Prp40 WW domains recognise other phosphorylation patterns in the CTD or sequences contained in the four imperfect repeats of yeast CTD. It could also be that the presence of other Prp40 domains is necessary for an interaction of the Prp40 WW domain pair with the CTD repeats used here. Nevertheless, our findings add important details for the involvement of Prp40 in linking transcription and splicing.

6.5 Concluding remarks

While WW domains were at first found to interact with well-defined PPxY motifs (Sudol, 1996), the discovery of other binding motifs such as the PPΨΨP (Chan *et al.*, 1996) and CTD motifs (Lu *et al.*, 1999) has challenged the view of a single consensus binding motif and ligand orientation for WW domains. CTD motifs encompass less prolines than PPxY motifs, and surprisingly they were found to bind in an orientation that is reverse with respect to the orientation of PPxY motifs in the WW domain binding sites (Macias *et al.*, 1996; Verdecia *et al.*, 2000). The binding mode of PPΨΨP motifs to WW domains can be expected to be even more complex, since the symmetry of the poly-proline helix allows these motifs to bind in both orientation, *i.e.* an orientation that resembles PPxY motifs and an orientation that resembles phospho-SerPro motifs (Macias, unpublished results). This intricate scenario underlines the need of three-dimensional structures of WW domains in complex with different binding motifs to fully understand the ligand recognition and specificity of different WW domains in general and within the different groups of WW domains in particular.

In summary, the data presented in this Chapter provide evidence that both Prp40 WW domains have the same ligand preferences and recognise both PPΨΨP and PPxY motifs. The observed interactions between the Prp40 WW domains and PPxY motif peptides derived from BBP and Prp8, respectively, supports the view that the Prp40 WW domains function in cross-intron bridging by interacting with BBP and Prp8. Although the observed chemical shift changes cannot be directly related to binding affinities, they are indicative of weak binding in the cases of the PPxY motifs. Given that the residues in the second binding pocket of the Prp40 WW domains might be sub-optimal for interactions with PPxY motifs, these weak interactions could, however, be important to keep a large complex as the spliceosome sufficiently dynamic and facilitate rapid responses to external stimuli. Accordingly, many other protein-interaction domains in multi-protein complexes have been shown to

bind proline-rich ligands with remarkably low affinity, yet nevertheless high specificity (Ball *et al.*, 2002).

Furthermore, in this study the WW domains of the yeast splicing factor Prp40 have been shown for the first time to interact with PPΨΨP motif containing peptides, which are known to bind to the WW domains of the mammalian Prp40 orthologue FBP11 (Bedford *et al.*, 1997). The Prp40 WW domains can therefore now be classified as group II WW domains not only based on sequence similarity, but also based on their ligand binding preferences and their potential to compete with SH3 domains for the same targets. A database search for *Sc.* proteins encompassing PPΨΨP motifs yielded 20 matching proteins, including formins and cytoskeleton organising proteins. This suggests that not all interaction partners of the Prp40 WW domains may have been identified yet.

Interestingly, no interaction was observed between the Prp40 WW domains and the CTD repeats used in this study, although positive control experiments with the Rsp5 WW2 and the Ess1 WW domain revealed binding. A detailed comparison of the binding sites of the Prp40 WW domains with those of Pin1 and Rsp5 shows that the two crucial hydrogen bond networks to the phospho-CTD cannot be formed in the Prp40 WW domains. This makes an interaction between the Prp40 WW domains and phosho-CTD sequences unfavourable and hence unlikely. However, it cannot be excluded that the Prp40 WW domains recognise other phosphorylation patterns in the CTD or other sequences as those contained in the four imperfect repeats of the yeast CTD. On the other hand, the previously suggested function of Prp40 in linking transcription and splicing may not be mediated by the WW domain pair directly, but might require the presence of other Prp40 domains or auxiliary proteins.

6.6 References

Aasland, R., Abrams, C., Ampe, C., Ball, L., Bedford, M. T., Cesareni, G., Gimona, M., Hurley, J. H., Jarchau, T., Lehto, V.-P., Lemmon, M. A., Linding, R., Mayer, B. J., Nagai, M., Sudol, M., Walter, U. & Winder, S. J. (2002). Normalization of nomenclature for peptide motifs as ligands of modular protein domains. *FEBS Letters,* **513**, 141–144.

Abovich, N. & Rosbach, M. (1997). Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell,* **89**, 403–412.

Ball, L. J., Jarchau, T., Oschkinat, H. & Walter, U. (2002). EVH1 domains: Structure, function and interactions. *FEBS Letters,* **513**, 45–52.

Bedford, M. T., Chan, D. C. & Leder, P. (1997). FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J.* **16**, 2376–2383.

Chan, D. C., Bedford, M. T. & Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J.* **15**, 1045–54.

Chang, A., Cheang, S., Espanel, X. & Sudol, M. (2000). Rsp5 WW domains interact directly with the carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **275**, 20562–71.

Corden, J. L. (1990). Tails of RNA polymerase II. *Trends Biochem. Sci.* **15**, 383–7.

Dahmus, M. (1996). Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Bio. Chem.* **271**, 19009–12.

Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science,* **266** (5188), 1241–7.

Gavva, N. R., Gavva, R., Ermekova, K., Sudol, M. & Shen, C. J. (1997). Interaction of WW domains with hematopoetic transcription factor p45/NF-E2 and RNA polymerase II. *J. Biol. Chem.* **272**, 24105–8.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M. & Eck, M. (2000). Stucture of a WW domain containing fragment of dystrophin in complex with b-dystroglycan. *Nat. Struct. Biol.* **7**, 634–38.

Kanelis, V., Rotin, D. & Forman-Kay, J. D. (2001). Solution structure of a Nedd4 WW domain-ENaC peptide complex. *Nat. Struct. Biol.* **8** (5), 407–12.

Kay, B. K., Williamson, M. P. & Sudol, M. (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231–241.

Lu, P. J., Zhou, X. Z., Shen, M. S. & Lu, K. P. (1999). Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science,* **283**, 1325–1328.

Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M. & Oschkinat, H. (1996). Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature,* **382**, 646–649.

Macias, M. J., Wiesner, S. & Sudol, M. (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Letters,* **513**, 30–37.

Maniatis, T. & Tasic, B. (2002). Alternative pre-mRNA splicing and the proteome expansion in metazoans. *Nature,* **418**, 236–43.

Maroney, P. A., Romfo, C. M. & Nilsen, T. W. (2000). Functional recognition of the 5' splice site by the U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol. Cell,* **6**, 317–328.

Morris, D. P. & Greenleaf, A. L. (2000). The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **275** (51), 39935–43.

Morris, D. P., Phatnani, H. P. & Greenleaf, A. L. (1999). Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-end formation. *J. Biol. Chem.* **274**, 31583–7.

Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: Graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166–170.

Pires, J. R., Taha-Nejad, F., Toepert, F., Ast, T., Hoffmuller, U., Schneider-Mergener, J., Kuhne, R., Macias, M. J. & Oschkinat, H. (2001). Solution structures of the YAP65 WW domain and the variant L30K in complex with the peptides GTPPPPYTVG, N-(n-octyl)-GPPPY and PLPPY and the application of peptide libraries reveal a minimal binding epitope. *J. Mol. Biol.* **314** (5), 1147–56.

Proudfoot, N. J., Furger, A. & Dye, M. J. (2002). Integrating mRNA processing with transcription. *Cell,* **108**, 501–512.

Sudol, M. (1996). Structure and function of the WW domain. *Prog. Biophys. Molec. Biol.* **65**, 113–32.

Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T. & Noel, J. P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat. Struct. Biol.* **7**, 639–643.

Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: A program to generate schematic diagrams of portein-ligand interactions. *Prot. Eng.* **8**, 127–34.

Wu, X., Chang, A., Sudol, M. & Hanes, S. D. (2001). Genetic interactions between the ESS1 prolyl-isomerase and the RSP5 ubiquitin ligase reveal opposing effects on RNA polymerase II function. *Curr. Genet.* **40**, 234–42.