Inter-molecular interactions in E3 Ubiquitin ligases

Smad proteins function as intracellular signalling effectors for the TGF-β superfamily of secreted polypeptides. As a result of their capacity to bind DNA and to induce transcriptional responses through interactions with other transcription factors, Smads act as transcription factors. Smads are modular proteins containing two conserved MH1 and MH2 domains, but they differ in the linker sequence connecting these domains. Accumulating evidence suggests that E3 ubiquitin ligases are critical regulators of transcription factors and growth factor receptors. Ubiquitination occurs through a three-step process involving

![Image of protein NMR spectroscopy](image_url)

**Figure 1.** Comparison of RhoGAPFF1 with other FF domains. (a) Superimposition of the typical $3_{10}$ helix (in red) of the FBP11FF1 domain (in dark blue) and the extended α helix (in yellow) that replaces the $3_{10}$ helix in the RhoGAPFF1 domain (in sky blue). The residues that form part of this extended α helix are marked in a yellow box in the alignment in (b). (b) Sequence alignment of FF domains from human p190-A RhoGAP, CA150, and FBP11. The alignment was generated with ClustalX and edited manually.
ubiquitin-activating (E1), ubiquitin-conjugating (E2), and ubiquitin ligase (E3) enzymes. Only the HECT domain-containing E3 ligases recognise and also directly catalyse the transfer of ubiquitin to the substrate. Thus, much research effort has focused on the identification of motifs in the targets that are recognised by the ligases or on studying whether a given ligase recognise one or more targets. In the case of ubiquitin ligases belonging to the Nedd4 family and the Smads, the interaction between the ligase and the target is driven by contacts from the WW domain of the ligase towards the PY motif of the Smad proteins. However, additional contacts from other regions in the proteins or from auxiliary proteins forming transient complexes may control the specificity of the interactions. We seek to elucidate the factors involved in controlling this specificity. However, in order to fully understand how the interactions occur at a molecular level, the actors needed and their order, a combined approach of in vivo experiments with detailed structural work is required.

In this context, Joan Massagué and co-workers identified Nedd4L as the ubiquitin ligase responsible for the polyubiquitination of Smad2/3. They also mapped the interaction domains of Nedd4L and Smad3 using a series of expression vectors encoding several fragments of Nedd4L and Smad3. When expressed in HEK293T cells, the second WW domain (WW2) of Nedd4L bound to the Smad3 linker region, whereas the other three WW domains, the C2 domain, and the HECT domain did not. Mutation of the PY motif (PPGY to AAGY) abolished this interaction, as did mutation of the four-linker phosphorylation sites in Smad3. Furthermore, using Smad3 constructs with individual mutations in these phosphorylation sites or with mutation of all these sites but one, they determined that T179 is the only phosphorylation site required for the Smad3-Nedd4L interaction. T179 (T220 in Smad2) lies directly upstream of the PY motif, suggesting that the WW2 domain of Nedd4L specifically recognises a phosphothreonine-PY (pT-PY) motif in Smad2/3.

On the basis of this information, and in order to quantify the interactions shown to occur in vivo, we measured the affinity of the four individual Nedd4L WW domains for synthetic peptides comprising 13 amino acids, containing the T-PY motif of Smad2 or Smad3 with either a threonine or a phosphothreonine residue. Isothermal titration calorimetry analysis revealed high affinity of the WW2 domain for the pT-PY motif.
peptides ($K_d = 7.8 \mu M$ and $4.1 \mu M$, respectively). This affinity is among the highest reported to date for a WW-PY domain interaction. The affinity of WW2 for the unphosphorylated T-PY motif was 7-15-fold lower. The Nedd4L WW3 domain also preferentially bound to the phosphorylated T-PY motifs, but with lower affinity than the WW2 domain. The WW1 and WW4 domains bound even more weakly and with no preference for the phosphorylated T-PY motifs. Interestingly, Smad1 (and Smad5) also contains a conserved T-PY motif. However, this threonine residue was not phosphorylated in vivo under any of the agonist or antagonist stimuli tested and it was poorly phosphorylated by CDK8/9 in vitro. The Smurf1 WW2 domain binds a synthetic peptide of 13 residues including the T-PY motif of Smad1 with a $K_d$ of 32 $\mu M$ and the phosphorylated Smad3 pT-PY motif with a $K_d$ of 36 $\mu M$. These values are in agreement with the observation that Smurf1 plays a minor role in Smad3 turnover and it requires contacts with the phosphorylated SerPro cluster in order to target Smad1. A detailed analysis of these interactions, including the determination of the different complexes between Nedd4L WW domains and PY motifs of Smads, is also underway.

Unravelling the protein-protein interaction scenario during splicing: implication of FF domains

FF domains are protein-protein interaction modules of about 70 amino acids and are often found in several copies. They are present in three protein families: the splicing factors FBP11, Prp40 and URN1, the transcription factors CA150, and the p190RhoGTPase-related proteins. However, the simplicity of their distribution contrasts with the difficulty to define their biological roles. Indeed, for each FF domain studied, there appears to be a ligand that does not contain conserved features when compared to others previously characterised.

**p190RhoGAP FF1 domain shields its phosphorylation site in the domain core**

p190-A and -B Rho GAPs (guanosine triphosphatase activating proteins) are the only cytoplasmatic proteins containing FF domains. In p190-A Rho GAP, the region containing the FF domains has been implicated in binding to the transcription factor TFII-I. Moreover, phosphorylation of Tyr308 within the first FF domain inhibits this interaction. Because the structural determinants governing this mechanism were unknown, we sought to solve the structure of the first FF domain of p190-A Rho GAP (RhoGAPFF1) and to study the potential impact of phosphorylation on the structure. We found that RhoGAPFF1 does not fold with the typical $\alpha_1-\alpha_2-3\alpha(10)-\alpha_3$ arrangement of other FF domains. In addition, we observed that specific contacts between residues in the first loop and the fourth helix are indispensable for the correct folding and stability of this domain.

The structure also revealed that Tyr308 contributes to the domain hydrophobic core. Furthermore, the residues that compose the target motif of the platelet-derived growth factor receptor alpha kinase form part of the alpha 3 helix. We observed that the phosphorylation reaction requires a previous step including domain unfolding, a process that occurs at 310 K. The unfolding capacity of this FF domain was not observed in other domains used as controls. Furthermore, in the absence of phosphoryla-
tion, the temperature-dependent RhoGAPFF1 folding/unfolding process was reversible. However, phosphorylation caused an irreversible destabilisation of the RhoGAPFF1 structure, which probably accounts for the inhibitory effect that it exerts on the TFII-I interaction. Our results link the capacity of a protein domain to be phosphorylated with conformational changes in its three-dimensional structure.

Phosphorylation sites in proteins are often found in the loops or in linkers connecting domains. Instead, one of the phosphorylation sites of RhoGAP forms part of the protein core, and it is inaccessible to kinases in the folded state. Some proteins may have developed a double-check system, shielding phosphorylation sites in the protein core. Thus, the inhibitory role that phosphorylation plays in this regulatory process appears to cover the energetic cost of shifting the folding-unfolding equilibrium towards the unfolded state in a non-reversible manner. The removal of the phosphate group allows the system to recover the basal structure.

**Solution structure of the fourth FF domain of yeast Prp40 splicing factor**

Prp40 protein was originally identified as a suppressor of 5’ end U1 RNA point mutations. [1] Prp40 is a U1 snRNP-associated protein that participates in the early steps of yeast pre-messenger RNA splicing. Prp40 associates with the branch-binding point protein to bring the 5’ end splicing site and the intron branch point into spatial proximity. In addition, Prp40 has been implicated in the binding to the phosphorylated C-terminal domain (herein referred as phospho-CTD) of RNA polymerase II through regions involving the WW and FF domains. However, a subsequent study on the structure of the Prp40 WW domain pair also showed that, in the absence of additional FF domains, the WW domains do not interact with the phospho-CTD repeats.

We determined the solution structure of the first FF domain of Prp40 in 2006. That study also examined the binding of Prp40FF1 to the splicing factor Clf1 and to a pair of bisphospho-CTD repeats. The binding site for the association with the first TPR motif of Clf1 involves helices 2 and 3, and the N-terminal half of helix 3. In contrast, no interaction was detected for the Prp40FF1 domain with the phospho-CTD repeats and for the Prp40FF4 domain with the TPR motif of Clf1.

In this study we report the solution structure of the Prp40FF4 domain. Furthermore, prompted by the observation that the charge distribution of the FBP11FF1 region involved in the interaction with the bisphospho-CTD repeats is partially conserved in Prp40FF4, we also examined whether this domain interacted with the phospho-CTD repeats; however, no binding was detected under our experimental conditions.

---

**Scientific output**

**Publications**


**Research networks and grants**

Spanish Ministry of Science and Innovation, BFU2008-02795 (2009-2011)

Principal investigator: Maria Macias

**Collaborations**

*Inter- and intra-molecular interactions in E3 ubiquitin ligases: Recognition of new proline-rich motifs by WW domains and its implications in protein degradation and transcription*

Joan Massague, Memorial Sloan-Kettering Cancer Center (New York, USA)

*Structural studies of cell-penetrating peptides (γ-peptides) based on proline derivatives and of somatostatin analogues*

Antoni Riera, IRB Barcelona (Barcelona, Spain); Miriam Royo, Barcelona Science Park (Barcelona, Spain)

*Unravelling the protein-protein and protein-RNA scenario during splicing: Exon skipping*

Juan Valcárcel, Center for Genomic Regulation (Barcelona, Spain)