Interactions of Smads and WW domains of proteins involved in TGF-ß signalling

Dissertation zur Erlangung des akademischen Grades des

Doktors der Naturwissenschaften (Dr. rer. nat)

Eingereicht im Fachbereich Biologie, Chemie, Pharmazie

der Freien Universität Berlin

vorgelegt von

Nina Görner

aus Mülheim an der Ruhr

2011

This work has been carried out at the department of structural biology under the supervision of Dr. Maria J. Macias from January 2007 to May 2011.

Parts of this thesis have been published or will be published in the near future:

Goerner N, Aragon E, Gao S, Massagué J, Macias M: Nedd4L and Smurf2 E3 ubiquitin-ligases use different binding modes to recognize Smad7. *In preparation.*

Aragon E, Goerner N, Zaromytidou, Xi Q, Escoberdo A, Massagué J, Macias M: A Smad action turnover switch operated by WW domain readers of a phosphoserine code. *Genes Dev* 2011, **25:** 1275-1288. http://dx.doi.org/10.1016/j.molcel.2009.09.043

Gao S, Alarcon C, Sapkota G, Rahman S, Chen PY, Goerner N, Macias MJ, Erdjument-Bromage H, Tempst P, Massague J: Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF-ß signaling. *Mol Cell* 2009, **36**:457-468.
http://dx.doi.org/10.1101/gad.2060811

1. Gutachter: Prof. Dr. Hartmut Oschkinat, Forschungsinstitut für Molekulare Pharmakologie Campus Berlin-Buch, Rober-Rössle-Str. 10, 13125 Berlin, Germany.

2. Gutachter: Dr. Maria J. Macias, Institut de Recerca Biomèdica, Parc Científic de Barcelona, Calle Baldiri Reixac 10, 08028 Barcelona, Spain.

Disputation am 15.12.2011

Acknowledgements

This work would not have been possible to complete without the advice and support of many people.

First of all I would like to thank my supervisor Dr. Maria Macias for the possibility to work in her group of biomolecular NMR spectroscopy at the IRB. I wish to thank her for her encouragement, guidance and support throughout my thesis and I am particularly grateful for the opportunity to work in the joint project with Joan Massague's group.

Further I would like to thank Professor Dr. Hartmut Oschkinat for his support as external university supervisor of my thesis.

I am indebted to my colleagues for the stimulating and fun environment in the laboratory. I am especially grateful to Lida Ruiz for introducing me to molecular biology and to Eric Aragon for his work on cloning, protein expression and acquiring and processing NMR spectra. Pau Martin deserves special mention for his help with the structure calculations and with all kind of computational problems or questions concerning the technical aspects of biomolecular NMR. I further would like to acknowledge Albert Escobedo and Tiago Gomes for their help on peptide synthesis and my colleagues Dr. Begoña Morales, Dr. Roman Bonet and Constanze Braasch who accompanied parts of my Ph.D period.

I wish to thank Dr. Miriam Royo for her support in peptide synthesis and for being a member of my thesis advisory committee. Dr. Rafel Prohens for helping me with the ITC experiments and for valuable discussions and Dr. Irene Fernández for her kind assistance in mass spectrometry and her wise advice.

I further wish to thank the IRB for financial support by rewarding me with the PhD fellowship as well as for the good memories of the department retreats and the many cool off sessions that were organized. My time at the IRB was made enjoyable in large part due to the many friends and groups that became a part of my life. Particularly I would like to thank Maximilian

Becker and my friend Sabine Klischies for the emotional support, entertainment and caring she provided.

Lastly I wish to thank my family, my sisters Birgit and Karin and most importantly my parents for their constant support and love.

To them I dedicate this thesis.

Summary

Smad proteins are the mediators of the TGF-ß signalling pathway that controls cell growth and specific cell differentiations. Defects in this pathway can lead to rapid cell growth resulting in cancer diseases. In response to TGF-ß signalling, Smad2 and 3 are activated and transferred with the help of Smad4 to the nucleus, where they regulate the transcription of target genes. The inhibitor Smad7 can negatively regulate the TGF-ß pathway by triggering TGF-ß receptor I degradation. Smad proteins interact with other proteins through their linker region, which connects the DNA binding domain named MH1 with a TGF-ß-receptor-binding domain known as MH2 domain. The Smad2/3 linker region consists of several (S/T)P motifs that can be phosphorylated by different kinases. The Smad7 linker region is not phosphorylated but contains as Smad2 and 3 a PPxY (PY) motif that is recognized by WW domains. WW domains are small binding domains present in around 170 proteins of variable function involved in many cellular processes including transcription, splicing or protein degradation among others.

In the main part of this work the interactions of R-Smad3 with the proteins Nedd4L and Pin1 were studied to understand the role of the Smad3 linker phosphorylation and its influence on WW domain binding. In the second part the interactions of the inhibitor Smad7 and the proteins Nedd4L and TAZ were characterized to obtain more information about their role in TGF-ß inhibition. In both parts the aim was to understand how specificity can be achieved by the WW domains. The structures of six WW domain - Smad complexes were solved and similarities and differences were pointed out.

At the start the interaction of Nedd4L and the phosphorylated Smad2/3 linker, which leads to Smad2/3 degradation, has been studied. Using isothermal titration calorimetry (ITC) and nuclear magnetic resonance spectroscopy (NMR) it could be shown that of all four Nedd4L WW domains the second domain binds to the phosphorylated Smad2/3 pT[PY] motif with the highest affinity. The NMR structure of the Nedd4L WW2 – Smad3 pT[PY] complex revealed that the PPPxY site of the ligand is bound with all proline residues in *trans* configuration and the tyrosine residue accommodated in the second loop of the WW domain.

The novelty of the complex resides in the binding of the phosphate group, which is specifically recognized by amino acids K378 and R380 in the first loop region. This new interaction has been corroborated by Nedd4L WW2 point mutations and by ^{31}P-NMR titration experiments.

Furthermore, the binding of the phosphorylated Smad3 linker to the WW domain of the peptidyl-prolyl isomerase Pin1 has been studied. Pin1 catalyses the *cis-trans* isomerisation of prolines specifically preceded by phosphorylated serines or threonines. The K_D values obtained by ITC binding studies showed that the Pin1 WW domain strongly interacts with the Smad3 pT179P motif while bindings to the bi-phosphorylated Smad3 pS204P-pS208P and pS208P-pS213P peptides are weaker or barely observable as in the case of the monophosphorylated Smad3 pS204P peptide. To understand these differences in the binding affinities the structure of the Pin1 WW - Smad3 pT179P complex and of the Pin1 WW - Smad3 pS208P-pS213P complex was solved by solution NMR. It could be shown that in both complexes only one p(S/T)P motif is bound, also in *trans*-configuration, by the Pin1 WW domain involving amino acids of the first ß-stand and that the affinity is enhanced through the presence of a negatively charged amino acid at position x-1 and of a proline at position x-2 of the p(S/T)P motif. The influence of the phosphate group on binding the Pin1 WW domain was pointed out by ^{31}P titrations. When compared to the Nedd4L WW2 recognition of the Smad3 pT179[PY] motif the Pin1 WW - Smad3 pT179P complex reveals a different binding mode. On the one hand the phosphate group is bound at different locations within the domains on the other hand the peptide is bound from the N- to C-terminus by the Pin1 WW domain and in opposite direction from the C- to N-terminus by the Nedd4L WW2 domain. The different modes of Smad3 recognition may reflect the different outcomes of these two interactions.

ITC measurements of Nedd4L WW2 and Smad3 pT179[PY] as well as of Pin1 WW and Smad3 pS208P-pS213P at different proton concentrations revealed that in both cases the binding affinity increases with the pH values. The highest affinity was observed at pH 8.0. This information was used to optimize the conditions for the NMR experiments, but could also play a role in the regulation of protein binding in the physiological environment of the cell.

In order to understand the role of the other WW domains present in Nedd4L, Smad3 fragments including the PY motif and other pSP sites were titrated with the three possible

Nedd4L WW pairs. Of the three consecutive pairs present in Nedd4L, only the WW2-WW3 pair showed high affinity for the extended Smad3 sequence. However, the determination of the structure of the Nedd4L WW2-WW3 module in solution bound to the Smad3 linker peptide pT[PY]-pS-pS (Smad3 176-211 segment tri-phosphorylated at T179, S204 and S208) posed a challenge owing to the complexity introduced by the poor NMR signal dispersion of the 80 amino acid region connecting WW2 and WW3 domains. To simplify the interpretation of the NMR data, the WW2-WW3 module was prepared as two separate fragments, which allowed the use of sequential isotope labelling. A disulfide-bond protein ligation strategy was applied to connect these two fragments. Using a stepwise assignment approach the complex structure with the Smad3 176-211 segment was studied in detail.

In the second part of this work the question was raised how inhibitor Smad7 is recognized by the WW domain containing proteins Nedd4L and TAZ. Smad7 recruits Nedd4L for TGF-ß receptor I degradation in the cytoplasm while the interaction with TAZ mainly occurs in the nucleus. Through ITC binding studies of the individual Nedd4L WW domains and a 15 amino acid Smad7 linker peptide containing the PY motif it could be shown that the second WW domain is the preferred Smad7 binding domain. Solution NMR based structure determination of the Nedd4L WW2 domain bound to the Smad7[PY] peptide revealed that the high affinity is due to additional contacts of the C-terminal part of Smad7 and the ß1 sheet region of the Nedd4L WW2 domain and electrostatic interactions of the Smad7 N-terminus and the loop 1 region of the domain. Further it could be shown by ITC experiments that this affinity is not enhanced by the presence of additional Nedd4L WW domains.

The co-factors YAP and its paralogue TAZ consist of two (YAP) and one (TAZ) WW domain(s) and have also been reported to interact with Smads. However, the interaction between YAP and Smad7 results in enhanced TGF-ß receptor degradation whereas binding of TAZ and Smad7 has not been studied yet. The sequence alignment of the WW domains of TAZ and YAP reveals a higher sequence identity between YAP WW1 and TAZ WW than between the YAP WW1 and WW2 domains. ITC experiments carried out in this thesis revealed that the first WW domain in YAP binds the 15 amino acid Smad7 peptide with a higher affinity than the second WW domain and equally well as TAZ WW. Structures of the YAP WW1 domain bound to PY motifs have been previously described but in this work the first structure for the WW domain of the human TAZ bound to a PY ligand was determined.

The TAZ WW - Smad7[PY] complex is very similar to its Nedd4L WW2 counterpart in the PY motif-containing region and in the C-terminal extension, but differ in the way in which the Smad7 N-terminus is recognized.

As a conclusion, the findings presented in this work expand the known structural and functional versatility of WW domains as protein-protein interaction modules. Further this work shows that the interaction of Nedd4L and Pin1 with Smad3 is regulated by linker phosphorylation up- and downstream of the PY motif, which is carried out by special kinases in a precise order. Nedd4L binds to the Smad3 pT[PY] motif and two additional phosphorylation sites within Smad3 linker. The Pin1 WW domain only requires the Smad3 pTP motif with a single phosphorylation site for binding, which probably acts as the anchoring place for the prolyl-isomerase protein. Thus, Pin1 WW may be able to participate in the regulation of Smad3 as an active player of the Smad3 and Nedd4L interaction. Further, Nedd4L achieves high specificity by binding Smad3 following a two-step cooperative mode that involves the two vicinal domains WW2 and WW3. This ability of WW domains to recognize both pT[PY] and pS-pS motifs was previously unknown.

The interaction of Nedd4L and Smad7 results in inhibition of the TGF-ß pathway through TGF-ß receptor I degradation. In this work it could be demonstrated that the single Nedd4L WW2 domain is sufficient to interact with the Smad7[PY] motif with high affinity. The individual WW domain of the adaptor protein TAZ also interacts with the inhibitor Smad but slightly weaker. Both complexes are very similar in the region containing the PY motif and also the C-terminal extension but they differ in the recognition of the N-term site by the domain. These differences may reflect the distinct functional outcomes of the two interactions. TAZ was previously known as a TGF-ß signalling agonist but in this work it is shown that TAZ also can play an inhibitory role in TGF-ß signalling as reported for its paralogue YAP.

Zusammenfassung

Smad Proteine sind Signalüberträger innerhalb des TGF-ß Signalweges, welcher das Zellwachstum und die Zelldifferenzierung kontrolliert. Fehler innerhalb dieses Signalweges können zu schnellem Zellwachstum und damit zu Krebs führen. Durch die TGF-ß Signale werden Smad2 and 3 aktiviert und mit Hilfe von Smad4 in den Zellkern transportiert, wo sie die Gentranskription bestimmter Gene regulieren. Inhibitor-Smad7 kann durch den Abbau des TGF-ß Rezeptors I den TGF-ß Signalweg negativ beeinflussen. Smad Proteine wechselwirken mit anderen Proteinen über eine *Linker*-Region, welche die DNA-bindende MH1 Domäne mit der TGF-ß Rezeptor-bindenden MH2 Domäne verbindet. Innerhalb des Smad2/3-Linkers befinden sich mehrere (S/T)P Motive, welche von verschiedenen Kinasen phosphoryliert werden können. Der Smad7-Linker wird nicht phosphoryliert, enthält jedoch wie Smad2 und 3 ein PPxY (PY) Motiv, das von WW Domänen erkannt wird. WW Domänen sind kleine Bindungsdomänen, die in ca. 170 Proteinen verschiedener Funktionen identifiziert wurden und an vielen Zellprozessen beteiligt sind, wie unter anderem der Transkription, dem Spleißen oder dem Proteinabbau.

Im Hauptteil dieser Arbeit wurde die Wechselwirkung zwischen R-Smad3 und den Proteinen Nedd4L und Pin1 untersucht, um die Bedeutung der Smad3-Linker Phosphorylierung und deren Einfluss auf die Bindung zu WW Domänen zu verstehen. Im zweiten Teil wurde die Bindung zwischen dem Inhibitor-Smad7 und den Proteinen Nedd4L und TAZ charakterisiert, um unsere Kenntnis über deren inhibitorische Funktion innerhalb des TGF-ß Signalweges zu erweitern. Das Ziel in beiden Teilen war es zu verstehen, wie Spezifität von den WW Domänen erreicht werden kann. Die Strukturen von sechs Komplexen aus WW Domänen und Smads wurden gelöst und Gemeinsamkeiten und Unterschiede aufgezeigt.

Zunächst wurde die Bindung zwischen Nedd4L und dem phosphorylierten Smad2/3-Linker untersucht, welche den Smad2/3 Abbau zur Folge hat. Mittels Isothermer Titrationskalorimetrie (ITC) und Kernspinresonanzspektroskopie (NMR) konnte gezeigt werden, dass von allen vier Nedd4L WW Domänen die zweite Domäne das phosphorylierte Smad2/3 pT[PY] Motiv am stärksten bindet. Die NMR Struktur des Nedd4L WW2 –

Smad3 pT[PY] Komplexes zeigte, dass alle Prolinreste des PPPxY Teils des Liganden in *trans*-Konfiguration vorliegen und, dass der Smad3 Tyrosinrest von WW2 Aminosäuren in der Nähe der zweiten Schleife gebunden wird. Die Besonderheit dieses Komplexes ist die Wechselwirkung mit der Phosphatgruppe, welche von den Aminosäuren K378 und R380 innerhalb der ersten Schleife gebunden wird. Diese neue Art der Bindung konnte durch Punktmutationen innerhalb der Nedd4L WW2 Domäne und durch ^{31}P-NMR-Titrationsexperimente bestätigt werden.

Des Weiteren wurde die Bindung zwischen dem phosphorylierten Smad3-Linker und der WW Domäne der Peptidyl-Prolyl-Isomerase Pin1 untersucht. Pin1 katalysiert die *cis-trans* Isomerisierung von Prolinen, denen ein phosphoryliertes Serin oder Threonin in der Sequenz vorangeht. Bindungsstudien mittels ITC zeigten, dass die Pin1 WW Domäne das Smad3 pT179P Motiv am stärksten bindet während Bindungen zu dem zweifach phosphorylierten Smad3 pS204P-pS208P und pS208P-pS213P Peptiden schwächer oder wie im Fall des monophosphorylierten pS204P Peptids kaum zu beobachten waren. Um die unterschiedlichen Bindungsstärken zu verstehen, wurden die Strukturen des Pin1 WW – Smad3 pT179P und des Pin1 WW – Smad3 pS208P-pS213P Komplexes mittels NMR gelöst. Es konnte gezeigt werden, dass in beiden Komplexen nur ein p(S/T)P Motiv, ebenfalls in *trans*-Konfiguration, von der Pin1 WW Domäne gebunden wird, wobei Aminosäuren des ersten ß-Stranges an der Bindung der Phosphatgruppe beteiligt sind. Des Weiteren wurde gezeigt, dass die Bindung verstärkt wird durch negativ geladene Aminosäuren in Position x-1 sowie durch die Anwesenheit eines Prolinrestes in Position x-2 ausgehend vom p(S/T)P Motiv. Der Einfluss der Phosphatgruppe auf die Bindung wurde mittels ^{31}P-Titration verdeutlicht. Im Vergleich zum Nedd4L WW2 – Smad3 pT[PY] Komplex ist ein anderer Bindungsmodus im Falle der Pin1 WW – Smad3 pTP Wechselwirkung zu erkennen. Zum einen wird die Phosphatgruppe an unterschiedlichen Stellen von den Domänen gebunden, zum anderen wird das Peptid von der Pin1 WW Domäne vom N- zum C-Terminus gebunden und in umgekehrter Richtung vom C- zum N-Terminus von Nedd4L WW2. Diese strukturellen Unterschiede im Bindungsmodus könnten die verschiedenen Funktionen der beiden Wechselwirkungen widerspiegeln.

ITC Messungen der Nedd4L WW2 – Smad3 pT179[PY] Wechselwirkung als auch der Pin1 WW – Smad3 pS208P-pS213P Bindung bei unterschiedlichen Protonenkonzentrationen

zeigten, dass in beiden Fällen die Bindungsstärke mit dem pH Wert steigt. Die höchste Affinität wurde für pH 8.0 ermittelt. Diese Information wurde zur Optimierung der NMR-Experimente genutzt, kann aber auch eine Rolle bei der Regulierung von Protein-Protein Wechselwirkungen in der physiologischen Umgebung der Zelle spielen.

Um die Funktion der übrigen WW Domänen in Nedd4L zu verstehen, wurden Smad3 Fragmente mit dem PY Motiv und weiteren pSP Seiten zu den drei möglichen Nedd4L WW Paaren titriert. Von den drei aufeinanderfolgenden WW Paaren in Nedd4L, zeigte lediglich das WW2-WW3 Paar eine hohe Affinität zum Smad3-Linker (Smad3 176-211 Segment dreifach phosphoryliert in Position T179, S204 und S208). Die Strukturbestimmung des Nedd4L WW2-WW3 – Smad3 pT[PY]-pS-pS Komplexes stellte aufgrund der schlechten NMR-Signaldispersion der 80-Aminosäure-Region, welche die WW2 und die WW3 Domänen miteinander verbindet, eine große Herausforderung dar. Um die Interpretation der NMR Daten zu vereinfachen, wurde das WW2-WW3 Paar in zwei getrennten Fragmenten hergestellt, wodurch eine sequenzielle Isotopenmarkierung ermöglicht wurde. Mittels chemischer Ligation wurden die beiden Fragmente durch eine Disulfidbrücke miteinander verbunden. Mit diesem schrittweisen Ansatz konnte die Struktur des Nedd4L WW2-WW3 – Smad3 Komplexes im Detail untersucht werden.

Im zweiten Teil dieser Arbeit wurde die Frage aufgeworfen wie der Inhibitor-Smad7 durch die WW Domänen der Proteine Nedd4L und TAZ erkannt wird. Im Zytoplasma bindet Smad7 mit Nedd4L und bewirkt den Abbau des TGF-ß Rezeptor I während TAZ hauptsächlich im Zellkern vorhanden ist. ITC Bindungsstudien der individuellen Nedd4L WW Domänen und einem Smad7-Linker Peptid aus 15 Aminosäuren inklusive dem PY Motiv zeigten, dass die zweite WW Domänen die bevorzugte Smad7 Bindungsdomäne ist. Die hohe Bindungsstärke wird von zusätzlichen Kontakten des Smad7 C-Terminus und der ß1-Strangregion der Nedd4L WW2 Domäne hervorgerufen, sowie durch elektrostatische Wechselwirkungen des Smad7 N-Terminus und der ersten Schleife der WW2 Domäne. Dies konnte mittels NMR basierende Strukturanalyse des Nedd4L WW2 – Smad7[PY] Komplexes verdeutlicht werden. Darüber hinaus zeigten ITC Experimente, dass die Bindung durch die Präsenz weiterer Nedd4L WW Domänen nicht verstärkt wird.

Der Co-Faktor YAP und sein Paralog TAZ besitzen zwei (YAP) und eine (TAZ) WW Domäne(n) und wurden ebenfalls als Smad Bindungspartner beschrieben. Die Bindung zwischen YAP und Smad7 führt zu verstärktem TGF-ß Rezeptor I Abbau, während die Wechselwirkung zwischen TAZ und Smad7 bisher noch nicht untersucht wurde. Eine Sequenzanilierung der YAP und TAZ WW Domänen zeigt eine höhere Sequenzidentität zwischen YAP WW1 und TAZ WW als zwischen den beiden YAP Domänen. In dieser Arbeit konnte durch ITC Experimente gezeigt werden, dass die YAP WW1 Domäne das Smad7-Linkerpeptid stärker bindet als die WW2 Domäne und zwar mit einer ähnlichen Bindungsstärke wie sie zwischen der TAZ WW Domäne und Smad7 ermittelt wurde. Strukturen der YAP WW1 Domäne im Komplex mit PY Motiven wurden bereits in der Literatur mehrfach beschrieben allerdings ist dies das erste Mal, dass die Struktur einer WW Domäne des humanen TAZ Proteins mittels NMR bestimmt wurde. Im Vergleich mit Nedd4L binden sowohl TAZ WW als auch Nedd4L WW2 das Smad7 PY Motiv und den verlängerten C-Terminus sehr ähnlich jedoch sind Unterschiede in der Wechselwirkung mit dem Smad7 N-Terminus zu erkennen.

Zusammenfassend erweitern die in dieser Arbeit präsentierten Ergebnisse unser Wissen über die strukturelle und funktionelle Vielseitigkeit der WW Domänen als Protein-Protein Wechselwirkungsmodule. Darüber hinaus zeigt diese Arbeit, dass Smad3-Linker-Phosphorylierung, welche von verschiedenen Kinasen in einer bestimmten Reihenfolge durchgeführt wird, die Interaktion mit Nedd4L und Pin1 reguliert. Nedd4L bindet das Smad3 pT[PY] Motiv und zwei weitere Phosphorylierungsstellen innerhalb des Smad3-Linkers. Die Pin1 WW Domäne benötigt lediglich eine Phosphorylierungsstelle, welche möglicherweise als Anker für die Prolyl-Isomerase dient. Pin1 WW Domäne könnte daher an der Smad3 Regulierung beteiligt sein und die Nedd4L – Smad3 Wechselwirkung durch *cis-trans* Isomerierung des Smad3-Linkers beeinflussen. Des Weiteren konnte gezeigt werden, dass Nedd4L durch einen kooperativen, zwei Stufen Modus der benachbarten Domänen WW2 und WW3 eine hohe Bindungsspezifität gegenüber Smad3 erreicht. Diese Fähigkeit der WW Domänen sowohl pT[PY] als auch pS-pS Motive zu erkennen war bislang unbekannt.

Die Bindung zwischen Nedd4L und Smad7 führt zur TGF-ß Inhibierung durch den Abbau des TGF-ß Rezeptors I. In dieser Arbeit konnte gezeigt werden, dass die zweite Nedd4L WW Domäne ausreicht, um das Smad7[PY] Motiv mit hoher Affinität zu binden. Die WW Domäne

des Adapter-Proteins TAZ interagiert ebenfalls mit dem Inhibitor-Smad7 aber mit einer etwas schwächeren Affinität. In beiden Komplexen ist die Bindung des PY Motivs und der C-terminalen Verlängerung sehr ähnlich, sie unterscheiden sich jedoch in der Erkennung des Smad7 N-Terminus. Diese Differenzen könnten die unterschiedlichen Funktionen der beiden Wechselwirkungen widerspiegeln. TAZ war bisher als TGF-ß Agonist bekannt, in dieser Arbeit konnte jedoch gezeigt werden, dass TAZ auch eine hemmende Funktion im TGF-ß Signalweg einnehmen kann wie bereits für seine Paralog YAP berichtet wurde.

Table of Contents

Introduction

1 Introduction

1.1 TGF-ß signalling and Smads

TGF-ß (transforming growth factor beta) signalling is a communication system able to transfer a signal from outside the cell through the membrane and the cytoplasm to the nucleus where it regulates the expression of target genes. TGF-ß signalling regulates proliferation, cell differentiation, migration, apoptosis, cell division and cell fate, thus malfunction of this pathway can result in the progression of autoimmune diseases or cancer.

The signalling cascade is initiated by multifunctional cytokines (TGF-ß growth factors), which interact and thus activate the TGF-ß receptors. The first TGF-ß growth factor TGF-ß1 was discovered in the early 1980s. To date the TGF-ß growth factor family has grown significantly to 42 members in humans, seven in drosophila and four in Caenorhabditis elegans (C. elegans) (Massague 1998). The family members in vertebrates are further divided into two branches, the bone morphogenic protein (BMP)/growth and differentiation factor (GDF)/Muelerian inhibiting substance (MIS) subfamily and the TGF-ß/Actin/Nodal subfamily. Receptors of the TGF-ß family are located at the cell membrane and are complexes formed by a cell surface receptor domain and an intracellular Ser/Thr-kinase domain. The Ser/Thr-receptors can be distinguished between type I and type II receptors depending on their transmembrane proteins. After receptor type II activation upon ligand binding, the receptor type I is bound and phosphorylated by the receptor type II kinase domain. In the human genome seven distinct type I and five different type II receptors are known but the TGF-ß receptor type II (TßRII) only binds TGF-ß growth factor and only incorporates TGF-ß receptor type I (TßRI) in this complex. In contrast, the type I and type II receptors for BMP-like growths factors bind ligands independently of each other and in different receptor combinations (Groppe et al. 2008).

The signal is further propagated by the activated type I receptor through binding and phosphorylation of the intracellular mediator proteins called Smads (small mothers against decapentalplegic) (Massague 1998; Massague and Chen 2000; Attisano and Wrana 2002; Shi and Massague 2003; Feng and Derynck 2005; Massague and Gomis 2006). Three distinct

classes of Smad proteins exist: the receptor regulated-Smads (R-Smads) including Smad1/2/3/5 and 8, the common partner-Smad (Co-Smad) Smad4 and the inhibitor-Smads (I-Smads) Smad6 and 7. Smad proteins consist of two globular domains, the Mad-homology1 (MH1) DNA binding domain and the Mad-homology2 (MH2) domain connected by a flexible linker that contains multiple phosphorylation sites. The MH1 domain is stabilized by a tightly bound zinc atom and is highly conserved in Co- and R-Smads but not in I-Smads. The MH2 domain is conserved in all Smads and plays an important role in the interaction with the receptor. The activated type I receptor phosphorylates the R-Smad C-tail sequence of repeated Ser-X-Ser (SXS) motifs resulting in R-Smad activation. TGF-ß receptors type I recognize Smad2/3, whereas Smads1/5/8 bind to BMP receptors (Shi and Massague 2003).

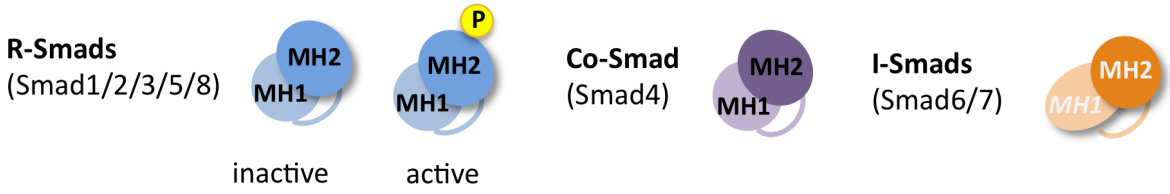


Figure 1.1: Scheme of the three different classes of Smad proteins: R-Smads in the inactive and active (phosphorylated at the C-tail) state, Co-Smad and I-Smads.

Activated R-Smads are transferred to the nucleus with the help of Co-Smad4 by forming heteromeric Co-Smad4 – R-Smad complexes. In the nucleus the Smad complex can associate with one of many DNA binding partners and various transcriptional co-activators or co-repressors resulting in positive or negative regulation of gene expression (Massague 1998; Massague and Chen 2000; Attisano and Wrana 2002; Shi and Massague 2003; Feng and Derynck 2005; Massague et al. 2005; Massague and Gomis 2006; Ross and Hill 2008).

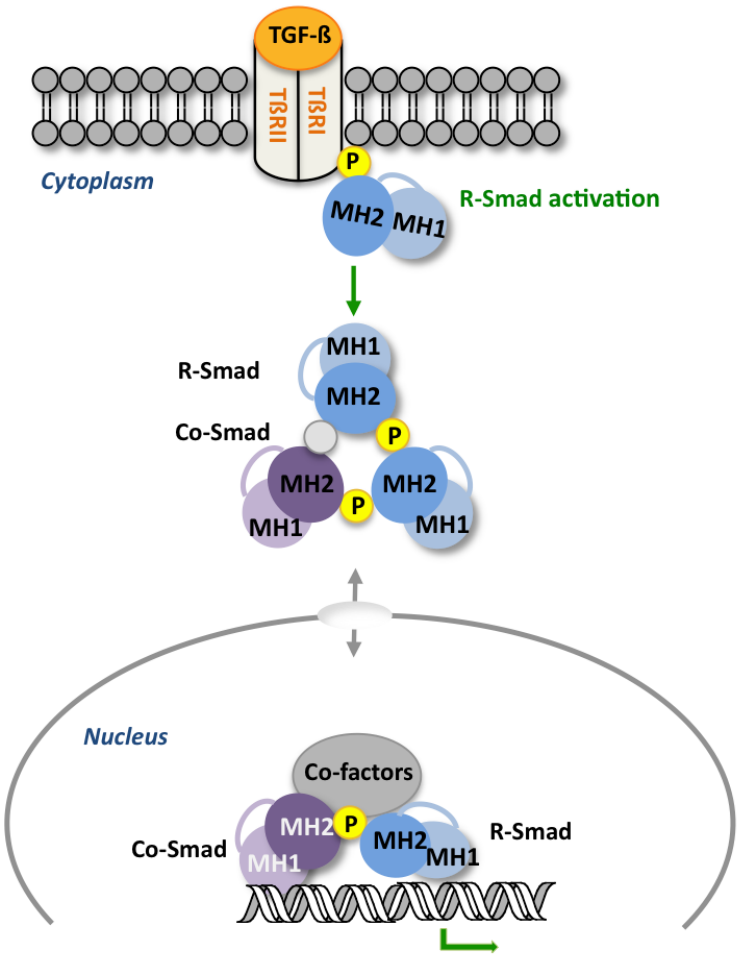


Figure 1.2: R-Smad activation in TGF-ß signalling: Phosphorylated TßRI receptor activates R-Smad by C-tail phosphorylation. Actviated R-Smad forms a complex with the Co-Smad that is transferred to the nucleus where it activates the transcription of target genes with the help of transcription Co-factors.

1.1.1 R-Smad linker region

The linker region of R-Smad proteins contains several phosphorylation sites. Different kinases are able to phosphorylate the linker resulting in both, positive and negative regulation of the signalling pathway. Mitogen-activated protein kinases (MAPKs) can negatively regulate TGF-ß or BMP signalling through phosphorylation of (S/T)P motifs within the R-Smad linker region in the cytoplasm as respond to cellular stress (Grimm and Gurdon 2002; Fuentealba et al. 2007; Sapkota et al. 2007). Following TGF-ß or BMP signalling, in the nucleus the activated R-Smads through C-tail phosphorylation (Wang et al. 2009a), are further phosphorylated at their linker region by cycline-dependent kinases (CDKs)8/9 which enhance the transcriptional action of Smads (Matsuura et al. 2004; Mori et al. 2004; Alarcon et al. 2009; Matsuzaki et al.

2009; Millet et al. 2009; Wang et al. 2009a). CDK8/9 are components of a transcriptional mediator complexe that regulate RNA polymerase II during transcription. The linker is also phosphorylated by glycogen synthase kinase-3 ß (GSK3ß). GSK3ß is found in both cytoplasm and nucleus and recognizes a previously phoshorylated substrate through a specific binding pocket. Subsequently GSK3ß can phosphorylate a serine or a threonine at position n – 4 after CDK8/9 phosphoruylation (Cohen and Frame 2001; Rayasam et al. 2009) and switches the Smad linker from one that favors Smad action to one that favors Smad destruction (Aragon et al. 2011).

The linker region of all R-Smads contains in addition to the phosphorylation sites a PY motif (PPxY, where x could be any amino acid), which is recognized by WW containing proteins, some of which are E3 ubiquitin ligases. In Smad2/3 three of the phosphorylation sites are located downstream of the PY motif, whereas the forth one lies directly adjacent to the PY motif. In Smad1/5/8 all phosphorylation sites precede the PY motif (Sapkota et al. 2007; Alarcon et al. 2009).

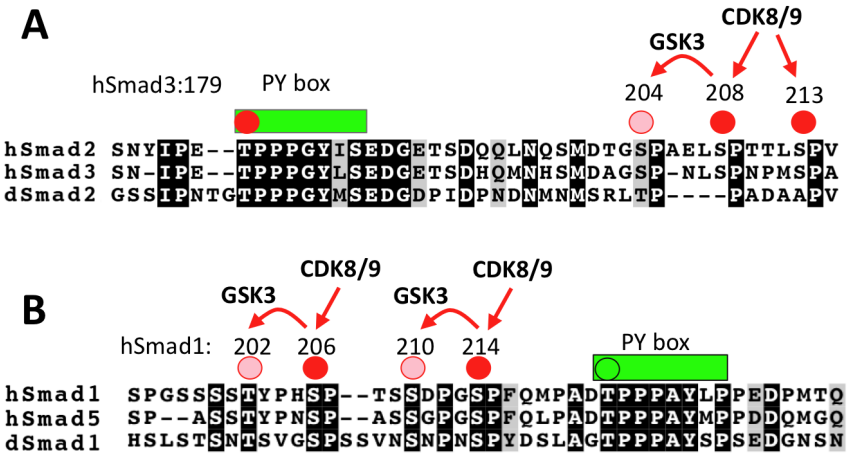


Figure 1.3: Scheme of the R-Smad linker region indicating the location of the PY motif and phosphorylation sites. (A) Smad2/3 (B) Smad1/5, h: human, d: drosophila.

1.2 R-Smad binding proteins with WW domains

The R-Smad linker region is recognized by WW domain containing proteins through interactions with the PY motif. WW domains are the smallest independently folded natural domains. They fold as triple stranded anti-parallel ß-sheet and bind to Pro-rich sequences containing p(S/T)P and PPxY motifs. They exist in proteins as single domains or as several copies. The name of WW domains refers to two highly conserved Trp residues about

20 amino acids apart from each other. Although these two Trp play an important role in domain stability and ligand binding, in some cases one of the Trp is replaced by another aromatic acid like Phe or Tyr (Bork and Sudol 1994; Macias et al. 1996; Koepf et al. 1999; Macias et al. 2002). Regarding their binding motifs they are divided into four groups. WW domains that belong to group I recognize PY motifs as present in the R-(and I-)Smad linker region. WW domains of group II interact with PPLP motifs, the group III of WW domains binds poly-Pro ligands containing Arg (PR motifs) and group IV of WW domains recognizes phosphorylated Ser or Thr residues followed by a Pro (p(S/T)P motifs) (Sudol and Hunter 2000; Kato et al. 2004). A comparison of the binding properties of about 200 WW sequences have revealed that more than half (56%) of the WW domains recognize PY motifs, whereas only 4% WW fell into group IV (Kato et al. 2002; Kato et al. 2004).

Table 1.1: Classification of WW domains.

Group	Motif	Abundance	Examples	
I	PY	56%	Hect type E3-ubiquitin ligase	(Nedd4L, Smurf1/2)
II	PPLP	27%	Formin binding protein 11/28	(FBP11, FBP28)
III	PolyP-R		Formin binding protein 30	(FBP30)
IV	pS/pT-P	4%	Peptidyl-prolyl isomerase	(Pin1)
Unclassified		13%		

1.2.1 Hect type E3-ubiquitin ligases

Hect type E3-ubiquitin ligases trigger the degradation of Smads by the ubiquitin proteosome pathway. Hect type E3-ubiquitin ligases are together with the zinc-binding RING (really interesting new gene) finger E3 ligases the main regulator of the ubiquitination cascade that is catalyzed by three enzymes. The activation enzyme E1 binds ubiquitin, a small, extremely well conserved protein among eukaryota, in an ATP dependent reaction. The activation is carried out through a Cys residue of the E1 enzyme that forms a thiol ester bond to a Gly carboxyl group located at the ubiquitin C-terminal. Subsequently in a transacylation reaction the E2 conjugating enzyme picks up the activated ubiquitin from the E1 enzyme. Finally the ubiquitin is directly transferred from the E2 enzyme to a Lys ε-amino group of the substrate by the E3 ligase (Hochstrasser 2009). Ubiquitination controls biological signalling in many different ways depending whether the substrate is mono- or polyubiquitination and depending on the type of the polyubiquitin chain. Ubiquitin consists of seven Lys residues, which can all participate in

the ubiquitin chain formation. Polyubiquitination by Lys48 ubiquitin chains results in substrate degradation by the 26S proteosome.

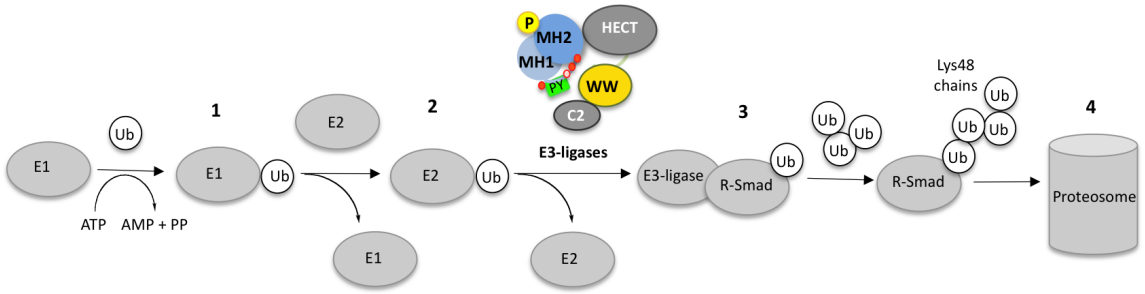


Figure 1.4: Ubiquitination pathway with R-Smad as substrate: (1) Activation of ubiquitin (Ub) by the E1 enzyme. (2) Transfer of the activated Ub to E2. (3) Substrate (R-Smad) binding by E3 ligase and transfer of Ub from E3 to R-Smad. (4) Polyubiquitination (Lys48 chains) of R-Smad resulting in proteosomal degradation.

Hect type E3-ubiquitin ligases have been localized in the cytoplasm, plasma membrane and in the nucleus (Chen and Matesic 2007). They interact with many substrates and their malfunction has a role in several diseases like liddle's syndrome, or cancer (Scheffner and Staub 2007; Bernassola et al. 2008). All Hect type E3-ubiquitin ligases have a similar architecture. They consist of a N-terminal C2 domain for membrane association, two to four WW domains for substrate recognition and a catalytic HECT (Homologous to E6-AP C-terminus) domain (Ingham et al. 2004). It has been shown that Smurf1 (Smad ubiquitin regulatory factor 1) mainly inhibits BMP signalling by binding to Smad1/5/8, whereas Nedd4L, Smurf2 and WWP1 (WW domain containing protein 1) interact with Smad2/3 and are antagonists of TGF-ß signalling (Zhang et al. 2001). The interaction between Nedd4-like E3 ligases and R-Smads is enhanced by R-Smad linker phosphorylation (Sapkota et al. 2007). Phosphorylation of a Tyr/Thr/Ser residue within the WW binding site or between WW domains can decrease or inhibit substrate recognition (Debonneville et al. 2001; Morales et al. 2007).

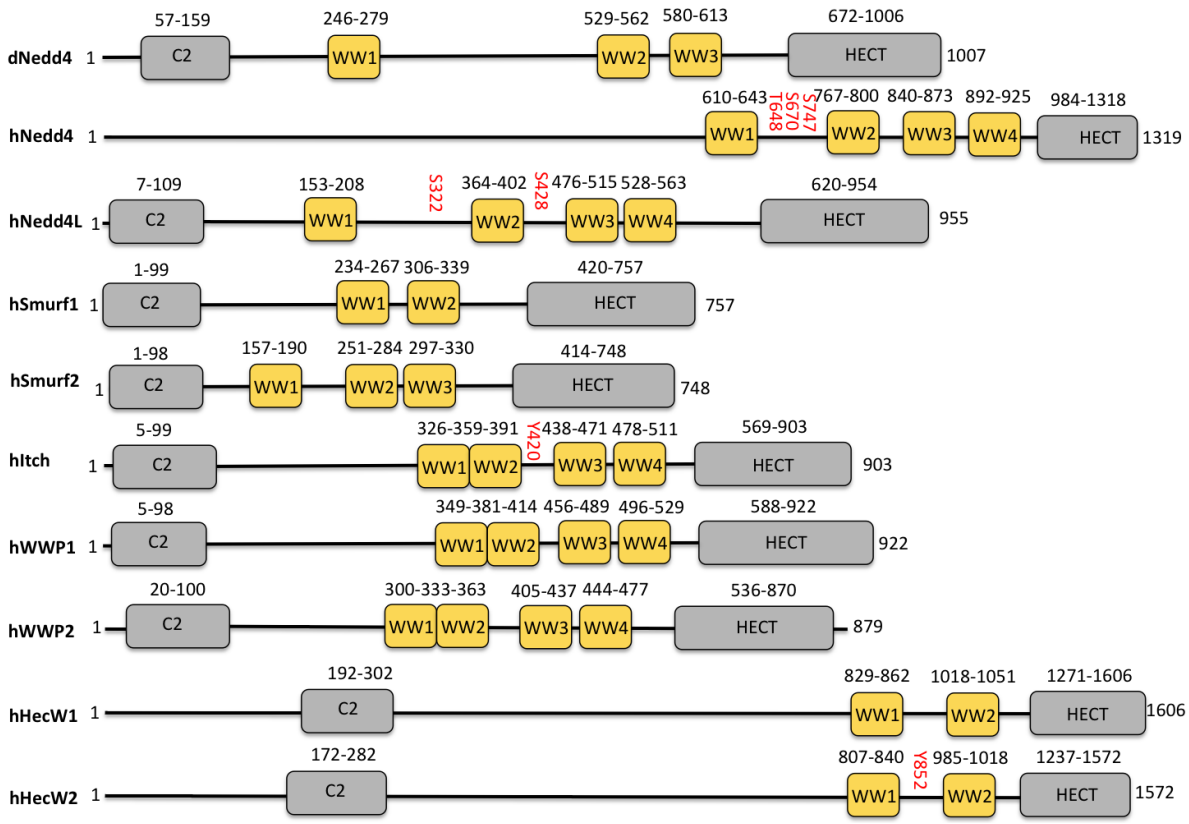


Figure 1.5: Members of Hect type E3-ligases (h: human, d: drosophila) with the typical topology of a N-terminal C2 domain, two to four WW domains and a catalytic HECT domain. Phosphorylation sites between WW domains are in red.

1.2.2 E3-ubiquitin ligase Nedd4L

The E3-ubiquitin ligase Nedd4L (Nedd4-like) was discovered in 2001, ten years after Nedd4 (Neural-precursor-cell-expressed, developmentally down regulated gene 4), and displaced Nedd4 as main physiological regulator of the epithelial sodium channel (ENaC) (Asher et al. 2003). Up to date several Nedd4L substrates have been identified (Kuratomi et al. 2005; Yang and Kumar 2010). As a channel regulator, Nedd4L is able to bind the subunits of ENaC through its WW domains with WW3 being the major binding domain (Asher et al. 2003; Fotia et al. 2003). It also has been reported that phosphorylation of the ENaC γ subunit increases binding to the WW domains of both Nedd4 and Nedd4L proteins (Shi et al. 2002a; Asher et al. 2003). The presence of several WW domains in Nedd4L is a common feature of all Hect type E3-ubiquitin ligases (Figure 1.5) and it raises questions about how specificity is achieved and if they work independently or organized in a cooperative way.

1.2.3 Peptidyl-prolyl isomerase Pin1

Pin1 (Protein interacting with NIMA (never in mitosis A-1)) is a member of the peptidyl-prolyl isomerase (PPIase) family that recognizes p(S/T)P motifs of phosphorylated proteins and catalyses prolyl cis-trans isomeration (Lu et al. 1996). Proline residues can adopt *cis* or *trans* configuration in a protein backbone with distinct functional outcome. The transition can occur spontaneously at slow rate or is catalysed by *cis-trans* isomerases. Phosphorylation further slows down the proline *cis-trans* conversion. Pin1 is the only phosphorylation dependent prolyl *cis-trans* isomerase and contains a N-terminal group IV WW domain that recognizes the p(S/T)P motif and a catalytic PPIase domain at the C-terminus. The domains are connected through a Gly-rich, flexible linker (Figure 1.6).

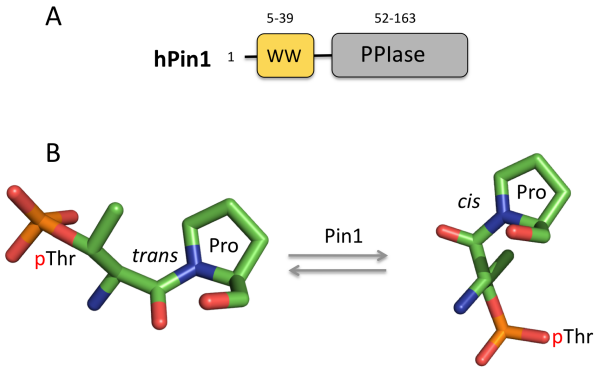


Figure 1.6. (A) Schematic representation of Pin1 with one WW domain for substrate recognition and a catalytic PPIase domain, h: human. (B) Proline *cis-trans* isomeration of a pTP motif catalyzed by Pin1.

The first crystal structure of the whole 18 kDa Pin1 protein was solved in 1997 (Ranganathan et al. 1997). Structural studies of Pin1 WW domain bound to a double phosphorylated peptide of the C-terminal domain (CTD) of RNA polymerase II (Pol II) large subunit (Verdecia et al. 2000) and NMR studies of the Pin1 WW domain bound to single phosphorylated peptides derived from the mitotic phosphatase Cdc25 and microtubule-associated human tau protein (Wintjens et al. 2001) have been accomplished shortly after. Multiple cell cycle regulatory proteins are Pin1 binding partners, thus it was originally classified as a cell cycle protein. To date many other Pin1 substrates were identified, which are involved in different cellular processes like cell growth and immune response. Pin1 is overexpressed in several cancer tissues but also plays a role in human diseases like Alzheimer's disease and asthma (Wulf et al. 2005; Lippens et al. 2007; Lu et al. 2007; Lu and Zhou 2007; Yeh and Means 2007). The mechanisms of Pin1 binding and isomerisation reactions remain unclear at the molecular level and its precise function in the cell is controversial. Pin1 is able to stabilize and activate or to destabilize and degrade a given target and for several pathways Pin1 can act both as positive or

negative regulator. Among the various Pin1 targets the Smad proteins were recently identified (Nakano et al. 2009). It has been shown that the Pin1 WW domain can recognize the phosphorylated linker region of Smad2 and Smad3 resulting in their degradation by enhancing the binding between an E3 ubiquitin ligase and Smad2/3 (Nakano et al. 2009). Based on these findings it has been suggested that Pin1 negatively regulates TGF-ß signalling like Nedd4L. Especially the Pin1 WW domain, which is responsible for substrate recognition, is an interesting target to study the role of Pin1 in TGF-ß signalling.

1.3 I-Smad7 binding proteins with WW domains

In their basal state the inhibitory Smad6/7 are mainly localized in the nucleus and upon TGF-ß stimulation they move to the cytoplasm (Itoh et al. 1998). In the cytoplasm Smad6/7 can negatively regulate TGF-ß signalling in two ways: I-Smads bind Co-Smad4 to block the binding site for the activated R-Smads and thus retains R-Smad translocation to the nucleus (Figure 1.7 (A)) or they can inhibit R-Smad activation by blocking the binding site at receptor I (Figure 1.7 (B)). Smad7 recruits E3-ubiquitin ligases like WWP1, Smurf2, and Nedd4L or Smurf1 to the TGF-ß or BMP receptors resulting in receptor-degradation a drastic way of inhibiting the signalling from membranes to the nucleus (Figure 1.7 (C)) (Itoh et al. 1998; Kavsak et al. 2000; Komuro et al. 2004; Kuratomi et al. 2005).

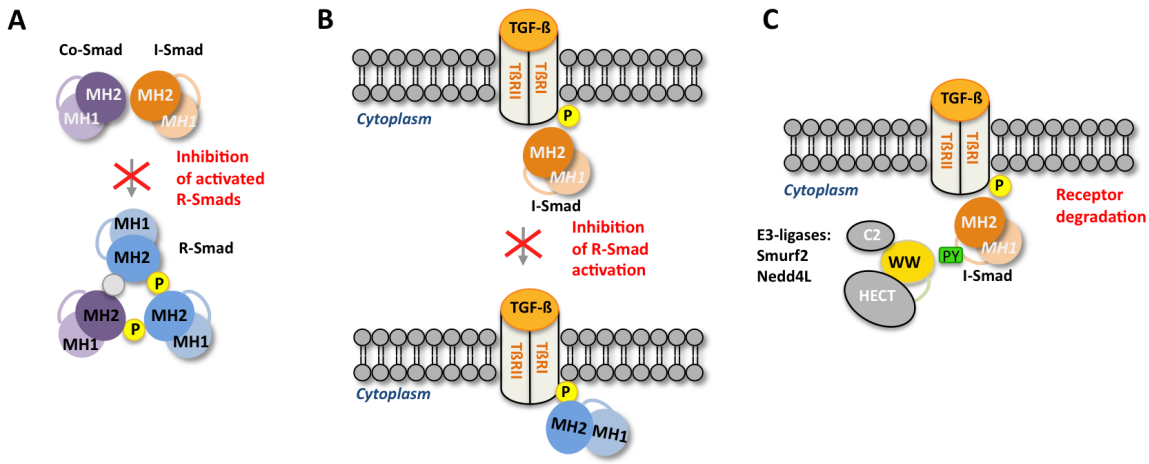


Figure 1.7: Schematic representation of TGF-ß inhibition by I-Smad6/7. (A) Inhibition of activated R-Smads by Smad4 binding, (B) Inhibition of R-Smad activation by-TßRI binding, (C) TßRI degradation by E3-ligase binding.

1.3.1 Transcription co-activators YAP and TAZ

YAP (Yes-kinase associated protein) and its paralogue TAZ (transcription co-activator with PDZ-binding motif) also known as WW domain containing transcription regulator 1 (WWtr1) are transcriptional co-activators conserved from fly to human (Yagi et al. 1999; Kanai et al. 2000).

YAP and TAZ induce organ growth and are controlled by the Hippo tumour suppressor pathway (Zhao et al. 2008). They have a similar topology that consists of one or two WW domains depending on the isoform, a N-terminal TEAD (transcriptional enhancer factor-domain protein)-binding domain and a C-terminal transcriptional activation domain (Wang et al. 2009b).

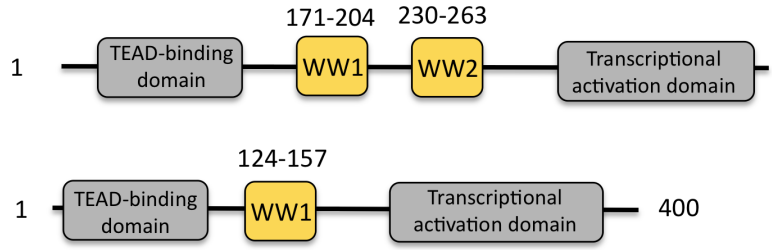


Figure 1.8: Schematic representation of human YAP and TAZ with TEAD-binding domain, WW domains and transcriptional activation domains. One example of each isoforms is represented.

Several YAP/TAZ binding partners have been identified such as transcription factors, 14-3-3 proteins and Smad proteins of the TGF-ß and BMP pathway (Kanai et al. 2000; Wang et al. 2009b). Upon BMP signalling YAP has been reported to interact with linker phosphorylated Smad1/5 (Alarcon et al. 2009) and can bind to the Smad4 - Smad2/3 complex in response to TGF-ß (Varelas et al. 2010). Apart from the interaction with R-Smads and Co-Smad4, YAP also binds the inhibitor Smad7 resulting in enhanced TßRI degradation (Ferrigno et al. 2002). TAZ has been reported to bind the Smad4 - Smad2/3 complex in response to TGF-ß (Varelas et al. 2008) but binding between TAZ and Smad7 has not been studied as yet. However, based on its sequence similarity with YAP and with the functional parallelism between both proteins, it is most likely that TAZ and YAP interact with Smad7 in a similar way.

Aims of this thesis

2 Aims of this thesis

In the course of their action in the nucleus, receptor regulated-Smads (R-Smads) undergo certain phosphorylation events that enable peak transcriptional activity but also mark the proteins for destruction. These findings present a paradox but also an opportunity to define how the delivery of TGF-ß signals is coupled to the turnover of the Smad signal transducers. R-Smads consist of a globular amino-terminal MH1 domain with DNA-binding activity, a carboxy-terminal MH2 domain that mediates key protein-protein interactions, and an inter-domain linker region with a conserved cluster of phosphorylation sites adjacent to a PY motif. The inhibitor-Smads (I-Smads) contain a more divergent MH1 domain and as the R-Smads the MH2 domain and the PY motif within the linker. The presence of the common regions facilitates the competition of R-Smads and I-Smads for receptor and ligands and it accounts for the inhibitory role of I-Smads.

The objective of this thesis was to characterize the interactions between the TGF-ß signal mediators Smads and binding partners with WW domains to decipher the rules that govern the specificity of the interactions. To achieve these objectives the work has been divided into two main goals.

The first aim was to characterize the interaction of the R-Smad3 linker with either the WW domains of the E3 ligase Nedd4L or with the unique WW domain of the prolyl *cis-trans* isomerase Pin1 using isothermal titration calorimetry and nuclear magnetic resonance in solution. This aim was achieved by dissecting the problem in three parts:

- Bindings of the four individual WW domains of Nedd4L and the PY motif were determined and the impact of phosphorylation site upstream of the PY motif was studied.
- Bindings of the Pin1 WW domain and the Smad3 p(S/T)P motifs within the Smad3 linker were studied to elucidate Smad3 recognition by Pin1.
- Bindings of the Nedd4L WW domain pairs to extended regions of the Smad3 linker was studied to understand how specificity is achieved.

The second aim was to study the interaction between the I-Smad7 and WW domains of two TGF-ß proteins in their role as TGF-ß inhibitors. A short Smad7[PY] linker peptide was synthesized to study the interactions with the individual Nedd4L WW domains as well as the Nedd4L WW domain pairs. Further the binding of the unique WW domain of transcription co-activator TAZ and Smad7 was characterized for the first time.

Differences and similarities between Nedd4L binding of Smad3 and Smad7 on atomic level were pointed out and possible roles of the proteins Pin1 and TAZ in TGF-ß signalling were proposed.

Methods and Materials

3 Methods and Materials

3.1 Cloning of WW domains

3.1.1 Cloning of the Nedd4L WW domains

Nedd4L WW domains (WW1 173-208, WW2 364-402, WW3 476-515, WW4 528-563) were cloned in pET (premier *E. coli* expression) systems with kanamycin resistance, tobacco etch virus (TEV) protease cleavage site, N-terminal His_6-tag and with (pETM30) or without (pETM11) glutathion-S-transferase (GST)-tag (Sheng Gao, Joan Massagué's group, Memorial Sloan-Kettering Cancer Center, NY). In both cases NcoI and HindIII sites were used.

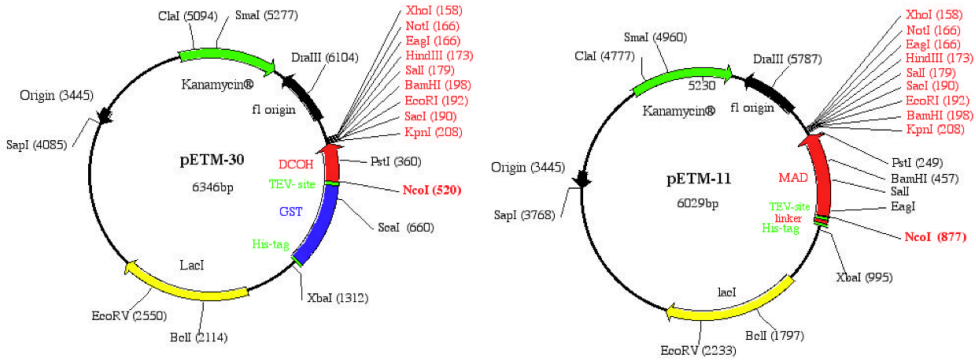


Figure 3.1 Plasmids pETM30 and pETM11.

All molecular weights (Mw) presented in this thesis are isotopic average masses calculated with the program *ProtParam* of the proteomic server *ExPASy (http://www.expasy.ch/tools/protparam.html)*. All extinction factors (ε) are calculated with the program *ProtParam/ExPASy* at 280 nm.

Table 3.1: Sequences of the recombinant Nedd4L WW domains.

Nedd4L	Sequence	ε [M^{-1} cm^{-1}]	Mw [g/mol]
WW1	GAMEPPLPPG WEEKVDNLGR TYYVNHNNRT TQWHRPSLMD	13980	4708.2
WW2	GAMATPGLPS GWEERKDAKG RTYYVNHNNR TTTWTRPIMQ LAE	13980	4906.4
WW3	GAMAWQSFLP PGWEMRIAPN GRPFFIDHNT KTTTWEDPRL KFPVH	16500	5281.9
WW4	GAMGPLPPGW EERIHLDGRT FYIDHNSKIT QWEDPRLQN	12490	4576.0

3.1.2 Cloning of the Nedd4L WW2 mutants

All Nedd4L WW2 domain mutants were generated (Eric Aragon, Maria J. Macias group, IRB Barcelona) using the Nedd4L WW2 wild type as template. Mutations were introduced via polymerase chain reaction (PCR) site-directed mutagenesis with a pair of complementary primers (forward and reverse) containing the desired mutation. The entire plasmid (pETM11/30) was amplified by *Pfu* DNA polymerase (Agilent Technologies, CA, USA) generating a circular DNA. Unlike *Taq* DNA polymerase, applied in normal PCR, *Pfu* DNA polymerase corrects misincorporation errors and requires longer elongation time and less PCR cycles. The DNA is not amplified exponentially, rather linearly. For this reason the template DNA remaining in the mixture after the reaction needs to be removed. The restriction endonuclease DnpI selectively digests the template DNA by cleaving only at methylated DNA sequences (5'-GmeATC-3'). The biosynthesized methylated vector with the Nedd4L WW2 domain sequence is digested, but the *in vitro* generated vector with the mutated Nedd4L WW2 domain remains (Hutchison et al. 1978).

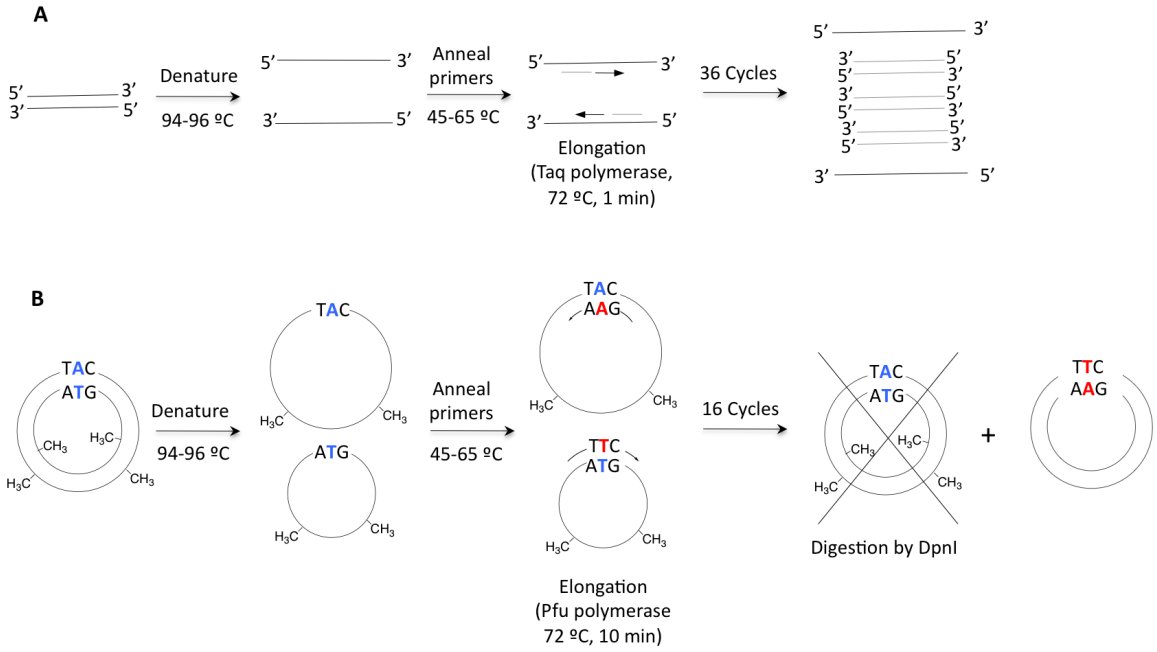


Figure 3.2: (A) Normal PCR with cDNA template. (B) PCR site-directed mutagenesis.

The vector was purified with the MinElute PCR Purification Kit (QIAgen, Hilden, Germany) and transformed into DH5α cells (Invitrogen, Carlsbad, CA, USA) as described in chapter 3.2.1 for the transformation into BL21 cells. The bacterial cells were dispersed on an agar plate, which remained at 37 ºC overnight to let the cells form colonies. One colony was picked,

transferred into 5 ml lysogeny broth (LB-medium) and let grown over night at 37 ºC. The plasmid DNA was extracted from the bacterial cell suspensions using a miniPrep protocol (QIAgen, Hilden, Germany). This procedure takes advantage of the fact that plasmids are relatively small, supercoiled DNA molecules and bacterial chromosomal DNA is larger and less supercoiled. Cells are lysed under alkaline conditions, which denatures both nucleic acid and proteins. When the solution is neutralized by the addition of potassium acetate, chromosomal DNA and proteins precipitate because it is impossible for them to renature correctly (Bimboim H C 1979). In contrast plasmids renature correctly and stay in solution.

Table 3.2: Sequences of the recombinant Nedd4L WW2 mutants.

Mutation	Sequence	Mw [g/mol]
K378/E	GAMATPGLPSG WEERKDA**E**GR TYYVNHNNRT TTWTRPIMQL AE	4907.3
R380/E	GAMATPGLPSG WEERKDAKG**E** TYYVNHNNRT TTWTRPIMQL AE	4879.3
K378R380/EE	GAMATPGLPSG WEERKDA**E**G**E** TYYVNHNNRT TTWTRPIMQL AE	4880.2

3.1.3 Cloning of the Nedd4L WW2-WW3, WW1-WW2 and WW3-WW4 constructs

The Nedd4L pair of domains WW1-WW2 (173-400), WW2-WW3 (364-512) and WW3-WW4 (476-563) were cloned by Eric Aragon using the Nedd4L DNA in a pC2-HA plasmid provided by Joan Massagué's group. All ligation reactions were performed within 15 min at 25 ºC with three-fold DNA excess with respect to the vector.

Table 3.3: Sequences of the recombinant Nedd4L pairs WW2-WW3, WW1-WW2 and WW3-WW4.

Nedd4L	Sequence	ε [M^{-1} cm^{-1}]	Mw [g/mol]
WW2-WW3	GAMATPGLPS GWEERKDAKG RTYYVNHNNR TTTWTRPIMQ LAEDGASGSA TNSNNHLIEP QIRRPRSLSS PTVTLSAPLE GAKDSPVRRA VKDTLSNPQS PQPSPYNSPK PQHKVTQSFL PPGWEMRIAP NGRPFFIDHN TKTTTWEDPR LKFPVH	26.470	17320.3
WW1-WW2	GAMAPPLPPG WEEKVDNLGR TYYVNHNNRT TQWHRPSLMD VSSESDNNIR QINQEAAHRR FRSRRHISED LEPEPSEGGD VPEPWETISE EVNIAGDSLG LALPPPPASP GSRTSPQELS EELSRRLQIT PDSNGEQFSS LIQREPSSRL RSCSVTDAVA EQGHLPPPSA PAGRARSSTV TGGEEPTPSV AYVHTTPGLP SGWEERKDAK GRTYYVNHNN RTTTWTRPIM QLAE	34950	23978.3
WW3-WW4	GAMAQSFLPP GWEMRIAPNG RPFFIDHNTK TTTWEDPRLK FPVHMRSKTS LNPNDLGPLP PGWEERIHLD GRTFYIDHNS KITQWEDPRL QN	23490	10752.1

3.1.4 Cloning of the Nedd4L WW2-WW3 point mutants

The Nedd4L WW2-WW3 mutants were cloned from the Nedd4L WW2-WW3 pETM11 construct via PCR site-directed mutagenesis by Eric Aragon as described in chapter 3.1.2. Point mutations are indicated in red.

Table 3.4: Sequences of the Nedd4L WW2-WW3 mutants.

Nedd4L WW2-3 mutants			
Mutation	**Sequence**	**ε [M^{-1} cm^{-1}]**	**Mw [g/mol]**
V384Y	GAMATPGLPS GWEERKDAKG RTYY**Y**NHNNR TTTWTRPIMQ LAEDGASGSA TNSNNHLIEP QIRRPRSLSS PTVTLSAPLE GAKDSPVRRA VKDTLSNPQS PQPSPYNSPK PQHKVTQSFL PPGWEMRIAP NGRPFFYDHN TKTTTWEDPR LKFPVH	29450	17519.5
I496Y	GAMATPGLPS GWEERKDAKG RTYYVNHNNR TTTWTRPIMQ LAEDGASGSA TNSNNHLIEP QIRRPRSLSS PTVTLSAPLE GAKDSPVRRA VKDTLSNPQS PQPSPYNSPK PQHKVTQSFL PPGWEMRIAP NGRPFF**Y**DHN TKTTTWEDPR LKFPVH	27960	17455.5
I496Y R492A	GAMATPGLPS GWEERKDAKG RTYYVNHNNR TTTWTRPIMQ LAEDGASGSA TNSNNHLIEP QIRRPRSLSS PTVTLSAPLE GAKDSPVRRA VKDTLSNPQS PQPSPYNSPK PQHKVTQSFL PPGWEMRIAP NG**A**PFF**Y**DHN TKTTTWEDPR LKFPVH	27960	17370.3
I496Y R486A	GAMATPGLPS GWEERKDAKG RTYYVNHNNR TTTWTRPIMQ LAEDGASGSA TNSNNHLIEP QIRRPRSLSS PTVTLSAPLE GAKDSPVRRA VKDTLSNPQS PQPSPYNSPK PQHKVTQSFL PPGWEM**A**IAP NGRPFF**Y**DHN TKTTTWEDPR LKFPVH	27960	17370.3

3.1.5 Cloning of the Pin1 WW domain

The Pin1 WW domain (1-41) was cloned via synthetic (template free) PCR (Eric Aragon) that involves the use of partially overlapping, complementary oligonucleotides. The non-overlapping part of the oligonucleotide is used by the *Taq* DNA polymerase (Fermentas Life Sciences, Burlington, Canada) as a template for complementary expansion of the DNA chain. The DNA is built up from the centre outwards. The DNA clone was introduced into the pETM30 vector by ligation reaction. This approach was used to remove some rare codons present in this insert.

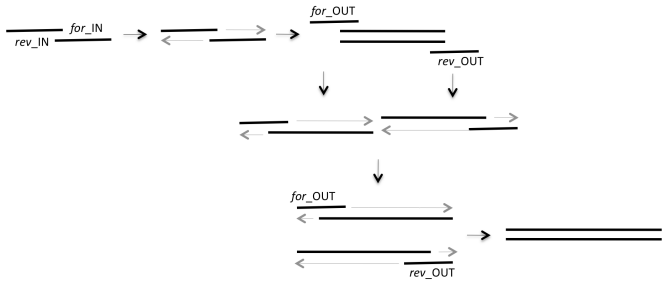


Figure 3.3: Synthetic PCR using four primers: forward/reverse inside (for_/rev_IN) forward/reverse outside (for_/rev_OUT).

3.1.6 Cloning of the YAP/TAZ WW domains

cDNA templates for YAP (WW1 163-206, WW2 227-266) were obtained from the Memorial Sloan-Kettering Cancer Center, New York (J. Massagué's group) while the TAZ WW domain (116-160) was cloned from cDNA. Inserts were introduced in pETM30 or pETM11 (Eric Aragon).

Table 3.5: Sequence of Pin1 WW, YAP WW1/2 and TAZ WW.

Protein	Sequence	ε [M^{-1} cm^{-1}]	Mw [g/mol]
Pin1 WW	GAMADEEKLP PGWEKRMSRS SGRVYYFNHI TNASQWERPS GNS	13980	4928.4
YAP WW1	GAMEIPDDVP LPAGWEMAKT SSGQRYFLNH IDQTTTWQDP RK	12490	4789.3
YAP WW2	GAMGPLPDGW EQAMTQDGEI YYINHKNKTT SWLDPRLDPR	13980	4603.1
TAZ WW	GAMGQSYDVT DELPLPPGWE MTFTATGQRY FLNHIEKITT WQDPRKAMN	13980	5644.3

DNA sequencing was carried out by the "Unidad de Genómica, Servicios Científico-Técnicos", university of Barcelona.

3.2 Expression and purification of protein domains

3.2.1 Transformation

The transformation of the expression vector into the plasmid DNA of the host cell was accomplished by the heat shock method. 1 µl DNA clone was added to 100 µl competent BL21(DE3) *E. coli* cells (Novagen, Darmstadt, Germany) and cooled down on ice for 20 min before heating up to 42 ºC for 45 s without moving. Super optimal broth with catabolite repression (SOC medium) was added (300 µl) and the cells were grown for 45 min at 37 ºC with maximal stirring velocity of 1400 rpm. The cells were centrifuged and the SOC medium was removed. A small volume of around 100 µl SOC medium was left to resuspend the cell

pellet and to disperse it on an agar plate. The agar plate was incubated overnight at 37 ºC to obtain colonies.

3.2.2 Recombinant protein expression

Unlabelled, ^{15}N-labelled, ^{13}C-, ^{15}N-labelled and ^{13}C, ^{15}N and ^{2}H-labelled proteins were expressed in E. coli BL21 (DE3) in LB medium or minimal medium (M9), using either $^{1}H_2O$ or $^{2}H_2O$ (99.89%, CortecNet) enriched with $^{15}NH_4Cl$ and/or D-[^{13}C] glucose as the sole sources of carbon and nitrogen respectively (Marley et al. 2001). Protein expression was initiated by adding isopropylthiogalactoside (IPTG) to a final concentration of 0.5 mM and carried out throughout the night. IPTG releases the lac-repressor, so that T7 polymerase and therefore also the target gene can be expressed. The expression was controlled by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

3.2.3 Cell lysis

After overexpression of the fusion proteins, cultures were centrifuged at 4 ºC at 3000 × g during 30 min. *E. coli* cells with transformed pETM30 vector were resuspended in phosphate buffered saline (PBS buffer), whereas cells containing the pETM11 vector were resuspended in Ni^{2+}-column buffer A. Solutions were lysed either with a French press or by sonication and centrifuged for 20 min at 4 ºC and 23000 × g.

3.2.4 Protein purification

Expressed fusion proteins containing a His_6-tag-GST (pETM30 vector) were purified using a GST-affinity-column and eluted with GST elution buffer, as described by the manufacturer. The fusion protein was incubated over night at 4 ºC with the tobacco etch virus protease (TEV to cleave the fusion protein from the protein of interest (Kapust et al. 2002). The reaction was finished using a protease inhibitor phenylmethanesulphonyl fluoride (PMSF, 1 mM final concentration). The GST was removed using a second purification step with a Ni^{2+}-affinity-column. Finally, the protein of interest was purified by gel filtration on a Superdex^{TM} 30/75/200 column (GE Healthcare Life Science, Uppsala, Sweden). Fusion proteins containing only the His_6-tag (pETM11 vector) were purified with the Ni^{2+}-affinity-column. The fusion protein bound to the Ni^{2+}-affinity-column was first washed with Ni^{2+}-column washing buffer containing a NaCl concentration of 1.0 M. Then the protein was eluted from the column with the Ni^{2+}-column elution buffer containing 50 mM ethylenediaminetetraacetic

acid (EDTA). Expression and cleavage of the recombinant proteins was monitored by SDS-PAGE and mass spectrometry.

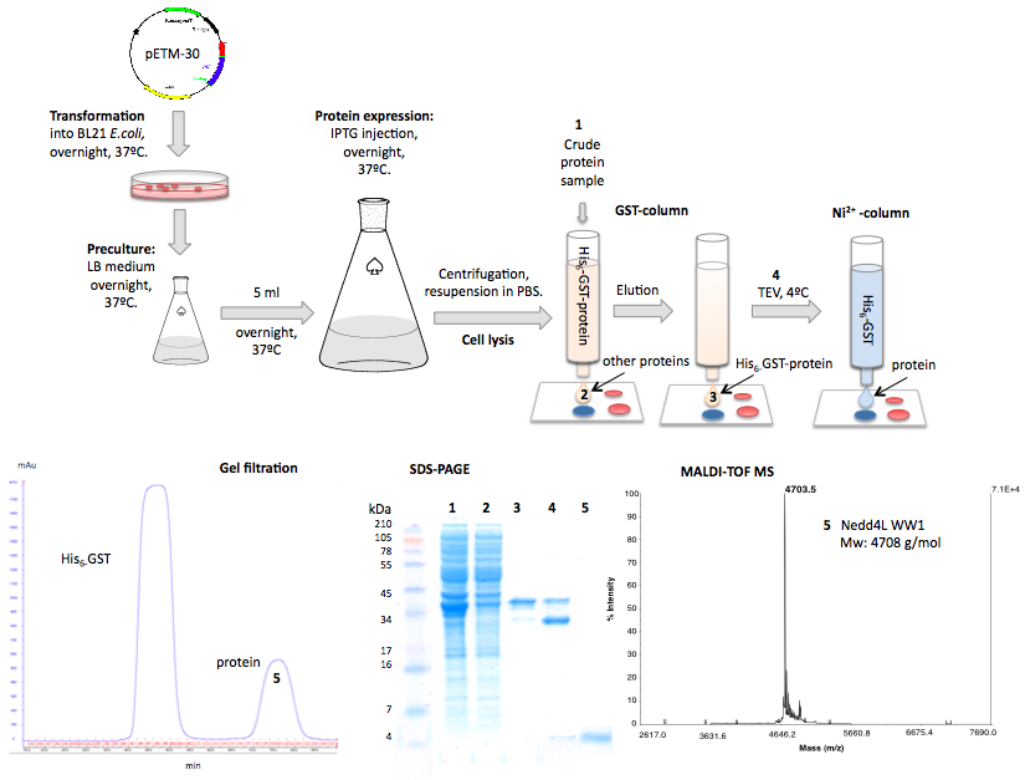


Figure 3.4: Above: Scheme of recombinant protein (Nedd4L WW1) expression and purification by GST- and Ni^{2+}-column. Below: gel filtration purification profile, SDS-PAGE and MALDI-TOF MS spectra of Nedd4L WW1 purification.

3.2.5 Determination of protein concentrations

Protein concentrations were determined by measuring the UV absorption of the protein sample on a Shimadzu UVmini-1240 UV-VIS spectrophotometer. Radiation absorption by proteins in the near UV depends on the amount of the aromatic amino acids Tyr with the absorption maximum at 275 nm and Trp with the absorption maximum at 280 nm. Since each Nedd4L WW domain consists of two Trp residues, the protein concentrations can be calculated by applying the Lambert-Beer law (3.1) after measuring the absorption at 280 nm.

$$C = \frac{A_{280}}{\varepsilon \cdot d} \tag{3.1}$$

C: concentration, A_{280}: absorption at 280 nm, ε: extinction coefficient [M^{-1} cm^{-1}], d: length of the cuvette (1 cm).

3.2.6 Buffers and solutions/material I

Cell culture reagents were purchased from Conda (Madrid, Spain) all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) or Merck (Darmstadt, Germany).

Buffers and solutions I

LB medium (1 l)	
Bactotryptose	10 g
Yeast extract	5 g
NaCl	10 g
Kanamycin (25 mg/ml)	1 ml
	(Autoclaved)

LB-agar plates	
Agar in autoclaved	
LB medium	1.5% (w/v)
Kanamycin	25 μg/ml

PBS buffer 10x (1 l) pH 7.4	
NaCl	80 g
KCl	2 g
$Na_2H\ PO_4$	14.4 g
KH_2PO_4	2.4 g

GST elution buffer (50 ml) pH 8	
L-Glutathion	10 mM
Tris-HCl pH 8.0	50 mM
NaCl	100 mM
	(Always prepare fresh)

SOC medium (500 ml)	
Bactotryptose	10 g
Yeast extract	2.5 g
NaCl [5M]	1 ml
KCl [1M]	1.25 ml
$MgCl_2$ [1M]	5 ml
$MgSO_4$ [1M]	5 ml
Glucose [1M]	10 ml
	(Autoclaved)

Trace elements 100x (1 l)	
EDTA	5 g
$FeCl_3 \times 6\ H_2O$	833 mg
$ZnCl_2$	84 mg
$CuCl_2 \times 2\ H_2O$	13 mg
$CoCl_2 \times 6\ H_2O$	10 mg
H_3BO_3	10 mg
$MnCl_2 \times 6\ H_2O$	1.6 g
	(Adjust pH to 7.5 after adding EDTA)

M9 medium 10x (1 l)	
$Na_2H\ PO_4$	60 g
KH_2PO_4	30 g
NaCl	5 g
$^{14}NH_4Cl/^{15}NH_4Cl$	5 g

M9 wash 10x (1 l)	
$Na_2H\ PO_4$	60 g
KH_2PO_4	30 g
NaCl	5 g

Buffers and solutions I			
M9 medium 1x (1 l)		^{13}C-labelled M9 medium 1x (1 l)	
M9 medium 10x	100 ml	M9 medium 10x	100 ml
Trace elements 100x	10 ml	Trace elements 100x	10 ml
Glucose 20%	20 ml	$^{13}C_6$-glucose 20%	10 ml
MgCl or $MgSO_4$ [1M]	1 ml	MgCl or $MgSO_4$ [1M]	1 ml
$CaCl_2$ [1M]	300 µl	$CaCl_2$ [1M]	300 µl
Biotin [1mg/ml]	1 ml	Biotin [1mg/ml]	1 ml
Thiamin [1mg/ml]	1 ml	M9 medium 10x	100 ml
Kanamycin [1mg/ml]	1 ml	Trace elements 100x	10 ml
Ni^{2+}-column buffer A		Ni^{2+}-column washing buffer	
NaCl	150 mM	NaCl	1 M
Tris-HCl	20 mM	Tris-HCl	20 mM
Imidazole pH 8.0	10 mM	Imidazole pH 8.0	10 mM
Ni^{2+}-column elution buffer		Tris-HCl buffer	
NaCl	150 mM	NaCl	100 mM
Tris-HCl	20 mM	Tris-HCl	20 mM
Imidazole pH 8.0	10 mM	NaN_3	500 mM
EDTA	50 mM		Adjust pH
Sodium phosphate buffer		Bis-tris buffer	
NaCl	100 mM	NaCl	100 mM
Na H_2PO_4	20 mM	Bis-Tris	20 mM
Na_2HPO_4	20 mM	NaN_3	500 mM
NaN_3	500 mM		(Adjust pH)

3.3 Peptide synthesis

Solid phase peptide synthesis (SPPS) was developed in the 1960s by R.B. Merrifield (Nobel laureate in 1984) (Merrifield 1962; Merrifield 1963). The basic concept is the growth of a peptide on a solid polymer through repeated cycles of amino acid coupling and the removal of an N-terminal protection group. SPPS is typically carried out in direction from the C- to the N-terminus using amino acids with two sets of protection groups, "temporarily" protected at the N-terminus and "permanently" protected at the side chain. A great deal of research has been devoted to improve the solid support system, coupling reagents, protection groups and

solvents. Fmoc (9*H*-fluoren-9-ylmethoxycarbonyl), first described as "temporary" protection group in SPPS by L.A. Carpino (Carpino 1972), is currently widely used due to its stability in acidic conditions while is easy to remove under mild basic conditions, such as 20% piperidine in dimethylformamide (DMF). Amino acids Arg, Asn, Asp, Cys, Glu, Gln, His, Lys, Ser, Thr, Trp and Tyr are additionally protected at their side chain through "permanent" protection groups 2,2,4,6,7-pentamethyldihydro-benzofuran-5-sulfonyl (Pbf), *t*-butoxycarbonyl (Boc), *tert*-butyl (tBu) or trityl (Trt) to prevent amino acid polymerisation during synthesis. These protection groups are stable in basic conditions but easy to release at low pH so that they are removed only in the final deprotection step. The amino acids are coupled through a nucleophilic attack of the deprotected amine group (N-terminus of aa_{x+1}) at the carbonyl group of aa_x. To accelerate the nucleophilic substitution the carboxylic acid has to be activated by additives like carbodiimides or triazolos. Triazolos like *N*-hydroxy-9-azabenzotriazole (HOAt) or *N*-hydroxybenzotriazole (HOBt) are favoured over carboiimides, which can cause racemization of the amino acid. Triazolos form an active ester, which is less reactive than carboiimides and have a lower tendency to racemize. Other triazolos like *N*,*N*-methylmethan-aminium hexafluorophosphate (HATU) or benzotriazol-1-yl-*N*-oxy-tris(pyrrolidino)phosphonium hexafluorophosphate (PyPOB) avoid the carbodiimids completely by introducing the active ester as a phosphonium salt of a non-nucleophilic anion. Combinations of HATU and HOAt or HOBt are also successfully applied. Final cleavage of the peptide from the resin and deprotection of the side chain protection groups is performed in one step under acidic conditions using 95% trifluoroacetic acid (TFA), 2.5% deionized water and 2.5% triisopropylsilan (TIS) as a scavenger to quench reactive carbocations (Pearson D A 1989). After cleavage the peptides were precipitated by the addition of cold *tert*-butylmethylether and the peptide taken up in 20-30 ml aqueous acetonitrile (50%). The crude peptides were either directly purified by reverse phase high performance liquid chromatography (RP-HPLC) or stored at -20 °C after lyophilisation. SPPS was done manually or automatically with and without microwave-assistance. All peptides synthesized for this work have been acetylated before cleavage and have the following composition:

$$H_3CCONH\text{-}aa_1\text{-}aa_{1+x}\text{-}COOH$$

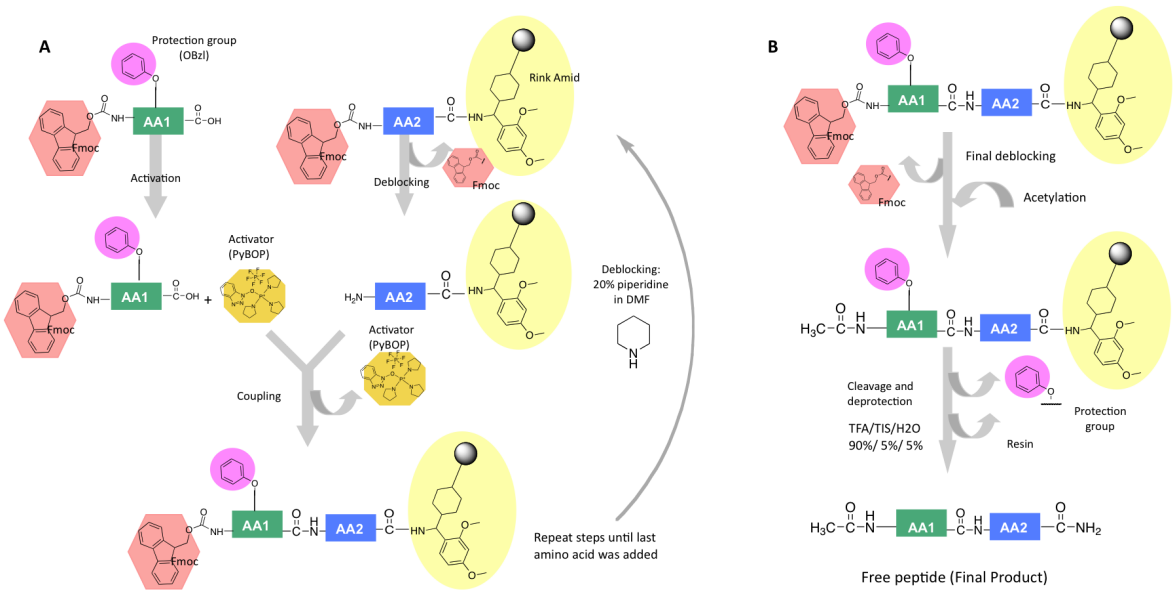


Figure 3.5: Fmoc-SPPS: (A) Fmoc-SPPS coupling step. AA2 as the last amino acid in sequence is attached to the Rink Amide AM resin. Fmoc group (AA2) is removed prior to AA1 coupling with PyBOP as activator. AA1 side chain is protected by OBzl. (B) Final Fmoc removal, peptide cleavage from the resin and deprotection of AA1.

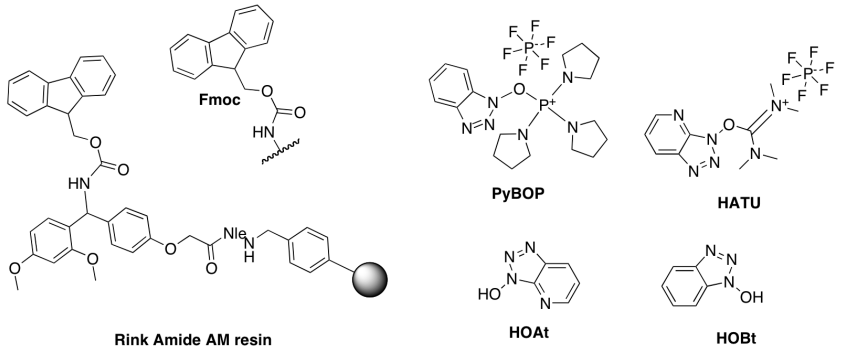


Figure 3.6: Chemical structures of Fmoc-SPPS compounds: Rink Amide AM resin, the Fmoc-group and activators PyBOB, HATU, HOAt and HOBt.

3.3.1 Manual Fmoc-solid phase peptide synthesis

Small peptides (up to 10 aa) were manually synthesized on a 0.1 mmol scale (100 mg, 0.068 mmol, substitution: 0.68 mmol/g) Rink Amide AM resin (200-400 mesh), a polystyrene based matrix supplied with the Rink Amide linker. The resin was loaded into a polypropylene syringe fitted with a polyethylene porous disk. After previous washing (3 x 1 min) with DMF, the Fmoc-group was removed with 20% piperidine in DMF for 15 min. Coupling was performed during up to one hour with a six-fold excess of Fmoc protected L-amino acids

(Fmoc-aa-OH). As coupling reagents 0.45 M HATU or PyBOP in DMF and 12 equivalence of Hünig's base *N,N*-diisopropylethylamine (DIEA) were used. DMF washes (3 x 1 min) were preformed between deprotection and coupling steps. After coupling the final aa (first in sequence) and removal of the Fmoc group, the unprotected N-terminus was acetylated for 30 min with 1 ml acetic anhydrid solution (50 µl Ac_2O, 85 µl DIEA, 865 µl DMF). Peptides were cleaved from the resin with 5 ml of a 95% TFA solution containing TIS and H_2O (1:1) during up to 4 h. Cold *tert*-butylmethylether was slowly added to precipitate the peptides. Before lyophilization of the crude peptide, the solution was centrifuged for 30 min at 4 ºC and 4000 rpm, the supernatant decanted and the remaining ether evaporated with nitrogen. Coupling steps and the removal of the Fmoc group were monitored with the colorimetric ninhydrin (Kaiser)-test, which is based on the reaction of ninhydrin with primary amino groups forming a blue adduct (Sarin et al. 1981). A saturated chloranil solution in toluol was used for the detection of secondary amines (Vojkovsky 1995). Solvents and soluble reagents were removed by suction.

3.3.2 Automatic Fmoc-solid phase peptide synthesis

Peptides in the 10-20 aa range were synthesized automatically with or without microwave assistance. Automated peptide synthesis without microwave assistance was performed on an automated synthesizer on a 0.1 mmol scale (147 mg) Rink Amid resin (substitution: 0.68 mmol/g), 10 equivalence of Fmoc-aa-OH. Coupling was carried out for 30 min with a mixture of HATU and HOAt (0.45 mM) in DMF in presence of DIEA. Deprotection was repeated four times and each step was monitored by a conduction profile of the Fmoc leaving group. Acetylation and final cleavage was carried out as described in chapter 3.3. ChemMatrix (0.1 mmol scale, substitution: 0.62 mmol/g) was used as solid support for microwave assisted Fmoc-SPPS. ChemMatrix is a 100% polyethylene glycol (PEG)-based polymer cross-linked through primary ether bonds, which allows high chemical and thermal stability and very good loading (Garcia-Martin et al. 2006). Unlike the polystyrene based resin, the ChemMatrix does not contain any linker. After previous washing with DMF (3 x 1 min) and dichloromethane (3 x 1 min) the Fmoc-Rink Amide linker (three equivalence) was incorporated manually with PyBOP (three equivalence) and DIEA (six equivalence) for two hours. The synthesis was carried out with a five-fold molar excess of 0.2 M Fmoc-aa-OH dissolved in DMF. A mixture of HATU and HOBt (0.5 M) in DMF was used as coupling reagents in presents of DIEA. Coupling reactions were preformed for five minutes at 40 W and a maximum temperature of

80 ºC and controlled with the software *Pepdriver*. Microwave energy has been adopted in SPPS during the last 20 years to enhance coupling reactions (Hui-Ming Yu 1992). It activates directly any molecule with a dipole moment and allows rapid heating at the molecular level (Palasek et al. 2007).

3.3.3 Synthesis of phosphopeptides

Phosphopeptides were synthesized as described above using phosphorylated serine (Fmoc-Ser(PO(OBzl)OH)-OH) or phosphorylated threonine (Fmoc-Thr(PO(OBzl)OH)-OH). Depending on the peptide length either the manual or automatic approach was used.

Table 3.6: Peptides synthesized by Fmoc SPPS.

Peptide	Sequence	Mw [g/mol]
Smad3 linker region	PETPPPGYLSEDG	1399.5
	PEpT$_{179}$PPPGYLSEDG	1479.5
	PEpT$_{179}$PPPGYLSEDGETSD	1911.9
	IPEpT$_{179}$PPPG	927.9
	MDAGpS$_{204}$PN	811.7
	AGpS$_{204}$PNLpS$_{208}$PNP	1154.0
	NLpS$_{208}$PNPMpS$_{213}$PA	1228.2
	IPEpT$_{179}$PPPGYLSEDGETSDHQMNHSMDAGpS$_{204}$PNLpS$_{208}$PNP	4102.1
	IPEpT$_{179}$PPPGYLSEDGETSDHQMNHSMDAGSPNLpS$_{208}$PNP	4022.1
Smad2 linker region	PETPPPGYISEDG	1399.5
	PEpT$_{220}$PPPGYISEDG	1479.5
Smad7 linker region	ELESPPPPYSRYPMD	1819.0

3.3.4 Peptide purification

All peptides were purified using reverse phase high performance liquid chromatography (RP-HPLC) on a ÄKTApurifier10 (GE Healthcare Life Science, Uppsala, Sweden) with 215 nm (to detect the amide bond), 280 nm (to detect Trp and Tyr aromatic side chains) or with 256 nm (Phe) UV detection. Analytic HPLC was performed at a flow rate of 1 ml/min on a Vydac C18 column (5 μm particle size, 4.6 x 250 mm) packed with silica particles chemically bonded with a short, branched alkylsilane. This rigid and sterically hindered stationary phase is stable at low pH and high temperatures. Semipreparative HPLC was performed at a flow rate of 5 ml/min on a preparative Vydac C18 or C4 Sephasil column (15-20 μm particle size, 300 Å pore size, 250 x 22 mm). All runs used linear gradients of 5% acetonitrile in deionized water and 0.05%

TFA (solvent A) vs. 70% acetonitrile plus 0.05% TFA (solvent B). For small peptides (7-15 aa) a gradient of around 15-40% solvent B over 20 min was used; for peptides up to 40 aa a flat gradient over 30-40 min was created with UNICORN software. All buffers and solvents were degassed using Durapore® membrane filters (0.22 μm pores, Milipore, Billerica, MA, USA) before use. Samples were filtered prior injection through polyetrafluoroethylene syringe filters (0.2 μm pores). Mass spectrometric analysis was routinely applied to all synthetic peptides.

3.3.5 Mass spectrometry

Matrix-assisted laser desorption ionisation - time of flight - mass spectrometry (MALDI-TOF-MS) was introduced as an ionization method in the late 1980s (Karas and Hillenkamp 1988). Over the years it has emerged as a powerful analytical technique that permits the mass analysis of higher molecular weight compounds and was performed to measure the mass of both expressed proteins and synthesized peptides (Zenobi 1998). Spectra were acquired with a MALDI-TOF/TOF MS 4700 Proteomics Analyser (Applied Biosystems, Foster City, CA, USA) and processed and analysed with Data Explorer 4.8 software. Dried droplet sample spotting was performed with peptide and protein samples (0.5-1 μl, ~10 pmol/μl) and 1 μl matrix on a MALDI sample plate (Applied Biosystems, Foster City, CA, USA). Nonphosphorylated peptides were co-crystallized with α-cyano-4-hydroxycinnamic acid (CHCA) and a smaller volume of peptide, phosphopeptides were co-crystallized with 2,5-dihydroxybenzoic acid (DHB). Sinapic acid (SA) was used as matrix to analyse protein masses (Kussmann Martin 1997). MS spectra were acquired in positive reflector mode. The negative reflector mode was used occasionally for measuring the mass of phosphopetides. The linear mode was used for masses > 5 kDa. Typically 500 shots per spectrum were accumulated.

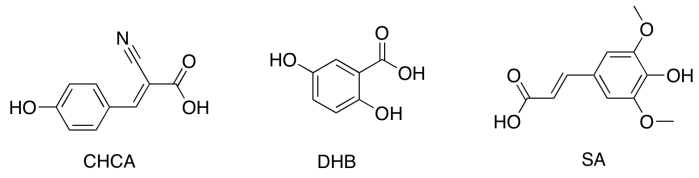


Figure 3.7: Chemical structure of MALDI matrices. CHCA used for peptides and proteins up to 5 kDa, DHB for phosphopeptides and SA for proteins > 5 kDa.

3.3.6 Determination of peptide concentrations

Determination of peptide concentrations was performed after dissolving the solid peptides in water and neutralizing with 1 M sodium hydroxide. Concentrations of peptide samples

containing Tyr were determined by applying the Lambert-Beer law on UV spectroscopy measurements as described in chapter 3.2.5. The concentration of peptide samples lacking Tyr where determined by amino acid analysis carried out at the "Unidad de Técnicas Separativas de Análisis, Servicios Científico-Técnicos", University of Barcelona. The procedure basically includes a peptide hydrolysis in acidic conditions (6 M hydrochloric acid, 24 h, 110 ºC) followed by a separation of the amino acids by a HPLC system equipped with a fluorescence detector. Met undergoes oxidation during this process. Gln and Asn are deamidated, resulting in Glu and Asp, respectively.

3.3.7 Buffers and solutions / material II

All Fmoc amino acids and the Rink Amid resin were purchased from Novabiochem (Laufelfingen, Switzerland). All other chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) or Merck (Darmstadt, Germany).

Buffers and solutions II

HPLC buffer A	
ACN in H_2O	5% (v:v)
TFA	0.05%

HPLC buffer B	
ACN in H_2O	70% (v:v)
TFA	0.05%

Acetic anhydrid solution (1 ml)	
Ac_2O	50 µl
DIEA	85 µl
DMF	865 µl

Acidic cleavage solution	
TFA	95.0%
H_sO milliQ	2.5%
TIS	2.5%

CHCA matrix MALDI-TOF	
CHCA	20 mg/ml
ACN in H_2O	50%
TFA	0.1%

DHB matrix MALDI-TOF	
DHB	20 mg/ml
ACN in H_2O	50%
Ortho-phosphoric acid	2.5%

SA matrix MALDI-TOF	
SA	20 mg/ml
ACN in H_2O	50%
TFA	0.1%

3.4 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a thermodynamic method widely used for the characterisation of biomolecular interactions such as protein-protein or protein-ligand interactions. When substances interact the binding energy in form of heat is either absorbed (endothermic reaction) or released (exothermic reaction). The measurement of this heat allows accurate determination of association constants (K_A), reaction stoichiometry (N), enthalpy (ΔH) free energy (ΔG) and entropy (ΔS). The thermodynamic parameters are related to each other by the following equations (3.2) and (3.3).

$$\Delta G = \Delta H - T\Delta S \tag{3.2}$$

$$\Delta G = -RT \ln K_A \tag{3.3}$$

The calorimeter consists of two highly efficient thermal conducting cells, a reference cell and the sample cell (volume: 1.43 ml both) surrounded by an adiabatic jacket to avoid any heat exchange with the environment. Sensitive thermocouple circuits detect temperature differences between the two cells and the jacket. To maintain identical temperatures heaters are located on all three components to be activated when necessary. The reference cell is applied with a constant power (15 μcal/s, used for all experiments) that directs a feedback circuit to activate the heater located on the sample cell during each injection. The peptide (300 μl, 0.4-7.0 mM) was titrated into the sample cell containing the protein solution (0.02-0.40 mM) and the time-dependent input of power required to maintain equal temperatures in the sample and reference cell is observed.

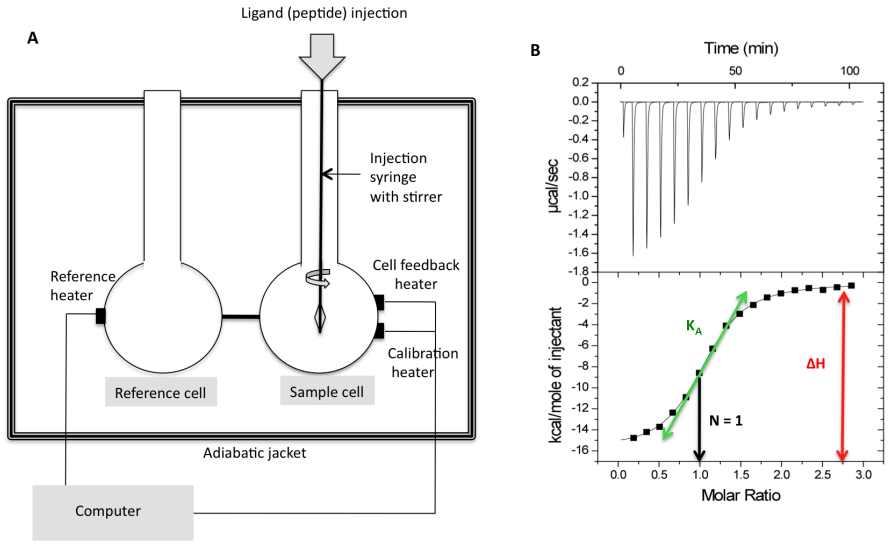


Figure 3.8: (A) Scheme of the ITC instrument. (B) ITC profile and fitting curve with association constant K_A, stoichiometry N and reaction enthalpy ΔH.

Concentrations of protein and peptide samples were calculated using equation (3.4) with *c:* unitless constant (values between 1-1000 are recommended), K_A: approximately expected binding constant in M^{-1}, [M]: sample concentration in the cell and N: stoichiometry of the reaction.

$$c = K_A[M]N \tag{3.4}$$

Large *c* values result in very sharp transition; saturation may be achieved in a single injection of ligand. Binding isotherms with low *c* values are very broad transitions that approach linearity, which makes it difficult to identify the equivalence point (Pierce et al. 1999). ITC experiments were conducted at 15 and 25 ºC with a MicroCal™ VP-ITC system (GE Healthcare Life Science, Uppsala, Sweden). Injections of 10 µl of ligand solution were added from a computer-controlled micro-syringe (volume: 300 µl) at an interval of four minutes into the sample solution of the binding partner with stirring at 300 rpm (up to a three- to four-fold final ligand excess). A control ITC measurement was done by a titration of the peptide into the sample buffer. Dissociation constant, entropy, enthalpy and standard deviations were derived after subtracting the raw data from the control measurement. Automatic subtraction with the MicroCal™ software led to the same results. Therefore the automatic approach was used in all other ITC experiments. All steps of the data analysis were performed using the software ORIGIN 7.0. Data were fitted assuming a *single set of sites* to calculate the binding affinity K_A. All binding affinities in this work are given as dissociation constants K_D, the reciprocal value of

the binding constant K_A, equation (3.5). The error ΔK_D was calculated by error propagation with Δ: absolute error, equation (3.6).

$$y = \frac{1}{x} \Rightarrow K_D = \frac{1}{K_A} \tag{3.5}$$

$$\Delta y = \frac{dy}{dx} \cdot \Delta x \Rightarrow \Delta K_D = -\frac{\Delta K_A}{K_A^2} \tag{3.6}$$

The protein-ligand reaction product remaining in the sample cell could be recycled by reverse phase HPLC purification using a Vydac analytic C18 column. Two clearly separated peaks could be monitored corresponding to the peptide (first peak) and the protein (second peak). The purified peak fractions were characterized by MALDI-TOF mass spectrometry.

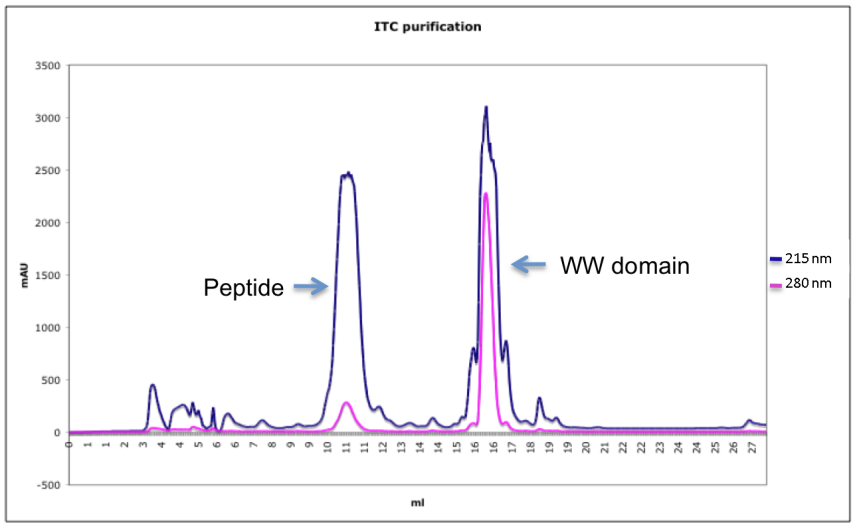


Figure 3.9: RP-HPLC purification of the protein-peptide complex after the ITC experiment. Detection at 215 and 280 nm.

3.5 Chemical ligation of the Nedd4L WW2-WW3 fragment

To investigate the structural characteristics of this protein pair a segmental labelling approach was selected that requires the protein being expressed as two fragments; each fragment is produced with a different labelling pattern, and then combined by chemical ligation. For the ligation we used the formation of a disulfide bridge built from the side chains of two cysteine residues, one at the C-terminus of fragment 1 (WW2-connector-C471) and the other at the N-terminus of fragment 2 (C472-WW3) using mutant I496Y as template. In order to guide the reaction towards the formation of hetero-disulfide WW2-WW3 product, the C472-WW3 site was activated prior to the ligation reaction with 2,2'-dipyridyl disulfide (Baca et al. 1995). The

pyridylsulfenyl-cysteine-WW3 product was identified as a new peak with a higher retention time in the RP-HPLC profile of the reaction, and its mass was corroborated by MALDI-TOF mass spectroscopy. The product was purified by semi-preparative RP-HPLC and lyophilized. $^{2}H,^{15}N$-labelled Nedd4L WW2-linker-C471 (0.4 µmol) and $^{1}H,^{15}N$-labelled Nedd4L PyS-(C472)-WW3 domain (0.3 µmol) were dissolved in 600 µl of 6 M guanidine-HCl and sodium acetate buffer (20 mM, 100 mM NaCl pH 4.5), and stirred overnight at 4 ºC. The ligation-product was identified by RP-HPLC/MALDI-TOF MS. After purification, the product was lyophilized.

3.5.1 Cloning of the Nedd4L WW2-linker-Cys and Cys(Gly)-WW3 constructs

Nedd4L WW2-linker-Cys was cloned via PCR site-directed mutagenesis (chapter 3.1.2) using the Nedd4L WW2-WW3 construct in pETM11 (chapter 3.1.4) as template. A stop code was added directly downstream of the Nedd4L WW2-linker-Cys sequence to prevent the expression of Nedd4L WW3 in *E.coli*. Nedd4L Cys(Gly)-WW3(I496Y) mutant was cloned via PCR site-directed mutagenesis (chapter 3.1.2) using the Nedd4L WW3 construct in pETM30 (chapter 3.1.1) as template. The amino acid Gly preceding the Cys residue in the beginning of the Nedd4L WW3 sequence remains after removing the His_6-tag with TEV protease (chapter 3.2.4). The His_6-tag of the Nedd4L WW2-linker-Cys domain was not removed.

Table 3.7: Constructs for chemical ligation of Nedd4L WW2 and Nedd4L WW3. In blue His-tag, in red point mutations.

Nedd4L	Sequence	ε [M^{-1} cm^{-1}]	Mw [g/mol]
His_6-tag-WW2-linker-Cys ($^{2}H,^{15}N$)	**MKHHHHHHPM SDYDIPTTEN LYFQ**DAMATP GLPSGWEERK DAKGRTYYVN HNNRTTTWTR PIMQLAEDGA SGSATNSNNH LIEPQIRRPR SLSSPTVTLS APLEGAKDSP VRRAVKDTLS NPQSPQPSPY NSPKP**C**	15470	16445.8
Gly-Cys-WW3(I496Y) (^{15}N)	G**C**VTQSFLPP GWEMRIAPNG RPFF**Y**DHNT KTTTWEDPRL KFPVH	12490	5239.8
His_6-tag-WW2-linker-Cys-Cys(Gly)-WW3(I496Y) ($^{2}H,^{15}N/^{15}N$)	**MKHHHHHHPM SDYDIPTTEN LYFQ**DAMATP GLPSGWEERK DAKGRTYYVN HNNRTTTWTR PIMQLAEDGA SGSATNSNNH LIEPQIRRPR SLSSPTVTLS APLEGAKDSP VRRAVKDTLS NPQSPQPSPY NSPKP**CC**(G)VTQ SFLPPGWEMR IAPNGRPFF**Y** DHNTKTTTWE DPRLKFPVH	26470	21668.7

3.5.2 Activation of the Nedd4L Cys(Gly)-WW3 domain

The Nedd4L Gly-Cys-WW3 domain (2 μmol) in 6 M guanidine hydrochloride (GuHCl) solution of 1.8 ml acetoacetic acid (20 mM, 100 mM NaCl), pH 4.0 was slowly added to 22-fold excess of 2,2'-dipyridyl disulfide (10 mg, PySSPy, Sigma-Aldrich, St. Louis, MO, USA) dissolved in 400 μl 90% ACN/9.9% H_2O/0.1% TFA and stirred for 3 h at 25 ºC (Baca M. 1995). The reaction was monitored by analytic RP-HPLC with a gradient of 25-50% solvent B over 30 min. Activated Nedd4L Gly-Cys-WW3 was identified as a new peak at a longer retention time and purified on a semi-preparative C4 column. Finally, the product was lyophilized and the mass determined by MALDI-TOF MS (chapter 3.3.5).

NMR Spectroscopy

4 NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is one of the most important structure determination methods at atomic resolution based on the magnetic property of particular nuclei. NMR is used for structure determination of small molecules as well as macromolecules and can provide additional information about dynamics, binding, stability or protein folding. At the beginning of the 20th century the foundation of modern NMR spectroscopy was laid by the development of quantum mechanics leading to the discovery of the magnetic property of the nucleus. Max Planck (NP 1918) proposed in the year 1900 the quantization of energy $E = h\nu$ depending on the frequency ν and the Planck constant $h = 6.6 \times 10^{-34}$ Js to explain the experimental observed failure of the classic mechanic on atomic level. Twenty years later the nuclear spin was observed in the Stern-Gerlach experiment as a deflection of the particles in the presence of a magnetic field. In 1938 Isidor Isaak Rabi (NP 1944) published a method to detect nuclear magnetic resonance of ^{7}Li (Rabi 1938). Parallel works of Felix Bloch and Edward Mills Purcell (NP 1952) and the contribution of other scientists established the technique during the following decades. In 1964 Richard Ernst (NP 1991) introduced the pulsed Fourier transform (FT) technique that is nowadays used in modern NMR spectroscopy instead of the continuous flow technique of slow radio waves. The first three-dimensional structure of a globular protein by solution NMR has been determined 1984 by Kurt Wüthrich (NP 2002) (Williamson et al. 1985) who developed the sequential resonance assignment (chapter 4.6.1) used in this work. Up to date several new NMR techniques have been established like imaging solid and in vivo NMR with application in chemistry, biology, medicine and material science. Technical improvements led from the first 40 MHz magnet to a 1 GHz (2009) magnet which is located in Lyon, France. In this work multidimensional solution NMR was applied to determine three-dimensional structures of proteins domains and protein-peptide complexes and to study protein-peptide binding.

4.1 Principles

The principle ideas and equations of nuclear magnetic resonance are presented in this chapter. Each nucleus has a nuclear magnetic moment μ described by the gyromagnetic ratio

γ (10^7 rad s^1 T^{-1}) of the nucleus, the nuclear spin quantum number I and the reduced Planck constant ($h/2\pi$).

$$\mu = \gamma \cdot I \cdot \hbar \tag{4.1}$$

The nuclear spin quantum number of most nuclei is ½ and has $2I + 1$ non-degenerated energy levels in a magnetic field resulting in two spins with opposite orientation, called spin α (magnetic quantum numbers m=+1/2) and spin ß (m=-1/2).

Table 4.1: Examples of spin ½ nuclei and their gyromegnetic ratio γ (Hore P.J. 2000).

Spin ½	γ [10^7 rad s^{-1} T^{-1}]
^{1}H	26.7
^{15}N	-2.7
^{13}C	6.7
^{19}F	25.2
^{31}P	10.8

The energy E_m of the spin depends on the magnetic quantum number m, the magnetic field B_0 that is experienced by the nucleus and on the gyromagnetic ratio γ of the nucleus.

$$E_m = -m\hbar\gamma B_0 \tag{4.2}$$

The energy difference between the two states is described by

$$\Delta E = E_\beta - E_\alpha \tag{4.3}$$

Resonance occurs when the nucleus experiences energy of exactly this frequency, called Larmor frequency ω_0 (rad s^{-1}) or ν_0 in Hz.

$$\omega_0 = \gamma \cdot B_0 \tag{4.4}$$

$$\nu_0 = \frac{\gamma B_0}{2\pi} = \frac{\omega_0}{2\pi} \tag{4.5}$$

In the sample the nucleus experiences a local field B_0 that is slightly different to the applied magnetic field B. This difference is represented by the shielding constant σ, which has its origin in the different electron density depending on the structure of the sample.

$$B_0 = (1 - \sigma) \cdot B \tag{4.6}$$

The variation in the electron distribution results in different Lamor frequencies for the same nucleus at distinct positions in the molecule and is described by the chemical shift δ (in parts per million).

$$\delta = \frac{\omega - \omega_0}{\omega_0} \times 10^6 \tag{4.7}$$

The magnetic moments of the nuclei themselves can also affect each other by transferring magnetisation from one spin to another. Magnetisation can be transferred via through bond coupling mediated by the electrons participating in the bond that connects the two nuclei. The coupling is described by the coupling constant *J* (Hz). The value of *J* does not depend on the applied magnetic field and is very small compared to the chemical shift δ. Two spins can further interact with each other through space. This dipolar coupling provides important information for 3D structure determination (chapter 4.3.1).

The NMR signal can be generated and analysed in the most simple way by acquiring a one dimensional NMR experiment.

4.2 One-dimensional (1D) NMR experiments

In the applied magnetic field *B* there is a slight excess of spins in the lower energy level, which leads to a bulk magnetisation M_z of the sample parallel to *B* along the z-axis. In order to acquire a 1D NMR experiment a radio frequency (RF) pulse, which is an oscillating magnetic field along the x-axis, was applied. This 90°-pulse tilts the magnetisation of the sample from the z-axis into the x-y plane and provokes a big magnetisation because all spins are in phase. In the course of time the spins loose phase and the x-y magnetisation decays exponentially depending on the time. This loss of x-y magnetisation after a 90° pulse is called relaxation and has two reasons. One is the spin lattice or longitudinal relaxation, which is the return to thermal equilibrium with the magnetisation M_z around the z-axis depending on T_1. Time T_1 usually takes a few seconds but can also take microseconds, milliseconds, minutes or hours. The other reason is the transversal relaxation depending on T_2, which describes the return to its equilibrium value of zero in the x-y plane. In many experiments the pulses are so short that one can ignore the transversal relaxation while the RF field is present. During time T_2, which usually takes a few seconds, magnetisation is transferred from one spin to the other via

through-bond and through-space coupling. The decaying signal due to relaxation is measured after amplifying it and is called free induction decay (FID). The FID is an ensemble of all decaying frequencies due to the distinct chemical environments of the nuclei. To extract the different resonance frequencies out of the signal, the time dependent FID is transformed into a frequency dependent signal through Fourier transformation. Information about stability and folding of the protein and relative protein concentrations were obtained in this work from one-dimensional NMR spectra. For additional information on interactions between spins double or multiple irradiation experiments were carried out. This requires the use of a second irradiation of a particular resonance for a given time t. The extension of the measurements into a second or third dimension was necessary in order to determine peptide sequences and to solve 3D protein structures.

4.3 Two-dimensional (2D) NMR experiments

There is a basic scheme, which is used for all 2D NMR experiments. It consists of a preparation period, an evolution period, a mixing period and a detection period. On the axis the chemical shifts of the observed nuclei are plotted, the diagonal usually corresponds to the 1D spectrum and the cross-peaks indicate the existence of an interaction between two spins. The type of interaction depends on the experiment. In correlation spectroscopy (COSY) the cross peaks arise from through-bond correlations, while in nuclear Overhauser effect spectroscopy (NOESY) they arise from through-space correlations.

4.3.1 Nuclear Overhauser effect

NOE signals are due to through-space interactions (dipolar coupling) in which magnetisation is transferred through cross-relaxation between two nuclei. NOE measurements are extremely valuable for 3D structure determination because the intensity is proportional to the inverse-sixth power of the inter-nuclear distance. NOE intensities fall off rapidly with increasing distance between the nuclei. Nevertheless, proton distances up to 5 Å can be obtained by NOESY. The general concept is to disequilibrate one spin and to determine the effect of this on the z-magnetisation of the other spin. In double irradiation 1D NMR experiments NOE signals can be observed if the behaviour of the z-magnetisation of spin1 is affected by the amount of z-magnetisation of spin2. Therefore, a selective RF pulse at the position of the frequency of spin1 is applied. If spin2 is close in space its magnetisation is decreased. In multi-dimensional NOESY NOEs are observed as cross peaks (Evans 1995; Hore P.J. 2000; Keeler

2002). NOE spectra were used in combination with total correlation spectra for 3D structure determination of proteins carried out in this work as well as for sequential resonance assignment of peptides (chapter 4.6.1).

4.3.2 Total correlation spectroscopy

The total correlation spectroscopy (TOCSY) is a type of COSY experiment that creates correlation between all protons within a given spin system. For protein or peptide assignment it means that all protons of an amino acid can be identified. Correlations are only seen between protons of the same amino acid but not to protons of a different amino acid. Magnetisation is transferred from spin1 via spin2 to spin3 and vice versa from spin3 to spin1 via spin2. The transfer is usually interrupted by heteroatoms like oxygen. The key feature of the TOCSY experiment is the suppression of the chemical shift while J-coupling is acting. This is achieved through a process called spin locking. During this isotopic mixing the spins become temporarily equivalent and magnetisation is exchanged through J-coupling. The range of magnetisation transfer can be increased by changing the spin-lock time (mixing time) from a short spin-lock time of 20 ms to a long spin-lock time of 80 or 120 ms (Evans 1995; Hore P.J. 2000; Keeler 2002).

4.3.3 Heteronuclear single quantum coherence

The heteronuclear single quantum coherence (HSQC) experiment is a two-dimensional heteronuclear NMR experiment, which is widely used in protein NMR to identify all protons directly bound to a nitrogen atom. Since the protein backbone consists of amid bonds, each amino acid, except for proline, can be identified as a cross peak in the HSQC spectrum. The cross peaks represent one bond correlations between ^{15}N and ^{1}H. ^{15}N labelled protein samples are essential for this method. The 1D ^{1}H-spectrum is plotted on the x-axis against the 1D ^{15}N-spectrum on the y-axis. HSQC spectroscopy is a relatively rapid experiment to identify a characteristic protein pattern or to detect protein-peptide interactions. The unlabelled peptide can be titrated to the protein to study possible changes in the chemical shift of the peaks. A movement of the cross peaks in different directions is an indication for binding because the environment of the nuclei in the binding interface changes upon ligand binding. In this work titration experiments were carried out in several titrating steps up to a peptide excess of three-fold. Then the HSQC spectrum of the free protein was compared to the HSQC spectrum of

the protein sample plus the titrated peptide. Measurements were performed at 285 or 295 K with a protein concentration of 0.25 mM.

4.4 The magnet and signal detection

The NMR experiments were acquired on a Bruker AdvanceIII 600-MHz spectrometer of 18 Tesla equipped with a z pulse field gradient unit and either a triple (^{1}H, ^{13}C, ^{15}N) or quadruple (plus ^{31}P) resonance probe head. The superconducting wire is cooled to a temperature close to absolute zero, has a typically length of several kilometres and is made of $(NbTaTi)_3Sn$. To generate a homogeneous magnetic field shim coils that create small magnetic fields to cancel out an inhomogeneity are positioned within the magnet. The RF coils induce the RF pulse and also detect the transversal magnetisation to give the FID. All data acquired for this work were processed using NMRPipe/NMRDraw software (Delaglio et al. 1995) and analysed with XEASY/CARA (Bartels Christian 1995).

4.5 Sample preparation

For all NMR experiments aqueous sample solutions containing 10% D_2O for the lock signal were used with a protein concentration of about 1.0 mM for structural studies and sample volume of 300-500 µl. Samples in 100% D_2O solution were also prepared to acquire NMR spectra of nonexchangeable protons. These NMR experiments are useful to assign NOEs involving aromatic amino acids and help in the identification of resonances close to the water frequency. Deuterium incorporation of nonexchangeable protons was applied for Nedd4L WW2-linker-Cys domain to facilitate the assignment of the WW3 site and also for the WW2-WW3 sample used to assign the backbone of the pair.

4.6 Protein structure determination

The determination of a protein structure by solution NMR is a stepwise approach leading to the final 3D structure. Several multidimensional, homo- and hetero-nuclear NMR spectra of unlabelled and isotope-labelled protein samples need to be acquired under different conditions. The first step of the structure determining process is a sequential resonance assignment via combination of through-bond and through-space correlations. In the second step secondary structure elements like through-space connectivity between protons and torsion angles are determined. Based on these experimental restrictions the 3D structure is finally calculated and validated (Wuthrich 1990; Clore and Gronenborn 1998).

4.6.1 Sequential resonance assignment

The sequential resonance assignment refers to the primary protein structure, the protein backbone. Through heteronuclear 3D NMR techniques the H^N chemical shift of each residue can be correlated to the $^{13}C_A$ and $^{13}C_B$ chemical shifts of the same and the preceding residue. A ^{15}N and ^{13}C isotope labelled sample is crucial for this technique.

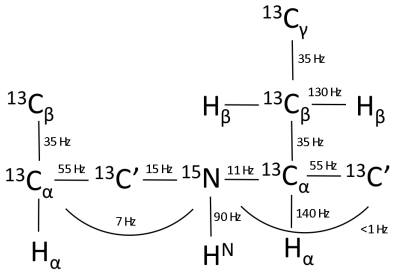


Figure 4.1: Spin system of the protein backbone with coupling constants.

In the HNCO experiment the magnetisation is transferred through heteronuclear J-coupling from the amide proton H^N through nitrogen (^{15}N) to the carbonylcarbon (^{13}CO) of the preceding residue. Is the magnetisation further transferred to the $^{13}C_A$ and $^{13}C_B$ carbon the experiment is called CACB(CO)NH (the chemical shift of the spins in brackets in not involved). The CBCA(CO)NH experiment obtains the chemical shifts $^1H^N$ and $15N^H$ of amino acid x and the chemical shifts $^{13}C_A$ and $^{13}C_B$ of amino acid x-1. The CBCANH experiment obtains the chemical shifts $^1H^N$ and $15N^H$ and $^{13}C_A$ and $^{13}C_B$ of amino acid x and the chemical shifts $^{13}C_A$ and $^{13}C_B$ of amino acid x-1 from the preceding residue. With the information of the two spectra the protein backbones were assigned sequentially. All other side chain carbon resonance assignments were obtained by the (H)CC(CO)NH experiment. Lateral protons and amid protons were assigned through 3D HCCH-TOCSY (12 ms mixing time), 3D ^{15}N NOESY and 3D ^{13}C NOESY experiments with 100-150 ms mixing times depending on the protein size (Evans 1995).

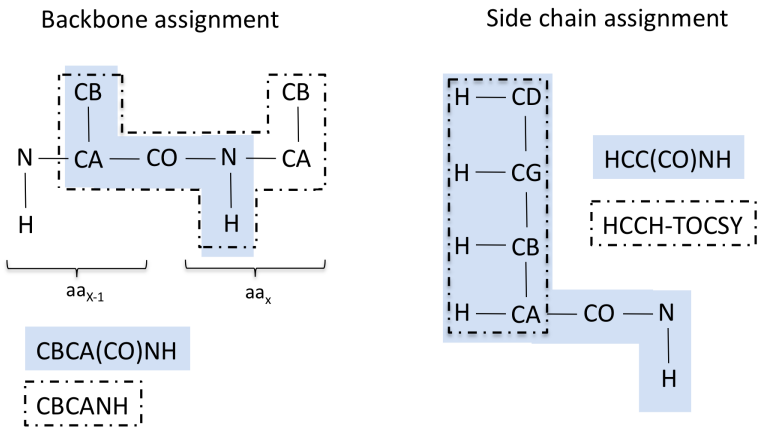


Figure 4.2: Protein assignment strategy for backbone and side chains.

Finally, the list of assigned atoms was transferred with an in-house script, written by Pau Martin, to the homonuclear 2D TOCSY (50 ms mixing time for single WW domains, 30 ms for WW domain pairs) and 2D NOESY (120-180 ms mixing time) spectra to identify ^{1}H-^{1}H the 3D structure determining NOEs. Peptides were assigned via NOE-based sequential resonance assignment using homonuclear 2D TOCSY and 2D NOESY spectra.

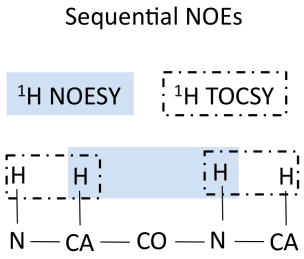


Figure 4.3: Sequential resonance assignment via homonuclear TOCSY and NOESY spectra.

4.6.2 Structure calculation and refinement

The final protein structures or protein peptide complex structures were determined using inter- and intra-molecular NOEs obtained from the previously described experiments. With the experimental data the structures were calculated using the version 1.2 of the crystallography & NMR system (CNS 1.2) (Brunger 2007) together with version 1.2 of the programme ambiguous restraints for interactive assignment (Aria 1.2) (Nilges et al. 1997) applied with the in-house script *Structcalc*, written by Pau Martin. All calculated structures were submitted to water refinement and were ranked based on minimum values of energy-terms and violations. All complex structures determined in this work consist of WW domains. Hence, secondary

structure elements like hydrogen bonds and dihedral information are equal to those previously published (Macias et al. 1996). Between 100-200 structures were calculated that represents a good fit of the experimental data. Of those the 20 lowest energy structures were selected and evaluated with the program PROCHECK-NMR (Laskowski et al. 1996).

Results

5 Results

5.1 Interaction of the R-Smad3 linker and WW domains

5.1.1 Binding of the Smad2/3 linker and Nedd4L WW domains

Nedd4L has been identified as Hect type of E3-ubiquitin ligase that selectively recognizes linker phosphorylated Smad2 and 3 in the TGF-ß pathway and triggers Smad2/3 degradation (Kuratomi et al. 2005; Gao et al. 2009). Based on these findings it has been suggested that Nedd4L regulates the TGF-ß signalling pathway in a similar manner as the E3 ubiquitin ligase Smurf1 in the BMP pathway (Zhu et al. 1999). The PY motif and one or more phosphorylated p(S/T)P sites within the Smad1 linker region are required for Smurf1 binding to Smad1 (Sapkota et al. 2007). Smad2 and 3 also contain a PY motif and four phosphorylation sites but arranged in a distinct manner with respect to Smad1 (Figure 1.3).

Nedd4L contains four WW domains all belonging to the group of PY binders. Since Smad2 and 3 contain a unique PY site the goal was to characterize which of the four potential WW candidates is the best binder in terms of affinity. To achieve this purpose independently expressed domains were prepared and a set of Smad3 peptides containing the PY motif were synthesized and titrated to each domain. ITC experiments were used to quantify the interactions.

To start with the search of the binding motif, a window in the sequence of 13 amino acids was selected with or without phosphorylation at T179 (Figure 5.1). The peptides were synthesized by manual Fmoc-SPPS as described in chapter 3.3.1.

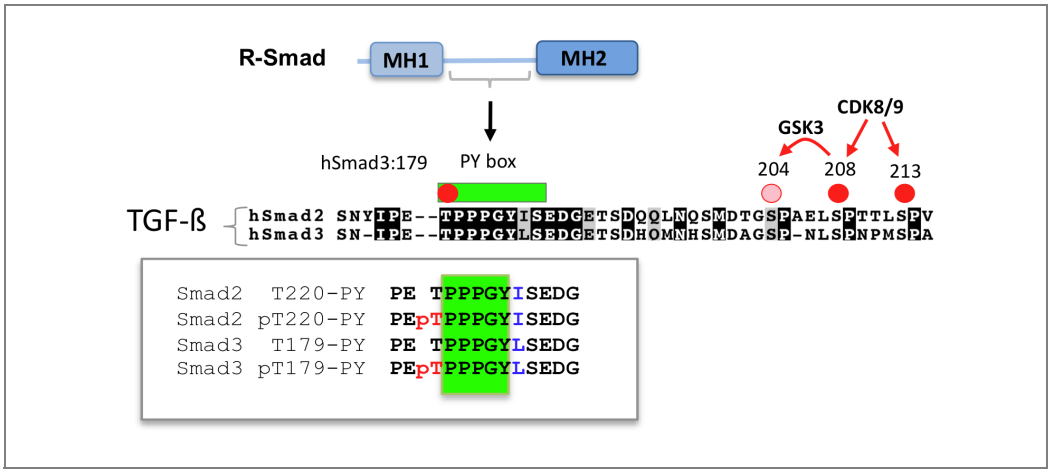


Figure 5.1: Linker sequence of hSmad2/3 with PY motif and phosphorylation sites. Box below: Name and sequence of the Smad2/3 linker peptides containing the (p)T[PY] motif synthesized by Fmoc-SPPS with the phosphorylated threonine in red, the PY motif in green and the amino acids I226 (Smad2) and L185 (Smad3) in blue.

During Fmoc-SPPS the first amino acid, the N-terminal proline of the polyproline motif and the phosphorylated threonine were usually coupled twice. During the purification step of the synthesized phosphopeptides incomplete deprotection of the phosphate group was detected by RP-HPLC. The reaction side product could be identified as an additional peak at longer retention time. Characterization by mass spectrometry indicated a difference to the peptide mass of Δm = 90 g/mol corresponding to the mass of the phosphate protection group OBzl. Deprotection was completely carried out by adding TFA/H_2O/TIS (95%/2.5%/2.5%) to the crude peptide. The remaining protection groups were removed using standard conditions.

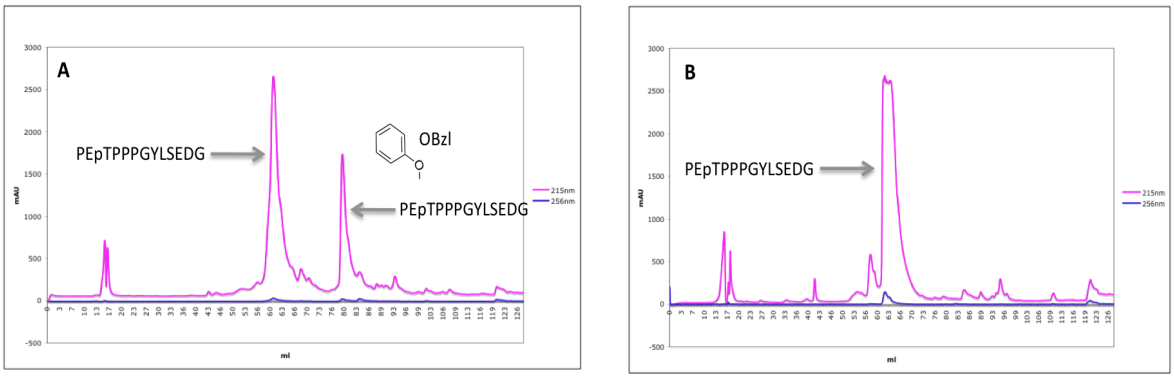


Figure 5.2: RP-HPLC purification profile with detection of 215 and 280 nm. (A) Peptide peak and peak of peptide with phosphate protection group at longer retention time. (B) Peptide peak after additional deprotection. Peak at longer retention time disappeared.

Smad2/3 peptides within the selected frame only differ in one amino acid: I226' in Smad2 which corresponds to L185' in Smad3. The obtained binding affinities of Smad2 (p)T220[PY]

and Smad3 (p)T179[PY] peptides towards each Nedd4L WW domain are almost identical, suggesting that the I to L change does not have a critical role on binding. The ITC experiments revealed that all Nedd4L WW domains bind the Smad2/3 PY motif with dissociation constants in the 10-200 μM range. Nedd4L WW2 and WW3 bind the unphosphorylated peptides with a K_D around 50 μM two- to three-fold stronger than WW1 and WW4 (K_D ~ 110-150 μM). Different results are obtained for the phosphorylated Smad2/3 pT[PY] peptides. Nedd4L WW2 binds the pT[PY] motif eight- to 12-fold stronger than the nonphosphorylated T[PY] motif with affinities of K_D = 6.7 ± 0.3 μM (Smad2) and K_D = 4.1 ± 0.3 μM (Smad3). Nedd4L WW3 domain binds the pT[PY] motif three-fold stronger (Smad2: K_D = 12.8 ± 0.6 versus 46.7 ± 4.4 μM, Smad3: K_D = 17.2 ± 1.7 versus 46.1 ± 2.4 μM). Nedd4L WW1 and WW4 bind weakly to both, the phosphorylated pT[PY] and nonphosphorylated T[PY] motifs at which Nedd4L WW4 even binds slightly weaker to the phosphorylated peptide (Table 5.1).

Table 5.1: Binding affinities between Nedd4L WW domains and Smad2/3 linker peptides measured by ITC (sodium phosphate buffer pH 6.5 at 25 ºC). ITC profiles corresponding to (A) (B) (C) and (D) are shown in **Figure 5.3**.

Peptide		**Nedd4L**			
		WW1	**WW2**	**WW3**	**WW4**
		K_D [μM]	K_D [μM]	K_D [μM]	K_D [μM]
Smad3	T179[PY]	156.5 ± 31	50.0 ± 2.8 (**A**)	46.1 ± 2.4	133.2 ± 41
	pT179[PY]	104.7 ± 8.1	4.1 ± 0.3 (**B**)	17.2 ± 1.7	202.4 ± 8.3
Smad2	T220[PY]	132.8 ± 10	52.9 ± 6.3 (**C**)	46.7 ± 4.4	108.9 ± 16
	pT220[PY]	133.9 ± 5.3	6.7 ± 0.3 (**D**)	12.8 ± 0.6	220.3 ± 22

The ITC experiments reveal preferential binding between Nedd4L WW2 and the phosphorylated Smad3/2 pT[PY] motif. This result is in good agreement with the *in vivo* experiments performed by J. Massagué and co-workers (Memorial Sloan-Kettering Cancer Center, NY). Through coimmunoprecipitation experiments they could show that the Nedd4L WW domains and in particular the second domain is the preferred binding region for the Smad2/3 linker specifically when the linker is phosphorylated at the threonine residue that precedes the PY motif, (pT179[PY] in Smad3 and pT220[PY] in Smad2) (Gao et al. 2009).

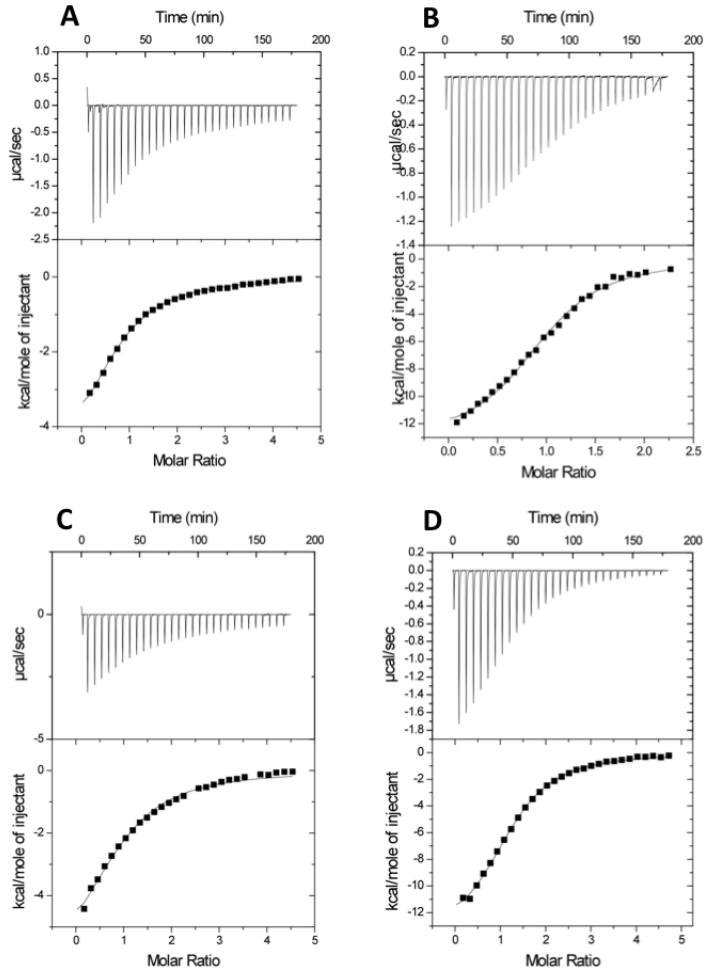


Figure 5.3: ITC profiles of Nedd4L WW2 and (A) Smad3 T[PY], (B) Smad3 pT179[PY], (C) Smad2 T[PY] and (D) Smad2 pT220[PY].

NMR titrations were performed of all four ^{15}N labelled Nedd4L WW domains and the Smad2/3 (p)T[PY] motif. Thereby the peptide was titrated in small amounts (around 10 µl) up to a peptide excess of three-fold to the protein sample as described in chapter 4.3.3. For each domain and in order to identify the cross peaks in the $^{1}H^{15}N$-HSQC spectra backbone CBCA(CO)NH and CBCANH experiments in the free and the bound state were previously acquired. Through sequential assignment (chapter 4.6.1) the resonances of the Nedd4L WW domains were assigned and the information transferred to the two-dimensional $^{1}H^{15}N$-HSQC spectra. The corresponding backbone assignments of the WW2 domain bound to Smad2 pT[PY] are included in Figure 5.4, the remaining domains are shown in the appendix. The NMR titration experiments are in agreement with the ITC experiments. Titration of the Smad2/3 pT[PY] peptide to the Nedd4L WW2 domain results in intense chemical shift perturbations, which are barely observable in the case of Nedd4L WW1 and WW4 titrations

(Figure 5.4 and Figure 5.5). The most affected backbone amides are T392 and T391 (T33 and T32 in the histogram Figure 5.5), which are located on the ß3-stand. Amide protons of ß2-stand residues are also shifted but to a lesser extent.

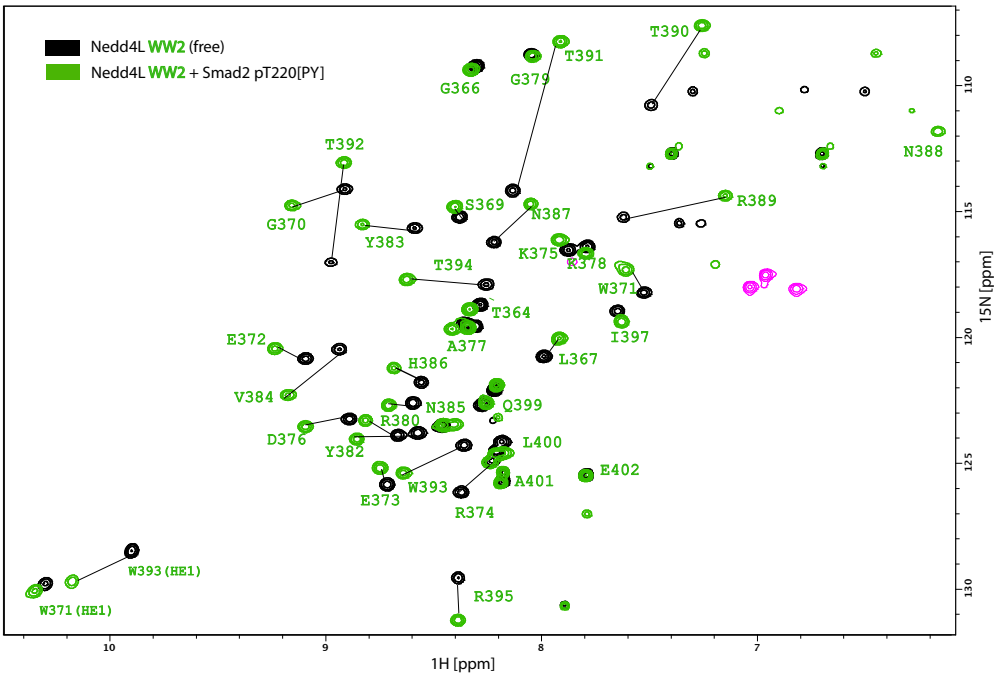


Figure 5.4: ^{1}H ^{15}N-HSQC spectrum of the Nedd4L WW2 domain in the free (black) and bound state (green) after saturation with the Smad2 pT220[PY] peptide.

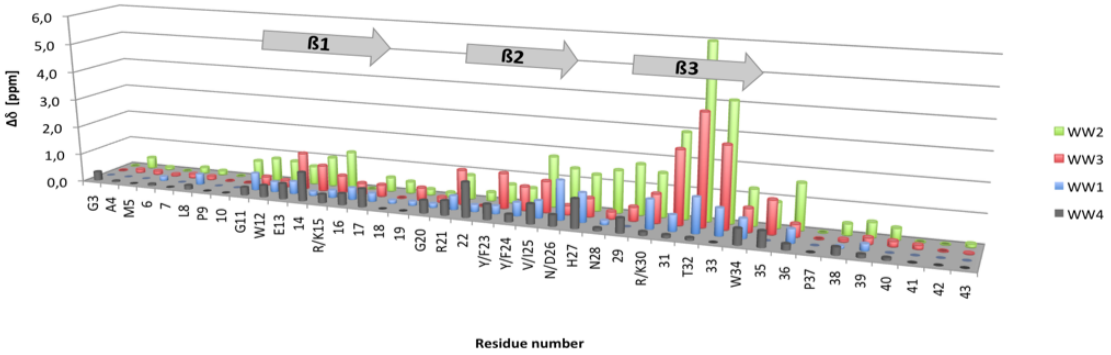


Figure 5.5: Histogram of the average chemical shift changes observed for Nedd4L WW1-4 amide backbone protons after saturation with the Smad2/3 pT[PY] peptides with respect to the free domains. Residues are uniformly labelled to simplify a comparison of the WW domains.

5.1.2 Structural determination of the Nedd4L WW2 – Smad3 pT[PY] complex

The complete assignment of the Nedd4L WW2 domain bound to the Smad3 pT179[PY] ligand was obtained starting from backbone assignment using sequential resonance assignment strategy and then adding information of the side chains from HCCH-TOCSY and ^{15}N-TOCSY experiments. NOEs were obtained from the analysis of ^{13}C and ^{15}N-edited NOESY experiments and from sets of homonuclear TOCSY and NOESY experiments. The structure was refined as described in chapter 4.6 and the statistics of the calculation are shown in the appendix Table 7.6. The peptide was identified in the bound state through sequential- and intramolecular NOEs in 2D ^{1}H-^{1}H TOCSY and NOESY spectra. Structure determination revealed that the Nedd4L WW2 domain binds the Smad3 pT[PY] peptide in negative direction from the C- to N-terminus and that all amino acids are bound in *trans* configuration.

The aromatic residues Y382 of the ß2-strand together with T391 and W393 within the ß3-strand of the WW2 domain form the so-called XP groove, which binds the P181' and P182' of the Smad3 PY motif. NOEs between the aromatic protons of W393 and Y382 and the residues P181' and P182' could be unambiguously identified. Y184' is bound in a hydrophobic pocket formed by V384, H386 and R389. NOEs between the protons HG2/3 of the P182' pyrrolidin ring and the D376 side chain as well as T391 lateral methyl group and α proton HA1 of G183' could additionally be identified.

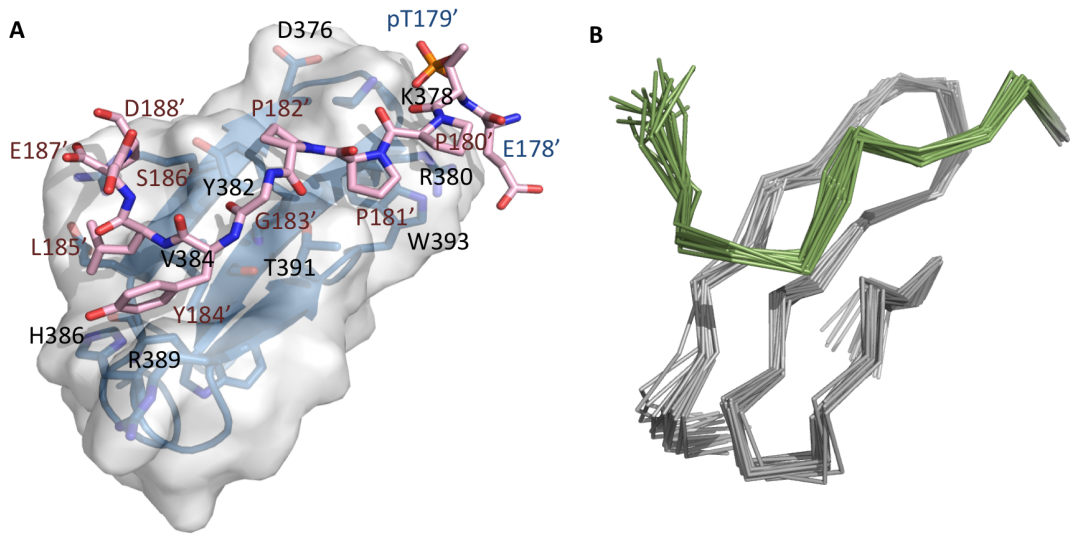


Figure 5.6: (A) Lowest-energy structure of the Nedd4L WW2 – Smad3 pT179[PY] complex shown as a semitransparent solid surface representation with the secondary structure elements shown as a blue cartoon and the peptide as pink sticks. Residues located in the binding interface are labelled as well as the peptide residues. (B) Backbone superposition of the 20 best structures after water refinement (domain in grey and ligand in green).

The Nedd4L WW2 – Smad3 pT[PY] complex is similar to other group I WW domain-PY complexes in the region describing the PY motif recognition site (Kanelis et al. 2001; Pires et al. 2001). In Nedd4L, all WW domains posses a XP groove and a hydrophobic pocket and are therefore PY binder.

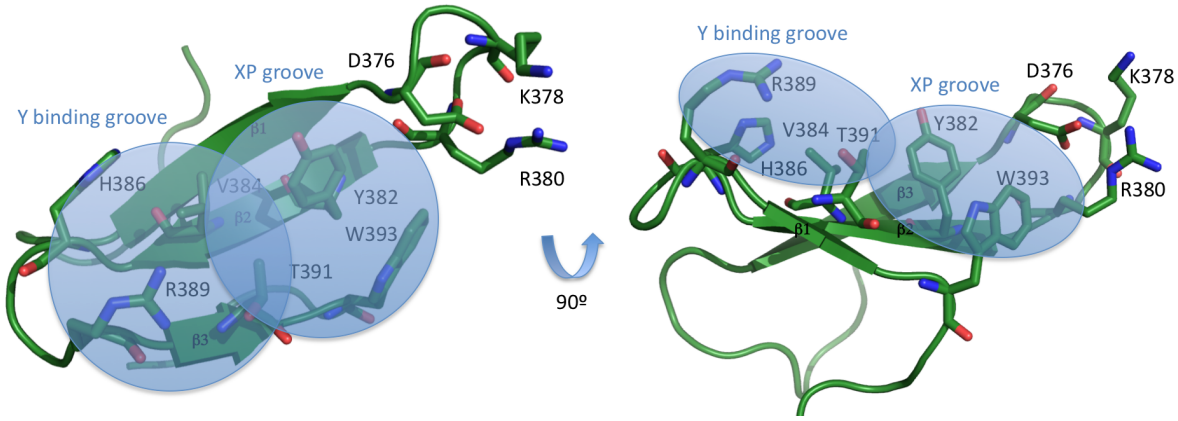


Figure 5.7: Structure of the Nedd4L WW2 domain without peptide. Residues H396, R389 and V384 form the Tyr binding groove, the XP groove is formed by residues Y382 and W393.

However phosphorylation of T179' upstream of the Smad3 PY motif remarkably increases the affinity to the Nedd4L WW2 domain. In the Nedd4L WW2 - Smad3 pT179[PY] complex pT179' is located close to loop 1 residues K378 and R380. Potential electrostatic interactions between side chains K378/R380 and the pT179' phosphate group could explain the high binding affinity of this complex.

Table 5.2: Inter- and intramolecular peptide NOEs indentified through 2D TOCSY/NOESY assignment of the Nedd4L WW2 - Smad3 pT179[PY] complex (nomenclature of protons is presented in appendix 7.6).

Smad3 pT179[PY] - Nedd4L WW2					
^{180}PEpTPPPGY^{185}ISEDG		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
pT179'	HN			E178'	*HN*
					HA
	HA			P180'	HD2
					HD3
	HB	W393	HZ2		
	QG2	W393	HH2	E178'	HA
			HD1		
			HZ2		
P181'	HA	W393	HH2	P182'	HD2
			HD1		HD3
			HZ2		
	HB2	W393	HN		
			HH2		
			HD1		
			HZ2		

Smad3 pT179[PY] - Nedd4L WW2					
^{180}PEpTPPPGY^{185}ISEDG		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	HD2	W393	HE1	P180'	HA
	HD3	W393	HE1	P180'	HA
	HG2	W393	HN		
			HH2		
			HD1		
			HZ2		
	HG3	W393	HZ2		
	HA			G183'	HN
	HB3	Y382	QD		
			QE		
P182'	HD2	Y382	QE	P181'	HA
	HD3		QD		HA
	HG2/3	D376	HA		
			HB2/3		
G183'	HN			P182'	HA

Smad3 pT179[PY] - Nedd4L WW2					
180 185 PEpTPPPGYISEDG		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	HA1	T391	QG2	Y184'	HN
Y184'	HN	V384	QG2	G183'	HA1
		T391	QG2		HA2
	HA	V384	QG1	L185'	HN
			QG2		QD1
		T391	QG2		
	HB2/3	T391	QG2	L185'	QD1
		V384	QG2		HN
		H386	HE1		
	QD	H386	HA	L185'	QD1
			HE1		
		V384	QG2		
			QG1		
		R389	HD2/3		

Smad3 pT179[PY] - Nedd4L WW2				
180 185 PEpTPPPGYISEDG		Inter-molecular NOEs		Intra-molecular NOEs
AA	Proton			
			HA	
			HG2/3	
			HB2/3	
		T391	QG2	
	QE	R389	HA	
			HG2/3	
			HD2/3	
			HB2/3	
		V384	QG2	
			QG1	
		T391	QG2	
		H386	HA	
			HE1	

5.1.3 The Smad3 pT179 phosphate group is directly involved in Nedd4L WW2 binding

In order to characterize the role of the pT179' phosphate group in the interaction with Nedd4L WW2, one-dimensional ^{31}P-NMR experiments were performed. NMR spectra of the phosphorylated Smad3 linker peptide (Smad3 pT179[PY]) bound to the Nedd4L WW2 domain and in the unbound state were acquired at two different pH values. The phosphate signal of the bound and unbound peptide shifts upfield in both cases at lower pH. This chemical shift pertubation is more pronounced in the case of the unbound peptide with a difference of 4.0 ppm (from -0.5 ppm to 3.5 ppm), whereas the bound peptide only shifts 2.0 ppm (from 2.5 ppm to 4.5 ppm). The phosphorus of the phosphate group in the deprotonated state is deshielded resulting in a chemical shift towards downfield. At pH 6.0 hydrogen bonds are formed to the phosphate group leading to an upfield shift of the ^{31}P peak. In the bound state this effect is weaker because the phosphate group is less accessible by the solvent protons and the phosphate signal only shifts slightly. The phosphorus signal of the free to the bound state is shifted by 3.0 ppm at pH 6.0, which again reveals that the phosphorus nucleus does not experience the same field in both cases. The ^{31}P-NMR experiments confirmed a direct role of the pT179' and the WW domain, as previously suggested by biochemical experiments.

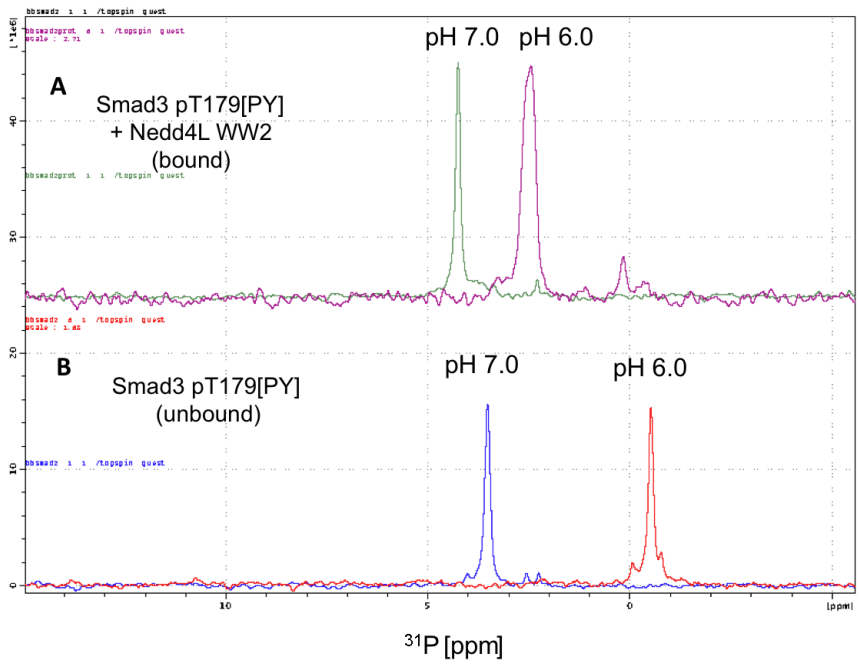


Figure 5.8: 1D ^{31}P-NMR spectra of (A) the bound Smad3 pT179[PY] peptide and (B) the free Smad3 pT179[PY] peptide at pH 6.0 and 7.0.

As mentioned before in chapter 5.1.2, two positively charged amino acids (R380 and K378) in the loop 1 region of Nedd4L WW2 were identified as possible binding partners for the phosphate group. Sequence alignment of the four Nedd4L WW domains reveals that R380 is conserved in all WW domains, whereas K378 is only present in Nedd4L WW2 (Figure 5.9). In the case of the Nedd4L WW3 domain the amino acid Asn is situated at the same position as K378 in Nedd4L WW2, Nedd4L WW1 contains the aliphatic residue Leu and Nedd4L WW4 the negatively charged Asp at the same position. Comparing the binding affinities of all Nedd4L WW domains - pT[PY] complexes (Table 5.1) binding becomes weaker from WW2 (K_D = 4.1 ± 0.3 µM) to WW3 (K_D = 17.2 ± 1.7 µM) to WW1 (K_D = 104.7 ± 8.1 µM) to WW4 (K_D = 202.4 ± 8.3 µM). These findings suggest that the K378 side chain (WW2) participates in phosphate binding, whereas an aliphatic or negatively charged residue (WW1 and WW4, respectively) reduce Smad2/3 pT[PY] binding.

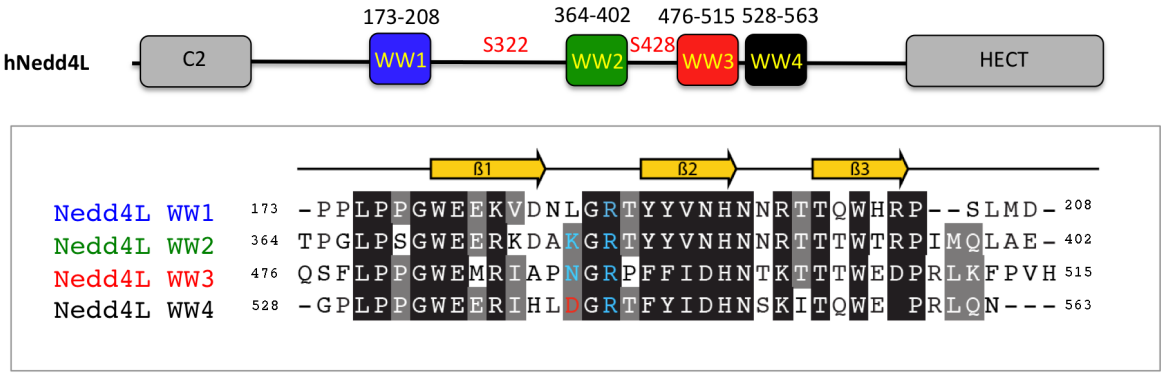


Figure 5.9: Sequence alignment of Nedd4L WW1-4. Alignments were performed using *ClustalW* (Thompson et al. 1994) and BoxShade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

To study the effect of R380 and K378 on binding the phosphate group of pT[PY] motif, three Nedd4L WW2 mutants (Figure 5.10, Figure 5.11) were cloned, expressed and purified as described in chapter 3.1.2. Either one or both K378 and R380 residues were mutated to the negatively charged glutamic acid. To exclude that potential differences in the interaction are due to changes in the three-dimensional structure of the mutants, the structure of each mutant was determined. These structures revealed that the overall fold of the mutated WW domain was not affected.

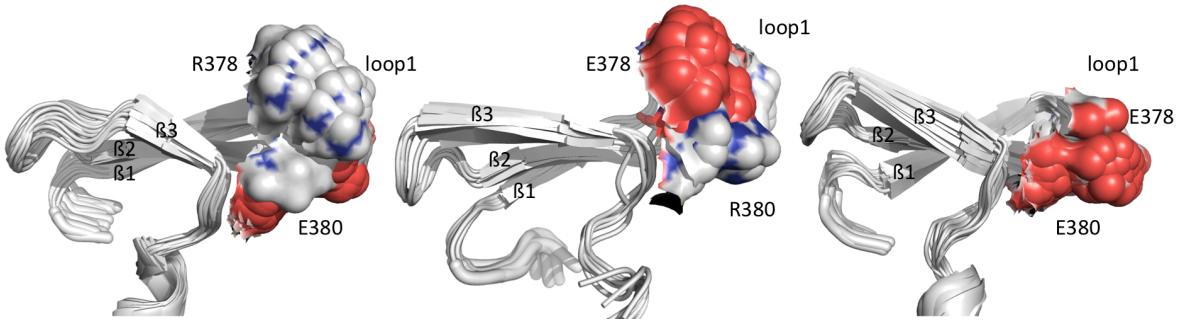


Figure 5.10: Structure of Nedd4L WW2 loop 1 mutants: Superposition of the ten lowest energy structures after water refinement (backbone in cartoon representation, loop 1 residues 378 and 380 in surface electrostatic representation).

Binding affinities of the Nedd4L WW2 mutants and phosphorylated (pT179[PY]) and nonphosphorylated (T179[PY]) Smad3 peptides were measured. ITC experiments reveal that all mutations decrease the binding between Smad3 pT179[PY] and the Nedd4L WW2 domain. This confirms the prediction that both loop 1 residues are important for binding the pT179' phosphate group. Glutamic acid mutation decreases binding of the phosphorylated Smad3 pT[PY] peptides and both Nedd4L WW2 K378/E and R380/E mutants by five- to six fold whereas binding to the Nedd4L WW2 K378R380/EE mutant is decreased by around ten-

fold. In case of the nonphosphorylated Smad3 T[PY] peptide and the three mutants binding is increased by around two- to four-fold. The ITC experiments reveal that a substitution of K378 and R380 for the shorter and negatively charged Glu stabilizes the nonphosphorylated T179', whereas K378 and R380 are both important for binding the pT179' phosphate group.

Table 5.3: K_D values of the Nedd4L WW2(mutants) - Smad3 (p)T[PY] complex formation (sodium phosphate buffer pH 6.5 at 25 °C).

	Nedd4L WW2			
Smad3	**K378 R380 (WT)**	**K378 E380**	**E378 R380**	**E378 E380**
	K_D [µM]	K_D [µM]	K_D [µM]	K_D [µM]
T179[PY]	59.9 ± 9.0	16.1 ± 7.8	28.5 ± 3.8	14.4 ± 3.4
pT179[PY]	4.1 ± 0.3	20.0 ± 8.1	23.5 ± 3.2	44.6 ± 17

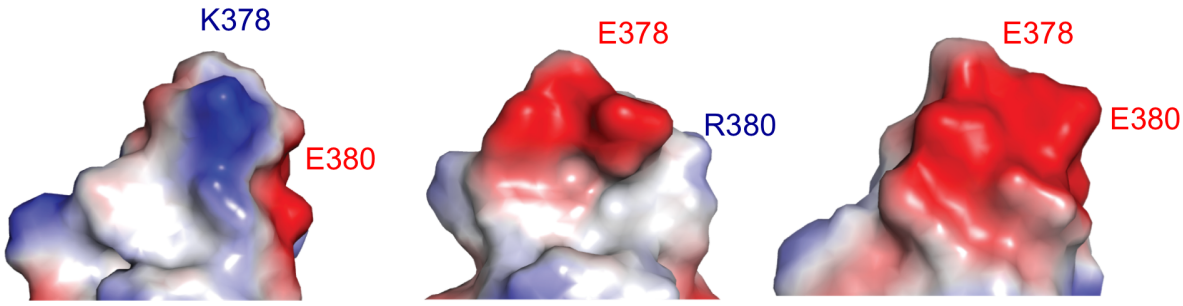


Figure 5.11: Surface electrostatic representation of the Nedd4L WW2 loop 1 mutants K378/E, R380/E and K378R380/EE.

5.1.4 The phosphorylated Smad3 linker interacts with the Pin1 WW domain

Pin1 catalyses the *cis/trans* isomerisation of phosphorylated p(S/T)-P motifs specifically. Recently Pin1 has been identified as binding partner for Smad2/3 (Nakano et al. 2009). It has been shown that Pin1 is able to recognize the phosphorylated Smad3 linker and that it enhances binding between Smad3 and E3-ligases, resulting in Smad3 ubiquitination (Nakano et al. 2009). Pin1 recognizes its substrate through an unique WW domain that binds p(S/T)P motifs, whereas the *cis/trans* isomeration is carried out by the catalytic domain. All phosphorylation sites T179', S204', S208' and S213' within the Smad3 linker are followed by a proline, hence, the Smad3 linker contains four potential binding sites for the Pin1 WW domain.

To study the interaction between the Pin1 WW domain and the Smad3 p(S/T)P motifs seven short peptides (ranging from 7 to 13 aa) were synthesized by Fmoc SPPS and the binding affinities between these peptides and the Pin1 WW domain were studied by ITC and NMR.

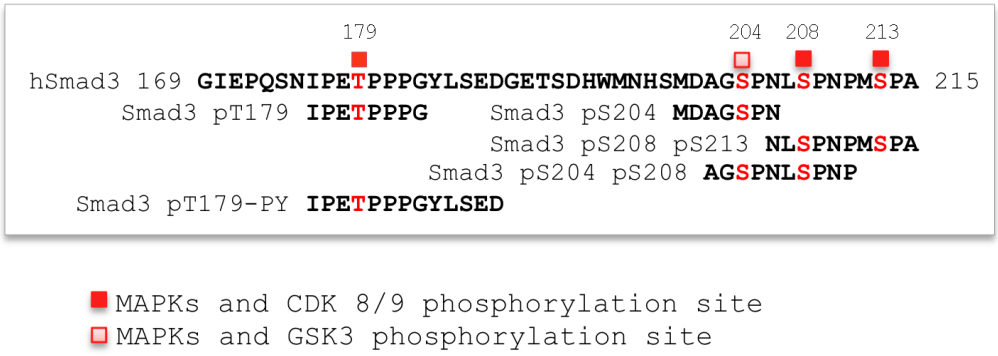


Figure 5.12: Sequences of the Smad3 linker peptides used for Pin1 WW binding studies. Phosphorylated amino acids in red.

5.1.5 Smad3 pT179P is the preferred Pin 1 WW binding site

The dissociation constants of the Pin1 WW – Smad3 p(S/T)P interaction obtained by ITC binding studies lie in the high µM range typical for group IV WW domains (Macias et al. 2002). The pT179P motif is the preferred binding site for Pin1 WW domain (K_D = 14.2 ± 2.0 µM), whereas all other peptides bind two- to six-fold weaker. The aromatic amino acid F25 located within the ß2 strand inhibits binding of a PY motif. For this reason there is no difference in the binding affinity of the Smad3 pT179 and pT179[PY] peptides observable (K_D = 14.2 ± 2.0 and 12.6 ± 0.9 µM, respectively). As expected the Pin1 WW domain does not bind the nonphosphorylated Smad3 T179[PY] peptide.

Table 5.4: K_D values Pin1 WW + Smad3 p(S/T)P motifs (sodium phosphate buffer, pH 6.5, 25 °C).

Peptide	Pin1 WW
	K_D [µM]
Smad3 T179[PY]	NB
Smad3 pT179[PY]	12.6 ± 0.9
Smad3 pT179	14.2 ± 2.0
Smad3 pS204	84.8 ± 29.2
Smad3 pS204-pS208	49.0 ± 11.9
Smad3 pS208-pS213	23.4 ± 7.7

Additionally, NMR titration experiments of the Pin1 WW domain and peptides Smad3 pT179 and Smad3 pS204 were performed. HSQC spectra were acquired of the ^{15}N labelled unbound Pin1 WW domain before titrating the peptides up to a peptide excess of three-fold to the domain. Triple resonance NMR spectroscopy of the ^{15}N and ^{13}C labelled Pin1 WW samples in the free and the bound state was acquired to assign the backbone amides. In the case of the Pin1 WW – Smad3 pT179 complex chemical shift pertubation occurs, whereas in the Pin1 WW – Smad3 pS204 complex it is barely observable. Other than observed for Nedd4L WW domains the most affected amide protons are located on ß1- and ß2-strand of the Pin1 WW domain. The largest chemical shift change was found for the backbone amide proton of R17 (R17 HN: 8.05 ppm (bound) vs. 8.46 ppm (unbound), complex: Pin1 WW-Smad3 pT179) indicating its presence in the direct vicinity of the binding site.

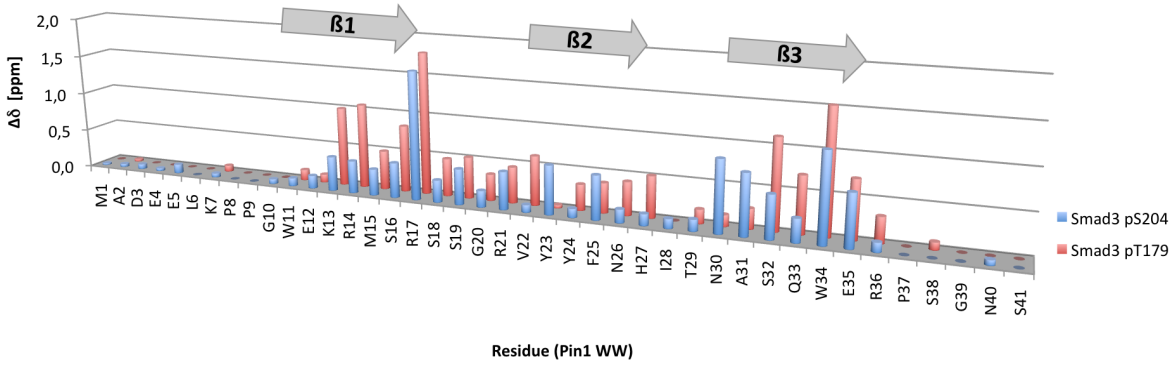


Figure 5.13: Histrogram of the average chemical shift changes observed for Pin1 WW domain amide backbone protons after saturation with the Smad3 pT179 (red) and Smad3 pS204 (blue) with respect to the free domains.

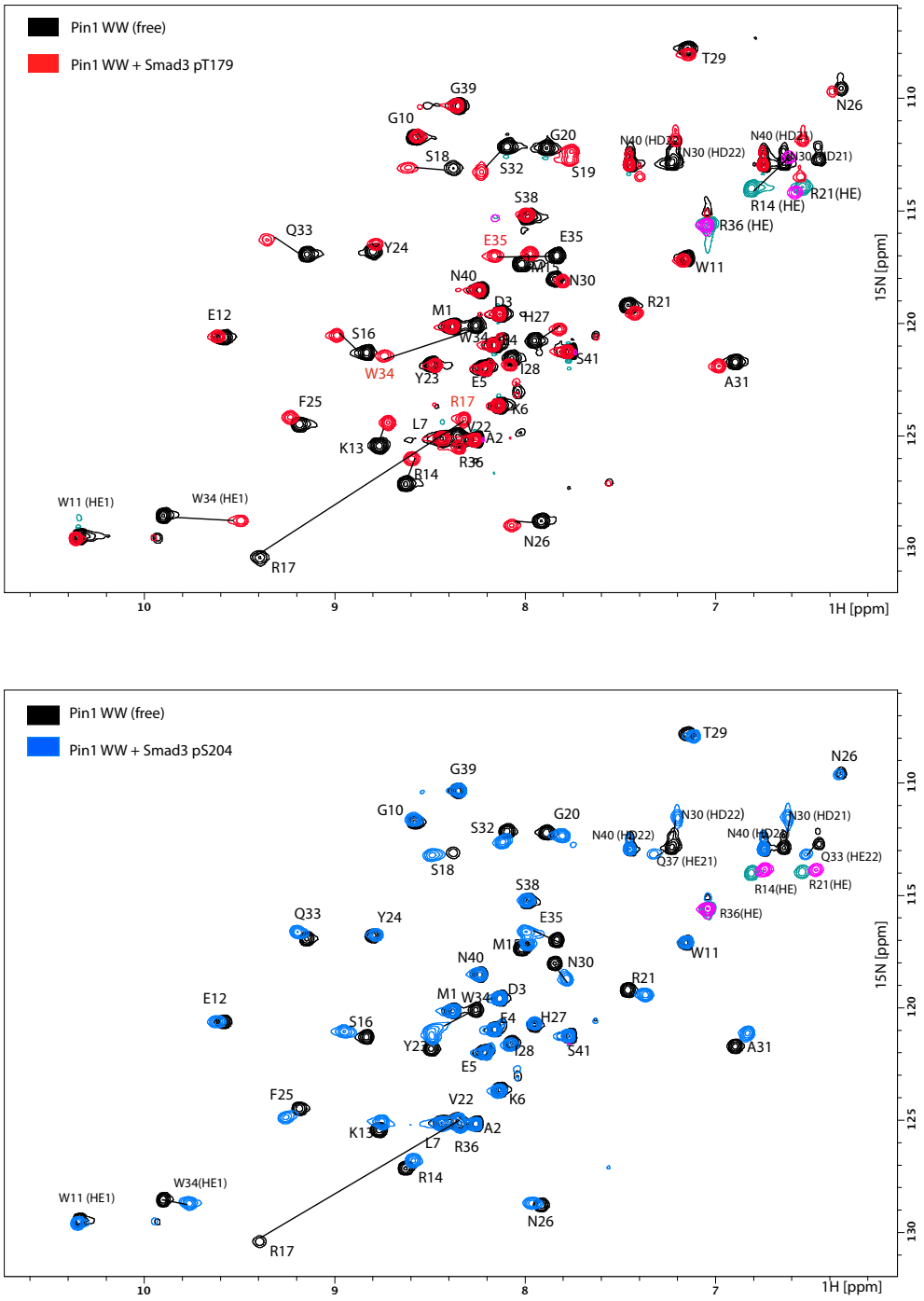


Figure 5.14: $^{1}H^{15}N$-HSQC spectra of Pin1 WW domain in the free and bound state with Smad3pT179 (above) and Smad3 pS204 (below).

5.1.6 The Smad3 pT179 phosphate group is directly involved in Pin1 WW binding

In order to illustrate the impact of the Smad3 pT179 phosphate group on Pin1 WW binding, one-dimensional ^{31}P-NMR experiments were performed. The ^{31}P-NMR spectrum of a 0.5 mM

Smad3 pT179P peptide sample in bis-tris buffer with 0.5 mM sodium phosphate buffer as ^{31}P reference was acquired at pH 7.0. Pin1 WW domain was added to the sample until saturation. As illustrated in Figure 5.15 the ^{31}P reference peak remains at the same position, whereas the Smad3 pT179 peak moves downfield during the titration experiment.

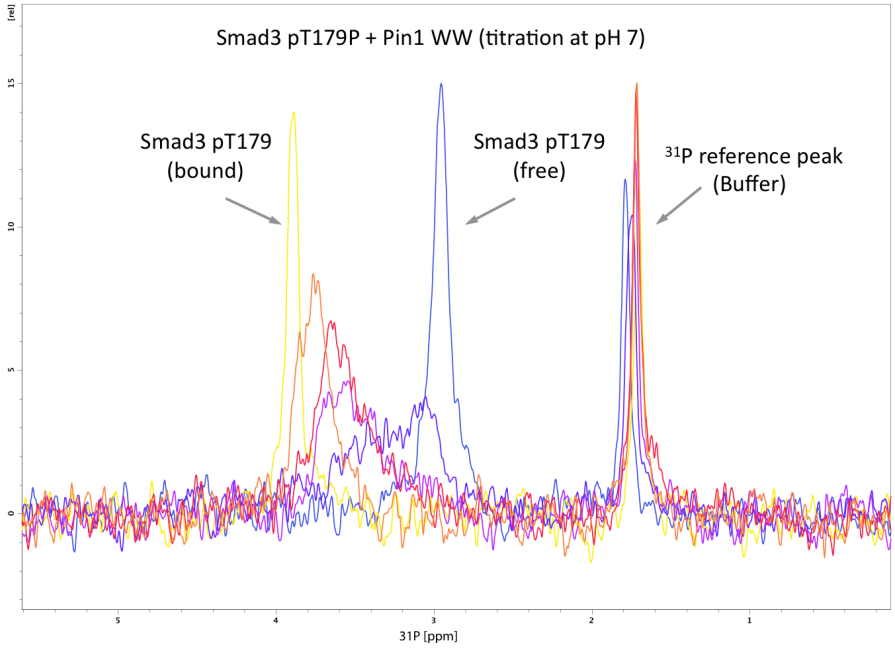


Figure 5.15: 1D ^{31}P NMR titration experiment of the free Smad3 pT179P peptide and the Pin1 WW domain. The transition from the unbound to the bound state of the peptide is illustrated by a chemical shift change from low to downfield, whereas the reference peak barely shows chemical shift pertubation.

In a second experiment the same sample was divided into two parts and the pH was adjusted to a higher pH of 8.0 and a lower pH of 6.0 by adding the same volumes of bis-tris buffer pH 9.0 and 5.0, respectively (with 0.5 mM sodium phosphate buffer as ^{31}P reference). NMR spectra of Pin1 WW – Smad3 pT179P at pH 6.0, 7.0 and 8.0 were superimposed (Figure 5.16). The reference peak is affected by the different proton concentrations and shifts upfield at low pH and downfield at high pH. ^{31}P of pT179 experiences a different field because it is bound to Pin1 WW and is not exposed to the environment as the ^{31}P nucleus of sodium phosphate within the buffer. Electron density is withdrawn from the phosphorus atom by deprotanation of the phosphate group at a higher pH. This deshielding effect results in chemical shift pertubation towards downfield. In the bound state the solvent effect is weaker because the phosphate group is less accessible by the solvent protons and the phosphate signal only shifts slightly downfield.

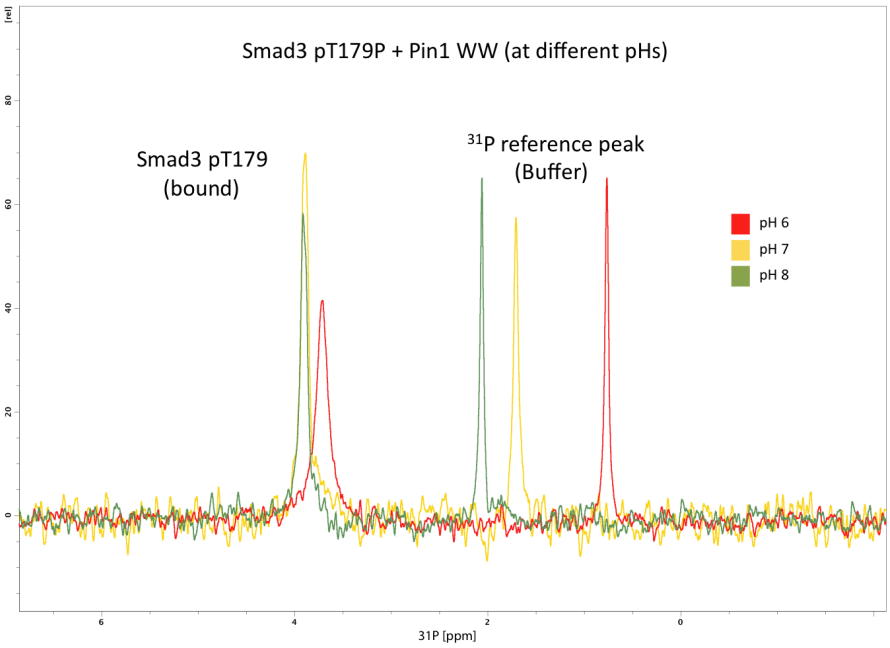


Figure 5.16: Superimposition of 1D ^{31}P NMR spectra of the Pin1 WW – Smad3 pT179P complex at pH 6.0 (red), 7.0 (yellow) and 8.0 (green).

5.1.7 Structural determination of the Pin1 WW - Smad3 pT179 complex

The assignment of most resonances of the Pin1 WW bound to the Smad3 pT179 ligand was obtained through backbone and side chain assignment using standard triple resonance experiments (chapter 5.1.5). The peptide was assigned in the free and bound state using 2D ^{1}H^{1}H TOCSY and ^{1}H^{1}H NOESY spectra. The bound state was defined on the bases of the intramolecular NOEs detected with the protein. Typical ring current shifts due to the presence of aromatic amino acids are observed in resonances belonging to the protein, which are typical of WW domains. For instance, the vicinity of the W11 indole ring contributes to the upfield shifts of N26 ß-protons (HB2/3: 1.76/-0.95 ppm), P37 ß-protons (HB2/3: 0.40 and 0.59 ppm) and γ-protons (HG2/3: -0.28/0.37 ppm). In the same way a large upfield shift close to the aromatic ring of Y24 is observed at R36 α-proton (HA: 2.51 ppm). The final structure calculation was performed following the protocol described in chapter 4.6. Structures were analysed with the PROCHECK-NMR and the statistics obtained are included in the appendix Table 7.6.

The Pin1 WW domain binds the Smad3 pT179 peptide in *trans* configuration from the N- to the C-terminus typical for group IV WW domains (Lu et al. 1999; Sudol and Hunter 2000; Verdecia et al. 2000; Zarrinpar and Lim 2000; Macias et al. 2002). P180' is bound between the aromatic residues Y23 of the ß2-strand and W34 of the ß3-strand. Intermolecular NOES could be further indentified between W34 and proline P181'. The aromatic residue F25 is located at the end of the ß2 strand and interacts with the upstream part of the peptide. In all Nedd4L WW domains F25 is replaced by an aliphatic residue (Val or Ile) that binds Y' of the Smad2/3 PY motif. In case of the Pin1 WW domain the presents of F25 inhibits Y' binding instead it binds P177' upstream of the Smad3 p(S/T)P motif. Electrostatic interaction between R14 (ß1-strand) and the E178' side chain as well as R17 (loop 1) and the phosphate group of pT179' explains the high binding affinity of the Smad3 pT179 peptide. The side chain and the backbone amid of R17 form the phosphate-binding pocket as observed in Pin1 WW complexes with other p(S/T)P motifs (Verdecia et al. 2000; Wintjens et al. 2001) and observed as a large chemical shift pertubation of the R17 backbone amide in HSQC spectra of all Pin1 – Smad3 p(S/T)P complexes. It has been shown that a replacement of the R17 by Ala decreases the binding by six-fold (Verdecia et al. 2000).

Table 5.5: Inter- and intramolecular peptide NOEs indentified through 2D TOCSY/NOESY assignment of the Pin1 WW - Smad3 pT179 complex (for nomenclature of protons see appendix).

Smad3 pT179[PY] – Pin1 WW					
^{176}IPEpTPPPG^{182}		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
P177'	HA			E178'	HN
	HG2/3	Y23	QQ/QE		
		F25	QD/QE		
	HB2/3	Y23	QD/QE		
E178'	HN			P177'	HA
	HA			pT179'	HN
		Y23	QE		
	HB2/3	Y23	QD/QE		
	HG2/3	Y23	QD/QE		
pT179'	HA			P180'	HD2
					HD3
	QG2	Y23	QE	E178'	HA
			QD		
P180'	HA			P181'	HD2
					HD3

Smad3 pT179[PY] – Pin1 WW					
^{176}IPEpTPPPG^{182}		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	HB2	Y23	QE		
			QD		
		W34	HZ2		
			HH2		
	HD2	W34	HE3	P179'	HA
P181'	HA	W34	HE1		
	HB2	W34	HE3		
			HE1		
			HH2		
			HZ2/3		
	HG2	W34	HE3		
			HE1		
	HG3	W34	HE3		
			HH2		
	HD2/3			P180'	HA

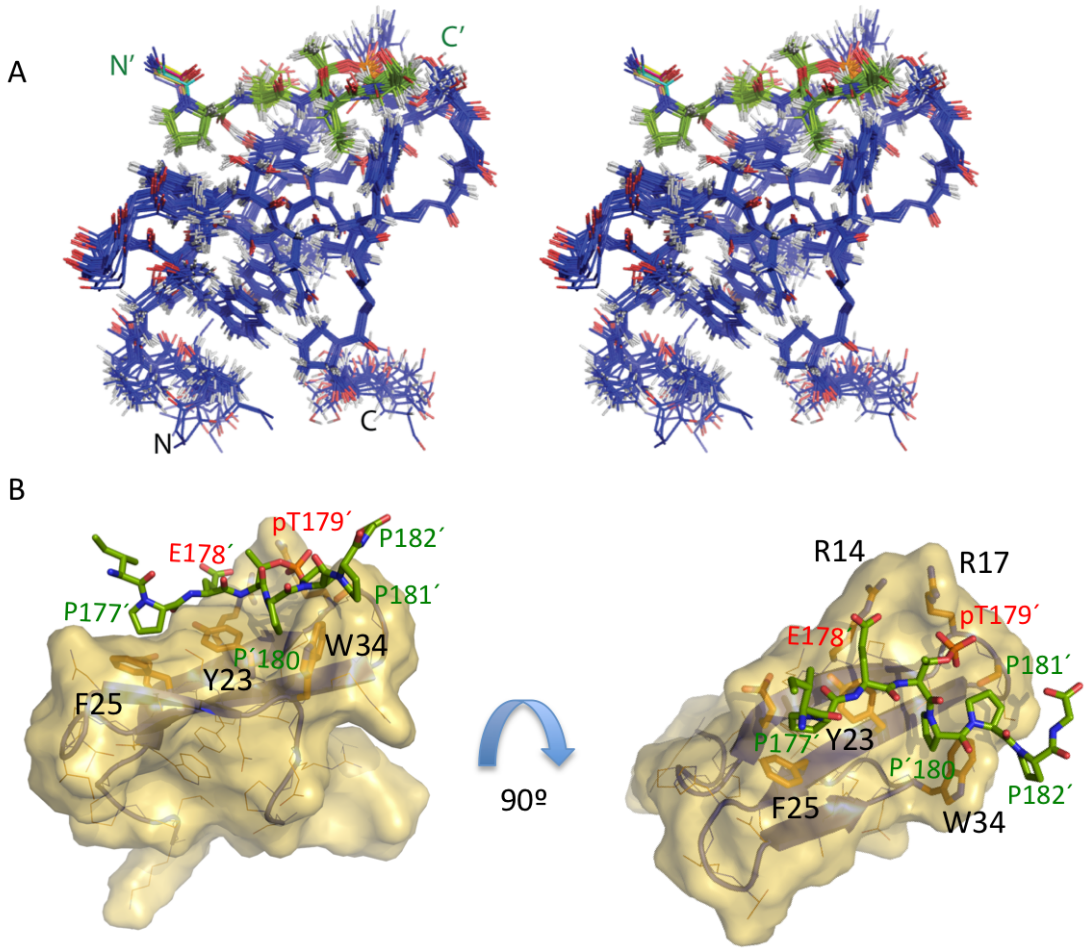


Figure 5.17: Solution structure of Pin1 WW in complex with Smad3 pT179 peptide. (A) Stereo view of the best-fit backbone superposition of the ten lowest energy structures after water refinement with the Smad3 pT179 ligand (domain in blue and ligand in green). (B) Surface representation of the lowest energy structure with the ligand in green. Residues E178' and pT179' are labelled in red.

5.1.8 Pin1 WW binds the pS213P motif of the Smad3 pS208-pS213 peptide

As presented in Table 5.4, Pin WW domain interacts with four regions of Smad3 linker that differ in the number of phosphorylation sites and in the spacing between them. Smad3 pS208-pS213 contains two phosphorylation sites and interacts with Pin1 WW stronger than Smad3 pS204 and Smad3 pS204-pS208 but weaker than Smad3 pT179. To describe the differences in affinity at a molecular level, 2D ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY spectra were acquired to investigate the interaction of Pin1 WW domain and Smad3 pS208-pS213 peptide, which is the second best binder. The assigned peaks of the free Pin1 WW domain previously obtained using triple resonance experiments were used as the starting point to assign the ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY spectra of the Pin1 WW – Smad3 pS208-pS213 complex. The peptide was identified as in the pTP complex using the sequential resonance assignment

strategy. As described for the Pin1 WW – Smad3 pT179 complex, the peptide is bound from the N- to the C-terminus. Intermolecular NOEs could be identified between the C-terminal part of the peptide, but no NOEs between N-terminal pS208' and the WW domain could be detected. P214' is bound by the aromatic rings of W34 and Y23 while side chains of ß1 strand residues R14 and R17 coordinate the pS213' phosphate group. Additional contacts between W34 and A215', pS213' and Y23, W34 as well as between the aromatic ring of residue F25 and P209', P211' were identified.

Table 5.6: Inter- and intramolecular peptide NOEs indentified through 2D TOCSY/NOESY assignment of the Pin1 WW – Smad3 pS208-pS213 complex (for nomenclature of protons see appendix).

Smad3 pS208-pS213 + Pin1 WW					
^{210} ^{215} NLpSPNPMpSPA		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
P209'	HA			N210'	HN
	HB2/3	F25	QD	N210'	HN
			QE		
			HZ		
	HG2			N210'	HN
P211'	HD3	F25	QD		
			QE		
	HB2/3	F25	QD		
			QE		
pS213'	HA			P214'	HD2/3
	HB2	Y23	QD		
			QE		
		W34	HZ2		
			HH2		
			HZ3		

Smad3 pS208-pS213 + Pin1 WW					
^{210} ^{215} NLpSPNPMpSPA		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
P214'	HB2/3	W34	HZ2		
			HD1		
		Y23	QE		
			QD		
	HD2/3	W34	HH2		
			HZ2		
		Y23	QE		
			QD	pS213'	HA
A215'	HA	W34	HE1		
			HZ2		
	QB	W34	HD1		
			HE1		
			HH2	P214'	HG2
			HZ2		

In both Pin1 complexes with Smad3 pT179 and with Smad3 pS208-pS213 peptides, P177' and P211' respectively are bound in a cavity formed by F25 and Y23 aromatic rings (I**P**EpTPPPG and **N**LpSPN**P**MpSPA sequences) but this proline is absent in Smad3 pS204P and in Smad3 pS204-pS208 sites (MD**A**GpSPN and **A**GpSP**N**LpSPNP, respectively). It is most likely that the higher affinity of the first two peptides, compared to the second pair, is due to the additional contacts that the Pro can provide, when compared to the sequences containing an Ala or Asn at position n - 2 with respect to the p(S/T)P site.

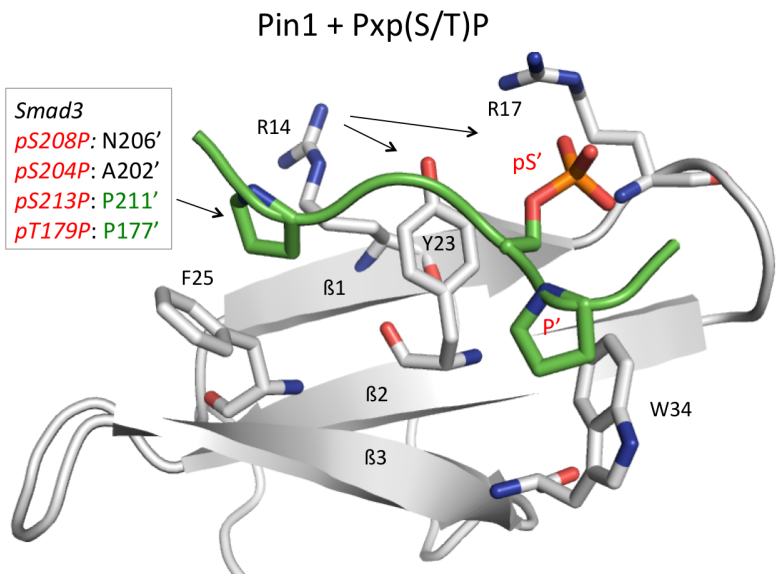


Figure 5.18: Ligand binding strategy of the Pin1 WW domain. P211' of the Smad3 pS213P motif (P177' in pT179P) increases binding affinity. R17 coordinates the p(S/T)' phosphate group. R14 either participates in phosphate binding or interacts with the residues directly upstream of p(S/T)'.

In Pin1 - Smad3 pT179 the phosphate group is bound by R17 while R14 interacts with the E178' carboxyl group. In Pin1 - Smad3 pS208-pS213 both residues R14 and R17 bind the phosphate group (Figure 5.19). It is most likely that the negatively charged amino acid E178' increases the binding affinity by facilitating secondary electrostatic interaction. Depending on the residue located directly upstream of p(S/T), R14 may either interact with this amino acid or participate in the phosphate group binding.

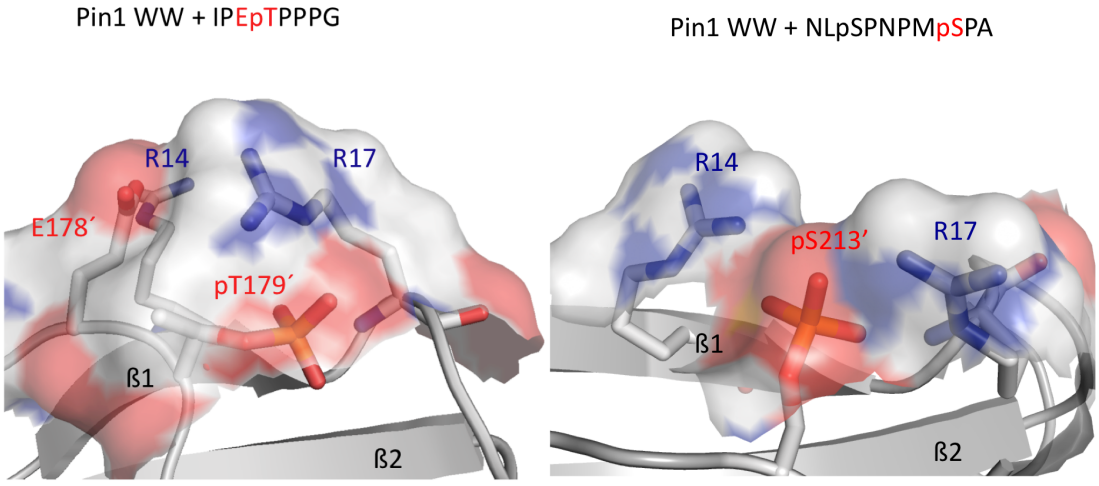


Figure 5.19: Representation of the Pin1 WW ß1 and loop 1 region. The p(S/T)' phosphate group is either bound by R17 and R14 (pS213') or only by R17 (T179'), whereas R14 binds E178'.

5.1.9 Binding affinity between Smad3 and Nedd4L WW2/Pin1 WW depends on the pH

The chemical environment of a protein solution can modulate to a given extend the binding to a ligand. In order to study the effect of pH variations on Smad3 binding to Nedd4L WW2 and Pin1 WW, the variation of the affinity was determined using ITC experiments measured at four pH values ranging from 6.0 to 7.5. To perform these experiments the bis-tris buffer was selected due to its buffer capacity in a wide pH range.

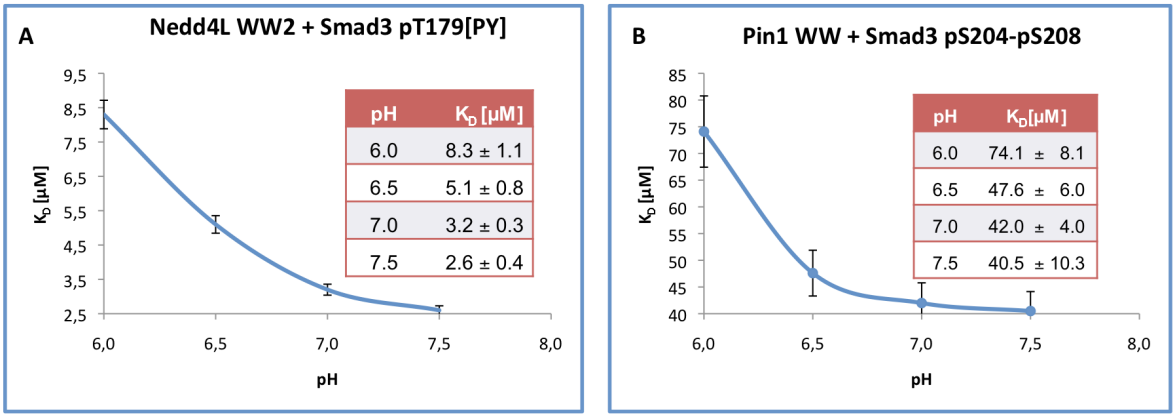


Figure 5.20: Binding affinities at different pH values of (A) Nedd4L WW2 and Smad3 pT179[PY], (B) Pin1 WW and Smad3 pS204-pS208. Decaying curve represents the pH dependence of the K_D values.

ITC experiments reveal that binding increases for both WW domains at higher pH. In the case of Nedd4L WW2 from K_D = 8.3 ± 1.1 µM (pH 6.0) to K_D = 2.6 ± 0.4 µM (pH 7.5). In the case of Pin1 WW from K_D = 74.1 ± 8.1 µM (pH 6.0) to K_D = 40.5 ± 10.3 µM (pH 7.5). The K_D values of 5.1 ± 0.8 µM (Nedd4L WW2 + Smad3 pT179[PY]) and 47.6 µM ± 6.0 (Pin1 WW + Smad3 pS204-pS208) obtained at pH 6.5 are in good agreement with those obtained from previous ITC experiments presented in Table 5.1 and Table 5.4.

Nedd4L WW2 contains a highly conserved residue H386 located in the tyrosine-binding pocket. Residue His at the end of ß2-sheet is critical in all group I WW domains. It has been shown that the binding affinity decreases significantly when His is replaced by Ala (Macias et al. 1996). If the His residue is uncharged it can form a hydrogen bond to the hydroxyl group of the aromatic Tyr within the PY motif. Protonation of His by lowering the pH results in weaker ligand binding as it has been observed for Smurf2 WW3 (Chong et al. 2006).

Pin1 WW also contains a His residue close to loop 2, but the vicinity of the aromatic residue Phe inhibits Tyr binding of the PY motif. In the case of Pin1 the stronger binding at higher

pH may arise from intense electrostatic interaction between the bivalent phosphate group and R17 (R14) side chains. When the phosphate group is totally deprotonated, electrostatic interactions between positively charged amino acids of the Pin1 WW domain and the phosphate group are stronger. This effect might also play a role in the case of Nedd4L WW2.

5.2 The Nedd4L WW2-WW3 pair binds the extend Smad3 linker

A common feature of E3 ligases is the presence of several copies of WW domains. WW domains can bind individually to different targets with different functional outcome or they can bind their targets in a cooperative way as WW domain pairs (Sudol et al. 2005; Ramirez-Espain et al. 2007; Bhattacharya et al. 2008). In order to study the binding of the Nedd4L WW pairs and the Smad3 linker, the three possible WW1-WW2, WW2-WW3 and WW3-WW4 pairs were produced and two C-terminal extended Smad3 peptides were synthesized by automatic Fmoc SPPS with ChemMatrix as resin. The first amino acid was introduced in a double coupling step as well as the phosphorylated amino acids. The incorporation of phosphorylated amino acids was controlled by a mini cleavage that was carried out with a small amount of resin. The phosphorylated amino acid was either coupled again manually or the synthesis was continued depending on the results of the mini cleavage obtained by mass spectrometry. Coupling of the N-terminal proline P180' of the polyproline sequence was repeated. Several purification steps had to be carried out by HPLC to separate the product peak from the side products. The synthesized peptides included the pT179[PY] motif as well as phosphorylated site S208' or both phosphorylated sites S204' and S208' (Figure 5.21). Phosphorylation of S213' was neglected as *in vivo* experiments carried out by J. Massagué's group (personal communication, Memorial Sloan-Kettering Cancer Center, NY) showed that Ala mutation of S213' had no effect on Nedd4L binding, whereas mutation of S204' and S208' diminished the binding of Smad3 to Nedd4L in HEK293 cells.

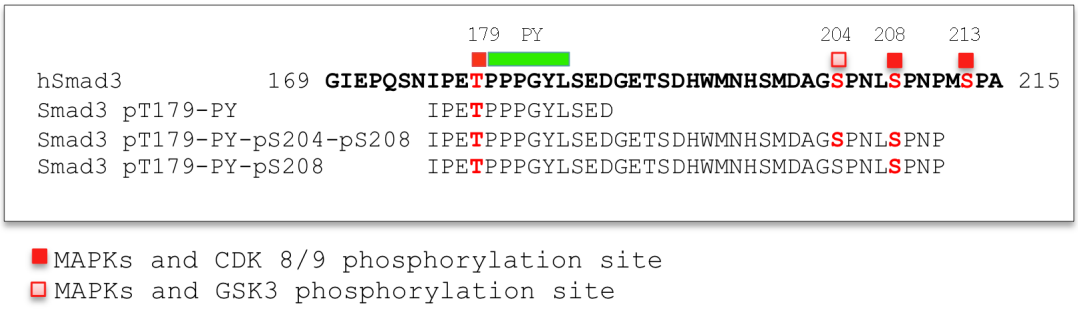
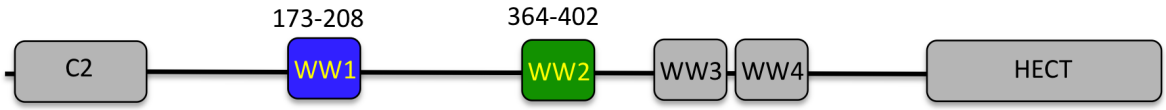


Figure 5.21: Sequence of Smad3 linker peptides synthesized by Fmoc SPPS for ITC binding studies with Nedd4L WW pairs. Phosphorylated amino acids are shown in red.

5.2.1 Binding of the Nedd4L WW1-WW2 pair and Smad3



The Nedd4L WW1-WW2 pair contains a long linker of around 150 amino acids resulting in a molecular mass of almost 24 kDa. The HSQC spectrum of the Nedd4L WW1-WW2 domain pair was acquired and showed a poor dispersion of the amide peaks. Superposition of the Nedd4L WW1 HSQC spectrum (appendix 7.2.3) illustrates that the dispersed peaks correspond to the WW1 domain, whereas characteristic WW2 domain peaks cannot be identified. Signal dispersion did not improve upon Smad3 pT179[PY]-pS204-pS208 titration.

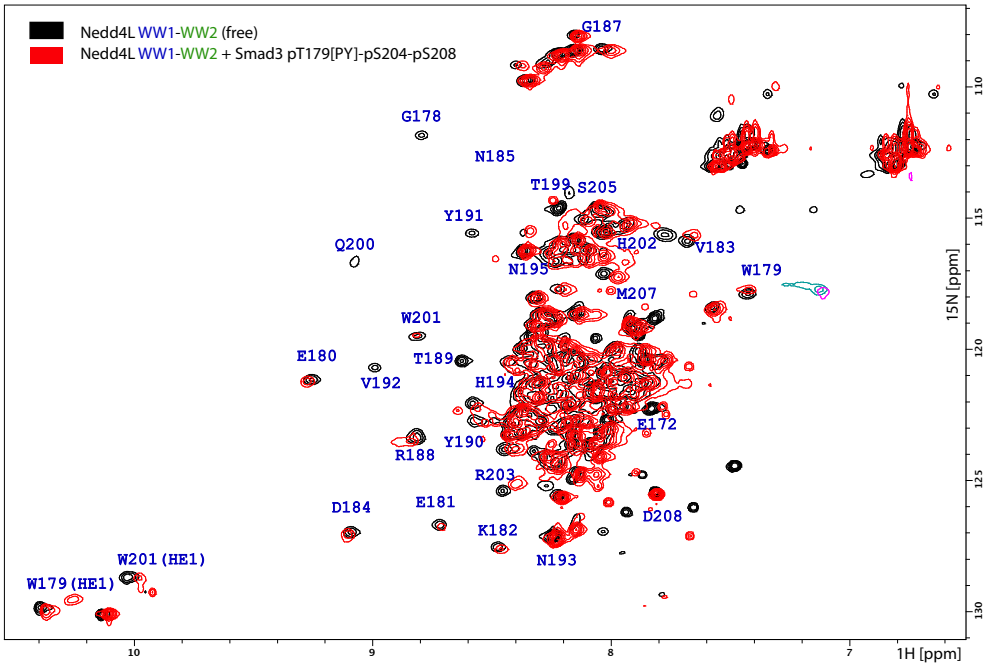


Figure 5.22: HSQC spectrum of Nedd4L WW1-WW2 before (black) and after (red) Smad3 pT179[PY]-pS204-pS208 titration. WW1 residues are labelled in blue.

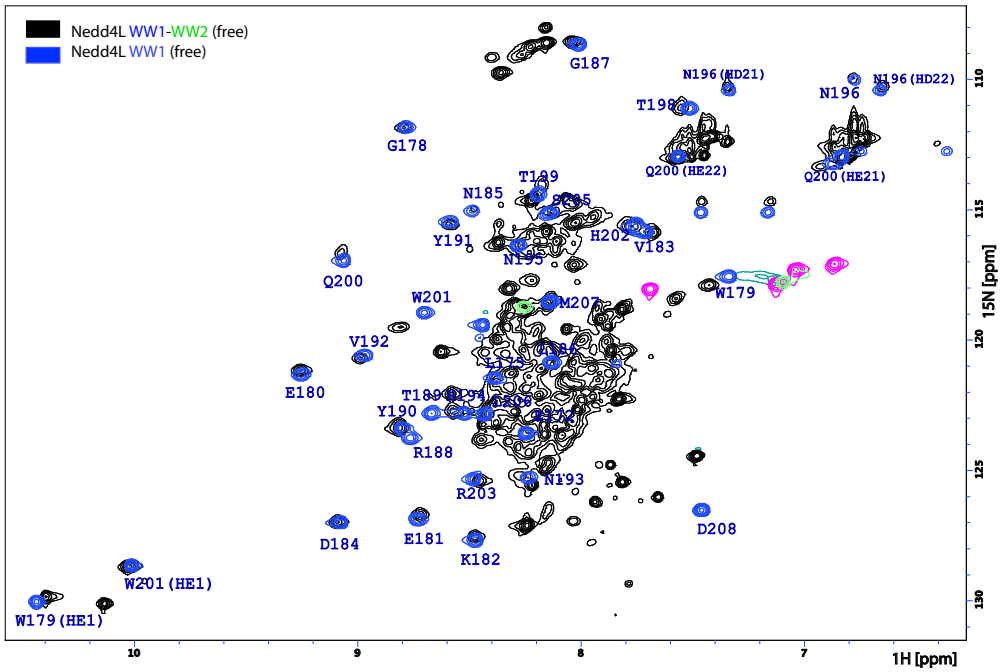
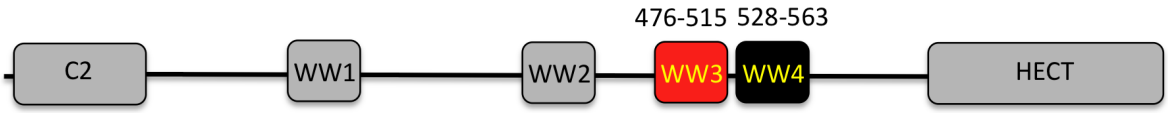


Figure 5.23: Superposition of the HSQC spectrum of the unbound Nedd4L WW1-WW2 pair in black and the spectrum of the single unbound Nedd4L WW1 in blue.

5.2.2 Binding of the Nedd4L WW3-WW4 pair and Smad3



The Nedd4L domains WW3 and WW4 are connected by a short linker of 12 amino acids resulting in a construct with molecular weight of 11 kDa for the WW3-WW4 pair. The HSQC spectrum of the WW3-WW4 pair is well dispersed and triple resonance experiments were acquired to assign the backbone resonances. Upon the addition of increasing amounts of Smad3 pT179[PY]-pS204-pS208 peptide, chemical shift pertubations are only detected for residues located in the WW3 domain and barely observed for WW4, suggesting that only the WW3 domain interacts with the peptide. A comparison of the chemical shifts observed in the WW3 residues to these obtained from a titration with the independent WW3 domain and a short peptide containing the pT[PY] motif suggests that the residues affected in the WW3-WW4 constructs concentrated in the WW3 domain and corresponds to an interaction with the pT[PY] motif.

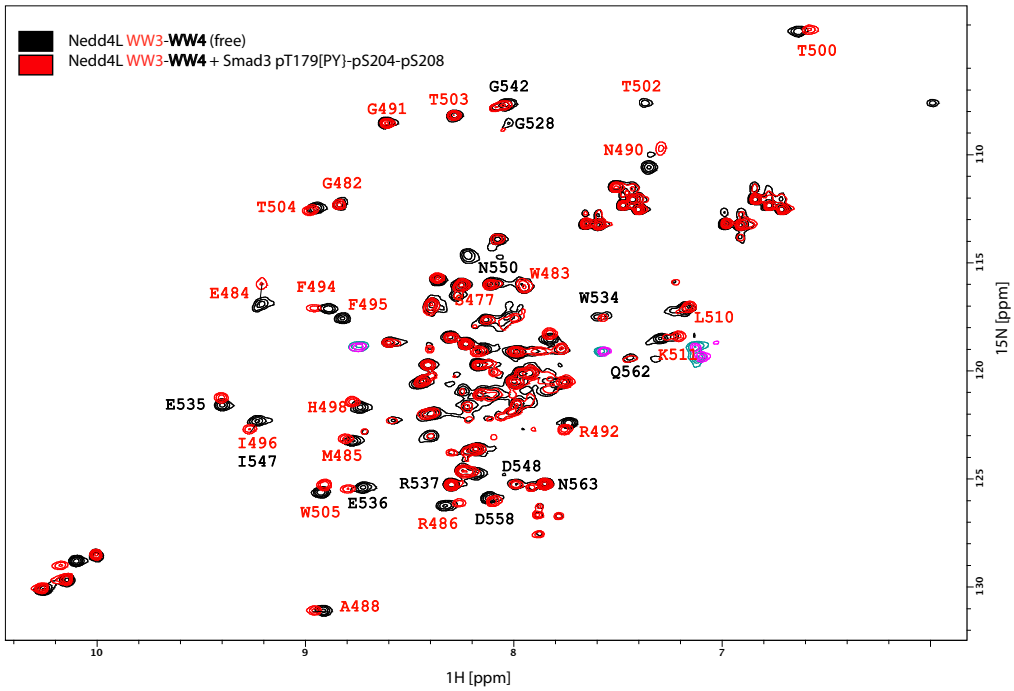
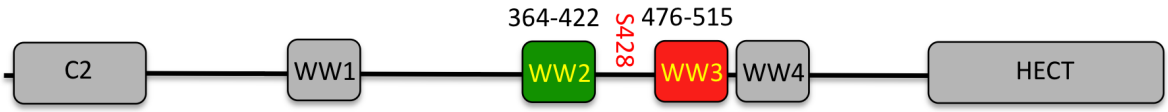


Figure 5.24: Superposition of the HSQC spectra of the Nedd4L WW3-WW4 pair before (in black) and after titration of peptide Smad3 pT179[PY]-pS204-pS208 (in red). Residues that correspond to the WW3 domain are labelled in red; those that correspond to the WW4 domain are labelled in black.

In summary, NMR titration experiments of Nedd4L WW1-WW2 and WW3-WW4 pairs demonstrated that under the experimental conditions used, neither WW1-WW2 nor WW3-WW4 pairs benefit from additional contacts to the phosphorylated C-terminal part of the Smad3 linker.

5.2.3 Binding of the Nedd4L WW2-WW3 pair and Smad3



Nedd4L domains WW2 and WW3 are connected by a long linker of about 70 amino acids including a Ser at position 428 that inhibits Nedd4L activity upon phosphorylation (Gao et al. 2009). However, during *in vitro* ITC and NMR binding experiments Nedd4L WW2 and WW3 could compete with one another for the pT[PY] site since both are pT[PY] binders (K_D = 4.1 ± 0.3 and 17.2 ± 1.7 μM, respectively for the individual domains). *In vivo* a one to one complex between Nedd4L and Smad3 is expected. In order to avoid binding artefacts such as two Smad3 molecules bound to the WW2-WW3 pair, a point mutation at each WW domain

was introduced. Thus, two WW2-WW3 pair samples were prepared, mutant WW2 (V384Y)-WW3 and mutant WW2-WW3 (I496Y), named WW2m-WW3 and WW2-WW3m, with the aim of blocking the PY binding site of WW2 domain in the WW2m-WW3 mutant or that of WW3 in the WW2-WW3m counterpart. HSQC experiments were acquired of both Nedd4L WW2m-WW3 and WW2-WW3m mutants and the backbone residues were identified using triple resonance experiments run on a triple labelled sample. We also used a superposition of the HSQC spectra of the unbound independent domains to that of the pair (Figure 5.4 and Figure 7.3). NMR titration experiments of mutant Nedd4L WW2m-WW and Smad3 pT[PY]-pS204-pS208 showed that chemical shift changes are concentrated in WW3, suggesting that only WW3 binds to the pT[PY] site, while WW2 shows chemical shifts identical to the free domain (Figure 5.25).

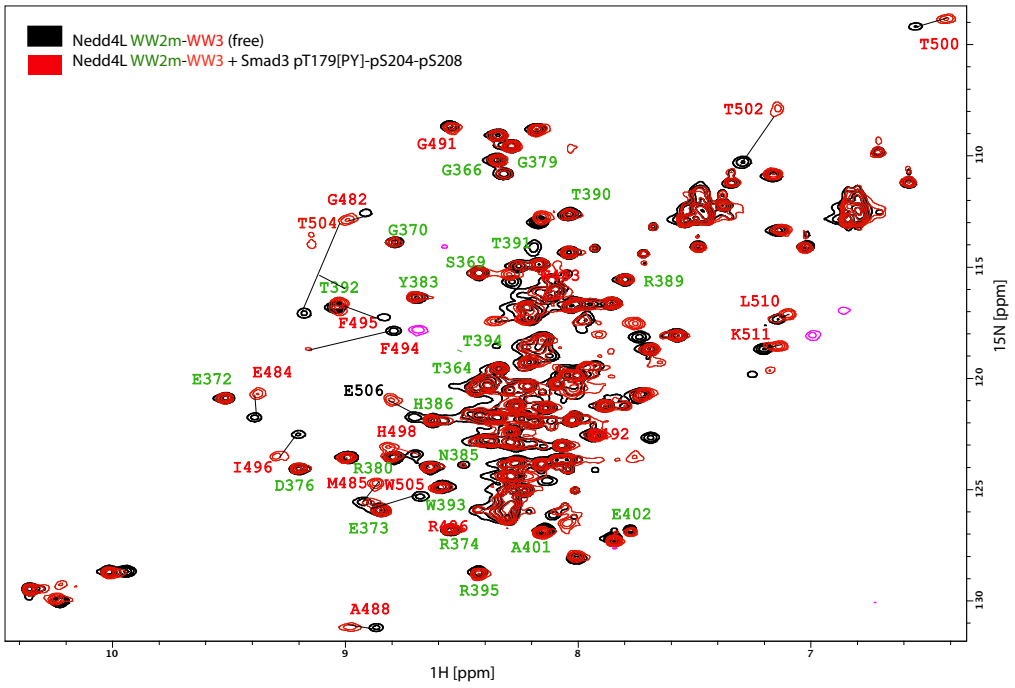


Figure 5.25: HSQC spectra of Nedd4L WW2m-WW3 reference (black) and last point of the titration with Smad3 pT179[PY]-pS204-pS208 (red). Labelled peaks in red correspond to affected WW3 residues, WW2m residues are labelled in green.

NMR titrations with mutant WW2-WW3m showed chemical shift changes in both WW domains. A superposition of the HSQC spectrum of the single WW2 domain bound to the pT[PY] motif to that of the pair revealed that the WW2 residues affected in the pT[PY] complex are also affected in the WW2-WW3m pair when bound to Smad3 pT179[PY]-pS204-pS208 peptide (Figure 5.26).

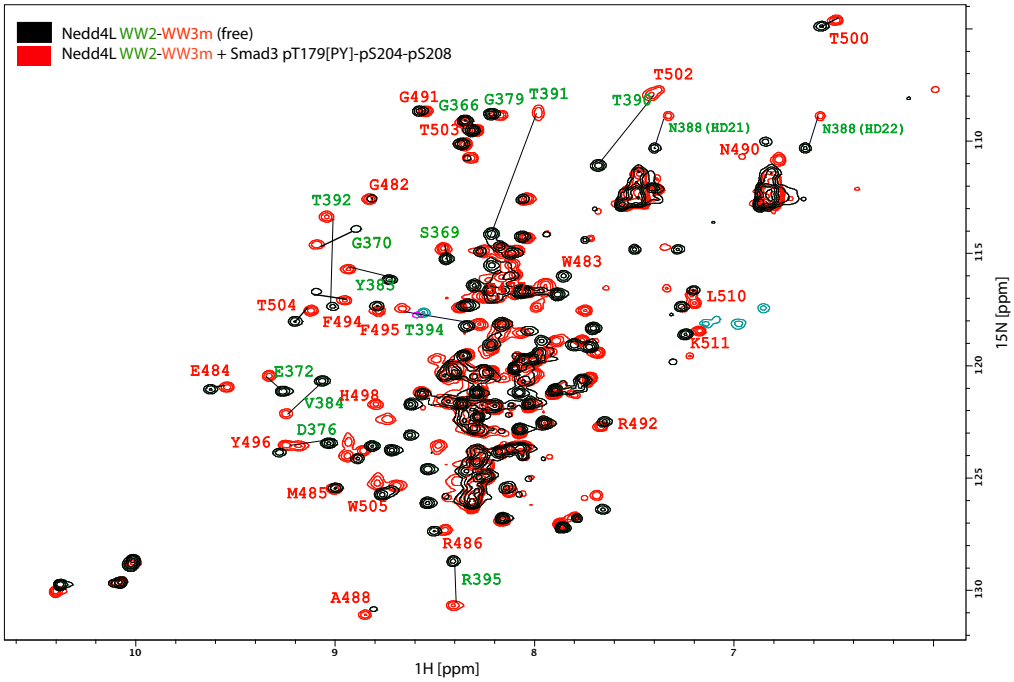


Figure 5.26: HSQC spectra of Nedd4L WW2-WW3m reference (black) and last point of the titration with Smad3 pT179[PY]-pS204-pS208 (red). Labelled peaks in green correspond to affected WW2 residues, WW3m residues are labelled in red.

The remaining amide peaks do not match with the peaks of the free WW3m domain. A backbone assignment of the WW2-WW3m pair revealed that R486, E484, F494, R492, T504 and W505 residues, which correspond to WW3m, are affected upon binding suggesting that they participate in additional contacts to the extended C-terminal part of the peptide.

In order to study the impact of phosphorylation sites pS204' and pS208' on binding the Nedd4L WW2-WW3m pair, ITC binding experiments with peptides Smad3 pT179[PY] (1), Smad3 pT179[PY]-pS208 (2) and Smad3 pT179[PY]-pS204-pS208 (3) were carried out. Binding of the Smad3 pT179[PY]-pS208 peptide increased by almost three-fold (K_D: 8.0 ± 0.7

vs. 2.9 ± 0.2 μM) and by ten-fold (K_D = 0.8 ± 0.1 μM) when bound to the tri-phosphorylated Smad3 pT179[PY]-204-pS208 peptide (Table 5.7).

Table 5.7: ITC values: Nedd4L WW2-WW3m + Smad3 linker peptides (sodium phosphate buffer pH 6.5 at 25 ºC).

	Nedd4L WW2-WW3m				
		K_D [μM]	N	ΔS	ΔH [kcal/mol]
1	Smad3 pT179[PY]	8.0 ± 0.7	1.20	-0.3	-7.1 ± 0.2
2	Smad3 pT179[PY]-pS208	2.9 ± 0.2	1.01	-14.8	-12.0 ± 0.1
3	Smad3 pT179[PY]-pS204-pS208	0.8 ± 0.1	0.94	-23.0	-15.2 ± 0.5

All complex formation reactions were energetically favourable (negative ΔH values) but entropically unfavourable (negative ΔS values). The loss of entropy might be due to motional restrictions upon peptide binding and was particularly obtained for the bi- and tri-phosphorylated peptides (ΔS = -14.8 and -26.0, respectively). The obtained parameters are in agreement with the hypothesis that both, Nedd4L WW2 and WW3m domains interact with the long Smad3 peptides in a cooperative way that constrains the movement of the domains. In the case of the Nedd4L WW2-WW3m - Smad3 pT179[PY] complex formation entropy changes barely occurs with ΔS = -0.3 (Figure 5.27).

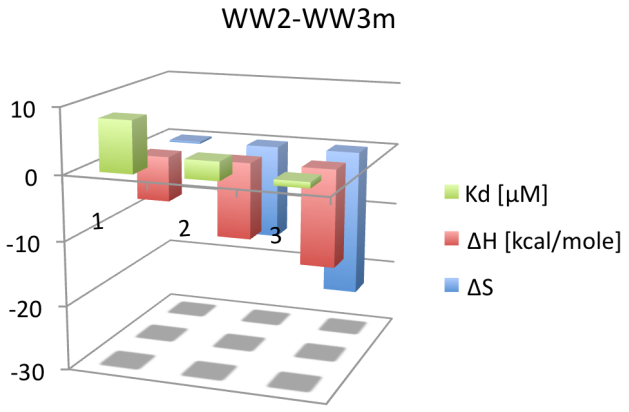


Figure 5.27: ITC fitting parameters for Nedd4L WW2-WW3 with (1) Smad3 pT179[PY], (2) Smad3 pT179[PY]-pS208 and (3) Smad3 pT179[PY]-pS204-pS208.

5.2.4 Chemical ligation of Nedd4L WW2 ($^{2}H^{15}N$) and Nedd4L WW3m (^{15}N)

The Nedd4L WW2 domain binds the Smad3 pT179[PY] motif with the highest binding affinity of all Nedd4L WW domains. ITC experiments of the Nedd4L WW2-WW3m pair and C-terminal extended Smad3 peptides revealed that binding is further increased by adding the bi-phosphorylated peptide and reaches a even lower dissociation constant of 0.8 ± 0.1 μM with the tri-phosphorylated Smad3 peptide. After the identification of the residues affected in the NMR titration experiments it is proposed that Nedd4L recognizes the Smad3 linker in a way that WW2 binds the pT[PY] motif while WW3 interacts with the downstream part of the Smad3 linker peptide.

The Nedd4L WW2-WW3 pair contains 156 amino acids resulting in a molecular weight of around 17 kDa. The high molecular weight and moreover the presence of the almost 70 aa long linker between both domains makes the task of the complex structure determination a challenge. To simplify the interpretation of the NMR data the WW2-WW3m tandem was expressed separately to allow sequential isotope labelling. The domains were then connected using chemical ligation through a disulfide bridge formed by two Cys side chains located at the C-terminal of WW2 and the N-terminal of WW3 (chapter 3.5).

Activation of Nedd4L Cys(Gly)-WW3 (^{15}N)

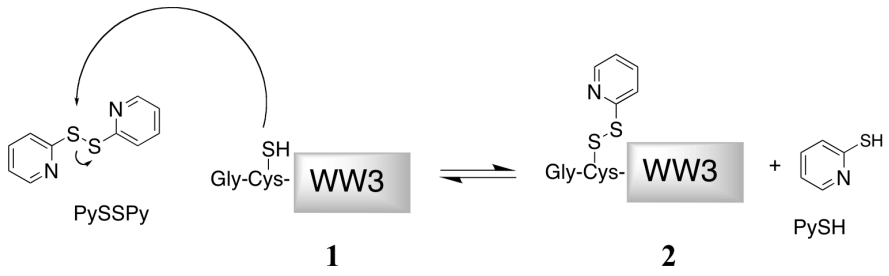


Figure 5.28 Scheme of Nedd4L Gly-Cys-WW3m (^{15}N) **1** activation: Cys lateral thiol group nucleophilic attacks the PySSPy disulfid bridge resulting in activated Nedd4L Gly-*Cys-WW3m (^{15}N) **2** and PySH.

In the first step of the chemical ligation the Nedd4L WW3m domain (Gly-Cys-WW3m (^{15}N) **1**) was activated with 2,2'-dipyridyldisulfide (PySSPy) for 3 h at 25 ºC (chapter 3.5.2). The reaction product Gly-*Cys-WW3m (^{15}N) **2** could be indentified by RP-HPLC as a peak at longer retention time. The product was characterized by MALDI-TOF MS spectrometry with a mass of 5334 g/mol. This mass represents an increase of 110 units with respect to the value

of the starting Gly-Cys-WW3m (^{15}N) molecule. The mass difference of 110 units corresponds to the molecular weight of 2-thiopyridin bound to the side chain of the N-terminal Cys residue. The measured masses (5225 and 5334 g/mol) of Gly-Cys-WW3m (^{15}N) and Gly-*Cys-WW3m (^{15}N) are slightly (0.3%) smaller than the calculated masses (5239 and 5349 g/mol, respectively). This difference is due to a 99.7% incorporation of isotope ^{15}N. Errors generated by the 4800 *Plus* MALDI TOF/TOF instrument used for this measurements are usually around 0.05%.

Table 5.8: Calculated vs. measured molecular weights of the activation reaction.

Nedd4L	Mw (measured) [g/mol]	Mw (calculated) [g/mol]	Δ [g/mol]	%
Gly-Cys-WW3m (^{15}N)	5225	5239	14	99.7
Gly-*Cys-WW3m (^{15}N)	5334	5349	15	99.7

Excess of PySSPy appeared as a prominent peak at a longer retention time than the product peak. Gly-*Cys-WW3m(^{15}N) could be separated most effectively from PySSPy with a Vydac C4 column. Between the product and PySSPy peaks a peak corresponding to the WW3m-Cys-Cys-WW3m(^{15}N) dimer was identified by MALDI-TOF MS spectrometry.

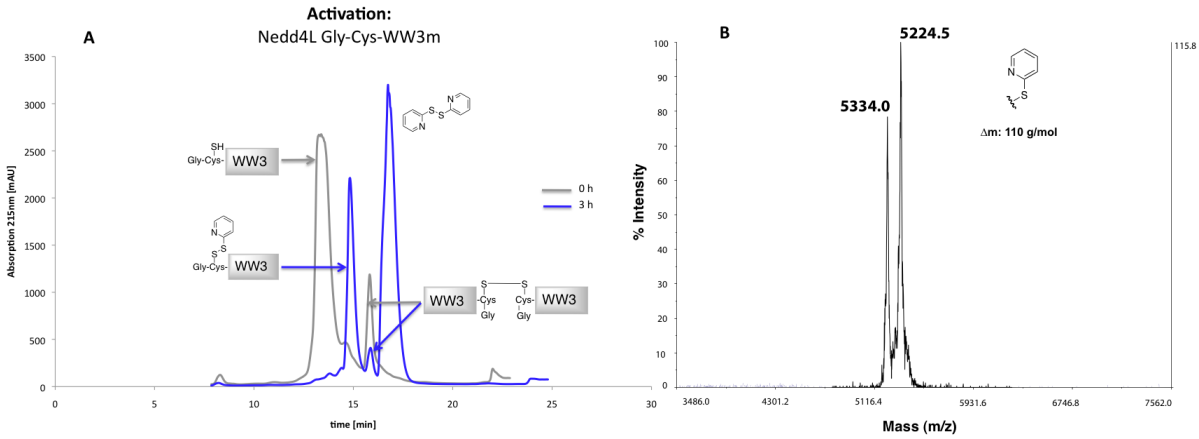


Figure 5.29: (A) RP-HPLC profile of the activation reaction. Grey: Nedd4L Gly-Cys-WW3m(^{15}N) sample before the reaction, blue: Gly-*Cys-WW3m(^{15}N) + PySSPy (after 3h reaction time). (B) MALDI-TOF MS spectrum of Nedd4L Gly-Cys-WW3m(^{15}N) and activated Gly-*Cys-WW3m(^{15}N).

Chemical ligation of Nedd4L WW2-linker-Cys ($^{2}H/^{15}N$) and *Cys(Gly)-WW3m (^{15}N)

In the ligation reaction, a disulfide bridge between the activated WW3m domain (Gly-*Cys-WW3m(^{15}N) **2**) and the WW2-linker domain (WW2-linker-Cys(^{2}H,^{15}N) **3**) was formed (Figure 5.30).

Figure 5.30: Scheme of the chemical ligation reaction: Nedd4 WW2-linker-Cys(^{2}H,^{15}N) **3** nucleophilic attack Nedd4L Gly-*Cys-WW3m(^{15}N) **2** resulting in the ligation product Nedd4L WW2-linker-Cys-Cys(Gly)-WW3m(^{2}H,^{15}N/^{15}N) **4.**

The ligation reaction was monitored by analytic RP-HPLC on a Vydac C18 column after 3, 6 and 21 h. The new peak corresponding to the product appeared at a retention time that lies between those of the two unbound domains. During the reaction course the peak product is increasing while the reactants are continuously decreasing. In addition, a prominent peak at short retention time is continuously growing and was identified as the Gly-*Cys-WW3m(^{15}N) leaving group PySH.

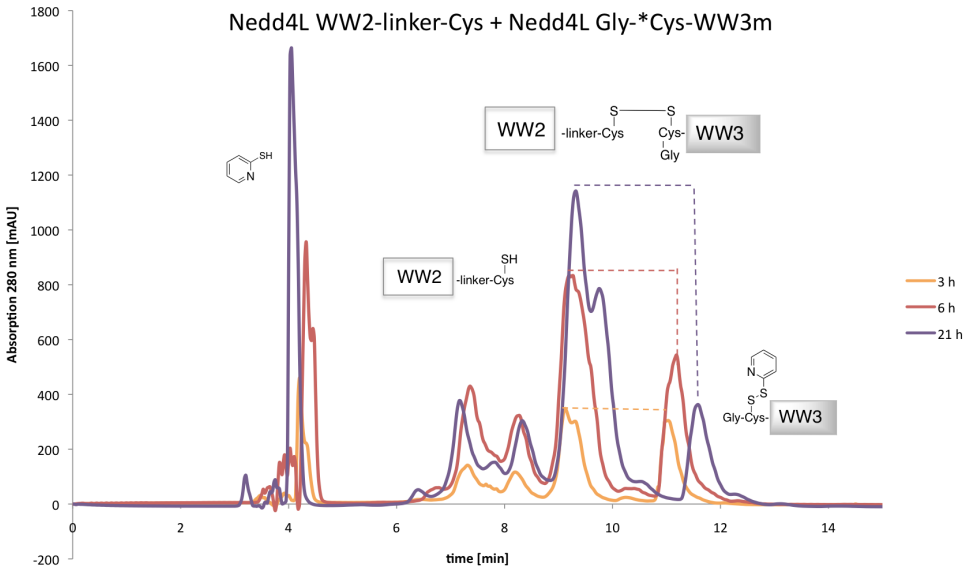


Figure 5.31: RP-HPLC chromatograms of the chemical ligation after 3h (yellow), 6h (red) and 21h (purple) reaction time. Increasing peaks: leaving group PySH (retention time: 4 min), product peak (retention time: 9-10 min, shoulder: product with and without His_6-tag). Decreasing peaks: Nedd4 WW2-linker-Cys(^{2}H,^{15}N) (retention time: 6-9 min, shoulder: WW2-linker-Cys(^{2}H,^{15}N) with and without His_6-tag), Gly-*Cys-WW3m(^{15}N) (retention time: 11-12 min).

The measured mass (5324 g/mol) of remaining Gly-*Cys-WW3m (^{15}N) after the ligation reaction is as previously observed in the activation reaction slightly (0.5%) smaller than the calculated masses (5349 g/mol). This corresponds to a high ^{15}N incorporation rate of 99.5% (99.7% obtained previously). The incorporation of isotope ^{2}H is smaller resulting in double labelled samples of 97.4% for WW2-linker-Cys (^{2}H,^{15}N) and of 97.9% for the ligation product WW2-linker-Cys-Cys(Gly)-WW3m (^{2}H,^{15}N/^{15}N). The mass difference is detectable by mass spectrometry but the incorporation rate is high enough to obtain a 1 mM sample for the desired NMR experiments.

Table 5.9: Calculated vs. measured molecular weights of the chemical ligation components.

Nedd4L	Mw [g/mol] (measured)	Mw [g/mol] (calculated)	Δ Mw [g/mol]	%
WW2-linker-Cys (^{2}H,^{15}N)	16022	16446	444	97.4
Gly-*Cys-WW3m (^{15}N)	5324	5349	25	99.5
WW2-linker-Cys-Cys(Gly)-WW3m (^{2}H,^{15}N/^{15}N)	21221	21669	448	97.9

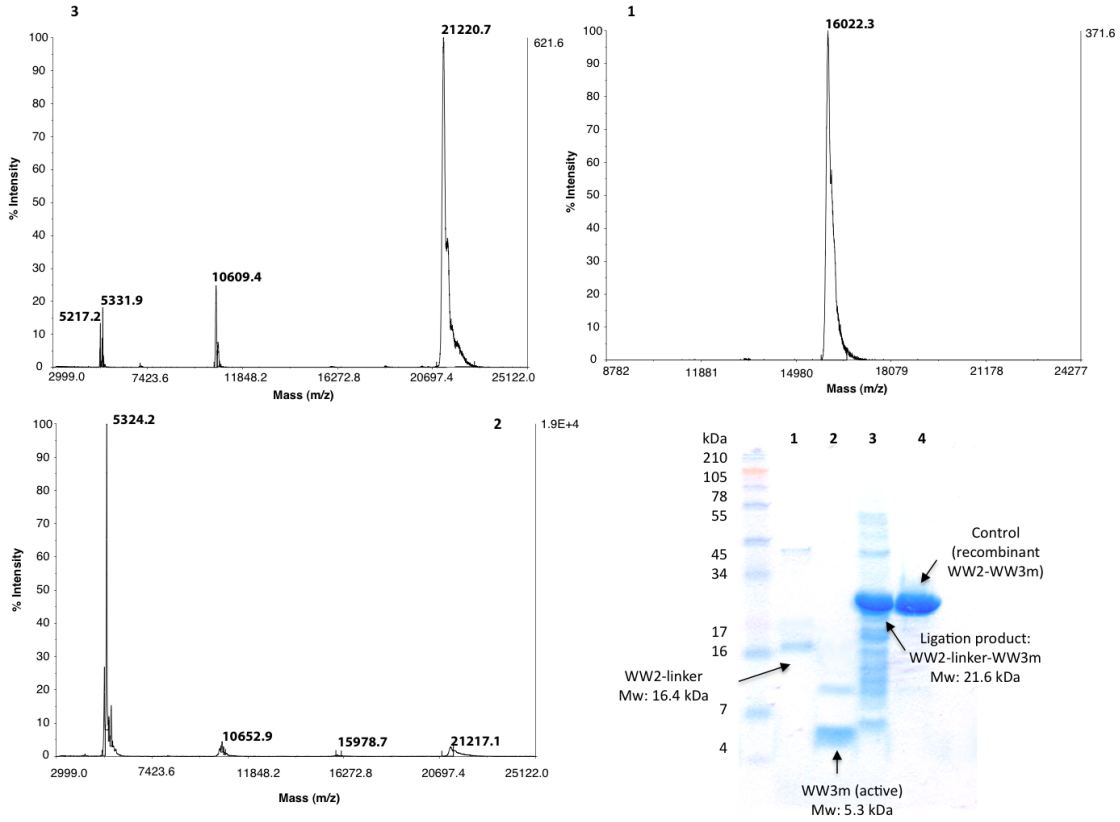


Figure 5.32: MS spectra and SDS PAGE of native chemical ligation. (1) WW2 linker, (2) activated WW3m, (3) WW2-linker-WW3m ligation product and (4) recombinant Nedd4L WW2-WW3m as control.

The Nedd4L WW2-WW3m sample obtained by chemical ligation was compared with the wild type Nedd4L WW2-WW3m by superposing both HSQC spectra. As shown in Figure 5.33 both chemical shift patterns are almost identical.

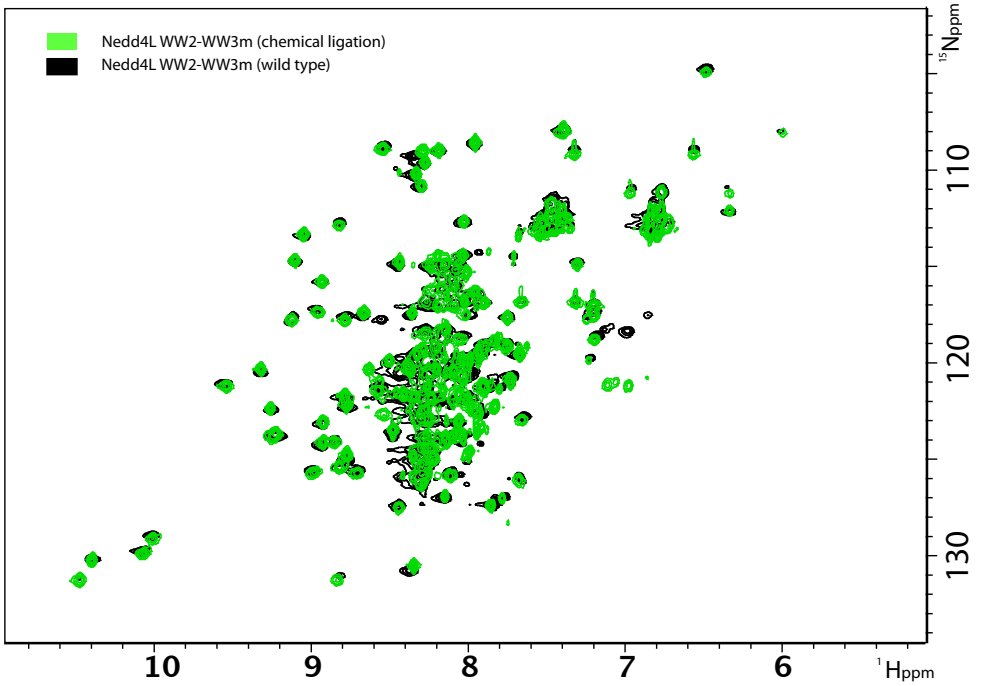


Figure 5.33: Superposition of the HSQC spectra of the Nedd4L WW2-WW3m sample obtained by chemical ligation in green and the Nedd4L WW2-WW3m wild type in black.

5.2.5 Structure determination of the Nedd4L WW3 – Smad3 pS-pS complex

The sequential isotope labelling approach allows to determine Nedd4L WW3m domain binding of the Smad3 pS-pS part within the WW2-WW3m - Smad3 pT179[PY]-pS204-pS208 complex. The backbone and side chain chemical shifts of the free Nedd4L WW3m domain were determined by multidimensional heteronuclear NMR spectroscopy and the information was transferred to the HSQC, ^{15}N-edited NOESY and homonuclear TOCSY/NOESY spectra of the Nedd4L WW2-linker(^{2}H,^{15}N)-WW3m(^{15}N) ligation product. The Smad3 pT179[PY]-pS204-pS208 peptide was added up to three-fold peptide excess and the 2D TOCSY/NOESY spectra of the Nedd4L WW2-linker(^{2}H,^{15}N)-WW3m(^{15}N) - Smad3 pT179[PY]-pS204-pS208 complex were acquired. NOEs of the WW3 domain and amino acids of the downstream part of the tri-phosphorylated Smad3 peptide could be detected. The side chain of loop 1 residues N490 and the guanidinium group of residue R492 bind the phosphate group of pS208', which also contacts the aromatic ring of F494. Additional contacts of P209' to the aromatic ring of W505

were identified. pS204' is bound by the guanidinium group ß1-sheet residue R486 and the hydrophobic part of E484. P205' pyrrolidin ring is accommodated between the side chains of E484, R486 and Y496.

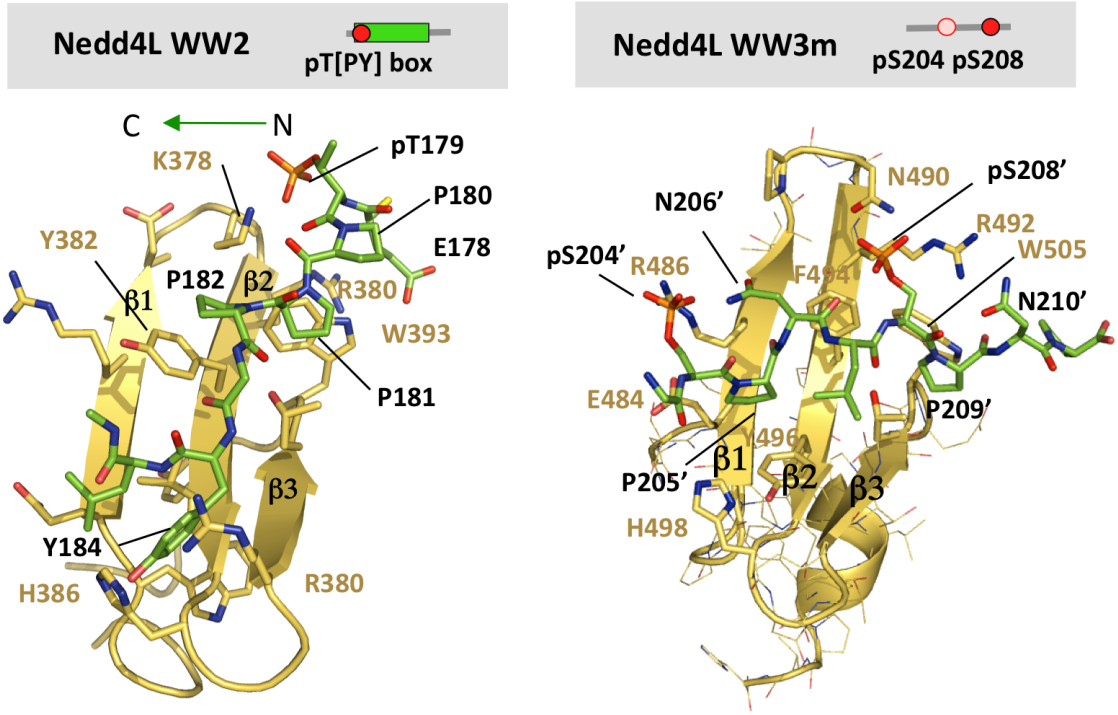


Figure 5.34: Cartoon representation of Nedd4L WW2 bound to Smad3 pT179[PY] and Nedd4L WW3m bound to the downstream part of Smad3 containing pS204 and pS208. WW domains are illustrated in gold with key residues labelled in black, Smad3 ligand is shown in green with residues labelled in black.

In order to confirm the relevance of residues R486 and R492 on Smad3 binding two Ala point mutants Nedd4L WW2-WW3m(R486A) and WW2-WW3m(R492A) were prepared (Figure 5.35).

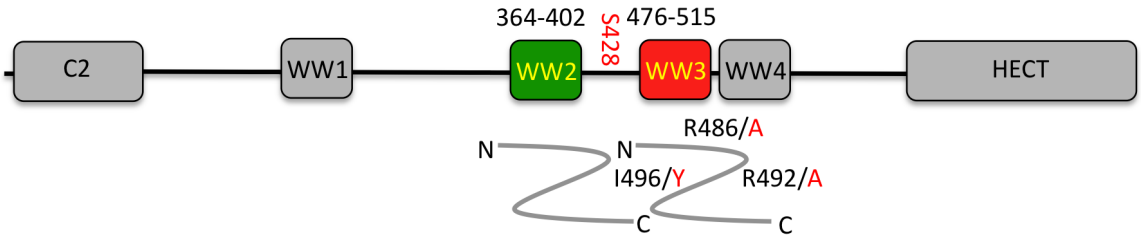


Figure 5.35: Schematic representation of Nedd4L with the WW2-WW3m constructs in colour. Point mutations located at WW3m, WW2-WW3m interdomain phosphorylation site in red.

ITC binding experiments confirmed the impact of the two positively charged ß1- and 2 strand Nedd4L WW3 residues R486 and R492. Alanine point mutations of either R486 or R492 decreased binding of both mutants almost equally by around two-fold from K_D = 0.8 ± 0.1 μM (WT) to K_D = 1.6 ± 0.1 and 1.4 ± 0.1 μM, respectively (Table 5.10).

Table 5.10: ITC parameters: Nedd4L WW2-WW3m mutants + Smad3 pT179[PY]-pS204-pS208 (sodium phosphate buffer pH 6.5, 25 °C).

Nedd4L	**Smad3 pT179[PY]-pS204-pS208**	
	K_D [μM]	N
WW2-WW3m(WT)	0.8 ± 0.1	0.9
WW2-WW3m(R486A)	1.6 ± 0.1	1.0
WW2-WW3m(R492A)	1.4 ± 0.1	0.9

5.3 Binding of the I-Smad7 linker and WW domains

5.3.1 Smad7 linker binds Nedd4L WW2 domain with high affinity

The inhibitory Smad7 can negatively regulate the TGF-ß pathway with the help of E3 ubiquitin ligases (Kavsak et al. 2000). It has been shown that the E3 ubiquitin ligases Smurf1 and Smurf2 can recognize Smad7 and transfer the inhibitory Smad from the nucleus to the cytoplasm. In the cytoplasm the E3 ligase – Smad7 complex binds to the activated TGF-ß receptor type I and triggers the degradation of the heteromeric TGFßI/II receptor complex by ubiquitination (chapter 1.3). Binding of E3 ubiquitin ligases and Smad7 occurs between the WW domain region of the ligase and the Smad7 linker. The linker region of Smad7 contains like R-Smads2/3 a PY motif that can be recognized by the WW domains but Smad7 does not consist of a preceding phosphorylation site. A yeast two-hybrid screening of a HeLa cDNA library and a full-length Smad7 as bait led a few years ago to the identification of Nedd4L as a binding partner (Kuratomi et al. 2005). In order to identify the regions involved in the Nedd4L-Smad7 interaction, ITC binding assays were performed between recombinant proteins containing WW segmentes and a 15-residue Smad7 peptide.

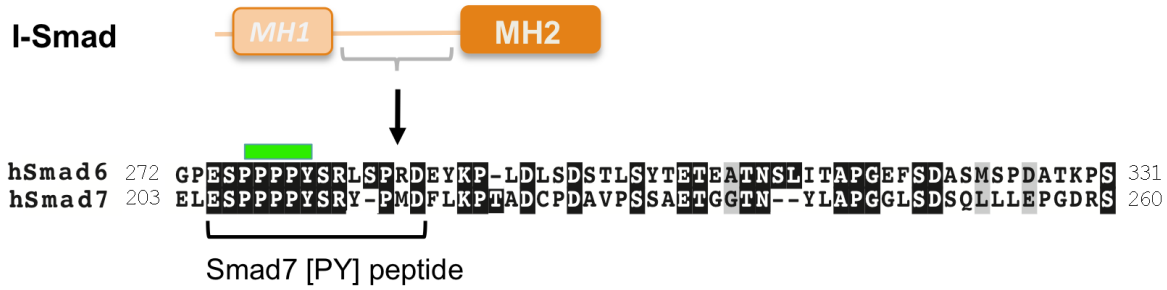


Figure 5.36: Sequence alignment of the Smad6 and Smad7 linker. Smad7[PY] peptide (E203'- D217') was synthesized by Fmoc SPPS. Alignments were performed using *ClustalW* (Thompson et al. 1994).

For the experiments Nedd4L fragments containing either each of the four independent domains or the three consecutive WW1-WW2, WW2-WW3 and WW3-WW4 pairs were used. The ITC experiments revealed that the four independent domains bound the peptide with a one to one protein-peptide ratio stoichiometry and dissociation constants of 23.6 ± 3.6, 4.3 ± 0.2, 8.0 ± 0.3 and 12.4 ± 1.8 µM, respectively. Slightly weaker values were observed for the WW1-WW2, WW2-WW3 and WW3-WW4 pairs, with K_D values of 18.2 ± 7, 5.2 ± 0.3 and

16.9 ± 1.0 μM, respectively. In these cases the stoichiometry is one protein - one ligand except for WW2-WW3, which is of one protein - two peptides bound to the recombinant WW pairs, something very unlikely to occur *in vivo*. To prevent the presence of this artefact WW2-WW3 the same constructs as in chapter 5.2.3 were used containing point mutations at each WW domain, with the aim of reducing the interaction of one of the WW domains to the PY site, WW2(V384Y)-WW3 (WW2mWW3) and WW2-WW3(I496Y) (WW2WW3m). According to these values Nedd4L WW2 is the preferred Smad7[PY] binding domain.

Table 5.11: K_D values of the interaction between Nedd4L WW1-4 and Smad7[PY] (sodium phosphate buffer pH 6.5, 25 ºC).

Nedd4L		**Smad7[PY]**
		K_D [μM]
WW1		23.6 ± 3.6
WW2	**(A)**	4.3 ± 0.2
WW3		8.0 ± 0.3
WW4		12.4 ± 1.8
WW1WW2 (wt)		18.2 ± 7.0
WW2WW3 (wt)	**(B)**	5.2 ± 0.3
WW2WW3m	**(C)**	5.6 ± 0.2
WW2mWW3	**(D)**	7.5 ± 0.2
WW3WW4 (wt)		16.9 ± 1.0

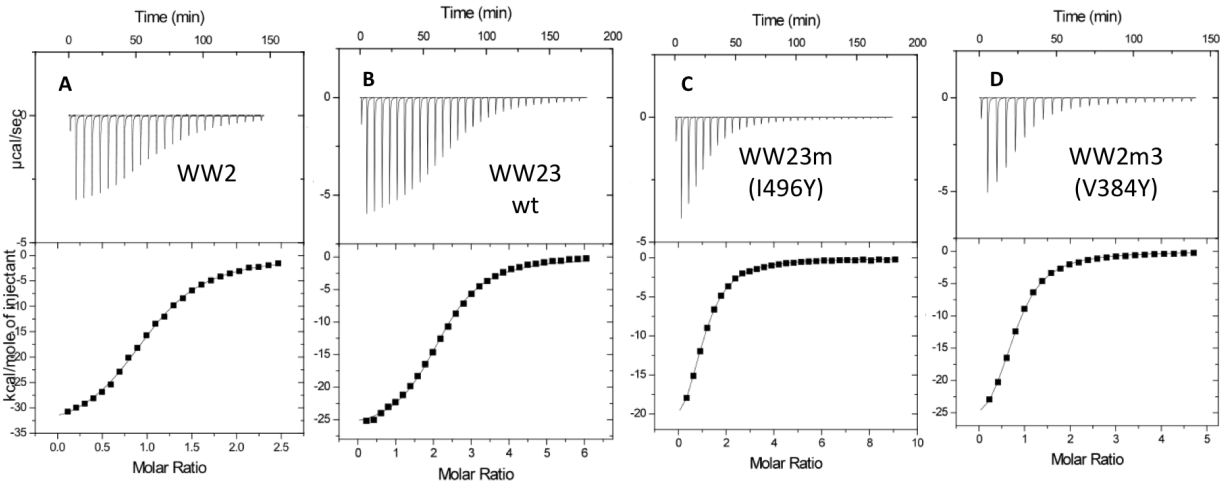


Figure 5.37: ITC profiles of Smad7[PY] titrated to Nedd4L (A) WW2, (B) WW23, (C) WW23m and (D) WW2m3.

The ITC experiments indicate that Nedd4L WW2 binds the nonphosphorylated Smad7[PY] ligand with a similar binding affinity as that of the phosphorylated Smad3 pT179[PY] peptide (K_D = 4.3 ± 0.2 vs. 4.1 ± 0.3 μM, respectively) although Smad7 lacks a phosphorylation site upstream of the PY motif.

To corroborate these interactions and to identify the domain-binding site, NMR-based titrations were performed using ^{15}N-labeled recombinant proteins and the Smad7 peptide and the titrations were followed by ^{15}N-^{1}H-HSQC experiments. Once the equilibrium is close to the 1:1.5 protein: peptide ratio, new chemical shift values are observed in residues present in the three strands, suggesting that the peptide is accommodated using a large binding surface.

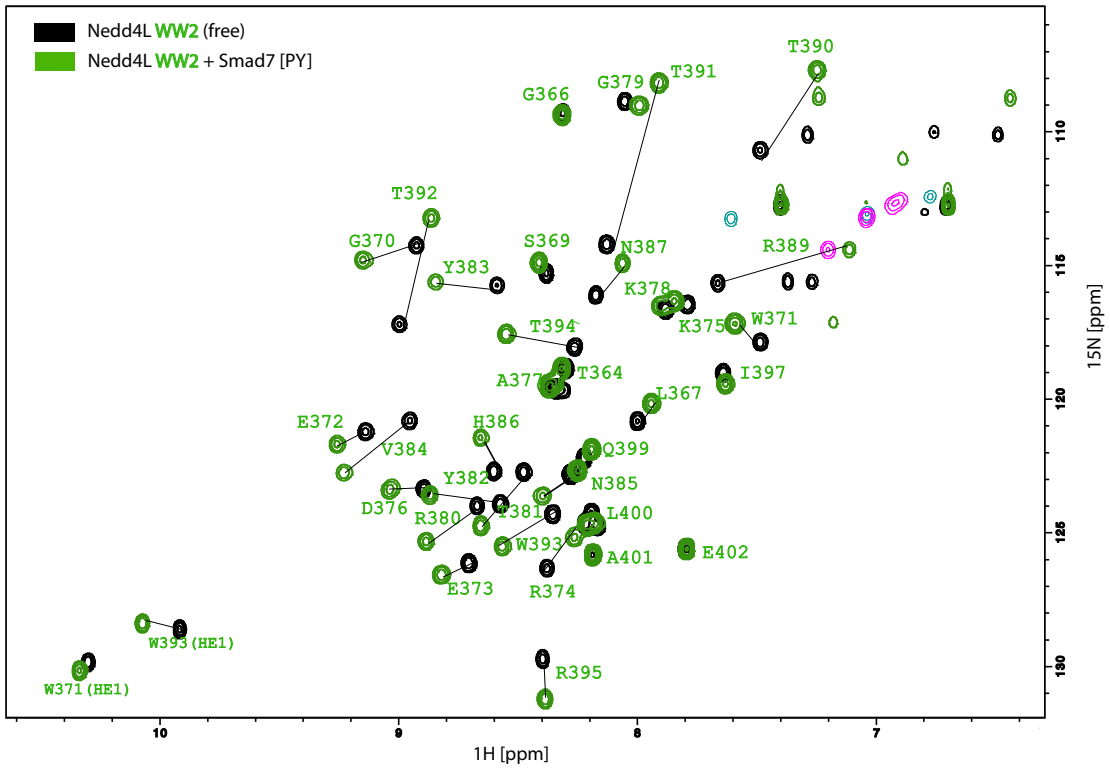


Figure 5.38: ^{1}H ^{15}N-HSQC spectrum of the Nedd4L WW2 domain in the free (black) and the bound state (green) after saturation with the Smad7[PY] peptide.

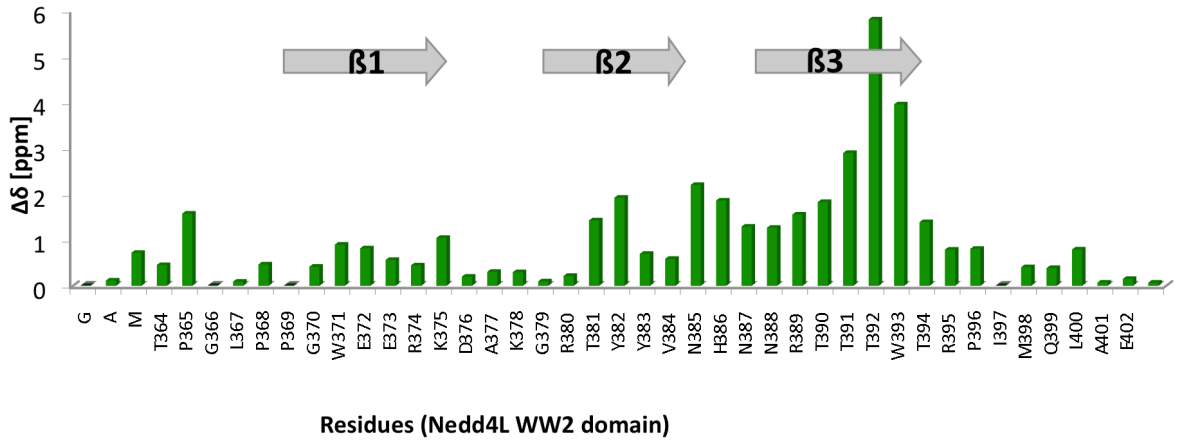


Figure 5.39: Histogram of the average chemical shift changes observed for Nedd4L WW2 backbone protons after saturation with the Smad7[PY] peptide with respect to the free domain.

5.3.2 Structural determination of the Nedd4L WW2 - Smad7[PY] complex

Triple resonance NMR spectroscopy was applied to assign the WW2 domain in the Nedd4L WW2 - Smad7[PY] complex, whereas homonuclear spectra were used to assign the Smad7 peptide and its contacts with the WW domain.

The structure of the Nedd4L WW2 domain bound to the Smad7[PY] ligand adopts the typical WW fold, consisting of an anti-parallel triple stranded ß-sheet stabilized by a network of hydrogen bonds and side chain interactions. However, additional intermolecular NOEs between domain and peptide residues could be identified. Residue Y211' of the Smad7 PY motif is located in the tyrosine-binding pocket of the WW2 domain, formed by V384, H386, T391, R389 and the T390 methyl group, whereas P207', P208' and P209' contact the aromatic protons of W393 and Y382 that define the XP binding site. NOEs between T392 α-, β-protons and γ-(methyl)-protons and P208' could unambiguously be identified as well as NOEs between D376 α-, β-protons and P209' pyrrolidine protons ß and γ (Table 5.14). Overall all these NOEs localize the PY motif of the Smad7 peptide on the surface of the domain and define the peptide orientation.

In this structure, apart from the above mentioned interactions which are consistent with structures of other WW domain complexes (Macias et al. 2000), additional intra-molecular NOEs between the residues located downstream of the PY motif in the ligand and residues located in the first and second strands of the domain were identified. The amide proton of

S212' contacts T391 and V384 methyl groups as well as abundant backbone and side chain interactions between the R213' and H386 residues. Furthermore the R213' side chain interacts with both aromatic rings of the Y211' and Y214' residues, thus stabilizing the ligand structure. This interaction occurs because the Smad7 fragment forms a turn in the Y211' R213' Y214' region. The turn is further stabilized by backbone interactions of R213' and Y214', as well as by contacts with the side-chain of R374 located in the first strand and the P215' pyrrolidine ring. A similar turn was also observed in the Smurf2WW3–Smad7[PY] (Chong et al. 2006).

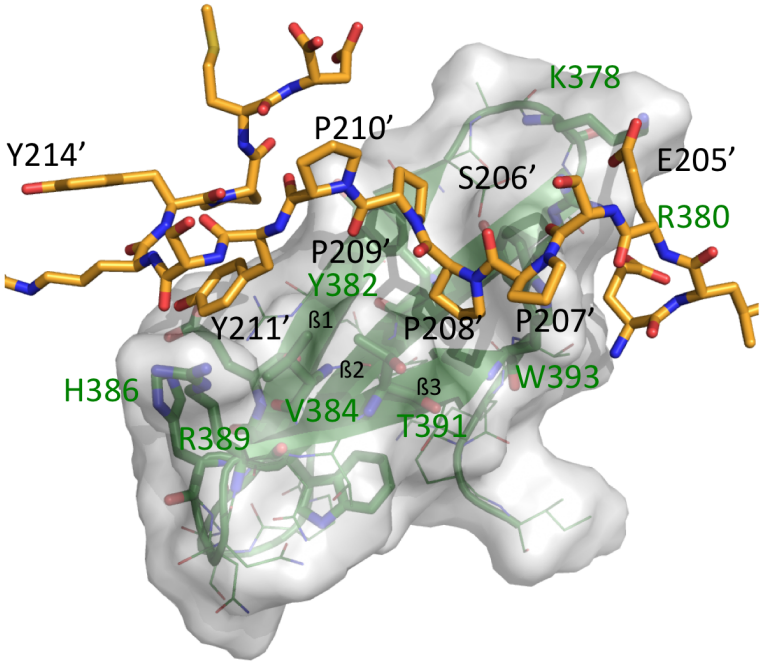


Figure 5.40: Lowest-energy structure of the Nedd4L WW2 – Smad7[PY] complex shown as a solid surface/cartoon representation with domain in green and the peptide as sticks in yellow. Peptide residues are labelled in black, domain residues are labelled in green.

Through NMR structural analysis intermolecular NOEs of nine Nedd4L WW2 residues (D376, Y382, V384, H386, R389, T390, T391, T392 and W393) and seven Smad7[PY] residues (P207' P208', P209', Y211', S212', R213' and P215') were identified. These contacts provide a very large binding interface that explains the high binding affinities towards all Nedd4L WW domains. The Smad7[PY] complex interaction seems to be unique regarding the binding interface. A Smad3 pT[PY] complex formed using a C-terminus extended segment with three additional amino acids (Smad3 pT179[PY]-long) did not increase the binding affinity to the Nedd4L WW2 domain (Table 5.12). 2D TOCSY/NOESY spectra of the Nedd4L WW2 –

Smad3 pT[PY]-long complex have been assigned but additional inter- or intramolecular NOEs like observed for Nedd4L WW2 – Smad7[PY] complex were absent.

Table 5.12: K_D values of Nedd4L WW2 + Smad7 vs. Smad3 pT179[PY]-long (phosphate buffer pH 6.5 at 25 °C).

Peptide	Sequence	Nedd4L WW2
		K_D [µM]
Smad2 pT220[PY]	PEpTPPPGYISEDG	6.7 ± 0.3
Smad3 pT179[PY]	PEpTPPPGYLSEDG	4.1 ± 0.3
Smad3 pT179[PY]-long	PEpTPPPGYLSEDGETS	6.5 ± 0.4
Smad7 [PY]	ELESPPPPYSRYPMD	4.3 ± 0.2

The structure of the complex further revealed that the N-terminal part of Smad7[PY] fragment interacts with residues located in the first loop of the Nedd4L WW2 domain. Interestingly, in the Nedd4LWW2 – Smad7[PY] complex, most of these contacts are detected between the side chain of E205' and two positively charged residues of loop 1, Lys378 and Arg380, whereas in the Nedd4LWW2 – Smad3 pT179[PY] complex contacts were observed with the phosphate group of T179'. Although the electrostatic interactions are conserved in both complexes (pT179' in Smad3 and E205' in Smad7 are negatively charged) the position of these residues in each peptide is different. In Smad3 the phosphorylated residue is directly upstream of the polyproline site (position -1) while in Smad7 the glutamic acid is in position -2.

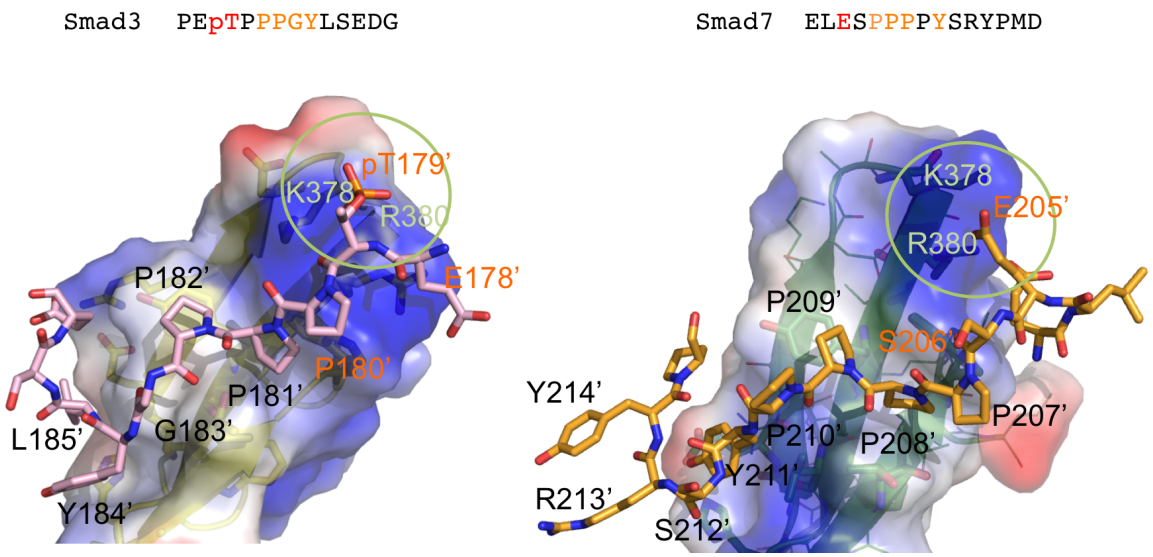


Figure 5.41: Surface electrostatic representation of the Nedd4L WW2 - Smad3 pT179[PY] complex vs. Nedd4L WW2 – Smad7 complex. K378 and R380 coordinate the phosphate group of Smad3 pT179 or bind E205' in Smad7[PY].

The complex structure reveals an electrostatic interaction between the side chains of E205' and Arg380 and Lys378 that contributes to the final high binding affinity measured in the Nedd4LWW2–Smad7[PY] interaction. To dissect the relevance of these contacts the interaction of protein variants with the Smad7 peptide were measured using ITC in which either R380 or K378 or both R380/K378 were replaced by a negatively charged residue (Glu). In the case of the R380E mutant, the binding affinity became around four-fold weaker (K_D = 15.4 ± 3.5 μM), around five-fold weaker in the case of the K378E mutant (K_D = 23.5 ± 1.5 μM) and decreased by more than six-fold (K_D = 27.3 ± 2.2 μM) when both residues mutated. Thus, these results reveal that the N-terminal part of the peptide, where the E205' residue is located, interacts with the domain. (Table 5.13).

Table 5.13: K_D values of Nedd4L WW2 mutants and Smad7[PY] (sodium phosphate buffer pH 6.5 at 25 ºC).

Nedd4L WW2	Smad7[PY]
	K_D [μM]
K378 R380 (WT)	4.3 ± 0.2
R380E	15.4 ± 3.5
K378E	23.5 ± 1.5
K378E R380E	27.3 ± 2.2

Table 5.14: Inter- and intramolecular peptide NOEs of the Nedd4L WW2 – Smad7 complex (for proton nomenclature see appendix 7.6).

Smad7[PY] - Nedd4L WW2					
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	HD2/3			S206'	HA
	HA	W393	HE1	P208'	HD2/3
			HD1		
P207'	HB2	W393	HE1		
	HG2/3	W393	HE1		
			HD1		
			HZ2		
	HA	W393	HE1	P209'	HD2/3
P208'			HH2		
			HD1		
			HZ3		
		Y382	QD		
	HD2/3	W393	HE1	P207'	HA
			HN		
			HD1		
			HZ3		
			HZ2		
		T392	HA		
	HG2	W393	HE1		
			HN		
			HD1		
		T392	HN		
	HG3	W393	HE1		

Smad7[PY] - Nedd4L WW2					
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
			HE3		
			HH2		
			HZ2		
		T392	HA		
	HB2/3	W393	HE1		
			HN		
			HD1		
			HZ2		
		Y382	QD		
	HB2	T392	HA		
	HA	Y382	QE	P210'	HD2/3
	HD2/3	W393	HH2	P208'	HB3
			HZ2		HG3
			HZ3		
		D376	HB2/3		
P209'		Y382	QD/E		
	HB2/3	W393	HH2		
		Y382	QE		
	HG2/3	Y382	QE	P208'	HG3
		D376	HB2/3		
			HA		
Y211'	HN	V384	QG2	P209'	HA
		T391	QG2	S212'	HN

Smad7[PY] - Nedd4L WW2					
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	HA	V384	QG1/2	P209′	HB2
		T391	QG2	S212′	HN
	HB2/3	V384	QG1/2	S212′	HN
	QD/E	H386	HA	S212′	HA
		R389	HA	R213′	HB2/3
			HB2/3		
			HG2/3		
			HD3		
			QG2		
		V384	HB	R213′	HA
			QG1/2	P210′	
		T390	QG2		
	HN	T391	QG2	Y211′	HN
		V384	QG2		QD/E
S212′					HA
					HB2/3
	HA				QD/E
	HN	H386	HN	Y214′	HN
					QD
				Y211′	QD/E
				M216′	QE
					HB3
	HA	H386	HE1	Y214′	HN
			HN	Y211′	QD/E
R213′			HD2		
	HD2/3			Y214′	QD/E
	HB2/3	H386	HE1	Y211′	QD/E
				Y214′	HN
				Y214′	QD/E
	HG2/3	H386	HE1	Y214′	QD/E
				Y214′	HN
Y214′	HN			R213′	HN
					HB2/3
					HG2/3
				P215′	HD2/3

Smad7[PY] - Nedd4L WW2					
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
	QD			P215′	HD2/3
				R213′	HD2/3
					HB2/3
					HG2/3
				M216′	HG2
	QE			P215′	HD2/3
				R213′	HD2/3
					HB2/3
					HG2/3
	HA			P215′	HD2/3
	HB2/3			P215′	HD2/3
	HB3			P215′	HA
	HB2			P215′	HG2
	HA			M216′	HN
				Y214′	HB3
	HD2/3			Y214′	HN
P215′					QD/E
					HB2/3
	HB2	Y382	QD	M216′	HN
	HB3	Y382	QD		
	HN			D217′	HN
				P215′	HA
	HA			D217′	HN
M216′	HB2/3			D217′	HN
				R213′	HN
	HG2/3			D217′	HN
				Y214′	QD
	HN			M216′	HN
					HA
D217′					HG2/3
					HB2/3
				P215′	HB2

5.3.3 Smad7 linker interacts with YAP/TAZ WW domains

Smad7 has been reported to interact with the transcription co-activator YAP (Ferrigno et al. 2002). Interaction mainly occurs between the Smad7 linker region and YAP WW domains. The transcription co-activator TAZ is a YAP paralogue that has also been implicated in modulating the TGF-ß /BMP signalling (Varelas et al. 2008). Based on the sequence similarity observed between the WW domains of both TAZ and YAP proteins, it is most likely that TAZ could also act as a potential Smad7 binding partner. Sequence alignment of YAP WW1, YAP WW2 and TAZ WW domains reveals a higher sequence identity between YAP WW1 and TAZ WW than between YAP WW1 and YAP WW2 (Figure 5.42). Human YAP exists as two isoforms that only differ in the number of WW domains they contain (one or two respectively). Human TAZ protein has so far been characterized as a unique isoform with one WW domain but in Oryzias Latipes (Medaka fish) a second TAZ isoform with two WW domains has been identified (Webb et al. 2011). In this part of the thesis the known human isoform of TAZ with a single WW domain was studied regarding its interaction with Smad7.

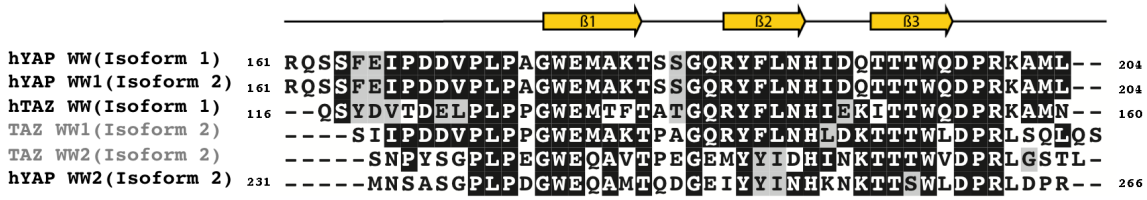


Figure 5.42: Sequence alignment of the human isoforms of YAP/TAZ WW domains in black and TAZ isoform (Oryzias Latipes) with two WW domains in grey. Alignments were performed using *ClustalW* (Thompson et al. 1994) and BoxShade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

To compare the affinities of the 15-residues Smad7[PY] peptide (ELESPPPPYSRYPMD, Figure 5.36) and YAP/TAZ WW domains ITC binding assays were performed. The ITC experiments reveal that the three WW domains bind Smad7 in the low μM range. The YAP WW1 domain binds Smad7[PY] with a K_D of 6.9 ± 0.3 μM more than three-fold stronger than YAP WW2 (K_D = 25.0 ± 3.4 μM). TAZ WW interacts with Smad7[PY] in agreement with the sequence similarity to YAP WW1 with a similar dissociation constant of 8.1 ± 0.4 μM (Table 5.15 and Figure 5.43).

Table 5.15: K_D values of Smad7[PY] bound to TAZ WW, YAP WW1 and YAP WW2 (Tris buffer pH 7.2 at 15 °C).

Peptide	TAZ WW	YAP WW1	YAP WW2
	K_D [μM]	K_D [μM]	K_D [μM]
Smad7	8.1 ± 0.4 **(A)**	6.9 ± 0.3 **(B)**	25.0 ± 3.4 **(C)**

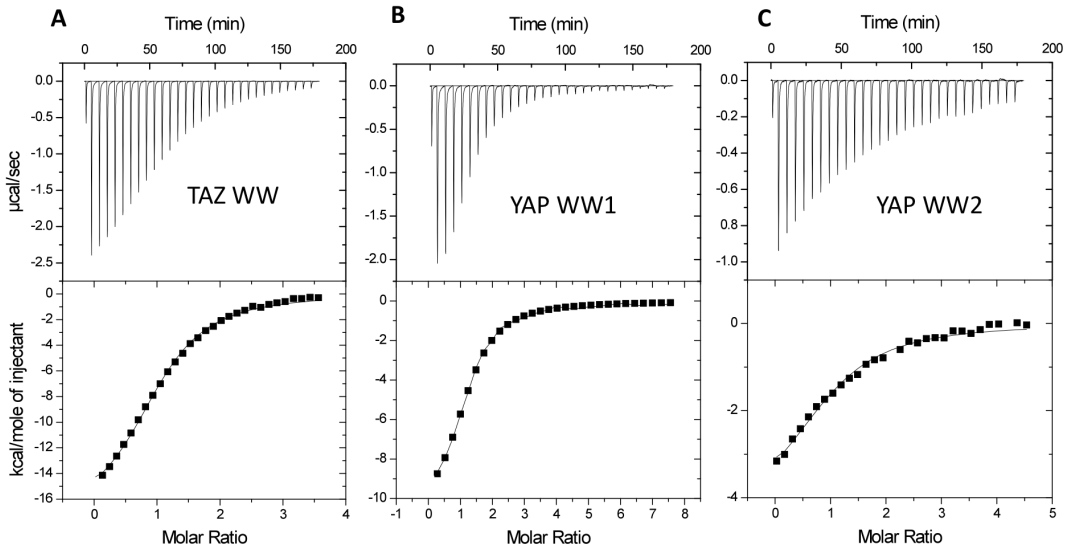


Figure 5.43: ITC profiles of Smad7[PY] and (A) TAZ WW, (B) YAP WW1 and (C) YAP WW2.

To corroborate the ITC results NMR binding experiments using HSQC spectroscopy were performed. The backbone amides of the free and bound TAZ WW1 domain were assigned using a standard ^{13}C, ^{15}N, ^{1}H triple resonance NMR spectroscopy. Residues located on the ß2- and ß3-strand of the TAZ WW domain but also ß1-strand residues are affected upon the addition of increasing amounts of the Smad7[PY] peptide. The largest chemical shift is obtained for the amide proton of T151.

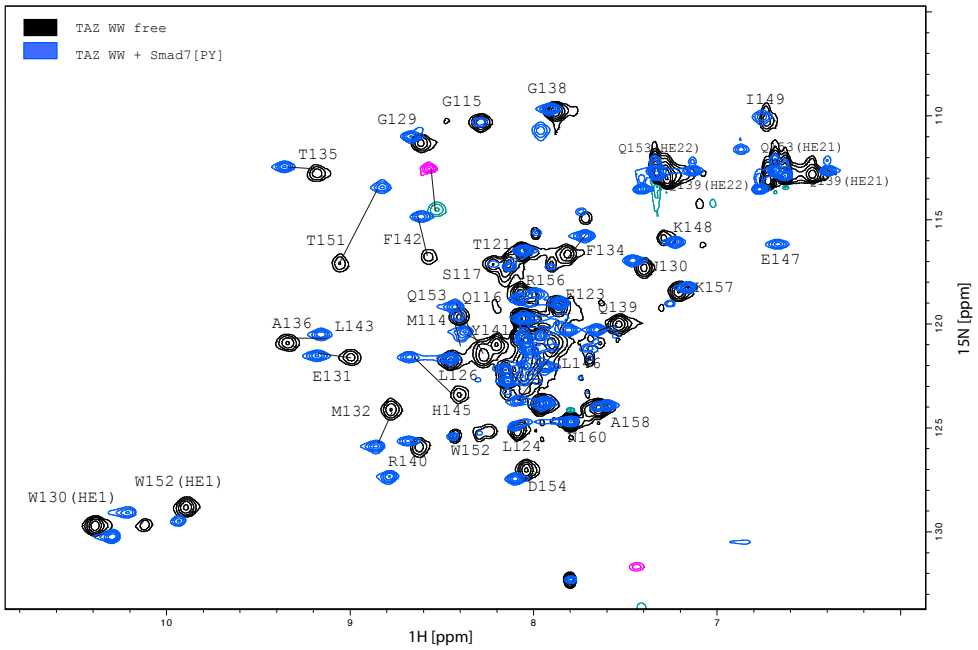


Figure 5.44: $^{1}H^{15}N$-HSQC spectrum showing the TAZ WW amide backbone protons in the free state (black) and the bound state after saturation with Smad7[PY] (blue).

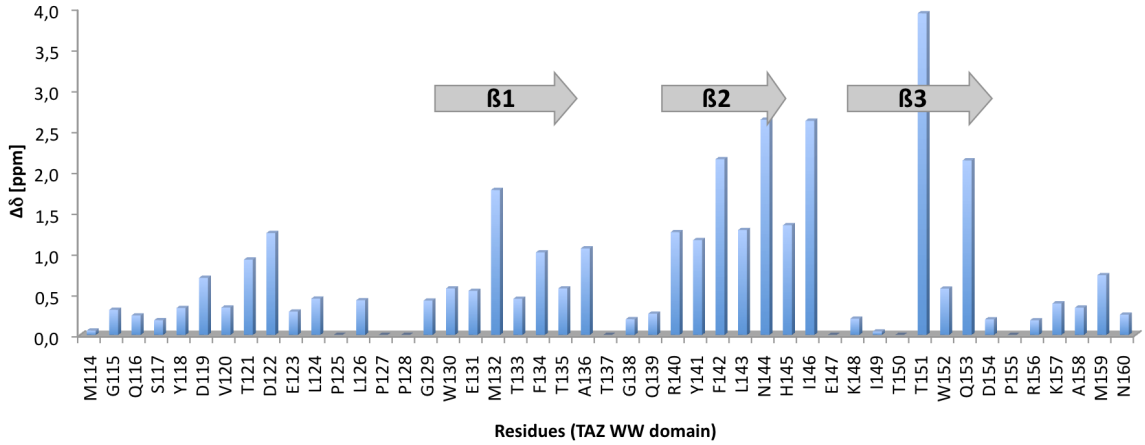


Figure 5.45: Histogram of the chemical shift changes observed for TAZ WW amide backbone protons after saturation with Smad7[PY] with respect to the free TAZ WW domain.

5.3.4 Structural determination of the TAZ WW – Smad7[PY] complex

In order to determine the 3D structure of the TAZ WW – Smad7[PY] complex the side chains were assigned and inter- and intramoleclar NOEs were identified in the 2D TOCSY/NOESY spectra. As displayed in Figure 5.46 the TAZ WW domain folds with the characteristic triple stranded anti-parallel ß-sheet and interacts with the Smad7[PY] motif using the typical XP and tyrosine binding cavity. Thus, Y211' is accommodated in the cavity formed by the residues L143, H145, K148, T149 and T150 while the XP site is formed by the aromatic residues Y141 and W152 that bind the pyrrolidine rings of P209' and P208'. In addition to these conserved interactions, contacts between the H145 imidazole ring and the R213' side chain as well as to the vicinal Y211' ring could be detected as well as contacts between the S212' amide proton and the methyl group of T150. C-terminal P115' is bound by side chains of ß1-stand residues E131, T133 and ß2 residue L143. These contacts as well as contacts involving Y211', R213' and Y214' (P210', Y211', S212', Y214' and M116') side chains were also described for the Nedd4L WW2 – Smad7[PY] complex. Thus, Smad7 also forms its characteristic loop structure when bound to the TAZ WW domain. Differences between both complex lie in the binding of the negatively charged N-terminus of Smad7. In case of Nedd4L WW2 positively charged amino acids located in the loop 1 region bind this part whereas in case of the TAZ WW domain strong NOEs were obtained for E205´and the Q139 side chain as well as the W152 indole ring of the second and third strand, respectively.

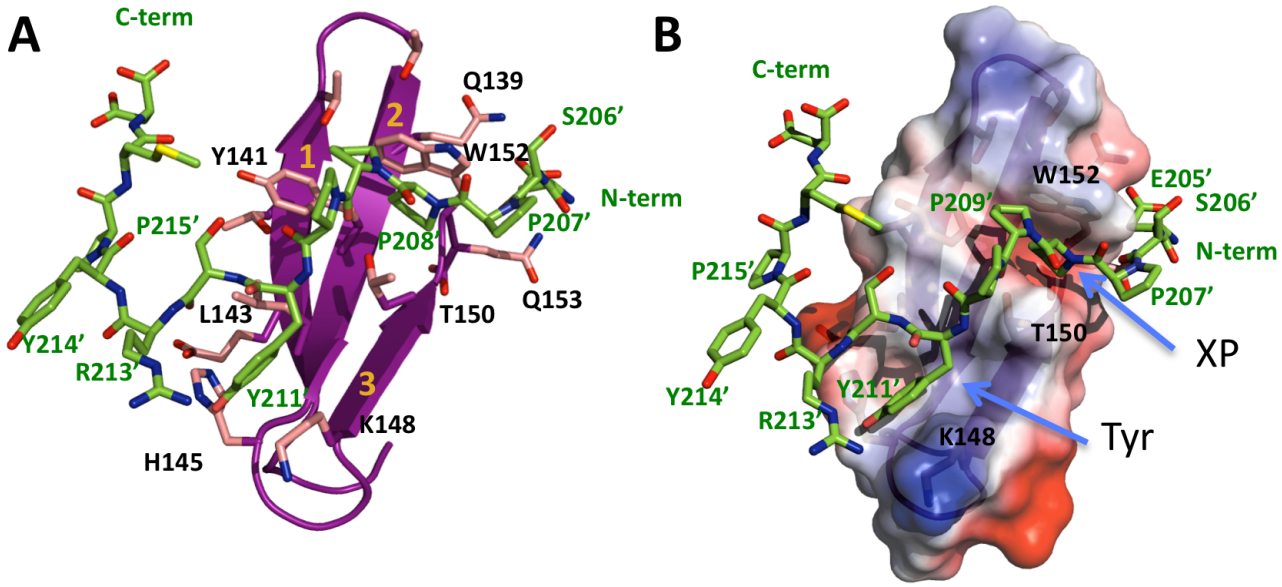


Figure 5.46: Cartoon representation (A) and charge distribution (B) of the lowest-energy structure of the TAZ WW - Smad7[PY] complex. The Smad7[PY] peptide is shown in green

Table 5.16: Inter- and intramolecular Smad7[PY] peptide NOEs indentified through NMR assignment of the TAZ WW – Smad7[PY] complex (for proton nomenclature see appendix 7.6).

Smad7[PY] — TAZ WW						
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
E205′	HA			S206′	HN
	HB3	W152	HZ3	S206′	HN
	HB2	W152	HH2	S206′	HN
			HZ3		
		Q139	HB2/3		
	HG2/3			S206′	HN
		Q139	HG2/3		
S206′	HB2/3			P207′	HD2/3
	HN			E205′	HA
					HB2/3
					HG2
P207′	HD2/3			S206′	HA
	HA			P208′	HD2/3
P208′	HA	W152	HZ2	P209′	HD2/3
			HH2		
			HZ3		
	HD2/3	W152	HE1	P207′	HA
			HD1		
			HZ2		
			HH2		
		T150	QG2		
	HG2/3	Y141	QE		
		T151	HN		
		W152	HN		
			HE3		
			HZ3		
			HH2		
			HZ2		
	HB2/3	T151	HN		
		W152	HD1		
			HZ2		
			HH2		
			HE3		
			HE1		
		T150	QG2		
P209′	HA	Y141	QD/E	P210′	HD2/3
		T150	QG2		
	HD2/3	Y141	QD/E	P208′	HB3
		W152	HZ2		HG3
			HH2		
			HZ3	R213′	HG2
	HB2/3	T150	QG2	R213′	HG2
		Y141	QD/E		
	HG2/3	T150	QG2	P208′	HG3
		Y141	QD/E		
P210′	HA	T150	QG2	Y211′	HN
					HB2/3
				S212′	HN
				R213′	HN
	HD2/3			Y211′	HN
				P209′	HA
	HB2/3			Y211′	HN
				S212′	HN
				Y213′	HN
	HG2/3			Y211′	HN
Y211′	HN	T150	QG2	P210′	HD2/3
					HG2/3
					HB2/3

Smad7[PY] — TAZ WW					
205 210 215 ELESPPPPYSRYPMD		Inter-molecular NOEs		Intra-molecular NOEs	
AA	Proton				
S212′	HA	T150	QG2	S212′	HN
	HB2/3	T150	QG2	P210′	HA
					HB2/3
	QD/E	I149	HA		
		H145	HA		
			HB2/3	S212′	HN
		T150	QG2		
		K148	HE2	S212′	HA
			HA		
			HB2/3		
			HG2/3		
		L143	QD1/2		
			HG		
		T150	QG2		
	HN	T150	QG2	Y211′	HA
					HB2/3
				P210′	HB3
					HA
	HA			R213′	HN
	HB2/3			R213′	HN
R213′	HN			Y214′	QD/E
					HB2/3
				Y211′	QD/E
				S212′	HA
					HB2/3
	HA	H145	HN	Y214′	HN
			HD1	Y214′	QD/E
	HB2/3			Y214′	HN
	HG2/3			Y214′	QD/E
				P209′	HB2/3
					HD2/3
Y214′	HN			R213′	HA
					HB2/3
	HB2/3			R213′	HN
	QD/E			R213′	HA
				R213′	HG/3
				P215′	HD2/3
P215′	HA			M216′	HN
	HD2/3			Y214′	HA
					HN
					QE/D
				M216′	HA
	HB2			M216′	HN
	HG2/3	T133	QG2		
M216′	HA			D217′	HN
	HN			P215′	HA
					HD2/3
					HB2
	HB2/3			D217′	HN
	HG2/3			D217′	HN
D217′	HN			M216′	HA
					HG2/3
					HB2/3

Discussion

6 Discussion

Signalling pathways are complex signal transduction systems with multiple components involved that interact with each other to regulate the transfer of a signal. The interaction of those compounds can be controlled through specific cellular locations or through posttranslational modifications such as phosphorylation. TGF-ß signal mediators Smad2/3 are located in the cytoplasm and nucleus, phosphorylated Smad2/3 proteins are mainly found in the nucleus as well as the inhibitor Smad7. Smad action is controlled by the E3 ubiquitin ligase Nedd4L that can recognize both, the phosphorylated Smad3 linker and the Smad7 linker through its WW domain region. Adaptor proteins Pin1 and TAZ/YAP might regulate Nedd4L action by enhancing binding either to Smad3 in case of Pin1 or to Smad7 in case of YAP/TAZ.

6.1 Interaction of Smad3 and the WW domains of Nedd4L and Pin1

The first part of this work was devoted to study binding of the Smad3 linker region containing a PY motif and four phosphorylation sites to the WW domains of the E3 ubiquitin ligase Nedd4L and peptidyl prolyl *cis/trans* isomerase Pin1.

6.1.1 Selectivity in the Nedd4L - Smad3 interaction

All Nedd4L WW domains individually recognize the PY motif of the nonphosphorylated Smad3 linker. However, when the Smad3 linker is phosphorylated directly upstream the PY motif the binding affinity between Nedd4L WW2 and Smad3 increases remarkable. The three remaining Nedd4L WW domains share 66% (WW1), 42% (WW3) and 49% (WW4) sequence identities with the WW2 domain but the interaction with the Smad3 pT[PY] motif is reduced for WW3 and extremely weak for WW1 and WW4. The binding affinities can be summarized as WW2 > WW3 > WW1 > WW4. The E3 ligase Nedd4 is a close relative of Nedd4L and interacts with a PY motif of the epithelial sodium channel subunit through its WW domains. It has been reported that the interaction between Nedd4 and the ENaC subunit is enhanced by more than three-fold, when the threonine residue directly downstream of the PY motif is phosphorylated (Shi et al. 2002b). Nedd4L also interacts with a PY motif of the ENaC subunits and induces their ubiquitination. Interestingly in this context, Nedd4L WW3 and WW4 interact with the PY motif, whereas WW1 and WW2 do not interact or play a minor role

(Asher et al. 2003; Fotia et al. 2003; Henry et al. 2003). It has also been reported that phosphorylation of the threonine upstream of the PY motif enhances binding to a Nedd4L WW3-WW4 construct, but binding studies of the individual domains do not exist.

Based on the additional role of Nedd4L in ENaC regulation it is suggested that Nedd4L interacts with different targets through different WW domains and distinct requirements on target phosphorylation. In this work it was shown that in TGF-ß signalling Nedd4L selectively recognizes the Smad3 pT[PY] motif though its WW2 domain. This result shows that Nedd4L can distinguish between activated linker phosphorylated Smad3 and ground state nonphosphorylated Smad3.

The structure of the Nedd4L WW2 - Smad3 pT[PY] complex, which is the first 3D structure of a group I WW domain bound to phosphorylated p(S/T)[PY] motif was determined by multidimensional NMR experiments in solution. Hereby it was demonstrated that the peptide binding site is formed in the typical manner of group I WW domains by a XP groove and a tyrosine binding groove. Specificity for binding the phosphate group was achieved by two positively charged residues R380 and K378 in the loop 1 region. In all Nedd4L WW domains loop 1 residue R380 is conserved, but K378 is replaced by a slightly polar amino acid N510 (WW3), aliphatic residue L206 (WW1) and an acidic amino acid D561 (WW4). Binding affinities decrease in exactly this order K378 (WW2) > N510 (WW3) > L206 (WW1) > D561 (WW4).

In response to TGF-ß activation of Smad3 through C-tail phosphorylation occurs after around 10 min. Another 10 min later the threonine residue of the T[PY] motif is phosphorylated by CDK8/9, followed by phosphorylation of the other phosphorylation sites in the following order T179' > S208' > S213' > S204', where S204' is phosphorylated by GSK3ß (Gao et al. 2009). In contrast, stimulation through UV radiation, osmotic stress or epidermal growth factor (EGF) that induces strong MAPK-mediated phosphorylation shows different preferences for phosphorylation. All other sites were phosphorylated before T179' that occurs at a less prominent extent (S208' > S213' > S204' > T179') (Kretzschmar et al. 1999; Gao et al. 2009). Hence, Nedd4L is able to distinguish between Smad3 phosphorylated by CDK8/9 and by MAPKs through selective binding of Smad3 pT[PY] by Nedd4L WW2. Smad3 linker phosphorylation by MAPKs does not require previous Smad3 activation through C-tail

phosphorylation. Phosphorylation by MAPKs occurs mainly in the cytoplasm and may be recognized by different E3-ubiquitin ligases.

6.1.2 Role of Pin1 in Smad3 recognition

Pin1 is the unique peptidyl prolyl isomerase that can isomerase the phosphorylation sites (p(S/T)P motifs) as they are present in the phosphorylated Smad3 linker. A recent study reports that Pin1 (like Nedd4L) only binds TGF-ß induced, C-tail phosphorylated Smad3 within the linker region. Mutation of the Smad3 C-tail region inhibits Pin1 binding and shows that Pin1 interacts selectively with activated Smad3 (Matsuura et al. 2010). The interaction of Nedd4L and Smad3 requires *trans* configuration of the Smad3 pT-PY motif as it has been shown by NMR structure determination in this work. Pin1 may inhibit binding between Smad3 and Nedd4L through *trans* to *cis* conversion that puts a kink in the Smad3 linker. Contrarily, Pin1 may enhance binding between Nedd4L and Smad3 through *cis* to *trans* isomeration of the phosphorylation sites, as it has been suggested also for the E3 ubitquitin ligase Smurf2 (Nakano et al. 2009). Interestingly, the WW domain of Pin1 predominantly binds the same Smad3 phosphorylation site that is recognized (with higher affinity) by Nedd4L WW2. Dissociation constants obtained for Pin1 WW and Smad3 peptides containing one to two p(S/T)P motifs lie between 12-85 μM, as reported for other Pin1 WW domain ligands (Verdecia et al. 2000). According to the data presented in this work the Pin1 WW domain binds Smad3 p(S/T)P motifs in the following order: pT179P > pS213P > pS208P > pS204P. Pin1 only requires one p(S/T)P motif, double phosphorylation does not enhance binding. Structure determination of the Pin1 WW – Smad3 pT179P complex revealed that Pin1 itself binds the ligand in *trans* configuration. *Cis* to *trans* conversion of the other Smad3 p(S/T)P motifs may then be carried out by the Pin1 catalytic domain and subsequently Nedd4L can interact with the Smad3 linker. Pin1 may further control phosphorylation and dephosphorylation of the Smad3 linker since most of the kinases require *trans* configuration of the ligand.

Structure determination of the Pin1 WW - Smad3 pS208-pS213 complex reveal that only the pS213P motif interacts with the Pin1 WW domain. The reason for the preference in pS213P binding is a Pro residue located upstream of the PY motif at position x-2. This proline is present in Smad3 pT179P and pS213P motifs. In Smad3 pS204P and pS208P motifs the Pro is absent explaining the weaker binding affinities that were obtained. In the same way it has been

reported that the Pin1 WW domain binds peptides YpSPRpSPS and YSPRpSPS of the RNA polymerase II subunit's C-terminal domain equally well regardless of the peptide being doubly or singly phosphorylated (Verdecia et al. 2000). Residues R14 and R17 were identified as amino acids mainly involved in phosphate binding. In both complexes (Pin1 WW – Smad3 pT179P and Pin1 WW – Smad3 pS208P-pS213P) R17 coordinates the phosphate group with its backbone amide and the guanidinium side chain. R14 either interacts with the phosphate group (Pin1 WW – Smad3 pS208P-pS213P) or with a negatively charged amino acid (E178') upstream of the p(S/T)P motif (Pin1 WW – Smad3 pT179P).

6.1.3 ^{31}P NMR experiments of phosphorylated Smad3 peptides

The impact of the pT179' phosphate group on binding Nedd4L WW2 or Pin1 WW could be demonstrated in three different experiments. In the first experiment, the 1D ^{31}P spectrum of the free peptide and of the Nedd4L WW2 – Smad3 pT179[PY] complex, was acquired at different pH values. The ^{31}P signal corresponding to the free peptide was more affected by pH changes than that of the bound one. In a second experiment the bound peptide (Pin1 WW – Smad3 pT179P) was compared to a sodium phosphate reference that was added to the buffer and the spectra were acquired at different pH values. Similar to the first experiment, the ^{31}P signal corresponding to the buffer was affected by the pH changes, whereas the ^{31}P signal of the Smad3 pT179P peptide is barely shifted. In a third experiment, the transition from the free peptide to the bound peptide was demonstrated by titrating the Pin1 WW domain to Smad3 pT179P. In this experiment, the ^{31}P sodium phosphate reference peak did not shift, but the peaks corresponding to the free and the bound state of the peptide were separated by 1.0 ppm. The transition peaks between the first and the last point of the titration were broader due to the presence of both populations. With these experiments it was shown that ^{31}P NMR is a useful tool to study phosphopeptides and the impact of the phosphate group on binding.

6.1.4 Smad3 - WW domain binding is affected by pH changes

Nedd4L WW2 and Pin1 WW belong to different classes of WW domains and bind their ligands in a distinct ways. However, a pH change from pH 6.0 to 7.5 results in both cases in stronger binding. The same observation has been made by intrinsic tryptophan fluorescence spectroscopy of Nedd4L WW3 and the PY motif of the ENaC subunit. Between pH 7.0 and 6.5 no changes in binding affinity could be observed but a slight decrease was measured at pH 6.0 (Kanelis et al. 2001). A possible explanation for this observation lies in the interaction

between the His residue within the hydrophobic binding groove that builds hydrogen bonds to the hydroxide group of the tyrosine of the PY motif. To facilitate this interaction, the His residue needs to be uncharged, a feature that occurs at a high pH. This favourable contribution only occurs in Nedd4L WW2 – PY complexes because Pin1 WW cannot bind PY motifs. The explanation for the interaction with phosphorylated residues may have its grounds in the electrostatic balance between the bivalent negatively charged phosphate group and the positively charged residues close to it. The formation of the bivalent phosphate anion is optimized at higher pH and is probably the basis of the affinity increase in both, Nedd4L WW2 and Pin1 WW, complexes. It is therefore important to control the pH value during all binding experiments, in order to obtain reproducible and interpretable results. A series of ITC binding experiments can be useful before starting NMR structure determination since they can contribute to the identification of the best conditions for the complex formation.

6.1.5 Cooperativity of WW domains in Nedd4L - Smad3 interaction

In E3 ubiquitin ligases the region containing the WW domains is the most variable and accounts for the selectivity that the different ligases display towards their target proteins. Still, the details for target recognition are an open question, and the information that is collected reflects a more complex scenario than perhaps initially expected. Most of the target motifs contain a PY binding site and the selectivity in the target recognition arises from the ability to use either a single domain or a domain pair to recognize extended PY sequences that contain additional binding sites. Some E3 ligases seem to recognize short sequences in their target proteins, as is the case of Itch and they do so using a unique WW domain (Morales et al. 2007). In drosophila it has been shown that the WW3-WW4 tandem of the Nedd4-like E3 ubiquitin ligase suppressor of deltex (Su(dx)) operates in a way that ligand binding to one WW domain enhances binding to the nearby WW domain (Jennings et al. 2007). The Nedd4 like E3 ubiquitin ligase in saccharomyces cerevisiae yeast Rsp5 (reverses SPT-phenotype protein 5) is also suggested to mediate ligand binding in a cooperative way through the tandem domains WW2-WW3 (Bhattacharya et al. 2008). A recent study reported cooperative binding of Nedd4L WW2-WW3 to the PY motif of the serum- and glucocorticoid-regulated kinase Sgk1 (Wiemuth et al. 2010).

The kinase Skg1 not only binds Nedd4L it also phosphorylates the E3 ligase at two positions: Residue S428 is phosphorylated within the linker connecting WW2 and WW3 and residue

S322 between WW1 and WW2 is phosphorylated to a lesser extent. Skg1 mediated phosphorylation of Nedd4L reduces the affinity towards the ENaC subunit by creating a binding site for 14-3-3 proteins (Yaffe et al. 1997; Debonneville et al. 2001; Bhalla et al. 2005; Snyder 2005; Nagaki et al. 2006; Liang et al. 2008). In TGF-ß signalling Skg1 might inhibit in a similar way the binding of Nedd4L and Smad3 by phosphorylation of residue S428 between the WW2 and the WW3 domains which strengthens the result obtained in this thesis that Nedd4L uses a the WW2-WW3 domain pair to expand the binding interface to recognize the extended Smad3 linker, including the PY motif and three phosphorylation sites. Sites T179' and S208' are phosphorylated by kinases CDK8/9 and it was recently shown that these kinases also create binding sites that are preferentially recognized by WW domain-containing Smad transcription factors (Aragon et al. 2011). However, additional Smad3 phosphorylation at S204' mediated by GSK3ß is required for Nedd4L binding as demonstrated in this thesis. In case of the E3 ubiquitin ligase Smurf2, Smad3 binding is not altered upon GSK3ß-mediated Smad3 linker phosphorylation (Millet et al. 2009). It is therefore suggested that GSK3ß triggers Smad3 destruction, which is carried out selectively by the Nedd4L WW2-WW3 pair after transcription activity. Each domain recognizes a binding site in the target following a two-step process, which starts with a high affinity interaction (WW2 and pT179[PY]) further improved with the recognition of the second site (WW3 and pS204-pS208) that accounts for both an enhanced affinity and specificity.

Chemical ligation was successfully carried out to obtain a Nedd4L WW2-WW3 sample with distinct isotopic labelled domains for structural studies. Through this approach it could be shown that the WW3 loop 1 amino acid R486 interacts with the pS204' phosphate group, whereas residues R492 and N490 coordinate pS208'. Alanine mutations of R486 and R492 reduced binding to the tri-phosphorylated Smad3 peptide to the same amount.

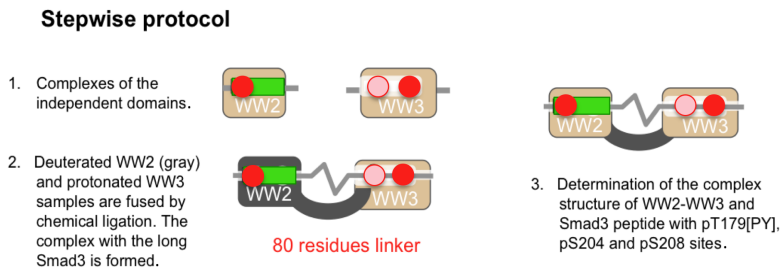


Figure 6.1: Stepwise protocol that was followed to obtain the complex structure of Nedd4L WW2-WW3 and Smad3 pT179[PY]-pS204-pS208.

In summary, the binding studies of the Smad3 linker and proteins Pin1 and Nedd4L reveal that Smad3 recognition by Nedd4L is carried out by the WW2-WW3 domain pair and depends on Smad3 linker phosphorylation. This interaction might be enhanced through *cis* to *trans* isomeration of the Smad3 p(S/T)P sites by Pin1 where its WW domain acts as an anchor for Smad3.

The following mechanism is suggested: After TGF-ß induced Smad3 activation and localisation to the nucleus, the Smad3 linker is phosphorylated by CDK8/9 at position T179'. Subsequently, phosphorylation of S208' by CDK8/9 induces S204' phosphorylation by GSK3ß. The Pin1 WW domain binds the Smad3 pT179P motif and facilitates *cis* to *trans* isomeration of phosphorylation sites S204' and S208' by the catalytic domain. Nedd4L recognizes the phosphorylated *all-trans* Smad3 linker with the WW2-WW3 pair and promotes Smad3 degradation (Figure 6.2).

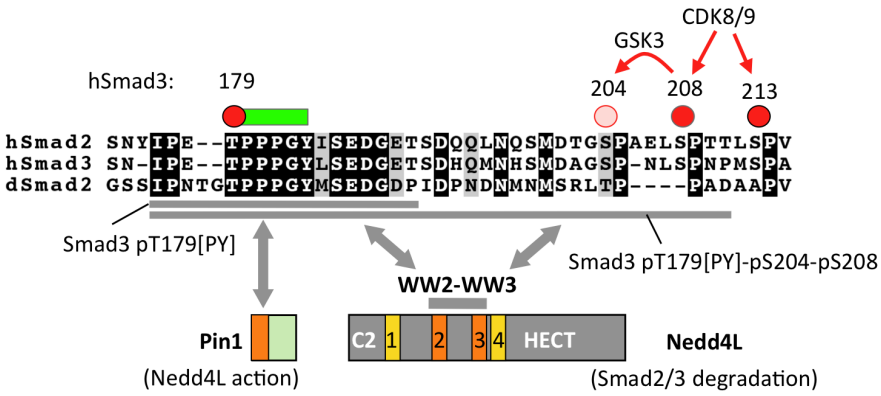


Figure 6.2: Scheme of Smad3 recognition by Nedd4L and Pin1 after TGF-ß induced Smad3 linker phosphorylation.

6.2 Interaction of Smad7 and the WW domains of Nedd4L and TAZ

The second part of this thesis was carried out to understand the binding mode of the interactions between the inhibitor Smad7 and the WW domains of E3 ubiquitin ligase Nedd4L as well as the WW domain of transcriptional co-activator TAZ to obtain more information about the role of both proteins in TGF-ß inhibition.

6.2.1 Binding of Smad7 and Nedd4L

The linker region of I-Smad6 (BMP) and I-Smad7 (TGF-ß) contains a PY motif, which is present in all Smad proteins except Co-Smad4. Several E3 ubiquitin ligases have been reported to interact with Smad7 and negatively regulate TGF-ß pathway like WWP1 and Smurf2 and

Nedd4L (Kavsak et al. 2000; Komuro et al. 2004). From the structural and biochemical data presented in this work it can be concluded that Nedd4L binds to the Smad7[PY] fragment through contacts exclusively involving the second WW domain. The high affinity of the Nedd4L WW2 – Smad7[PY] complex is explained as the sum of contacts including the canonical PY motif, the extended C-terminus that aligns on the first strand of the WW2 domain and an electrostatic balance between the glutamic acid at position -2 with respect to the PY motif in Smad7 and two positively charged residues located in loop 1 of the WW domain. A different situation has been described for the Smurf2 and Smad7 interaction, where two WW domains form a composed structure that recognizes a short fragment of Smad7 (Chong et al. 2010). The Smurf2 WW2-WW3 tandem interacts with Smad7 in a way that WW3 binds the Smad7[PY] motif and the WW2 participates in additional contacts to Smad7. This coupled WW domain configuration enhances binding by over 20-fold with respect to the interaction of the single Smurf2 WW3 domain and Smad7. The isolated Smurf2 WW2 is instable and does not interact with Smad7 (Chong et al. 2006). It could be shown in this work that in the case of Nedd4L binding affinities towards the Smad7[PY] motif are not enhanced by a WW pair formation. Therefore, it seems that Smurf2 and Nedd4L have found two efficient but distinct mechanisms to recognize the Smad7[PY] ligand since both ligases interact with the Smad7 fragment with affinities in the low micromolar range (Smurf2 pair with 1.7 μM and Nedd4L WW2 domain with 4 μM). Hence, depending on the context WW domains may play their role either solo or in an orchestrated manner.

6.2.2 Binding of Smad7 and TAZ

The multifunctional adaptor protein TAZ and YAP have been reported to enhances transcriptional activity in responds to TGF-ß signalling by binding R-Smads and Co-Smad4 (Varelas et al. 2008). YAP however, also can play an inhibitory role by binding Smad7 and enhancing Smad7 and Nedd4L mediated TßRI degradation (Ferrigno et al. 2002).

ITC experiments carried out in this work showed that the TAZ WW domain binds Smad7 with almost the same binding affinity as the YAP WW1 domain. Structure determination by solution NMR revealed that Smad7 is bound by the TAZ WW domain by residues located on ß1, ß2 and ß3 strands as in the case of the Nedd4L WW2 – Smad7[PY] complex. The loop formation of Smad7 that resulting in a larger binding interface was observed in both complexes and seems to be characteristic for this peptide. Both WW domains bind the

negatively charged amino acid E205' at the Smad7 N-terminus but in Nedd4L WW2 this interaction involves loop 1 residues whereas in TAZ WW ß-stand 2 and 3 residues are engaged.

Nedd4L and TAZ both bind Smad7 with a single WW domain in the low µM range but most likely they will not compete for Smad7 binding since Nedd4L is mainly located in the cytoplasm and TAZ in the nucleus. In the nucleus Smad7 can compete with R-Smads for TAZ (and YAP) resulting in inhibition of the gene transcription. When shuttled to the cytoplasm Smad7 can interact with the TßRI receptor and contribute to its inhibition and degradation by recruiting E3 ubiquitin ligases like Nedd4L.

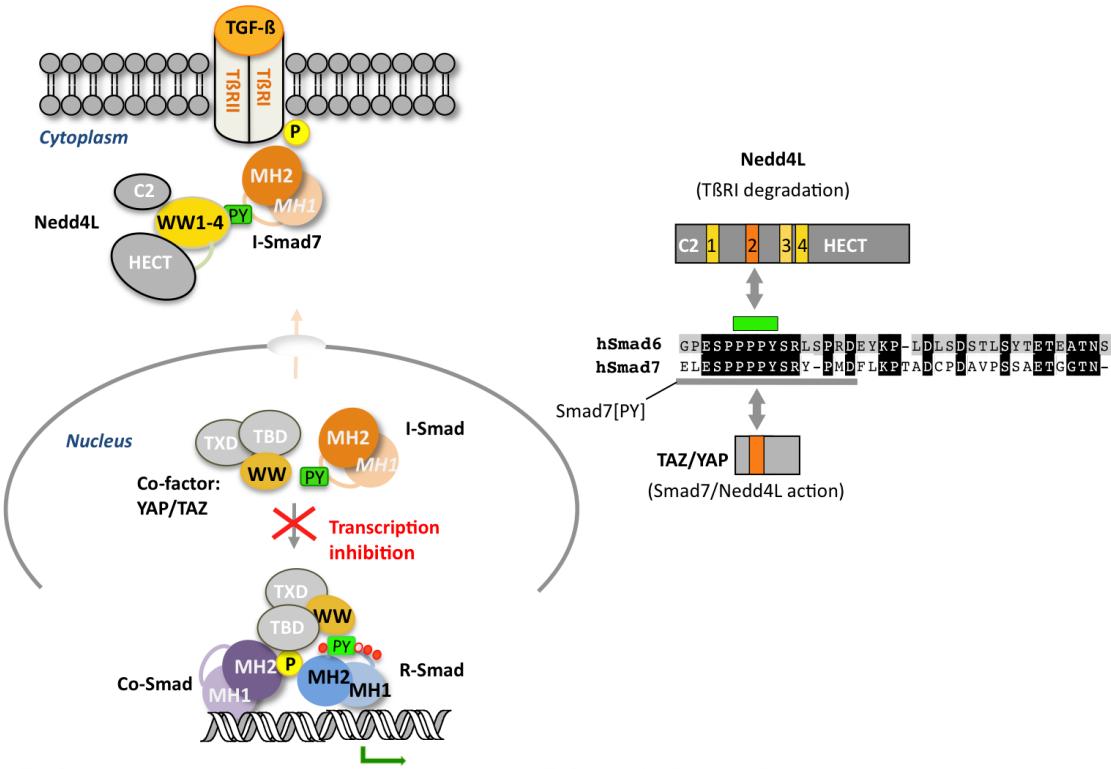


Figure 6.3: Scheme of TGF-ß inhibition mediated by TAZ, Smad7 and Nedd4L In the nucleus Smad7 can bind TAZ to inhibit the formation of the transcriptional complex with Co-Smad4 and R-Smads. TAZ WW shuttles Smad7 from the nucleus to the cytoplasm to facilitate TRßI degradation by Smad7 and Nedd4L.

In this work three complexes, Nedd4L WW2 - Smad3 pT[PY], Nedd4L WW2-Smad7[PY] and TAZ WW - Smad7[PY], were described that are similar in binding the PY region but differ in binding the C- and N-terminal part of the ligand resulting in different functional outcome.

Pin1 has been described as p(S/T)P binding motif and hereby it was shown on the bases of complexes Pin1 WW -Smad3 pTP and Pin1 WW - Smad3 pSP-pSP that only one phosphate group is bound by the domain by residues located at the ß-strand 1. CDK 8/9 and GSK3-mediated phosphorylation of the Smad3 linker prime the protein for degradation by Nedd4L after transcription activity. A cooperative binding mode for the Nedd4L WW2-WW3 pair could be determined where the interaction with the Smad3 pS-pS site and Nedd4L WW3 was unexpected from predictions based on analysis of the amino acid sequence.

Appendix

7 Appendix

7.1 Amino acids

Table 7.1: Amino acids used for Fmoc SPPS.

Amino acid	3-/1-letter code		Formula	MW (g/mol)
Alanine	Ala	A	Fmoc-Ala-OH	311.3
Argenine	Arg	R	Fmoc-Arg(Pbf)-OH	648.8
Asparagine	Asn	N	Fmoc-Asn(Trt)-OH	596.7
Aspartic acid	Asp	D	Fmoc-Asp(OtBut)-OH	411.5
Cysteine	Cys	C	Fmoc-Cys(Trt)-OH	585.7
Glutamic acid	Glu	E	Fmoc-Glu(OtBu)-OH	425.5
Glutamine	Gln	Q	Fmoc-Gln(Trt)-OH	610.7
Glycine	Gly	G	Fmoc-Gly-OH	297.3
Histidine	His	H	Fmoc-His(Boc)-OH	477.5
Isoleucine	Ile	I	Fmoc-Ile-OH	353.4
Leucine	Leu	L	Fmoc-Leu-OH	353.4
Lysine	Lys	K	Fmoc-Lys(Boc)-OH	468.5
Methionine	Met	M	Fmoc-Met-OH	371.5
Phenylalanine	Phe	F	Fmoc-Phe-OH	387.4
Proline	Pro	P	Fmoc-Pro-OH	337.4
Serine	Ser	S	Fmoc-Ser(tBu)-OH	383.4
Threonine	Thr	T	Fmoc-Thr(tBu)-OH	341.4
Tryptophan	Trp	W	Fmoc-Trp(Boc)-OH	526.6
Tyrosine	Tyr	Y	Fmoc-Tyr(tBu)-OH	459.6
Valine	Val	V	Fmoc-Val-OH	339.4
Phosphorylated amino acids				
Phosphoserine	pSer,	pS	Fmoc-Ser(PO(OBzl)OH)-OH	497.4
Phosphothreonine	pThr,	pT	Fmoc-Thr(PO(OBzl)OH)-OH	511.5

7.2 Nedd4L

7.2.1 Full-length hNedd4L protein sequence

Table 7.2: Sequence of full-length Nedd4L with WW1 (blue), WW2 (green), WW3 (red), WW4 (black). Skg1 site in yellow.

	1	11	21	31	41	51	
1	ILRVKVVSGI	DLAKKDIFGA	SDPYVKLSLY	VADENRELAL	VQTKTIKKTL	NPKWNEEFYF	60
61	RVNPSNHRLL	FEVFDENRLT	RDDFLGQVDV	PLSHLPTEDP	TMERPYTFKD	FLLRPRSHKS	120
121	RVKGFLRLKM	AYMPKNGGQD	EENSDQRDDM	EHGWEVVDSN	DSASQHQEEL	PPPPLPPGWE	180
181	EKVDNLGRTY	YVNHNNRTTQ	WHRPSLMDVS	SESDNNIRQI	NQEAAHRRFR	SRRHISEDLE	240
241	PEPSEGGDVP	EPWETISEEV	NIAGDSLGLA	LPPPPASPGS	RTSPQELSEE	LSRRLQITPD	300
301	SNGEQFSSLI	QREPSSRLRS	CSVTDAVAEQ	GHLPPPSAPA	GRARSSTVTG	GEEPTPSVAY	360
361	VHTTPGLPSG	WEERKDAKGR	TYYVNHNNRT	TTWTRPIMQL	AEDGASGSAT	NSNNHLIEPQ	420
421	IRRPRSLSSP	TVTLSAPLEG	AKDSPVRRAV	KDTLSNPQSP	QPSPYNSPKP	QHKVTQSFLP	480
481	PGWEMRIAPN	GRPFFIDHNT	KTTTWEDPRL	KFPVHMRSKT	SLNPNDLGPL	PPGWEERIHL	540
541	DGRTFYIDHN	SKITQWEDPR	LQNPAITGPA	VPYSREFKQK	YDYFRKKLKK	PADIPNRFEM	600
601	KLHRNNIFEE	SYRRIMSVKR	PDVLKARLWI	EFESEKGLDY	GGVAREWFFL	LSKEMFNPYY	660
661	GLFEYSATDN	YTLQINPNSG	LCNEDHLSYF	TFIGRVAGLA	VFHGKLLDGF	FIRPFYKMML	720
721	GKQITLNDME	SVDSEYYNSL	KWILENDPTE	LDLMFCIDEE	NFGQTYQVDL	KPNGSEIMVT	780
781	NENKREYIDL	VIQWRFVNRV	QKQMNAFLEG	FTELLPIDLI	KIFDENELEL	LMCGLGDVDV	840
841	NDWRQHSIYK	NGYCPNHPVI	QWFWKAVLLM	DAEKRIRLLQ	FVTGTSRVPM	NGFAELYGSN	900
901	GPQLFTIEQW	GSPEKLPRAH	TCFNRLDLPP	YETFEDLREK	LLMAVENAQG	FEGVD	

7.2.2 Pattern of secondary structure NOEs (Nedd4L WW domains)

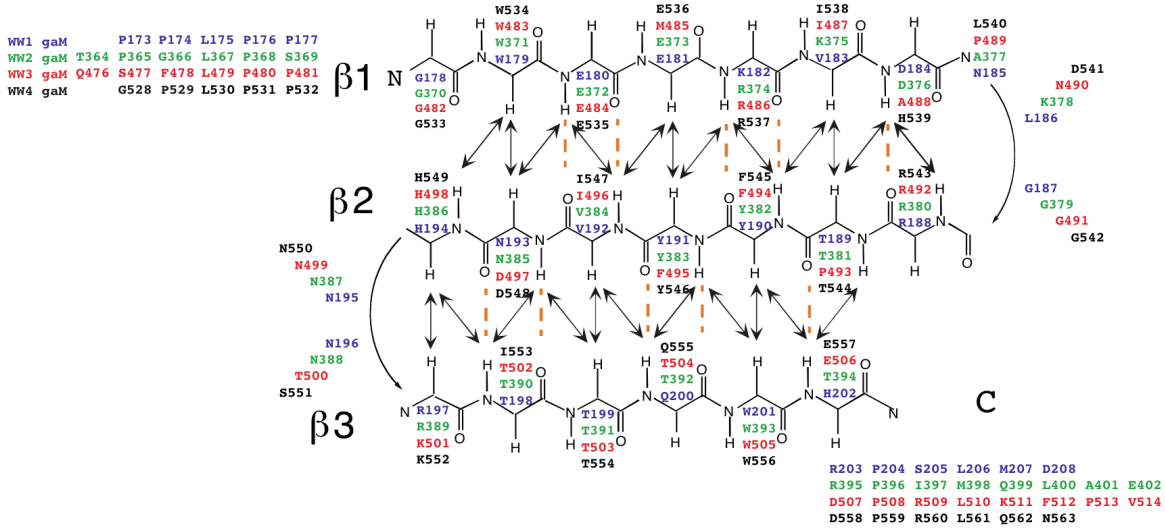


Figure 7.1: Pattern of secondary structure NOEs (represented as black arrows) for Nedd4L WW domains WW1 (blue), WW2 (green), WW3 (red) and WW4 (black). Hydrogen bonds are illustrated as orange broken lines.

7.2.3 NMR titrations: Nedd4L WW domains + Smad2/3 pT[PY]

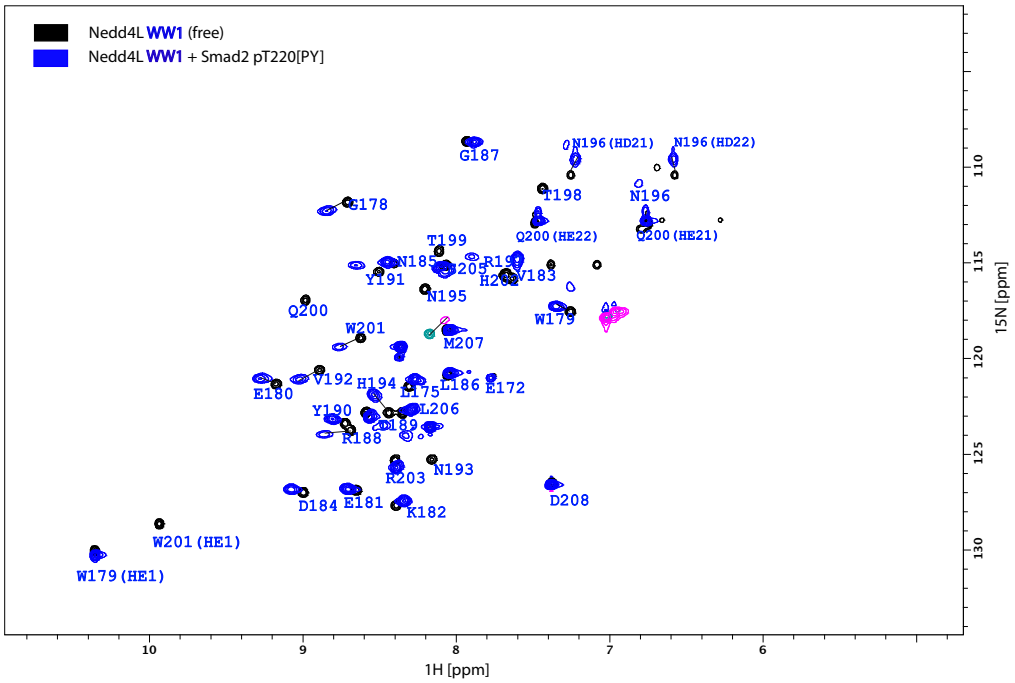


Figure 7.2: ^{1}H ^{15}N-HSQC spectrum of the Nedd4L WW1 domain in the free (black) and bound state (blue) after saturation with Smad2 pT220[PY] peptide.

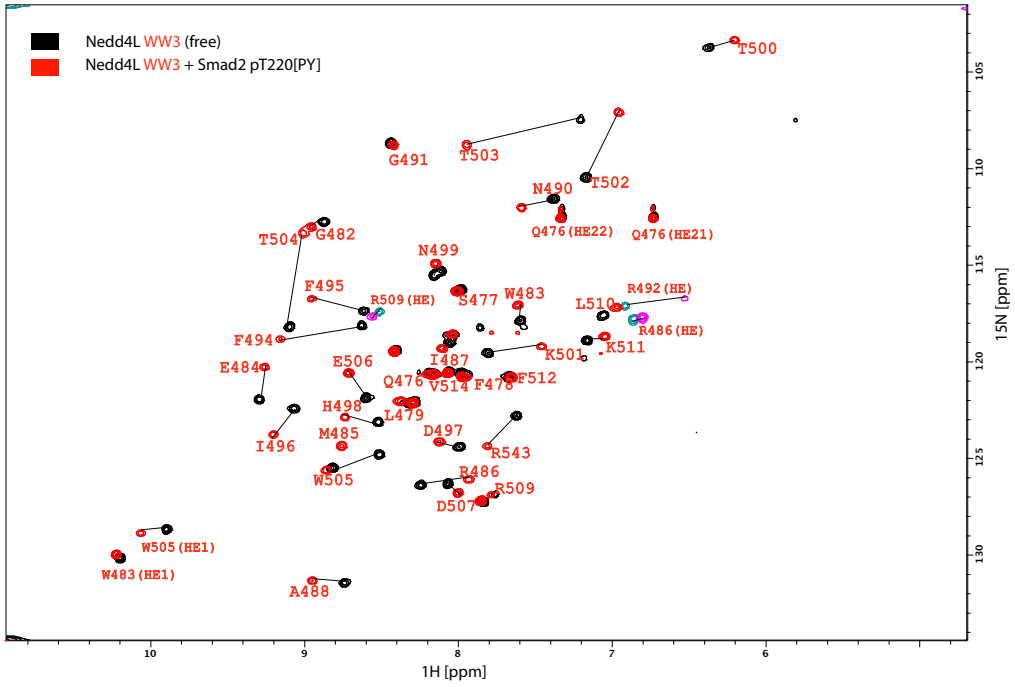


Figure 7.3: ^{1}H ^{15}N-HSQC spectrum of the Nedd4L WW3 domain in the free (black) and bound state (red) after saturation with Smad2 pT220[PY] peptide.

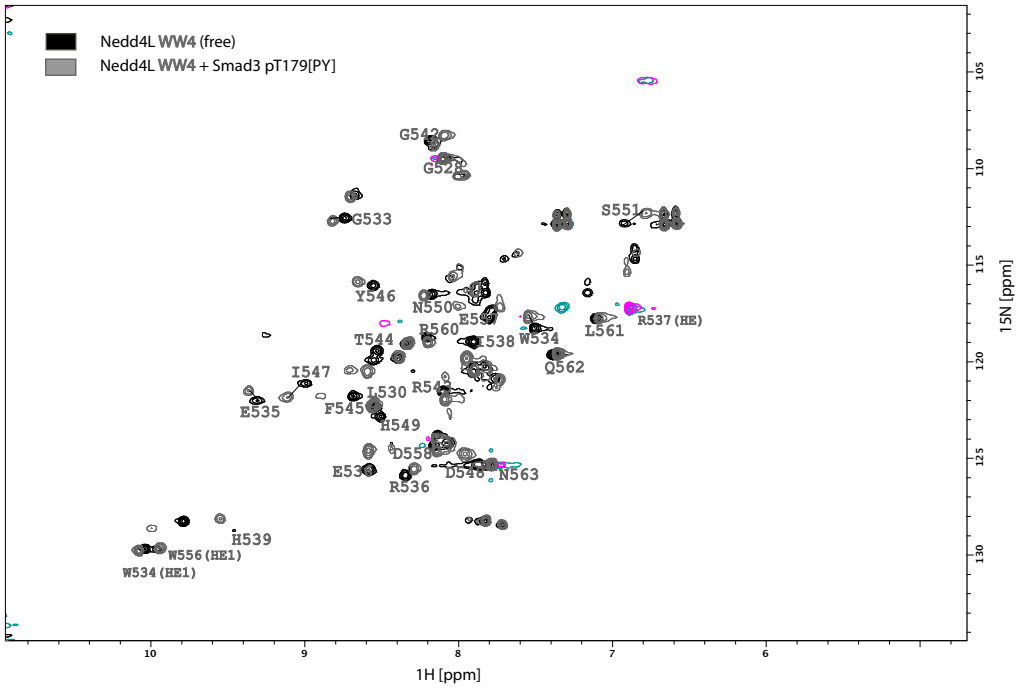


Figure 7.4: ^{1}H ^{15}N-HSQC spectrum of the Nedd4L WW4 domain in the free (black) and bound state (grey) after saturation with Smad3 pT179[PY] peptide.

7.3 Pin1

7.3.1 Full-length hPin1 protein sequence

Table 7.3: Sequence of full-length hPin1 with WW domain in bold black.

	1	11	21	31	41	51	
1	**MADEEKLPPG**	**WEKRMSRSSG**	**RVYYFNHITN**	**ASQWERPSGN**	**S**SSGGKNGQG	EPARVRCSHL	60
61	LVKHSQSRRP	SSWRQEKITR	TKEEALELIN	GYIQKIKSGE	EDFESLASQF	SDCSSAKARG	120
121	DLGAFSRGQM	QKPFEDASFA	LRTGEMSGPV	FTDSGIHIIL	RTE		

7.3.2 Pattern of secondary structure NOEs (Pin1 WW domain)

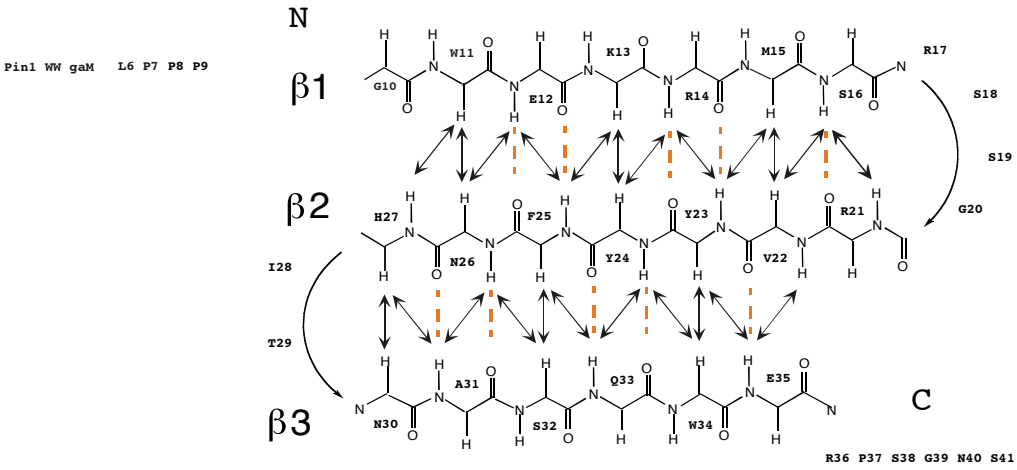


Figure 7.5: Pattern of secondary structure NOEs (represented as black arrows) for the Pin1 WW domain. Hydrogen bonds are illustrated as orange broken lines.

7.4 YAP and TAZ

7.4.1 Full-length hYAP protein sequence

Table 7.4: Sequence of full-length hYAP with WW1 domain in blue and WW2 in bold black.

```
      1           11          21          31          41          51
    1 MDPGQQPPPQ  PAPQGQGQPP  SQPPQGQGPP  SGPGQPAPAA  TQAAPQAPPA  GHQIVHVRGD   60
   61 SETDLEALFN  AVMNPKTANV  PQTVPMRLRK  LPDSFFKPPE  PKSHSRQAST  DAGTAGALTP  120
  121 QHVRAHSSPA  SLQLGAVSPG  TLTPTGVVSG  PAATPTAQHL  RQSSFEIPDD  VPLPAGWEMA  180
  181 KTSSGQRYFL  NHIDQTTTWQ  DPRKAMLSQM  NVTAPTSPPV  QQNMMNSASG  PLPDGWEQAM  240
  241 TQDGEIYYIN  HKNKTTSWLD  PRLDPRFAMN  QRISQSAPVK  QPPPLAPQSP  QGGVMGGSNS  300
  301 NQQQQMRLQQ  LQMEKERLRL  KQQELLRQAM  RNINPSTANS  PKCQELALRS  QLPTLEQDGG  360
  361 TQNPVSSPGM  SQELRTMTTN  SSDPFLNSGT  YHSRDESTDS  GLSMSSYSVP  RTPDDFLNSV  420
  421 DEMDTGDTIN  QSTLPSQQNR  FPDYLEAIPG  TNVDLGTLEG  DGMNIEGEEL  MPSLQEALSS  480
  481 DILNDMESVL  AATKLDKESF  LTWL
```

7.4.2 Full-length hTAZ protein sequence

Table 7.5: Sequence of full-length hTAZ with WW domain in bold black.

```
      1           11          21          31          41          51
    1 MNPASAPPPL  PPPGQQVIHV  TQDLDTDLEA  LFNSVMNPKP  SSWRKKILPE  SFFKEPDSGS   60
   61 HSRQSSTDSS  GGHPGPRLAG  GAQHVRSHSS  PASLQLGTGA  GAAGSPAQQH  AHLRQQSYDV  120
  121 TDELPLPPGW  EMTFTATGQR  YFLNHIEKIT  TWQDPRKAMN  QPLNHMNLHP  AVSSTPVPQR  180
  181 SMAVSQPNLV  MNHQHQQQMA  PSTLSQQNHP  TQNPPAGLMS  MPNALTTQQQ  QQQKLRLQRI  240
  241 QMERERIRMR  QEELMRQEAA  LCRQLPMEAE  TLAPVQAAVN  PPTMTPDMRS  ITNNSSDPFL  300
  301 NGGPYHSREQ  STDSGLGLGC  YSVPTTPEDF  LSNVDEMDTG  ENAGQTPMNI  NPQQTRFPDF  360
  361 LDCLPGTNVD  LGTLESEDLI  PLFNDVESAL  NKSEPFLTWL
```

7.4.3 Pattern of secondary structure NOEs (TAZ WW domain)

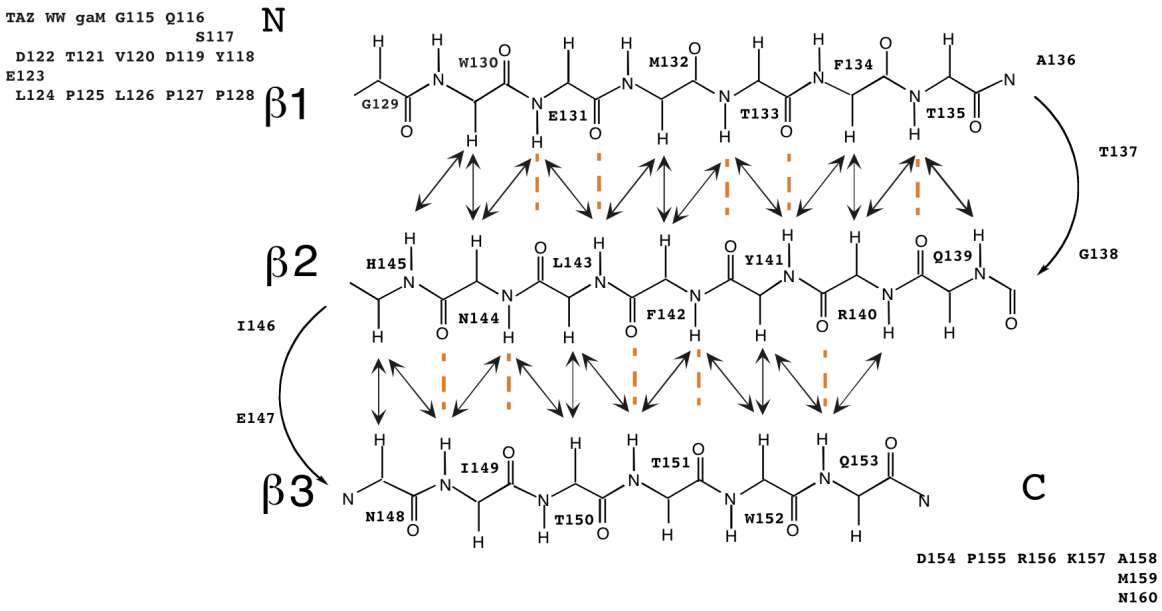


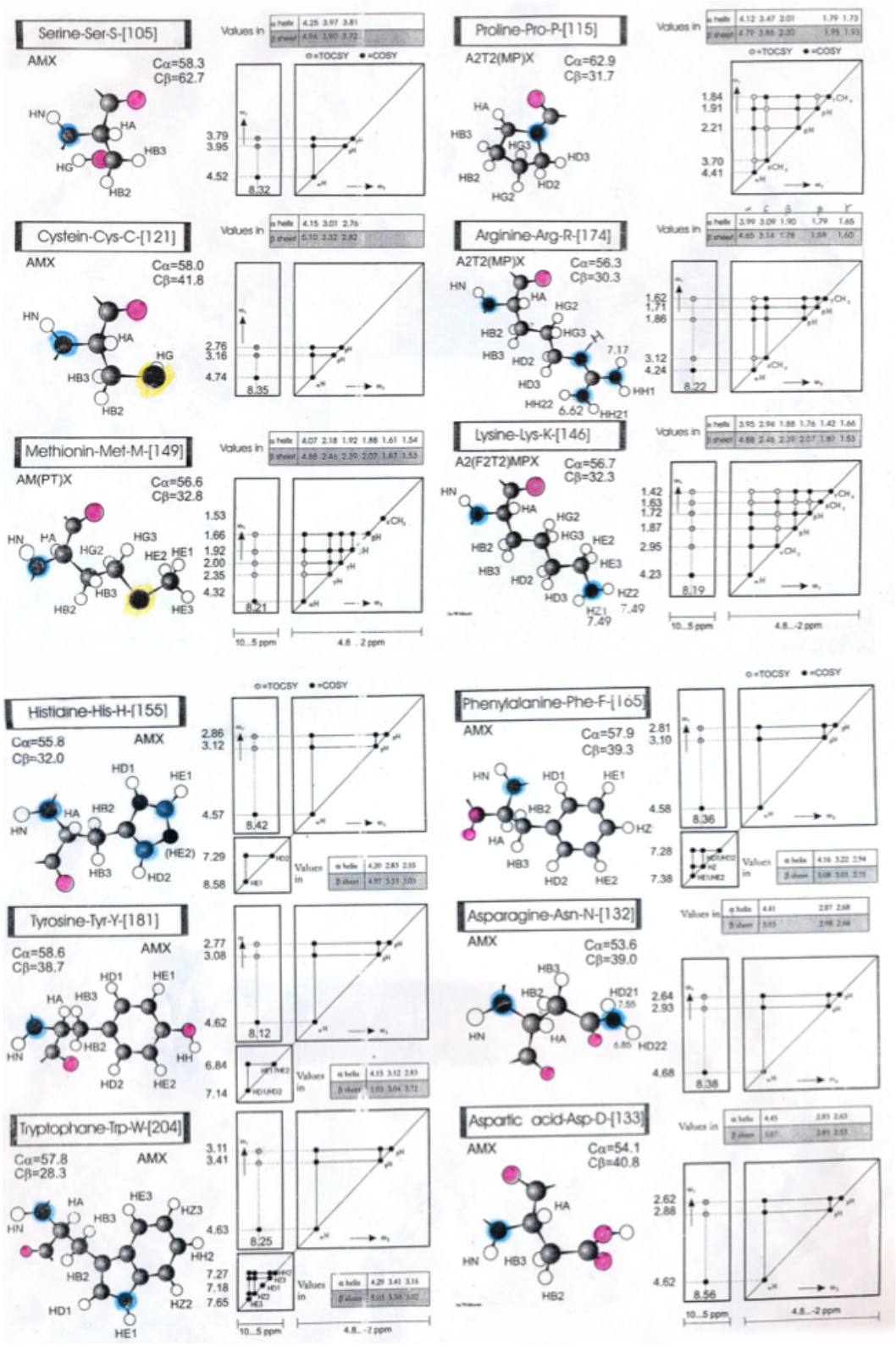
Figure 7.6: Pattern of secondary structure NOEs (represented as black arrows) for the TAZ WW domain. Hydrogen bonds are illustrated as orange broken lines.

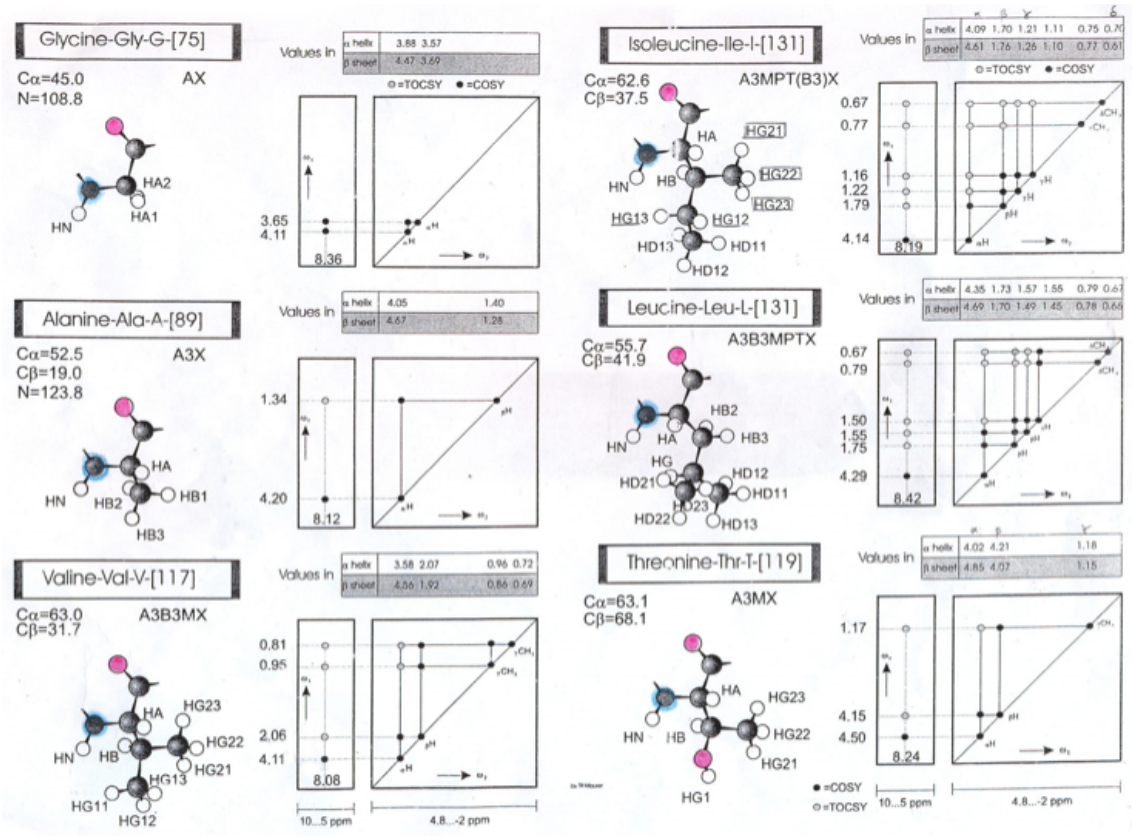
7.5 Structure statistics

Table 7.6: Structure statistic of the NMR refined complexes analysed with PROCHECK. (a) <SA> refers to the ensemble of the 20 structures with the lowest energy. (b) No distance restrains in any of the structures included in the ensemble was violated by more than 0.3 Å.

Ligand	hSmad3 pT179[PY]		hSmad3 pS204-pS208	hSmad3 pS208-pS213	hSmad7 [PY]	
Domain	hNedd4L WW2	hPin1 WW	hNedd4L WW3	hPin1 WW	hNedd4L WW2	hTAZ WW1
Restraints used for calculation <SA>[(a)]						
Sequential	249	167	204	232	275	257
Medium range	47	56	53	110	85	105
Long range	298	215	237	428	325	305
Dihedrals	84	62	78	68	86	67
Hydrogen bonds	10	10	10	10	10	10
Root-mean-square deviation [Å] from experimental [(b)]						
NOE ($x10^3$) [Å]	22.5 ± 0.5	15.9 ± 0.7	17.7 ± 0.3	11.3 ± 1.1	4.7 ± 0.2	10.4 ± 0.2
Bonds ($x10^3$) [Å]	14.2 ± 0.3	8.4 ± 0.3	11.2 ± 0.1	4.4 ± 0.2	4.7 ± 0.4	8.1 ± 0.2
Angles [°]	1.56 ± 0.04	1.08 ± 0.03	1.20 ± 0.03	0.61 ± 0.03	0.66 ± 0.04	0.89 ± 0.02
Ramachandran plot statistics [%]						
Residues in most favoured regions	80.6	84.8	88.6	82.6	85.9	86.1
Residues in allowed regions	19.3	12.7	11.6	13.2	11.8	13.6
CNS potential energy [kcal/mol]						
Total energy	-847.9 ± 40.9	-751 ± 26.9	-1060 ± 28.7	-1487 ± 49	-1372 ± 46	-888.9 ± 18
Electrostatic	-2002 ± 47.5	-1553 ± 42.5	-2153 ± 37.4	-1741 ± 59	-1656 ± 60	-1382 ± 32
Van der Waals	80.6 ± 20.4	8.8 ± 13.1	49.1 ± 14.5	-125.8 ± 15	-125.6 ± 10	-41.1 ± 9.6
Bonds	154 ± 7.9	50.8 ± 3.8	104.7 ± 2.8	14.7 ± 1.4	17.0 ± 1.5 ±	44.9 ± 2.1
Angles	209.9 ± 20.2	229 ± 15.6	369.1 ± 13.3	76.5 ± 7.1	92.5 ± 9.7	150.2 ± 8.2

7.6 Proton nomenclature for NMR assignment




Glycine-Gly-G-[75]
Cα=45.0
N=108.8
AX
HA2
HA1
HN
Values in
α helix 3.88 3.57
β sheet 4.47 3.69
⊖=TOCSY ●=COSY
3.65
4.11
8.36
Alanine-Ala-A-[89]
Cα=52.5
Cβ=19.0
N=123.8
A3X
HA
HN
HB2
HB1
HB3
Values in
α helix 4.05 1.40
β sheet 4.67 1.28
1.34
4.20
8.12
Valine-Val-V-[117]
Cα=63.0
Cβ=31.7
A3B3MX
HA
HG23
HG22
HN
HB
HG21
HG13
HG11
HG12
Values in
α helix 3.58 2.07 0.96 0.72
β sheet 4.56 1.92 0.86 0.69
0.81
0.95
2.06
4.11
8.08
10...5 ppm
4.8...-2 ppm
Isoleucine-Ile-I-[131]
Cα=62.6
Cβ=37.5
A3MPT(B3)X
HA
HG21
HN
HB
HG22
HG23
HG13
HG12
HD13
HD11
HD12
Values in
α helix 4.09 1.70 1.21 1.11 0.75 0.70
β sheet 4.61 1.76 1.26 1.10 0.77 0.61
⊖=TOCSY ●=COSY
0.67
0.77
1.16
1.22
1.79
4.14
8.19
Leucine-Leu-L-[131]
Cα=55.7
Cβ=41.9
A3B3MPTX
HB2
HN
HA
HB3
HG
HD21
HD12
HD11
HD23
HD22
HD13
Values in
α helix 4.35 1.73 1.57 1.55 0.79 0.67
β sheet 4.69 1.70 1.49 1.45 0.78 0.66
0.67
0.79
1.50
1.55
1.75
4.29
8.42
Threonine-Thr-T-[119]
Cα=63.1
Cβ=68.1
A3MX
HA
HG23
HN
HB
HG22
HG21
HG1
Values in
α helix 4.02 4.21 1.18
β sheet 4.65 4.07 1.15
1.17
4.15
4.50
8.24
●=COSY
⊖=TOCSY
10...5 ppm
4.8...-2 ppm
Glutamine-Gln-Q-[146]
AM(PT)X
Cα=56.2
Cβ=30.1
HB3
HG3
HN
HA
HB2
HG2
HE21
6.79
7.60
HE22
Values in
α helix 4.02 2.44 2.33 2.15 2.03
β sheet 4.68 2.25 2.10 2.07 1.88
⊖=TOCSY ●=COSY
1.97
2.14
2.25
2.37
4.31
8.23
Glutamic Acid-Glu-E-[147]
AM(PT)X
Cα=56.7
Cβ=29.7
HG3
HG2
HN
HA
HB3
HB2
Values in
α helix 4.00 2.39 2.29 2.15 2.03
β sheet 4.65 2.26 2.19 2.03 1.93
⊖=TOCSY ●=COSY
2.00
2.11
2.28
2.35
4.22
8.42
10...5 ppm
4.8...-2 ppm
Carbon Chemical Shift Statistics
Cα
Cβ
Cγ
Cδ
Cε
from BMRB restricted data (2003) by Jeremy Jonker
V Val
Y Tyr
W Trp
T Thr
S Ser
P Pro
F Phe
M Met
K Lys
L Leu
I Ile
H His
G Gly
E Glu
Q Gln
C Cys
ox red
D Asp
N Asn
R Arg
A Ala
70 60 50 40 30 20 10 ppm

Glossary

Abbreviations & Units

Å	Ångstrøm (10^{-10} m)
aa	Amino acid
Ac_2O	Acetic anhydride
Amp	Ampillicin
Aria	Ambiguous restraints for interactive assignment
Bis-Tris	Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
BMP	Bone morphogenic protein
Boc	*t*-Butoxycarbonyl
Bzl	Benzyl
C. elegans	Caenorhabditis elegans
CDK8/9	G1 cycline-dependent kinases 8 and 9
cDNA	Complimentary DNA
CKIP-1	Casein kinase-2 interacting protein-1
CNS	Crystallography & NMR system
Co-Smad	Common partner-Smad
COSY	Correlation spectroscopy
CTD	C-terminal domain
Da	Dalton (g/mol)
DCM	Dichloromethane
DHB	2,5-Dihydroxybenzoic acid
DIEA	*N,N*-Diisopropylethylamine
DMF	*N,N*-Dimethylformamid
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
ENaC	Epithelial sodium channel

Abbreviations & Units

ExPASy	Expert protein analysis system
FBP	Formin binding protein
FID	Free induction decay
Fmoc	9*H*-Fluoren-9-ylmethoxycarbonyl
GDF	Growth and differentiation factor
GSK3-ß	Glycogen synthase kinase 3-beta
GST	Glutathion S-transferase
GuHCl	Guanidine hydrochloride
HATU	*N*,*N*-Methylmethan-aminium hexafluorophosphate
HCCA	α-Cyano-4-hydroxycinnamic acid
HECT	Homologous to E6-AP C-terminus
HOAt	*N*-Hydroxy-9-azabenzotriazole
HOBt	*N*-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single-quantum correlation
I-Smad	Inhibitor-Smad
IPTG	Isopropyl-ß-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_A	Association constant
KAN	Kanamycin
K_D	Dissociation constant
Km	Kilometer
LB	Lysogeny broth
LC	Liquid chromatography
M	Molar (g/mol)
MALDI	Matrix-assisted laser desorption ionization
MAPK	Mitrogen-activated protein kinases
MH1/2	Mad-homology1/2
MIS	Muelerian inhibiting substance
MS	Mass spectrometry
Mtt	4-Methyltrityl

Abbreviations & Units

Mw	Molecular weight
N	Stoichiometry
NaOH	Sodiumhydroxid
Nedd4	Neural-precursor-cell-expressed, developmentally down regulated gene 4
Nedd4L	Neural-precursor-cell-expressed, developmentally down regulated gene 4-like
NMR	Nuclear magnetic resonance
NOE	Nuclear overhauser effect
NOESY	Nuclear overhauser effect spectroscopy
OD	Optical density
Pbf	2,2,4,6,7-Pentamethyldihydro-benzofuran-5-sulfonyl
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylen glycol
pH	Pondus hydrogenium
Pin1	Protein interacting with NIMA (never in mitosis A)-1
PMSF	Phenylmethanesulphonyl fluoride
Pol II	RNA polymerase II
PyBOP	Benzotriazol-1-yl-*N*-oxy-tris(pyrrolidino)phosphonium hexafluorophosphate
PyS	2-Thiopyridine
PySSPy	2,2'-Dipyridyldisulfide
R	Gas constant (8.31 J/mol K)
R-Smad	Receptor regulated-Smad
RF	Radio frequency
RING	Really interesting new gene
RNA	Ribonucleic acid
RP	Reverse phase
rpm	Rounds per minute
Rsp5	Reverses SPT-phenotype protein 5
SA	Sinapinic acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sgk1	Serum- and glucocorticoid-regulated kinase

Abbreviations & Units

Smad	Small mothers against decapentalplegic
Smurf	Smad ubiquitin regulatory factor
SOC	Super optimal broth with catabolite repression
SPPS	Solid phase peptide synthesis
Su(dx)	Suppressor of deltex
TAZ	Transcription co-activator with PDZ-binding motif
tBu	*t*-Butyl
TEAD	Transcriptional enhancer factor-domain
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TGF-ß	Transforming growths factor beta
TIS	Triisopropylsilan
™	Unregistered trade mark
TOCSY	Total correlation spectroscopy
TOF	Time of flight
Tris	Tris(hydroxymethyl)aminomethane
Trt	Trityl
TßRI	TGF-ß receptor type I
TßRII	TGF-ß receptor type II
Ub	Ubiquitin
WT	Wild type
WWP1	WW domain containing protein 1
WWRT1	WW domain containing transcription regulator 1
YAP	Yes-kinase associated protein
γ	Gyromagnetic ratio
ε	Extinction factor
μ	Nuclear magnetic moment

List of Figures

List of Tables

References

Alarcon, C., Zaromytidou, A.I., Xi, Q., Gao, S., Yu, J., Fujisawa, S., Barlas, A., Miller, A.N., Manova-Todorova, K., Macias, M.J. et al. 2009. Nuclear CDKs drive Smad transcriptional activation and turnover in BMP and TGF-ß pathways. *Cell* **139**(4): 757-769.

Aragon, E., Goerner, N., Zaromytidou, A.I., Xi, Q., Escobedo, A., Massague, J., and Macias, M.J. 2011. A Smad action turnover switch operated by WW domain readers of a phosphoserine code. *Genes Dev* **25**(12): 1275-1288.

Asher, C., Sinha, I., and Garty, H. 2003. Characterization of the interactions between Nedd4-2, ENaC, and sgk-1 using surface plasmon resonance. *Biochim Biophys Acta* **1612**(1): 59-64.

Attisano, L. and Wrana, J.L. 2002. Signal transduction by the TGF-ß superfamily. *Science* **296**(5573): 1646-1647.

Baca, M., Muir, T.W., Schnolzer, M., and Kent, S.B.H. 1995. Chemical Ligation of Cysteine-Containing Peptides: Synthesis of a 22 kDa Tethered Dimer of HIV-1 Protease. *J Am Chem Soc* **117**: 1881–1887.

Baca M., M.T.W., Schnölzer M, Kent S. B. H. 1995. Chemical Ligation of Cysteine-Containing Peptides: Synthesis fo a 22 kDa Tethered Dimer of HIV-1 Protease. *J Am Chem Soc* **117**: 1881-1887.

Bartels Christian, X.T.-h., Billeter Martin, Güntert Peter and Wüthrich Kurt 1995. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules *Journal of Biomolecular NMR* **6**: 1-10.

Bernassola, F., Karin, M., Ciechanover, A., and Melino, G. 2008. The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell* **14**(1): 10-21.

Bhalla, V., Daidie, D., Li, H., Pao, A.C., LaGrange, L.P., Wang, J., Vandewalle, A., Stockand, J.D., Staub, O., and Pearce, D. 2005. Serum- and glucocorticoid-regulated kinase 1 regulates ubiquitin ligase neural precursor cell-expressed, developmentally down-regulated protein 4-2 by inducing interaction with 14-3-3. *Mol Endocrinol* **19**(12): 3073-3084.

Bhattacharya, S., Zoladek, T., and Haines, D.S. 2008. WW domains 2 and 3 of Rsp5p play overlapping roles in binding to the LPKY motif of Spt23p and Mga2p. *Int J Biochem Cell Biol* **40**(1): 147-157.

Bimboim H C, D.J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA *Nucleic acids research* **7**: 1513-1523.

Bork, P. and Sudol, M. 1994. The WW domain: a signalling site in dystrophin? *Trends Biochem Sci* **19**(12): 531-533.

Brunger, A.T. 2007. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**(11): 2728-2733.

Carpino, L.A. 1972. The 9-Fluorenylmethoxycarbonyl Amino-Protecting Group. *J Org Chem,* **37**(22): 3404-3409.

Chen, C. and Matesic, L.E. 2007. The Nedd4-like family of E3 ubiquitin ligases and cancer. *Cancer Metastasis Rev* **26**(3-4): 587-604.

Chong, P.A., Lin, H., Wrana, J.L., and Forman-Kay, J.D. 2006. An expanded WW domain recognition motif revealed by the interaction between Smad7 and the E3 ubiquitin ligase Smurf2. *J Biol Chem* **281**(25): 17069-17075.

Chong, P.A. 2010. Coupling of tandem Smad ubiquitination regulatory factor (Smurf) WW domains modulates target specificity. *Proc Natl Acad Sci U S A*.

Clore, G.M. and Gronenborn, A.M. 1998. Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol* **16**(1): 22-34.

Cohen, P. and Frame, S. 2001. The renaissance of GSK3. *Nat Rev Mol Cell Biol* **2**(10): 769-776.

Debonneville, C., Flores, S.Y., Kamynina, E., Plant, P.J., Tauxe, C., Thomas, M.A., Munster, C., Chraibi, A., Pratt, J.H., Horisberger, J.D. et al. 2001. Phosphorylation of Nedd4-2 by Sgk1 regulates epithelial Na(+) channel cell surface expression. *EMBO J* **20**(24): 7052-7059.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**(3): 277-293.

Evans, J.N.S. 1995. Biomolecular NMR Spectroscopy. *Oxford University Press*.

Feng, X.H. and Derynck, R. 2005. Specificity and versatility in TGF-ß signaling through Smads. *Annu Rev Cell Dev Biol* **21**: 659-693.

Ferrigno, O., Lallemand, F., Verrecchia, F., L'Hoste, S., Camonis, J., Atfi, A., and Mauviel, A. 2002. Yes-associated protein (YAP65) interacts with Smad7 and potentiates its inhibitory activity against TGF-ß/Smad signaling. *Oncogene* **21**(32): 4879-4884.

Fotia, A.B., Dinudom, A., Shearwin, K.E., Koch, J.P., Korbmacher, C., Cook, D.I., and Kumar, S. 2003. The role of individual Nedd4-2 (KIAA0439) WW domains in binding and regulating epithelial sodium channels. *FASEB J* **17**(1): 70-72.

Fuentealba, L.C., Eivers, E., Ikeda, A., Hurtado, C., Kuroda, H., Pera, E.M., and De Robertis, E.M. 2007. Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell* **131**(5): 980-993.

Gao, S., Alarcon, C., Sapkota, G., Rahman, S., Chen, P.Y., Goerner, N., Macias, M.J., Erdjument-Bromage, H., Tempst, P., and Massague, J. 2009. Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF-ß signaling. *Mol Cell* **36**(3): 457-468.

Garcia-Martin, F., Quintanar-Audelo, M., Garcia-Ramos, Y., Cruz, L.J., Gravel, C., Furic, R., Cote, S., Tulla-Puche, J., and Albericio, F. 2006. ChemMatrix, a poly(ethylene glycol)-based support for the solid-phase synthesis of complex peptides. *J Comb Chem* **8**(2): 213-220.

Grimm, O.H. and Gurdon, J.B. 2002. Nuclear exclusion of Smad2 is a mechanism leading to loss of competence. *Nat Cell Biol* **4**(7): 519-522.

Groppe, J., Hinck, C.S., Samavarchi-Tehrani, P., Zubieta, C., Schuermann, J.P., Taylor, A.B., Schwarz, P.M., Wrana, J.L., and Hinck, A.P. 2008. Cooperative assembly of TGF-ß superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Mol Cell* **29**(2): 157-168.

Henry, P.C., Kanelis, V., O'Brien, M.C., Kim, B., Gautschi, I., Forman-Kay, J., Schild, L., and Rotin, D. 2003. Affinity and specificity of interactions between Nedd4 isoforms and the epithelial Na+ channel. *J Biol Chem* **278**(22): 20019-20028.

Hochstrasser, M. 2009. Origin and function of ubiquitin-like proteins. *Nature* **458**(7237): 422-429.

Hore P.J., J.J.A., Wimperis S. 2000. NMR: The Toolkit. *Oxford Chemistry Primers*.

Hui-Ming Yu, S.-T.C., Kung-Tsung Wang. 1992. Enhanced coupling efficiency in solid-phase peptide synthesis by microwave irradiation. *J Org Chem,* **57**(18): 4781-4784.

Hutchison, C.A., 3rd, Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P., and Smith, M. 1978. Mutagenesis at a specific position in a DNA sequence. *J Biol Chem* **253**(18): 6551-6560.

Ingham, R.J., Gish, G., and Pawson, T. 2004. The Nedd4 family of E3 ubiquitin ligases: functional diversity within a common modular architecture. *Oncogene* **23**(11): 1972-1984.

Itoh, S., Landstrom, M., Hermansson, A., Itoh, F., Heldin, C.H., Heldin, N.E., and ten Dijke, P. 1998. Transforming growth factor ß1 induces nuclear export of inhibitory Smad7. *J Biol Chem* **273**(44): 29195-29201.

Jennings, M.D., Blankley, R.T., Baron, M., Golovanov, A.P., and Avis, J.M. 2007. Specificity and autoregulation of Notch binding by tandem WW domains in suppressor of Deltex. *J Biol Chem* **282**(39): 29032-29042.

Kanai, F., Marignani, P.A., Sarbassova, D., Yagi, R., Hall, R.A., Donowitz, M., Hisaminato, A., Fujiwara, T., Ito, Y., Cantley, L.C. et al. 2000. TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *EMBO J* **19**(24): 6778-6791.

Kanelis, V., Rotin, D., and Forman-Kay, J.D. 2001. Solution structure of a Nedd4 WW domain-ENaC peptide complex. *Nat Struct Biol* **8**(5): 407-412.

Kapust, R.B., Tozser, J., Copeland, T.D., and Waugh, D.S. 2002. The P1' specificity of tobacco etch virus protease. *Biochem Biophys Res Commun* **294**(5): 949-955.

Karas, M. and Hillenkamp, F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**(20): 2299-2301.

Kato, Y., Ito, M., Kawai, K., Nagata, K., and Tanokura, M. 2002. Determinants of ligand specificity in groups I and IV WW domains as studied by surface plasmon resonance and model building. *J Biol Chem* **277**(12): 10173-10177.

Kato, Y., Nagata, K., Takahashi, M., Lian, L., Herrero, J.J., Sudol, M., and Tanokura, M. 2004. Common mechanism of ligand recognition by group II/III WW domains: redefining their functional classification. *J Biol Chem* **279**(30): 31833-31841.

Kavsak, P., Rasmussen, R.K., Causing, C.G., Bonni, S., Zhu, H., Thomsen, G.H., and Wrana, J.L. 2000. Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF-ß receptor for degradation. *Mol Cell* **6**(6): 1365-1375.

Keeler, J. 2002. Understanding NMR Spectroscopy. *Wiley*.

Koepf, E.K., Petrassi, H.M., Ratnaswamy, G., Huff, M.E., Sudol, M., and Kelly, J.W. 1999. Characterization of the structure and function of W --> F WW domain variants: identification of a natively unfolded protein that folds upon ligand binding. *Biochemistry* **38**(43): 14338-14351.

Komuro, A., Imamura, T., Saitoh, M., Yoshida, Y., Yamori, T., Miyazono, K., and Miyazawa, K. 2004. Negative regulation of transforming growth factor-ß (TGF-ß) signaling by WW domain-containing protein 1 (WWP1). *Oncogene* **23**(41): 6914-6923.

Kretzschmar, M., Doody, J., Timokhina, I., and Massague, J. 1999. A mechanism of repression of TGF-ß/ Smad signaling by oncogenic Ras. *Genes Dev* **13**(7): 804-816.

Kuratomi, G., Komuro, A., Goto, K., Shinozaki, M., Miyazawa, K., Miyazono, K., and Imamura, T. 2005. Nedd4-2 (neural precursor cell expressed, developmentally down-regulated 4-2) negatively regulates TGF-ß (transforming growth factor-ß) signalling by inducing ubiquitin-mediated degradation of Smad2 and TGF-ß type I receptor. *Biochem J* **386**(Pt 3): 461-470.

Kussmann Martin, N.E., Rahbek-Nielsen Henrik , Haebel Sophie , Rossel-Larsen Martin, Jakobsen Lene , Gobom Johan, Mirgorodskaya Ekatarina , Kroll-Kristensen Anne , Palm Lisbeth, Roepstorff Peter. 1997. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry Sample Preparation Techniques Designed for Various Peptide and Protein Analytes. *Journal of Mass Spectrometry* **32**(6): 593-601.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**(4): 477-486.

Liang, X., Butterworth, M.B., Peters, K.W., Walker, W.H., and Frizzell, R.A. 2008. An obligatory heterodimer of 14-3-3-ß and 14-3-3-ε is required for aldosterone regulation of the epithelial sodium channel. *J Biol Chem* **283**(41): 27418-27425.

Lippens, G., Landrieu, I., and Smet, C. 2007. Molecular mechanisms of the phospho-dependent prolyl cis/trans isomerase Pin1. *FEBS J* **274**(20): 5211-5222.

Lu, K.P., Finn, G., Lee, T.H., and Nicholson, L.K. 2007. Prolyl cis-trans isomerization as a molecular timer. *Nat Chem Biol* **3**(10): 619-629.

Lu, K.P., Hanes, S.D., and Hunter, T. 1996. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **380**(6574): 544-547.

Lu, K.P. and Zhou, X.Z. 2007. The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat Rev Mol Cell Biol* **8**(11): 904-916.

Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. 1999. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **283**(5406): 1325-1328.

Macias, M.J., Gervais, V., Civera, C., and Oschkinat, H. 2000. Structural analysis of WW domains and design of a WW prototype. *Nat Struct Biol* **7**(5): 375-379.

Macias, M.J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. 1996. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* **382**(6592): 646-649.

Macias, M.J., Wiesner, S., and Sudol, M. 2002. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* **513**(1): 30-37.

Marley, J., Lu, M., and Bracken, C. 2001. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**(1): 71-75.

Massague, J. 1998. TGF-ß signal transduction. *Annu Rev Biochem* **67**: 753-791.

Massague, J. and Chen, Y.G. 2000. Controlling TGF-ß signaling. *Genes Dev* **14**(6): 627-644.

Massague, J. and Gomis, R.R. 2006. The logic of TGF-ß signaling. *FEBS Lett* **580**(12): 2811-2820.

Massague, J., Seoane, J., and Wotton, D. 2005. Smad transcription factors. *Genes Dev* **19**(23): 2783-2810.

Matsuura, I., Chiang, K.N., Lai, C.Y., He, D., Wang, G., Ramkumar, R., Uchida, T., Ryo, A., Lu, K., and Liu, F. 2010. Pin1 promotes transforming growth factor-ß-induced migration and invasion. *J Biol Chem* **285**(3): 1754-1764.

Matsuura, I., Denissova, N.G., Wang, G., He, D., Long, J., and Liu, F. 2004. Cyclin-dependent kinases regulate the antiproliferative function of Smads. *Nature* **430**(6996): 226-231.

Matsuzaki, K., Kitano, C., Murata, M., Sekimoto, G., Yoshida, K., Uemura, Y., Seki, T., Taketani, S., Fujisawa, J., and Okazaki, K. 2009. Smad2 and Smad3 phosphorylated at both linker and COOH-terminal regions transmit malignant TGF-ß signal in later stages of human colorectal cancer. *Cancer Res* **69**(13): 5321-5330.

Merrifield, R.B. 1962. Peptide synthesis on a solid polymer. *Fed Proc Amer Soc Exp Biol* **21**: 412.

Merrifield, R.B. 1963. Solid Phase Peptide Synthesis II. The Synthesis of a Tetrapeptide. *J Am Chem Soc* **85**: 2149-2154.

Millet, C., Yamashita, M., Heller, M., Yu, L.R., Veenstra, T.D., and Zhang, Y.E. 2009. A negative feedback control of transforming growth factor-ß signaling by glycogen synthase kinase 3-mediated Smad3 linker phosphorylation at Ser-204. *J Biol Chem* **284**(30): 19808-19816.

Morales, B., Ramirez-Espain, X., Shaw, A.Z., Martin-Malpartida, P., Yraola, F., Sanchez-Tillo, E., Farrera, C., Celada, A., Royo, M., and Macias, M.J. 2007. NMR structural studies of the ItchWW3 domain reveal that phosphorylation at T30 inhibits the interaction with PPxY-containing ligands. *Structure* **15**(4): 473-483.

Mori, S., Matsuzaki, K., Yoshida, K., Furukawa, F., Tahashi, Y., Yamagata, H., Sekimoto, G., Seki, T., Matsui, H., Nishizawa, M. et al. 2004. TGF-ß and HGF transmit the signals through JNK-dependent Smad2/3 phosphorylation at the linker regions. *Oncogene* **23**(44): 7416-7429.

Nagaki, K., Yamamura, H., Shimada, S., Saito, T., Hisanaga, S., Taoka, M., Isobe, T., and Ichimura, T. 2006. 14-3-3 Mediates phosphorylation-dependent inhibition of the interaction between the ubiquitin E3 ligase Nedd4-2 and epithelial Na+ channels. *Biochemistry* **45**(21): 6733-6740.

Nakano, A., Koinuma, D., Miyazawa, K., Uchida, T., Saitoh, M., Kawabata, M., Hanai, J., Akiyama, H., Abe, M., Miyazono, K. et al. 2009. Pin1 down-regulates transforming growth factor-ß (TGF-ß) signaling by inducing degradation of Smad proteins. *J Biol Chem* **284**(10): 6109-6115.

Nilges, M., Macias, M.J., O'Donoghue, S.I., and Oschkinat, H. 1997. Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from ß-spectrin. *J Mol Biol* **269**(3): 408-422.

Palasek, S.A., Cox, Z.J., and Collins, J.M. 2007. Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. *J Pept Sci* **13**(3): 143-148.

Pearson D A, B.M., Bake M L, Guindon C A. 1989. Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis *Tetrahedron Letters* **30**(21): 2739-2742.

Pierce, M.M., Raman, C.S., and Nall, B.T. 1999. Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**(2): 213-221.

Pires, J.R., Taha-Nejad, F., Toepert, F., Ast, T., Hoffmuller, U., Schneider-Mergener, J., Kuhne, R., Macias, M.J., and Oschkinat, H. 2001. Solution structures of the YAP65 WW domain and the variant L30 K 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): 1147-1156.

Rabi, Z., Millman, Kusch. 1938. *Phys Rev* **53**: 318.

Ramirez-Espain, X., Ruiz, L., Martin-Malpartida, P., Oschkinat, H., and Macias, M.J. 2007. Structural characterization of a new binding motif and a novel binding mode in group 2 WW domains. *J Mol Biol* **373**(5): 1255-1268.

Ranganathan, R., Lu, K.P., Hunter, T., and Noel, J.P. 1997. Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* **89**(6): 875-886.

Rayasam, G.V., Tulasi, V.K., Sodhi, R., Davis, J.A., and Ray, A. 2009. Glycogen synthase kinase 3: more than a namesake. *Br J Pharmacol* **156**(6): 885-898.

Ross, S. and Hill, C.S. 2008. How the Smads regulate transcription. *Int J Biochem Cell Biol* **40**(3): 383-408.

Sapkota, G., Alarcon, C., Spagnoli, F.M., Brivanlou, A.H., and Massague, J. 2007. Balancing BMP signaling through integrated inputs into the Smad1 linker. *Mol Cell* **25**(3): 441-454.

Sarin, V.K., Kent, S.B., Tam, J.P., and Merrifield, R.B. 1981. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. *Anal Biochem* **117**(1): 147-157.

Scheffner, M. and Staub, O. 2007. HECT E3s and human disease. *BMC Biochem* **8 Suppl 1**: S6.

Shi, H., Asher, C., Chigaev, A., Yung, Y., Reuveny, E., Seger, R., and Garty, H. 2002a. Interactions of ß and γ ENaC with Nedd4 can be facilitated by an ERK-mediated phosphorylation. *J Biol Chem* **277**(16): 13539-13547.

Shi, H., Asher, C., Yung, Y., Kligman, L., Reuveny, E., Seger, R., and Garty, H. 2002b. Casein kinase 2 specifically binds to and phosphorylates the carboxy termini of ENaC subunits. *Eur J Biochem* **269**(18): 4551-4558.

Shi, Y. and Massague, J. 2003. Mechanisms of TGF-ß signaling from cell membrane to the nucleus. *Cell* **113**(6): 685-700.

Snyder, P.M. 2005. Minireview: Reuglation of Epithelial Na+ Channel Trafficking. *Endocrinology* **146**: 5079-5085.

Sudol, M. and Hunter, T. 2000. NeW wrinkles for an old domain. *Cell* **103**(7): 1001-1004.

Sudol, M., Recinos, C.C., Abraczinskas, J., Humbert, J., and Farooq, A. 2005. WW or WoW: the WW domains in a union of bliss. *IUBMB Life* **57**(12): 773-778.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**(22): 4673-4680.

Varelas, X., Sakuma, R., Samavarchi-Tehrani, P., Peerani, R., Rao, B.M., Dembowy, J., Yaffe, M.B., Zandstra, P.W., and Wrana, J.L. 2008. TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal. *Nat Cell Biol* **10**(7): 837-848.

Varelas, X., Samavarchi-Tehrani, P., Narimatsu, M., Weiss, A., Cockburn, K., Larsen, B.G., Rossant, J., and Wrana, J.L. 2010. The Crumbs complex couples cell density sensing to Hippo-dependent control of the TGF-ß-Smad pathway. *Dev Cell* **19**(6): 831-844.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. 2000. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **7**(8): 639-643.

Vojkovsky, T. 1995. Detection of secondary amines on solid phase. *Pept Res* **8**(4): 236-237.

Wang, G., Matsuura, I., He, D., and Liu, F. 2009a. Transforming growth factor-{ß}-inducible phosphorylation of Smad3. *J Biol Chem* **284**(15): 9663-9673.

Wang, K., Degerny, C., Xu, M., and Yang, X.J. 2009b. YAP, TAZ, and Yorkie: a conserved family of signal-responsive transcriptional coregulators in animal development and human disease. *Biochem Cell Biol* **87**(1): 77-91.

Webb, C., Upadhyay, A., Giuntini, F., Eggleston, I., Furutani-Seiki, M., Ishima, R., and Bagby, S. 2011. Structural features and ligand binding properties of tandem WW domains from YAP and TAZ, nuclear effectors of the Hippo pathway. *Biochemistry* **50**(16): 3300-3309.

Wiemuth, D., Lott, J.S., Ly, K., Ke, Y., Teesdale-Spittle, P., Snyder, P.M., and McDonald, F.J. 2010. Interaction of serum- and glucocorticoid regulated kinase 1 (SGK1) with the

WW-domains of Nedd4-2 is required for epithelial sodium channel regulation. *PLoS One* **5**(8): e12163.

Williamson, M.P., Havel, T.F., and Wuthrich, K. 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. *J Mol Biol* **182**(2): 295-315.

Wintjens, R., Wieruszeski, J.M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., and Landrieu, I. 2001. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J Biol Chem* **276**(27): 25150-25156.

Wulf, G., Finn, G., Suizu, F., and Lu, K.P. 2005. Phosphorylation-specific prolyl isomerization: is there an underlying theme? *Nat Cell Biol* **7**(5): 435-441.

Wuthrich, K. 1990. Protein structure determination in solution by NMR spectroscopy. *J Biol Chem* **265**(36): 22059-22062.

Yaffe, M.B., Rittinger, K., Volinia, S., Caron, P.R., Aitken, A., Leffers, H., Gamblin, S.J., Smerdon, S.J., and Cantley, L.C. 1997. The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91**(7): 961-971.

Yagi, R., Chen, L.F., Shigesada, K., Murakami, Y., and Ito, Y. 1999. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO J* **18**(9): 2551-2562.

Yang, B. and Kumar, S. 2010. Nedd4 and Nedd4-2: closely related ubiquitin-protein ligases with distinct physiological functions. *Cell Death Differ* **17**(1): 68-77.

Yeh, E.S. and Means, A.R. 2007. PIN1, the cell cycle and cancer. *Nat Rev Cancer* **7**(5): 381-388.

Zarrinpar, A. and Lim, W.A. 2000. Converging on proline: the mechanism of WW domain peptide recognition. *Nat Struct Biol* **7**(8): 611-613.

Zenobi, R.K., R. 1998. Ion Formation in MALDI Mass Spectrometry. *Mass Spectrom Rev* **17**: 337-366.

Zhang, Y., Chang, C., Gehling, D.J., Hemmati-Brivanlou, A., and Derynck, R. 2001. Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase. *Proc Natl Acad Sci U S A* **98**(3): 974-979.

Zhao, B., Lei, Q.Y., and Guan, K.L. 2008. The Hippo-YAP pathway: new connections between regulation of organ size and cancer. *Curr Opin Cell Biol* **20**(6): 638-646.

Zhu, H., Kavsak, P., Abdollah, S., Wrana, J.L., and Thomsen, G.H. 1999. A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. *Nature* **400**(6745): 687-693.