Mechanisms underlying the regulation of Nedd4-family E3 Ubiquitin ligases

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Natalia Ruetalo Buschinger
aus Montevideo, Uruguay

Tübingen 2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	08.12.2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Dr. Silke Wiesner
2. Berichterstatter:	Prof. Dr. Thilo Stehle

Table of contents

Summary

Ubiquitination is a post-translational modification that involves the covalent attachment of one or several ubiquitin (Ub) molecules to a substrate protein. Initially, it was considered as a mechanism to control the abundance of proteins in the cell. These days it is evident that ubiquitination fulfils many other roles, being involved in a wide array of cellular functions. Among the enzymes that participate in the ubiquitination pathway, E3 ligases stand out since they define the specificity of the reaction and determine which type of Ub chains is attached to the substrate. This in turn, defines the fate of the substrate inside the cell. From the approximately 600 E3 ligases that the human genome encodes, 28 belong to the HECT (Homologous to the E6-AP C terminus) family of E3s, a particular group that has intrinsic catalytic activity by forming a thioester intermediate before Ub is transferred to the substrate. Numerous studies associate aberrant expression or mutations in HECT-type E3s with various human diseases including cancer, where they can act both as oncoproteins or tumor suppressors. It is thus of utmost importance to understand in detail catalytic mechanism and the regulation of HECT-type Ub ligases.

Nedd4-family Ub ligases are a subgroup of HECT E3s that contain an N-terminal C2 domain, two to four central WW domains and a C-terminal HECT domain. For a few members of this family it has been shown that they are kept in an auto-inhibited conformation that prevents untimely ubiquitination. However, the mechanistic details were not clear. Here, I studied the basis of C2-mediated auto-inhibition in two members of the Nedd4-family: Smurf2 and Nedd4. Using NMR spectroscopy and biochemical assays I found that binding of the C2 domain to the HECT domain impairs E2-E3 transthiolation, by locking the enzyme in an incompetent catalytic conformation. Moreover, C2 binding abolishes non-covalent interaction of the HECT domain with Ub, which is known to be an important step in Ub chain elongation.

The Smurf1 E3 ligase is an E3 that shares more that 70% sequence identity with Smurf2. However, it was not clear whether Smurf1 is regulated by the same C2:HECT auto-inhibitory mechanism as Smurf2. My studies revealed that although the C2:HECT binding surface is conserved between Smurf1 and Smurf2 and the Smurf1 C2 domain inhibits HECT activity *in trans*, the full length (FL) enzyme is constitutively active. I was able to show that this strikingly difference between Smurf1 and Smurf2 is due to the lack of the WW1 domain in Smurf1, that plays a role in enhancing C2:HECT binding affinity.

Finally, Ca^{2+}-mediated release of the C2 domain is as a mechanism for Nedd4 E3 ligase activation. I showed that the Nedd4 C2 domain binds calcium through the conserved calcium binding region and thereby competes with HECT domain interaction. However, I could also show that other members of the Nedd4 family such as Smurf1, Smurf2 and the yeast homologous Rsp5 are unable to interact with calcium. This suggests that Ca^{2+}-

mediated release of the C2 domain from the HECT domain as an activation mechanism is likely restricted to a small subset of Nedd4 proteins.

Overall, Nedd4-family enzymes play key roles in the establishment and regulation of both developmental and carcinogenic processes. Hence, profound studies of the catalytic mechanisms and detailed analysis of protein-specific regulation of these enzymes as presented in this thesis, contribute significantly to the understanding of HECT ligases but also may have implications for the development of pharmaceutical inhibitors.

Zusammenfassung

Ubiquitinierung ist eine posttranslationale Modifikation, die mit der kovalenten Übertragung eines oder mehrerer Ubiquitinmoleküle auf ein Substratprotein einhergeht. Ursprünglich wurde angenommen, dass es sich hierbei um einen Mechanismus zur Mengenkontrolle von Proteinen in der Zelle handelt. Heutzutage gilt es allerdings als erwiesen, dass die Ubiquitinierung eine Vielzahl an Funktionen erfüllt, die in einer großen Zahl von zellularen Prozessen eine wichtige Rolle spielen. Unter den Enzymen, die während der Ubiquitinierung involviert sind, spielen die E3-Ligasen eine besonders wichtige Rolle, da sie die Reaktionsspezifität festlegen und somit den Ubiquitinkettentyp bestimmen. Der Ubiquitinkettentyp wiederum bestimmt das Schicksal des Substratproteins. Im menschlichen Genom sind etwa 600 E3-Ligase kodiert von denen 28 zur Familie der HECT-Ligasen (Homologous tot he E6AP C terminus) gehören. Alle Mitglieder dieser Genfamilie besitzen eine intrinsische katalytische Aktivität, da sie mit Ubiquitin (Ub) als Reaktionsintermediat einen Thioester ausbilden, bevor es auf das Substrat übertragen wird. In einer Vielzahl von Studien assoziieren eine fehlerhafte Expression oder Mutationen von HECT E3-Ligasen mit verschiedenen Krankheiten wie z.B. Krebs wobei sie sowohl als Onkoproteine als auch Tumorsupressoren auftreten können. Deshalb ist es von äußerster Wichtigkeit, dass die Details des Katalysemechanismus und der Regulation der HECT-Ubiquitin-Ligasen untersucht und verstanden werden.

Die Ub-Ligasen der Nedd4-Familie sind eine Untergruppe der HECT-E3-Ligasen und bestehen aus einer N-terminalen C2-Domäne, zwei bis vier WW-Domänen und einer C-terminalen HECT-Domäne. Für einige Mitglieder dieser Genfamilie wurde bereits gezeigt, dass sie in einer autoinhibitorischen Konformation vorliegen die schlussendlich die Ubiquitinierungsreaktion unterbindet. Allerdings waren die Details des zugrundeliegenden Mechanismus nicht bekannt. In dieser Doktorarbeit habe ich die Grundlagen der C2-vermittelten Autoinhibition zweier Mitglieder der Nedd4-Familie, Smurf2 und Nedd4, untersucht. Mit Hilfe von NMR-Spektroskopie und biochemischen Untersuchungen habe ich herausgefunden, dass die Bindung der C2- an die HECT-Domäne die E2-E3-Transthiolierung verhindert indem sie das Enzym in einer katalytisch inaktiven Konformation fest hält. Ferner verhindert die C2-Bindung die nicht-kovalente Interaktion der HECT-Domäne mit Ub, die einen wichtigen Schritt in der Verlängerung von Ub-Ketten darstellt.

Die Smurf1 E3-Ligase ist ein E3-Enzym, das mehr als 70 % Sequenzidentität mit Smurf2 aufweist. Allerdings war bis heute unbekannt, ob Smurf1 über denselben C2:HECT Autoinhibitionsmechanismus wie Smurf2 verfügt. Meine Studien haben aufgezeigt, dass das Smurf1 Volllängeenzym konstitutiv aktiv ist obwohl die C2:HECT Bindungsoberfläche zwischen Smurf1 und Smurf2 konserviert ist und die Smurf1 C2-Domäne die HECT-

Aktivität *in trans* inhibiert. Ich konnte zeigen, dass dieser gravierende Unterschied zwischen Smurf1 und Smurf2 auf die Abwesenheit der WW1 Domäne in Smurf1 zurück zu führen ist, die die Bindeaffinität zwischen C2- und HECT-Domäne erhöht.
Darüber hinaus ist die Ca^{2+}-abhängige Freisetzung der C2-Domäne ein Aktivierungsmechanismus für Nedd4-E3-Ligasen. Ich konnte zeigen, dass Calcium an die C2-Domäne von Nedd4 mittels einer konservierten Calciumbindedomäne bindet und so mit der Interaktion zwischen C2- und HECT-Domäne konkurierrt. Allerdings konnte ich auch nachweisen, dass andere Mitglieder der Nedd4-Familie, wie zum Beispiel Smurf1, Smurf2 und das Hefehomolog Rsp5 nicht in der Lage sind mit Calcium zu interagieren. Diese Ergebnisse deuten darauf hin, dass der Aktivierungsmechanismus über die Ca^{2+}-abhängige Freisetzung der C2-Domäne von der HECT-Domäne wahrscheinlich auf eine kleine Untergruppe der Nedd4-Proteine beschränkt ist.
Global betrachtet spielen die Enzyme der Nedd4-Familie eine Schlüsselrolle bei der Etablierung und Regulation von entwicklungsbiologischen und krebsassoziierten Prozessen. Daher sind intensive Studien des Katalysemechanismus und eine detaillierte Analyse der proteinspezifischen Regulation dieser Enzyme so wie sie hier in dieser Studie vorgestellt werden ein integraler Bestandteil zum Verständnis der HECT-Ligasen und haben möglicherweise Auswirkungen auf die Entwicklung pharmazeutischer Wirkstoffe.

Abbreviations

1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
ATP	Adenosine triphosphate
CBR	Calcium binding region
CDS	Coding sequence
C-lobe	C-terminal lobe
CSP	Chemical shift perturbation
C-t	Carboxyl terminus
Cys (C)	Cysteine
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
DTT	Dithiothreitol
ECL	Enchanced chemiluminescence
EGTA	Ethylene glycol tetraacetic acid
ER	Endoplasmic reticulum
FD	Fast digest
FID	Free induction decay
FT	Fourier transformation
FL	Full length
Fw	Forward (primer)
GAT	GGA and Tom1 domain
Glu (E)	Glutamic acid
Gln (Q)	Glutamine
Gly (G)	Glycine
h	Hour (s)
HA	Hemagglutinin
His (H)	Histidine
HMQC	Heteronuclear multiple quantum coherence
HSQC	Heteronuclear single quantum coherence
Hz	Hetrz
IM	Isolucine/Methionine
IMAC	Immobilised metal affinity chromatography
IP_3	Inositol trisphosphate
IPTG	Isopropyl-d-1-thiogalactopyranoside
Ile (I)	Isoleucine
kDa	Kilo Dalton

Kd	Dissociation constant
Koff	Dissociation rate
Kon	Association rate
L	Ligand
LB	Luria Broth
Leu (L)	Leucine
Lys (K)	Lysine
Met (M)	Methionine
mM	Milimolar
min	Minute (s)
ms	Milisecond
µM	Micromolar
µs	Microsecond
nM	Nanomolar
NMR	Nuclear magnetic resonance
N-lobe	N-terminal lobe
N-t	Amino terminus
OD	Optical density
P	Protein
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PD10	Protein desalting column
Phe (F)	Phenylalanine
Pro (P)	Proline
PTM	Post-translational modification (s)
PVDF	Polyvinylidene fluoride
q.s.	Sufficient quantity
RE	Restriction enzime
RF	Radio frequency
Rv	Reverse (primer)
ROI	Region of interest (s)
SAP	Shrimp alkaline phosphatase
SDS-PAGE	Sodium dodecysulfate- polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
Ser (S)	Serine
SI(ngle)	Single HECT
TBS	Tris-buffered saline
Thr (T)	Threonine
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimized spectroscopy
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine

UBS	Ubiquitin binding surface
UPS	Ubiquitin-Proteasome system
Val (V)	Valine
WT	Wild type

Name of Domains and Proteins

A20	Ataxin3
AREL1	Apoptosis-resistant E3 ubiquitin protein ligase 1
BH3	Bcl-2 (B-cell lymphoma 2) homology 3
BMP	Bone morphogenetic proteins
Bsd2	Metal homeostatis protein BSD2
CUE	Coupling of ubiquitin to ER degradation
DUBs	Deubiquitylating enzymes
Dvl2	Dishevelled segment polarity protein 2
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
E6AP	E6-associated protein
EGFR	Epidermal growth factor receptor
Hh	Hedgehog
HACE1	HECT domain and ankyrin repeat-containing E3 ubiquitin-protein ligase 1
hHARI	Human homolog of ariadne
hHR23	Human homologues of the Rad23 proteins
HECT	Homologous to the E6-AP C terminus
HECW1	HECT, C2 and WW domain-containing protein 1
HECW2	HECT, C2 and WW domain-containing protein 1
HERC	HECT and RDL containing protein
HOIP	HOIL-1L interacting protein
HOIL-1L	Heme-oxidized IRP2 ubiquitin ligase 1
HRP	Horseradish peroxidase
HUWE1	HECT, UBA and WWE domain-containing protein 1
IBR	In between RINGs
Itch	E3 ubiquitin-protein ligase Itchy homolog
LMP-1	LIM mineralization protein-1
LUBAC	Linear ubiquitin chain assembly complex
MBP	Maltose-binding protein
MIU	Motif interacting with ubiquitin
NEMO	NF-kappa-B essential modulator
NDFIP1	Nedd4 family-interacting proteins 1
NDFIP2	Nedd4 family-interacting proteins 1
Nedd4	Neuronal precursor cell-expressed developmentally downregulated gene 4

Nedd4L	Nedd4-Like
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
Notch	Neurogenic locus notch homolog protein
NZF	Nlp4 Zinger Fing
NusA	Transcription termination/antitermination protein (N utilization substance A)
Parkin	Parkinson juvenile disease protein 2
PKC	Protein kinase C
cPLA2	Calcium-dependent phospholipase A2
PLCs	Phospholipase C
RAD18	E3 ubiquitin-protein ligase RAD18
Rap80	Receptor-associated protein 80
RBR	RING between RINGs
RDL	RCC1 (Regulator of chromosome condensation 1) like domain
RING	Really interesting new genes (s)
Rsp5	Reverses SPT-phenotype protein 5
SSQ	Heat shock protein SSQ1, mitochondrial
SHARPIN	SHANK (SH3 and multiple ankyrin repeat domains 3) associated RH domain interactor
Smad7	Mothers against decapentaplegic homolog 7
Smurf1	Smad ubiquitin regulatory factor 1
Smurf2	Smad ubiquitin regulatory factor 2
SPRY	SPla and the RYanodine receptor
SUMO	Small ubiquitin-like modifier
TAB2	TGF-beta-activated kinase 1 and MAP3K7-binding protein 2
TEV	Tobacco etch virus
TGF-β	Tumor grow factor β
TNFα	Tumor necrosis factor α
TRIP12	E3 ubiquitin protein ligase TRIP12
UIM	Ubiquitin-interacting motif
USP	Ubiquitin specific protease
Ub	Ubiquitin
UBA	Ubiquitin-associated
Ub^a	Acceptor ubiquitin
UBAN	Ubiquitin binding in ABIN and NEMO domain
Ub^d	Donor ubiquitin
UBD	Ubiquitin binding domain
UBE3C	Ubiquitin protein ligase E3C
UBM	Ubiquitin binding motif
UBL	Ubiquitin-like
UBP	Ubiquitin binding proteins
UBR5	E3 ubiquitin protein ligase UBR5
UBZ	ubiquitin-binding zinc finger

UEV	Ubiquitin E2 variant
VHS	Vps27, Hrs, and STAM
Wnt	Wingless-related integration site
WWP1	WW Domain containing E3 ubiquitin protein ligase 1
WWP2	WW Domain containing E3 ubiquitin protein ligase 2

1. General Introduction

Ubiquitination

1.1 Ubiquitin and its conjugation mechanism

Ubiquitination is a post-translational modification (PTM) which involves covalent attachment of a ubiquitin (Ub) molecule to a target protein. Ub is expressed from four distinct genes and translated as head to tail repeats, which are then cleaved into free Ub molecules (Heride et al., 2014). It is a 76 amino acids protein, highly conserved through eukaryotes, which adopts a compact β-grasp fold with a flexible carboxyl-terminus (C-t) tail. It possesses different hydrophobic surfaces, which allows its recognition by several proteins during the different process in which it is involved (Varshavsky, 2012).

Ubiquitination was first described as a mechanism to control the abundance of proteins in the cell by targeting ubiquitinated proteins for degradation. Nowadays, it is known to be implicated in almost all cellular functions including transcription, DNA repair, protein stability, signal transduction, protein trafficking, internalization and lysosomal targeting, alteration of sub-cellular distribution and cell cycle control (Weissman, 2001). The process can be described as an enzymatic cascade comprising three enzymes that act consecutively (Figure 1). First, a Ub-activating enzyme (E1), activates the C-t of a Ub moiety in an ATP-dependent manner and tethers it to a catalytic Cysteine (Cys) through a thioester bond. This E1-Ub intermediate binds to a Ub-conjugating enzyme (E2), which catalyzes the transfer of Ub to its active site Cys in a transesterification reaction. The E2 binds then to an Ub ligase (E3) and in the final step the Ub is transferred to a substrate. The result is the formation of an isopeptide linkage between the C-t Glycine (Gly) of Ub and the ε-amino group of the target Lysine (Lys) (Fanga & Weissman, 2004).

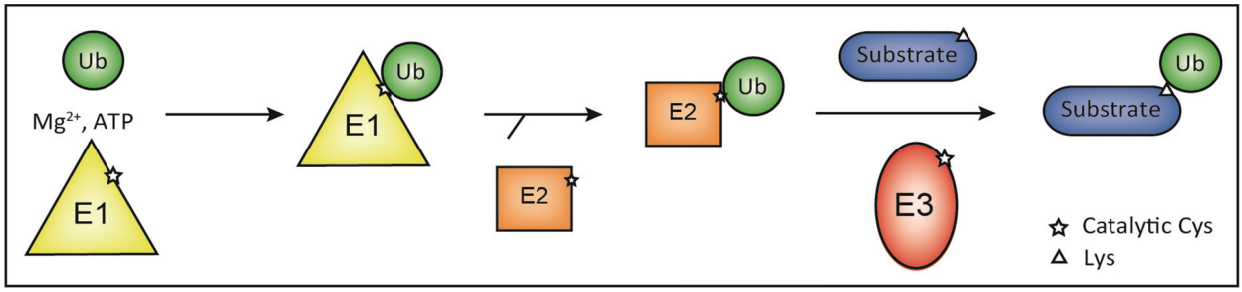


Figure 1. General overview of the ubiquitination cascade. The three enzymes E1, E2 and E3 act consecutively to generate a ubiquitinated substrate.

The outcome of the ubiquitination process is the generation of mono-, multi- or poly-ubiquitinated substrates. The modified substrates serve as signals that are interpreted by the cell in a dynamic process, meaning that they can be further modified by the addition of other PTMs, the shorting or extension of the chain, etc. Therefore, the ubiquitination

process as a whole consists of three steps known as writing, reading and erasing; the coordination of the three give rise to the first layer of complexity of the "ubiquitin code".

1.2 Ubiquitin code

Mono-ubiquitination is defined as the attachment of one Ub molecule to a specific substrate. In mammalian cell lines explains almost 60% of all conjugated Ub (Heride et al., 2014). Mono-ubiquitination is mainly involved in membrane trafficking, endocytosis and viral budding. Proteins can also be mono-ubiquitinated at multiple sites (multi-ubiquitination). A myriad of signals can be generated when the first Ub is further modified at any one of its seven Lys (Figure 2A) to form poly-ubiquitin chains (Figure 2B). Additionally, a Met1-linked or 'linear' chain can be generated if a new Ub molecule is linked to the amino terminus (N-t) of the already attached Ub (Komander & Rape, 2012).

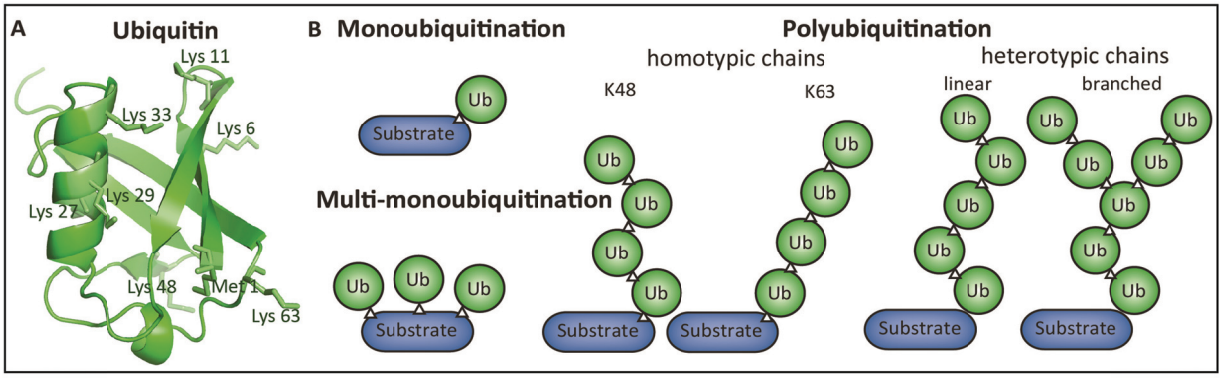


Figure 2. Ubiquitination outcome. **A)** Crystallographic structure of Ub, where the seven Lys residues and the N-t Met are shown (PDB ID:1UBQ, Vijay-Kumar et al., 1987). **B)** According to the linkage used and the amount of Ub attached the substrate can be mono-, multi- or poly-ubiquitinated. Poly-ubiquitin chains can be homotypic or heterotypic.

Lys48-linked Ub chains are the most common linkage type in eukaryotes and their main role is to target proteins to the proteasome for degradation. Their function was discovered early on since their levels increase rapidly when the proteasome is inhibited (Hershko A & Ciechanover, 1998). In contrast, Lys63 chains, which are second in terms of abundance, have non-degradative roles concerning the proteasome. They regulate the activation of the NF-κB transcription factor, DNA repair, innate immune responses, clearance of damaged mitochondria, and protein sorting. K63-linked chains can also accomplish their functions when they are not attached to substrate proteins, similar to second messengers (Swatek & Komander, 2016). Ub chains built using other Lys linkages are known as atypical, and their roles are not completely deciphered. For Met1-linked chains, it is known that they are exclusively assembled by the linear Ubiquitin chain assembly complex (LUBAC). They play essential roles in inflammatory and immune responses, by regulating

the activation of the transcription factor NF-κB. In the case of Lys11-linked chains, it has been established that they act as additional proteasomal degradation signal, used specifically in cell cycle regulation. Interestingly, M1- and K11-linked chains seem to be counterparts to the canonical ones; while both M1- and K63-linked conjugates arrange the assembly of protein complexes, K11- and K48-linked chains can drive proteasomal degradation. The other four types (Lys6, Lys27, Lys29 and Lys33) are even less studied and little information is available (Yau & Rape, 2016). Lys6 and Lys27 seem to be associated with DNA repair events, Parkin-mediated mitophagy and immune response. Lys29 is an inhibitor of the Wnt signaling pathway, while Lys33 has been implicated in post-Golgi protein trafficking (Akutsu et al., 2016) (Table 1).

Table 1. Biological function of an ubiquitinated protein according to the Ub-chain linkage.
UBD: Ubiquitin binding domain. The name of a representative protein that contains the mentioned domain is shown in brackets. -: no information available. * The name of domains and proteins are described in the "List of abbreviations" section.

Linkage Type	Signaling outcome	UBD (*)
mono-Ub	Membrane trafficking; endocytosis; viral budding	UBA, UIM, NZF CUE, MIU, VHS, GAT, UBZ, UBC, UEV, UBM
Lys6	DNA damage response; Parkin-mediated mitophagy	-
Lys11	Poly-ub chains: human cell cycle control/hypoxia	-
	Branched-chains: strong proteolytic degradation signal	-
Lys27	Nuclear translocation; DNA damage response	-
Lys29	Wtn/b-catenin signaling	-
Lys33	Post-golgi protein traffiking	-
Lys48	Canonical signal for proteasome degradation	UIM (Ataxin3)
		UBA (hHR23)
		S5a/Rpn13 (19S)
Lys63	Endocytosis; protein trafficking; innate immunity	UIM (Rap80)
		NZF (TAB2/A20)
		UBZ (RAD18)
		S5a/Rpn13 (19S)
Met-1	Inflammatory and immunity response; NF-kb signaling	UBAN (NEMO)

From a structural point of view it is possible to distinguish two conformations that the different types of chains can adopt. One is a compact conformation, in which each Ub moiety interacts with the next one in the chain (Lys6, Lys11, Lys33 and Lys48-linked). The other is an open conformation, where no binding surface between Ub is present, apart from the linkage site (Lys 29, Lys63 and Met1-linked chains) (Akutsu et al., 2016; Komander & Rape, 2012).

In addition to homotypic (one linkage type), heterotypic chains can also be formed (Figure 2). In this case, chains that were linked using one type of linkage are extended by a second type. Another option instead, is that the Ub molecule already attached is ubiquitinated at a different Lys residue, forming a so-called branched structure. Branching allows E3 ligases to increase the local Ub concentration explaining the strong proteolytic signal of branched chains (Yau & Rape, 2016). Detailed studies demonstrate that while homotypic Lys11-chains do not bind to the proteasome, heterotypic Lys11/Lys48-polyubiquitin chains induce potent degradation signals (Grice & Nathan, 2016). Early studies in the field defined the Lys48-linked tetra-ubiquitin as the canonical signal to trigger degradation (Thrower et al., 2000). However, new data suggest that chain branching could convert any non-degradative chain into a degradation signal. In fact, a "ubiquitination threshold" model has been proposed, in which instead of having a specific type of chain, the amount of poly-ubiquitin is a relevant factor to determine degradation (Swatek & Komander, 2016).
A second layer of complexity can be added to the system in the case that Ub is modified with other PTMs, for instance an UBL (Ubiquitin Like) modifier such as SUMO (Sriramachandran & Dohmen, 2014), or when small chemical modifications such as phosphorylation or acetylation are introduced (Ohtake et al., 2015; Wauer et al., 2015). All Ub moieties, even in complex topologies, could undergo these modifications generating an essentially unlimited number of combinations. Following Ub conjugation, the cell needs to reliably read and interpret each type of poly-ubiquitin chain and act in consequence. In order to do so, the immediate decoders are the Ub receptors, Ub binding proteins (UBPs) that contain one or several Ub binding domains (UBDs) (Grabbe et al., 2011) (Table 1). At least 20 different families of UBDs (e.g. UBA, UIM, NZF and CUE) present in hundreds of human proteins have been described and they often occur in tandem repeats. The majority of them recognize the canonical Ub hydrophobic patch centered on Ile44, although a small number can interact with other surfaces such as the Ile36 patch, the Phe4 patch, the Asp58 patch, the TEK box or a flexible loop (Kulathu & Komander, 2012). By recognizing a particular Ub surface, UBPs couple substrate ubiquitination to a downstream event. However, the mechanism underlying the recognition of each specific linkage is not fully understood. The vast majority of UBPs interact with mono-ubiquitinated proteins, with affinities in the μM range. The challenge gets tougher when they need to identify distinct poly-ubiquitin chains. One way to distinguish among chains is using the existing difference in distance between the Ub molecules in each chain type. Many proteins possess multiple UIMs domains with a defined spacer between them in order to do that. Another mechanism is to differentiate chain flexibility. NZF domains can differentiate Lys63- from Met1-linked chains despite their structural similarities; Lys63-linked chains can bend allowing a perpendicular way of interaction with the same NZF domain that is not possible for Met1-linked Ub chains. Additionally, the context of the linkage (e.g. the sequence in

the vicinity) as well as a combination of different patches in the molecules of the same chain can also be used for specific recognition. Finally, in the last couple of years new studies have supported the idea of "induced fit selection", where the binding of each UBD type would be able to change/select a particular Ub conformation (Husnjak & Dikic, 2012; Komander & Rape, 2012).

To prevent ubiquitination from being constitutively active, modifications can be reversed by a specific group of Cys proteases called Deubiquitinating enzymes (DUBs), which can cleave Ub isopeptide bond at the end or within the chain. Some are in charge of protecting Ub from being degraded, which is essential in order to maintain sufficient levels of free Ub inside the cell. Others disassemble chains independently of the linkage, but are rather substrate specific, such as Ubiquitin specific protease (USPs). On the other hand, several DUBs show specificity towards one linkage type. Finally, it is often observed that DUBs and E3s coordinate their "writing and erasing" functions by binding to each other (Heride et al., 2014).

1.3 Cellular roles of Ubiquitination

Clearly, the main task of the ubiquitination pathway is the proteasomal degradation. The elimination of unwanted proteins occurs via the conserved Ubiquitin-Proteasome system (UPS), where the ubiquitination enzymes are key players, since the fate of ubiquitinated substrates depends both on the length and the type of chain linkages. Another key player is the 26S proteasome, who consists of two main regions: the proteolytic 20S chamber and the 19S regulatory particle. After ubiquitination, the modified protein can be target to the proteasome. This can be done directly by the 19S regulatory particle subunits or otherwise it can be mediated by proteasomal shuttle factors. These proteins contribute with the recognition and transport of ubiquitinated targets which need to reach the proteasome from distant locations. They function is critical, since they prevent chain disassembly during the transfer, preserving the original signal. When proteins destined for degradation reached the entrance of the proteasome, de-ubiquitination takes place. As a result, the Ub molecule is recycled and the protein unfolded and translocated to the proteolytic chamber for its destruction (Grabbe et al., 2011; Grice & Nathan, 2016). On the other hand, mono-ubiquitination or Lys63- as well as Met1- linked chains results in the formation of non-proteolytic signals that are involved in roles such as: recruiting proteins to participate in particular signaling pathways, attract trafficking factors that change substrate localization, or to control substrate activity.

The way different E3s target substrates varies significantly. Initial studies have revealed that many E3 ligases use short linear sequences called degrons to localize their substrates. Degrons normally adopt extended conformations to facilitate the interaction. Moreover, PTMs of these sequences play a role in recognition. In addition, specific amino acids, exposed hydrophobic surfaces, misfolded chains or even other molecules such as sugars

can be used to identify substrates (Mizushima et al., 2007). Another way of substrate recruitment consists of direct or indirect substrate binding through regular protein-protein interaction domains. Finally, a different mechanism consist of recruiting the E3s to a specific location in the cell, where then the ligase is going to ubiquitinate any available target (Buetow & Huang, 2016).

1.4 E3 ligases

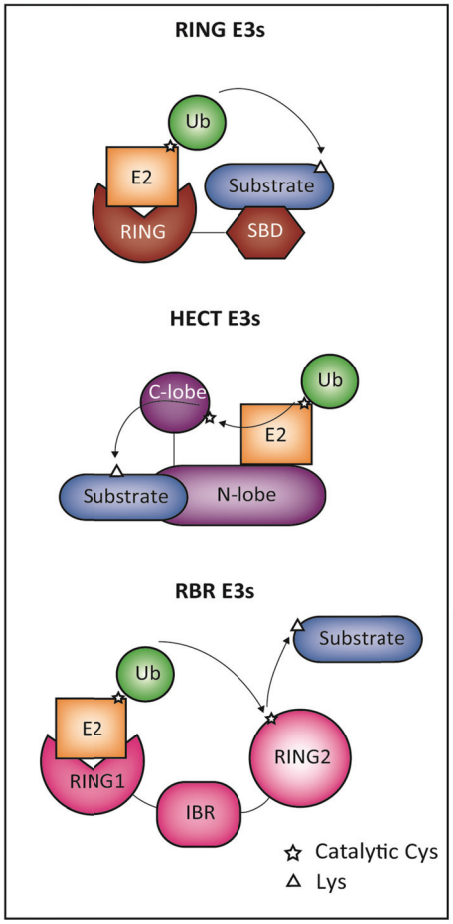


Figure 3. Mechanisms of E3 catalysis. Schematic representation of Ub transference from the E2 enzyme to the final substrate for the different classes. SBD: substrate binding domain. IBR: in between Rings.

The E3 ligases play a determinant role in the ubiquitination process not only because they define the specificity of the pathway by recognizing specific substrates, but also because their genetic alteration or abnormal expression can lead to several pathological disorders (Bernassola et al., 2008). In order to target the great deal of potential substrates inside the cell, the human genome encodes more than 600 E3s (Rotin & Kumar, 2009).

1.4.1 Classification of E3 ligases

Based on their structural features and their catalytic mechanisms, E3 ligases are grouped into three classes: Really Interesting New Genes (RINGs; also includes U-box E3s); Homologous to the E6AP (E6-associated protein) C-Terminus (HECT E3s); and RING Between RINGs (RBR E3s). Although all E3s carry out the final step of isopeptide formation, they differ in both structure and mechanism. While RING E3s catalyze the direct transfer of Ub from the E2 to the substrate by simultaneously binding to both of them, HECT and RBR E3s ubiquitinate substrates in a two-step reaction in which Ub is transferred from the E2 to an active site Cys in the E3, and then from the E3 to the substrate (**Figure 3**) (Metzger et al., 2012).

1.4.1.1 RING E3 ligases

In humans, RING E3s are the largest class of E3 ligases. They contain a RING domain, required for both E2~Ub (~:thioester) recruitment and stimulation of Ub transfer. Potentially, they can bear several additional domains, including the ones in charge of substrate recruitment. Studies showed that E2~Ub thioester intermediates are highly dynamic and adopt open conformations. When the RING E3 is present, the E2~Ub intermediate embraces a close conformation in which the Ub is

located in proximity of the RING. The RING E3 binds then both the E2 and the Ub molecule, which is still bound to the E2 through the thioester bond. This holds the C-t tail of Ub in a favorable conformation for catalysis. Basically, RING E3s work as scaffold proteins that facilitate the ubiquitination reaction by bringing together the E2~Ub and the substrate. For RING E3s, the outcome of the reaction is defined mainly by the E2 that participates in each particular reaction, whereas which Lys is accessible for modification is defined by the E3 (Berndsen & Wolberger, 2014, Buetow & Huang, 2016).

1.4.1.2 HECT E3 ligases

From the approximately 600 human E3s, only 28 belong to the HECT class. They are named after the conserved C-t of the E6AP HECT domain, the founding member of this E3 class. Based on the N-t region HECT E3s can be further divided into 3 groups: the Nedd4 family, the HERC family and other HECT or SI(ngle)-HECT E3s (Rotin & Kumar, 2009). HECT E3s catalyze two distinct reactions: the transference of an Ub molecule from the E2 active site to the catalytic Cys located in the HECT domain, forming a thioester intermediate, in a so-called transthiolation reaction. Subsequently, the HECT~Ub thioester is transferred to a Lys residue in the substrate forming an isopeptide bond. Of note, this Lys can also be located in the HECT domain, leading to auto-ubiquitination (Berndsen & Wolberger, 2014).

1.4.1.3 RBR-domain E3 ligases

To date, this group consists of 14 proteins in humans, from which PARKIN, HHARI and HOIP are the most studied. The mechanism they use to perform Ub ligation is different from the previous two classes, although it shares features with both of them. From a structural point of view, RBRs possess two RING domains (RING1 and RING2) linked by a conserved sequence called in between RINGs (IBR). The RING1 domain works similarly to a canonical RING domain, recruiting the E2~Ub. On the contrary, the RING2 domain resembles a HECT domain, since it bears a catalytic Cys able to form a thioester intermediate. Although the general mechanism is similar to HECT E3s, RBR substrate ubiquitination seems to be catalyzed differently. The best example is provided by the LUBAC complex, which is able to generate peptide-bond formation between the C-t of the donor Ub molecule and the Met1 of the acceptor one, generating linear Ub chains (Berndsen & Wolberger, 2014; Buetow & Huang, 2016).

1.4.2 The HECT family in detail

1.4.2.1 Catalytic mechanism

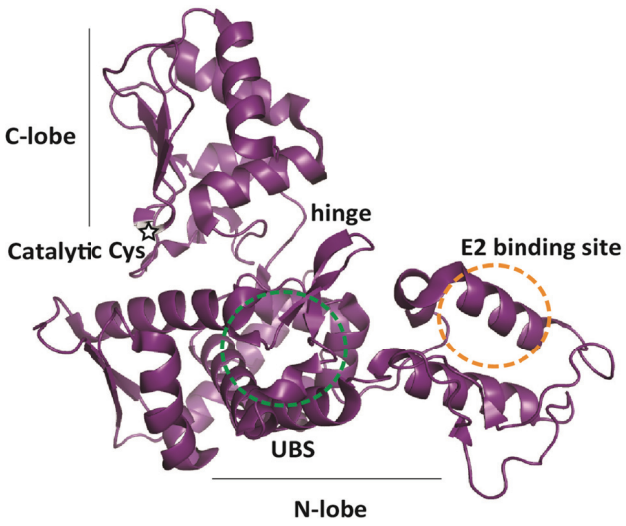


Figure 4. Main features of the HECT domain. As an example, the crystal structure of Smurf2 HECT domain (PDB: 1ZVD), showing the most relevant characteristics conserved through the family was chosen. UBS: Ubiquitin binding surface.

HECT domains consist of two lobes (an N- and a C-t lobe) that are connected by a flexible hinge loop (Figure 4). This linker allows for multiple relative conformations, what has been proved critical for the catalytic activity (Kamadurai et al., 2009; Maspero et al., 2011, 2013; Ogunjimi et al., 2010).

Briefly, the N-lobe interacts with the E2~Ub thioester (Figure 5A-B). Subsequently, Ub is transferred to the catalytic Cys on the C-lobe in a transthiolation reaction (Figure 5C). Finally, the HECT~Ub thioester is juxtaposed to a substrate Lys (Figure 5D), to which Ub is transferred (Figure 5E).

In more detail, several structures of HECT E3s show that the position of the C-lobe relative to the N-lobe can vary. In addition, due to the large distances observed between the E2´s and HECT domain´s catalytic Cys, it is highly probable that conformational changes are required in order to allow Ub transference from the E2 to the E3 (Figure 5A) (Huang et al., 1999; Ogunjimi et al., 2005; Verdecia et al., 2003).

The crystal structure of Nedd4L HECT domain in complex with E2~Ub thioester (Kamadurai et al., 2009) trapped the pre-transthiolation state (Figure 5B). The E2~Ub thioester is recruited by the N-lobe. Upon rotation around the hinge loop, the C-lobe binds the donor Ub (Ub^d) bringing together the two catalytic Cys (Figure 5B) promoting the formation of a HECT~Ub intermediate (Figure 5C). It is possible that a transthiolation mechanism mediated by rotation around the hinge loop is share by other HECT E3s as well. A crystal structure of the Nedd4 HECT-Ub^d complex showed that after transthiolation the Ub^d interface with the C-lobe is maintained (Figure 5C), but the C-t tail of Ub is locked in an extended conformation, ready for catalysis. Interestingly, RING and HECT E3s seems to conserve the same thioester-activating mechanism, since the constraining of the Ub C-t position appears in both of them (Zheng & Shabeck, 2017). Following transthiolation, the HECT~Ub thioester is placed close to the substrate Lys (Figure 5D).

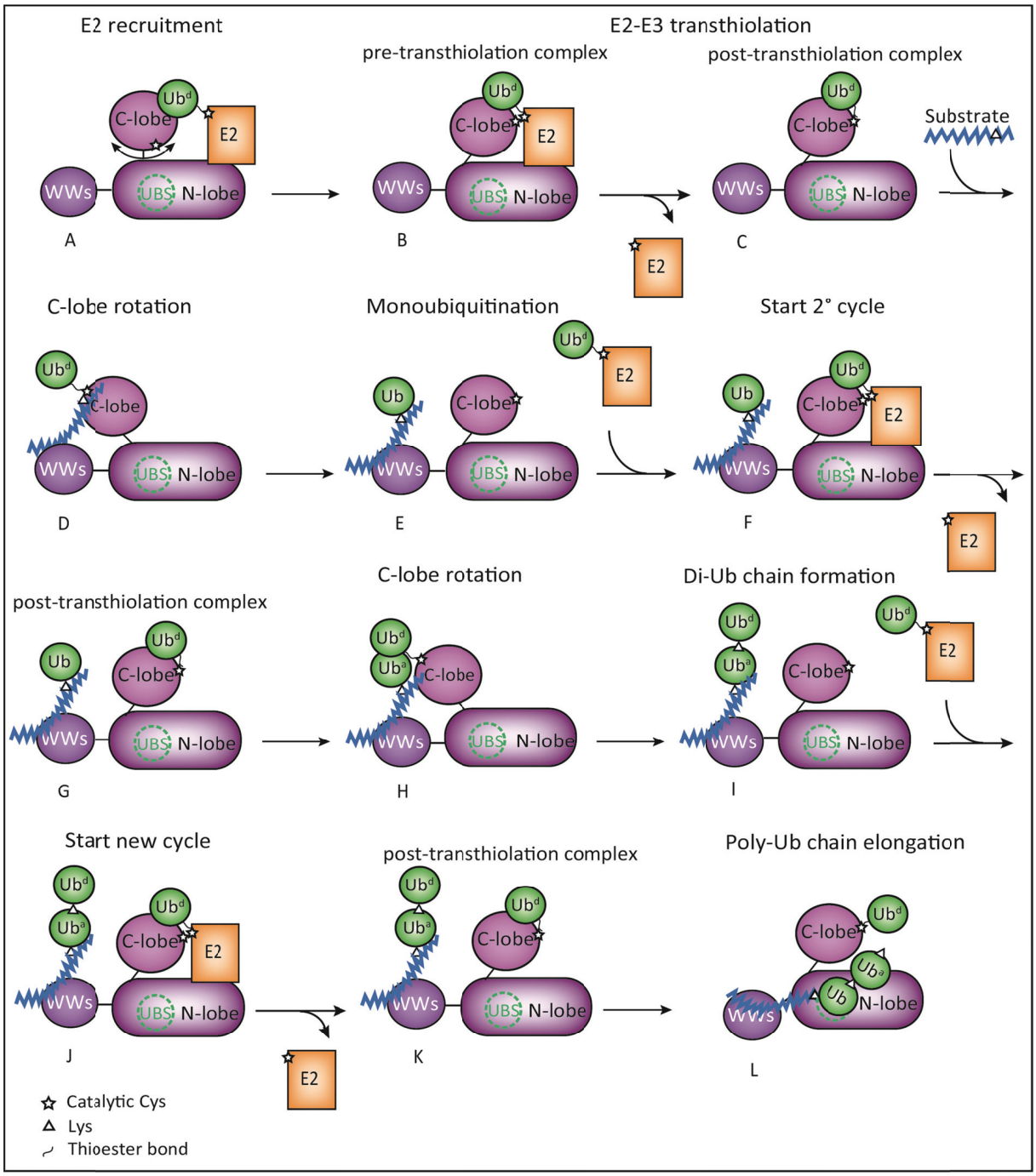


Figure 5. HECT Mechanism of ubiquitination. The C-lobe can rotate freely (**A**) until the N-lobe binds an E2~Ub thioester. In this moment, the C-lobe rotates, placing the catalytic Cys next to that of the E2 (**B**). The transthiolation process occurs: The Ub forms a thioester with the catalytic Cys in the C-lobe (**C**) and the E2 is freed. Next, the C-lobe rotates again to juxtapose the catalytic Cys of the HECT domain next to the substrate (**D**), to promote isopeptide bond formation between the Ub C-t and the ε-amino group of the target Lys. The substrate is mono-ubiquitinated (**E**). A new ubiquitination cycle can occur (**F-I**); a new Ub molecule is transferred from the E2 to the E3 and from it to a specific Lys in the Ub^a, which is now attached to the substrate (**I**). Finally, the UBS plays an important role in Ub chain elongation, probably by keeping the ubiquitinated substrate in close proximity to the HECT domain, increasing enzyme processivity (**L**).

In Nedd4, apart from the catalytic Cys (C867), a highly conserved Asp (D900) and a His located near the catalytic Cys (H865) are essential for substrate ubiquitination but not for transthiolation, helping in the positioning of the acceptor Ub (Ub^a) (Maspero et al., 2013). Previous reports also showed that a Phe at position -4 is needed for isopeptide formation (Salvat et al., 2004). In agreement with this, the Rsp5-Ub-substrate structure published by Kamadurai et al., 2013, clarified its function. The "-4 Phe" sits in an N-lobe pocket and mediates the inter-lobe contacts anchoring them in an orientation suitable for substrate ubiquitination. Of note, in this conformation the C-lobe of Rsp5 is rotated 130°C about the hinge loop in comparison with the Nedd4-Ub^d structure. This -4 Phe, conserved in all HECT E3s, has a clear role in substrate ubiquitination and poly-ubiquitin chain formation, but it is not required for the formation of the HECT~Ub intermediate. Moreover, the conservation of the N-lobe pocket among the Nedd4 proteins, suggests that the N- and C-lobe orientation required for Ub transfer is carried out by a similar mechanism in the entire subfamily (Buetow & Huang, 2016).

Additionally, the majority of the Nedd4 family members interact non-covalently with Ub through a surface called ubiquitin-binding site (UBS), which is located on the N-lobe of the HECT domain. The UBS is known to be essential for poly-ubiquitin chain elongation, but its mutation shows no effect on E2-E3 transthiolation or mono-ubiquitination. Since this UBS is remote from the thioester Ub in all available HECT structures, it probably cannot act as an Ub^a in Ub chain elongation, but rather serves as a site for binding ubiquitin-modified substrates, promoting processivity (Kamadurai et al., 2013; Maspero et al., 2011; Maspero et al., 2013; Ogunjimi et al., 2010). When Ub is ready for catalysis, the E3 ligase positions a substrate Lys next to the HECT~Ub thioester for ligation (Figure 5D). If the substrate is released, then the reaction results in mono-ubiquitination (Figure 5E). Instead, more Ub molecules can be attached to the first one, leading to poly-ubiquitination. This reaction can be divided in two stages: the initial transference of Ub to a substrate Lys called initiation, and the following rounds, where the Ub^d is attached to a Lys on the first Ub^a, called elongation (Figure 5F-I). Basically, the process of poly-Ub chain formation consists of a HECT E3 enzyme which faces an E2 to receive the Ub, after what turns around to pass it to the substrate. This sequential Ub addition mechanism is used by Nedd4-family E3s (Kim & Huibregtse, 2009), while preliminary results suggest that this may not be the case for the rest of HECT enzymes (Fajner et al., 2017). How the following Ubs are attached to the substrate it is still unknown. One possibility is that when the second Ub is already attached to the substrate (Figure 5J), the C-lobe does not need to rotate anymore (Figure 5K) since the UBS present in the N-lobe would help to maintain the growing Ub chain in proximity as well as to position the Ub^a in the appropriate orientation to elongate the chain (Figure 5L). This idea is supported by the fact that the UBS is essential for poly-ubiquitination, but not mono-ubiquitination (Maspero et al., 2011; Ogunjimi et al., 2010).

In the case of mono-ubiquitination, the E3 needs to select one particular Lys from the ones available in the substrate, which is many times achieved by defining a specific distance between the catalytic Cys and the ubiquitination site on the substrate. Alternatively, for multi-monoubiquitination and chain elongation, the E3 cannot use this strategy since this distance is varying while the chain is growing. In this case, a specific Lys from the Ub^a has to be selected for ligation of the Ub^d. To date, it is not known whether a specific motif indicates which would be the outcome of the process.

For HECT E3s, it is known that the linkage-specificity is a feature determined by the HECT domain, independently of the E2 used. The exchange of the last 60 residues of the Rsp5 C-lobe for the E6AP corresponding region has an effect on the chain specificity, and switches chain specificity from K63 linkage to K48 (Kim et al., 2011). This means that this part of the C-lobe is necessary to determine the linkage type. Moreover, Maspero et al., 2013 showed that just the substitution of the last three/four residues alter the type of the chains produced by Nedd4. Probably, the last amino acids of the HECT domain, together with the determinants in the C-lobe are what define chain specificity. However, further studies are necessary to understand this mechanism in a more detail.

1.4.2.2 Nedd4 family E3s

In humans, the Nedd4-family consist of nine members (Nedd4, Nedd4L, Itch, Smurf1, Smurf2, WWP1, WWP2, HECW1 and HECW2), which share a common domain architecture (Figure 6). Another relevant HECT E3 is scRsp5, the Nedd4 homologous protein in yeast. All of them possess an N-t C2 domain, 2–4 WW domains and the C-t catalytic HECT domain. The C2 domain (approx. 140 amino acids) is able to target proteins to membranes by phospholipid binding as well as to participate in protein–protein interactions; either in a Ca^{2+}-dependent or -independent manner (see chapter 4.3). The WW domains, named after the presence of two Trp residues, are small protein–protein interaction modules, usually around 40 amino acids. They interact mainly with Proline (pro) rich PPxY or LPxY motifs, being responsible for the interaction with substrates and adaptor proteins.

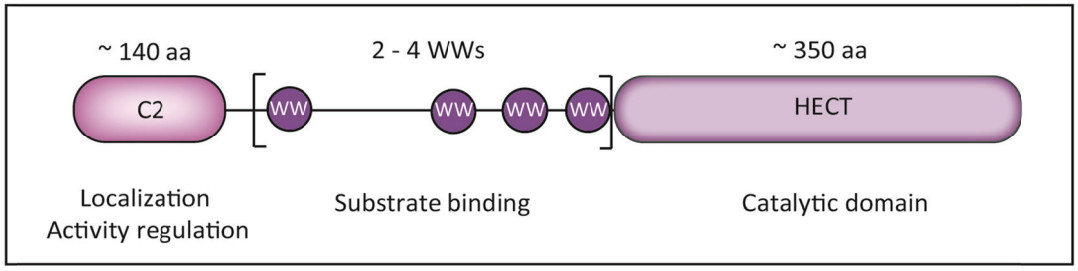


Figure 6. Domain organization of Nedd4-family members. All human nine members share the same organization. This particular example is drawn on scale, based on the founding member of the family Nedd4.

Nedd4 E3s present overlapping but distinct repertoires of substrates, and it seems that the specificity is determined by the difference in ligand affinity displayed by the WW domains. Even more, as they are present in tandem repeats, they can co-operate to fine-tuned WW-ligand interactions (Zou et al., 2015). Finally, the HECT domain, that consist of ~350 amino acids, which main feature is the ability to form an HECT~Ub thioester intermediate. By modulating the stability, localization and function of key players in diverse cellular pathways including TGF-β, TNFα, WNT, Notch, EGFR, p53/p73/p63, Hippo and Hh signaling, Nedd4 proteins define a multitude of essential cellular phenomena. They participate in endocytosis and trafficking of membrane proteins through mono-ubiquitination. In addition, they build poly-Ub chains mainly through K63 linkages, playing a role in the stability of transmembrane receptors as well as intracellular substrates. It has also been shown that they can play opposing roles as tumor promoting or suppressing factors according to different cellular contexts (Bernassola et al., 2008, Zou et al., 2015).
Nedd4 is the first member of the family to be described and it is named after the product of the neural precursor cell-expressed developmentally downregulated gene 4. It is localized in the cytoplasm and is expressed in most tissues. Nedd4 is involved in the regulation of cellular homeostasis by down-regulating the activity of epithelial Na^+ channel (ENaC) and other ion channels found in the brain, thereby affecting fundamental processes such as embryonic development and animal growth (Boase & Kumar, 2015).
There are two Smurf (Smad Ubiquitin Regulatory Factors) proteins in mammalian cells, called Smurf1 and Smurf2. They were originally described as negative regulators of the Smad (receptor-regulated mothers against decapentaplegic) proteins. They also down-regulate TGF-β and BMP signaling pathways, resulting in the degradation of critical players within those pathways. They are also involved in osteoblast differentiation, cell motility and polarity regulation.
Smurf1 and Smurf2 seem to have dissimilar molecular functions, despite the fact that they display 70% of sequence identity and some redundancy in their substrate repertoire. Smurf1 has been frequently associated with apoptosis and cancer metastasis. For Smurf2, it was shown that its deletion results in alterations in gene expression, DNA damage response, and genomic integrity. On the contrary, it also has a role as a tumor suppressor (David et al., 2013).

1.4.2.3 Other HECT families

The other two HECT families do not show a conserved domain organization N-t to the HECT domain as is the case of the Nedd4 family. The HERC family, which has 6 members in humans, is defined by the presence of RLDs domains (Regulator of chromosome condensation 1 (RCC1) like domains). The small members (HERC3-6, approx. 100 kDa) have only one RLD domain while the large ones (HERC1-2, approx. 500 kDa) contain

multiples RLDs. In addition, they can bear extra domains such as SPRY and WD40 domains. RLDs are characterized by seven repeats of approx. 60 amino acids that adopt a seven-bladed β-propeller fold. From a functional point of view they work both as a guanine nucleotide-exchange factor for small GTPase as well as in the interaction with chromatin through histones H2A and H2AB (Fajner et al., 2017). Although they have been discovered more than 20 years ago, it is still not clear whether all of them act as real E3 ligases, and only a few substrates are known (Sánchez-Tena et al., 2016).

The SI(ngle)-family (13 members in humans) is composed of all the HECT E3s which possess neither a WW nor a RLD domain. Some of the most relevant members are E6AP, HUWE1, HACE1, TRIP12 and UBR5, all of which play important roles in human physiology and in human diseases, such as cancer. Little is known about their structure and catalytic mechanism, with few exceptions. E6AP, whose HECT structure was the first to be solved, bears no other domain apart from the HECT, and it is able to assemble Lys48 linkages exclusively. HUWE1 is an extremely large protein of about 500 kDa, which has many extra domains like a UBA and a BH3 domain. Its HECT domain structure is also solved, and mass spectrometry analysis shows that it assembles K11 linkages (37%), K48 linkages (33%) and K6 linkages (26%) (Michel et al., 2017). Since this family includes diverse E3 proteins, each of them fulfill different roles through several types of Ub-chains: UBE3C was implicated in the assembly of Lys29 linkages and Lys48 linkages; UBR5 was suggested to ubiquitinate β-catenin with Lys29- or Lys11-linked ubiquitin chains; AREL1 produces K33 linkages as well as K11 and the ankryin repeat containing E3 HACE1 assemble Lys27-linked chains (Swatek and Komander, 2016; Yau and Rape, 2016). Importantly, the majority of the analysis used to determine linkage types are done *in vitro* and in the absence of *bona fide* substrates. It has to be taken into account that the presence of natural substrates might change the preference that the HECT enzymes show regarding linkage assembly (Kristariyanto et al., 2015).

1.4.3 E3 Activity regulation

Due to the large amount of processes in which it is involved and the numerous substrates it can modify, ubiquitination is considered nowadays of utmost importance (Zou et al., 2015). Relying on the proteasome/lysosome system it can virtually modulate every process in the cell, having also the ability of shutting down complete signal transduction networks. E3s define the specificity of the reaction by selecting the substrates that need to be modified. Due to these reasons, it is vital for the cell to regulate the activity of these enzymes, and different ways have been described how this is achieved in the cell.

Current evidence indicates that within the cell many E3s are kept in an inhibited state, which is mainly achieved by intramolecular interactions that restrict protein flexibility, impairing essential processes such as E2~Ub thioester recruitment (Buetow & Huang, 2016). This is particularly true for the Nedd4 family of HECT E3s (Fajner et al., 2017).

Interestingly, they have evolved slightly different mechanisms to carry out activation/ inactivation, despite that they all exhibited the same domain architecture. The details of auto-inhibition were first shown for Smurf2; the enzyme´s inhibition is mediated through an C2:HECT domain intramolecular interaction which takes place near the catalytic Cys, interfering with Ub thioester formation (Wiesner et al., 2007). Nowadays it is know that other members of Nedd4 are also kept in an auto-inhibited conformation (Chen et al., 2017; Courivaud et al., 2015; Mari*, Ruetalo* et al., 2014; Mund and Pelham, 2009; Wang et al. 2010; Wiesner et al., 2007; Zhu et al., 2017), while for Smurf1 there is still contradictory data regarding whether it is regulated by auto-inhibition or not (see 4.2.2). Subsequently, a mechanism that enables activation must take place. Smurf2 is activated by binding of the adaptor protein Smad7 to the WW domains, which releases the auto-inhibitory state (Ogunjimi et al., 2005), promoting Ub transfer activity. A different way of controlling E3 activity is phosphorylation, as is the case of Itch. Both Nedd4 as well as Nedd4L, are activated by Ca^{2+} binding to the C2 domain (Escobedo et al., 2014; Wang et al., 2010), which disrupts the interaction between the HECT and the C2 and target the protein to the plasma membrane (see 4.3.2).

E3 regulation involves a great variety of mechanisms that are not mutually-exclusive, since one ligase can be regulated by many of them. Those include other PTMs, such as the related Neddylation pathway but also other proteins, small molecules and non-protein ligands can serve as non-covalent regulators.

Considering the relevance of E3 ligases not only in physiological but also in pathological events, it is indispensable to reveal the molecular mechanisms by which they regulate their activity, including possible roles that other proteins or even small modifications can provide, in order to understand better the basis of diseases and search for promising treatments (Vittal et al., 2015).

NMR spectroscopy

2.1 Brief summary of NMR principles

Nuclear magnetic resonance (NMR) is a spectroscopic technique which takes advantage of the magnetic properties of particular nuclei to obtain detailed information about the structure and dynamics of molecules. In protein NMR, the most frequently used nuclei are ^{1}H, ^{13}C and ^{15}N, all of which possess a nuclear spin of ½, meaning that in a static magnetic field they have two energy levels, a high and low energy state. In a magnetic field, the nuclear spins precess around the direction of the static magnetic field at specific frequencies, the so-called Larmor frequency, giving rise to a bulk magnetization that is parallel to the direction of the magnetic field. Spin state transitions are achieved and by applying energy, i.e. by the absorption of a photon that matches exactly the energy difference between the two states that is related to the Larmor Frequency. This energy

transfer happens at a wavelength in the radio frequency (RF) range and therefore in NMR experiments, appropriate RF pulses are applied to the sample. An appropriate RF pulse thus excites all spins in the sample such that they precess with the same phase (coherence). In this way, the bulk magnetization is rotated to the plane that is perpendicular to the direction of the magnetic field and the precession of the spins is recorded as a function of time by inducing a detectable current in the receiver coil of the probe head. However, due to interactions with the environment and other spins the NMR signal decays over time and is thus called Free induction decay (FID). *Via* Fourier transformation (FT) the FID is converted into a frequency domain spectrum that can be used for analysis (Figure 7).

Figure 7. Primary signal of an NMR experiment. The time domain of the FID is converted into a frequency domain, resulting in a useful NMR spectrum. This conversion is completed by the mathematical process known as Fourier Transformation (FT). This figure is adapted from Keeler, 2002.

The frequency of resonance is determined by the effective magnetic field and electron shielding. As a result, using the resonance frequency, information about the nucleus´ chemical environment can be obtained. Since the chemical environment of nuclei in a protein is influenced by solvent exposure, secondary structure, aromatic ring currents, bond torsion angles, hydrogen bonds, etc., every molecule display its own set of frequencies, resulting in unique NMR spectra.

The nature of the excited nucleus, the selected pulse sequence and the magnetization pathway define the NMR experiment. The different types of 2D NMR experiments can be classified as: homonuclear, where magnetization transfer occurs between nuclei of the same type through bond or through space, or heteronuclear, where magnetization transfer occurs between different nuclei with one-bond correlation (HSQC or HMQC) or long-range correlation. In this thesis, I mainly used HMQC (Heteronuclear Multiple-Quantum Coherence) spectra and HSQC (Heteronculear Single Quantum Coherence) in some cases.

Another important phenomenon that governs an NMR experiment is the relaxation of the NMR signal. Relaxation refers to the process by which the spins return to equilibrium mainly as a consequence of the interaction with the environment. Returning to the ground

state is mediated by two types of relaxation: the transverse relaxation, due to the loss of coherence among nuclei that rotate at their own frequency, and spin-lattice (or longitudinal) relaxation due to the transference of energy from spins to the surrounding. Both types of relaxation influence the quality of the signal. The slower tumbling of larger proteins enhances the FID and gives rise to peak broadening. In this context, the size of the protein is a main limitation in NMR experiments, since larger proteins induce faster relaxation. However, different approaches are available to slow down the relaxation and extend NMR applications to larger proteins.

2.2 NMR strategies to study large proteins

There are two major problems associated with the analysis of large proteins by NMR: signal overlap and peak broadening. The first arise as a result of the numerous active nuclei in the sample, all of which will be observed in a limited spectral range; the spectrum gets overcrowded rendering the analysis virtually impossible. The second one is a consequence of the relaxation of transverse magnetization. As the size of proteins increases, the tumbling rates of biomolecules increase resulting in more abundant spin-spin interactions which in turn, generates faster relaxation of the NMR signal. Ultimately, this translates into spectra of poor sensitivity and resolution.

To overcome these limitations a number of strategies were developed, including specific isotope labeling schemes as well as particular NMR pulse sequences. Deuteration of biomolecules is one of the basic approaches (Sattler and Fesik, 1996). Here, the dipole-dipole interaction can be diminished by replacing protons (^{1}H) with deuterons (^{2}H). Since dipolar interactions are proportional to the square of their gyromagnetic ratios and the gyromagnetic ratio of ^{2}H is 6.5 times lower than the one of the proton, the benefit of the exchange is significant. A second advantage of deuteration, which should also be introduced in binding partners and by using heavy water (D_2O) as solvent, is the suppression of spin diffusion that takes place between protons from the protein and protons from the solvent. Heteronuclear experiments done with uniformly labeled proteins (^{15}N and ^{13}C isotopes) increase the limit to 25 kDa proteins. Combined with this isotopic scheme, deuteration expands the use of NMR up to 50 kDa (Cvetković & Sprangers, 2017).

The second approach consists of an optimized NMR pulse sequences which helps dealing with the slow tumbling of the high molecular weight proteins called transverse relaxation optimized spectroscopy (TROSY) (Pervushin et al., 1997), which opened the NMR field to study proteins with a molecular weight of up to 100 kDa. The intensity of the NMR signal in a particular experiment depends on the amount of initial magnetization as well as on the decay rate of this magnetization. While recording a spectrum, several magnetization terms are created which decay with different relaxation rates. The TROSY sequence is able to separate the coherences that relax slowly (sharp lines) form the ones that relax fast (broad lines) and to select only the slowly relaxing ones. As a result spectra of substantially

improved quality are obtained, in comparison to HSQC experiments, where both fast and slow relaxing coherence are merged. Although these two strategies have been extremely useful to study large proteins by NMR, it often happens that the negative effects of complexity still affect the recording of ^{1}H-^{15}N-based spectra.

2.3 The Methyl TROSY Experiment

The combination of ^{1}H,^{13}C-labeled methyl groups in an otherwise fully deuterated protein, together with the use of the TROSY pulsing sequence, is currently one of the best approaches for studying protein complexes of high molecular weight. This is also true for proteins of intermediate size, if the quality of the ^{1}H-^{15}N-TROSY is poor and/or their long-term stability not enough to assign H-N-bases spectra, as it is the case for the HECT domains used in this thesis.

The advantage of using isotopically labeled-methyl groups comes from their excellent relaxation properties which are defined by its three protons arranged symmetrically and the fast rotation around its three-fold symmetry axis. This leads to highly sensitive and well-resolved NMR signals (Wiesner & Sprangers, 2015). In addition, methyl groups are a good choice to study protein interactions. The risk of NMR invisible regions is low since methyl groups are evenly distributed in protein structures, being usually located in the hydrophobic interior of proteins and along binding surfaces. Also, since they are highly sensitive to changes in their chemical environment, subtle events such as side-chain rearrangements inside the binding surface will lead to detectable chemical shift changes.

2.3.1 Labeling Schemes

An ideal isotopic labeling scheme should consider abundance and distribution of the NMR active nuclei in the protein, in order to obtain the data of interest without overcrowding the spectra. To perform a methyl-TROSY experiment ^{1}H,^{13}C-labeled methyl groups are introduced in particular amino acids, while the rest of the protein should be uniformly deuterated. To deuterate the sample, the bacterial culture is grown in D_2O-based minimal medium with the addition of deuterated glucose. The introduction of $^{13}CH_3$ labeled groups can be done in single residues or in combinations of them, depending on the metabolic pathways of each amino acid. In some cases, it is possible to add a methyl labeled biosynthetic precursor before overexpression induction. However, if the precursor is involved in the metabolic route of another amino acid, this will lead to labeling of more than one amino acid type. Several methyl-labeled precursors and/or amino acids are now available allowing distinct combinations. In this thesis, the introduction of Ile labeled at δ_1 position was done by adding the precursor 2-ketobutyrate to the media. On the other hand, as the biosynthesis of Val (Valine) and Leu (Leucine) is connected, they methyl groups were labeled simultaneously by using 2-keto-3-isovalerate as a precursor. Finally, as the pathways of Met and Ala are also involved in the production of other molecules,

they were added before induction in their final forms Met-ε [$^{13}CH_3$], Ala-β [$^{13}CH_3$] (Kerfah et al., 2015).

2.3.2 Methyl Group Resonance Assignments

To perform a detailed NMR analysis it is necessary to have assigned spectra, that means, knowing which resonance or peak in the spectra belongs to which residue in the protein. Traditionally, the collection of data necessary to assign a spectrum, requires to record different types of NMR experiments, where the protein has to be highly concentrated and stable for a couple of days at the chosen temperature. Usually, this is restricted just to small-medium sized proteins.

In the case of methyl resonances, they can be assigned conventionally with H,N-based 3D spectra for proteins up to ~25 kDa proteins. For larger proteins there are two main strategies: the "divide and conquer" and the mutagenesis strategy. In the first one, since usually small blocks show better spectral quality than the full length (FL) protein, the domains are assigned independently and then the individual assignments are transferred to the whole spectra. Alternatively, each labeled methyl residue can be individually mutated to an NMR-invisible one. In this way, a direct comparison with the wild type (WT) spectra will show which peak is the missing residue, allowing its identification (Cvetković & Sprangers, 2017).

2.4 NMR applications

2.4.1 Protein-protein interaction studies

NMR spectroscopy is a highly sensitive tool to study biomolecular interactions since the resonance frequencies (chemical shifts) of the observed atomic nuclei depend on the local chemical environment. Ligand binding causes changes in the local chemical environment of those amino acids that are located in the binding pocket resulting in observable chemical shift perturbations (CSPs) for those residues. In this regard, NMR is unique since it can be used to study strong binding events with nM affinities up to weak interactions with dissociation constants in the mM range, providing detailed information about the binding surfaces of proteins and complexes in solution.

CSP mapping is a simple and widely used experimental technique for studying protein interactions and identifying binding interfaces (Williamson, 2013). A standard titration experiment consists of an isotopically labeled protein and an unlabeled ligand, which can be a small molecule or another protein. During the titration experiment the unlabeled ligand is added in consecutive steps to the NMR visible protein, modifying the local chemical environment of those amino acids that constitute the binding surface. Each titration step is followed by the acquisition of a new spectrum. When the interaction under study is between two proteins, experiments where each of them is labeled in turns can be done, in order to obtain information from both sides. CSPs can be measured rigorously;

therefore almost any authentic binding event will be detectable by this technique. After recording the different titration points, the analysis involves measuring how each peak moves throughout the titration. The magnitude of the CSP depends on the proximity of the residue to the ligand but also on local effects (e.g. exposure to aromatic residues in the ligand that induce ring current shifts). The residues which are not involved are not going to be affected by the presence of the ligand and, therefore, will not display CSPs (Figure 8). Of notice, it is important to remember that CSPs can also appear as a consequence of indirect effects, like conformational rearrangements, allosteric interactions or changes in conformational dynamics (Göbl et al., 2014).

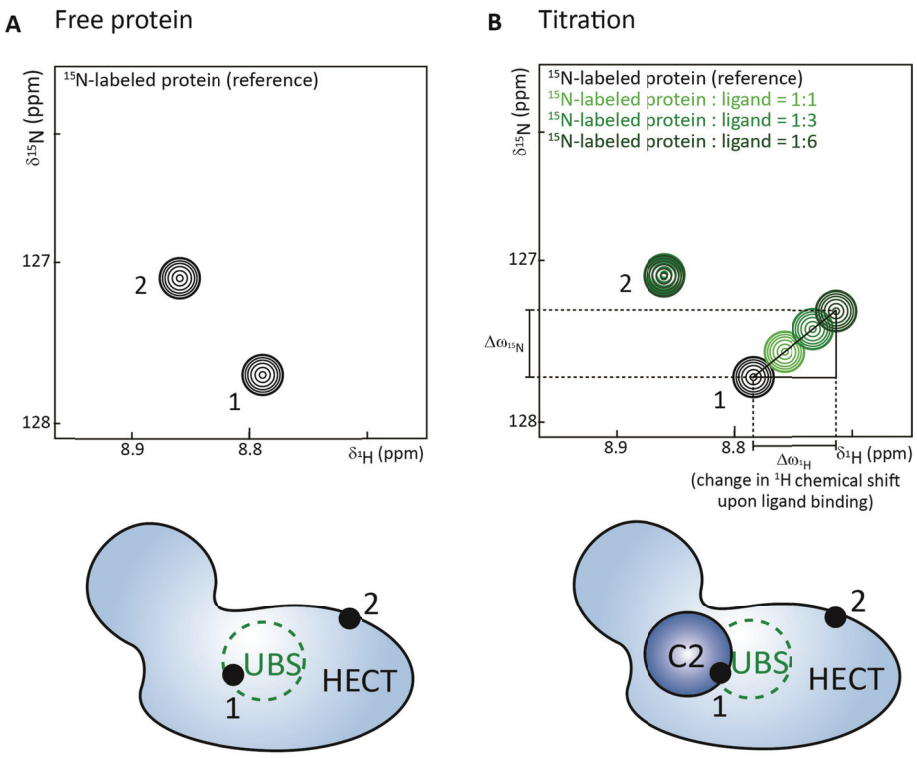


Figure 8. Basics of an NMR titration experiment. A) Before ligand addition a spectrum of the free labeled protein is recorded, which is referred as the reference spectrum. B) Upon ligand addition a binding event might happen. In this case, the chemical environment of the labeled protein will be altered. The residues located inside the binding site (1) will sense the change, leading to an observable CSP in the spectrum. A residue located outside the binding pocket (2) would not be affected and therefore its position in the titration spectra will remain the same as in the reference spectrum.

The CSPs observed during the titration provide information that allows us to characterize the interaction from a structural as well as from a thermodynamic and kinetic point of view. Interestingly, both aspects can be studied from the same data, which makes this technique powerful. Spectral assignments are needed in order to map the binding surface onto the structure. However, they are not necessary when only K_d values (dissociation

constants) for the estimation of the affinity are wanted. A typical set up involves the use ^{15}N-HSQC spectra since it is the easiest to assign, it is sensitive, and detects only one signal per amino acid (with a few exceptions). Nonetheless, ^{13}C-HSQC or HMQC spectra can be also used with the advantage that the shifts observed are less influenced by structural changes of the protein than for ^{15}N, being better indicators of the effects induced upon binding. To study larger proteins, ^{15}N-TROSY spectra are normally weak and crowded. A good option then is to use $^{13}CH_3$ labeling, since methyl side-chains are known to be involved in biomolecular interactions and are frequently located in interaction binding surfaces. It is possible though, that not enough methyl-containing residues are present in the system or region of interest. The "methionine scanning approach" can be used to overcome this problem (Stoffregen et al., 2012). Briefly, it involves the introduction of methyl-labeled Methionine (Met) instead of solvent exposed residues, one-at-a-time. The newly introduced methyl group is easily assigned by comparison of the WT and mutant spectra and serves as a reporter for ligand binding. Subsequently, upon ligand addition, a new spectrum is recorded. According to the position of the introduced Met three different possibilities arise: If the reporter Met is part of the binding interface its chemical environment will be affected, producing a CSP. If the reporter is located outside the binding pocket, then no correspondent CSP will be observed. Finally, if the residue mutated is essential for the interaction (a so-called hotspot) the binding will be disrupted; this situation is recognized since not only no CSP for the new resonance will be detected but also the naturally occurring methyl groups will display no CSP.

2.4.2 Chemical shift exchange

The dynamic component of proteins has proven to be fundamental for relevant biological processes such as enzyme catalysis and allosteric regulation. Indeed, molecular recognition inherently involves equilibrium of protein association and dissociation, and is a kinetic process that typically occurs in the microsecond to second time scale (Göbl et al., 2014). The reversible one-site binding of a protein (P) to its ligand (L) is characterized by two rate constants (k_{on} and k_{off}) for the association and the dissociation reaction, and the dissociation constant (K_d). K_d is an indicator of binding affinity and is equal to k_{off}/k_{on}, which in turn equals [P][L]/[PL], where [P], [L] and [PL] are the concentrations of the protein, the ligand and the complex formed by P and L. In a CSP experiment, the chemical shifts of the free and the bound protein are observed, with the chemical shift difference $\Delta\omega$ being the difference between these two resonances. The K_d is related to the exchange rate between the two states (K_{ex}).

Depending on K_{ex} of the interaction three different NMR time scales can be defined (Figure 9). One possibility is that the complex dissociates very slow, implying that K_{ex} is considerably smaller than the difference in chemical shifts between the free and the bound state ($K_{ex} << \Delta\omega$). This is reflected in the NMR spectra as two set of resonances, one for P

and one for PL (Figure 9A). Throughout the titration experiment, the peaks corresponding to the free protein are going to disappear (signal loss), and resonances for, the bound state arise as the population of the bound state increases (signal increase). These features define a regime which is known as slow chemical exchange. Usually the interactions are strong, with characteristic K_d values in the nM range.

If $k_{ex} \sim \Delta\omega$, the interaction is in the intermediate exchange regime (Figure 9B). Here the peaks move throughout the experiment, but they are broadened at intermediate titration steps and re-appear towards saturation. In some cases the peaks can become broad enough to disappear completely (Zuiderweg, 2002).

Finally, if $k_{ex} >> \Delta\omega$, the complex dissociates very fast. In this case, upon binding, only one peak per residue is observed, which displays chemical shifts changes according to the fraction of free protein present at each step. In addition, the signal is going to move continuously in the course of the titration until saturation is reached. This regime is known as fast chemical exchange and is depicted in Figure 9C. Also, it is characteristic of systems where the interactions are weak, with K_d values ranging from mM to high µM.

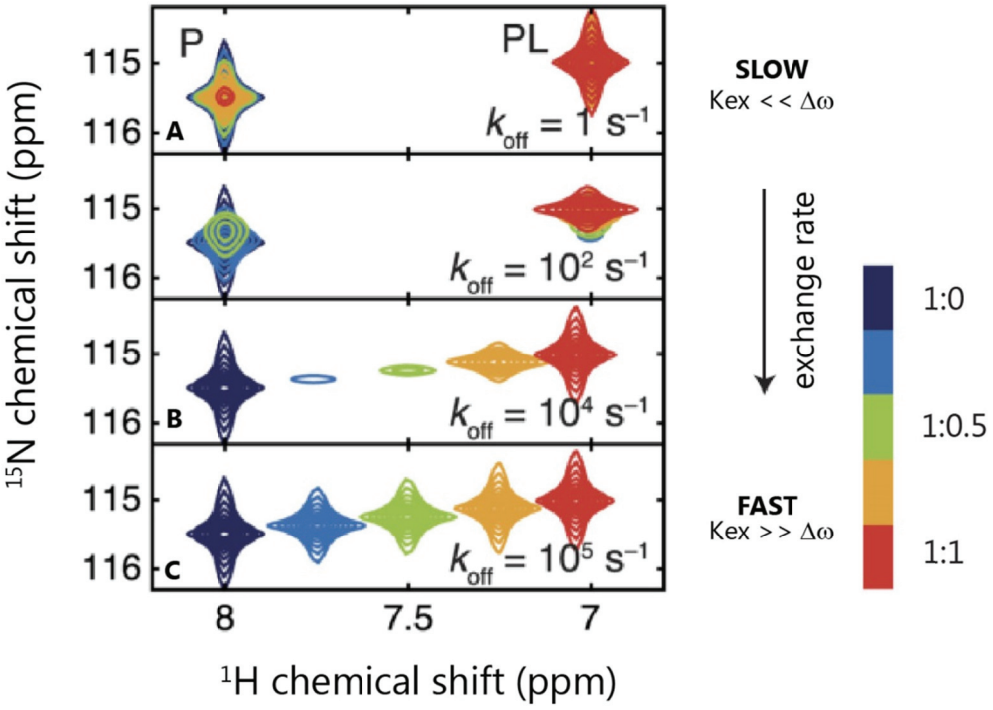


Figure 9. Chemical shift exchange regimes. Simulated titration of ^{1}H, ^{15}N-HSQC spectra for a protein-ligand interaction illustrating two-dimensional lineshape that may arise under various exchange regimes, from slow (A) to fast exchange (C). The color bar indicates the titration steps. P refers to the initial and PL to the final step of the titration. Kex: exchange rate. Δω: chemical shift. This figure was adapted from Waudby et al., 2016.

The identification of ligand binding sites from the titration data can be done almost independently of the exchange regime to which the studied system is subjected. However,

a complete analysis of a titration experiment is simplified when the system is in fast exchange, since following the signal of a peak from the free to the bound state is usually straightforward. In addition, traditional fitting of kinetic constants (such as K_d) using equations which depend on P and L concentrations is easily feasible. An accurate estimate of the K_d is achieved if many titration curves are fitted simultaneously. Moreover, the protein concentration should be close to the $0.5K_d$, ranging from $K_d/5$ to $5K_d$. The values obtained are reliable from approximately 1 µM up to 10 mM, "which is also close to the upper limit for biologically relevant affinities" (Vaynberg and Qin, 2006). Ideally, the concentration of the ligand should reach full saturation, but unfortunately this is not possible in many cases.

On the contrary, in the case of slow exchange a re-assignment of chemical shifts in the bound state is normally required, since it is impossible to recognize which signal in the reference spectra belongs to the signal at the end of the process. By measuring the intensity of the peaks that disappear and re-appear the binding constant can be calculated, although line broadening makes this process harder. In the intermediate exchange, although the peaks shift, the $\Delta\omega$ is less than what is predicted based on the simple weighted average. Generally, when the fitted K_d becomes stronger than about 10 µM the system is leaving the fast exchange.

An important consideration is that although signal broadening is a typical feature of the intermediate/slow exchange regime, it can also happen in fast exchange due to other reasons. One is slow tumbling, what happens when the ligand is of high molecular weight or if it leads to oligomerization/aggregation. In general, in this case all the signals of the labeled protein will broaden and never recover. The other reason could be exchange, which can be explained by different factors including that the protein undergoes a pre-equilibrium conformational change before binding, a structural rearrangement after complex formation or protein dimerization before or after binding. In all these cases, even considering that the binding of the ligand is fast, broadening is observed. In practical terms, when the calculated K_d value is lower than 10 µM and the regime is in fast exchange, judging from limited line broadening, the fitted values should be close to the true value (Williamson, 2013). In any case, the level of accuracy in the determination of affinities constants can be significantly increased by using lineshape fitting analysis.

2.4.2 2D line shape fitting

NMR line shape analysis is a well-established method for the quantitative analysis of titration data based upon the fitting of one-dimensional spectra, which can be used to study processes on timescales from 10 µs to 100 ms (Rao, 1989).

The analysis can be extended to two-dimensional experiments, e.g. ^{15}N-HSQC or HMQC experiments, for which a couple of additional features had to be taken into account. In the first place, each dimension can experience different chemical shift differences and hence

can be subject to different exchange regimes. A new software called TITAN (Waudby et al., 2016), able to perform 2D lineshape fitting, has several advantages over the regular 1D methods, regarding convenience and accuracy. Performing 2D spectra direct analysis, peak overlap is avoided, which is a general problem for biomolecules. The program allows selecting ROIs (region of interest) where one or more peaks display CSPs, and doing a global fitting, a robust tool for monitoring the quality of each fit. Another major benefit is that it tracks the relaxation of magnetization over the entire pulse sequence and performs calculations to take that into account. This is especially relevant when the system is in slow-intermediate exchange regimes, since distinct conformations of the protein caused by folding/unfolding reactions, dimerization and other association/dissociation reactions do not show the same line width. This can lead to the introduction of errors in the case of using 1D analysis methods. Finally, it is essential for a lineshape analysis to obtain reliable error values associated with the estimated parameters, for which TITAN performs a bootstrap error analysis method based on resampling of fitting residuals in two-dimensional blocks. A methyl-HMQC version of the program is available, which was used in this thesis.

2. Aims and significance of the project

The main aim of this thesis was to contribute to the Ub field by unraveling the mechanisms of HECT-mediated ubiquitination and its regulation on a structural and functional level. Ubiquitination plays a role in a wide array of biological processes. Therefore, understanding how this process is regulated and how E3 ligases carry out their functions on a structural level is of great relevance. Moreover, due to their role as central determinants of specificity in the ubiquitination process, E3s have emerged as critical factors in cancer development and thus are promising targets for drug development (Bernassola et al., 2008; David et al., 2013; Zou et al., 2015).

In this thesis, I focused on a particular group of HECT-type E3s: the Nedd4 family. In addition to their physiological relevance, they share a conserved domain architecture which facilitates a comparative analysis among their members. Using NMR spectroscopy and biochemical methods I investigated how the Nedd4-family members Nedd4, Rsp5, Smurf1 and Smurf2 regulate their activity.

Some open questions in the field were:

- How Nedd4-family E3 ligases regulate their activity in order to prevent unwanted ubiquitination?
- Is the already described auto-inhibitory mechanism present and conserved in all members of the family?
- How does the C2 domain mediate this inhibitory role on the HECT domain activity?
- Are there any conserved mechanisms regarding Nedd4 E3 ligase activation?

As a first step to address these questions, I analyzed in detail how the C2:HECT interaction down-regulates the activity of the Nedd4 and Smurf2. I was able to show that the C2 domain induces a catalytically incompetent HECT domain conformation that interferes with the ability of the HECT domain for transthiolation and to interact non-covalently with Ub, thereby inhibiting ligase processivity. Altogether, this provides a detailed explanation on how Nedd4 E3 ligases can regulate their function. These results are published in: "Mari S*, Ruetalo N*, Maspero E, Stoffregen M, Pasqualato S, Polo S, Wiesner S. (2014). Structural and Functional Framework for the Autoinhibition of Nedd4-Family Ubiquitin Ligases. Structure 22: 1639–1649". * First author/Equal contribution.

In a second project, I aimed to shed light on the regulation of Smurf1 activity. To this end, I performed a detailed structural characterization of Smurf1 intramolecular interactions as

well as Ub binding, and examined the activity of Smurf1 in comparison with the closely related protein Smurf2. Surprisingly, in contrast to Smurf2, Smurf1 is not subject to auto-inhibition. I could show that this difference in regulation results from the lack of the first WW domain in Smurf1. Moreover, I found that the Smurf2 WW1 domain interacts with the C2 domain enhancing the interaction and thus auto-inhibition of the HECT domain. In sum, these results showed that not all the members of the Nedd4-family undergo the same regulatory mechanisms and highlights the relevance of mechanistic studies in single proteins, even in highly related ones.

Finally, in a last project, I addressed whether calcium binding is a conserved mechanism of activation in Nedd4 family E3s by triggering the release of the C2 domain from the HECT domain. I found that Nedd4 C2 domain binds calcium through the conserved CBR and that this binding precludes HECT domain interaction, since the binding surfaces use in both cases partially overlap. Other members of the family such as Smurf1, Smurf2 and the yeast homolog Rsp5 do not bind calcium, restricting Ca^{2+}-mediated C2 releasing as an activation mechanism only for Nedd4 and Nedd4L proteins.

3. Materials and Methods

3.1 Constructs

An exhaustive list of all constructs used in the thesis is shown in the Table T2.

Table 2. List of constructs used in the different experiments, in alphabetical order by protein name. Within each protein, the constructs are ordered by decreasing number of residues. Color code: Pink: constructs used only in NMR experiments; yellow: constructs used only in activity assays; blue: constructs used in both types of experiments. N° res: Number of residues; MW: Molecular weight; iso.2: Isoform 2; C.O.: Codon optimized.

N°	Protein	Domain	Mutation	Residues	N° res	MW(kDa)	Tag	vector
1	Nedd4	HECT	ΔC-t	513-893	384	45.5	His	pProExHTb
2	Nedd4	C2	WT	16-153	142	16.5	His/NusA	pETM-10-60
3	Nedd4	C2	D35A	16-153	142	16.5	His/NusA	pETM-10-60
4	Nedd4	C2	D41A	16-153	142	16.5	His/NusA	pETM-10-60
5	Nedd4	C2	D93A	16-153	142	16.5	His/NusA	pETM-10-60
6	Nedd4	C2	N95A	16-153	142	16.5	His/NusA	pETM-10-60
7	Nedd4	C2	D101A	16-153	142	16.5	His/NusA	pETM-10-60
8	Smurf1	FL	WT (iso. 2)	1-731	733	83.6	His	pProExHTb
9	Smurf1	ΔC2	WT	146-731	589	67.5	His	pProExHTb
10	Smurf1	HECT	WT	345-731	391	45.5	His	pProExHTb
11	Smurf1	C2	C.O.	12-141	134	14.9	His/Z-domain	pETZ2-1a
12	Smurf2	FL	WT	1-748	752	86.5	His	pETM-11
13	Smurf2	FL	Y453A	1-748	752	86.5	His	pETM-11
14	Smurf2	FL	ΔC-4	1-744	748	86	His	pETM-11
15	Smurf2	FL	ΔC-4/Y453A	1-744	748	86	His	pETM-11
16	Smurf2	ΔWW1	WT	1-162; 186-754	732	84	His	pETM-11
17	Smurf2	ΔC2	WT	145-748	607	70.4	His	pProExHTb
18	Smurf2	ΔC2	ΔC-4	145-744	603	70	His	pProExHTb
19	Smurf2	HECT	WT	366-748	386	45.1	His	pProExHTb
20	Smurf2	HECT	E404A	366-748	386	45.1	His	pProExHTb
21	Smurf2	HECT	Y453A	366-748	386	45.1	His	pProExHTb
22	Smurf2	C2-linker-WW1	C151A	10-197	191	21.5	His/MBP	pETM-41
23	Smurf2	C2	WT	10-140	134	14.9	His/NusA	pETM-10-60
24	Smurf2	linker-WW1	C151A	145-197	56	6.4	His/MBP	pETM-41
25	Smurf2	WW1	WT	157-193	41	4.7	His/GST	pETM-30
26	Rsp5	C2	WT	2-137	140	15.5	His/NusA	pETM-10-60
27	Sf1:Sf2 WW1	1-158 Sf1, 158-193Sf2, 164-731 Sf1			764	87.3	His	pProExHTb
28	Sf1:Sf2 linker	1-142 Sf1, 141-365 Sf2, 344-731 Sf1			756	86.8	His	pProExHTb
29	Sf2: Sf1 linker	1-141 Sf2, 142-344 Sf1, 366-748 Sf2			729	83.2	His	pETM-11

A codon optimized gene fragment coding the Smurf2 C2 domain (amino acids 10-140) and Smurf1 C2 domain (12-141) were purchased from LifeTechnologies. A cDNA clone encoding human Smurf1 (Sf1) was obtained from BioScience ImaGenes. Nedd4 and Rsp5 coding sequence (CDS) were obtained from Dr. Daniel Rotin (University of Toronto) and Smurf2 from Dr. J.L. Wrana (Molecular Genetics, University of Toronto). Human E1 enzyme

MGC clone 4781 was obtained from SidNet (Toronto, Canada), UbcH7 from Dr. J.L. Wrana (Molecular Genetics, University of Toronto) and Ub from Dr. A.J. Wand (Department of Biochemistry & Biophysics, University of Pennsylvania).
All pETM-derived vectors were obtained from EMBL Heidelberg. pETM-10-60, is a modified version of the pETM-60 vector encoding an additional N-t His6-tag.
All other constructs were generated by PCR from the FL versions of hNedd4, scRsp5, hSmurf1 and hSmurf2.

3.2 Molecular Biology Methods

3.2.1 Cloning

3.2.1.1 PCR amplification

PCR (polymerase chain reactions) were performed according to the manufacturers' protocols using the polymerases Pfu (Promega) for gene amplification and Q5 (New England Biolabs) or KapaHiFi (peqlab) for site-directed mutagenesis and RF Cloning. Primers were order either from eurofins or SIGMA-ALDRICH.

3.2.1.2 Site-directed mutagenesis

Single or multiple point mutations were introduced into the protein CDS using the QuikChange strategy (Stratagene). QuikChange primers were designed using PrimerX software (Lapid & Gao, 2003). Reactions were carried out according to manufacturers´ instructions for the specific polymerase and specific characteristic of each reaction (primer melting temperature, vector and insert length, etc.). Typically, a QuikChange reaction contained: 50-100 ng of template DNA, 2,5 µL of each Fw and Rv primers (10 µM), 5X Buffer, 0,5 µL dNTPS (100 mM), 0,5 µL Polymerase and water q.s. to 50 µL. The PCR program included 20 cycles, using annealing temperatures between 50-65°C, according to each set of primers.
After the PCR reaction, the reaction mixture was treated with Fast Digest (FD) DpnI restriction enzyme (RE) (Termo Fisher Scientific) which degrades only methylated DNA (in this case only the parental DNA is methylated). The product of the reaction was purified using the PCR clean-up kit from Macherey-Nagel following the instructions of the manufacturer. 5 µL of the final product were transformed into DH5α cells. A negative control (PCR without polymerase) was included; it was subjected to DpnI treatment and transformed into cells. The absence of colonies in this plate functions as a strong indicator that the DpnI reaction was worked successfully.

3.2.1.3 Restriction Free (RF) cloning

To deleted or insert large DNA fragments or to clone an entire open reading frame in a different vector RF Cloning was used (Unger et al., 2010). Primers were designed with the

online program available at http://www.rf-cloning.org/. Reactions where set up as described in (Ent & Löwe, 2006). Briefly, a first PCR reaction is used to generate a fragment called "megaprimer", which bears the CDS of interest flanked by complementary regions to the destination vector. A typical RF PCR reactions contained: 40 ng of vector, 200 ng of megaprimer, 5X Buffer, 0,5 µL dNTPS (100 mM), 0,25 µL Polymerase and water q.s. to 25 µL. The PCR program included ~30 cycles, using annealing temperatures between 60-65°C.

The products of the first PCR reaction were analyzed by agarose gel electrophoresis. After the final PCR reaction, the product was treated with FD DpnI and purified using the PCR clean-up kit from Macherey-Nagel following the instructions of the manufacturer. 5 µL of the final product were transform into DH5α cells. A negative control (PCR without polymerase) was included; it was subjected to DpnI treatment and transformed into cells. The absence of colonies in this plate functions as a strong indicator that the DpnI reaction was worked successfully.

3.2.1.4 Restriction digest and DNA ligation

Alternatively to RF cloning, traditional restriction enzymes were used when suited better to the strategy. FD restriction enzymes (Termo Fisher Scientific) were used to digest both vectors and inserts according to the manufacturers' protocols. Vectors were dephosphorylated by adding shrimp alkaline phosphatase (SAP) (New England Biolabs). Typically, restriction digest reactions contained: 1 µg DNA, 1µL of each RE, 10X FastDigest buffer and water q.s. to 20 µL.

DNA Ligation was performed using T4 ligase following manufacturer protocols (New England Biolabs). A molar ration between insert and vector was calculated to perform the reaction in a 1:1 - 1:5 ratio. A negative control (reaction without insert) was always included. After incubation 5 µL of the final product was transform into DH5α cells. Typically, a ligation reaction contained: 100 ng of vector, from 1-5X molar excess of DNA insert, 1 µL T4 DNA ligase, 10×T4 DNA Ligase and water q.s. to 10 µL.

3.2.2 Plasmid transformation

Plasmid transformation into chemically or electro-competent *E. coli* was performed using either DH5α cells for plasmid amplification or chemically competent BL21-CodonPlus (DE3)-RIL cells for protein expression. 50 µL of competent cells were thawed on ice. In case of a PCR reaction 5 µL of PCR product was used, otherwise only 1 µL was used. Cells were incubated with DNA on ice for 30 min and subjected to a heat shock for 60 seconds at 42°C. Alternatively, they were incubated for 5 min. and electroporated. After the procedure, cells were incubated on ice for 2 additional min. Then, 500 µl of LB media was added and the cells were recovered at 37°C for another 30 min. (for vectors with

ampicillin-resistance) or 60 min. (for vectors that were kanamycin-resistant) before plating with the appropriate antibiotic. Plates were incubated overnight at 37°C.

3.2.3 Plasmid isolation and sequencing

Single colonies from agar plates were used to inoculate 7 mL LB cultures, which were grown overnight at 37 °C with the addition of the required antibiotic. Pellets were treated for plasmid isolation according to the NucleoSpin Plasmid EasyPure protocol from Macherey-Nagel. Elution of the plasmid from the column was performed with 30 µL of elution buffer. Plasmids of interest were analyzed by Sanger sequencing. ~80 ng of plasmid was added to a sequencing reaction mix that contained: 1 µL of primer, 5x sequencing buffer, 0,5 µL of BDT-mix (BigDye™Terminator-mix), and water q.s. to 10 µL. Completed reactions were submitted to the in-house sequencing facility for capillarity gel electrophoresis and computer analysis. Sequencing results were analyzed with Bioedit Sequence Alignment Editor (Hall, 1999).

3.3 Protein Methods

3.3.1 Protein Expression and Purification for biochemical assays

All proteins used in biochemical assays were expressed as His6-fusion proteins using *E. coli* BL21-CodonPlus (DE3) RIL cells (Stratagene) grown in Luria broth (LB) medium. Cultures were grown at 37 °C until an OD_{600} of 0.8 was reached, then the temperature was shifted to 18 °C and the cells were induced with 1 mM IPTG and grown overnight (16 h). Cells were harvested by centrifugation and lysed by sonication on ice (lysis buffer: 50 mM Tris-HCl pH 8, 150-300 mM NaCl, 0-5% glycerol, 10-30 mM imidazole, 1 mM DTT (Dithiothreitol). Protease inhibitors (Roche) were added for FL proteins. Cell debris was removed by centrifugation. Soluble recombinant proteins in the supernatants were purified by Immobilized metal affinity chromatography (IMAC). N-terminal solubility tags (NusA, MBP, or Z-domain) were cleaved with TEV protease and removed together with the His-tagged TEV protease via Ni-affinity and / or size exclusion chromatography (SEC). For the FL or chimeric proteins expressed from pETM-11 or pProExHTb vectors, the His-tag was not removed and after IMAC the buffer was exchange using PD10 columns. In the case of HECT domains the His-tag was removed by TEV protease and Ni-affinity when used for auto-ubiquitination and competition assays. On the contrary, no His-tag was removed for pull-down experiments. The final protein buffer for assays was 50 mM Tris pH 8.0, 200 mM NaCl, 2.5% glycerol, 1 mM DTT.

3.3.2 Protein Expression and Purification for NMR Spectroscopy

All proteins for NMR spectroscopy were expressed in *E. coli* BL21-CodonPlus (DE3) RIL cells (Stratagene). Unlabeled proteins used as titration ligands were grown in LB or in 90/10% D_2O/H_2O when the complexes to study exceeded 40 kDa. ^{15}N labeled-proteins

were expressed in M9 minimal medium (when smaller than 20 kDa), otherwise in 90/10% D_2O/H_2O using $^{15}NH_4Cl$ as sole sources of nitrogen. ^{13}C IM (Met-[ε $^{13}CH_3$]-, Ile-[δ1 $^{13}CH_3$]- or ILVAM- (Ile-[δ1 $^{13}CH_3$]-; Leu-[δ $^{13}CH_3$]-; Val-[γ $^{13}CH_3$]-; Ala-[β $^{13}CH_3$]-; Met-[ε $^{13}CH_3$]), HECT domains were expressed in 90/10% D_2O/H_2O or 100% D_2O M9 minimal medium using $^{14}NH_4Cl$ and $^{12}C^2H$ glucose as nitrogen and carbon source respectively. Pre-cultures were grown in LB and the amount of bacteria required for inoculation of a culture at an OD_{600}=0.4 spun down and re-suspended in the minimal media. Isotopic labeled amino acids (L-Methionine-methyl-$^1H^{13}C$, L-Alanine-3-$^2D^{13}C$) or precursors (Ile: 2-Ketobutyric acid-4-^{13}C; Val-Leu: 2-Keto-3-methyl-^{13}C-butyric-4-^{13}C acid) were added to the growth medium 1 h before induction, which was performed at an OD_{600}=0.8 between 20-25 °C with 1mM IPTG. The cells were grown overnight (~ 12-16 h). Cells were harvested by centrifugation and lysed by sonication on ice (lysis buffer: 50 mM Tris-HCl pH 8, 150-300 mM NaCl, 0-5% glycerol, 10-30 mM imidazole, 1 mM DTT). Cell debris was removed by centrifugation and purification performed by IMAC. Removal of His-tag and N-t solubility tags (NusA, MBP, GST or Z domain) was done by a 2nd IMAC and SEC when required.
Both, the Nedd4 C2 and HECT domains, were buffer exchanged into NMR buffer (50 mM Tris-HCl pH 7.0, 200 mM NaCl, 2.5% glycerol, 1 mM DTT). In the case of Smurf1 and Smurf2 experiments the HECT and C2 domains as well as all Smurf2 C2-WW1 and Ub were buffer exchanged into NMR buffer (20 mM NaP pH 6.5 or 7.5, 150 mM NaCl, 1 mM DTT). Finally, the C2 domains used in Ca^{2+} binding experiments were exchanged into the NMR buffer (20 mM Tris pH=7-7.5, 250 mM NaCl, 1mM EGTA).

3.3.3 SEC

When required, proteins where further purified by SEC. According to the protein size Superdex 75 (3-70 kDa) or 200 (10-600 kDa) 16/60 columns (GE) were used. The columns were used according to the instructions of the manufacturer on either an NGC (Biorad) or ÄktaPrime (GE) FPLC system. Fractions were collected in volumes of 0.5-1 mL. Elution profiles were monitored by UV absorption at 280nm and the presence and purity of the desired protein verified by SDS-PAGE.

3.3.4 PAGE

3.3.4.1 Tris-glycine SDS-PAGE

Relevant fractions after protein purification, level of protein purity as well as protein size were analyzed by SDS-PAGE. Gels were casted from a 40%, 29:1 mix of acrylamide: bisacrylamide (Sigma) using standard recipes (Laemmli, 1970). Proteins were loaded using 3X Laemmli buffer on gel with different acrylamide percentages (8, 12 or 16%) according to the characteristics of the sample. Electrophoresis was ran at 220 V and gels stained with Coomassie brilliant blue G-250.

3.3.4.2 Tris-tricine SDS-PAGE

This method is useful for separation of peptides and proteins under 10 to 15 kDa, what is usually not possible in the traditional Tris-glycine discontinuous gel system, due to the co-migration of SDS and small proteins. In order to obtain better results for Ub analysis, 11% Tris-tricine (Schägger, 2006) gels were used.

3.3.5 Western blot

Antibodies and their suppliers were: α-mouse mono-Ub (P4D1; Santa Cruz Biotechnology); α-mouse mono-HA (Sigma), and α-HRP-coupled mouse IgG secondary antibody (ThermoScientific).

Desired amounts of proteins were loaded onto SDS-PAGE or Tricine-SDS-PAGE for electrophoresis. Proteins were transferred to nitrocellulose (Santacruz Biotechnology) in the case of pull-down experiments or polyvinylidene fluoride (PVDF) membranes (Immobilion P, Millipore) for auto-ubiquitination assays in western blot transfer tanks (Hoefer system) that contained transfer buffer (250 mM Tris, pH 8.3, 192 mM Glycine, 20% v/v ethanol). PVDF membranes were previously activated by incubation in 100% MetOH for 1 min. at room temperature. Protein transfer was achieved at 50V for 1.5 hs or 20V overnight. For nitrocellulose membranes, Ponceau staining was carried out after protein transfer. Membranes were blocked for 1 h (or overnight) in 5% milk in Tris-buffered saline (TBS) supplemented with 0.05% Tween (TBS-T). After blocking, they were incubated with the primary antibody (α-Ub 1:1000; α-HA 1:5000) diluted in TBS-T 2.5% milk for 1.5 hs at room temperature, followed by three washes of 10 min. each in TBS-T. Next, membranes were incubated with the appropriate HRP-secondary antibody (α-mouse IgG 1:10000) diluted in TBS-T 2,5% milk for 1 h. Finally, membranes were washed three times in TBS-T and imaged using the enhanced chemiluminescence (ECL) method (Pierce). Images were acquired by using an Amersham Imager 600 (GE). After signal detection membranes were stained with InstantBlue (expedeon).

Western blot quantification was performed using the open source software ImageJ 1.x (by Wayne Rasband, 1997; Schneider et al., 2012) following the recommendations for densitometric analysis in the ImageJ User Guide IJ 1.46r.

3.4 NMR spectroscopy

All NMR data were collected on an 800 MHz Bruker Avance-III spectrometer, equipped with a room temperature probe head. The software used was TOPSPIN 2.1 (Bruker, 2008). All NMR data were processed and analyzed using the NMRPipe/NMRDraw program suite (Delaglio et al., 1995), and depicted with NMRView (OneMoonScientific).

3.4.1 Nedd4 Interaction studies

The NMR active sample was a 100 µM partially deuterated ^{15}N-labeled Nedd4 HECT ΔC-t domain. 2D ^{1}H,^{15}N-TROSY-based spectra were recorded at 30 °C before and after addition of a two-fold stoichiometric excess of unlabeled Nedd4 C2 domain or Ub, respectively.

3.4.2 Smurf1 and Smurf2 HECT CSP experiments

The NMR active sample was a 50 µM IM or ILVAM-labeled Smurf1 HECT domain and IM-labeled Smurf2 HECT domain. 2D ^{1}H,^{13}C-methyl-TROSY spectra were recorded at 30 °C, before and after addition of unlabeled Smurf1/Smurf2 C2 or Smurf2 C2WW1 domain respectively or monomeric Ub. The stoichiometric excess used is indicated in each figure.

3.4.3 Smurf1 HECT Resonance assignment

All Ile δ1 and Met ε-methyl groups of the Smurf1 HECT domain were mutated to Val or Leu. In total, 29 individual point mutants of the Smurf1 HECT domain were generated: I375V, I382L, I390V, M391V, M393V, M403V, M429V, I444V, I449L, M446V, I455V, I469V, M470V, I480V, I497V, I516V, I521V, I539V, M571V, I574V, I589V, I605V, I606V, I612V, I632V, I683V, I686V, I703V, I705V. 2D ^{1}H,^{13}C -methyl-TROSY spectra of 40 µM IM-labeled mutant Smurf1 HECT domains were recorded at 30 °C and compared to the WT spectrum.

3.4.4 Kd determination by two-dimensional lineshape fitting analysis.

Dissociation constants were calculated from ^{1}H,^{13}C-methyl-TROSY spectra using TITAN (Methyl version:cwaudby_titan_95C5769C3906) according to developer´s instructions and online documentation (Waudby et al., 2016). Spectra were acquired with 1024 and 120 points in the ^{1}H and ^{13}C dimensions respectively, and processed with exponential window functions with a line broadening of 4 Hz and 8 Hz. Spectra were zero-filled to 4096 and 480 points in the ^{1}H and ^{13}C dimensions, respectively. Errors were estimated with bootstrapping statistics on 100 replicas. Figures for lineshape analyses were prepared with TITAN.

3.4.5 C2 domain CSP experiments

The NMR active sample were 100 µM ^{15}N-labeled C2 domains: Nedd4, Smurf1, Smurf2 and Rsp5. 2D ^{1}H,^{15}N-HSQC spectra were recorded at 25 °C before and after addition of a buffered $CaCl_2$ solution (20 mM Tris pH=7, 250 mM NaCl and 1 mM EGTA). For Nedd4 C2 mutants the same buffer (pH =7 or 7.5) was used.

Chemical shift perturbations were analyzed with TOPSPIN and quantified as average chemical shift changes $\Delta\delta A_v = ((\Delta\delta(^{1}H))^2/5 + (\Delta\delta(^{15}N))^2)^{1/2}$ in p.p.m. The absolute value was plotted against the peak number.

For the Nedd4 HECT domain, 2D $^{1}H,^{15}N$-TROSY spectra of an 80 μM ^{15}N-labeled sample in the absence and presence of unlabeled C2 domain were recorded at 25 °C. After that, an excess of 20 mM $CaCl_2$ was added and a new spectra recorded.

3.5 Structure modeling and visualization

A homology model of the Smurf1 HECT domain were generated using HHPred based on the structure of the Smurf2 HECT domain (1ZVD) that shares 86% sequence identity. Alignments are made with Jalview (Waterhouse et al., 2009). Prediction of secondary structure elements was performed with Quick2D (Alva et al., 2016).
All structural representations were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.7.6.6 Schrödinger, LLC).

3.6 Functional assays

3.6.1 Ubiquitination and Competition Assays

Auto-ubiquitination assays were performed at room temperature in 25 μL reactions using ubiquitination buffer (25 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1 mM DTT, 5 mM $MgCl_2$, and 2.5 mM ATP), and purified enzymes (0.6 μM E1, 90 μM UbcH7 (E2), 2,5 μM E3s) and 65 μM Ub. Reaction mixtures were stopped at specific time points indicated in each figure by addition of Laemmli buffer containing 100 mM DTT. For SDS-PAGE analysis, samples were loaded on 8-12% acrylamide gels. Detection was performed either by immunoblotting with an α-Ub antibody. After signal detection, PVDF membranes were stained with Coomassie to show equal loading of proteins. Alternatively, fluorescent labeled Ubiquitin was used, kindly provided by Magnus Jäckl. In this case no Western blot is needed, since the fluorescent signal of Ub can be directly measured in an Imager.
Competition assays were performed for Smurf1 and Smurf2 HECT domains. In addition to the reaction mix used for the ubiquitination assay, an excess of Smurf1 or Smurf2 C2 domain or a C2-linker-WW1 fragment (12.5, 25, 50 or 75 μM) was added. Samples were incubated for 20 min. at room temperature and loaded on 12% acrylamide gels for SDS-PAGE and immunoblotting. An input gel (in the absence of E1 and E2) was run in each case on a 16% acrylamide SDS-PAGE to show the initial amounts of protein.

3.6.2 Thioester Assays

Thioester assays were done essentially as previously described (Wiesner et al., 2007). They were performed at room temperature in ubiquitination buffer using recombinantly expressed and purified enzymes (0.6 μM E1, 9 μM UbcH7, 4 μM Smurf2 proteins) and 13 μM HA-tagged Ub. To better visualize the reduction-sensitive Ub thioester, the Smurf2 WT, ΔC2 and Y453A enzymes were truncated by four residues from their C termini (''-4'')

as deletion of the conserved -4 Phe position impairs HECT-mediated Ub isopeptide-linkage without affecting thioester formation (Salvat et al., 2004). The reactions were divided in two after the indicated incubation times and stopped with SDS-PAGE loading buffer without DTT or with 100 mM DTT. Ub-modified proteins were detected by immunoblotting with α-HA antibody against HA-tagged Ub, whereas levels of unmodified proteins were depicted by Ponceau staining.

3.7 Pull-Down Assays

Bacterially expressed His6-tagged Smurf2 proteins were incubated at 2 mM with 250 ng synthetic K63-linked poly-Ub (Boston Biochem) for 2 hs at 4 °C in YY buffer (50 mM Na-HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 10% glycerol, 1% Triton X-100). After four washes with YY buffer, specifically bound proteins were resolved with Tris-Tricine PAGE (11%) and detected by immunoblotting using α-Ub antibody. Nitrocellulose membranes were then stained with Coomassie to show equal loading of proteins.

4. Results

Chapter 1: C2-meadiated auto-inhibition of Nedd4-family E3 ligases

4.1.1 Contribution

This work has been carried out in collaboration with the research group of Dr. Simona Polo at the Institute FIRC of Molecular Oncology (IFOM), Milan, Italy. The results were published in: "Mari S*, Ruetalo N*, Maspero E, Stoffregen M, Pasqualato S, Polo S, Wiesner S. (2014). Structural and Functional Framework for the Autoinhibition of Nedd4-Family Ubiquitin Ligases. Structure 22: 1639–1649". * Co-first author/Equal contribution.
All the results shown in the section 4.1.3 of this thesis were performed by me.

4.1.2 Introduction

The fact that the activity of the HECT-type Smurf2 Ub ligase is regulated by an auto-inhibitory mechanism, where the C2 domain interacts with the HECT domain, was already describe by Wiesner et al., 2007. There, the HECT binding surface on the C2 domain was characterized and a preliminary C2 interaction site on the HECT domain was mapped, showing that the C2 domain binds to the N-lobe on the HECT domain. In Mari*, Ruetalo* et al., 2014, a detailed characterization of the C2 binding surface from the HECT domain side was done, using "methionine scanning" NMR spectroscopy (Stoffregen et al., 2012). Briefly, TROSY spectra of the WT $^{1}H,^{13}C$- Ile-δ_1, Met-ε-methyl labeled Smurf2 HECT domain were recorded in absence and presence of four-fold excess of unlabeled C2 domain. This lead to the identification of five natural occurring Met and Ile residues

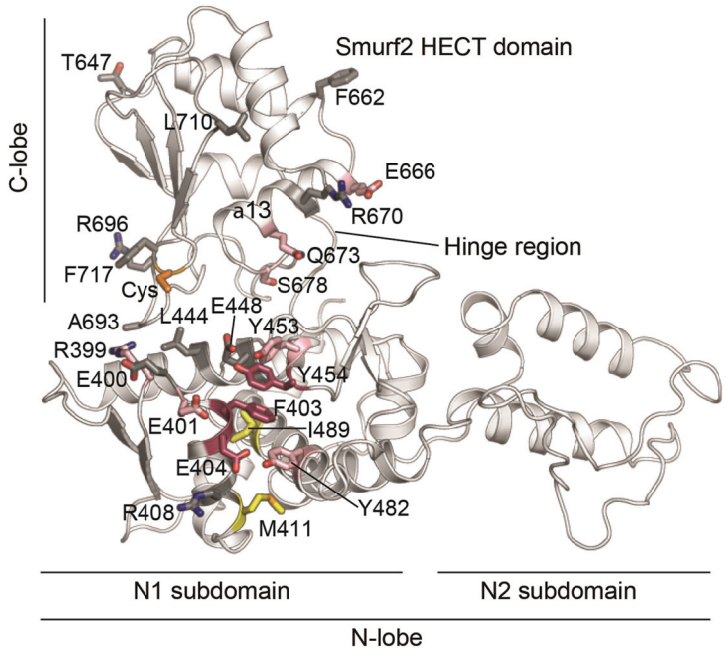


Figure 10. Chemical shift mapping of the Smurf2 C2 binding surface onto the HECT domain. The residues displaying CSPs are highlighted on the crystal structure of Smurf2 HECT domain (PDB ID: 1ZVD). ''Hot spot'', ''In'' and ''Out'' residues are shown in dark pink, light pink and dark gray respectively. Relevant CSPs on naturally occurring methyl groups (IM_{ref}) are shown in yellow and the catalytic Cys in orange. The Met scanning experiments were performed by Mira Stoffregen and Silke Wiesner. This figure is reproduced from Mari*, Ruetalo* et al., 2014 with permission from Elsevier.

involved in the interaction (I402, M411, M449, I489 and I626). The individual mutation of 21 solvent-exposed residues in the region, together with the $^{1}H,^{13}C$-Ile-δ_1, Met-ε-methyl TROSY spectra assignments (which were previously done by site-directed mutagenesis), made it possible to identify the C2 binding surface in detail (Figure 10). The data showed that the binding occurs through electrostatic and hydrophobic interactions by residues located on the N1 sub-domain of the N-t lobe of the Smurf2 HECT domain. The Ub non-covalent interaction surface has been shown to be essential in Ub chain elongation for Smurf2. Interestingly, the Nedd4 C2 domain binding surface, overlaps substantially with this previously mapped interface (Ogunjimi, et al., 2010). In 2012, preliminary results from Dr. Simona Polo´s group suggested that same auto-inhibitory mechanism may regulate Nedd4 activity. In that context, I aimed to find structural evidence of this mechanism in Nedd4 and to study how the C2-HECT interaction mediates the activity inhibition that was observed both for Smurf2 as well as for Nedd4.

4.1.3 Results

Structural analysis of Smurf2 HECT binding surfaces clearly showed a partial overlap between the C2 binding surface that our group mapped, and the non-covalent UBS previously described by several groups (French et al., 2009; Kim et al., 2011; Maspero et al., 2011; Ogunjimi et al., 2010) which is shown in Figure 11A-B.

Considering that the residues involved in the C2 binding pockets are highly conserved in many HECT domains of the Nedd4 family, I studied a possible C2-HECT interaction for Nedd4. A methyl-TROSY spectra of a $^{1}H,^{13}C$-Ile-δ_1, Met-ε-methyl labeled Nedd4 HECT domain was recorded. Unfortunately, the quality of the spectra was not good enough to performed titration experiments due to precipitation of the Nedd4 C2 domain at concentrations higher than a two-fold excess. Alternatively, I recorded $^{1}H,^{15}N$-TROSY NMR spectra of a partially deuterated ^{15}N-labeled Nedd4 HECT domain, in the absence and presence of a two-fold stoichiometric excess of unlabeled C2 domain. Numerous CSPs were detected, showing that the two domains are able to interact with each other (Figure 11C).

Next, I analyzed whether the overlap between the C2 domain and the UBS described for Smurf2 was also taking place in Need4. I performed a titration experiment using $^{1}H,^{13}C$-Ile-δ_1, Met-ε-methyl labeled Nedd4 HECT domain and a two-fold excess of Ub as a ligand. CSPs were observed, confirming that Nedd4 HECT domain interacts with Ub non-covalently (Figure 11D). Interestingly, an analysis of both spectra (Figure 11C-D) showed that some resonances of the Nedd4 HECT domain shifted upon addition of both ligands: Ub and the C2 domain. This means that the residues displaying CSPs in both spectra, marked with arrows in Figure 11C-D, are involved in the interaction with the C2 domain as well as Ub. In addition, they shifted into different directions, which is expected from the different chemical nature of both ligands. Since resonance assignments of the Nedd4 HECT domain were not available, the residues involved in Nedd4 HECT:Ub interaction could not be identified. None-the-less, this result demonstrates that as for Smurf2, the C2 and UBS surfaces do also overlap on the Nedd4 HECT domain.

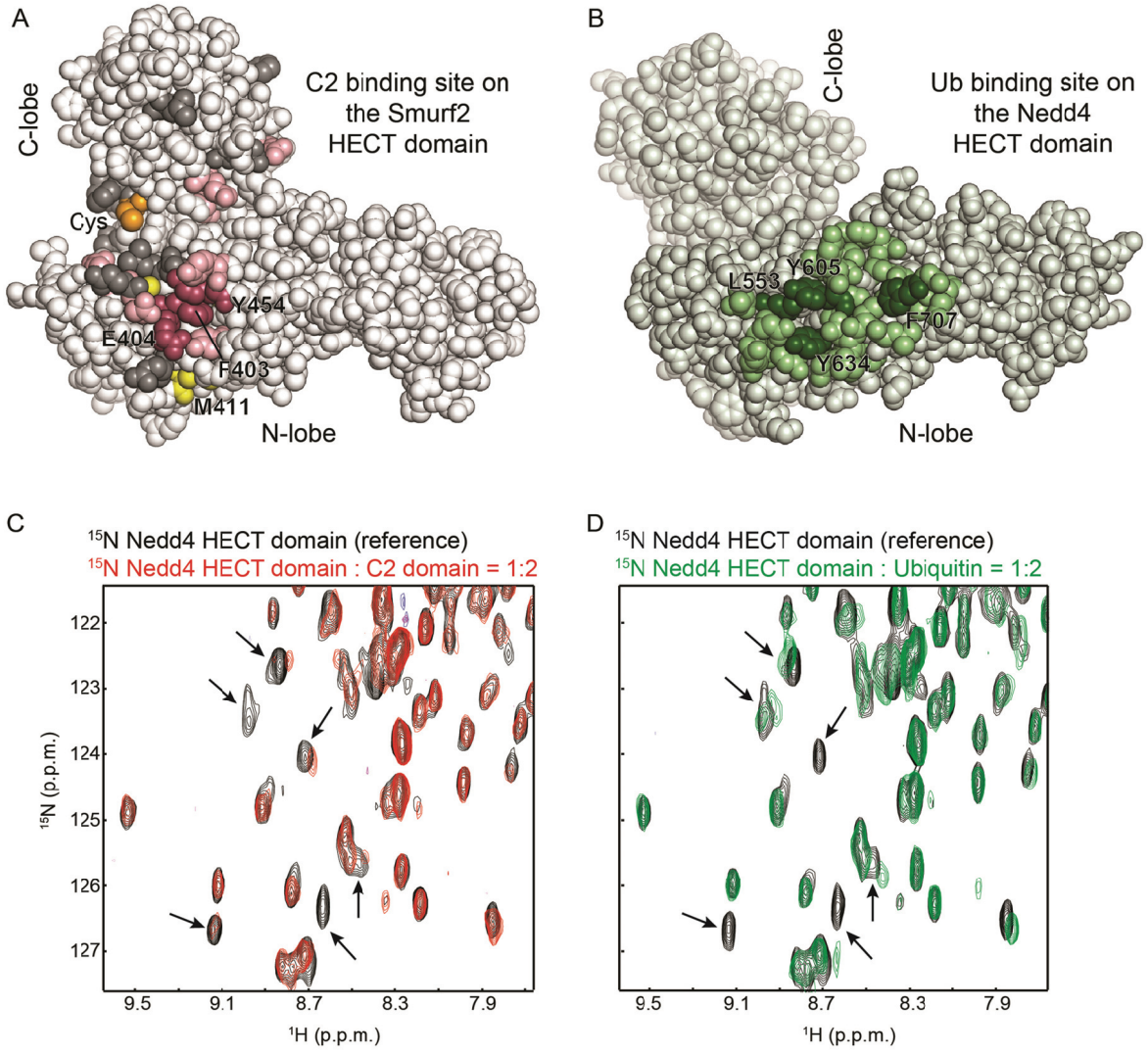


Figure 11. The C2 binding surface overlaps with the non-covalent UBS. A) The Smurf2 HECT domain is represented as spheres (PDB ID: 1ZVD), with the C2 binding surface color coded as in **Figure 10. B)** Nedd4 HECT crystal structure (PDB ID: 2XBB) is represented as spheres. Residues involved in non-covalent Ub binding are highlighted in light green, while residues impairing Ub binding in dark green, as describe previously by Maspero et al., 2011. **C)** Overlay of a selected region of the $^{1}H,^{15}N$-TROSY spectra of Nedd4 HECT domain in the absence (black) and presence of a two-fold excess of unlabeled Nedd4 C2 domain (red). **D)** As **C**, but in the absence (black) and presence of a two-fold excess of unlabeled Ub (green). The arrows mark the resonances affected by both ligands. This figure is reproduced from Mari*, Ruetalo* et al., 2014 with permission from Elsevier.

As shown by Wiesner et al., 2007 for Smurf2 and Mari*, Ruetalo* et al., 2014 for Nedd4 the deletion of the C2 domain cause a huge activation of the protein in comparison with the FL WT protein, as judged by their enhanced auto-ubiquitination kinetics. To study whether the C2-HECT interaction may inhibit HECT activity by regulating non-covalent Ub binding, I examined whether the Smurf2 C2 domain was able to interfere with Ub binding to the HECT domain. To do so, I performed pull-down assays with bacterially expressed Smurf2 FL, a mutant lacking the C2 domain (ΔC2) and the isolated HECT domain using commercially available K63-linked Ub chains, which mimic the physiological reaction product of Nedd4 and Smurf2 catalysis. Consistent with the structural data (**Figure 10**), the HECT domain and the ΔC2 mutant efficiently pull-down poly-Ub chains, while the FL enzyme was unable to bind poly-Ub chains (**Figure 12**).

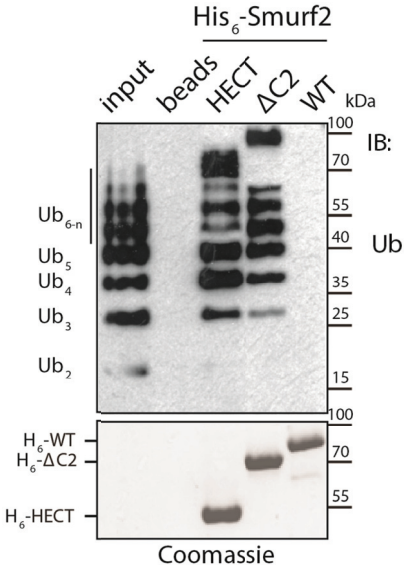


Figure 12. The C2:HECT interaction prevents non-covalent Ub binding to the HECT domain. Pull-down assays were conducted with the indicated His-tagged Smurf2 proteins and synthetic K63-polyUb chains (input). Poly-Ub chains were detected by immunoblotting (IB) with α-Ub antibody (top panels). Initial protein levels were confirmed by Coomassie staining (bottom panels). In contrast to the other two Smurf2 variants that lack the C2 domain, the FL Smurf2 protein is not able to pull-down poly-Ub chains. This figure is reproduced from Mari*, Ruetalo* et al., 2014 with permission from Elsevier.

To address the functional importance of the C2:HECT interaction in more detail, I generated two Smurf2 HECT mutants: E404A and Y453A. These residues were classified as "hot spot" (E404; meaning that this residue is crucial for the C2:HECT interaction) and "inside the binding pocket" (Y453) respectively for the C2-HECT interaction (**Figure 10**) (Mari*, Ruetalo* et al., 2014). To analyze whether the mutation was not interfering with

catalytic activity, I performed *in vitro* ubiquitination assays with the mutant HECT domains, and compared their activity with the WT HECT domain. As shown in Figure 13A, the mutant behavior is comparable to the WT. In addition, the functional relevance of the equivalent residues in Nedd4 (E554 and Y604) were tested by our collaborators, i.e. Ub binding capabilities as well as E2 to E3 transthiolation levels, showing that these mutants were fully functional. With this knowledge, I carried out *in vitro* auto-ubiquitination assays with Smurf2 comparing the FL protein, the ΔC2 mutant and the Y453A FL protein (Figure 13B). I observed that the Y453A mutant is virtually as active as the ΔC2 enzyme, while in the case of the FL WT enzyme, lower levels of auto-ubiquitination were observed. This result demonstrates that the disruption of the C2:HECT binding surface is sufficient to activate the enzyme, since the point mutant reaches the activity levels of the ΔC2 enzyme. In addition, the results of Figure 13B validate the previously mapped Smurf2 C2:HECT interaction surface (Figure 10), since the result showed for Y453 can be explained only if the residue is located inside the binding pocket.

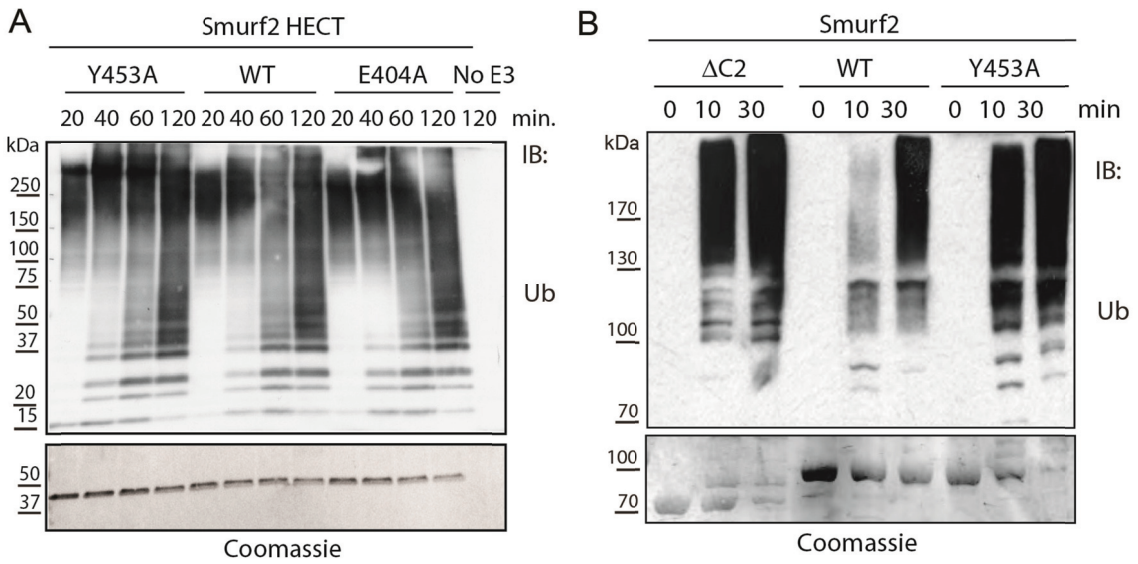


Figure 13. E3 auto-inhibition can be released by mutations in the C2:HECT interface.
In vitro ubiquitination assay using the indicated Smurf2 proteins. Auto-ubiquitination activity was detected by immunoblotting (IB) with α-Ub antibody (top panels). Initial protein levels were confirmed by Coomassie staining (bottom panels). **A)** Characterization of mutations in the C2:HECT interface. As shown in the figure, the auto-ubiquitination activity of the Smurf2 HECT domains is not affected by the point mutations Y453A and E404A. **B)** Mutation of the C2 binding surface in the context of FL Smurf2. Y453A mutant is virtually as active as the ΔC2, whereas low levels of auto-ubiquitination are observed for the FL WT enzyme, showing that the disruption of the C2:HECT binding surface is sufficient to activate the enzyme. Figure 13B is reproduced from Mari*, Ruetalo* et al., 2014 with permission from Elsevier.

Since deletion of the C2 domain was shown to enable thioester bond formation (Wiesner et al., 2007), I also performed thioester assays with FL WT Smurf2, the ΔC2 and the Y453A mutants. As shown is Figure 14, the ΔC2 and Y453A gain-of-function mutant, formed the Ub thioester significantly more efficiently than the WT protein.

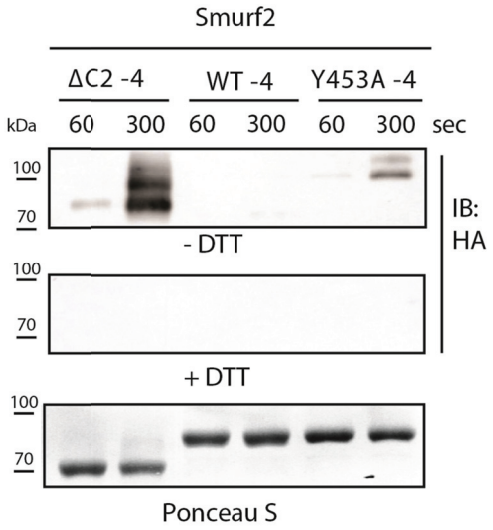


Figure 14. Ub thioester formation is enhanced by mutations in the C2:HECT interface. Thioester assays were performed using Smurf2 proteins carrying a C-t truncation of four amino acids (-4). This deletion severely impairs HECT-mediated Ub isopeptide-linkage without affecting thioester formation, which facilitates the detection of the Ub thioester (Salvat et al., 2014). Proteins were detected by immunoblotting with α-HA antibody against HA-tagged Ub. Initial protein levels were visualized by Ponceau staining. As shown in the upper two panels, the ubiquitinated species observed in the absence of DTT disappear in its presence, confirming the labile nature of thioester the bond. This figure is reproduced from Mari*, Ruetalo* et al., 2014 with permission from Elsevier.

Conclusions:

- The results demonstrate an inhibitory effect of the C2 domain on HECT catalytic activity for the Nedd4 family members Smurf2 and Nedd4. C2 domain binding to the HECT domain blocks the UBS, which is essential for poly-Ub chain formation.
- The activity of Smurf2 FL enzyme is increased through mutations in the HECT domain, which interfere with the C2:HECT interaction, releasing the auto-inhibition mechanism. A single point mutation in this surface (e.g. Y453A) is sufficient to generate this effect, since the point mutant activity resembles the deletion of the C2 domain.
- The presence of the Smurf2 C2 domain blocks the Ub transthiolation process, and thus enzyme activity.
- Comparable results were obtained for Nedd4 (Mari*, Ruetalo* et al., 2014) supporting the concept of structural and functional similarity inside the Nedd4-family.

Finally, considering my NMR and biochemical analyses, together with previous results, I can conclude that the Smurf2 and Nedd4 C2 domains down-regulate HECT activity by: 1) interfering with Ub transthiolation and 2) blocking non-covalent Ub binding to the HECT domain. A model for the auto-inhibition mechanism is depicted in figure 15.

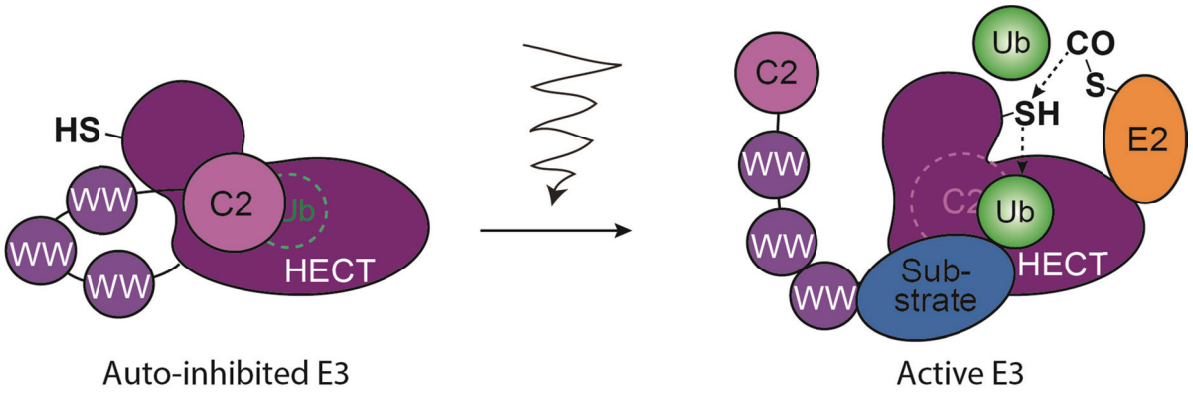


Figure 15. Model for C2-mediated down-regulation of Smurf2 and Nedd4. In the auto-inhibited form of the FL enzymes, the C2 domain induces a catalytic incompetent conformation of the HECT domain that prevents thioester formation. In addition, the C2 domain at least partially blocks the UBS (left). Upon activation the C2 domain is released from the HECT domain, which in turn allows for Ub transthiolation, catalysis and Ub chain elongation (right). This figure is adapted from Mari*, Ruetalo* et al., 2014.

4.1.4 Discussion

In the first part of this thesis, I characterized from a structural and functional perspective two members of the Nedd4-family E3 ligases. I studied how the C2 domain mediates an effect on HECT domain activity that leads to down-regulation of both Smurf2 and Nedd4. On the one hand, my pull-down assays showed that the C2 domain decreases HECT activity by blocking poly-Ub binding to the HECT domain, which is achieved by partially burying the UBS surface. On the other hand, C2 domain binding keeps the HECT domain in an inactive conformation, where E2-E3 transthiolation cannot occur. In Mari*, Ruetalo* et al., 2014 we showed that some of the residues participating in the C2 binding pocket belong to the C-lobe of Smurf2. By modeling different C-lobe conformations with respect to the N-lobe, taking into account the well-known flexibility of the hinge region that connect them, it was possible to define a continuous C2 interaction surface, where the C-lobe residues extend the negatively charged binding surface on the N-lobe. In this modeled conformation, the catalytic Cys on the E3 is located far away from the E2´s Cys, providing a structural explanation for the defective Ub transfer observed in the FL WT proteins.

Interestingly, the inhibition that the C2 domain exerts on the HECT domain is not fully effective. In ubiquitination assays, we consistently observed low levels of FL activity. This is probably due to a dynamic interaction between the C2 and the HECT domains, whit "on" and "off" states, resulting in basal activity. In the context of E3 ligase function, rather than achieving full inactivation of the enzyme, it may be sufficient to avoid the generation of a signal that would deplete the ligase itself or off-targets from the cell. In addition, this transient nature of the C2-HECT interaction might be beneficial if a fast response from these enzymes is needed. A consequence of this transient dissociation would be an eventual thioester formation. However, the C2 could still have the possibility to re-associate, and in this case to outcompete the Ub from the UBS, inhibiting chain formation. This then, might be a reason why the C2 domain has the potential to inhibit HECT domains at two different steps of the ubiquitination pathway.

Finally, the observation of a rather large overlapping region between the C2 binding surface and the UBS, would suggest that Ub could compete with the C2 domain for the HECT domain binding. The K_d values determined for different C2: or Ub:HECT interactions might support this hypothesis since they are in the same range (Table 3). However, the auto-ubiquitination assays showed that although a large stoichiometric excess of Ub with respect to the C2 domain is present in the assays, it is not enough to activate FL Smurf2. This fact evidences the strong effect that C2 domain can exert on the HECT domain as a result of an intramolecular interaction.

Chapter 2: Regulation of Smurf1 Activity: Differences and Similarities with Smurf2

4.2.1 Contribution

Silke Wiesner performed the titration experiments shown in Figure 22A and the NMR experiments to assign the linker-WW1 constructs. She assigned the C2-linker-WW1 region spectra shown in Figure 31A. Samira Anders performed the titration experiments shown in Figure 29 and Figure 31A. Magnus Jäckl recorded the data shown in Figure 30. The backbone resonance assignments of the Smurf1 C2 domain were available in our group, performed by Christine Wolf. All other experiments shown in the section 4.2.3 of this thesis were performed by me.

4.2.2 Introduction

Under basal cell conditions (in the absence of substrates or adaptor proteins), the catalytic activity of Nedd4-family E3s is down-regulated by intramolecular interactions that prevent premature auto- and substrate ubiquitination. In various Nedd4 members, such as Smurf2 (Mari*, Ruetalo* et al., 2014; Wiesner et al., 2007), Nedd4 (Mari*, Ruetalo* et al., 2014; Mund & Pelham, 2009) and Nedd4L (Escobedo et al., 2014; Wang et al., 2010) the N-t C2 domain interacts with the C-t HECT domain to inhibit ligase activity. This interaction precludes E2-E3 transthiolation and blocks the non-covalent UBS that is important for Ub chain elongation (Figure 15, 16) (Mari*, Ruetalo* et al., 2014; Wiesner et al., 2007). In the case of WWP1 (Courivaud et al., 2015) and WWP2 (Mund et al., 2015; Wiesner et al., 2007) *in vitro* ubiquitination assays have shown that it is a combination of C2 and WW domains that down-regulates activity. In the case of Itch instead, auto-inhibition does not involve the C2 domain, but is mediated by an intramolecular interaction between the two central WW domains and the HECT domain (Gallagher et al., 2006; Riling et al., 2015). Recently, two crystal structures provided an explanation on how the WW domains of Itch, WWP1 and WWP2 maintain their enzymes in the closed conformation. The linker between the WW2 and WW3 domains interacts with the HECT domain, impairing E2-E3 transthiolation (Chen et al., 2017; Zhu et al., 2017). In more detail, Chen et al., 2017 showed for WWP2 that a 26-residue long α-helix in the WW2-WW3 linker wraps around the HECT domain, making extensive contact with both the N and C-lobe (Figure 16). This interaction restrains the flexibility of the C-lobe, contributing to enzyme inhibition. In the case of Itch, the WW2 domain and the linker connecting the WW2 and WW3 domains (WW2L) bind specifically to the HECT domain, restricting its inter-lobe motions (Zhu et al., 2017). Detailed analysis of the interaction surfaces revealed that the region where the WW2L binds to the HECT domain partially overlaps with the UBS (Figure 16). An alignment

of all nine Nedd4-family members shows that the residues involved in the WW2L:HECT interaction for Itch are well conserved in WWP1/WWP2, but less so in other HECT domains. On the other hand, the majority of residues involved in C2:HECT for Smurf2 binding are more conserved in Smurf1, Nedd4 and Nedd4L (Zhu et al., 2017). Based on these previous studies, one can divide the mechanisms of Nedd4 auto-inhibition in two: 1) mediated by the C2 and the HECT domains, as in Smurf2, Nedd4 and Nedd4L and 2) mediated by the WW2-WW3 linker (referred to hereafter as WW2-3L) and the HECT domain as in WWP1, WWP2 and Itch (Figure 16). Of notice, a helix similar to the one present in WWP1/2 and Itch, but located in the WW1-WW2 linker, has an auto-inhibitory effect on Nedd4 (Chen et al., 2017), showing that even in this case the C2 domain might not act alone. So far, no information is available for HECW1 and HECW2, two members of the family that remain almost unstudied.

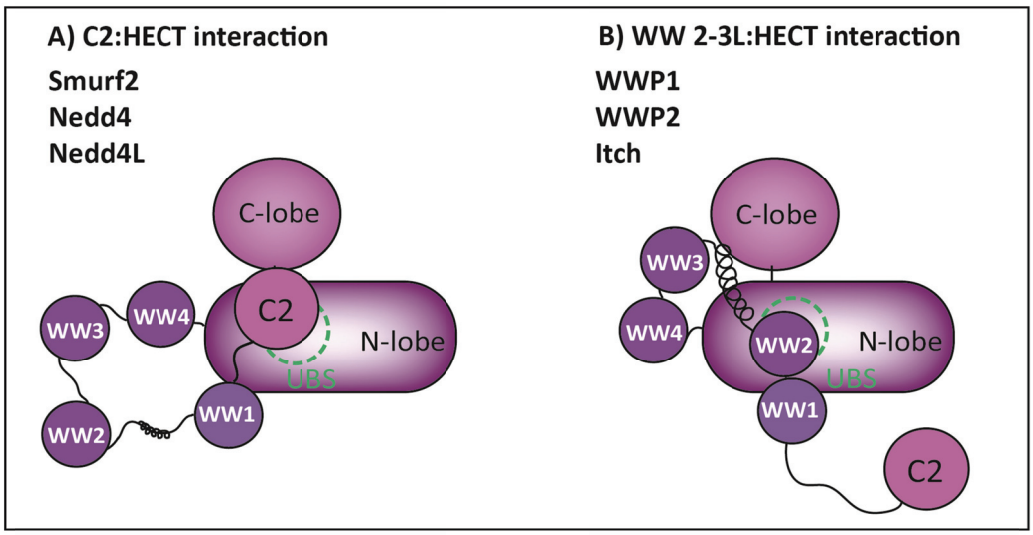


Figure 16. Modes of auto-inhibition within the E3 Nedd4-family. A) The auto-inhibition of Smurf2 and Nedd4 is achieved by two events mediated by the C2 domain. In the first place, the C2 domain locks the C-lobe of the HECT domain in a conformation that does not allow E2-E3 transthiolation. Secondly, since the C2 domain binds in an interface which partially overlaps with the UBS, the HECT domain capacity of binding Ub non-covalently is blocked. This auto-inhibitory mechanism has been observed in three members of the Nedd4-family: Smurf2, Nedd4 and Nedd4L. **B)** In WWP1, WWP2 and Itch the C-lobe is locked in an inactive conformation and the UBS is block, but in this case the WW2 domain and the linker between the WW2 and WW3 domains interacts with the HECT domain.

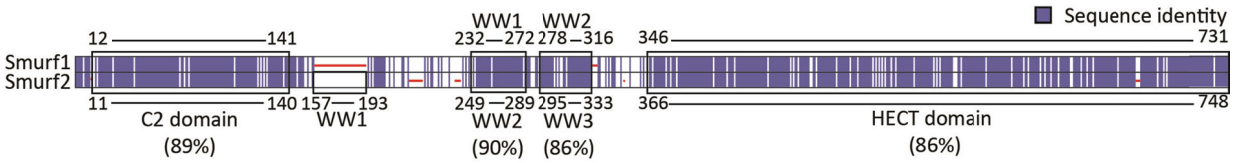


Figure 17. Sequence identity between Smurf1 and Smurf2. Identical residues are colored in blue. Alignment gaps are indicated with red lines. The regions comprising the C2, WW and HECT domains are highlighted with black boxes and the corresponding residue numbers are shown. In brackets the percentage of sequence identity is shown for each domain.

All previous studies demonstrate that despite the shared domain architecture and significant sequence similarity, Nedd4 family ligases have evolved distinct mechanisms of regulation. Smurf1 and Smurf2 share 70% sequence identity for the FL proteins with the major difference being the lack of the first WW domain in Smurf1 (Figure 17).
Considering only the C2 and HECT domains, the level of sequence identity between Smurf1 and Smurf2 rises to even almost 90% suggesting that the auto-inhibitory C2:HECT interaction should be conserved in Smurf1. However, reports on the catalytic activity in the literature disagree on Smurf1 activity regulation. In one study, Smurf1 is reported to be inhibited by a C2-mediated interaction (Wan et al., 2011), while in other studies Smurf1 seems to be fully active under steady-state conditions (Courivaud et al., 2015; Lu et al., 2011; Mund et al., 2015). My goal in this second project was to understand in detail the mechanism of Smurf1 regulation.

4.2.3 Results

To investigate whether the C2:HECT interaction is present in Smurf1, I performed CSP experiments with an ILVAM (Ile δ_1, Leu δ_1, Val γ, Ala β and Met ε [$^{13}CH_3$]), but otherwise U-[^{2}H,^{12}C])-labeled Smurf1 HECT domain. To examine whether the Smurf1 C2 domain interacts with the HECT domain in *trans* (as two independent molecules in contrast to *cis* where the interaction is intramolecular as in the FL protein), I recorded 2D ^{1}H,^{13}C-correlation (HMQC-TROSY) spectra of the ILVAM-labeled HECT domain alone and in the presence of unlabeled Smurf1 C2 domain (Figure 18, S1A).

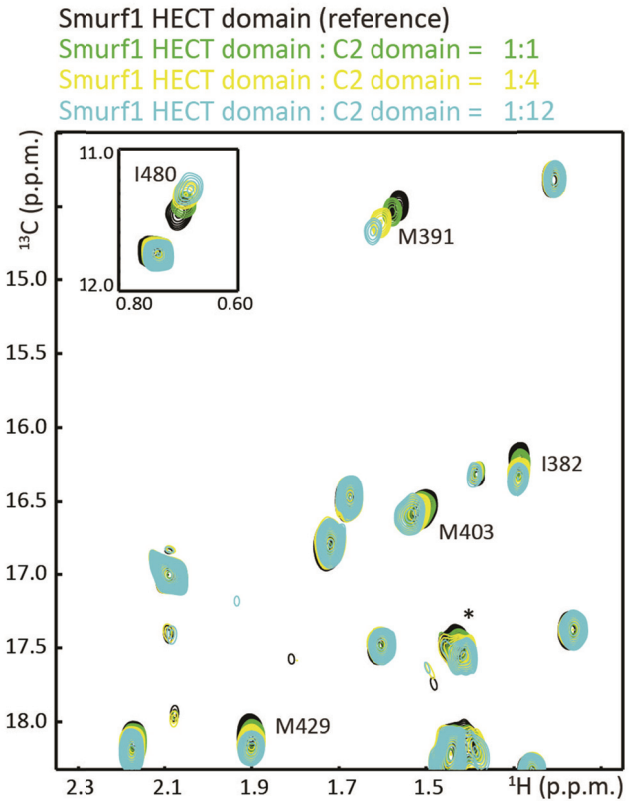


Figure 18. The Smurf1 HECT domain interacts with the C2 domain. Overlay of selected regions of ^{1}H,^{13}C- ILVAM methyl TROSY spectra of -labeled Smurf1 HECT domain in the absence (black) and presence of increasing amounts of Smurf1 C2 domain (one-fold (green), four-fold (yellow) and 12-fold (cyan) stoichiometric excess). Residues involved in the interaction are assigned and labeled. The resonance marked with an asterisk is an unassigned residue. According to the resonance position in the spectra it is probably an Ala.

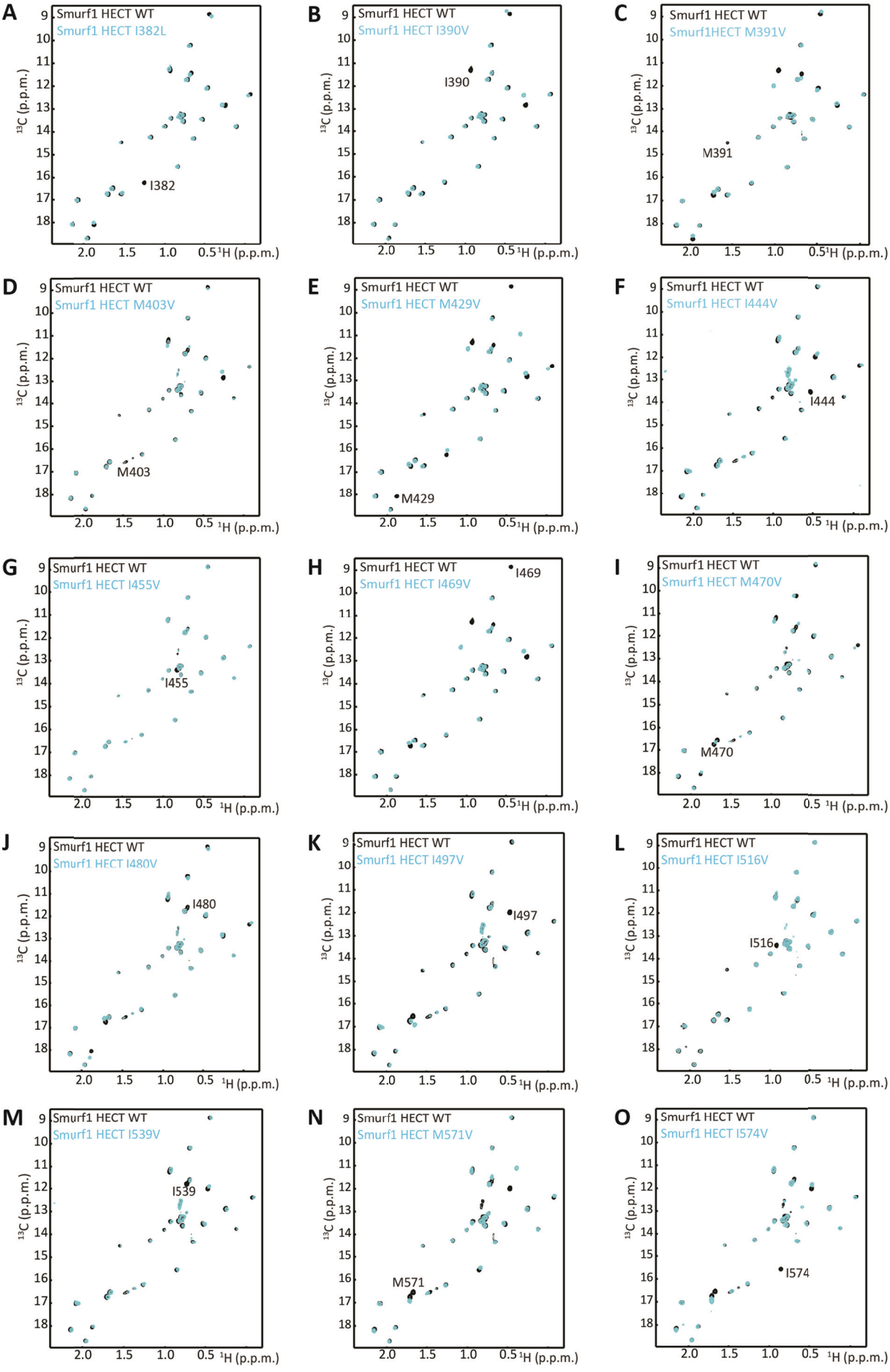

A
Smurf1 HECT WT
Smurf1 HECT I382L
I382
B
Smurf1 HECT WT
Smurf1 HECT I390V
I390
C
Smurf1 HECT WT
Smurf1HECT M391V
M391
D
Smurf1 HECT WT
Smurf1 HECT M403V
M403
E
Smurf1 HECT WT
Smurf1 HECT M429V
M429
F
Smurf1 HECT WT
Smurf1 HECT I444V
I444
G
Smurf1 HECT WT
Smurf1 HECT I455V
I455
H
Smurf1 HECT WT
Smurf1 HECT I469V
I469
I
Smurf1 HECT WT
Smurf1 HECT M470V
M470
J
Smurf1 HECT WT
Smurf1 HECT I480V
I480
K
Smurf1 HECT WT
Smurf1 HECT I497V
I497
L
Smurf1 HECT WT
Smurf1 HECT I516V
I516
M
Smurf1 HECT WT
Smurf1 HECT I539V
I539
N
Smurf1 HECT WT
Smurf1 HECT M571V
M571
O
Smurf1 HECT WT
Smurf1 HECT I574V
I574
13C (p.p.m.)
1H (p.p.m.)

Figure 19. Assignment of Ile δ_1 and Met ε methyl resonances through point mutations. Overlay of the $^{1}H,^{13}C$-methyl TROSY spectra of IM-labeled Smurf1 WT HECT domain (black) and the indicated HECT mutants (cyan) used for resonance assignment. The disappearing peaks are labeled.

This revealed significant CSPs for a number of methyl groups in the Smurf1 HECT domain and demonstrates that the Smurf1 HECT domain can interact with the C2 domain in *trans*. To map the observed CSPs onto the structure of the Smurf1 HECT domain, I generated a homology model based on the sequence of the ~86% identical Smurf2 HECT domain (PDB ID: 1ZVD) and assigned approx. 85% of the IM-methyl peaks in the HMQC spectrum. The rest of the peaks could not be assigned due to spectra overlap. All Ile and Met residues on the HECT domain (29 residues) were mutated to Leu or Val *via* point mutations and the disappearance of the corresponding peak was used to assign the residue (**Figure 19**). For those spectra where more than one peak displayed CSP the assignment choice was done considering unambiguous assignment of the rest of the peaks. With this in hand, I could show that the C2 domain interacts with the HECT domain on the same surface on the N-lobe as in Smurf2 (**Figure S1C**) (**Figure 20**).

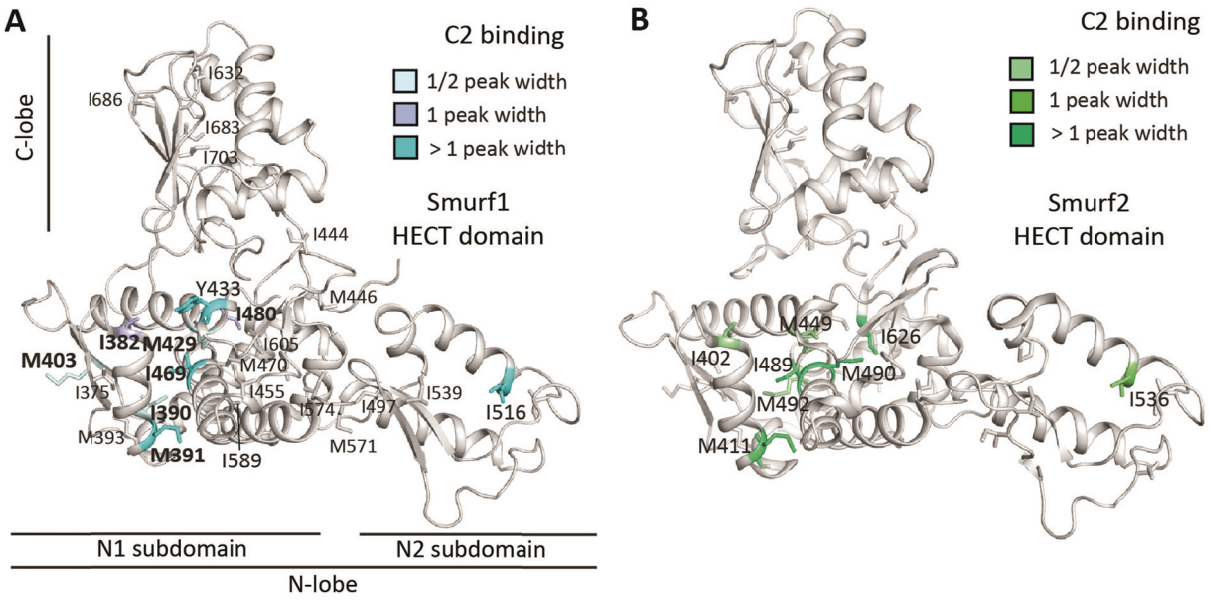


Figure 20. The Smurf1 HECT domain interacts with the C2 domain on the same surface as in Smurf2. A) Chemical shift mapping on a ribbon representation of the Smurf1 HECT homology model that was generated based on the crystal structure of the Smurf2 HECT domain (PDB ID: 1ZVD). Residues exhibiting CSPs upon C2 domain binding are labeled in bold and color code form pale cyan to cyan according to the increasing size of the CSP observed. Other assigned peaks are labeled in regular font. **B)** Similar to **A**, but for Smurf2, displaying only the peaks involved in the interaction, color-coded from pale green to dark green.

To substantiate this result, I introduced a Y433M mutation in the Smurf1 HECT domain that is located within the Smurf2 C2 binding surface as an additional NMR reporter. The additional peak resulting from the mutation was easily identified by comparison with the WT spectrum (Figure 21A). Addition of the C2 domain to this mutant resulted in an observable CSP of the peak corresponding to the Y433M mutation (Figure 21B) confirming the mapped C2 domain binding surface on the Smurf1 HECT domain.

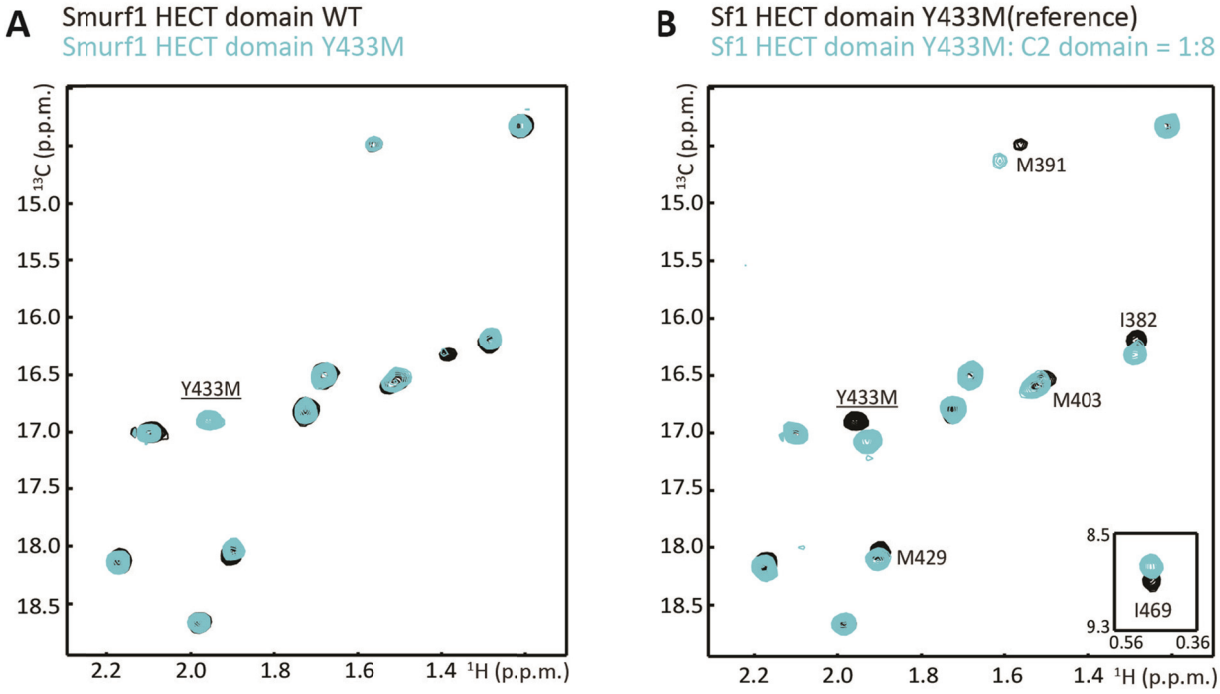


Figure 21. Introduction of a reporter Met confirms the C2 binding surface on the Smurf1 HECT domain. (A) Overlay of the Met region of the ^{1}H,^{13}C-methyl TROSY spectra of IM-labeled Smurf1 WT HECT domain (black) and the Smurf1 HECT Y433M mutant showing the additional peak arising from the Met mutation (cyan). **B)** As **A**, but for the Smurf1 HECT Y433M mutant in the absence (black) or presence of an eight-fold stoichiometric excess of C2 domain (cyan).

To further characterize the interaction between the Smurf1 C2 and HECT domains, the HECT binding surface on the C2 domain was mapped. To do so, a 2D ^{1}H,^{15}N-correlation spectrum of a ^{15}N-labeled Smurf1 C2 domain sample was recorded (Figure 22A). Since the assignments were available in our group (unpublished data from Christine Wolf) I was able to identify the residues showing CSPs upon stepwise addition of unlabeled HECT domain. The interaction with the HECT domain occurs through hydrophobic and charged residues in the Smurf1 C2 domain located in the β1-β2 and β3-β4 loop, the αI-helix and the αI-β6 loop (Figure 22B) with F30, F31, R32 and L58 displaying the largest CSPs (Figure 22A). This surface matches exactly the one described for Smurf2 (Wiesner et al., 2007) (Figure 22C), once again showing how the C2:HECT interaction is conserved in these two proteins.

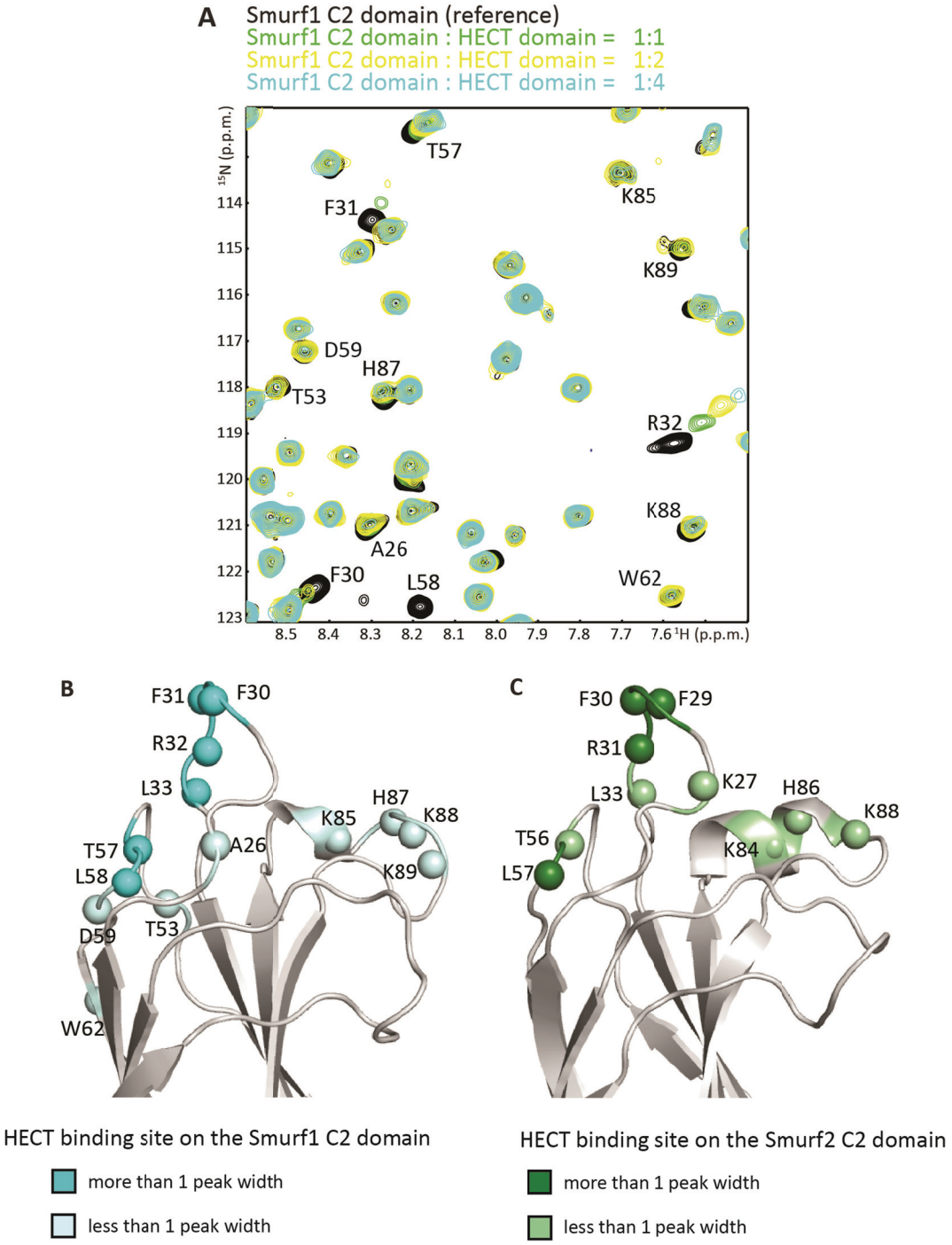


Figure 22. Chemical shift mapping of the Smurf1 C2:HECT interaction. **A)** Overlay of selected regions of $^{1}H,^{15}N$-HSQC spectra of Smurf1 C2 domain in the absence and presence of increasing amounts of Smurf1 HECT domain (one-fold (green), four-fold (yellow) and 12-fold (cyan) stoichiometric excess). The titration experiment was performed by Silke Wiesner. **B)** Chemical shift mapping of the Smurf1 C2 domain (PDB ID: 3PYC) binding site on the HECT domain structure. Spheres represent the nitrogen atoms of affected residues. **C)** As **B** but for Smurf2 (PDB ID: 2JQZ) (data from Wiesner et al., 2007).

As mentioned before, numerous HECT domains contain a UBS that promotes Ub chain elongation and, in the case of Smurf2 and Nedd4, overlaps with the C2 interaction site blocking poly-ubiquitination activity. I therefore examined whether Smurf1 also interacts with Ub in a non-covalent manner. Indeed, the IM-labeled Smurf1 HECT domain exhibited numerous CSPs upon addition of unlabeled monomeric Ub, demonstrating that Smurf1 can interact with Ub in a non-covalent manner (Figure 23A, S1B). I mapped the residues displaying CSPs onto the Smurf1 HECT homology model and could show that the same residues are involved in Ub interaction as in Smurf2 (Figure 23B-C, S1D). Moreover, as in Smurf2, many residues (I382, M391, M403, M429, I469 and I480) are involved in both Ub and C2 domain binding demonstrating that these binding surfaces overlap on the Smurf1 HECT domain.

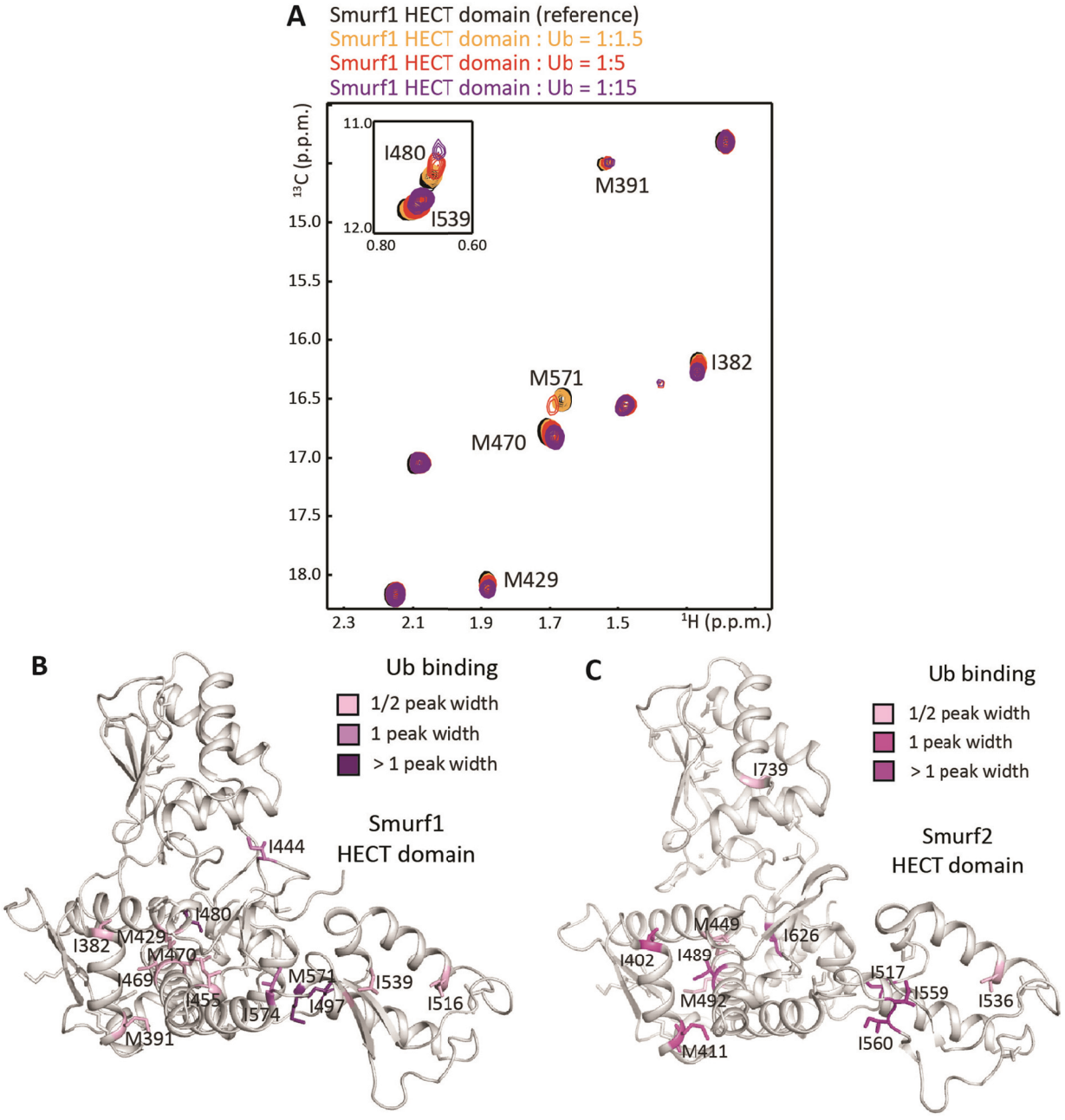


Figure 23. The Smurf1 HECT domain interacts with Ub on the same surface as Smurf2. A) Overlay of selected regions of $^{1}H,^{13}C$-methyl TROSY spectra of IM-labeled Smurf1 HECT domain in the absence (black) and presence of increasing amounts of Ub domain (1.5-fold (orange), five-fold (red) and 15-fold (purple) stoichiometric excess). Residues involved in the interaction are assigned and labeled. **B)** Chemical shift mapping of the Ub binding surface on the homology model of the Smurf1 HECT domain. Residues displaying CSPs upon Ub binding are color code form light pink to violet according to the increasing size of the CSP observed. **C)** As **B** but for Smurf2 (PDB ID: 1ZVD), color-coded from light pink to magenta.

I could thus show that the Smurf1 C2 and the HECT domains interact in *trans* and that their binding surfaces are highly conserved in comparison to Smurf2. The same holds true for the interaction with Ub. To obtain quantitative insights into these interactions, I determined dissociation constants (K_d values) for the C2 domain and Ub interactions with the Smurf1 and Smurf2 HECT domains by NMR 2D lineshape fitting, using the chemical shift titrations discussed above (**Figure S1**). For each of the titration experiments, 3D plots of the fitting were generated and a bootstrap analysis to determine the error was performed. One significant peak for Smurf1 C2 and Ub interaction (I382) was chosen as an example of the analysis done. A side-by-side view with the observed CSPs and the fits is shown in **Figure 24A** and **C**; while the 3D plots of the peaks are depicted in **Figure 24E** and **G**. Identical plots are shown for Smurf2 residue I402 in **Figures 24B, D, F** and **H**. The K_d values obtained are reported in the **Table 3**.

Table 3. Dissociation constants (K_d) obtained for the interactions of the Smurf1 and Smurf2 HECT domains with the respective C2 domain and Ub. The K_d values are expressed in µM.

Ligand	Sf1 HECT	Sf2HECT
Ubiquitin	302 ± 11	393 ± 2
C2 domain	347 ± 10	269 ± 6

This quantitative analysis showed that Smurf1 binds with almost the same affinity to the C2 domain than to Ub, while Smurf2 seems to bind slightly stronger to the C2 domain than to Ub. More importantly, the values for Smurf1 and Smurf2 are highly similar. To summarize, as suggested by the high level of sequence identity of the Smurf1 and Smurf2 C2 and HECT domains, I found that the binding capabilities of the HECT domains in terms of residues involved and their affinities towards Ub and the respective C2 domains are conserved between Smurf1 and Smurf2. It should be noted, however, that the C2:HECT interactions in the FL E3s can be expected to be significantly stronger than those with Ub, since the former are intramolecular interactions.

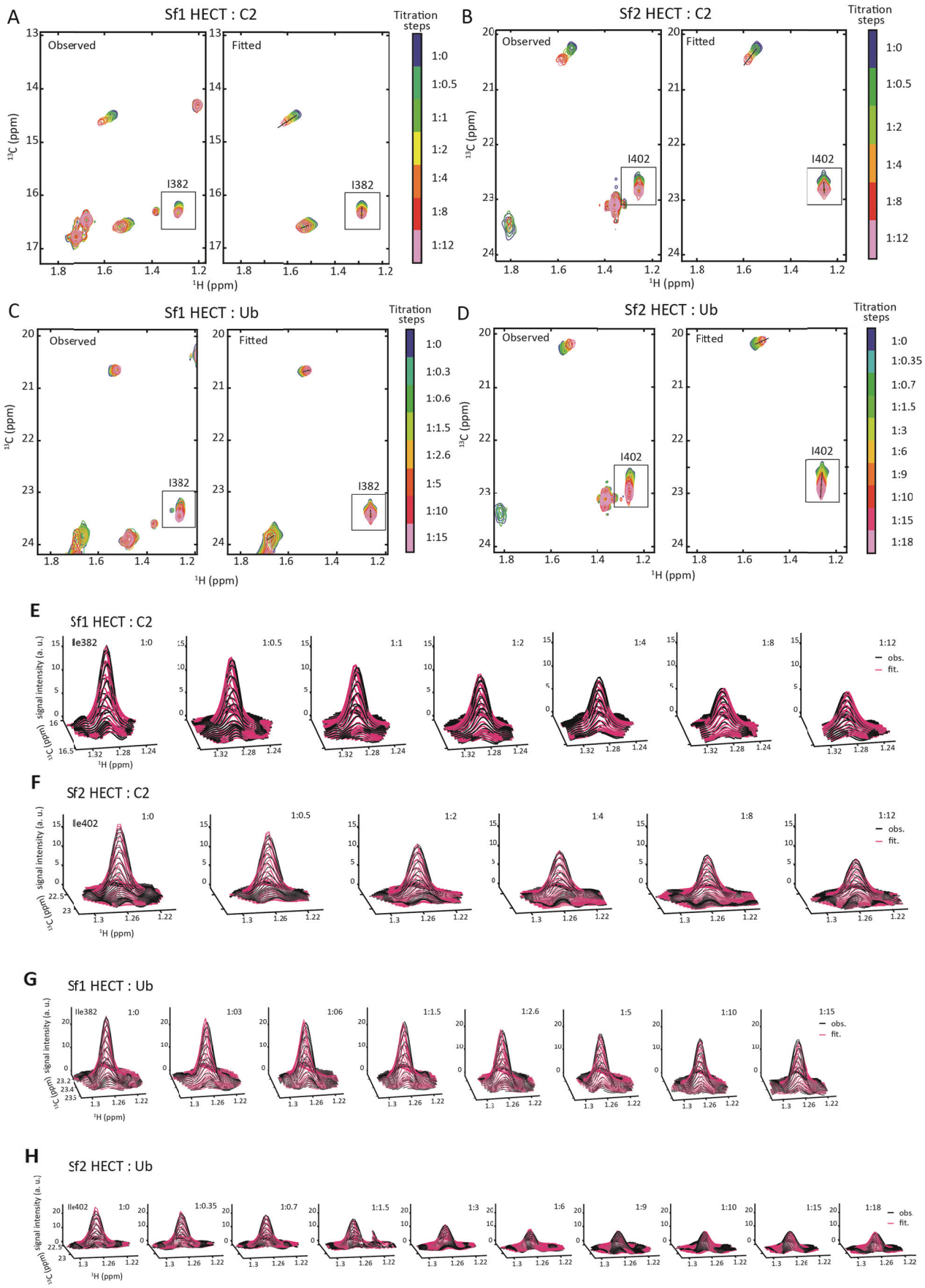

A
Sf1 HECT : C2
Observed
Fitted
I382
Titration steps
1:0
1:0.5
1:1
1:2
1:4
1:8
1:12
^{13}C (ppm)
^{1}H (ppm)
B
Sf2 HECT : C2
I402
C
Sf1 HECT : Ub
1:0.3
1:0.6
1:1.5
1:2.6
1:5
1:10
1:15
D
Sf2 HECT : Ub
1:0.35
1:0.7
1:3
1:6
1:9
1:18
E
Sf1 HECT : C2
Ile382
signal intensity (a. u.)
obs.
fit.
F
Sf2 HECT : C2
Ile402
G
Sf1 HECT : Ub
H
Sf2 HECT : Ub

Figure 24. 2D line shape fitting analyses of the Smurf1 or Smurf2 HECT interactions with their respective C2 domain or Ub. A) Representative region of the $^{1}H,^{13}C$-methyl TROSY titration experiment of the Smurf1 HECT domain with the C2 domain (**Figure S1A**). The observed as well as the fitted data are indicated in the left and right spectrum as indicated. Titration steps are color coded as indicated in the figure. The δ_1-methyl resonance of Ile382, which is displayed in the panels below, is boxed. **B)** As **A**, but for the Smurf2 HECT:C2 interaction using the δ_1-methyl resonance of Ile402 as an example (**Figure S1C**). **C)** As **A**, but for the Smurf1 HECT:Ub interaction (**Figure S1B**). **D)** As **A**, but for the Smurf2 HECT:Ub interaction (**Figure S1D**). **E)** 2D line shape fit for the $^{1}H,^{13}C$-methyl TROSY titration experiment of the Smurf1 HECT domain and C2 domain. Each titration point is shown for the δ_1-methyl resonance of Ile382; black indicates the observed data whereas pink indicates the fit. **F)** As **E**, but for the interaction between the Smurf2 HECT and C2 domains, displaying the observed and fit data for Ile402. **G)** As **E**, but for the interaction between Smurf1 HECT domain and Ub. **H)** As **F**, but for the interaction between the Smurf2 HECT domain and Ub.

As discussed in the first chapter of the thesis (4.1), Smurf2 is kept in a closed conformation due to an intramolecular interaction between the C2 and the HECT domains. From the extremely similar way in which the Smurf1 and Smurf2 HECT domains interact with their respective C2 domains and with Ub, it can be expected that the C2 domain also plays an inhibitory role in Smurf1. However, previous reports showed that Smurf1 FL is active (Courivaud et al., 2015; Lu et al., 2011; Mund et al., 2015; Wan et al., 2011).

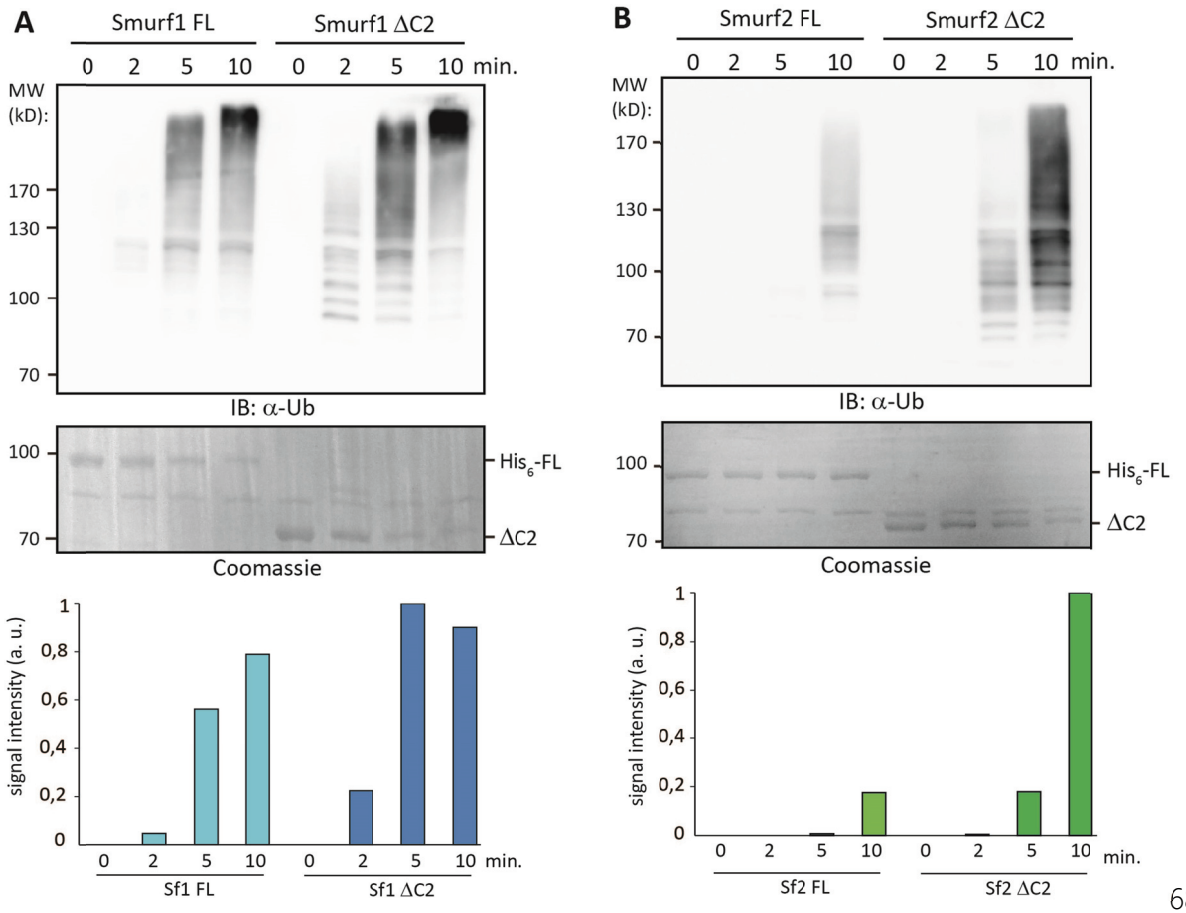


Figure 25. The C2 domain has only a minor effect on the activity of FL Smurf1. *In vitro* ubiquitination assays using WT and ΔC2 Smurf1 **(A)** and Smurf2 **(B)** proteins. All proteins were bacterially expressed and purified. Auto-ubiquitination activity was detected by immunoblotting (IB) with α-Ub antibody (top panels). Initial protein levels were confirmed by Coomassie staining (bottom panels). For each western blot, a densitometry analysis was done to quantify the observed signal. While the levels of ubiquitinated FL Smurf1 are only 10% less than for the ΔC2 variant, ΔC2 Smurf2 is almost six times more active than FL Smurf2.

To investigate whether the C2 domain can exert an auto-inhibitory effect in Smurf1, I analyzed the auto-ubiquitination activities of recombinant WT Smurf1 and a ΔC2 mutant and compared them to the corresponding Smurf2 proteins. Interestingly, I observed that at the time point of 10 min., the ΔC2 Smurf1 was only slightly (12%) more active than the FL protein (Figure 25A), while deletion of the C2 domain in Smurf2 resulted in a drastic increase in auto-ubiquitination activity (Figure 25B), being both results consistent with previous reports (Lu et al., 2011; Mari*, Ruetalo* et al., 2014; Wiesner et al., 2007). Therefore, despite the fact that Smurf1 C2 domain can interact with the HECT domain in *trans* (Figure 18), it only exerts a minor inhibitory effect on the Smurf1 FL activity.
To examine whether the Smurf1 C2 domain has the potential to modulate HECT domain activity in *trans*, I performed competition assays using Smurf1 HECT domain in the absence and presence of increasing amounts of C2 domain. The results showed that although the C2 domain does not regulate the activity of Smurf1 in the FL context, the C2 domain is indeed able to inhibit HECT domain activity in *trans* to a similar degree as in Smurf2 (Figure 26). This demonstrates that the Smurf1 C2 domain can interact and inhibit HECT domain activity in *trans*, but cannot exert an auto-inhibition in *cis*.

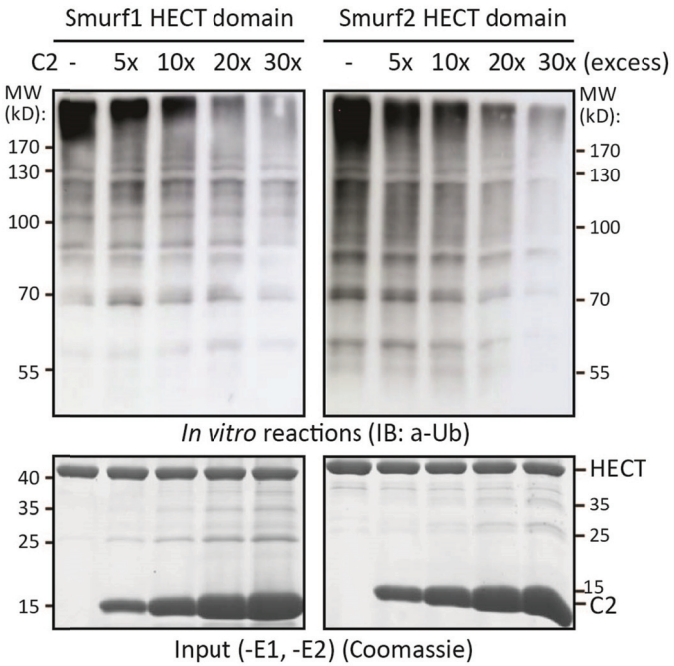


Figure 26. Smurf1 C2 domain is able to inhibit HECT domain activity in *trans*. *In vitro* auto-ubiquitination assay using Smurf1 (left panel) and Smurf2 (right panel) HECT domains in the absence and presence of increasing stoichiometric amounts of the respective C2 domain as indicated. Reaction samples were separated by SDS-PAGE and auto-ubiquitination activity detected by immunoblotting (IB) with α-Ub antibody (top panels). Protein levels in the starting material were confirmed by Coomassie staining (bottom panels). The level of inhibition showed by the Smurf1 C2 domain on the HECT activity is comparable to the inhibition exerted by Smurf2 C2 on its HECT domain.

Since the major difference between Smurf1 and Smurf2 lie in the region linking the C2 and HECT domains, the results suggested that this region plays a role in regulating Smurf activity. To assess the importance of this region in mediating auto-inhibition, I designed chimeric Smurf1 and Smurf2 proteins (Figure 27) and examined their auto-ubiquitination activities.

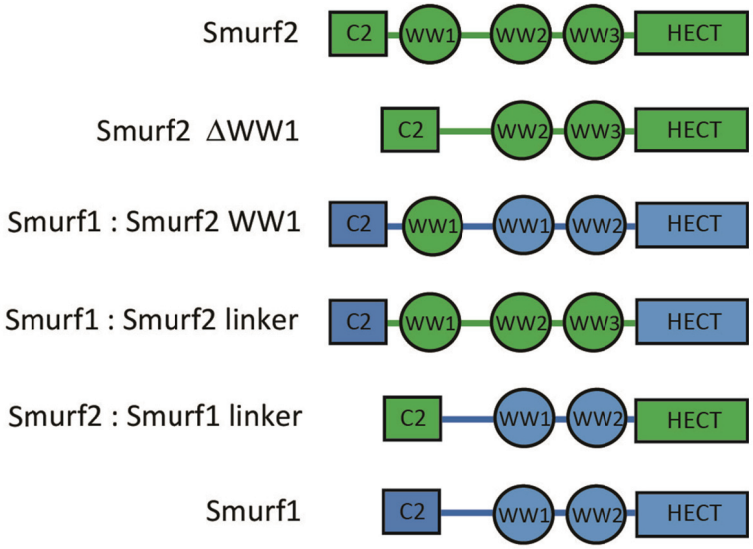


Figure 27. Schematic representation of the chimeric Smurf1 and Smurf2 proteins used for auto-ubiquitination assays. Proteins are displayed according to their increasing auto-ubiquitination activity (top to bottom). Green: Smurf2. Blue: Smurf1.

To this end, I generated a Smurf1 variant that contained the linker region of Smurf2 and *vice versa*. Interestingly, I observed that substitution of the Smurf1 linker with that of Smurf2 led to a significant loss in auto-ubiquitination activity as compared to WT Smurf1 and thus enabled Smurf1 auto-inhibition (Figure 28A). Accordingly, Smurf2 with a Smurf1 linker exhibited a strikingly higher auto-ubiquitination activity than WT Smurf2 demonstrating that the Smurf1 linker region does not allow for auto-inhibition (Figure 28B). This emphasizes that the Smurf1 C2 domain is capable of inhibiting HECT domain activity, but that the Smurf1 linker region does not allow for the C2 and HECT interaction to occur in Smurf1. Since the most significant difference between the Smurf1 and Smurf2 is the presence of an additional WW domain (WW1) in Smurf2 (Figure 17), I examined the

importance of the WW1 domain for Smurf auto-inhibition. To this end, I introduced the Smurf2 WW1 domain in Smurf1 and generated a Smurf2 variant lacking the WW1 domain (ΔWW1) (Figure 27). Auto-ubiquitination assays with these proteins showed that the presence of the Smurf2 linker enables auto-inhibition of FL Smurf1 (Figure 28A), both in the case where the entire linker was exchange or when only the WW1 was deleted. Accordingly, deletion of the WW1 domain led to Smurf2 activation, although the level of activation was lower than for Smurf2 carrying the Smurf1 WW linker region (Figure 28B). In brief, these results demonstrate that the region linking the C2 and HECT domains, and in particular the Smurf2 WW1 domain, play an important role in mediating Smurf2 auto-inhibition.

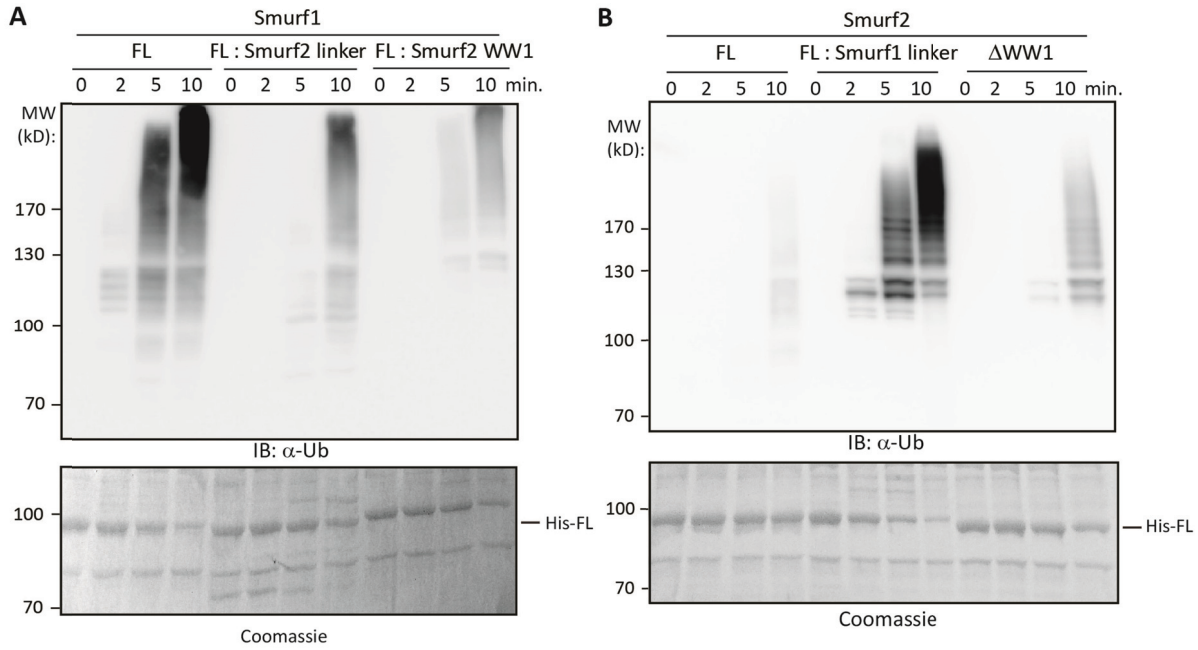


Figure 28. The region connecting the C2 and HECT domains plays a role in Smurf2 auto-inhibition. *In vitro* auto-ubiquitination assays using the indicated bacterially expressed and purified WT and chimeric Smurf1 and Smurf2 proteins. **(A)** Smurf1 carrying the Smurf2 linker or the Smurf2 WW1 domain is auto-inhibited. **(B)** The introduction of the Smurf1 linker in Smurf2 or deletion of the WW1 domain results in Smurf2 activation. Otherwise as Figure 25.

Collectively, my results demonstrate a role for the WW1 domain in enhancing an intramolecular C2:HECT interaction and thereby mediating Smurf2 auto-inhibition. However, it was previously shown that the Smurf2 WW1 domain does not interact directly with the HECT domain (Wiesner et al., 2007). Moreover, the crystal structure of the WW2-WW3-HECT construct did not reveal sufficient electron density to resolve the WW domains (Ogunjimi et al., 2005). Therefore, the WW2-3 tandem does not seem to adopt a fixed orientation with respect to the HECT domain in the WW2-3-HECT construct.

To uncover the mechanism by which the WW1 domain in Smurf2 was mediating auto-inhibition, whether the Smurf2 C2 domain interacts with the WW1 domain was examined. However, no chemical shift changes in the 2D $^{1}H,^{15}N$-correlation spectra of the C2 domain upon addition of a 12-fold excess of the WW1 domain were detected (Figure 29). This demonstrates that the C2 domain and the WW1 domain do not interact in *trans*.

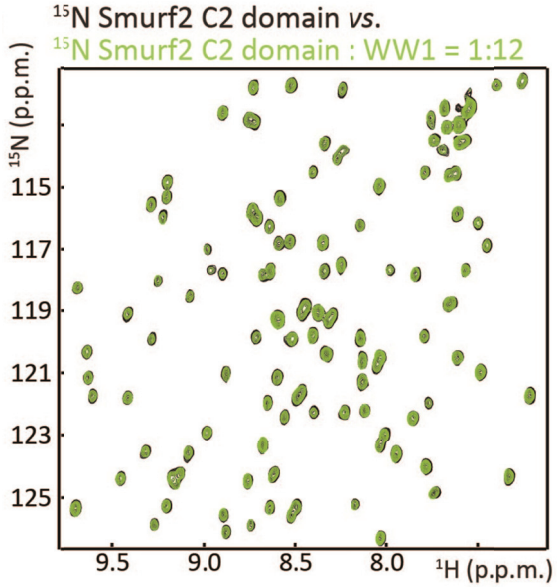


Figure 29. Smurf2 C2 domain cannot interact directly with the WW1 domain. Overlay of a representative region of the $^{1}H,^{15}N$-correlation spectra of ^{15}N-labeled Smurf2 C2 in the absence (black) or presence of a 12-fold stoichiometric excess of WW1 domain (light green). No significant CSPs are observed showing that the WW1 domain does not interact with the C2 domain. The titration experiment was performed by Samira Anders.

In conclusion, all these data suggest that the Smurf2 WW domains on their own do not interact with the HECT domain and thus may rather aid in positioning the C2 domain for HECT domain interaction. To test this hypothesis, a Smurf2 construct comprising the C2 domain, the linker and the WW1 domain (C2-linker-WW1) was generated and subsequently 2D $^{1}H,^{13}C$-correlation spectra of the IM-labeled Smurf2 HECT domain were recorded in absence and presence of C2-linker-WW1 construct (Figure 30A, S1E). Interestingly, I observed that overall the same residues experienced CSPs as upon addition of the C2 domain alone. However, the changes of the chemical shifts were significantly larger than with the C2 domain alone. Moreover, for some peaks such as M411 and I489, disappeared at intermediate titration steps (1:0.5 and 1:1 or 1:1 respectively) re-appearing close to saturation. While the C2:HECT interaction was clearly in the fast exchange regime for the C2 domain alone, the C2-linker-WW1 construct interacted with the HECT domain in the intermediate exchange regime. To determine the K_d of the C2-linker-WW1:HECT interaction, I used 2D NMR line shape fitting as describe above. The fitting of the same residue used in the case of the C2 domain analysis (Ile402) as well as the 3D plots are

shown in **Figure 30B** and **C** respectively. The K_d for the Smurf2 HECT:C2-linker-WW1 interaction, reported in **Table 4**, was 15 µM, in comparison to the 269 µM K_d obtained for HECT:C2 interaction. This shows that the presence of the linker between the C2 and the WW1 domain (referred to hereafter as C2-WW1L) and the WW1 domain enhances the C2:HECT interaction by at least 20-fold.

Table 4. Dissociation constant (K_d) obtained for the interaction of the Smurf2 HECT domain with the C2-linker-WW1 construct. The K_d value is expressed in µM.

Ligand	Sf2 HECT
C2-WW1 domain	15 ± 1

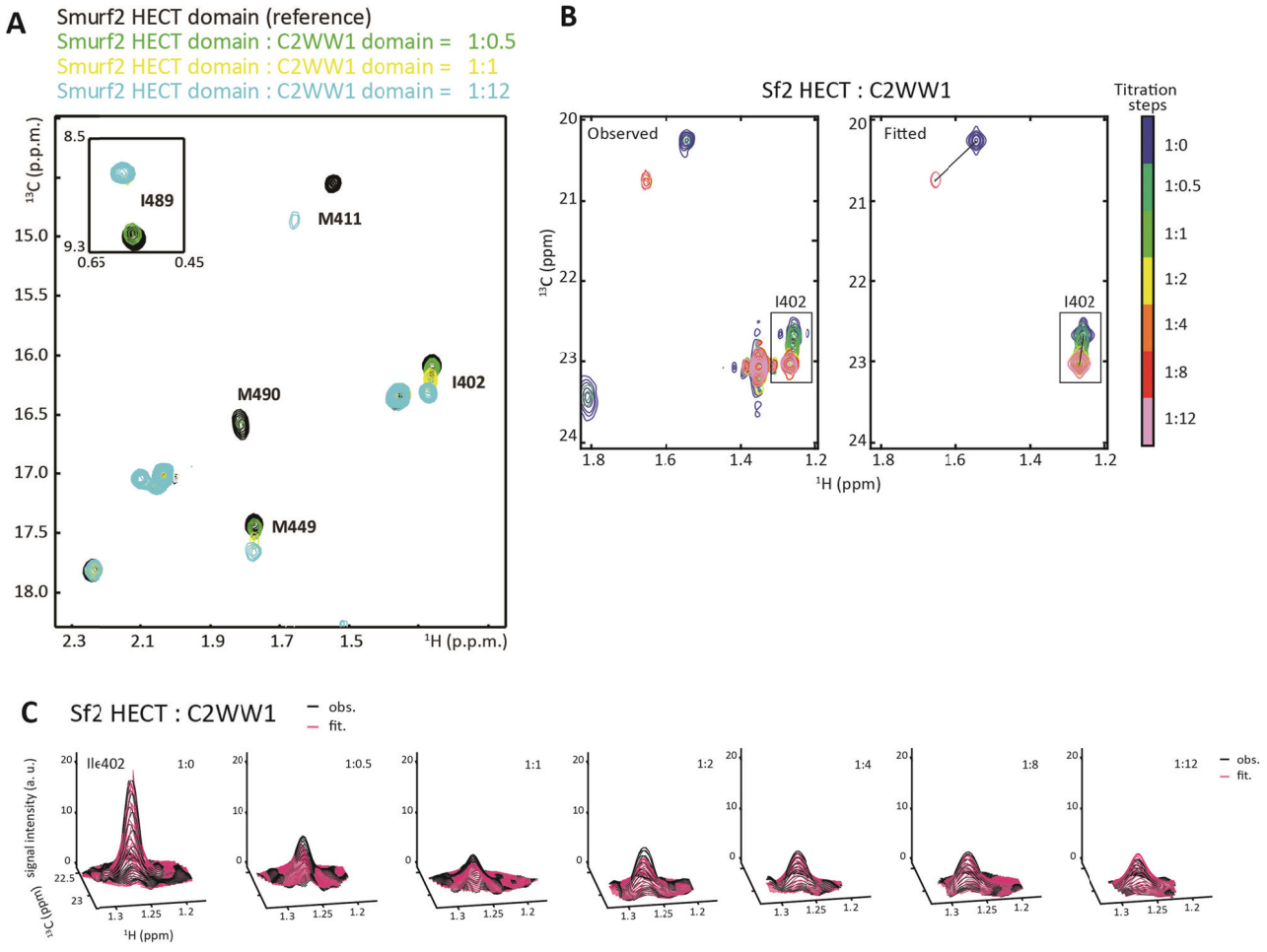


Figure 30. The Smurf2 HECT domain interacts stronger with the C2-linker-WW1 construct than with the C2 domain alone. A) Overlay of a selected region of the $^{1}H,^{13}C$-methyl TROSY titration experiment of the Smurf2 HECT domain in the absence (black) or presence of increasing amounts of the Smurf2 C2-linker-WW1 construct (half stoichiometric excess (green), one-fold (yellow) and 12-fold (cyan) stoichiometric excess of the C2-WW1 construct. The titration experiment was recorded by Magnus Jäckl. **B)**

Representative region of the $^{1}H,^{13}C$-methyl TROSY titration experiment of the Smurf2 HECT domain and C2-linker-WW1 construct (Figure S1E). The observed as well as the fitted data are shown on the left and on the right, respectively. Titration steps are color coded as indicated in the figure. The δ_1-methyl resonance of residue Ile402, which is displayed in the panels below, is boxed. C) 2D line shape fit for the $^{1}H,^{13}C$-methyl TROSY titration experiment of the Smurf2 HECT domain with the C2-linker-WW1. Each titration point for the δ_1-methyl resonance of Ile402; black indicates the observed data whereas pink indicates the fit.

Next, I aim to understand how the C2-WW1L and the WW1 domain mediate the affinity increase for the C2:HECT interaction. In order to do so, ^{15}N-HSQC NMR spectra of Smurf2 C2-linker-WW1, C2 domain and linker-WW1 were recorded (performed by Samira Anders). Assignments for the C2 domain spectrum were already available (Wiesner et al., 2007) and the NMR experiments used to assign the linker-WW1 construct were carried out by Silke Wiesner. The assignment of entire C2-linker-WW1 spectra was done by a "divide" and "conquer" approach, transferring sets of assignments both from the C2 domain and the linker-WW1 construct. By comparing the three spectra, I was able to detect CSPs appearing as a consequence of an intramolecular interaction inside the region C2-linker-WW1 (Figure 31A).

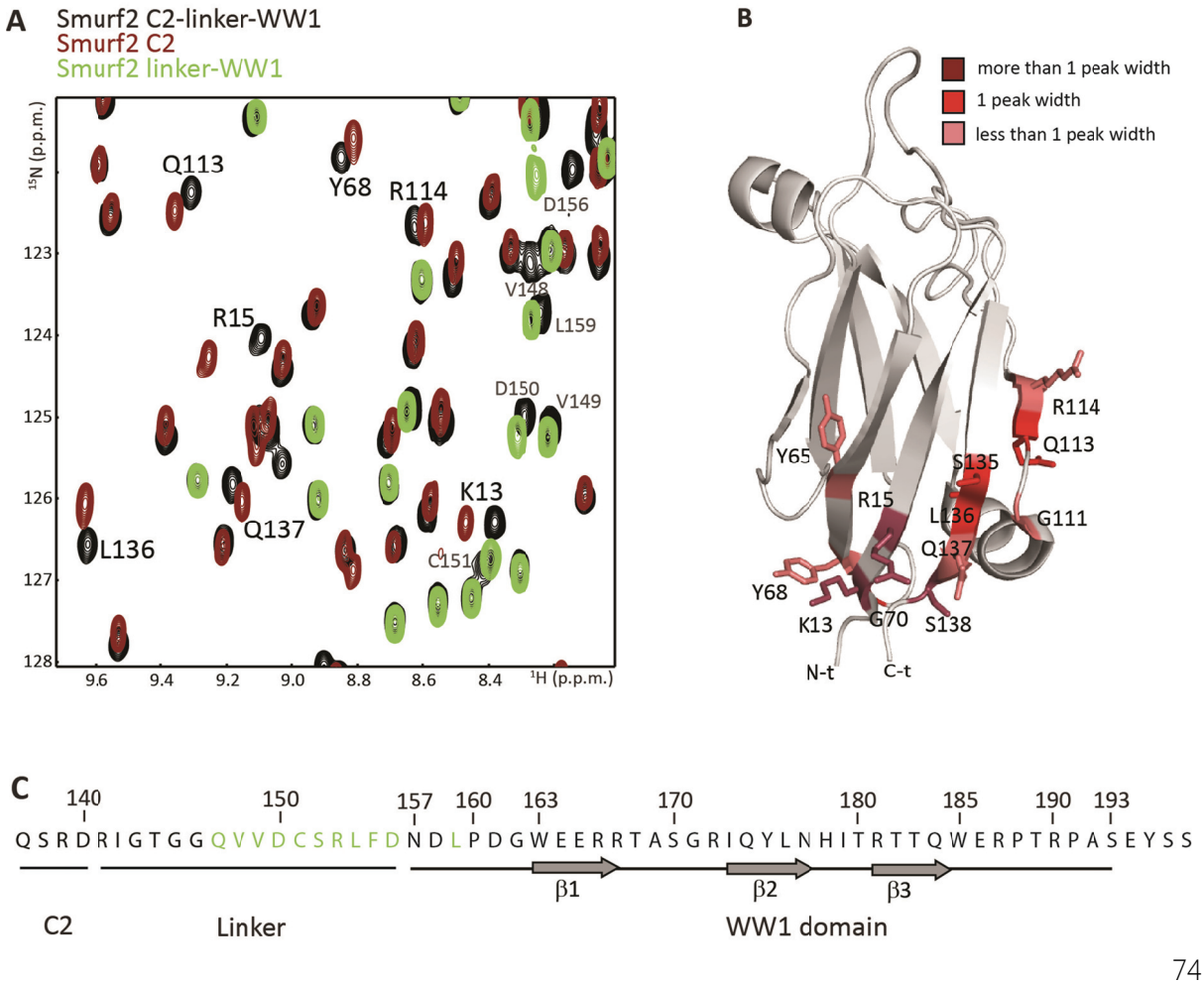


Figure 31. The C2-WW1L region interacts with the C2 domain in *cis*. A) Overlay of a selected region of three ^{15}N-HSQC NMR spectra: Smurf2 C2-linker-WW1 (black), C2 domain (dark red) and linker-WW1 (light green). The residues displaying CSPs, which belong to the C2 domain, are labeled in black while the ones which belong to the C2-WW1L region are labeled in grey with a smaller font size. The spectra in A were recorded by Samira Anders and the assignments were carried out by Silke Wiesner. B) The CSPs observed in A are mapped on the Smurf2 C2 domain structure (PDB ID: 2JQZ). They are labeled and color-coded from rose to dark red according to the increasing size of the CSPs observed. N-t and C-t are also indicated. C) The CSPs that mapped to the linker region are shown on the protein sequence. The last residues of the C2 domain, the C2-WW1L and the entire WW1 domain sequences are shown. The residues affected are colored in light green. The sequence bears a C151A mutation, which was introduced to improved spectra quality.

Analysis of the CSPs allowed me to map on the Smurf2 C2 domain structure the region that was affected by the linker-WW1 domain. The binding surface between the C2 domain and the linker-WW1 is shown in figure 31B. This region is located at the bottom of the C2 domain, in contrast to the HECT binding surface, which is located at the upper part of the domain. The residues involved in the interaction are arranged in four β strands: β1 (K13, R15), β4 (Y65, Y68), β7 (Q113, R114) and β8 (S135-S138), creating a binding surface that is situated at one side of the C2 domain. CSPs were also found to affect the C2-WW1L: 10 out of the 16 amino acids forming the C2-WW1L region displayed CSPs. These residues are found at the end of the linker, ranging from 147 to 156. In addition, one amino acid at the beginning of the WW1 domain was also shifting (Figure 31C). Altogether, this data showed that the C2 domain of Smurf2 interacts with the C2-WW1L region.
To examine the role of the WW1 domain and the C2-WW1L region on HECT domain activity, I performed competition assays where I compared auto-ubiquitination of the Smurf2 HECT domain in the absence and presence of increasing amounts of the C2-linker-WW1 construct and the C2 domain alone (Figure 32).

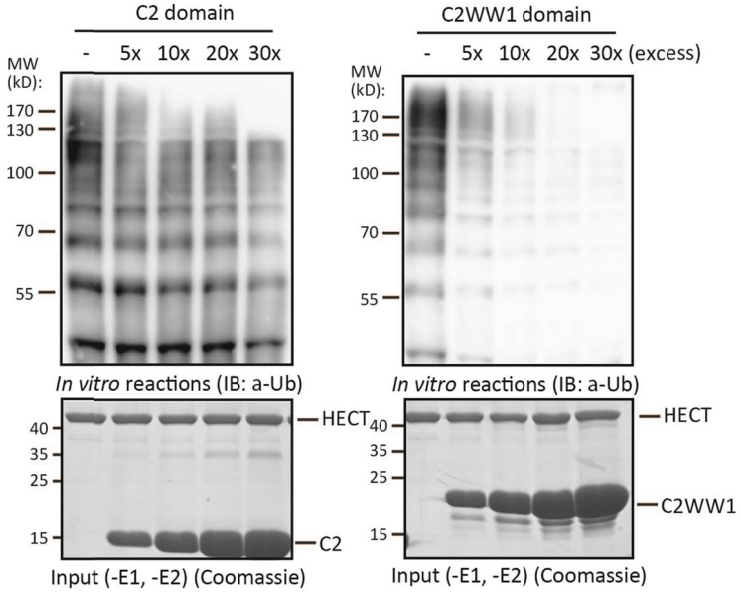


Figure 32. The WW1 domain and the C2-WW1L potentiate the C2-mediated inhibition of the Smurf2 HECT domain. *In vitro* auto-ubiquitination of the Smurf2 HECT domain in the absence and presence of a stoichiometric excess of the C2 domain (left panel) and the C2-linker-WW1 construct (right panel) as indicated. All proteins were bacterially expressed and purified. Reactions were prepared using fluorescent-labeled Ub that was used for detection. Samples were separated by SDS-PAGE (top panels). Protein levels in the starting material were confirmed by Coomassie staining (bottom panels). The inhibition effect exerted by the C2 domain on Smurf2 HECT activity is enhanced significantly by the presence of the linker-WW1 region.

Consistent with the increased affinity of the C2-linker-WW1 construct for the HECT domain (**Table 3** vs. **Table 4**), the auto-ubiquitination signal dropped to almost undetectable levels for the C2-linker-WW1 construct already at a five-fold stoichiometric excess, while the C2 domain at the same level had only a minor inhibitory effect on HECT domain activity. This confirms that the WW1 domain and the C2-WW1L region play an important role in mediating Smurf2 auto-inhibition.

Conclusions:

- The Smurf1 HECT domain interacts with the C2 domain in *trans* and contains a non-covalent UBS that overlaps with the C2 interaction surface.
- Both surfaces, the C2 binding surface and the UBS, are highly conserved between Smurf1 and Smurf2.
- The Smurf1 C2 domain is capable of inhibiting the HECT domain in *trans.* None-the-less, I found that the FL Smurf1 enzyme is not inhibited by a C2:HECT interaction.
- I found that the difference in regulation between Smurf1 and Smurf2 stems from the lack of the WW1 domain in Smurf1.
- Smurf2 WW1 domain interacts in *trans* neither with the C2 nor with the HECT domain. In contrast, addition of the C2-linker-WW1 constructs to the HECT domain shows a significantly decrease in HECT activity in comparison with the effect exerted by the C2 alone.
- The C2-WW1L region and the WW1 domain mediate an intramolecular interaction with the C2 domain and thereby increase the affinity of the C2 domain for the HECT domain (**Figure 33**).

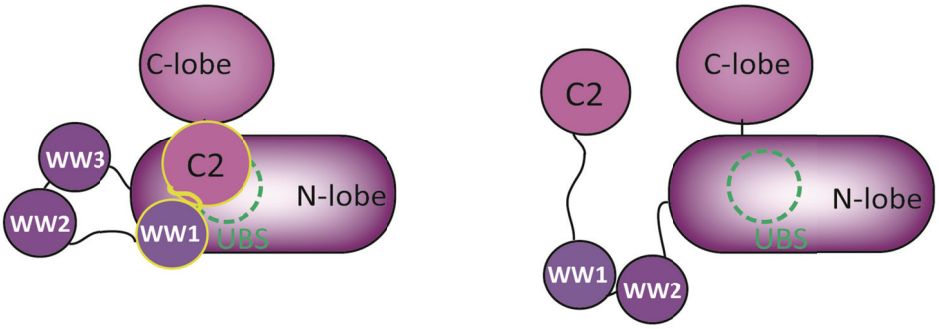


Figure 33. Model of the WW1 region in mediating Smurf2 auto-inhibition. The C2-WW1L region enables a closed, auto-inhibited Smurf2 conformation where the presence of the WW1 domain increases the affinity of the C2:HECT domain interaction, which binds on a surface overlapping with the UBS (left). The lack of the WW1 domain in Smurf1 prevents the C2 domain from interacting with the HECT domain in *cis*, resulting in an open and constitutively active Smurf1 enzyme (right).

4.2.4 Discussion

The mechanisms underlying auto-inhibition have been investigated for numerous Nedd4 family members. This revealed that both, the C2 domain and the WW domains, can mediate down-regulation of ligase activity. Surprisingly, despite their common domain architecture and high levels of sequence conservation within the C2, WW and HECT domains, the mode of auto-inhibition seems to be distinct for the individual Nedd4-family members. Unfortunately, structural analysis of FL Nedd4 ligases has so far been precluded due to the large linker regions that connect the C2, WW and HECT domains. In the case of the highly related Ub ligases Smurf1 and Smurf2, I found that despite sharing ~70% sequence identity, Smurf1 is not down-regulated by a C2-HECT interaction. This is the case even though the Smurf1 C2 domain can interact with the HECT domain and inhibit its activity in *trans*. This differential regulation can be attributed to the presence of an additional WW domain (WW1) in Smurf2.

Most eukaryotic proteins consist of domains that individually enable interactions and / or catalytic activity. This modular protein architecture lies at the heart of all signal transduction pathways. Although protein domains are in principle structurally and functionally independent, their activities can be modulated through intramolecular interactions with other domains or linker regions. While the knowledge on how linker regions regulate the activity of catalytic domains is still limited, numerous studies have elucidated the mechanisms underlying the modulation of enzymatic activity by interaction domains. In a subset of Nedd4 ligases, the N-t C2 domain interacts directly with the

catalytic HECT domain to inhibit ubiquitination activity (Figure 16A). In others, the linker region in between the central WW domains influences ligase activity interacting extensively with the N and the C-lobe of the HECT domain, which impairs E2-E3 transthiolation (Chen et al., 2017; Zhu et al., 2017) (Figure 16B). I found that the WW1 domain contributes to the inhibition of Smurf2 activity. Although it was already known that the Smurf2 C2 domain inhibits the HECT activity, I showed in this thesis that the presence of WW1 domain significantly enhances this effect. Although the isolated WW1 domain does not interact with the HECT or the C2 domain in *trans*, it enhances the affinity of the C2 domain to the HECT domain by a factor of 20 when coupled to the C2 domain (C2-linker-WW1 construct). An explanation can be given by the fact that the C2 domain interacts with the C2-WW1L. This might restrict the WW1 orientation in a way in which is able to help the C2 domain in its inhibitory role, by increasing C2:HECT domain affinity. This explanation fits with my model, since the C2:C2-WW1L intramolecular interaction maps to the bottom part of the C2 domain, leaving free the residues involved in HECT binding, located at the opposite side of the C2 domain. In addition, the result showed in Figure 29, where C2:WW1 do not interact with each other are also in agreement, since the interaction detected is mediated by the linker, which is not present in that experiment.

In contrast, Smurf1 lacks the WW1 domain and as a consequence the C2 domain is not able to interact or position itself appropriately to down-regulate HECT activity. Consistent with this, recombinant Smurf1 FL is a constitutively active E3 enzyme *in vitro*. Although Smurf1 has been reported in a previous study to be subject to C2-mediated inhibition (Wan et al., 2011), my results are fully consistent with at least three other studies that find that Smurf1 is not regulated by a C2:HECT interaction (Courivaud et al., 2015; Lu et al., 2011; Mund et al., 2015). The fact that Smurf1 is not inhibited by the C2 domain emphasizes the role of the WW1 domain, since I showed that the C2 is capable of exert inhibition in *trans*. In any case, it is unknown whether Smurf1 is indeed a constitutively active ligase under endogenous expression levels. Since Smurf1 plays important roles in key developmental processes, it seems unlikely that Smurf1 activity would not be regulated *in vivo*. Mechanisms of regulation, that may be obscured under *in vitro* or overexpression conditions, would be PTMs of Smurf1 or the presence of adaptor proteins that may inhibit Smurf1 activity in cells. Moreover, target-regulated expression or differential activity depending on the cellular localization of Smurf1 could also prevent premature target ubiquitination. In fact, the substrate specificities of Smurf ligases have been linked to both PTMs (Cheng et al., 2011; Narimatsu et al., 2009; Ozdamar et al., 2005) and cellular localization (Lu et al., 2011).

Lastly, given the high level of sequence identity between Smurf1 and Smurf2, the fact that these enzymes are regulated differently is surprising. This study thus emphasizes the importance of detailed mechanistic studies in order to decipher the molecular basis of

ligase activity. As Nedd4-family members are important regulators of developmental and carcinogenic processes, the detailed studies of the catalytic mechanisms and differential regulation of these enzymes as presented here, have direct implications not only for understanding their basic function but also for the design of novel ligase-selective therapeutics.

Chapter 3: Ca^{2+} Binding to the C2 domain as an activating mechanism for Nedd4 family E3 ligases

4.3.1 Contribution

Backbone resonance assignments for CSP mapping for the Smurf1 and Rsp5 C2 domains were performed by Christine Wolf and Silke Wiesner. Backbone resonance assignments for the Smurf2 C2 domain are published (Wiesner et al., 2007). The NMR structure shown in Figure 39A was solved by Vincent Truffault. All other experiments described in section 4.3.3 of this thesis were performed by me.

4.3.2 Introduction

There is now ample evidence that under basal conditions the majority of Nedd4-family ligases are in an inactive state which is mediated by intramolecular interactions (Chen et al., 2017; Courivaud et al., 2015; Mari*, Ruetalo* et al., 2014; Mund and Pelham, 2009; Wang et al. 2010; Wiesner et al., 2007; Zhu et al., 2017). This C2 or WW2-3L:HECT interaction comprises the main way to regulate Nedd4 E3 ligases function, crucial for maintaining cell homeostasis (see interaction 4.2.2). On the other hand, in order to activate the ligase, a mechanism to disrupt the correspondent interaction must exist. Nedd4 E3 ligases have developed different strategies. For Smurf2 it has been described that binding to an adaptor protein release the inhibition. Smad7, a *bona fide* Smurf2 substrate, binds to the HECT domain of Smurf2, out-competing the C2 domain. In addition, Smad7 recruits the proper E2, leading to enzyme activation (Ogunjimi et al., 2005). Another example of activation by adaptor proteins was described for Nedd4 and Itch, mediated by the Nedd4 family-interacting proteins (Ndfip1/Ndfip2). They activate the catalysis by binding to the WW domains, but multiple PY–WW interactions are required to activate the protein. The binding introduces a conformational constraint that disrupts the correspondent intramolecular interaction (C2 domain or WW2-3L region with the HECT domain), releasing the HECT domain to perform catalysis (Mund & Pelham, 2009). Later on, Zhu et al., 2017 confirmed that in order to increase auto-ubiquitination of Itch, the Ndfip1-Itch interaction must be mediated by three PY motifs and three correspondent WW domains in the ligase, since the interaction affinity of each WW domain alone is not sufficient. They hypothesized the fourth WW, which is then "free" would be the one mediating substrate recruitment. The same mechanism, mediated by Ndfips holds true for WWP2 (Riling et al., 2015). In addition, auto-inhibition of WWP2 is relieved by another adaptor protein called Dvl2. The requirement of multiple WW domains to be involved in the process is also needed here, since Dvl2 can fulfil its function only after polymerization (Mund et al., 2015). For the highly related WWP1 protein not much is known; it seems that

it becomes activated by Smad7 (Courivaud et al., 2015), as it happens for Smurf2, but how exactly this happens, whether Smad7 binds to the HECT or the WW1 domains was not addressed. A different mechanism consists of activation by Tyrosine (Tyr) phosphorylation. The first example was published for Itch (Gallagher et al., 2006). The mechanism involves the phosphorylation of two Tyr residues located in the WW2-3L, upon which the interaction with the HECT domain is disrupted (Zhu et al., 2017). Exactly the same strategy was defined for WWP2 (Chen et al., 2017). Phosphorylation also regulates Nedd4, but in this case the Tyr residues modified are located one in the HECT domain and the other one in the C2 domain, disrupting the auto-inhibition. These Tyr residues are conserved in many Nedd4-family members except for Smurf1 and Smurf2, precluding regulation by phosphorylation at these sites for these proteins (Persuad et al., 2014). Finally, the ability of the C2 domain to interact with specific ligands can also trigger activation of E3s, as it was described for Nedd4 proteins. Wang et al., 2010 showed that calcium binding to the C2 domain disrupts the C2:HECT domain interaction, leading to activation of the protein. The effect of calcium was significantly enhanced by the addition of a membrane-rich fraction, what suggest that the mechanism *in vivo* probably involves the disruption of the C2:HECT interaction upon Ca^{2+} influx and the translocation to the plasma membrane. Furthermore, Escobedo et al., 2014, showed that the C2 domain of Nedd4L binds Ca^{2+} and IP_3 through the same binding surface where the HECT domain interacts. In addition, they suggested that competition between IP_3 and Ca^{2+} and the HECT domain comprises a mechanism to activate Nedd4L. Due to the fact that the C2 domain is present and conserved in all members of the Nedd4 family, Ca^{2+} binding might represent a general mechanism to release the auto-inhibition state in the Nedd4-family.

C2 domains are independently folded modules of about 130 amino acids, present in a wide variety of proteins, involved mainly in signal transduction and membrane trafficking (Corbalan-Garcia & Gómez-Fernández, 2014). The C2 domain structure consists of a compact β-sandwich formed by two four-stranded β-sheets. A circular permutation defines the two topologies that C2 domains can adopt: class I (or S-family), where the N and C-t tails are located at the top of the domain (Synaptotagmins, PKCα and β) and class II (or P-family) where the N and C-t are located at the bottom (cPLA2, PLCs, PKC δ and ε, Nedd4 family) (Rizo & Südhof, 1998) (Figure 34). The top of the molecule is referred as the region where calcium binds to the C2 domain, and thus the bottom is located at the opposite site. C2 domains can bind both Ca^{2+} and phospholipids. Ca^{2+}-binding is mostly mediated by a set of conserved Asp residues located in three loops (loop 1, 2 and 3) at the top of the molecule, defining the so called calcium binding region (CBR). Of note, the residues involved in HECT binding are also located at the top loop region. On the contrary, C2 domain binding to target phospholipids is achieved by a combination of electrostatic and hydrophobic interactions. Despite the fact that all C2 domains share the

common structural β-sandwich, they displayed different ligand selectivity (Corbalan-Garcia & Gómez-Fernández, 2014). Since phospholipid binding can be regulated by Ca^{2+}, C2 domains may be referred to as Ca^{2+}-dependent lipid binding domains. However, trough evolution two different forms, a Ca^{2+}-dependent and Ca^{2+}-independent forms have diverged (Rizo & Südhof, 1998). The ability of C2 domains to interact and respond to different Ca^{2+} concentrations and lipid composition allows them to participate in many signal transduction processes and thus to play a key role in many cellular functions. The goal of this last chapter was to investigate the Ca^{2+}-C2 interactions as a possible mechanism for the regulation of Nedd4 family E3 ligase activity.

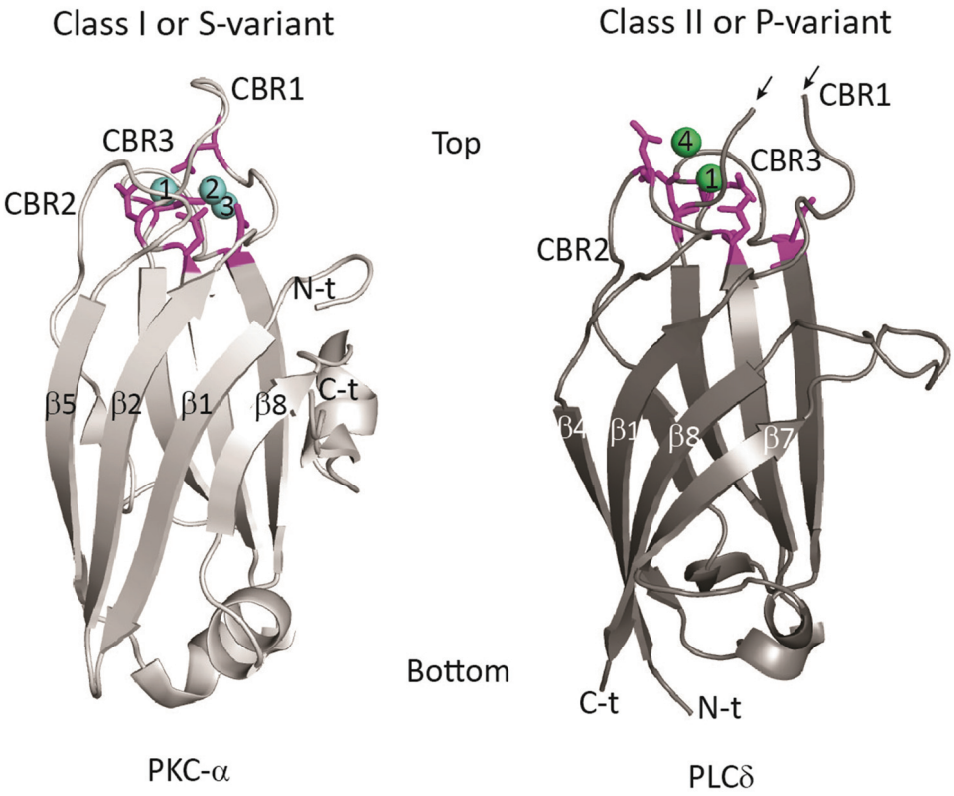


Figure 34. Structural classification of C2 domains. Two different topologies are observed among C2 domains: class I (left), where both the N and C-t are located at the top of the domain and class II (right), where both ends are located at the bottom. The loop region where calcium binds is used to define the top of the molecule, and the bottom is then the opposite region. Two canonical examples were selected: PKC-α (PDB ID: 1DSY; Verdaguer et al., 1999) (Class I) and PLCδ (PDB ID: 1DJI; Essen et al., 1997) (Class II). The arrows indicate a loop which is missing in the crystal structure. The three loops that can interact with Ca^{2+} are labeled as CBR1-3. Asp residues are color in magenta and Ca^{2+} ions in cyan for class I and green for class II.

4.3.3 Results

NMR titration experiments were used to investigate the ability of the Nedd4 members Nedd4, Smurf1, Smurf2 and Rsp5 to bind Ca^{2+}. To this end, I first recorded 2D $^{1}H,^{15}N$-HSQC spectra of the ^{15}N labeled Nedd4 C2 domain alone and in presence of increasing amounts of $CaCl_2$ (Figure 35). The overlay of the different titration points with the reference sample showed strong and numerous CSPs, which increased with each step of the titration up to a 200-fold stoichiometric excess, confirming that the Nedd4 C2 domain interacts with Ca^{2+} as has been shown previously for the highly similar Nedd4L C2 domain (Escobedo et al., 2014). Unfortunately, resonance assignments of the Nedd4 C2 domain was not possible due to protein aggregation at high concentrations and therefore the Ca^{2+} binding surface could not be mapped directly.

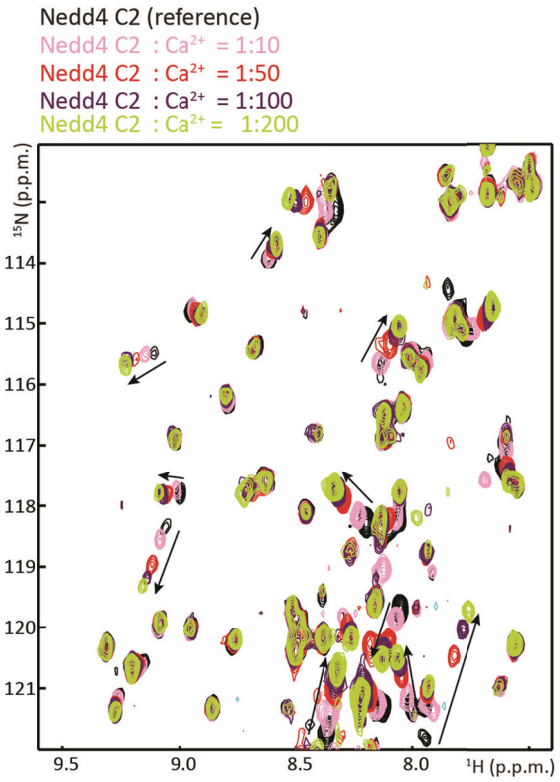


Figure 35. The C2 domain of Nedd4 interacts with Ca^{2+} in solution. Overlay of a representative region of the $^{1}H,^{15}N$-HSQC spectra of a ^{15}N-labeled Nedd4 C2 domain in the absence (black) and presence (rose (1:10), red (1:50), violet (1:100) and lime(1:200) stoichiometric excess of $CaCl_2$. The arrows mark the direction of the most significant CSPs.

Alternatively, I mutated the five conserved Asp and Asn residues located in the CBR to Ala (D35A, D41A, D93A, N95A and D101A respectively), and recorded the spectra of each mutant C2 domain alone and in presence of a 200-fold excess of $CaCl_2$ (Figure 36A-C, E-H). To analyze the effect of the mutations, I selected four peaks displaying CSPs for Nedd4 C2 domain spectra, before and after the addition of $CaCl_2$. The position of each of them was measured and the difference (in p.p.m.) between the original position (without

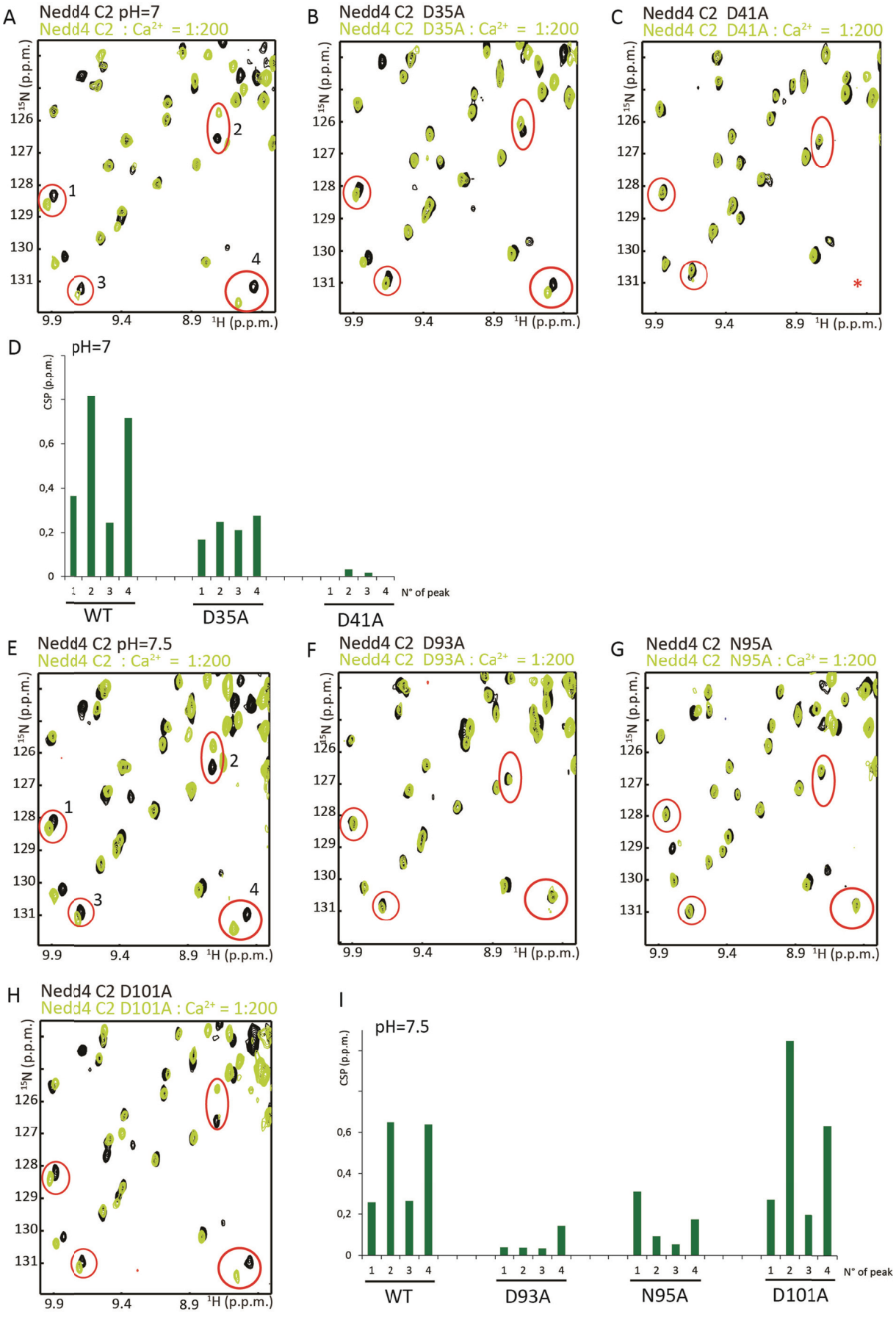

A
Nedd4 C2 pH=7
Nedd4 C2 : Ca2+ = 1:200
B
Nedd4 C2 D35A
Nedd4 C2 D35A : Ca2+ = 1:200
C
Nedd4 C2 D41A
Nedd4 C2 D41A : Ca2+ = 1:200
15N (p.p.m.)
1H (p.p.m.)
D
pH=7
CSP (p.p.m.)
N° of peak
WT
D35A
D41A
E
Nedd4 C2 pH=7.5
Nedd4 C2 : Ca2+ = 1:200
F
Nedd4 C2 D93A
Nedd4 C2 D93A : Ca2+ = 1:200
G
Nedd4 C2 N95A
Nedd4 C2 N95A : Ca2+ = 1:200
H
Nedd4 C2 D101A
Nedd4 C2 D101A : Ca2+ = 1:200
I
pH=7.5
WT
D93A
N95A
D101A

Figure 36. Asp and Asn residues are involved in calcium interaction. A-C) Overlay of a selected region of the $^{1}H,^{15}N$-HSQC spectra of ^{15}N-labeled Nedd4 C2 domain **A)** WT (pH=7), **B)** D35A and **C)** D41A mutants in the absence (black) and presence (lime) of an 200-fold stoichiometric excess of $CaCl_2$. The residues used for analysis are numbered and circled in red. The asterisk in **C** indicates that the resonance is missing. **D)** The plot shows, for each of the selected peaks, the difference in CSP before and after addition of $CaCl_2$, comparing the WT Nedd4 C2 domain and two mutants: D35A and D41A. Titrations experiments were recorded using a pH=7 buffer. **E-H)** As for **A-C** but for Nedd4 C2 domain **E)** WT (pH=7.5), **F)** D93A, **G)** N95A and **H)** D101A mutants. **I)** As **D**, but for Nedd4 C2 domain WT and three mutants: D93A, N95A and D101A. Titrations experiments were recorded using a pH=7.5 buffer.

calcium) and the final position (with calcium) was calculated as the chemical shift average. This analysis was performed for every mutant and the results plotted in Figure **D** and **I**. Of note, a buffer pH=7 was used for recording the spectra of mutants D35A and D41A and pH=7.5 for mutants D93A, N95A and D101A. Thus, a WT spectrum for each pH is used as the reference for the correspondent set of mutants. A reduction in CSP on the plot is an indicator of the relevance of a specific residue in Ca^{2+} binding, since it means that with the introduction of the mutation the C2 domain is less or no longer able to interact with Ca^{2+}. Analyzing plots **D** and **I**, it is possible to conclude that D41 and D93 are essential for the C2: Ca^{2+} interaction, since the four peaks considered showed a significant reduction of the CSP or no shift at all. N95 and D35 are relevant for the interaction, showing a significant reduction in CSP, while D101 is not playing a relevant role since the mutation did not affect the binding capabilities of the WT Nedd4 C2 domain. In the case of mutant D101A the peak number 2 shifts more than the correspondent peak in the WT, which might happen due to some local rearrangement in the mutant.

To investigate whether Ca^{2+}-binding to the C2 domain could indeed compete with HECT domain binding, I recorded a 2D $^{1}H,^{15}N$-TROSY spectra of a ^{15}N-labeled Nedd4 HECT domain as reference and I added a 2-fold excess of unlabeled C2 domain to form a complex (**Figure 37A**). Next, an excess of $CaCl_2$ was added to the sample to analyze whether Ca^{2+} would reverse the CSPs in the HECT domain that were induced by C2 domain binding (**Figure 37B**). The high molecular weight of the complex (approx. 60 kDa) compromised the quality of the spectra. None-the-less it was possible to observed that the resonances that have shifted or disappeared as a consequence of the C2:HECT interaction, re-appear or shift back to the original position after $CaCl_2$ was added. This shows that the presence of Ca^{2+} interferes with the HECT:C2 interaction. In sum, these structural data corroborates that Ca^{2+} binding can trigger the disruption of the auto-inhibition mechanism of human Nedd4 (Wang et al., 2010).

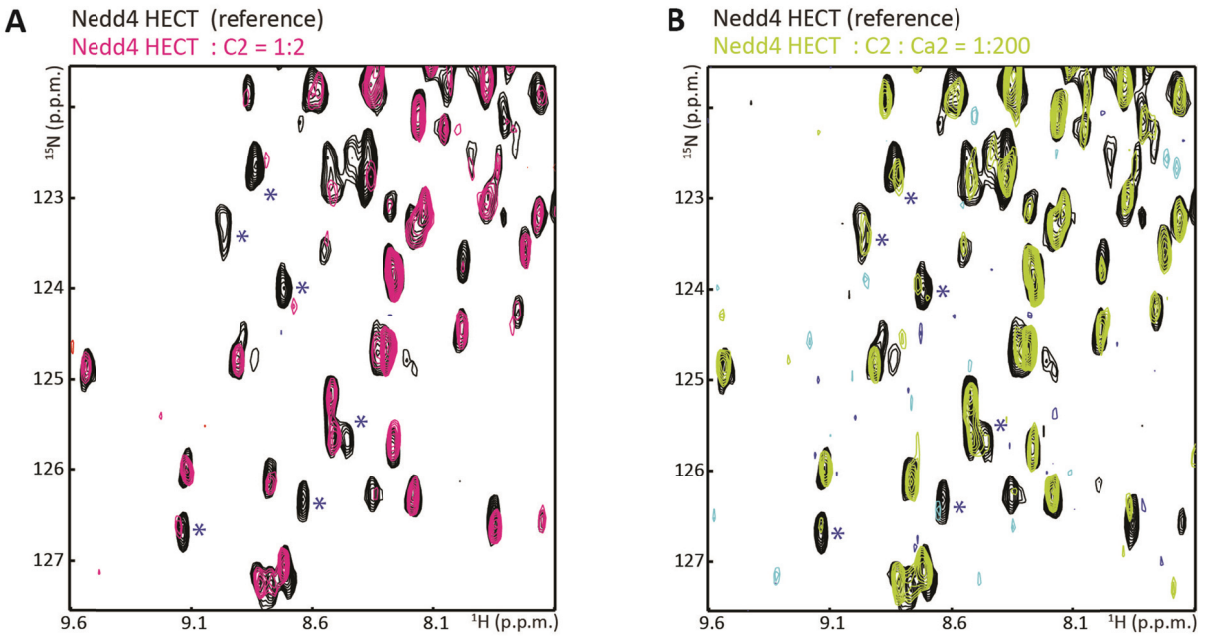


Figure 37. The Nedd4 C2:HECT interaction is impair by Ca^{2+}. **A)** Overlay of a representative region of the $^{1}H,^{15}N$-HSQC spectra of ^{15}N-labeled Nedd4 HECT domain in the absence (black) and presence (magenta) of a two-fold stoichiometric excess of the C2 domain. The blue asterisks mark the six HECT resonances that shifted or disappeared upon C2 domain interaction, according to Mari* Ruetalo* et al., 2014. **B)** Overlay of the same region as **A** after addition of a 200-fold stoichiometric excess of Ca^{2+} (lime).

In order to study how conserved C2-Ca^{2+} binding is as an activation mechanism for Nedd4 E3 ligases, I investigated two other members of the human Nedd4 family, Smurf1 and Smurf2, and the yeast Nedd4 ligase Rsp5. As for Nedd4, I recorded the spectra of ^{15}N-labeled samples for Smurf1, Smurf2 and Rsp5 C2 domains in the absence and presence of increasing amounts of Ca^{2+} (**Figure 38**).

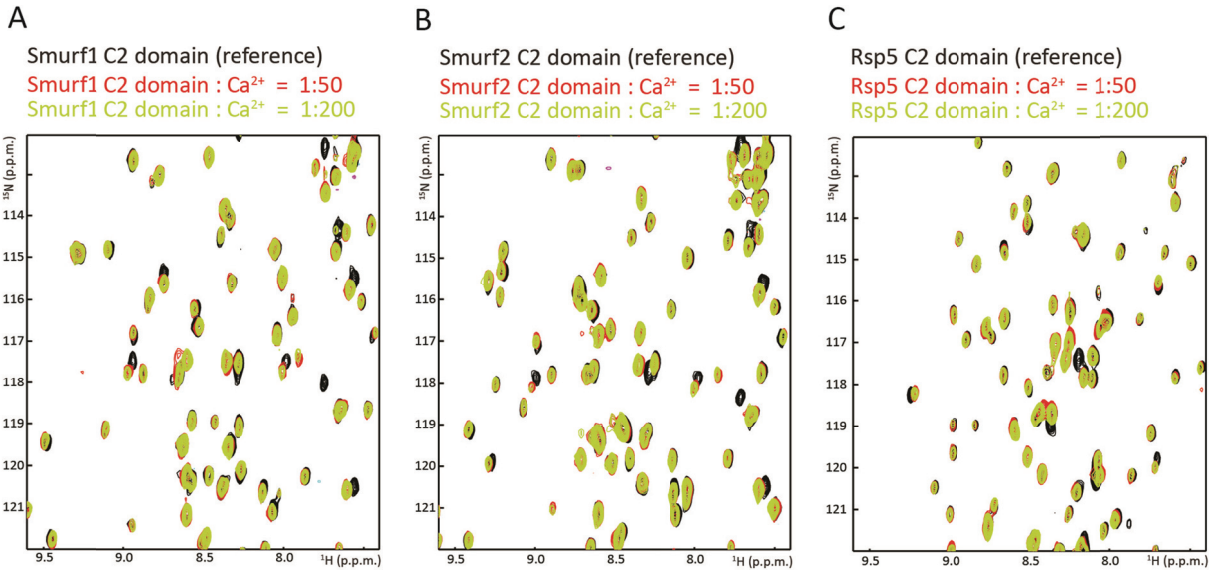


Figure 38. Smurf1, Smurf2 and Rsp5 C2 domains showed minor CSPs upon $CaCl_2$ addition. **A)** Overlay of a representative region of the $^{1}H,^{15}N$-HSQC spectra of ^{15}N-labeled Smurf1 C2 domain in the absence (black) and presence of a 50-fold (red) or 200-fold (lime) stoichiometric excess of $CaCl_2$. **B)** As **A** but for Smurf2. **C)** As **A** but for Rsp5.

For all three proteins only few CSPs were observed upon addition of a 200-fold stoichiometric excess of $CaCl_2$. This shows that these C2 domains have only very limited Ca^{2+} binding capabilities.

To gain structural insight into those functional differences between the Nedd4 and the Smurf1, Smurf2 and Rsp5 C2 domains, Silke Wiesner and Vincent Truffault solved the NMR structure of the Rsp5 C2 domain, while the solution structure of the Smurf2 C2 domain (PDB ID: 2JQZ; Wiesner et al., 2007) and the crystal structure of the Smurf1 C2 domain (PDB ID: 3YPC) were already available. For Smurf1 the backbone resonance assignments were already available in our group (performed by Christine Wolf). In Figure 39A an ensemble of the 10 lowest-energy NMR structures of the Rsp5 C2 domain is displayed. The NMR structure of the Rsp5 C2 domain revealed a canonical C2 domain β sandwich fold, comprising two four-stranded antiparallel β sheets.

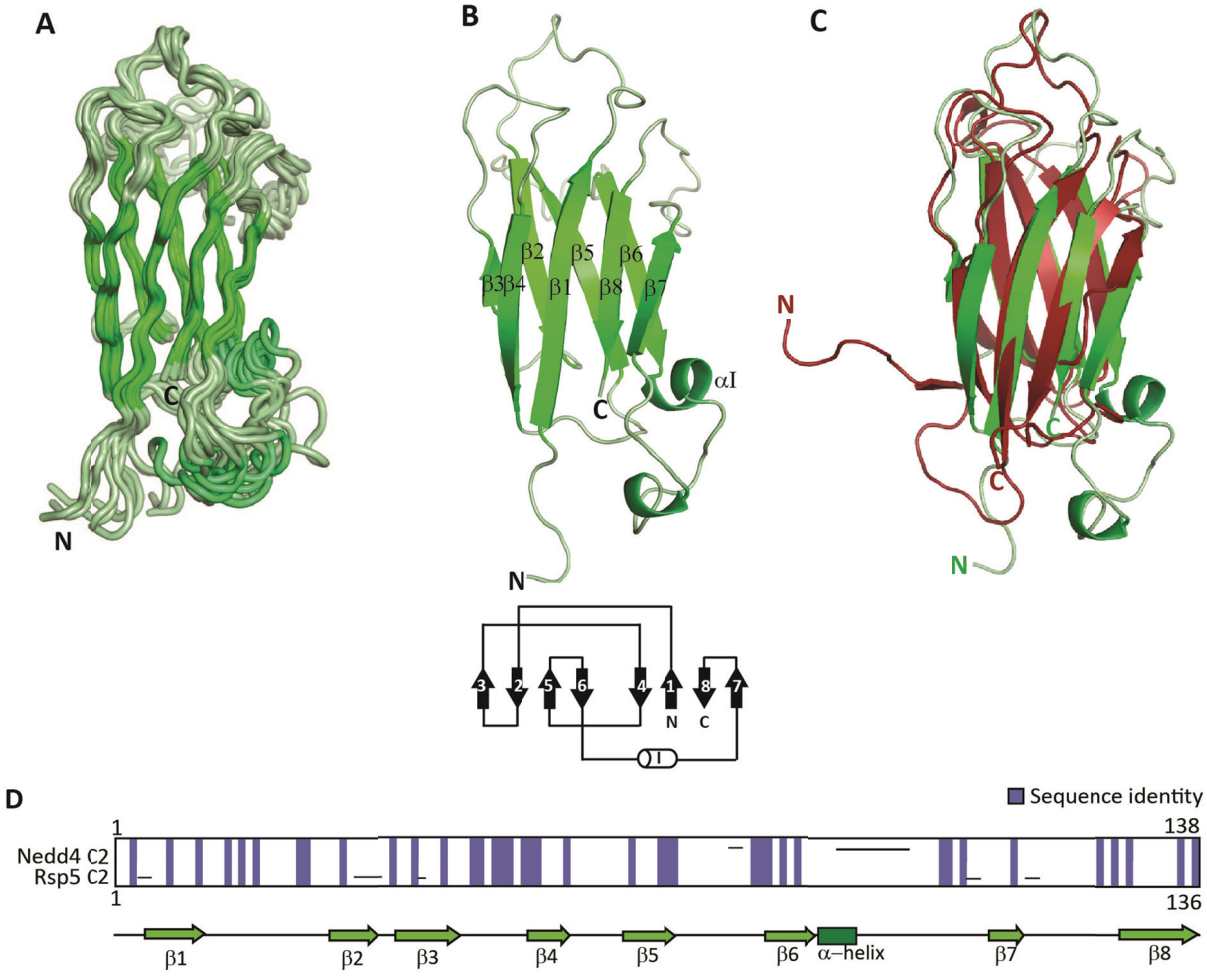


Figure 39. Solution structure of the Rsp5 C2 domain. A) Ensemble of the 10 lowest-energy NMR structures out of 50 structures calculated. The backbone is shown in pale green. α-helices and β-strands are color-coded in dark green and green respectively. The NMR structure was solved by Vincent Truffault. B) Ribbon representation of the lowest-energy structure color-coded as in A. A representation of the secondary structure topology of the Rsp5 C2 domain is indicated on the bottom. C) Structural superposition of the Rsp5 C2 domain (green) and the Nedd4 C2 domain (PDB:3B7Y) (purple), both belonging to the class II or P-family. D) Sequence alignment of Nedd4 and Rsp5 C2 domains. Identical residues are colored in blue and alignment gaps are indicated with a black line. A prediction of secondary structure elements for Rsp5 C2 domain was performed with Quick2D (Alva et al., 2016) and it is shown below the alignment.

Additionally, it contains one short α-helix that connects β strands 6 and 7. A secondary structure topology is shown in Figure 39B. The overlay of the Rsp5 and Nedd4 C2 domains (PDB ID: 3B7Y) showed that the core β sandwich fold is highly similar (Figure 39C), although the sequence conservation is only about 28% (Figure 39D). As for Nedd4, Smurf1 and Smurf2, the Rsp5 C2 domain belongs to the type II family where the N and C-t are located at the bottom of the structure.

With all resonance assignments and structures in hand, I mapped the residues that were affected by Ca^{2+} binding onto the respective C2 domain structures. For Smurf1, the conserved residues D29, D35, N56 and N82 do not display CSP. I found that the largest (though still minor in comparison with the Nedd4 C2 domain) CSPs stem from four Histidine (His) residues (H49, H65, H83 and H87) (Figure 40A). Two of them (H49 and H65) are located outside of the CBR. The other two (H83 and H87) lie in the loop 3, one of the known regions for Ca^{2+} interaction. However, His residues are very sensitive to small pH changes, which can happen upon the addition of $CaCl_2$ to the NMR sample. Smaller CSPs were also observed for a few other peaks (S50, N63, K84, K85, K88 and V129), but all of them were located next to one of the four His previously mentioned, and thus can be considered as indirect effects. No CSP was observed for the first loop, which is required for Ca^{2+} coordination. For Smurf2 the results were very similar to what was observed for Smurf1 (Figure 40B). D27 and D34, the conserved Asp residues that belong to the CBR1, as well as N55 from the CBR2, showed no CSP. As for Smurf1, four His residues (H48, H64, H82 and H86) displayed the most pronounced chemical shift changes. The latter two (H82 and H86) are located in the CBR3, while the first two (H48 and H64) are outside any CBR. Smaller CSPs were observed for residues C47, N81, K83, K87 and K88, all located next to His. Overall, considering the small number of residues affected, the magnitude of the shifts, the lack of CSPs in the CBR1 and the His sensitivity to pH changes, I can conclude that the C2 domain of Smurf1 and Smurf2 do not interact with Ca^{2+}.

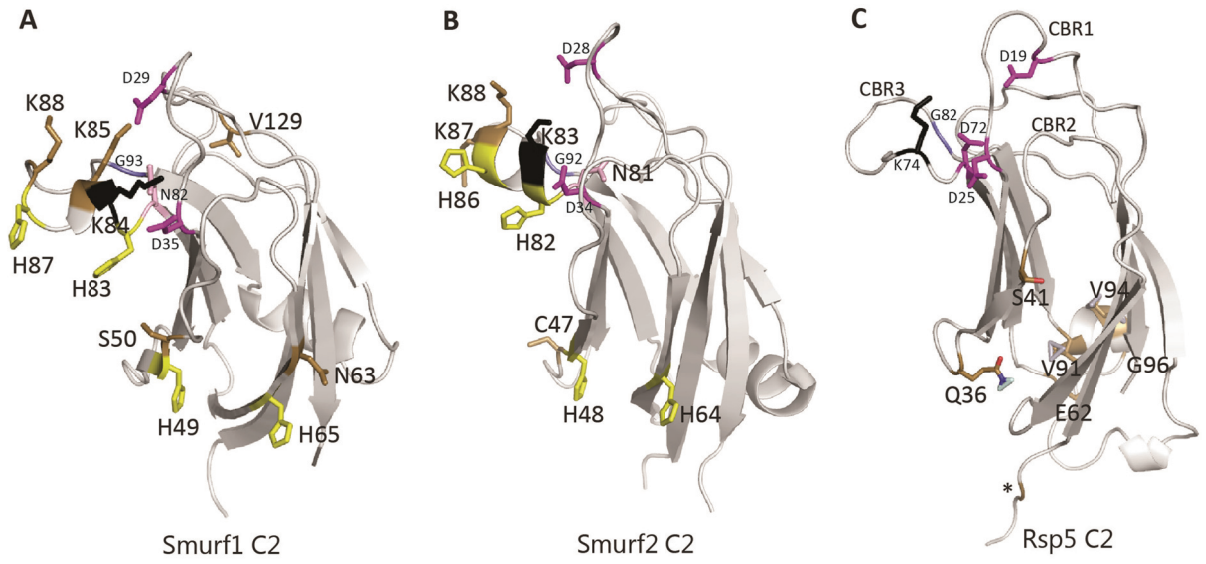


Figure 40. Residues displaying CSPs for the Smurf1, Smurf2 and Rsp5 C2 domain upon $CaCl_2$ addition. A) Ribbon representation of the Smurf1 C2 domain crystal structure (PDB ID: 3YPC). His residues affected by $CaCl_2$ addition are shown in yellow, while other residues displaying CSPs are color in pale yellow. **B)** As **A** but for the Smurf2 C2 domain (PDB ID: 2JQZ). **C)** As **A** but for the Rsp5 C2 domain. ***Gly** introduced by the restriction site. **A-C)** The positions corresponding to the five conserved Asp residues are shown color-code as in **Figure 41**. CBRs are labeled in Rsp5.

In the case of Rsp5, the observed CSPs (G*, Q36, S41, E62, V91, V94 and G96) were even smaller than those for Smurf1 and Smurf2 (**Figure 40C**). The largest ones were affecting two Gly residues (one resulting from the restriction site introduced for cloning) and none of them located in the conserved loops for Ca^{2+} coordination. Altogether, this shows that the Rsp5 C2 domain is not able to bind Ca^{2+}.

To understand the reason why some members of the Nedd4-family are able to bind Ca^{2+} and others do not, I made a sequence structural alignment to compare them to each other and to other C2 domains (**Figure 41**).

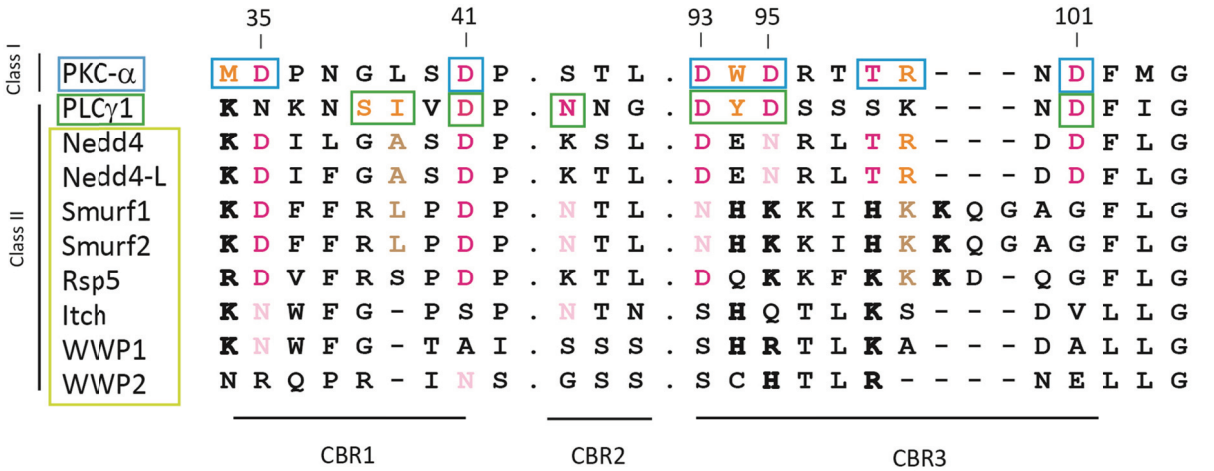


Figure 41. Sequence structural alignment of representative C2 domains together with Nedd4-family C2 domains. A canonical class I C2 domain (PCK-α, boxed in blue) and one that belongs to the class II family (PLCδ, boxed in green) were chosen for comparison. Members of the Nedd4-family are boxed in yellow. The C2 domain protein sequence of the three regions involved in calcium binding are shown and labeled as CBR1, CBR2 and CBR3. The dots indicate entire regions from the sequence that were not displayed, while the dashes indicate a gap in the alignment. PKC-α residues involved in Ca^{2+} binding are boxed in blue, while for PLCδ is done in green. The key residues for Ca^{2+} coordination are labeled in magenta, and conserved residues in light pink. In orange are labeled other residues that contribute to Ca^{2+} binding, while conserved residues in this position are shown in brown. Introduction of an extra positively charged residue is indicated in bold. The numbering above the alignment refers to the protein sequence of Nedd4.

PKC-α and PLCδ were chosen as a typical C2 domains from class I and class II respectively. Both of them share the main feature regarding calcium interaction: they have five Asp or Asn residues located in the CBR, four of which are located in the same positions. Other extra residues contribute to Ca^{2+} binding (four for PKC-α and three for PLCδ). The main difference is that in the class I only the CBR1 and CBR3 interact with calcium, while for C2 domains of the class II, the third CBRs is involved in calcium binding. This is share by the majority of the members of each family. The amount of Ca^{2+} ions that each protein binds it also seems to be a feature of each class, three Ca^{2+} ions in the case of class I and two for class II (Figure 34). Since the Nedd4-family C2 domains belong to the class II family, it is expected that they are more similar to PLCδ than to PKC-α. However, the alignment of the CBR1 region showed that there is a high level of conservation for all the Nedd4-family members studied (Nedd4, Smurf1, Smurf2 and Rsp5) with PKC-α, sharing two conserved Asp residues in that loop. In the CBR2, class II C2 domains such as PLCδ, bear a conserved Asn residue that participates in Ca^{2+} interaction; this is displayed only in Smurf1 and Smurf2. Finally, according to the CBR3, the Nedd4-family members studied can be group in two. On the one hand, Nedd4 (as well as Nedd4L) which shows a considerably high conservation with respect to PKC-α, having two Asp residues conserved out of three, and a Asn for the third one. On the other hand, Smurf1, Smurf2 and Rsp5 are more similar among each other and do not show residues involved in calcium binding in the appropriate positions. In addition, Smurf1 and Smurf2 bear four extra positively charge amino acids in the CBR3, while Rsp5 three more. This might increase the overall positive surface potential of the domains; further diminishing Ca^{2+} binding capabilities. In sum, the lack of conservation in the key regions for Ca^{2+} binding (CBRs) observed for Smurf1, Smurf2 and Rsp5 is probably the main reason why they cannot interact with calcium. Other members of the Nedd4 family such as Itch and the WWP ligases do not show any conservation at the three CBRs, suggesting that Ca^{2+} interaction is not involved in the regulation of their function.

Considering the features displayed in the alignment, Nedd4 seems to be more similar to class I C2 domains than to class II: the CBR1 bears two Asp residues as PKC-α, no Asp or Asn is found in the CBR2 and the CBR3 highly resembles PKC-α. The mechanism of Ca^{2+} coordination for the C2 domain of PKC-α was already described in Corbalan-Garcia & Gómez-Fernández, 2014. Its crystal structure (PDB ID: 1DSY), solved in presence of three Ca^{2+} ions, is shown in Figure 42A. The two Asp residues located in the CBR1 contribute to the coordination of Ca1 and Ca2. The three Asp residues in the CBR3 participate in the coordination of the three Ca^{2+} ions Ca1, Ca2 and Ca3. The crystal structure of Nedd4 (PDB ID:3B7Y) (Figure 42B, E) shows the presence of only two Ca^{2+} ions. Lack of electron density does not allow observing a region of the CBR1, which could be involved in the binding of a third one. However, ITC studies on the Nedd4L C2 domain (which shares an 83% sequence identity with the Nedd4 C2 domain) also suggests that it binds only to two Ca^{2+} ions (Escobedo et al., 2014). A superposition of PKC-α and Nedd4 C2 domain structures (Figure 42C), shows that the overall fold is highly similar. However, they only share position Ca1 to perform Ca^{2+} coordination, which is expected for proteins of these two classes. Figure 42D shows the crystal structure of PLCδ C2 domain (PDB ID:1DJI), which uses the position Ca1 but also Ca4 to bind calcium, as is generally seen for class II C2 domains.

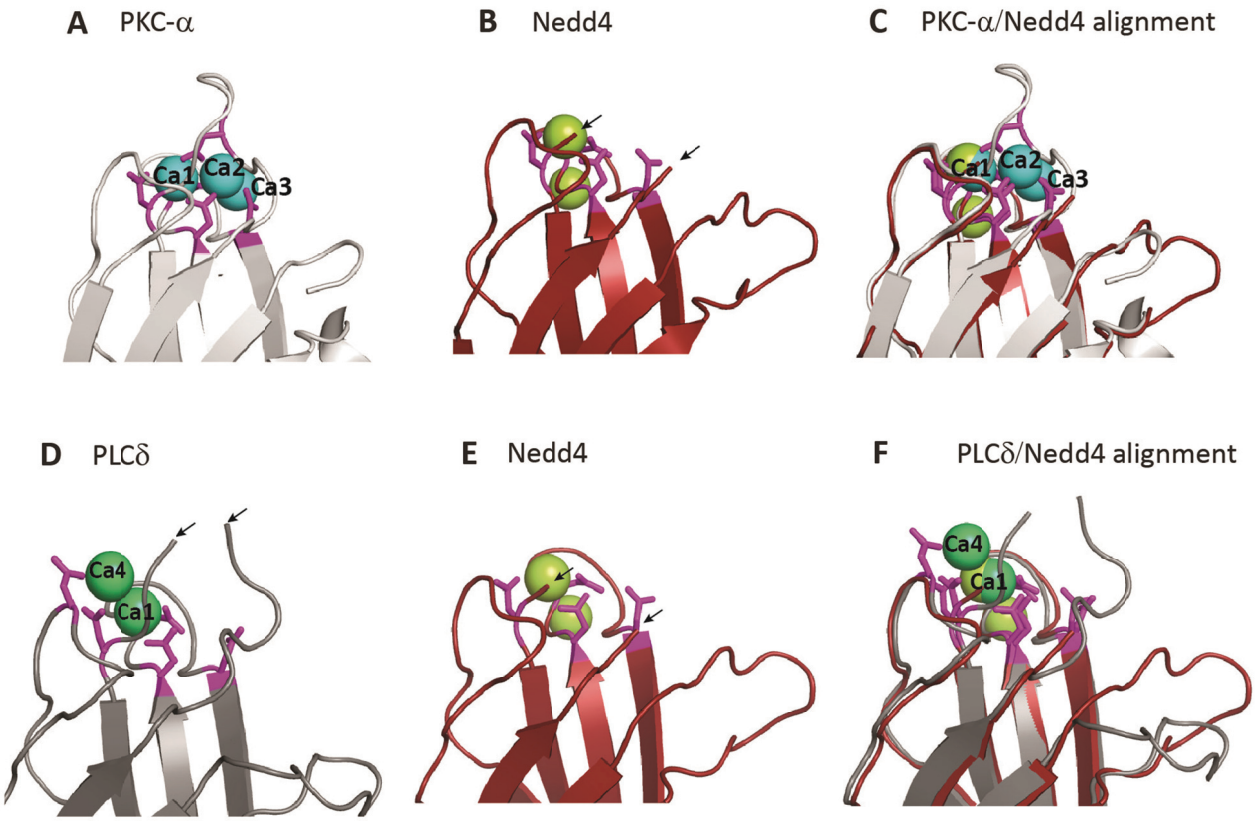


Figure 42. Calcium coordination of Nedd4 C2 domain in comparison with canonicals class I and II C2 domains. The color code is the same used in **Figure 34**. Nedd4 is color in red and the Ca^{2+} ions in lime. **A)** Ribbon representation of the CBR region of PKC-α C2 domain (PDB ID: 1DSY). The tree Ca^{2+} ions are indicated in the structure. **B)** As **A** but for Nedd4 C2 domain (PDB ID: 3B7Y). The arrows indicate a loop which is missing in the crystal structure. **C)** Structural superposition of PKC-α and Nedd4 C2 domains. **D)** As **A** but for PLCδ C2 domain (PDB ID: 1DJI). The arrows indicate a loop which is missing in the crystal structure. **E)** As **B** but aligned to overlap with PLCδ. **F)** Structural superposition of PLCδ and Nedd4 C2 domains.

Interestingly, a superposition of PLCδ and Nedd4 C2 domains (**Figure 42F**) shows that Nedd4 binds one Ca^{2+} ion in position Ca1, while the other does not seem to be in position Ca4 either. Whether Nedd4 C2 domain coordinates Ca^{2+} ions in a different why from both C2 domain classes would need further studies.

Conclusions:

- The Nedd4 C2 domain binds Ca^{2+} through conserved residues in the CBR and this interaction interferes with HECT domain binding.
- On the contrary, the C2 domains of Smurf1, Smurf2 and Rsp5 are not able to bind Ca^{2+}.
- Ca^{2+}-mediated Ub ligase activation thus does not play a role in Smurf1, Smurf2 and Rsp5. According to the structural alignment Itch, WWP1 and WWP2 are also probably not regulated by this mechanism that seems to be restricted to a small subset of Nedd4-family proteins.
- The lack of conservation in the CBR of Smurf1, Smurf2 and Rsp5 C2 domain is probably the main cause to explain why they are not able to coordinate Ca^{2+}.
- Nedd4 C2 domain might bind Ca^{2+} differently from the canonical members of class I and class II C2 domains.

4.3.4 Discussion

Several studies have addressed how Nedd4-family Ub ligases avoid untimely targeting of themselves or substrates by adopting an auto-inhibited conformation. It is thus important to understand how these ligases are activated in order to fulfill their functions. Previous studies (Escobedo et al., 2014; Wang et al., 2010) have shown that Ca^{2+} triggers activation of Nedd4 proteins. Ca^{2+} binding to the C2 domain relocate the Ub ligase to the plasma membrane, leaving the HECT domain in a "free" state, ready to perform catalysis. In this study I classified different C2 domains of Nedd4-family proteins regarding their Ca^{2+}

binding capabilities in order to evaluate how conserved this activation mechanism is within the Nedd4-family.

My NMR analysis show that, as Nedd4L, the Nedd4 C2 domain interacts with Ca^{2+} involving conserved residues located in the CBR. This is in agreement with the Nedd4 C2 structure, which co-crystalize with two Ca^{2+} ions (PDB ID: 3B7Y). In addition, I could show that peaks in the Nedd4 HECT domain that display CSPs upon C2 domain binding, reverse to the free state upon addition of Ca^{2+} to the HECT:C2 complex. This demonstrates that the C2 domain interacts both with the HECT domain and with Ca^{2+} through the same or at least a partially overlapping binding surface. This finding is consistent with results published by Escobedo et al., 2014, where the Nedd4L C2:HECT interaction was shown to dissociate in presence of Ca^{2+}. In contrast, Smurf1, Smurf2 and Rsp5 are not able to bind Ca^{2+}, therefore the mechanism of C2 release from the HECT domain that is triggered by Ca^{2+} does not play a role in their activity regulation. The small CSPs observed for their C2 domains upon calcium addition, are displayed mainly by residues located outside the CBR and especially near His, which are prone to be affected by small pH changes. Among Smurf1, Smurf2 and Rsp5 the most conserved with respect to a canonical C2 domain is Rsp5 (Figure 41). The fact that this domain, being the most conserved and possessing no His residue, is the one that shows less CSPs upon Ca^{2+} binding, supports the idea that the CSPs observed for Smurf1 and Smurf2 were not significant, but stem from changes induced by His residues.

The structural alignment based on the protein sequence of C2 domains (Figure 41) allows the understanding of some differences inside the Nedd4-family. The only two members that displayed a considerably high level of sequence conservation concerning the Asp/Asn residues in the CBRs, are the ones that can bind calcium: Nedd4 and Nedd4L. The other three studied proteins Smurf1, Smurf2 and Rsp5, showed less conservation, in particular concerning the CBR3. Indeed, CBR3 might be more relevant in terms of C2-Ca^{2+} interaction, since for PKC-α for example, the Asp residues belonging to the CBR3 are involved in the interactions with all Ca^{2+} ions (Ca1, Ca2 and Ca3) while the Asp residues in CBR1 participate in the coordination of only Ca1 and Ca2. Considering the CBR3, Smurf1, Smurf2 and Rsp5 are really similar among them, but not to the canonical C2 domains.

Nedd4 seems to share features with both classes of C2 domains; its topology corresponds to class II, though, its sequence conservation is higher with respect to class I C2 domains (Table 5).

Table 5. Nedd4 C2 domain shows more conservation to canonical members of class I than to class II C2 domains. Identical and similar residue percentages for Nedd4 C2 domain, in comparison to two of the canonical C2 domains of each class are shown. The alignment was performed in Jalview using protein sequences of the same length.

		Identities (%)	Positives (%)
class II	PKC-α	44	66
	Synaptotagmin-1	32	56
class I	PLCδ	33	49
	cPLA2	23	39

Considering Ca^{2+} coordination, Nedd4 binds two ions, one in position Ca1, which is shared by all C2 domains, and a second one, which does not seem to be in either of the other described positions (Ca2, Ca3 and Ca4). Whether Nedd4 displays particular features regarding Ca^{2+} binging that could have functional implications required further studies.

In sum, calcium mediated activation of Nedd4 E3 ligases is not a shared mechanism by all members of the Nedd4-family; on the contrary it is restricted to Nedd4 and Nedd4L proteins.

Nedd4 is activated then by the three known mechanisms: release of the intramolecular inhibition by binding to adaptor proteins (Ndfips) (Mund & Pelham, 2009), phosphorylation of Tyr residues located inside the C2 and the HECT domain (Persaud et al., 2014), and binding to Ca^{2+} using an overlapping surface with the HECT binding region. The Try^{43}, involved in phosphorylation, locates in the β2 strand of Nedd4 C2 domain near the CBR1 region, which may argued in favor of a common mechanism. However, to date there is no evidence that links calcium binding and phosphorylation for Nedd4. This holds true for Nedd4L, which is also regulated by the same three mechanisms (Escobedo et al., 2014; Mund & Pelham, 2009; Wang et al., 2010; Zhou & Snyder, 2005).

On the contrary, Smurf1, Smurf2 and Rsp5 are regulated neither by phosphorylation nor by Ca^{2+} influx. The Tyr residue which is involved in phosphorylation for the other members of the Nedd4-family (Nedd4, Nedd4L, WWP1, WWP2 and Itch) is not conserved in Smurfs and Rsp5, precluding this mechanism to happened. For Rsp5 it was described that an adaptor protein called Bsd2, orthologue of Ndfips in mammals, plays an activation role (Hettema et al., 2004).

Both for Smurf1 and Smurf2, the C2 domain still plays a role in membrane localization, but in a Ca^{2+} -independent manner. For Smurf2, it is necessary that Smad7 binds to the HECT domain to open the conformation, which allows then the C2 domain to bind membrane phospholipids. In the case of Smurf1 this opening state is not necessary since no auto-inhibition is present, in concordance with its default membrane localization (Lu et al., 2011).

Finally, different members of the Nedd4-family have evolved their one set of activation mechanism, involving also different adaptor proteins. This might allow the cell to differentially activate each of them when is required, which would not be possible if the same mechanism would applied for all of them. In addition, detailed understanding of how the activation works could be useful in developing ligase-specific therapeutics.

5. General discussion

Ubiquitination is a PTM that plays a role in virtually every cellular process by guiding the fate of hundreds of proteins inside the cell and is therefore crucial for cell homeostasis. Among the enzymes involved in this pathway, E3s stand out mainly for one reason: they define the specificity of the reaction. In other words, they choose the E2 from which they receive the Ub, and select the substrate which is going to be modified, thereby determining the outcome of the reaction. To date, it is estimated that the human genome encodes about 600 E3s, what helps to explain ubiquitination comprehensiveness. From those E3s, only 28 belong to the HECT family, whose main feature is their ability to form an Ub~thioester intermediate before they transfer Ub to the substrate. In my thesis, I focused on the Nedd4-family of HECT ligases, which has nine members in humans. Due to their relatively small size (~100 kDa) in comparison with other E3s and the high conservation level of their members, they form an interesting group of enzymes to study and characterize the mechanism of ubiquitination. In addition, Nedd4-family members have been associated with several human diseases; e.g. they play preponderant roles in cancer acting both as tumor suppressors or oncoproteins (Zou et al., 2015).

In view of their significance and the potential impact of their activity inside the cell, E3 ligases in general and Nedd4 E3s in particular, must be tightly regulated to avoid targeting themselves and/or their substrates unless it is required. E3 activity can be regulated at different levels including: E2 recruitment, substrate interaction, E3 processivity and through intermolecular as well as intramolecular interactions (Mari* Ruetalo* et al., 2014). In the case of the Nedd4-family, it is now clear that intramolecular interactions are essential for their regulation. When I started my PhD only a few examples of intramolecular interactions participating in activity regulation were available. The pioneer work published in Wiesner et al., 2007 established that Smurf2, one of the Nedd4-family members, is regulated by an intramolecular interaction between the C2 domain and the HECT domain that causes auto-inhibition of the enzyme. Itch was also described to be auto-inhibited, but through a different interaction, mediated by the WW and the HECT domains (Gallagher et al., 2006). Later, Wang et al., 2010 showed from a functional perspective that Need4 and Nedd4L are also auto-inhibited. In this context, the first part of my thesis (chapter 4.1) dealt with the molecular basis of the auto-inhibition mechanism in Smurf2 and Nedd4. Using structural and functional approaches I was able to explain how the C2:HECT auto-inhibitory mechanism works: The C2 domain binds to the HECT domain on a surface mapped in Mari*Ruetalo* et al., 2014, and this binding not only locks the HECT domain in a catalytically incompetent conformation where it cannot receive the Ub from the E2 (Figure 14), but also buries the UBS by partial surface overlapping, abolishing non-

covalent Ub binding (Figure 11 C-D, 12). These two events maintain FL Smurf2 and Nedd4 in an inhibited conformation, although a full inactivation is not reached (discussed further below). Of note, the same conclusions were obtained both for Smurf2 and Nedd4, showing a high level of conservation between these two ligases.

In the next couple of years, evidence was mounting that both the C2 domains and the WW-linker regions are involved in the regulation of many other members of the family (Chen et al., 2017; Courivaud et al., 2015; Mari* Ruetalo* et al., 2014; Mund et al., 2015; Riling et al., 2015; Zhu et al., 2017). All of them are described as being auto-inhibited. Differently from what was thought at the beginning of my work, the inhibition mechanism is not exactly the same for all the Nedd4-family members, but they have developed slightly protein-specific differences. Roughly, considering which domains mediate inhibition, two groups can be distinguished: A) C2-mediated inhibition, to which Smurf2, Nedd4 and Nedd4L belong, as already discussed and B) WW2-3L-mediated inhibition, as described for Itch, WWP1 and WWP2. Here, the inhibition is mediated by the interaction of a α-helix located in the linker between the WW2 and WW3 domain, which basically causes the same effect that the C2 domain does. It blocks the HECT domain and impairs E2-E3 transthiolation while part of the WW2 and the following linker partially occupy the UBS (Figure 16). A structural comparison as published in Zhu et al., 2017, shows that the residues in the HECT domain involved in auto-inhibition either in group A or B overlap with the UBS. However, these residues involved in auto-inhibition are not conserved among every member of the family, but just within each group. This helps to explain why Smurf2, Nedd4 and Nedd4L on one side and Itch, WWP1 and WWP2 on the other have common mechanisms.

This overview about the different proteins and mechanisms of the Nedd4-family does not include another relevant member, Smurf1. Few published studies are available (Lu et al., 2011; Courivaud et al., 2015; Mund et al., 2015; Wan et al., 2011) and no agreement about whether it is auto-inhibited was established. Therefore, I decided to study the activity and potential regulation of Smurf1 (Chapter 4.2). Considering the high level of sequence conservation with the related Smurf2 enzyme (Figure 17) one would assume that the C2:HECT auto-inhibition mechanism should be conserved between Smurf2 and Smurf1. Indeed, my results show that the Smurf1 HECT domain is able to interact with the C2 domain and Ub in a non-covalent manner (Figure 18, 22A, 23A). Both binding surfaces are partially overlapping on the HECT domain and are well conserved with respect to their Smurf2 counterparts (Figure 20, 22B-C, 23B-C). In addition, competition assays (Figure 26) showed that the Smurf1 C2 domain can inhibit HECT domain activity in *trans*. However, in auto-ubiquitination assays, Smurf1 FL was highly active, with the deletion of the C2 domain barely having an effect (Figure 25), consistent with other reports (Lu et al., 2011; Courivaud et al., 2015; Mund et al., 2015). Altogether, this shows that Smurf1 is the only

member of the family which is not regulated by an auto-inhibitory mechanism, at least *in vitro*. Given the relevance of E3 activity regulation in cells, it is reasonable to assume that a mechanism to regulate Smurf1 activity must be present in the cell, which is missing in *vitro*, such as the presence of specific proteins to inhibit Smurf1 or differential activity depending on cellular localization. Indeed, Lu et al., 2011 shows that the C2 domain has a role in Smurf1 localization to the plasma membrane, since deletion of the C2 domain changed it localization from being membrane-associated to a cytoplasmic distribution. Moreover, the C2 domain could also be involved in substrate selection. This proposed role for the C2 domain still awaits further structural evidence. PTMs are also linked to Smurf1 regulation; Cheng et al., 2011 showed that PKA dependent phosphorylation of Smurf1 can affect its affinity to substrates, reducing ubiquitination of some of them and increasing it for others. Finally, a potential mechanism of inhibition could be abolishing binding to its targets by blocking the WW domains, as it was described to happen when LMP-1 binds to Smurf1 (Cao & Zhang, 2013).

From a sequence point of view, the major difference between Smurf1 and Smurf2 is the linker region, between the C2 and the HECT domains, in particular the lack of one WW domain in Smurf1. The analysis of chimeric proteins (Figure 27), where the entire linker region was swapped, showed activation of Smurf2 and inhibition of Smurf1. In addition, the deletion of the WW1 domain from Smurf2 also activates the protein, but less efficiently (Figure 28). These results confirm the potential of the Smurf1 C2 domain to exert inhibition and led me to the conclusion that the Smurf1 linker does not allow for auto-inhibition, with the WW1 domain playing an important role in mediating C2:HECT inhibition. The Smurf2 WW1 domain was reported not to interact with the HECT domain (Wiesner et al., 2007) and our own results showed that it is not able to interact directly with the C2 domain either (Figure 29). I hypothesized then, that the Smurf2 WW1 domain might help the C2 domain to orient itself to perform inhibition. My NMR results showed indeed that the presence of the linker-WW1 fragment together with the C2 domain in an entire entity (C2-linker-WW1) changed the way in which the C2 domain interacts with the HECT domain, increasing the binding affinity ~20X (Figure 30). In addition, the effect of the Smurf2 linker-WW1 was confirmed in competition assays, where the HECT domain inhibition was strikingly higher for the C2-linker-WW1 than for the C2 domain alone (Figure 32). Analyzing the CSPs for the C2-linker-WW1 domains compared with the C2 domain and the linkerWW1 fragment, allowed me to map an interaction surface between the C2 domain and the C2-WW1L (Figure 31). Taken all the data together, I proposed a model where the interaction between the Smurf2 C2 domain and the C2-WW1L orients the WW1 domain to reach the HECT domain to inhibit its activity, together with the C2 domain. The C2-WW1L residues involved in the interaction with the C2 domain are quite conserved in Smurf1, highlighting the relevance of the WW1 domain, since the interaction

between the C2 domain and the C2-WW1L could in principle happen in Smurf1 too, but the lack of WW1 is still impairing auto-inhibition.

In Figure 43 I show the two major contributions that I made to the understanding of the mechanisms underlying auto-inhibition inside the Nedd4-family. Firstly, I confirmed and contributed with structural evidence to the fact that there is one member of the family, Smurf1, whose activity is not subject to auto-inhibition (Figure 43C). To date, this is the only Nedd4 member, while only two other members (HECW1 and HECW2) remain unstudied. Secondly, my results revealed that in the case of Smurf2 (Figure 43A), the auto-inhibitory interaction is mediated not only by the C2 domain, which locks the C-lobe and buries the UBS, but the WW1 domain plays an important role too. Interestingly, this role of the WW1 domain is what explains the difference between an auto-inhibited Smurf2 and an active Smurf1 enzyme.

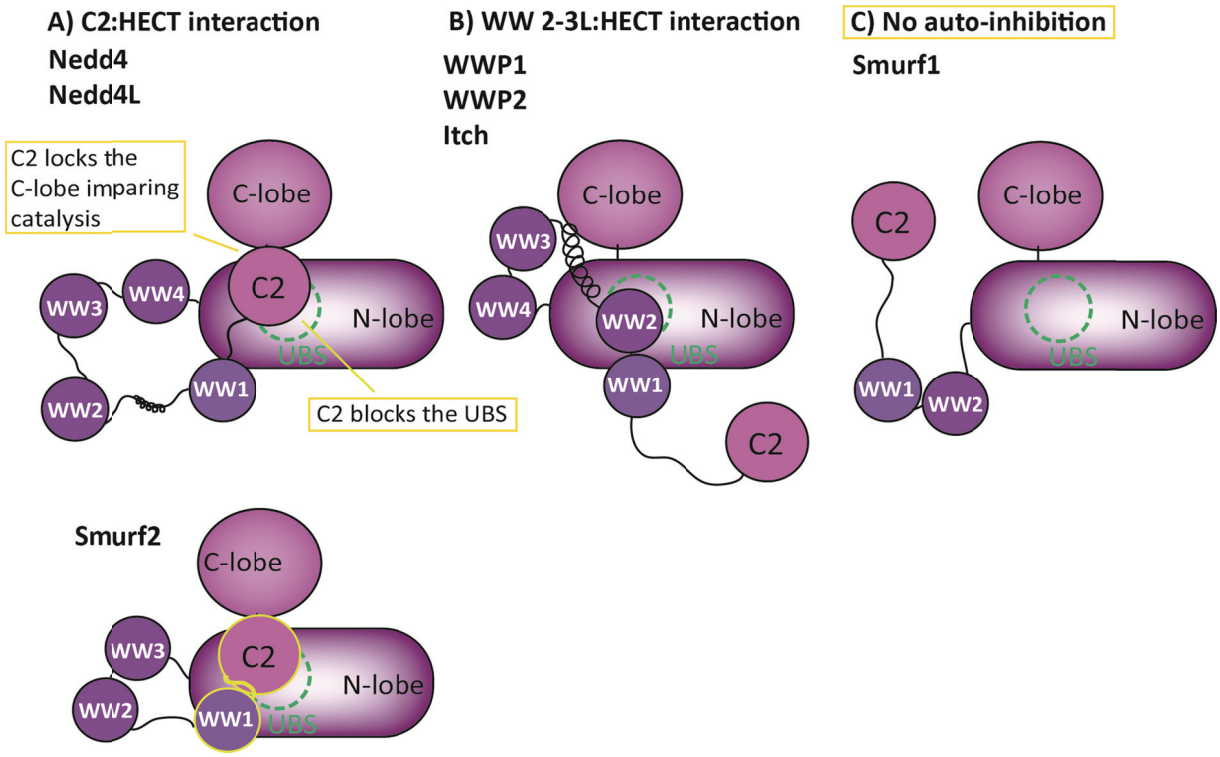


Figure 43. Update on the modes of auto-inhibition within the E3 Nedd4-family. Within the first group of Nedd4 ligases (A) a separate mechanism for Smurf2 can be distinguished, where the C2 domain binds to the C2-WW1L (marked in yellow), orienting the WW1 domain to bind to the HECT domain together with the C2 domain. As already mentioned, the binding of the entire C2-linker-WW1 shows a much stronger affinity to the HECT domain than the C2 domain alone. A new classification (C), highlighted in yellow, is added to the figure for Nedd4 members that lack an auto-inhibition mechanism, as it is the case for Smurf1, where the lack of a WW domain precludes C2:HECT interaction.

Zhu et al., 2017 claimed that the C2-mediated auto-inhibition in Nedd4 depends not only on the C2 domain, but that a α-helix located between the WW1 and WW2 domains

contributes to inhibition. It seems then, that the C2 domain is acting alone neither for Smurf2 nor for Nedd4 and that for each protein different regions can "help" the C2 domain to mediate down-regulation. The participation of other domains, e.g. WW domains, or linker regions in addition to the C2 domain, might allow the enzyme to control its activity tighter or at multiple levels. Another relevant point is that unstructured regions or linkers between domains can contribute significantly to activity regulation. Chen et al., 2017 and Zhu et al., 2017 showed that the linker between the WW2-3L domains is responsible for the inhibition interaction, what was only elucidated by crystallization of large linker-containing fragments. In addition, my NMR experiments revealed that an interaction between the C2 domain and the C2-WW1L is what orients the C2 and the WW1 domains. Furthermore, it is the presence of this short linker that increases the affinity of the interaction between the C2 and the HECT domains by 20-fold, contributing to the generation of the model, what would not have been possible considering only the direct interaction between the C2 and the HECT domain.

One of the biggest differences among Nedd4-family members is the number of WW domains, but also their distribution within the protein (Figure 44).

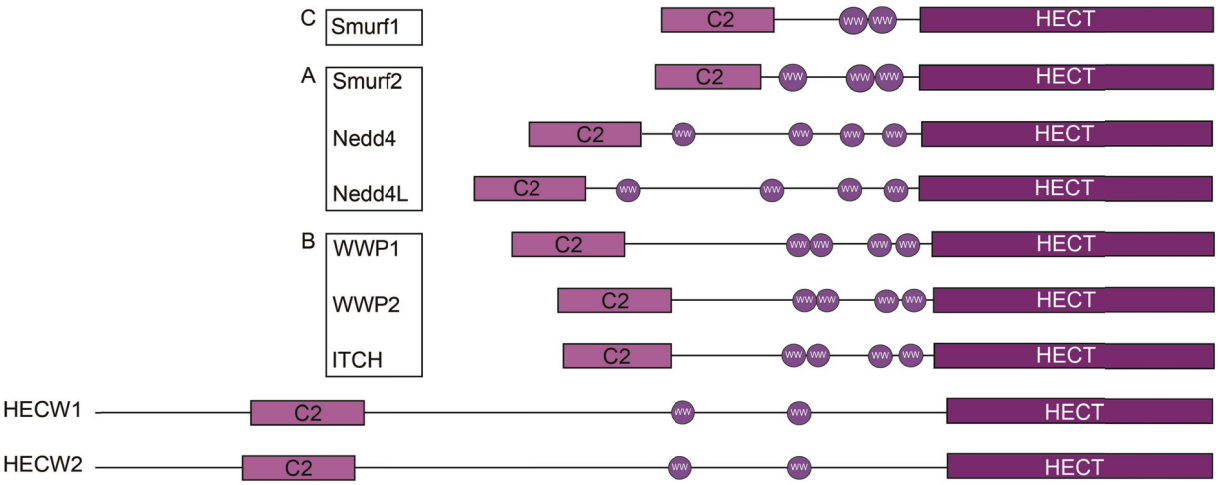


Figure 44. Domain organization within the Nedd4-family. The domains and the linker regions are drawn to scale. The three different groups of enzymes defined in Figure 43 are marked with boxes and the corresponding letter. Proteins belonging to each group have a more similar domain organization, in particular considering the distances between the C2 and the WW domains and among each WW domain.

Smurf1, the only constitutively active member of the family, bears only two WW domains, and it seems that this is precluding auto-inhibition. Smurf2 is the only member with 3 WW domains. All other nedd4 enzymes possess four WW domains. However, there seems to be a clear difference between enzymes in group A) formed by Smurf2, Nedd4 and Nedd4L where the first WW1 domain is located close to the C2 domain, with a linker

region of 17, 41 and 40 residues respectively, and enzymes in group B) that is composed of WWP1, WWP2 and Itch, where the first WW1 domain is far from the C2 domain having linker regions of 210, 159 and 148 residues respectively. Thus, I speculate that the number and location of the WW domains for each member determines a particular spatial configuration of domains, which facilitates that the C2 domain reaches the HECT domain (group A, Figure 44A) or alternatively the WW2-3L (group B, Figure 44B) fulfils this function. Finally, HECW1 and HECW2 contain only two WW domains like Smurf1, but on the contrary, they are double-size and carry a different domain distribution, what makes it difficult to predict what could happen in their cases.

Another aspect of regulation concerns enzyme activation. Since Nedd4-family members are kept in an inhibited conformation, they need a mechanism to release it. Basically, three main mechanisms have been described and all consist of different strategies to disrupt the interaction between the C2 domain or the WW2-3L region and the HECT domain. One of them involves adaptor proteins, which bind to the HECT domain and out-compete the inhibitory domain(s). This is the case for Smurf2, where Smad7 controls the E3 activity at multiple levels. It binds to the HECT domain disrupting the C2 domain binding and at the same time, recruits the E2 facilitating transthiolation (Ogunjimi et al., 2005). Activation by adaptor binding also happens for Itch and Nedd4, which interact with the adaptor protein NDFIP1-2 through their WW domains to release the HECT domain (Mund & Pelham, 2009). This strategy was also described for WWP2 and the protein Dvl2; upon Dvl2 polymerization it binds to the WW domains disrupting the interaction with the HECT domain and promoting HECT activity (Mund et al., 2015). For WWP1 not much is known; Courivaud et al., 2015 claims that it is activated by Smad7, but it is not clear whether this protein binds to the C2 or to the WW domains. A second mechanism is release of auto-inhibition through phosphorylation. The first example was described early on for Itch; phosphorylation of three residues in a Pro rich region activates the enzyme (Gallagher et al., 2006, Zhu et al., 2017). WWP2 was also found to be regulated by phosphorylation, when two Tyr residues present in the WW2-3L are modified the enzyme becomes active (Chen et al., 2017). Tyr phosphorylation in the HECT and the C2 domain disrupts their interaction for Nedd4, resulting in an active HECT domain (Persuad et al., 2014). Finally, a third mechanism, described for Nedd4 enzymes, is the Ca^{2+}-mediated C2 disruption and re-localization to the plasma membrane (Wang et al., 2010).

The overview of the different activation mechanism shows, that different from the auto-inhibition, it is not possible to group them according to particular features of each member. It seems that every protein can be subject to any of these activating mechanisms, as is the case for Nedd4 that is regulated by an adaptor protein, a PTM and Ca^{2+} influx. In this context, I studied the third mechanism described here, Ca^{2+}-mediated release of the C2 domain within the Nedd4 family (Chapter 4.3). My results confirm that

the Nedd4 C2 domain can interact with Ca^{2+} in solution (Figure 35, 36) through the conserved Asp and Asn residues located in the CBR. Also, I showed that the addition of $CaCl_2$ to the C2:HECT complex disrupts their binding (Figure 37). The NMR analysis of Smurf1, Smurf2 and Rsp5 (the yeast homolog to Nedd4) revealed that these C2 domains are not able to bind Ca^{2+} (Figure 38) and therefore most likely are not regulated by this mechanism. Ca^{2+}-mediated C2 release is thus restricted to Nedd4 and Nedd4L. From a structural alignment considering the CBR region from all Nedd4-family members (Figure 41), it is possible to explain that Smurf1, Smurf2 and Rsp5 do not coordinate Ca^{2+} as a consequence of loss of conservation in the CBR3. The other three members (Itch, WWP1 and WWP2) do not show any conservation in this region, reason why they were not included in the study. Of note, neither of them is inhibited by a C2-mediated mechanism, thus it is possible to speculate that the conservation of this region is less relevant for them, since an activation mechanism involving the C2 domain might not be useful.

As final considerations, the results of this thesis underline the relevance of performing detailed studies about the molecular mechanisms that govern the activity of individual proteins, even within a family such as the Nedd4-family of E3s, which display a remarkably high level of conservation in sequence, structure and domain organization. Huge activity differences as I found between Smurf1 (fully active) and Smurf2 (inhibited), in proteins that share more than 70% of amino acid sequence, were highly unexpected. Moreover, since these enzymes lie at the cross road between tumor suppression and oncogenesis, as well as other pathologies, understanding their differential regulation might be helpful for the design of specific therapeutics.

6. References

Akutsu M, Dikic I, Bremm A. (2016). Ubiquitin chain diversity at a glance. Cell Sci 129: 875-880.

Alva V, Nam SZ, Söding J, Lupas AN. (2016). The MPI Bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. pii: gkw348.

Bernassola F, Karin M, Ciechanover A, Melino G. (2008). The HECT Family of E3 Ubiquitin Ligases: Multiple Players in Cancer Development. Cancer Cell 14: 10-21.

Berndsen CE, Wolberger C. (2014). New insights into ubiquitin E3 ligase mechanism. Nat Struct Mol Biol 21(4):301-7.

Boase NA, Kumar S. (2015). NEDD4: The founding member of a family of ubiquitin-protein ligases. Gene 557(2): 113-22.

Buetow L, Huang DT (2016) Structural insights into the catalysis and regulation of E3 ubiquitin ligases. Nat Rev Mol Cell Biol 17, 626–642.

Cao Y, Zhang L. (2013). A Smurf1 tale: function and regulation of an ubiquitin ligase in multiple cellular networks. Cell Mol Life Sci 70 (13): 2305-17.

Chen Z, Jiang H, Xu W, Li X, Dempsey DR, Zhang X, Devreotes P, Wolberger C, Amzel LM, Gabelli SB, Cole PA. (2017). A Tunable Brake for HECT Ubiquitin Ligases. Mol Cell 66(3): 345-357.e6.

Cheng PL, Lu H, Shelly M, Gao H, Poo MM. (2011). Phosphorylation of E3 ligase Smurf1 switches its substrate preference in support of axon development. Neuron 69: 231-243.

Christianson JC, Ye Y. (2014). Cleaning up in the endoplasmic reticulum: ubiquitin in charge. Nat Struct Mol Biol 21(4):325-35.

Corbalan-Garcia S, Gómez-Fernández JC. (2014). Signaling through C2 domains: more than one lipid target. Biochim Biophys Acta 1838(6): 1536-47.

Courivaud T, Ferrand N, Elkhattouti A, Kumar S, Levy L, Ferrigno O, Atfi A, Prunier C. (2015). Functional Characterization of a WWP1/Tiul1 Tumor-derived Mutant Reveals a Paradigm of Its Constitutive Activation in Human Cancer. J Biol Chem 290: 21007-21018.

Cvetković M & Sprangers R. (2017). Methyl TROSY Spectroscopy to Study Large Biomolecular Complexes. Modern Magnetic Resonance 1-15.

David D, Nair SA, Pillai MR. (2013). Smurf E3 ubiquitin ligases at the cross roads of oncogenesis and tumor suppression. Biochim Biophys Acta 1835: 119–128.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR (3): 277-93.

Escobedo A, Gomes T, Aragón E, Martín-Malpartida P, Ruiz L, Macias MJ. (2014). Structural basis of the activation and degradation mechanisms of the E3 ubiquitin ligase Nedd4L. Structure 22(10): 1446-57.

Essen LO, Perisic O, Lynch DE, Katan M, Williams RL. (1997). A ternary metal binding site in the C2 domain of phosphoinositide-specific phospholipase C-delta1. Biochemistry 36 (10): 2753-62.

Fajner V, Maspero E, Polo S. (2017). Targeting HECT-type E3 ligases - insights from catalysis, regulation and inhibitors. FEBS Lett 591(17): 2636-2647.

Fanga S, Weissman AM. (2004). A field guide to ubiquitylation. Cell Mol Life Sci 61: 546–1561.

French ME, Kretzmann BR, Hicke L. (2009). Regulation of the RSP5 ubiquitin ligase by an intrinsic ubiquitin-binding site. J. Biol. Chem. 284: 12071–12079.

Gallagher E, Gao M, Liu YC, Karin M. (2006). Activation of the E3 ubiquitin ligase Itch through a phosphorylation-induced conformational change. Proc Natl Acad Sci U S A 103: 1717-1722.

Göbl C, Madl T, Simon B, Sattler M. (2014). NMR approaches for structural analysis of multidomain proteins and complexes in solution. Prog Nucl Magn Reson Spectrosc 80: 26-63.

Grabbe C, Husnjak K, Dikic I. (2011). The spatial and temporal organization of ubiquitin networks. Nat Rev Mol Cell Biol 12(5):295-307.

Grice GL, Nathan JA. (2016). The recognition of ubiquitinated proteins by the proteasome. Cell Mol Life Sci 73(18):3497-506.

Hall TA. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98.

Heride C, Urbé S, Clague MJ. (2014). Ubiquitin code assembly and disassembly. Curr Biol 24(6): R215-20.

Hershko A, Ciechanover A. (1998). The ubiquitin system. Annu Rev Biochem 67: 425-479.

Hettema EH, Valdez-Taubas J, Pelham HR. (2004). Bsd2 binds the ubiquitin ligase Rsp5 and mediates the ubiquitination of transmembrane proteins. EMBO J 23: 1279–1288

Huang L, Kinnucan E, Wang G, Beaudenon S, Howley PM, Huibregtse JM, Pavletich NP. (1999). Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. Science 286: 1321–1326.

Husnjak K, Dikic I. (2012). Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. Annu Rev Biochem 81: 291-322.

Kamadurai HB, Qiu Y, Deng A, Harrison JS, Macdonald C, Actis M, Rodrigues P, Miller DJ, Souphron J, Lewis SM, Kurinov I, Fujii N, Hammel M, Piper R, Kuhlman B, Schulman BA. (2013). Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. eLife 2: e00828.

Kamadurai HB, Souphron J, Scott DC, Duda DM, Miller DJ, Stringer D, Piper RC, Schulman BA. (2009). Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. Mol Cell 36: 1095–1102.

Keeler J. (2002). Understanding NMR Spectroscopy. University of Cambridge, Department of Chemistry. Wiley.

Kerfah R, Plevin MJ, Sounier R, Gans P, Boisbouvier J. (2015). Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. Curr Opin Struct Biol 32: 113-122.

Kim HC, Steffen AM, Oldham ML, Chen J, Huibregtse JM. (2011). Structure and function of a HECT domain ubiquitin-binding site. EMBO Rep 12(4): 334-41.

Kim HC, Huibregtse JM. (2009). Polyubiquitination by HECT E3s and the determinants of chain type specificity. Mol Cell Biol (12): 3307-18.

Komander D, Rape M. (2012). The ubiquitin code. Annu Rev Biochem 81: 203–229.

Kristariyanto YA, Choi SY, Rehman SA, Ritorto MS, Campbell DG, Morrice NA, Toth R, Kulathu Y. (2015). Assembly and structure of Lys33-linked polyubiquitin reveals distinct conformations. Biochem J 467(2):345-52.

Kulathu Y, Komander D. (2012). Atypical ubiquitylation - the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. Nat Rev Mol Cell Biol 13(8): 508-23.

Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage. Nature 15: 227(5259): 680-5.

Lapid C, Gao Y. (2003). PrimerX-Automated design of mutagenic primers for site-directed mutagenesis.

Lu K, Li P, Zhang M, Xing G, Li X, Zhou W, Bartlam M, Zhang L, Rao Z, He F.(2011). Pivotal role of the C2 domain of the Smurf1 ubiquitin ligase in substrate selection. J Biol Chem 286: 16861-16870.

Mari S*, Ruetalo N*, Maspero E, Stoffregen M, Pasqualato S, Polo S, Wiesner S. (2014). Structural and Functional Framework for the Autoinhibition of Nedd4-Family Ubiquitin Ligases. Structure 22: 1639–1649. * Co-first author.

Maspero E, Valentini E, Mari S, Cecatiello V, Soffientini P, Pasqualato S, Polo S. (2013). Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. Nat Struct Mol Biol 20: 696–701.

Maspero E, Mari S, Valentini E, Musacchio A, Fish A, Pasqualato S, Polo S. (2011). Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation. EMBO Rep 12: 342–349.

Metzger MB, Hristova VA, Weissman AM. (2012). HECT and RING finger families of E3 ubiquitin ligases at a glance. J Cell Sci 125: 531–537.

Michel MA, Swatek KN, Hospenthal MK, Komander D. (2017). Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. Mol Cell (17)30617-2.

Mizushima T, Yoshida Y, Kumanomidou T, Hasegawa Y, Suzuki A, Yamane T, Tanaka K. (2007). Structural basis for the selection of glycosylated substrates by SCFFbs1 ubiquitin ligase. Proc. Natl Acad. Sci USA 104: 5777–5781.

Mund T, Graeb M, Mieszczanek J, Gammons M, Pelham HR, Bienz M. (2015). Disinhibition of the HECT E3 ubiquitin ligase WWP2 by polymerized Dishevelled. Open Biol 5, 150185.

Mund T, Pelham HR. (2009). Control of the activity of WW-HECT domain E3 ubiquitin ligases by NDFIP proteins. EMBO Rep 10, 501-507.

Narimatsu M, Bose R, Pye M, Zhang L, Miller B, Ching P, Sakuma R, Luga V, Roncari L, Attisano L, Wrana JL (2009). Regulation of planar cell polarity by Smurf ubiquitin ligases. Cell 137: 295-307.

Ogunjimi AA, Wiesner S, Briant DJ, Varelas X, Sicheri F, Forman-Kay J, Wrana JL. (2010). The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. J Biol Chem 285: 6308–6315.

Ogunjimi AA, Briant DJ, Pece-Barbara N, Le Roy C, Di Guglielmo GM, Kavsak P, Rasmussen RK, Seet BT, Sicheri F, Wrana JL. (2005). Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. Mol Cell 19: 297–308.

Ohtake F, Saeki Y, Sakamoto K, Ohtake K, Nishikawa H, Tsuchiya H, Ohta T, Tanaka K, Kanno J. (2015). Ubiquitin acetylation inhibits polyubiquitin chain elongation. EMBO Rep 16: 192-201.

Ozdamar B, Bose R, Barrios-Rodiles M, Wang HR, Zhang Y, Wrana JL. (2005). Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity. Science 307: 1603-1609.

Persaud A, Alberts P, Mari S, Tong J, Murchie R, Maspero E, Safi F, Moran MF, Polo S, Rotin D. (2014). Tyrosine phosphorylation of NEDD4 activates its ubiquitin ligase activity. Sci Signal. 7(346): ra95.

Pervushin K, Riek R, Wider G, and Wuthrich K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A, 94: 12366-12371.

Rao BD. (1989). Nuclear magnetic resonance line-shape analysis and determination of exchange rates. Meth Enzymol 176: 279–311.

Riling C, Kamadurai H, Kumar S, O'Leary CE, Wu KP, Manion EE, Ying M, Schulman BA, Oliver PM. (2015). Itch WW Domains Inhibit Its E3 Ubiquitin Ligase Activity by Blocking E2-E3 Ligase Trans-thiolation. J Biol Chem 290: 23875-23887.

Rizo J, Südhof TC. (1998). C2-domains, structure and function of a universal Ca^{2+}-binding domain. J Biol Chem 273(26): 15879-82.

Rotin D, Kumar S. (2009). Physiological functions of the HECT family of ubiquitin ligases. Nat Rev 10: 398-409.

Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y. (2010). Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. J Struct Biol (1): 34-44.

Salvat C, Wang G, Dastur A, Lyon N, Huibregtse JM. (2004). The -4 phenylalanine is required for substrate ubiquitination catalyzed by HECT ubiquitin ligases. J Biol Chem 279(18): 18935-43.

Sánchez-Tena S, Cubillos-Rojas M, Schneider T, Rosa JL. (2016). Functional and pathological relevance of HERC family proteins: a decade later. Cell Mol Life Sci (10):1955-68.

Sattler, M. and Fesik, S.W. (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. Structure 4: 1245-1249.

Schägger H. (2006). Tricine–SDS-PAGE. Nature Protocols 1 (1): 16 – 22.

Schneider CA, Rasband WS, Eliceiri KW. (2012). "NIH Image to ImageJ: 25 years of image analysis". Nature methods 9 (7): 671-675.

Sriramachandran AM, Dohmen RJ. (2014). SUMO-targeted ubiquitin ligases. Biochim Biophys Acta 1843: 75-85.

Stoffregen MC, Schwer MM, Renschler FA, Wiesner S. (2012). Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. Structure 20: 573–581.

Swatek KN, Komander D. (2016). Ubiquitin modifications. Cell Res 26(4):399-422.

Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. (2000). Recognition of the polyubiquitin proteolytic signal. EMBO J 19 (1): 94-102.

van den Ent F, Löwe J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods 67(1): 67-74.

Varshavsky A. (2012). The Ubiquitin System, an Immense Realm. Annu Rev Biochem 81: 167–76.

Vaynberg J, Qin J. (2006). Weak protein-protein interactions as probed by NMR spectroscopy. Trends Biotechnol 24: 22–27.

Verdaguer N, Corbalan-Garcia S, Ochoa WF, Fita I, Gómez-Fernández JC. (1999). Ca(2+) bridges the C2 membrane-binding domain of protein kinase Calpha directly to phosphatidylserine. EMBO J 18 (22): 6329-38.

Verdecia MA, Joazeiro CA, Wells NJ, Ferrer JL, Bowman ME, Hunter T, Noel JP. (2003). Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. Mol Cell 11: 249–259.

Vijay-Kumar S, Bugg CE, Cook WJ. (1987). Structure of ubiquitin refined at 1.8 A resolution. J Mol Biol. 194 (3): 531-44.

Vittal V, Stewart MD, Brzovic PS, Klevit RE. (2015). Regulating the Regulators: Recent Revelations in the Control of E3 Ubiquitin Ligases. J Biol Chem 290(35): 21244-51.

Wan L, Zou W, Gao D, Inuzuka H, Fukushima H, Berg AH, Drapp R, Shaik S, Hu D, Lester C, Eguren M, Malumbres M, Glimcher LH, Wei W. (2011). Cdh1 regulates osteoblast function through an APC/C-independent modulation of Smurf1. Mol Cell 44: 721-733.

Wang J, Peng Q, Lin Q, Childress C, Carey D, Yang W. (2010). Calcium activates Nedd4 E3 ubiquitin ligases by releasing the C2 domain-mediated auto-inhibition. J Biol Chem 285: 12279-12288.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. (2009). Jalview version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25 (9): 1189-1191.

Waudby CA, Ramos A, Cabrita LD, Christodoulou J. (2016). Two-Dimensional NMR Lineshape Analysis. Sci Rep 25 (6): 24826.

Wauer T, Swatek KN, Wagstaff JL, Gladkova C, Pruneda JN, Michel MA, Gersch M, Johnson CM, Freund SM, Komander D. (2015). Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. EMBO J 34: 307-325.

Weissman AS. (2001). Themes and variations on ubiquitylation. Nat Rev 2: 169- 178.

Wiesner S., Sprangers R. (2015) Methyl groups as NMR probes for biomolecular interactions. Curr Opin Struct Biol, 35: 60-67.

Wiesner S, Ogunjimi AA, Wang H, Rotin D, Sicheri F, Wrana J, Forman-Kay JD. (2007). Autoinhibition of the HECT-Type Ubiquitin Ligase Smurf2 through Its C2 Domain. Cell 130: 651–662.

Williamson MP. Using chemical shift perturbation to characterize ligand binding. (2013). Prog Nucl Magn Reson Spectrosc 73: 1-16.

Yau R, Rape M. (2016). The increasing complexity of the ubiquitin code. Nat Cell Biol 18 (6): 579-86.

Zheng N, Shabek N. (2017). Ubiquitin Ligases: Structure, Function, and Regulation. Annu Rev Biochem 86: 129-57.

Zhou R, Snyder PM. (2005). Nedd4-2 phosphorylation induces serum and glucocorticoid-regulated kinase (SGK) ubiquitination and degradation. J. Biol. Chem 280: 4518–4523.

Zhu K, Shan Z, Chen X, Cai Y, Cui L, Yao W, Wang Z, Shi P, Tian C, Lou J, Xie Y, Wen W. (2017). Allosteric auto-inhibition and activation of the Nedd4 family E3 ligase Itch. EMBO Rep 18(9): 1618-1630.

Zou X, Levy-Cohen G, Blank M. (2015). Molecular functions of NEDD4 E3 ubiquitin ligases in cancer. Biochim Biophys Acta 1856(1): 91-106.

Zuiderweg ER. (2002). Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry 41: 1–7.

7. Acknowledgments

First of all, I would like to thank everyone that has supported me during these years, making this thesis possible.

I am especially grateful to my supervisor Dr. Silke Wiesner for giving me the opportunity to work in her group and offering me a project that turned out to be really interesting. I want to thank her for the support, the lab "freedom" and the detailed way in which she corrected my thesis, teaching me a lot in the process.

I am very thankful to Prof. Dr. Thilo Stehle for agreeing to be my university PhD supervisor. I am grateful to both him and Dr. Fulvia Bono for following my work during these years in our TAC meetings. I also want to thank Dr. Schwarzer and Dr. Proikas-Cezanne for being my PhD examiners.

I would also like to thank Dr. Simona Polo for the successful collaboration with her lab in Milan, Italy, and for the opportunity to learn there, together with Dr. Sara Mari, some of the main techniques regarding ubiquitination assays.

I really appreciate all the help and support that I have received from my colleagues in the lab. I would especially like to mention Mira for helping me with absolutely everything she could, even from the very beginning when we did not have the best understanding ☺. She was fundamental to my excellent time in Germany both inside and outside the lab. I also want to mention Magnus, Fabi, Carsten, Philip and Samira for the discussions, the nice talks, the help with experiments and all the small favors asked during this time. I also thank Samira for performing many experiments, some included in this thesis and some for the Smurf1 manuscript, Philip for the proofreading of some chapters of the thesis with really short notice, and Fabi for translating the abstract to German. I also want to thank previous lab members for providing a fun and friendly working environment, especially Timo, Matthias and Karolina, as well as Remco and all the people in his lab for sharing their time and many discussions together. I also want to include Vincent, for being available to help with the magnet when I needed it.

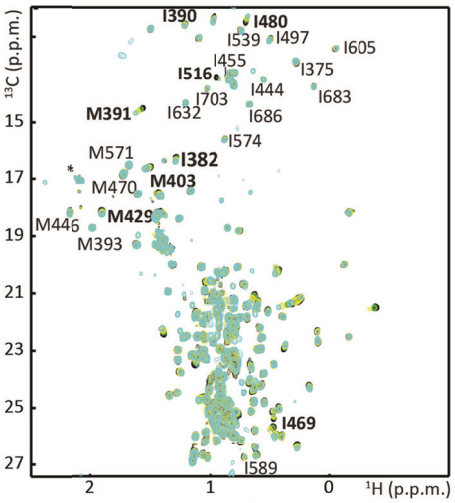
I am very grateful to Dagmar for all her support, from the interviewing process at the begging until the very last day of my PhD, answering immediately to all my needs and always being available for a talk.

I want to thank all my friends in Tübingen and abroad for always being present and helping in this long stay abroad. I am extremely thankful to my family, especially to my mother and my sister who are always there for me with their unconditional support.

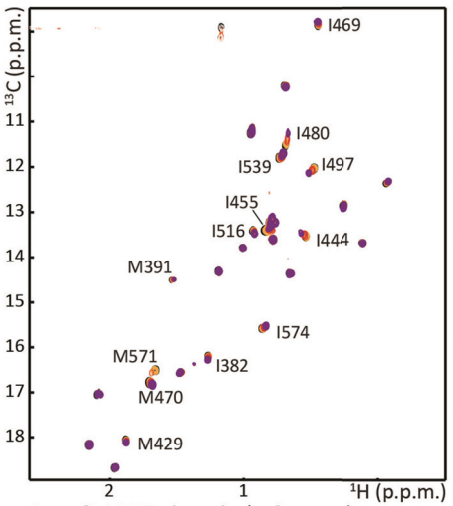
Finally I would like to dedicate this thesis to Santiago, without whom all this would never have been possible. I want to thank him for his support, from the initial decision to come to Germany at all, to every possible detail that one can imagine.

8. Supplemental Data

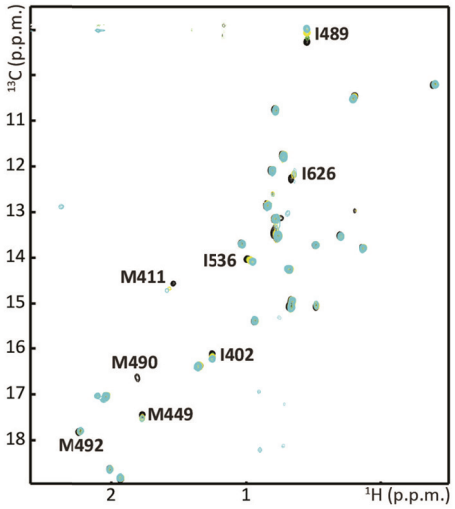
A Smurf1 HECT domain (reference)
Smurf1 HECT domain : C2 domain = 1:1
Smurf1 HECT domain : C2 domain = 1:4
Smurf1 HECT domain : C2 domain = 1:12



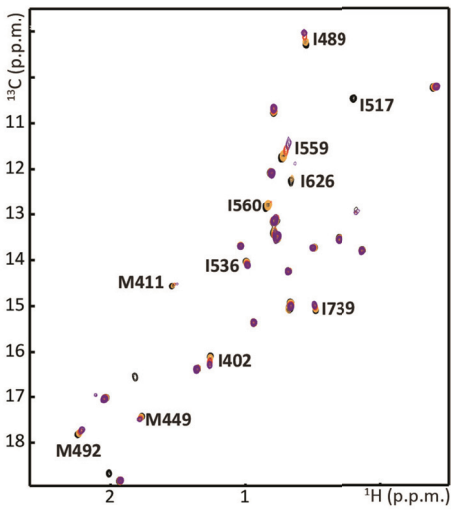
B Smurf1 HECT domain (reference)
Smurf1 HECT domain : Ub = 1:1.5
Smurf1 HECT domain : Ub = 1:5
Smurf1 HECT domain : Ub = 1:15



C Smurf2 HECT domain (reference)
Smurf2 HECT domain : C2 domain = 1:1
Smurf2 HECT domain : C2 domain = 1:4
Smurf2 HECT domain : C2 domain = 1:12



D Smurf2 HECT domain (reference)
Smurf2 HECT domain : Ub = 1:1.5
Smurf2 HECT domain : Ub = 1:6
Smurf2 HECT domain : Ub = 1:15



E Smurf2 HECT domain (reference)
Smurf2 HECT domain : C2WW1 domain = 1:0.5
Smurf2 HECT domain : C2WW1 domain = 1:1
Smurf2 HECT domain : C2WW1 domain = 1:12

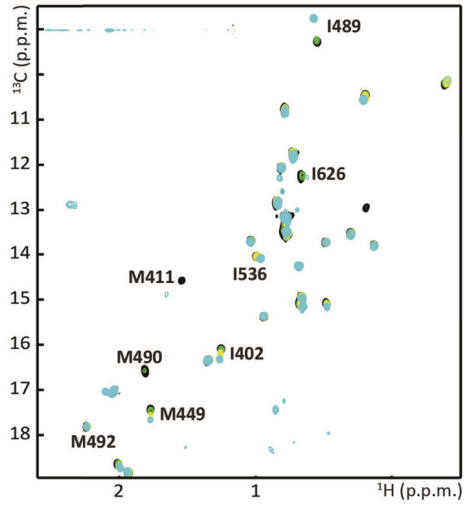


Figure S1. Full view spectra of titration experiments used for mapping interaction surfaces and K_d determination. $^{1}H,^{13}C$-methyl TROSY spectra using either an ILVAM-labeled sample (**A**) or IM-labeled of Smurf1 HECT domain (**B-E**). All residues involved in the interactions are assigned and labeled. **A)** Smurf1 HECT domain in the absence and presence of equimolar or increasing amounts of Smurf1 C2 domain (four-fold and 12-fold stoichiometric excess). **B)** Smurf1 HECT domain in the absence and presence of increasing amounts of Ub (1.5-fold, six-fold and 15-fold stoichiometric excess). **C)** Smurf2 HECT domain in the absence and presence of equimolar or increasing amounts of Smurf2 C2 domain (four-fold and 12-fold stoichiometric excess). **D)** Smurf1 HECT domain in the absence and presence of increasing amounts of Ub (1.5-fold, six-fold and 15-fold stoichiometric excess). **E)** Smurf2 HECT domain in the absence and presence of half equimolar, equimolar or 12-fold stoichiometric excess.

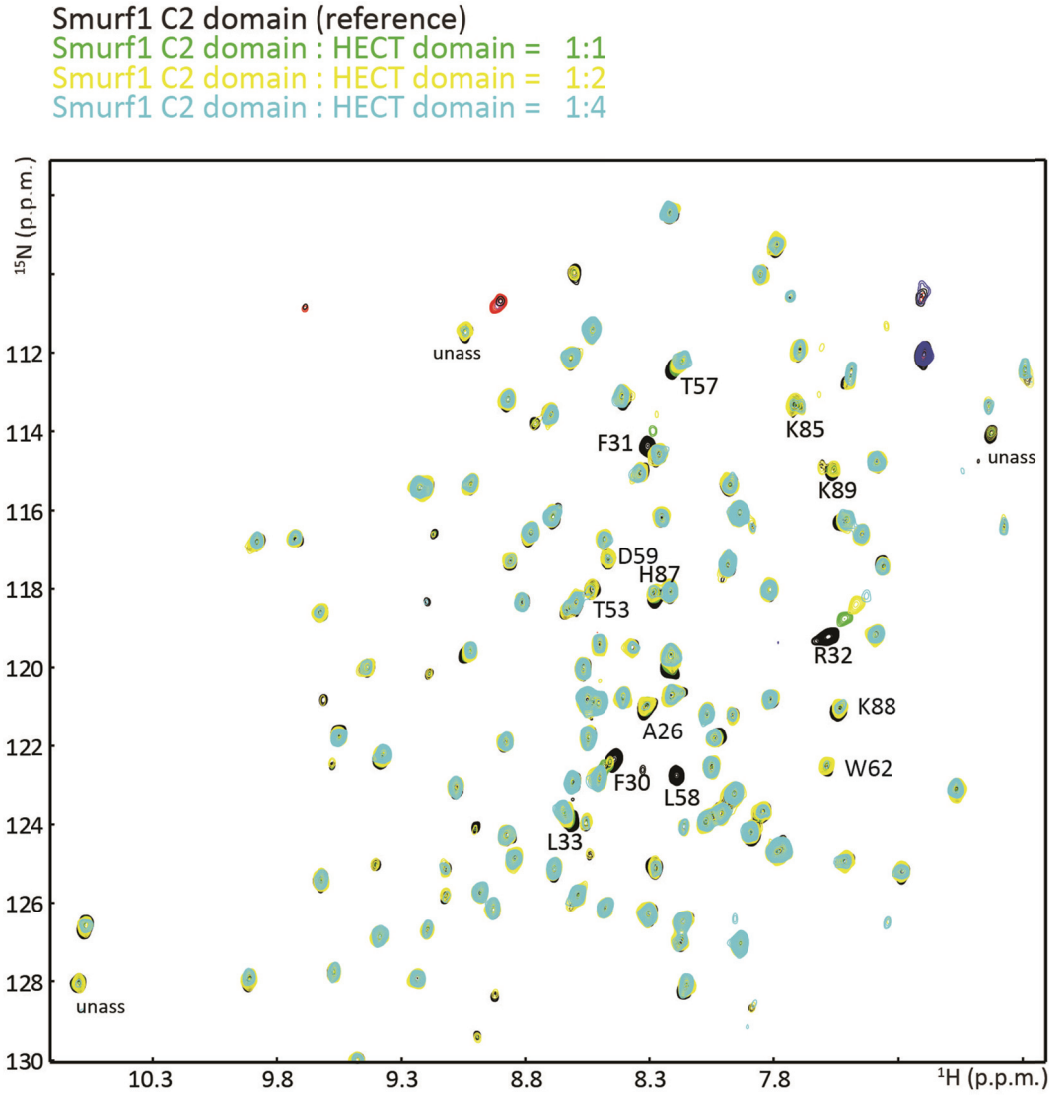


Figure S2. Full view spectra of titration experiment used for mapping Smurf1 C2: HECT interaction. $^{1}H,^{15}N$-HSQC spectra of an ^{15}N-labeled Smurf1 C2 domain in the absence and presence of equimolar or increasing amounts of Smurf1 HECT domain (two-fold and four-fold stoichiometric excess). All residues involved in the interaction are assigned and labeled.

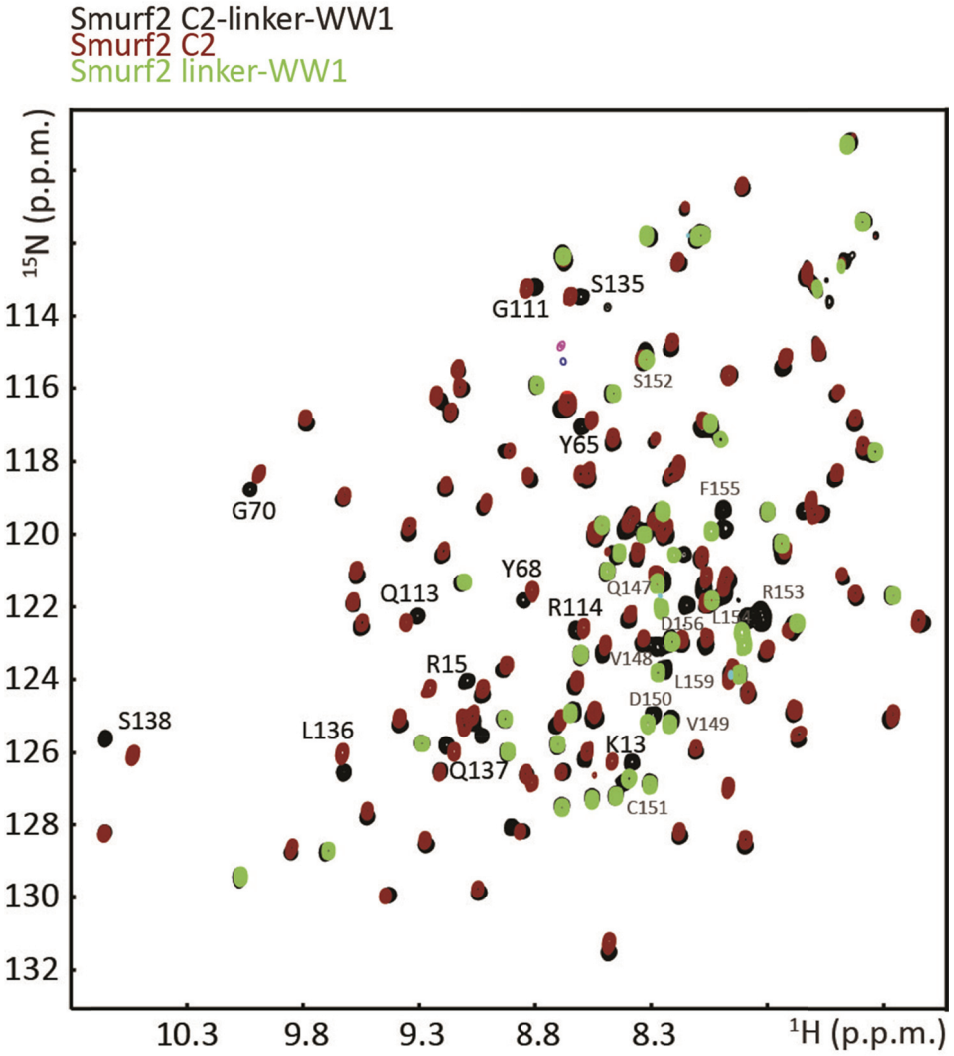


Figure S3. Full view of overlaid spectra used for mapping the binding site between the Smurf2 C2 domain and the C2-WW1L region. $^{1}H,^{15}N$-HSQC spectra of Smurf2 C2-linker-WW1 region (black), C2 domain (dark red) and linker-WW1 (light green). The residues displaying CSPs which belong to the C2 domain are labeled in black, while the ones which belong to the C2-WW1L region are labeled in grey with a smaller font size.