

SUBSTRATE RECOGNITION BY THE UBR FAMILY OF E3 UBIQUITIN LIGASES

By

Juliana Muñoz Escobar

Department of Biochemistry

McGill University

Montreal, Quebec, Canada

June 2017

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

©Juliana Muñoz Escobar, June 2017

Para Papá, Mamá y Paola

ABSTRACT

The N-end rule pathway controls the half-life of proteins in the cell based on the identity of their N-terminal residue. Recognition is mediated by E3 ubiquitin ligases called N-recognins, which directly bind destabilizing N-terminal residues (N-degrons) in substrates, catalyze the attachment of the ubiquitin molecule and direct proteasomal degradation. Primary N-degrons are classified in two groups: type 1, composed of basic residues (arginine, lysine and histidine), and type 2, composed of bulky hydrophobic residues. A specific domain termed the UBR-box recognizes type 1 N-degrons. The mammalian genome encodes seven UBR-box-containing proteins: UBR1 through UBR7. UBR1, UBR2, UBR4 and UBR5 bind N-degrons, whereas the enzymatic specificities of UBR3, UBR6 and UBR7 remain unclear. My studies have been focused on the structural and functional characterization of the UBR family of ubiquitin ligases. In particular, my interest has been to understand the molecular determinants that govern substrate recruitment by the UBR family in the N-end rule and the ubiquitin system overall.

First, I used a combination of biophysical and biochemical techniques to gain insight into how N-degrons are recognized and how the UBR-box domain optimizes binding of different ligands. Crystal structures of bound destabilizing peptides demonstrated that water molecules provide the structural plasticity required to bind different positively charged amino acids. I also revealed the ability of the UBR-box to bind methylated arginine and lysine peptides with high affinity and that the second residue is essential for binding. Finally, I demonstrated that a mutation present in the Johanson Blizzard syndrome changes the specificity for the second position by occluding the secondary pocket, thus decreasing substrate binding. These studies completed the molecular basis for N-degron recognition and revealed the plasticity of binding in the UBR-box.

Then, to understand the function of the UBR-box in other UBR proteins, I determined the crystal structure of the UBR-box from UBR6/FBXO11. The crystal structure forms a domain swapped dimer mediated by zinc coordination, where three independent protein chains come together to regenerate the topology of the monomeric UBR-box fold. Analysis of the structure suggests that the absence of N-degron binding arises from the lack of an amino acid binding pocket.

Lastly, I studied substrate-binding mechanisms in the UBR family that are not associated with the N-end rule. The C-terminus of UBR5 harbors a catalytic HECT domain and an adjacent MLLE domain. To understand the role of the MLLE domain in UBR5 I used NMR spectroscopy to characterize its binding properties. I identified a novel interaction with the adjacent HECT domain mediated by a PAM2-like sequence. This work confirmed the role of the MLLE domain of UBR5 in substrate recruitment and suggests a potential role of this domain in regulating catalytic activity.

My studies start to elucidate the great versatility of the UBR-box domain and the UBR family of ubiquitin ligases. Initially associated only with the N-end rule, these UBR-containing proteins reveal themselves as complex and multifunctional enzymes. From DNA damage response to gluconeogenesis and autophagy, the UBR family of E3 ligases exploits diverse mechanisms to control protein homeostasis in the cell.

RÉSUMÉ

La demi-vie des protéines cellulaires est contrôlée par une voie de dégradation basée sur la règle du N-terminal – elle dépend de l'identité du premier résidu de la protéine. Celui-ci est reconnu par les ubiquitine ligases E3 appelées N-recognines, qui se lient directement aux résidus N-terminaux déstabilisants (N-dégrons) des protéines substrats, et catalysent l'ajout de molécules d'ubiquitine, suivi de la dégradation par le protéasome. Il existe deux groupes de N-dégrons primaires : ceux de type 1 consistent d'acides aminés basiques (arginine, lysine, histidine) tandis que ceux de type 2 sont composés de résidus hydrophobes larges. Un domaine spécifique, le UBR-box, reconnaît les N-dégrons de type 1. Les mammifères possèdent sept protéines contenant ce domaine, appelées UBR1 à UBR7. UBR1, UBR2, UBR4 et UBR5 se lient aux N-dégrons, tandis que la spécificité enzymatique d'UBR3, UBR6 et UBR7 est inconnue. Mes recherches ont porté sur la caractérisation structurale et fonctionnelle de la famille d'ubiquitine ligases UBR. En particulier, je voulais élucider les déterminants moléculaires qui contrôlent le recrutement des substrats par la famille UBR dans le contexte de la règle du N-terminal et dans le système ubiquitine-protéasome en général.

Initialement, j'ai employé une combinaison de méthodes biophysiques et biochimiques pour déterminer les façons dont les N-dégrons sont reconnus, et comprendre comment le domaine UBR-box est optimisé pour la liaison à différents ligands. Les structures cristallines en complexe avec des peptides déstabilisants ont établi que des molécules d'eau rendent possible la liaison à différents acides aminés cationiques. J'ai aussi révélé la capacité de l'UBR-box à se lier en haute affinité à l'arginine et la lysine méthylées; le deuxième résidu du substrat est aussi essentiel à ce phénomène. De plus, j'ai démontré qu'une mutation causant le syndrome de Johanson-Blizzard affecte la spécificité à la deuxième position en bloquant la poche secondaire de liaison et affaiblissant la liaison au substrat. Ces travaux complètent la caractérisation de la base moléculaire de la reconnaissance des N-dégrons et révèlent la souplesse de liaison par l'UBR-box.

Ensuite, afin de clarifier la fonction de ce domaine dans les protéines autres que les N-recognines, j'ai déterminé la structure cristalline de l'UBR-box de UBR6/FBXO11. La structure

consiste en un dimère formé par échange de domaines supporté par la coordination d'ions de zinc. Le domaine complet UBR-box monomérique est ainsi reconstitué par trois chaînes de protéines indépendantes. L'analyse de cette structure suggère que l'incapacité de liaison aux N-dégrons est due à l'absence de poche de liaison aux acides aminés.

Finalement, j'ai investigué des mécanismes de liaison aux substrats non reliés à la règle du N-terminal. Le C-terminal d'UBR5 contient un domaine catalytique HECT ainsi qu'un domaine MLLE adjacent. Ce dernier, lorsqu'il est présent dans les protéines de liaison à la queue poly(A), participe à l'interaction entre protéines par la reconnaissance du motif PAM2. Le domaine HECT, par contre, catalyse directement le transfert de l'ubiquitine. Afin de déchiffrer le rôle du domaine MLLE d'UBR5, j'ai employé la résonance magnétique nucléaire pour caractériser ses propriétés de liaison. J'ai identifié une nouvelle interaction entre celui-ci et le domaine HECT, formée par l'intermédiaire d'une séquence semblable au PAM2. Ces résultats confirment le rôle du domaine MLLE d'UBR5 dans le recrutement des substrats, suggérant aussi une fonction de régulation de l'activité catalytique.

Mes recherches ont commencé à élucider la versatilité du domaine UBR-box et de la famille d'ubiquitine ligases E3 UBR. Celles-ci étaient initialement associées seulement à la règle du N-terminal, mais sont maintenant reconnues en tant qu'enzymes complexes et multifonctionnelles. Allant de la repose aux dommages de l'ADN à la néoglucogenèse et l'autophagie, ces enzymes emploient des mécanismes variés pour contrôler l'homéostasie des protéines cellulaires.

TABLE OF CONTENTS

ABSTRACT	iv
RÉSUMÉ	vi
TABLE OF CONTENTS	viii
LIST OF FIGURES	xi
LIST OF TABLES	xii
LIST OF ABBREVIATIONS	xiii
PREFACE	xv
CONTRIBUTIONS OF AUTHORS	xvi
ORIGINAL CONTRIBUTIONS TO KNOWLEDGE	xvii
ACKNOWLEDGMENTS	xx
CHAPTER ONE: GENERAL INTRODUCTION	22
1.1 The Ubiquitin System	22
1.1.1 Protein ubiquitylation	23
1.1.1.1 E2 ubiquitin conjugating enzymes	25
1.1.1.2 E3 ubiquitin ligases	26
1.1.1.2.1 RING type	27
1.1.1.2.1.1 Multi-subunit RINGs	28
1.1.1.2.1.2 SCF complex	28
1.1.1.2.2 HECT-type	31
1.1.1.2.2.1 Regulation of HECT-type ligases	32
1.1.1.2.3 RBR type	34
1.2 The N-End Rule Pathway	35
1.2.1 N-degrons	36
1.2.1.1 The Ac/N-end rule pathway	37
1.2.1.2 Pro/N-end rule pathway	39
1.2.1.3 The Arg/N-end rule pathway	40
1.2.1.3.1 The UBR family of ubiquitin ligases	41
1.2.1.3.1.1 The UBR-box domain	42
1.2.1.3.1.2 Type 1 binding site	43
1.2.1.3.1.2 Type 2 recognition	44
1.2.1.3.1.3.1 N-terminal methionine as a degradation signal	45
1.2.1.3.1.4 UBR1	46
1.2.1.3.1.5 UBR2	50
1.2.1.3.1.6 UBR3	51
1.2.1.3.1.7 UBR4	53
1.2.1.3.1.8 UBR5	55
1.2.1.3.1.9 UBR6	60
1.2.1.3.1.10 UBR7	61
1.3 The proteasome	62
1.4 Concluding remarks	64

CHAPTER TWO: BOUND WATERS MEDIATE BINDING OF DIVERSE SUBSTRATES TO A UBIQUITIN LIGASE.....	67
2.1 Summary	67
2.2 Introduction.....	68
2.3 Results.....	69
2.3.1 Ultra-high resolution structure of the human UBR-box in complex with arginine N-degron ..	69
2.3.2 Crystal structure of the human UBR-box in complex with a histidine N-degron	71
2.3.3 Post-translational modifications and N-degrons	74
2.3.4 A rigid scaffold that displays multiplicity of binding	78
2.3.5 Specificity of the second and third positions of the N-degron	80
2.3.6 Johanson Blizzard syndrome (JBS) and the second position in the N-degron.....	82
2.4 Discussion	87
2.4.1 Water as the sculptor of N-degron specificity	87
2.4.2 Hydrophobicity as the second driving force in N-degron modulation	88
2.4.3 JBS and the second position of the N-degron	89
2.5 Conclusions.....	90
2.6 Materials and Methods.....	91
2.6.1 Cloning, expression and purification.....	91
2.6.2 Crystallization	91
2.6.3 Data collection, structure solution and refinement.....	92
2.6.4 Thermal shift assays	93
2.6.5 Isothermal titration calorimetry.....	93
2.6.6 Quantification and statistical analysis	93
2.6.7 Data resources	94
2.7 Acknowledgements	94
CHAPTER THREE: CRYSTAL STRUCTURE OF THE UBR-BOX FROM UBR6/FBXO11 REVEALS A DOMAIN SWAPPING MEDIATED BY ZINC BINDING	97
3.1 Summary	97
3.2 Introduction.....	98
3.3 Results and Discussion	99
3.3.1 The UBR-box domain of UBR6 is a dimer in solution.....	99
3.3.2 Domain swapping in the UBR-box domain from UBR6	99
3.4 Materials and Methods.....	105
3.4.1 Protein expression and purification.....	105
3.4.2 SEC-MALS	105
3.4.3 Crystallization, data collection and structure determination	105
3.5 Acknowledgements	106
CHAPTER FOUR: THE MLLE DOMAIN OF THE UBIQUITIN LIGASE UBR5 BINDS TO ITS CATALYTIC DOMAIN TO REGULATE SUBSTRATE BINDING.....	109
4.1 Summary	109
4.2 Introduction.....	110
4.3 Results.....	113
4.3.1 GW182 interacts with the UBR5 MLLE domain.....	113
4.3.2 Structure of UBR5 MLLE bound to a PAM2 peptide.....	115
4.3.3 UBR5 binds Paip2.....	117
4.3.4 MLLE interacts with the HECT domain of UBR5.....	118
4.3.4 The N-terminal lobe of the HECT domain contains a PAM2-like sequence.....	121
4.4 Discussion	124
4.5 Materials and Methods.....	127

4.5.1 Protein expression, purification and peptide synthesis	127
4.5.2 Isothermal titration calorimetry measurements	127
4.5.3 Crystallization	128
4.5.4 Structure solution and refinement	128
4.5.5 NMR spectroscopy	129
4.5.6 Pull-down assays	129
4.6 Acknowledgements	129
CHAPTER FIVE: GENERAL DISCUSSION AND PERSPECTIVES	131
5.1 The UBR-box fold	131
5.2 UBR-box domains that do not bind N-degrons.....	132
5.3 Substrate recognition by the UBR-box	134
5.3.1 Proline has a dual role in the N-end rule.	134
5.3.2 UBR-box and N-domain, a cooperative interaction?	135
5.4 Internal degron recognition.....	135
5.5 UBR5 as a HECT- and CRL-type ligase.....	136
5.6 Concluding remarks	136

LIST OF FIGURES

Figure 1.1. Cascade of reactions in the ubiquitin proteasome system (UPS).	24
Description of the multi-step catalytic process that controls attachment of ubiquitin to a target substrate. Ubiquitylated proteins can have either of two fates: proteasomal degradation or change of subcellular location and regulation of activity.	24
Figure 1.2. E2:Ub intermediate in complex with a RING domain from E3 ligase.....	26
Figure 1.3. Structure of the SCF E3 ligase complex.....	29
Figure 1.4. HECT- vs. RING-type E3 ubiquitin ligases.....	30
Figure 1.5. Structure of the HECT domain.....	31
Figure 1.6. The mammalian N-end rule pathway.	37
Figure 1.7. N-degron specificity and domain architecture in the UBR family of ubiquitin ligases.	42
Figure 1.8. The 26S proteasome complex.	63
Figure 2.1. Ultra-high resolution structure of the UBR-box domain in complex with destabilizing peptide RLWS at 0.79 Å.	70
Figure 2.2. Histidine N-degron binds the negative pocket through a highly structured network of water molecules.	72
Figure 2.3. ITC affinity measurements for UBR-box domains bearing the V122L and E155A mutations with type 1 N-degron tetrapeptides.	73
Figure 2.4. ITC affinity measurements of UBR-box domain from UBR2 binding to type 1 N-degrons.	74
Figure 2.5. Methylated N-terminal arginine and lysine peptides bind to UBR-box domain with high affinity.....	77
Figure 2.6. The UBR-box of UBR2 forms a rigid scaffold.	79
Figure 2.7. The second position of type 1 N-degrons is important for binding to the UBR-box.	81
Figure 2.8. The V122L mutation in Johansson Blizzard syndrome prevents binding of N-degrons with bulky residues in the second position.	85
Figure 3.1. UBR-box from UBR6 forms different dimers in solution and in the crystal structure.	100
Figure 3.2. Dimer crystal structure emulates the monomeric UBR-box fold.....	102
Figure 4.1. The GW182 PAM2-like region interacts with UBR5 MLLE.	114
Figure 4.2. Crystal structure of the UBR5 MLLE-Paip1 peptide complex.	116
Figure 4.3. Binding of Paip2 to UBR5.	118
Figure 4.4 The UBR5 MLLE domain interacts with the HECT domain.....	120
Figure 4.5. The UBR5 MLLE domain recognizes a PAM2-like peptide from the UBR5 HECT domain.....	122
Figure 5.1. Sequence alignment of the UBR-box domains in the UBR family of proteins.....	131
Figure 5.2. Secondary pocket in the UBR-box domains.	132

LIST OF TABLES

Table 1.1. Mutations in the <i>UBR1</i> gene causing Johanson Blizzard syndrome.	48
Table 2.1. Dissociation constants, enthalpy and entropy changes for the UBR-box domain from UBR2 binding different tetrapeptides as measured by ITC.....	82
Table 2.2. Collection and refinement statistics for crystallographic data.....	86
Table 3.1. Collection and refinement statistics for crystallographic data.....	104
Table 4.1. Collection and refinement statistics for crystallographic data.....	123

LIST OF ABBREVIATIONS

ADMA – Asymmetrically double methylated arginine
ATE – Arginyltransferase or arginyl-tRNA protein transferase
ATP – Adenosine triphosphate
C-lobe – C-terminal lobe
CNS – Central nervous system
CRL – Cullin-RING E3 ligase
dNTPs – deoxynucleotides
E1 – Ubiquitin activating enzyme
E2 – Ubiquitin conjugating enzyme
E2:Ub – Ubiquitin conjugated to a E2 enzyme
E3 – Ubiquitin ligase
EDD – E3 identified by differential display
HECT – Homologous to the E6AP C-terminus
HECT:Ub – Ubiquitin conjugated to a HECT domain
HIV- Human immunodeficiency virus
HPV – Human papillomavirus
HYD – Hyperplastic discs
HSQC – Heteronuclear single quantum correlation NMR experiment
ITC – Isothermal titration calorimetry
JBS – Johanson-Blizzard syndrome
 K_d – Dissociation constant
MAPK – Mitogen-activated protein kinase
STAT2 – Signal transducer and activator of transcription 2
Met-(Φ) – N-terminal methionine followed by a hydrophobic residue
MetAPs – Methionine aminopeptidase
MMA – Monomethylated arginine
mRNA – messenger RNA
N-degron – N-terminal destabilizing residue

N-lobe – N-terminal lobe
NMR – Nuclear magnetic resonance
N-recogin – E3 ubiquitin ligase that recognizes N-degrons
NTAQ – N-terminal glutamine amidase
NTAN – N-terminal asparagine amidase
ORF – Open reading frame
Pri – polished rice or tarsal-less RNA
PRMT – Protein arginine methyltransferase
RING – Really interesting new gene domain
RNAi – RNA interference
SCF – SKP1-CUL1-F-box E3 ubiquitin ligase
SDMA – Symmetrically double methylate arginine
tRNA – Transfer RNA
TSA – Thermal shift assay
Ub – Ubiquitin
UBA – Ubiquitin-associated domain
UBC – Ubiquitin-conjugating domain
UBL – Ubiquitin-like domain
UBR-box – Ubiquitin recognin box domain
UPS – Ubiquitin proteasome system

PREFACE

This is a manuscript-based thesis, consisting of two published articles and one under review.

CHAPTER TWO

Muñoz-Escobar J, Matta-Camacho E, Cho C, Kozlov G, Gehring K. Bound waters mediate binding of diverse substrates to a ubiquitin ligase. *Structure*. 25(5): 719-729 (2017)

CHAPTER THREE

Muñoz-Escobar J, Kozlov G, Gehring K. Crystal structure of the UBR-box domain from UBR6/FBXO11 reveals a domain swapping mediated by zinc binding. *Protein Science, In Press*.

CHAPTER FOUR

Muñoz-Escobar J*, Matta-Camacho E*, Kozlov G, Gehring K. The MLLE domain of the ubiquitin ligase UBR5 binds to its catalytic domain to regulate substrate binding. *J Biol Chem*. 11; 290(37): 22841-50 (2015)

*Co-first author

CONTRIBUTIONS OF AUTHORS

CHAPTER TWO

Dr. Kalle Gehring and myself conceived this project. Dr. Edna Matta-Camacho obtained the crystal of the UBR-box from UBR1 in complex with MMA-IFS peptide. I designed and performed all the rest of experiments with the assistance of Cordelia Cho. Dr. Guennadi Kozlov assisted X-ray data collection and provided useful discussions. I prepared the manuscript, which was modified by Dr. Kalle Gehring.

CHAPTER THREE

Dr. Kalle Gehring and myself conceived this project. I designed and performed all experimental work. Dr. Guennadi Kozlov assisted during X-ray data collection and structure solving. I prepared the manuscript, which was modified by Dr. Kalle Gehring.

CHAPTER FOUR

This study was coordinated by Dr. Kalle Gehring. Dr. Guennadi Kozlov crystallized the MLLE domain in complex with GW182, Dr. Edna Matta-Camacho designed and performed pull-down assays with UBR5 and Paip2. Dr. Edna Matta-Camacho and I designed, performed and analyzed NMR experiments. I prepared the initial draft of the manuscript, which was modified by Dr. Kalle Gehring.

ORIGINAL CONTRIBUTIONS TO KNOWLEDGE

CHAPTER TWO

Bound waters mediate binding of diverse substrates to a ubiquitin ligase.

- Determined the crystal structure of human UBR-box from UBR2 in complex with Arg-Leu-Phe-Ser peptide at 0.8 Å resolution. This structure confirmed the essential role of water molecules in arginine binding.
- Determined the crystal structure of the human UBR-box from UBR2 in complex with His-Ile-Phe-Ser peptide at 1.5 Å resolution, which revealed the molecular determinants for histidine recognition by N-recognins and explained the lower affinity of histidine N-degrons compared to arginine and lysine.
- Determined the crystal structure of the human UBR-box from UBR2 in complex with the asymmetrically double methylated arginine peptide ^{ADMA}Arg-Ile-Phe-Ser that illustrated the binding mechanism of methylated N-degrons.
- Determined the crystal structure of the V122L mutant of human UBR-box from UBR2, which explained the reduced affinity of peptides for the UBR-box domain.
- The structures of arginine and histidine peptides along with methylated arginine highlighted the role of water molecules as adaptors of N-degron binding.
- The identity of the second position in mammalian N-degrons is essential for binding. Hydrophobic residues are preferred in the second position whereas proline disrupts binding to the UBR-box even in the presence of N-terminal arginine.
- Analysis of the V122L mutant from Johanson-Blizzard syndrome using thermal shift assays and isothermal titration calorimetry showed a considerable decrease in peptide binding, explaining the impairment of substrate ubiquitylation observed in yeast cells.

CHAPTER THREE

Structure of the UBR-box domain from UBR6/FBXO11 reveals a domain swapping mediated by zinc binding.

- Determined the crystal structure of human UBR-box from UBR6 at 2.2 Å resolution. The dimeric crystal structure reveals a unique form of domain swapping mediated by three zinc binding sites, where three independent protein chains come together to emulate the monomeric UBR-box fold.
- The UBR-box from UBR6 forms both, dimers and monomers in solution.
- Analysis of the conserved zinc fingers suggests that the absence of N-degron binding arises from the lack of binding pockets.

CHAPTER FOUR

The MLLE domain of the ubiquitin ligase UBR5 binds to its catalytic domain to regulate substrate binding.

- The MLLE domain of UBR5 directly binds and targets Paip2 for ubiquitylation and proteasomal degradation.
- The HECT domain interacts with the MLLE domain in UBR5. NMR HSQC experiments demonstrated that the MLLE domain directly binds the N-terminal lobe of the HECT domain.
- Binding between the HECT domain and the MLLE domain in UBR5 is disrupted by PAM2 peptides.
- A PAM2-like motif located in the N-terminal lobe of the HECT domain mediates the interaction with the MLLE domain. Mutation of the highly conserved phenylalanine residue in the PAM2 motif disrupts binding.

ACKNOWLEDGMENTS

Numerous people have contributed to the realization of this thesis. First, I would like to thank my supervisor Dr. Kalle Gehring, whom gave me the invaluable opportunity to come to his lab, first as a visitor and later as a PhD student. Rarely does one receive the trust and liberty needed to explore one's ideas and I was fortunate to find a research environment that supported my own exploration while providing the necessary guidance to become a scientist. I also thank him for the excellent advice and tutoring given during manuscript completion. I thank Dr. Guennadi Kozlov for all the advice, tutoring and support given throughout these years. His guidance and expertise were essential to complete my research. I thank him also for all his patience and the good times spent during experiments and synchrotron trips. Finally, I thank him for never letting me give up. This thesis would not have been possible without the support of my laboratory colleagues, past and present. Dr. Jingwei, Dr. Meng, soon-to-be-Dr. Marjan and Seby, thanks for making the lab such a fun place, for the advice, lab outings and lunch breaks. To Kathy Wong for the fun times, the shared frustrations and the science discussions.

To my RAC committee members, Dr. Wing, Dr. Trempe and Dr. Schmeing, for their advice and for helping me become a better scientist. I specially want to thank Dr. Martin Schmeing for all the countless times where his door was open to offer advice. I thank the NSERC Training Program in Bionanomachines for allowing me to be part of such great initiative. To CIHR Chemical Biology Scholarship, GRASP and PROTEO for funding and all the excellent training opportunities. I also thank the Department of Biochemistry for funding and travel awards.

I have to specially thank Dr. Edna Matta-Camacho, which was instrumental in the completion of my research. No tengo palabras para expresar el agradecimiento y la admiración que tengo por ti. Has sido mucho más que una supervisora, eres mi mentora y una de mis mejores amigas. Sin tu ayuda y consejo mi historia sería muy diferente. A Sara Lili, mi compañera de sueños, por haber alumbrado mis días en el laboratorio. Una de mis primeras mentoras que se convirtió en otra mejor amiga. Gracias por siempre motivarme a seguir adelante. A Dieguini, otro compañero de camino y estudiante brillante, que siempre estuvo ahí para alentarme y alegrarme el día con Amanda Miguel.

I also want to thank all my wonderful friends from the 4th floor Bellini. To Shane, Yazan and Diego for the stimulating science conversations, these discussions truly made the difference. To Kriss, Jon, Yogita and Geneviève for all the great memories, parties and scientific discussions. To Anna and Janice for a wonderful “crochet” club, I feel very lucky for having such a support network. To BGSS council, for giving me the opportunity to grow with such wonderful and bright group. Gracias a mi familia en Montreal, Cata y Cami por aceptarme en su casa sin conocerme. A Diana, Kathe, Pipe A, Pipe G, Marce, Saris, Gabo, Angie, Jill y Kpto, ustedes han sido parte del soporte y alegría necesarios para alcanzar mis metas.

To Love, for all the endless times you pushed me to be better, to continue and not give up, for letting me be my best self, you are truly one of the main reasons why I can write this today. I thank also Celia, Tony and Steve, for making me feel part of such a great family. Agradezco también a todos los Muñoz Zapata. Su apoyo fue esencial para perseguir mis sueños. Gracias por recibirme en sus casas, por ofrecer su ayuda y alentarme a seguir, todo hubiera sido mas difícil sin ustedes.

Finalmente, quiero agradecer a las personas mas importantes de mi vida y gracias a las cuales puedo decir que llegué tan alto: mi familia. A mi Papá, por enseñarme que el trabajo duro siempre paga y por ser un ejemplo de rectitud y responsabilidad, por ser mi mayor fan y empujarme a construir mis sueños: esto es para ti. A mi mamá, la mujer mas fuerte que conozco, que lo dio todo por nosotras y que es un ejemplo de entrega y lucha, que me educó para luchar por grandes cosas y convertirme en una mujer independiente: mis logros son para y por ti. Para mi hermana, por ser el más grande ejemplo de superación y por ser mi soporte en los tiempos difíciles, mi sueño fue siempre llegar a ser como tú: esto es para ti. A Luciana, el regalo más grande que hemos recibido, esto también es para ti.

CHAPTER ONE: GENERAL INTRODUCTION

1.1 The Ubiquitin System

Protein ubiquitylation is one of the most frequent post-translational modifications in eukaryotic cells (Khoury, Baliban & Floudas 2011). The 76-amino acid protein ubiquitin is covalently attached via its carboxy-terminal glycine to, in most cases, a lysine residue in the target protein. Ubiquitin was first discovered in 1974 as an abundant and ubiquitous protein present in all eukaryotic cells (Schlesinger, Goldstein & Niall 1975). Its role in protein degradation was later described through the discovery of the proteasome system, which was first observed as an ATP-dependent proteolytic pathway responsible for the degradation of abnormal proteins (Etlinger & Goldberg 1977).

In eukaryotes, an ATP-dependent multi-protein complex called the proteasome is responsible for the majority of proteolysis in the cell. Together with enzymes that catalyze the transfer of ubiquitin to substrates, the proteasome forms what is known as the ubiquitin proteasome system (UPS) (Schrader, Harstad & Matouschek 2009). Proteins are targeted for ubiquitylation and proteasomal degradation through specific sequences, conformational changes or posttranslational modifications. These are called degradation signals or degrons (Varshavsky 1991). The minimal signal for degradation was thought to consist of two elements: an acceptor lysine residue for ubiquitin conjugation and a recognition site for the E3 ubiquitin ligase enzyme. However, later discoveries have demonstrated the intricate nature of proteasomal regulation and have unveiled a diversity of mechanisms that control protein turnover.

The exponential growth of research in the field over the next decades exposed the versatility of ubiquitin as both, a degradation and regulatory signal. The ubiquitin molecule was revealed as a key element in the control of subcellular location and function in cases where its conjugation to a target protein did not lead to proteolysis. Degradation-independent functions of ubiquitylation range from membrane protein trafficking and endocytosis to protein localization and regulation of transcription factors in the nucleus (Acconcia, Sigismund & Polo 2009; Hammond-Martel, Yu & Affar el 2012; Schnell & Hicke 2003; Varshavsky 2017).

The nature of the ubiquitin modification has a direct impact on the activity and fate of the target protein. Ubiquitin contains seven lysine residues (K6, K11, K27, K29, K33, K48 and K63) and an N-terminal amino group that can be modified by ubiquitylation. Therefore, ubiquitin can be attached to proteins as a single moiety on one or multiple sites, or as a poly-ubiquitin chain that can be linked through any of its seven lysine residues or its amino terminal methionine (Peng et al. 2003; Rahighi et al. 2009; Xu et al. 2009). Ubiquitin polymers can contain mixed linkages within the same chain, providing versatility and specific functionality to the posttranslational modification. In general, modifications through K48 and K11 linkages are associated with proteasomal degradation, while K63 polymers or monoubiquitylation have non-proteolytic roles. All possible linkages have been found in cells. However, the importance of linkages through K6, K27, K29 or K33 is poorly understood.

The structural diversity of the ubiquitin chain is read by the cell as a code. Ubiquitin-binding partners recognize specific structural arrangements such as the distance between moieties or the relative orientation of ubiquitin molecules within the chain (Komander & Rape 2012). Furthermore, the fate of the target protein is also determined by the combination of other factors such as cellular localization, homeostasis and sensitivity to ubiquitin removal or deubiquitylation. A tight control of protein ubiquitylation requires the concerted interaction of all enzymes involved in the attachment of ubiquitin to the target protein.

Ubiquitin-like proteins (UBLs) are molecules that share sequence and structural similarity to ubiquitin. These molecules are also posttranslationally attached to other proteins using similar enzymatic cascades as ubiquitylation. SUMO, Nedd8, ISG15 and Atg8 are some of the identified UBLs in eukaryotes. Conjugation of UBLs to proteins controls a diverse range of cellular processes that go from regulation of CRL-mediated ubiquitylation (reviewed later in this chapter) to antiviral function, autophagy and splicing, among others (van der Veen & Ploegh 2012).

1.1.1 Protein ubiquitylation

Ubiquitin chains are formed by a three-step cascade of reactions that start with the activation of the ubiquitin molecule in an ATP-dependent manner (Hershko et al. 1983). The C-terminal glycine residue in ubiquitin is activated by the E1 activating enzyme, forming an ubiquitin

adenylate intermediate, followed by the formation of a thioester bond between a cysteine residue in the E1 and ubiquitin. Activated ubiquitin is then transferred to an active cysteine in the E2 conjugating enzyme in a thioester bond. The final step in the reaction is catalyzed by an E3 ubiquitin ligase (**Figure 1.1**). Humans encode eight different E1 activating enzymes (Schulman & Harper 2009), over 40 E2 conjugating enzymes (Wenzel, D. M., Stoll & Klevit 2011) and at least 600 E3 ubiquitin ligases (Li, W et al. 2008). This bottleneck structure of enzyme diversity in the cascade provides incremental specificity to the system as reactions progress. Moreover, E2 and E3 enzymes can act individually or in multi-protein complexes, responding to the need of targeting a broad range of substrates.

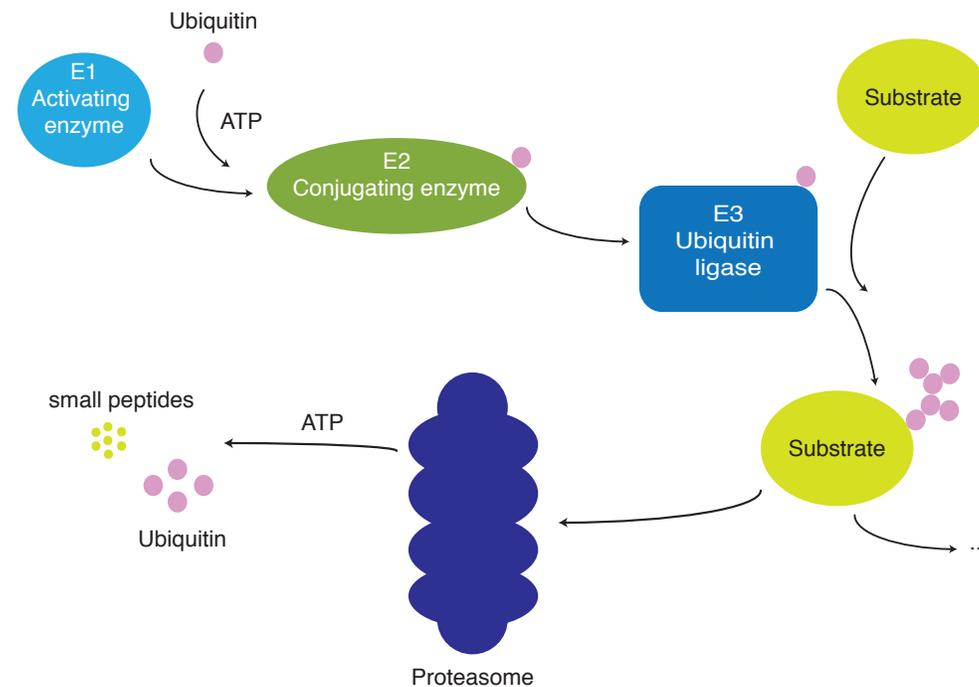


Figure 1.1. Cascade of reactions in the ubiquitin proteasome system (UPS).

Description of the multi-step catalytic process that controls attachment of ubiquitin to a target substrate. Ubiquitylated proteins can have either of two fates: proteasomal degradation or change of subcellular location and regulation of activity.

1.1.1.1 E2 ubiquitin conjugating enzymes

The E2 conjugating enzyme family is composed by 17 subfamilies of proteins based on phylogenetic analysis (Michelle et al. 2009). All E2 enzymes share a ~150-residue conserved catalytic domain, the Ubiquitin Conjugation (UBC) domain, which contains a catalytic cysteine residue that forms a thioester bond with ubiquitin and binds the E1 enzyme (Wenzel, D. M., Stoll & Klevit 2011). The E1-E2 interaction ensures specificity towards ubiquitin, neglecting ubiquitin-like modifiers (UBLs) such as SUMO, Nedd8, etc. (Schulman & Harper 2009).

E2 conjugating enzymes regulate ubiquitin chain topology and processivity of the ubiquitin transfer (**Figure 1.2**). Versatility in the branching structure and length of the ubiquitin chain work as a powerful regulatory code (Komander & Rape 2012), as various types of lysine and N-terminal linkages trigger completely distinct outcomes in the cell. Some E2s forms only one specific ubiquitin linkage while others generate different types of ubiquitin branching. Moreover, the decision of whether lysine in the ubiquitin moiety (chain elongation) or in the substrate (additional ubiquitylation site) will receive the next ubiquitin is made by the E2. Remarkably, ubiquitin chain initiation and elongation requires specific E2-E3 interactions, in some cases involving more than one E2 enzyme, each with distinct initiation and elongation roles (Christensen, Brzovic & Klevit 2007; Jin, L et al. 2008; Kleiger et al. 2009; Rodrigo-Brenni & Morgan 2007).

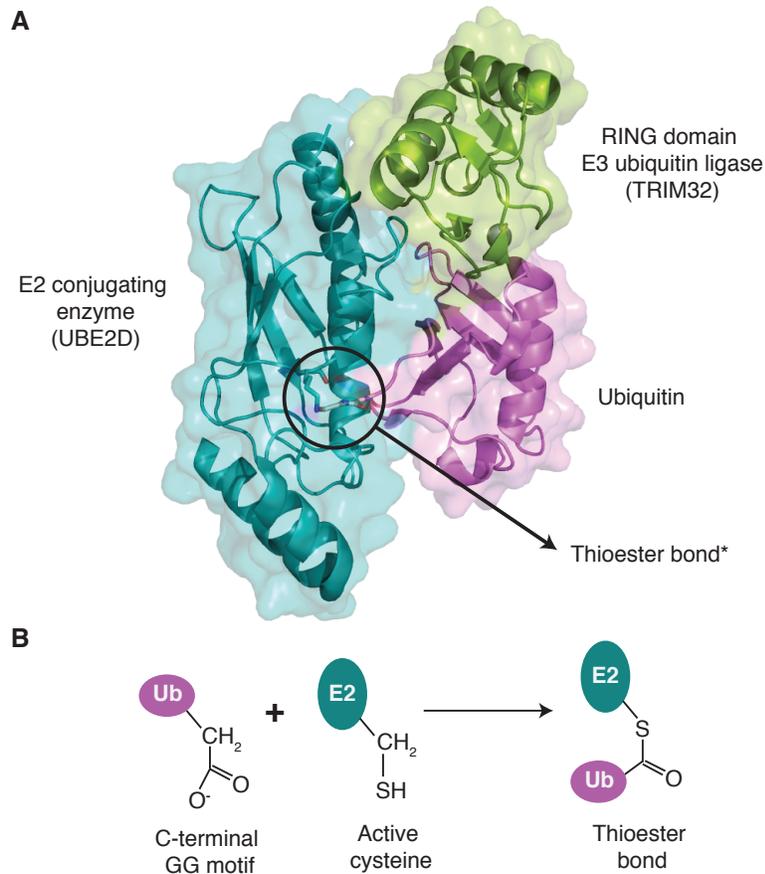


Figure 1.2. E2:Ub intermediate in complex with a RING domain from E3 ligase.

A. Structure of the UBE2D E2 conjugating enzyme covalently bound to ubiquitin. *C85K mutation allows isopeptide bond formation instead of thioester bond, a more stable intermediate for crystallization. PDB: 5FER. The RING domain of the E3 ubiquitin ligase TRIM32 is bound to the E2 enzyme. **B.** Thioester bond formation between the carboxyl-terminal group in the Gly-Gly motif of ubiquitin and the active cysteine (C85 in UBE2D) from the E2 conjugating enzyme.

1.1.1.2 E3 ubiquitin ligases

E3 ubiquitin ligases interact with the E2:Ub conjugate and the substrate protein to catalyze the formation of an isopeptide bond between the carboxy-terminus of ubiquitin and the ϵ -amino group of a lysine residue in the substrate (**Figure 1.1**). The selectivity of the pathway relies mainly on the ability of the E3 ligase to directly bind specific recognition signals or degrons in

target proteins. This heterogeneous class of enzymes is characterized by the presence of at least two functional domains: a domain that mediates interactions with the E2 conjugating enzyme and a specific substrate recognition domain. In general, E3s function in one of two mechanisms: they serve as catalytic intermediates for the ubiquitin transfer, or they assist as a scaffold for the direct linkage of ubiquitin from the E2 to the substrate. Based on the type of E2-binding domain and the mechanism involved in ubiquitin transfer, E3 ubiquitin ligases can be classified in three groups: RING type, HECT type, and RBR ligases.

1.1.1.2.1 RING type

The Really Interesting New Gene (RING) type is the most abundant group of E3 ubiquitin ligases. They function as mediators in the transfer of ubiquitin from the E2 enzyme to the substrate, serving as a scaffold to orient the charged E2 and its substrate. The E3 RING type is characterized by the presence of a RING or a U-box domain. The RING domain is a zinc-finger of 40-60 amino acids that coordinates two zinc atoms in a cross brace arrangement. In contrast, U-box domains do not coordinate zinc but adopt a similar fold to the RING domain. A remarkable characteristic of RING-type ligases is their ability to form dimers. Homo- and heterodimers have a distinctive capacity to regulate E2 interactions and activity, where the availability of the RING domain to interact with the E2 can be either enhanced or abolished upon dimerization (Brzovic et al. 2001; Dou et al. 2012; Liew et al. 2010; Poyurovsky et al. 2007; Uldrijan, Pannekoek & Vousden 2007; Zhang, L et al. 2011). Dimers are formed either through sequences found outside the RING domain or through the RING domain itself. Location of the RING domain in the primary sequence does not seem to determine the ability of the E3 to dimerize.

Substrates of RING-type E3 ligases are greatly diverse as are the numerous mechanisms to regulate their ubiquitylation. Regulation of RING-type ligases can occur posttranslationally, at the transcriptional level or via feedback mechanisms involving specific metabolites that interact with the ligase or its substrate (Lee, JN et al. 2006; Lipkowitz & Weissman 2011; Zhang, L et al. 2011). RING-type E3s display a variety of roles involving tightly regulated multi-step processes acting as tumour suppressors, oncogenes, activators of the DNA damage response or NF- κ B signalling, among others (Metzger et al. 2014). RING E3s can have multiple substrates and

different ligases can modify the same target protein. One particular case is that of the yeast Ubr1p protein, a RING E3 ligase that coordinates ubiquitylation of certain substrates along with the HECT E3 Ufd4. Interaction between the two ligases increases processivity of the ubiquitin transfer (Hwang, C-S et al. 2010).

1.1.1.2.1.1 Multi-subunit RINGs

Some RING-type ligases exist as an assembly of multiple proteins. One example is the Cullin-RING ligase (CRL) family. This group of enzymes exhibits great diversity in terms of subunit composition and function. Humans encode seven different cullin proteins (CUL1, 2, 3, 4A, 4B, 5 and 7), all sharing a common molecular architecture presumably to further expand the repertoire of substrates and diversify regulation. An active complex is formed by one or several cullin (CUL) proteins, a RING-containing protein called RBX1 (Roc1/Hrdt), and different protein adaptors that recognize and deliver substrates for ubiquitylation (Hua & Vierstra 2011). The identifier protein in these complexes is the CUL scaffold protein. The C-terminal region of the CUL protein interacts with the N-terminal domain of RBX1 to create a catalytic core. The C-terminal RING domain RBX1 docks the charged E2 promoting ubiquitin transfer directly to the substrate. The N-terminal region of the CUL subunit binds specific substrate adaptor receptors such as F-box and DCAF proteins, among others (**Figure 1.3**). Studies using a CRL-specific inhibitor suggest that CRL ligases mediate at least 20% of proteasomal-dependent protein turnover in mammalian cells (Soucy et al. 2009). CRLs are activated by the covalent attachment of a single ubiquitin-like protein named Nedd8 to a conserved lysine residue in the CUL subunit (Kamura et al. 1999). Cycles of attachment and removal of the Nedd8 moiety serve as modulators of ubiquitin activity. Neddylation enhances cullin-dependent ubiquitylation *in vitro*, while removal of Nedd8 from cullin decreases E2 recruitment and activation (Wu, Chen & Pan 2000).

1.1.1.2.1.2 SCF complex

The archetypal CRL is the SKP1-CUL1-F-box (SCF) complex (**Figure 1.3**). SCFs are critical regulators of cell cycle and its misregulation often results in oncogenic events (Frescas & Pagano 2008). The S-phase kinase-associated protein (SKP1) binds the N-terminus of CUL1 working as

a docking site for different substrate adaptors. The substrate binding subunits in the complexes are the members of the cyclin-F (F-box) family. F-box proteins share a conserved N-terminal ~40 amino-acid domain called the F-box (Skowyra et al. 1997) that directly interacts with SKP1. Each F-box protein consists of at least two major functional domains: the F-box domain and one or more C-terminal domains that bind substrates for ubiquitylation (Bai et al. 1996).

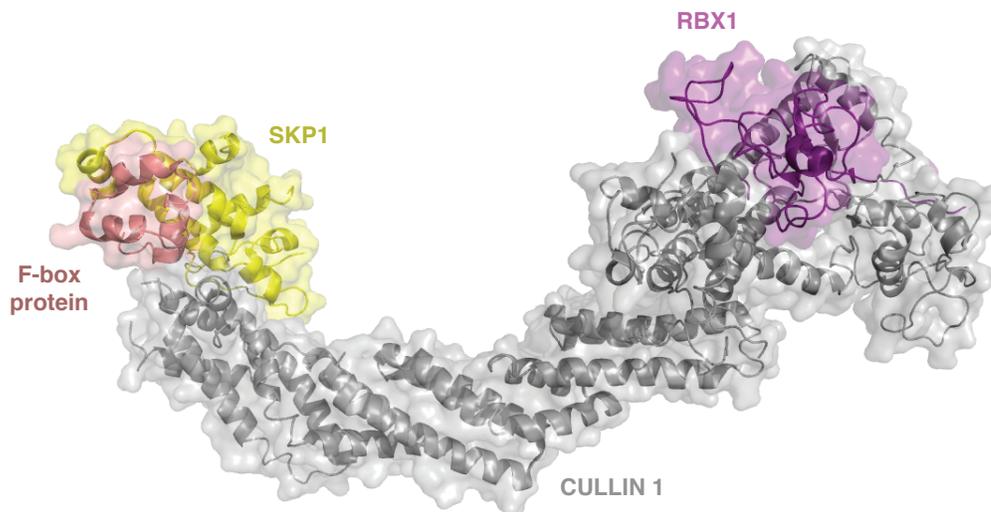


Figure 1.3. Structure of the SCF E3 ligase complex.

Crystal structure of the SKP1-F-box(SKP2)-Cullin1-RBX1 (SCF) E3 ligase complex. The SKP1 protein binds the F-box domain in SKP2 (F-box protein in salmon), the Cullin1 protein (CUL1, in grey) serves as a bridge between the F-box protein, which recognizes the substrate, and the RBX1 protein (purple), which recruits the charged E2 enzyme. PDB: 1LDK.

F-box proteins are classified in three groups according to their substrate binding domains: The FBXW family, characterized by the presence of WD40 repeat domains; the FBXL family, including SKP2, which contains leucine-rich repeat domains; and the FBXO family, with various substrate domains not fully characterized (Wang, Z et al. 2014). Substrate recognition by F-box proteins often requires post-translational modifications in the target protein. Phosphorylation of a degron motif in specific substrates is the most common regulatory mechanism for substrate recruitment (Barbash, Lee & Diehl 2011; van Drogen et al. 2006; Welcker et al. 2003; Wertz et al. 2011). In addition, some accessory proteins have been identified as mediators of substrate

recognition. One example is SKP2, a F-box protein that utilizes CKS1 to form contacts with a phosphodegron in p27 (Ganoth et al. 2001; Hao et al. 2005). F-box proteins also recognize glycosylated substrates for ubiquitylation through F-box-associated domains (Glenn et al. 2008). As cell cycle regulators, F-box proteins have a clear role in cancer development and tumorigenesis. However, recent studies have expanded the scope of the SCF complex beyond cell proliferation. Control of circadian rhythms, Parkinson's disease and viral infections are amongst the cellular process in which F-box proteins are involved (Skaar, Pagan & Pagano 2013).

A number of CRLs function as dimers. The prevalent dimerization mechanism observed in CRLs is mediated by the substrate-interacting subunit of the complex. Dimerization through the substrate adaptor protein (F-box protein in SCFs) utilizes specific domains to bring two CRL complexes together. Many F-box proteins form dimers, including Fbw7, Pop1, Pop2, Cdc4, Met30, Skp2, among others. In most cases, these dimers are mediated by a conserved domain located N-terminal to the F-box motif (Hao et al. 2007; Tang et al. 2007). Dimerization in general does not affect the affinity for the substrate but it enhances ubiquitin chain initiation and elongation, increasing the processivity of the complex (Bosu & Kipreos 2008).

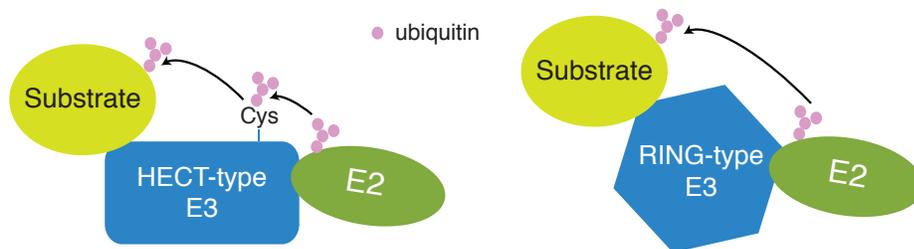


Figure 1.4. HECT- vs. RING-type E3 ubiquitin ligases

Schematic representation of the mechanisms for ubiquitin transfer by HECT- and RING-type E3 ubiquitin ligases. The HECT domain forms a thioester intermediate during the transfer of ubiquitin from the E2 to the substrate, while the RING domain serves as a scaffold for the transfer but does not directly bind ubiquitin.

1.1.1.2.2 HECT type

Distinct from RING type ligases, Homologous to the E6AP C Terminus (HECT) type ubiquitin ligases catalyze the transfer of ubiquitin to the substrate as intermediaries in a two-step reaction (Scheffner, Nuber & Huibregtse 1995). First, ubiquitin is transferred from the E2 active cysteine to a conserved cysteine residue in the HECT domain in a transthioesterification reaction. Then, the target lysine in the substrate attacks the thioester HECT:Ub link to form the isopeptide bond (**Figure 1.4**) (Huibregtse et al. 1995; Kamadurai et al. 2013). The human genome encodes 28 HECT-type E3 ubiquitin ligases varying from 80 to more than 500 kDa in size. The conserved HECT domain, located in the C-terminus of the protein, can be divided structurally into N-terminal and C-terminal lobes (**Figure 1.5**).

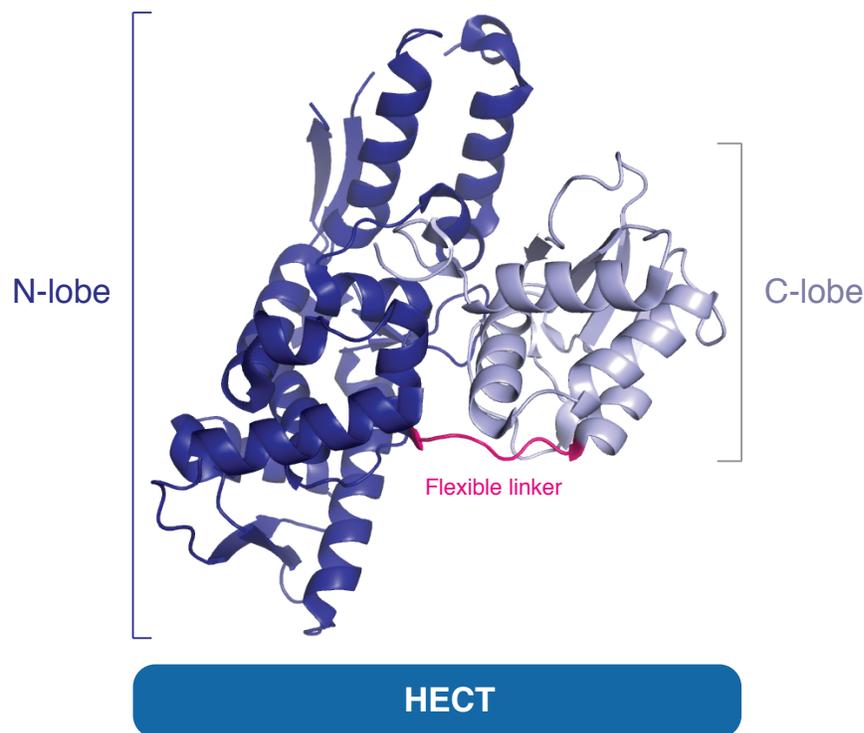


Figure 1.5. Structure of the HECT domain.

Crystal structure of the HECT domain from HUWE1 (3993-4374). N-terminal lobe (dark blue) and C-terminal lobe (light blue) are connected through a flexible linker (pink). PDB: 5LP8

The N-lobe directly interacts with the conjugated E2 enzyme (mainly UbcH7 and UbcH5 family members), while the C-lobe harbours the catalytic cysteine residue. Both lobes are linked through a flexible region that facilitates the relative positioning of E2:Ub, substrate and catalytic cysteine (Verdecia et al. 2003). The conjugation status of the E2 is a determinant factor in HECT-type ubiquitylation activity. The distance between the bound E2 enzyme and the active cysteine depends on the conjugation state of the bound E2 enzyme. Crystal structures of the HECT:E2 complexes for E6AP and WWP1 show a suboptimal distance between the two active cores that would preclude transthioesterification (Huang et al. 1999; Verdecia et al. 2003). In the NEDD4L:UbcH5:Ub complex structure (Kamadurai et al. 2009), the loaded E2 enzyme is positioned closer to the catalytic cysteine within a distance that promotes conjugation. In contrast to RING-type, the ability of HECT E3s to form thioester bonds permits chain elongation on the E3 (Scheffner & Kumar 2014). Given the flexibility of the HECT domain, conformational changes prompted by different E2:Ub topologies are expected to coordinate chain topology and elongation providing an additional regulatory mechanism.

1.1.1.2.2.1 Regulation of HECT-type ligases

Regulation of HECT-type E3s depends on the nature of their association with substrates. Different domains located in the N-terminal extensions of the E3s mediate substrate recruitment. Based on the presence of specific substrate-binding domains, HECT-type ligases are classified in three subfamilies: the NEDD4-like, which contains WW repeats; HERC, characterized by the presence of RLD domains and a final group of HECT ligases that contain different N-terminal domains (Scheffner & Kumar 2014).

Beyond direct interaction with its substrates, some HECT ligases utilize posttranslational modifications and auxiliary proteins to regulate ubiquitin transfer. One example is the hormone-dependent inhibition of NEDD4L. Insulin and aldosterone induce phosphorylation of NEDD4L, which leads to the recruitment of the adaptor protein 14-3-3. Binding of 14-3-3 prevents the interaction of the ligase with its cognate substrate ENaC, a sodium channel subunit, thus regulating sodium transport (Bhalla et al. 2005). ITCH, also a NEDD4-like ligase, is autoinhibited by an intramolecular interaction between its WW domains and the HECT domain. Ndfip1, an adaptor protein, binds to these WW repeats and releases the HECT domain prompting

ubiquitin transfer (Riling et al. 2015). A related mechanism was also reported for LRAD3, an activator for the ITCH ligase (Noyes et al. 2016). Similarly, WWP2 was shown to be autoinhibited by a linker segment located in between WW repeats. Phosphorylation of specific tyrosine residues within the linker reduces autoinhibition of the ligase while complete loss of the segment leads to hyperactivation and self-destruction by autoubiquitylation *in vivo* (Chen, Z et al. 2017). In addition, certain members of the α -arrestin family of proteins bind to the WW domains in NEDD4 ligases assisting in the ubiquitylation of substrates (Han, Kommaddi & Shenoy 2013; Shea et al. 2012).

Conformational changes and dimerization are important mechanisms for the control of ubiquitylation in HECT ligases. Also in the NEDD4 family, regulation of ubiquitin transfer can depend on oligomerization. Intramolecular interactions induced by ubiquitylation were observed to promote trimerization and subsequent ligase inactivation from yeast to humans (Attali et al. 2017). In HUWE1, inter- and intramolecular interactions induce conformational changes that regulate the activity of the E3. Structural elements located N-terminal to the HECT domain mediate dimerization of HUWE1. This self-association immobilizes the C-lobe and occludes its C-terminal region, which contains a phenylalanine residue with catalytic activity (Sander et al. 2017). In SMURF2, the C2 domain interacts with the HECT domain rendering the ligase inactive by impairing its ability to form thioester bonds and non-covalent interactions with ubiquitin (Mari et al. 2014). Binding of SMAD7, an auxiliary protein, enhances E2 binding to the N-lobe and disrupts the interaction between the C2 and the HECT domains releasing the autoinhibition (Wiesner et al. 2007).

Reflecting their diverse cellular functions, HECT E3 ligases have the ability to synthesize different types of ubiquitin chains (Kim, HC & Huibregtse 2009). For instance, E6AP synthesizes K48-linked ubiquitin chains while the NEDD4 family forms both K63 and K48-linked ubiquitin (Kim, HC & Huibregtse 2009). One of the unresolved questions is how the assembly of the ubiquitin chain progresses in the HECT enzyme. Two mechanisms have been proposed: assembly on the substrate by sequential attachment of single ubiquitin molecules or, polymerization on the catalytic cysteine in the C-lobe followed by a one-step transfer of the polymer to the substrate. Recent studies point at the prevalence of the first mechanism, where the

ability to build different chain topologies relies on the C-lobe. Moreover, some HECT ligases such as NEDD4, Rsp5 and SMURF2 display a non-covalent ubiquitin binding site in the N-lobe involved in promoting processivity in the chain transfer (Kim, HC et al. 2011; Maspero et al. 2011). It remains to be seen if this non-covalent ubiquitin binding site is also present outside the NEDD4 family.

1.1.1.2.3 RBR type

RING-between-RING (Wertz et al.) type enzymes were described in 2011 as the third type of E3 ubiquitin ligases when their ability to form thioester bonds was first observed (Wenzel, Dawn M. et al. 2011). RBR ligases display both RING- and HECT-like properties. They possess a RING domain while an active cysteine residue mediates the two-step transfer of ubiquitin from the E2 to the substrate (Morett & Bork 1999; van der Reijden et al. 1999). RBRs consist of two RING domains (RING1 and RING2) separated by an in-between-RING (IBR) domain. The RING2 domain has a distinct zinc-binding pattern that does not classify it as a true RING domain, however its RING name persists. Humans encode 12 RBR E3 ligases with similar domain architectures. RING1 binds the E2:Ub conjugate while the second RING2 domain (also called Rcat, required-for-catalysis) harbours the catalytic cysteine. The IBR domain (also called BRcat, benign-catalytic) has a similar fold to the RING2; however, it lacks catalytic cysteine residues (Spratt, Walden & Shaw 2014).

As in the case of the HECT and RING-types, RBR ligases possess protein-protein interaction domains that regulate substrate recruitment and catalytic activity. Remarkably, RBR enzymes appear to share a common mechanism of regulation in which the domains surrounding the conserved RBR domains inhibit catalytic activity (Duda et al. 2013; Riley et al. 2013; Smit et al. 2012; Trempe et al. 2013; Wauer & Komander 2013). In particular, Parkin is autoinhibited by its UBL domain, which binds to a region in between the IBR and RING2 domains. Several substrates of Parkin have been shown to bind the UBL suggesting that substrate recruitment contributes to the release of the inhibitory effect (Fallon et al. 2006; Tsai et al. 2003). The region between IBR and RING2 is also thought to prevent E2 binding to the RING1 further highlighting the need of activation (Riley et al. 2013; Trempe et al. 2013; Wauer & Komander 2013).

RBR proteins are also involved in important cellular processes such as transcription and translation, regulation of posttranslational modifications, protein degradation, cell cycle regulation and microbial infection. RBR enzymes are especially attractive targets for research and drug development given their involvement in maladies such as Parkinsonism, dementia and Alzheimer's disease (Eisenhaber et al. 2007).

1.2 The N-End Rule Pathway

After the discovery of ubiquitin and the cascade of reactions that control its conjugation, one of the key unresolved questions was how E3 ligases recognized substrates for degradation (Hershko et al. 1983). In 1986, the group of Alexander Varshavsky observed that β -galactosidase in *S. cerevisiae* had a remarkably different half-life in the cell depending on the identity of its N-terminal residue (Bachmair, Finley & Varshavsky 1986). Half-lives for X- β -galactosidase (X representing any one of 20 natural amino acids, except proline) ranged from 2 min to more than 20 h. These experiments shaped what is known as the "N-end rule", which relates the half-life of a protein in the cell with the identity of its N-terminal residue.

The last three decades of studies on the mammalian N-end rule have led to the elucidation of its fundamental role on homeostasis, with remarkably broad physiological functions that continue to be discovered. Cardiovascular development, neurodegeneration, gluconeogenesis, neural tube formation, spermatogenesis, apoptosis, oxygen, heme and NO sensing, and chromosomal stability are some of the processes that rely on the timely degradation of proteins based on their N-terminal residue (Varshavsky 2011). An even greater number of undiscovered substrates are likely to exist. One of the biggest challenges of protein degradation studies has been the insufficient knowledge of non-processive proteolytic events, as the conditional regulation of protein cleavage often leads to the generation of C-terminal fragments that are N-end rule substrates.

1.2.1 N-degrons

Destabilizing N-terminal residues are termed N-degrons, which identify the N-terminal degradation signals of proteins associated with the N-end rule pathway (Varshavsky 1991). In both, prokaryotes and eukaryotes, timely recognition of degradation signals is the major determinant of protein turnover. Surprisingly, prokaryotes contain a proteolytic system that resembles the recognition signals of the N-end rule pathway but without the conjugation of ubiquitin to the substrate (Dougan, D. A., Micevski & Truscott 2012). Bacteria and Archaea utilize ATP-dependent proteases such as ClpAP to degrade target proteins bearing specific N-degrons. In the N-end rule the recognition modules that bind N-degrons are termed N-recognins (Varshavsky 1996). In bacteria the ClpS protein, an N-recognin, acts as an adaptor for ClpAP by recognizing destabilizing N-terminal residues on target proteins and anchoring them to the protease to initiate their degradation (Dougan, David A. et al. 2002; Erbse et al. 2006). In eukaryotes, N-recognins are E3 ubiquitin ligases that directly recognize N-degrons in substrates and target them for degradation by the 26S proteasome.

For more than 20 years, the eukaryotic N-end rule classified only 13 natural amino acids as destabilizing. In 2010, a second branch of the N-end rule was discovered, termed the Ac/N-end rule pathway (Hwang, C-S, Shemorry & Varshavsky 2010). In this branch, acetylation of specific N-terminal residues is required for recognition. Moreover, a study published this year revealed the destabilizing nature of proline as an N-terminal residue in the context of gluconeogenesis (Chen, SJ et al. 2017). These findings described the third branch of the N-end rule pathway, the Pro/N-end rule, and catalogued all 20 natural amino acids as potentially destabilizing. Cellular compartmentalization of the N-recognin and its substrates as well as posttranslational modifications of the N-degrons determine the fate of the proteins in the N-end rule pathway.

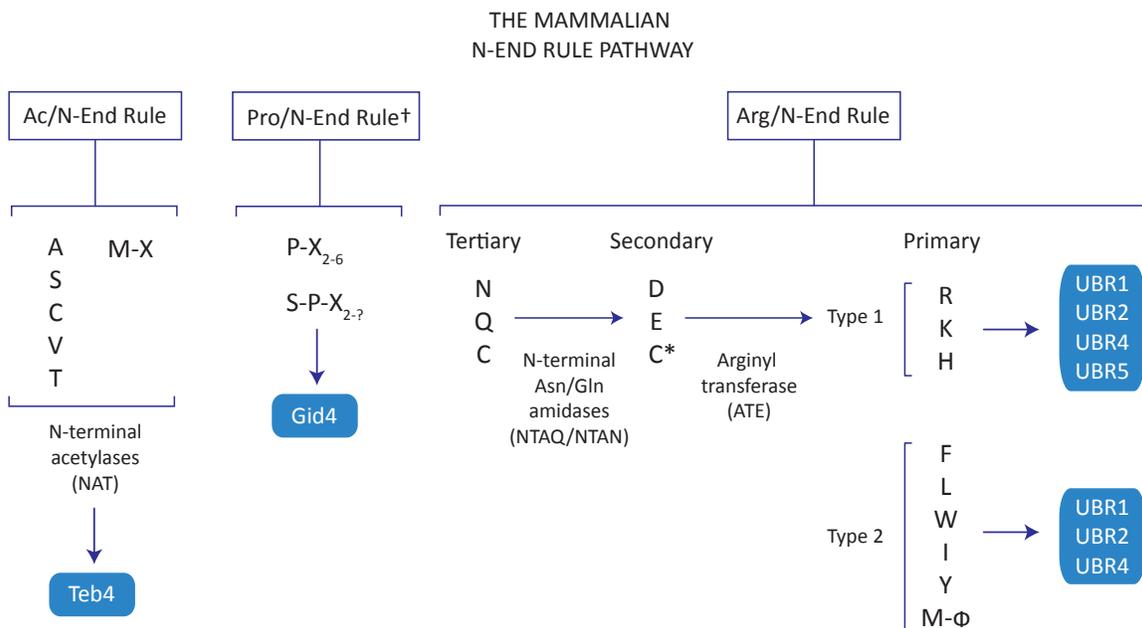


Figure 1.6. The mammalian N-end rule pathway.

Classification of the 20 natural amino acids as destabilizing N-terminal residues in the mammalian N-end rule pathway. X represents any of 20 natural amino acids, ϕ represents a hydrophobic residue. The corresponding N-recognins are indicated in blue boxes. † The Pro/N-end rule has been only described in *S. cerevisiae*. However, there is a highly conserved homolog of Gid4 suggesting the presence of the pathway in mammals as well. * Oxidized cysteine.

1.2.1.1 The Ac/N-end rule pathway

Acetylation is the second most abundant protein modification in eukaryotic cells (Khoury, Baliban & Floudas 2011). In eukaryotes, the N-terminal α -amino group of methionine is often acetylated by N-terminal acetyltransferases (Nt-acetylases) that cotranslationally modify the nascent protein (Gautschi et al. 2003). If the residue at the second position has a small side chain (Ala, Val, Ser, Thr, Cys, Gly and Pro), the N-terminal methionine is frequently cleaved off (Frottin et al. 2006). Moreover, removal of methionine often leads to acetylation of the resulting N-terminal α -amino group with N-terminal acetylation observed both, cotranslationally and posttranslationally *in vivo* (Helsens et al. 2011). No N-terminal deacetylases have been identified so it is thought that N-terminal acetylation is an irreversible modification (Starheim, Gevaert &

Arnesen 2012). Recently, the first lethal genetic disorder caused by a mutation in an N-terminal acetylase was reported (Rope et al. 2011). In the Ogden syndrome impairment of N-terminal acetylation causes developmental delays and cardiac arrhythmias.

The exact function of N-terminal acetylation has remained mostly elusive since its discovery in 1958 (Narita 1958). Recent studies have shed light into particular roles in individual proteins. Translocation to the endoplasmic reticulum, enhancement of protein complexes during neddylation and regulation of lysosomal localization are amongst the functions discovered to date (Behnia et al. 2004; Oh, Hyun & Varshavsky 2017; Scott et al. 2011; Setty et al. 2004). Since early studies on the proteasome, N-terminal acetylation was considered a metabolic stabilizer by way of blocking the degradation of otherwise metabolically unstable proteins (Hershko et al. 1984; Varshavsky 2011). However, the discovery of the Ac/N-end rule revealed the destabilizing nature of the modification and assigned it a regulatory role in protein turnover (**Figure 1.6**).

In the Ac/N-end rule pathway substrates are recognized when the α -amino group of N-terminal Ala, Val, Ser, Thr, Cys and Met residues is acetylated during translation (Hwang, C-S, Shemorry & Varshavsky 2010). This group of degradation signals is termed ^{Ac}N-degrons. Removal of N-terminal methionine by aminopeptidases (MetAPs) is conserved from bacteria to higher eukaryotes and occurs in ~70% of proteins (Giglione, Boularot & Meinnel 2004). MetAPs expose the degradation signal in the second residue while N-terminal acetylases generate the ^{Ac}N-degron on the nascent protein. Although essential and highly conserved throughout organisms, these groups of enzymes had an elusive physiological role before their identification as major components of the ubiquitin proteasome system through the N-end rule pathway (Lee, KE et al. 2016; Oh, Hyun & Varshavsky 2017). The Ac/N-end rule pathway was first discovered in yeast. N-terminal acetylation induced proteasomal degradation by the endoplasmic reticulum E3 ligase DOA10 and the cytosolic and nuclear E3 called Not4 (Hwang, C-S, Shemorry & Varshavsky 2010; Khmelinskii & Knop 2014; Kim, I et al. 2013; Shemorry, Hwang & Varshavsky 2013). Later the existence of the Ac/N-end rule in mammals was confirmed by studies on the E3 ligase Teb4 and its substrate Rgs2 (Park et al. 2015). Notably, hypertension-

associated mutations on Rgs2 in humans had an increased degradation rate by the Ac/N-end rule compared to wild type.

In the context of ^{Ac}N-degrons and given the prevalence of cotranslational N-terminal acetylation, what is the ultimate purpose of such massive production of short-lived proteins for the cell? The answer would be centered on quality control and regulation of stoichiometry. Varshavsky and colleagues propose a model based on the conditionality of ^{Ac}N-degrons, where the nascent degradation signal will only be recognized by a cognate E3 if the folding of the N-terminus is delayed, if protective interactions with associated proteins do not occur or if mutations perturb the correct conformation of the nascent protein (Varshavsky 2011). An example would be the stoichiometric control of targets involved in multi-protein complexes such as histones and ribosomal factors. When unassembled, these proteins are short-lived in the cells, due to their degradation through the N-end rule pathway.

As studies on the Ac/N-end rule continue to emerge, the scope of N-terminal recognition as a way of targeting proteins for degradation further expands. As nearly 90% of human proteins are N-terminally acetylated, it is only a matter of time before more roles of this pathway are discovered (Van Damme et al. 2012).

1.2.1.2 Pro/N-end rule pathway

Degradation of gluconeogenic enzymes is dependent on the action of the GID E3 ligase complex (Hammerle et al. 1998). The role of proline as an N-degron in this system was explored by Varshavsky and colleagues combining a modified two-hybrid binding assay and a novel method called promoter reference technique, which allows for gene-specific control of translation. These studies showed that fructose-1,6-bisphosphatase, isocitrate lyase and malate dehydrogenase were targeted for degradation through the recognition of their N-terminal proline residue. Phosphoenolpyruvate carboxykinase, which harbors a proline in the second position, was also shown to be targeted for degradation through recognition of its second residue (Chen, SJ et al. 2017). In addition, Gid4, a subunit of the GID E3 ubiquitin ligase complex in yeast was identified as the recognition component, or N-recogin, for the Pro/N-end rule pathway (**Figure 1.6**).

Discovery of the third branch of the N-end rule points at a highly specific role of N-terminal proline in the gluconeogenic pathway. Proline in the second position is a permissive residue for MetAPs, which would often lead to exposure of N-terminal proline in a variety of cytosolic proteins. Future studies will determine if other proteins not associated with gluconeogenesis are indeed targets of the Pro/N-end rule or are recognized by the GID ligase (Chen, SJ et al. 2017).

Even though the Pro/N-end rule has only been demonstrated in yeast, subunits of the GID E3 ubiquitin ligase complex are conserved in higher organisms, suggesting it might be conserved in higher eukaryotes. The physiological role and implications of a Pro/N-end rule in multicellular organisms remain to be revealed.

1.2.1.3 The Arg/N-end rule pathway

After the discovery of the Ac/N-end rule, the previously identified N-degrons were grouped into the Arg/N-end rule branch. The N-end rule pathway relies on the action of specific proteases that, upon the right stimuli, cleave the target proteins to expose specific N-degrons. The first protease that has the potential to act is the MetAP. During protein synthesis, MetAPs often remove the N-terminal translation initiator methionine revealing a new N-terminal residue or N-degron. However, MetAPs are restricted to act only if the second residue bears a small side chain. Endoproteases, on the other hand, display a variety of specificities towards binding and cleavage sites, exposing new N-degrons previously suppressed in internal sequences. The latter mechanism dominates the Arg/N-end rule, where proteases including caspases, calpains, secretases and separases, among others, create metabolically unstable fragments targeted by N-recognins (Brower, Piatkov & Varshavsky 2013; Ditzel et al. 2003; Piatkov, Brower & Varshavsky 2012; Rao et al. 2001).

The Arg/N-end rule classifies N-degrons on a hierarchical structure based on the number of posttranslational modifications required for recognition. Primary N-degrons directly bind the N-recognin while tertiary residues require two posttranslational modifications to prompt direct interaction. Glutamine, asparagine and cysteine are tertiary N-degrons (Gonda et al. 1989). N-terminal Asn/Gln amidases (NTAQ/NTAN) remove the amide group from the side chain

generating the secondary N-degrons aspartate and glutamate. In the case of N-terminal cysteine, specifically in plants and animals but not in yeast, oxidation by S-nitrosylation produces cyst-sulfenic, also a secondary N-degron (Hu, RG et al. 2005). All three, oxidised cysteine, aspartate and glutamate are targets of arginylation by arginyl-tRNA-protein transferases (ATE) (Kwon et al. 2002; Lee, MJ et al. 2005; Varshavsky 2011). Addition of the primary amino acid arginine to the N-terminus of secondary N-degrons allows direct recognition by specific N-recognins. Primary N-degrons are classified in two groups based on their chemical nature: type 1, which are the basic residues arginine, lysine and histidine, and type 2 that are the bulky hydrophobic amino acids phenylalanine, leucine, tryptophan, isoleucine and tyrosine (**Figure 1.6**).

1.2.1.3.1 The UBR family of ubiquitin ligases

The first identified N-recognin was Ubr1p in *Saccharomyces cerevisiae* (Bartel, Wunning & Varshavsky 1990). *UBR1* encodes a ~225 kDa E3 ubiquitin ligase with no significant sequence similarity to other yeast proteins while in mammals there are at least seven other similar proteins. The mammalian N-end rule ligases UBR1 and UBR2 were characterized by the groups of Alexander Varshavsky and H.Q. Han (Kwak et al. 2004; Kwon et al. 1998; Kwon et al. 2003; Kwon et al. 2001). Both proteins are 47% identical and have overlapping functionality with similar capabilities to bind N-degrons. Double *UBR1*^{-/-} and *UBR2*^{-/-} mouse strains die as early embryos in contrast with surviving *UBR1*^{-/-} or *UBR2*^{-/-} mutants.

Experiments involving fibroblasts derived from these *UBR1*^{-/-} and *UBR2*^{-/-} embryos showed that despite the lack of functional UBR1 and UBR2, the cells retained significant N-end rule pathway activity. This prompted the search for other mammalian E3s with the ability to recognize N-degrons. Affinity assays, iRNA experiments and proteomics identified UBR4 and UBR5 as N-recognins. Bioinformatics analysis showed the presence of a common motif among UBR1, UBR2, UBR4 and UBR5, a ~70-residue putative zinc-finger that authors named the ubiquitin recognin box (UBR-box) (Tasaki et al. 2005). Sequence alignments and phylogenetic analysis revealed seven UBR-containing proteins in mammals, numbered UBR1 through UBR7 (**Figure 1.7**). UBR proteins are heterogeneous in size and sequence but all contain domains that classify them as E3 ubiquitin ligases. These range from classical RING and HECT domains to multi-subunit adaptors such as the F-box. However, not all of the UBR-box proteins bind N-degrons.

Only UBR1, UBR2, UBR4 and UBR5 bind type 1 N-degrons: arginine, lysine and histidine (Tasaki & Kwon 2007; Tasaki et al. 2009).

1.2.1.3.1.1 The UBR-box domain

Pull down assays with recombinant proteins narrowed down the minimal sequence necessary to interact with the N-degrons (Tasaki et al. 2009). The UBR-box became the focus of structural studies aiming to explain the molecular determinants of binding to rationalize the specificity of the pathway. In 2010, the groups of Kalle Gehring and Hyun Kyu Song obtained the crystal structures of the bound and unbound UBR-box from human and yeast, respectively (Choi et al. 2010; Matta-Camacho et al. 2010). These structures confirmed the zinc-binding capabilities of the domain and revealed a new protein fold.

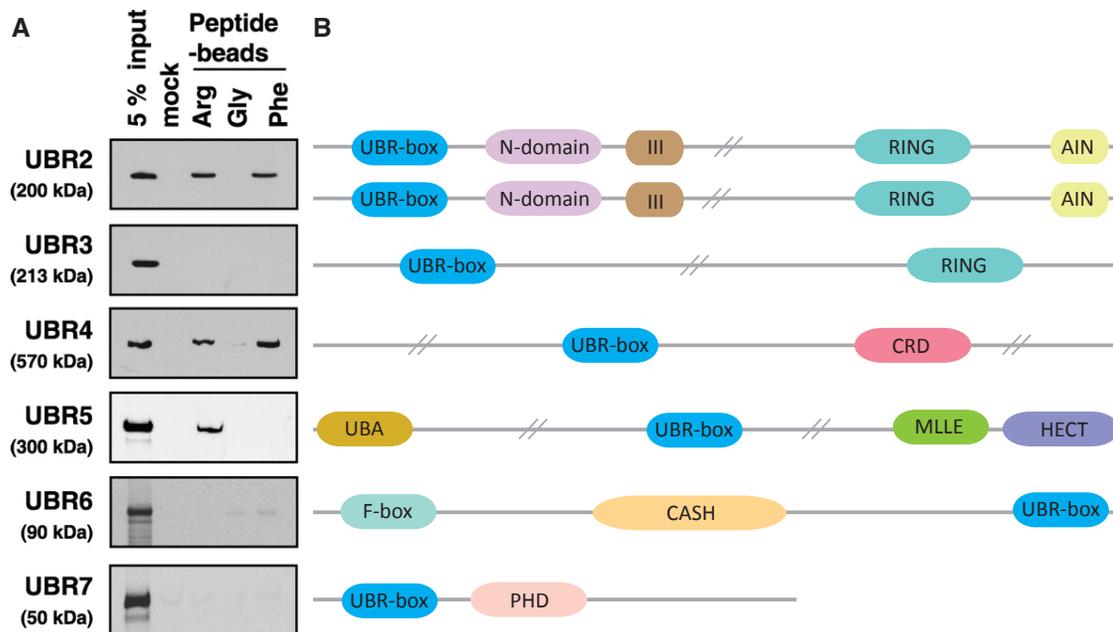


Figure 1.7. N-degron specificity and domain architecture in the UBR family of ubiquitin ligases.

A. Pull-down assay with overexpressed full-length UBR proteins: UBR2 and UBR3 in *S. cerevisiae*, UBR4, UBR5 in COS7 cells, and UBR6 and UBR7 in wheat germ lysate. X-peptides were used as bait. X represents arginine (type 1), glycine or phenylalanine (type 2) as N-terminal residues. UBR1 is not shown as it shares same specificity as UBR2. **B.** Domain architecture in the UBR family of proteins. Modified from Tasaki *et al* 2009.

The human UBR-box domains from UBR1 and UBR2 share a ~76% sequence similarity with nearly identical folds. The tertiary structure of the UBR-box is stabilized by three zinc ions that form two adjacent zinc fingers each coordinated in a tetrahedral arrangement. The first zinc finger has a novel coordination topology with one cysteine residue (Cys127) simultaneously coordinating two zinc atoms. The second zinc finger has the typical Cys₂-His₂ geometry with one zinc ion as the coordination element. These zinc-binding sites frame the N-degron binding pocket (see Chapter Two).

1.2.1.3.1.2 Type 1 binding site

The structures from human and yeast UBR-box with bound peptides revealed a conserved mechanism of type 1 N-degron binding. The peptides interact with two regions: a negatively charged pocket, which accommodates the N-terminal basic residue and, a hydrophobic pocket, which binds the amino acid in the second position. Isothermal titration calorimetry experiments revealed the specificity for N-degron binding. Peptides with an N-terminal arginine exhibited the highest affinity for the human UBR-box domain followed by peptides with lysine and histidine (see Chapter Two). In yeast, the UBR-box showed higher affinity for histidine rather lysine N-terminal peptides. The difference in specificity is due to subtle variations in the fold and interacting residues in the yeast UBR-box compared to human.

These studies also showed that free arginine interacts with the UBR-box but considerably weakly compared to dipeptides (Choi et al. 2010; Matta-Camacho et al. 2010). These observations agreed with early studies that demonstrated that dipeptides inhibit the N-end rule pathway in both *S. cerevisiae* and mammalian reticulocytes (Baker & Varshavsky 1991; Reiss, Kaim & Hershko 1988). Structural studies validated the importance of the second residue and explored the dependency of subsequent amino acids in the N-terminus of the substrate. Human UBR-box structures in complex with various arginine and histidine peptides show a conserved binding site for the second residue, while the third and fourth amino acids do not interact with the domain (Matta-Camacho et al. 2010) (see Chapter Two). In yeast, in addition to the conserved dependency for the second residue, the polypeptide backbone from the third position is also involved in hydrogen bond interactions with the UBR-box (Choi et al. 2010). These studies showed that the first two N-degron residues are the main determinants for binding; however, it

remains to be addressed if, in the context of full-length proteins, other regions further stabilize binding. In the initial structural study on the human UBR-box (Matta-Camacho et al. 2010), a preference for acidic residues in the second position was observed, while the yeast counterpart had a multi-specific pocket that accommodated hydrophobic, acidic and basic residues. Recent studies in the human UBR-box confirm the preference for hydrophobic residues in the second position (see Chapter Two).

1.2.1.3.1.2 Type 2 recognition

The structure and molecular determinants for type 2 N-degron binding in the N-end rule remain elusive. Bulky hydrophobic N-terminal residues constitute type 2 N-degrons. Phenylalanine, isoleucine, tryptophan, tyrosine and leucine are recognized by UBR1 (Ubr1p in yeast), UBR2 and UBR4 through a secondary binding site independent of type 1 binding. An ~80-residue fragment adjacent to the UBR-box in UBR1, named the N-domain (**Figure 1.7**), showed sequence and functional similarity to the ClpS adaptor, the bacterial N-recognin (Erbse et al. 2006; Tasaki et al. 2009). This region was also found in UBR4 although with significantly lower sequence similarity.

In the prokaryotic N-end rule, the ClpAP complex of *E. coli* governs N-degron binding (Wang, KH et al. 2008). Studies of the ClpS adaptor showed that a hydrophobic pocket recognizes the N-terminal α -amino group and the side chain of the N-degron, while also establishing contacts with the second residue in the substrate. Advances in the understanding of type 2 interactions and the structure of the N-domain will provide a platform for the design of inhibitors that not only will transform the research on protein degradation but also will have potential pharmacological applications.

Simultaneously with the discovery of the N-end rule in the 1980's, the design of competitive inhibitors demonstrated the potential of modulating protein degradation *in vitro* and *in vivo* (Kwon, Levy & Varshavsky 1999; Reiss, Kaim & Hershko 1988). Small molecular weight compounds have been shown to have an inhibitory effect on the N-end rule. The first example is the amino acid derivative leucine methyl ester, which was reported to inhibit type 2 binding *in vivo* using the mimic substrate Tyr- β -galactosidase (Baker & Varshavsky 1991). Recently, the

group of Min Jae Lee identified Phe-derived monomeric molecules, such as *p*-chloroamphetamine, as significant inhibitors of degradation for both, type 1 and type 2 substrates. The unexpected inhibitory effect on type 1 residues is not well understood. One possibility is that binding to the N-domain impacts the conformational freedom of the protein or the access of substrates to the UBR-box, which would have an indirect effect on type 1 recognition. Given its hydrophobic character, it seems unlikely that *p*-chloroamphetamine elicits its inhibitory effect by binding the UBR-box domain directly.

Recent efforts to develop more potent, less toxic and more effective inhibitors have utilized the proximity of the UBR-box domain and the N-domain in the primary sequence to design small molecules that target both sites simultaneously (Lee, JH et al. 2015). Even though these inhibitors, called heterovalent, have a higher efficacy compared to the classical dipeptides, further improvement is needed to optimize their binding. Heterovalent inhibitors target the UBR-box and the N-domain simultaneously by harbouring type 1- and type 2-like residues in the N-terminal and C-terminal ends of the molecule. Both residues are connected through a variable length linker. This geometry is expected to potentiate binding by maximizing the effective concentration of the ligands. One important feature is the length of this linker, which is optimized when it is the most similar to the spatial distance between the two binding sites. The group of Min Jae Lee optimized the linker length using *in silico* docking computation studies and *in vitro* assays (Jiang, YXL et al. 2013); however, it is suggested that these inhibitors could be further optimized with more detailed structural information on the N-recognins.

1.2.1.3.1.3.1 N-terminal methionine as a degradation signal

Recently yet another discovery in the N-end rule expanded the scope of substrates that can be targeted by the pathway. Kim *et al* demonstrated that yeast Ubr1p and mouse UBR1 and UBR2 recognize unacetylated N-terminal methionine as a N-degron if the second position is a hydrophobic residue (Met- ϕ) such as leucine, valine, alanine, phenylalanine, tyrosine, tryptophan or isoleucine (Kim, HK et al. 2014). Binding of the Met- ϕ N-degrons is thought to occur through the type 2 site known as N-domain (**Figure 1.7**). Remarkably, addition of dipeptides bearing arginine in the N-terminus enhanced polyubiquitylation of the substrate. This allosteric effect has been observed with other type 2 substrates (Varshavsky 2011). Among the substrates targeted

through this mechanism are misfolded proteins in the cytoplasm as well as transcription activator Msn4, hydroxyaspartate dehydratase Sry1, the Golgi-associated GTPase Arl3 and the subunit of the proteasome Pre5. This study suggests two modes of degradation for Met- ϕ proteins: 1) Unacetylated N-terminal methionine of a Met- ϕ protein can act as N-degron for the Arg/N-rule. 2) Acetylation of the N-terminal methionine of the Met- ϕ protein converts it into an ^{Ac}N-degron targeted by the Ac/N-end rule. One of the main conclusions of these findings is the plasticity of N-degrons.

1.2.1.3.1.4 UBR1

Pioneering work on the N-end rule and the UPS predicted the presence of recognition elements that would specifically bind N-terminal destabilizing residues in target substrates (Bachmair, Finley & Varshavsky 1986; Gonda et al. 1989; Reiss, Kaim & Hershko 1988). Mutational analysis in yeast allowed the isolation of the *UBR1* gene, the first N-recognin to be identified for the N-end rule (Bartel, Wunning & Varshavsky 1990). Yeast *ubr1* Δ cells resulted in slow growth, impaired sporulation and deficient import of di- and tripeptides (Alagramam, Naider & Becker 1995). These observations prompted the discovery of the first physiological function of the N-end rule: control of peptide import through degradation of Cup9, a transcriptional repressor of the Ptr2 peptide transporter (Byrd, Turner & Varshavsky 1998). In the positive feedback mechanism, imported dipeptides bearing destabilizing N-terminal residues bind to the Ubr1p protein leading to an allosteric activation of Cup9 degradation, therefore releasing Ptr2 repression and increasing peptide intake. These studies also revealed a third binding site in Ubr1p for an internal degron of Cup9 possibly located C-terminal to the N-domain (Turner, Du & Varshavsky 2000; Xia, Turner, et al. 2008). In the activation mechanism, dipeptides bind to the UBR-box and N-domain of Ubr1p causing the dissociation of the C-terminal autoinhibitory domain from the N-terminal region, thus exposing the three binding sites. The Ubr1p interaction with Cup9 is strongly increased only if both, type 1 and type 2 sites are occupied. Remarkably, this autoinhibitory mechanism was also observed in mammalian UBR1 (Du et al. 2002). Several E3 ubiquitin ligases are regulated by autoinhibitory mechanisms and posttranslational modifications. Phosphorylation of Ubr1p on Ser300 and Tyr277 in yeast by the Yck1/Yck2 kinase is required for the accelerated degradation of Cup9. This regulatory mechanism might be

present in the mammalian counterparts (Hwang, CS & Varshavsky 2008); however, there is no clear evidence yet of phosphorylation in human UBR1.

The N-end rule pathway also elicits an important control in the degradation of misfolded or damaged proteins. Ubr1p was the first ligase shown to be involved in cotranslational degradation of nascent proteins. Over 50% of newly synthesized chains never reach their full size before their degradation by the UPS. As partial folding occurs during ribosomal translation, a protein might expose degradation signals that become shielded only upon completion of correct folding. This mechanism facilitates quality control on protein synthesis, destroying polypeptides that fail to fold correctly (Turner & Varshavsky 2000; Verma et al. 2013). Varshavsky and colleagues proposed a similar model for the recently discovered ^{Ac}N-degrons. Moreover, Ubr1p was also found to be essential for targeting misfolded proteins for proteasomal degradation in the endoplasmic reticulum and cytoplasm (Eisele & Wolf 2008; Stolz et al. 2013; Theodoraki et al. 2012). This function is also attributed in mammals to Parkin and CHIP ligases.

UBR1 is ubiquitously expressed in all tissues. Mouse embryos show high expression of UBR1 in the branchial arches, tail and limb buds, while adults show the highest expression levels in skeletal muscle and heart (Kwon et al. 1998). UBR1^{-/-} mice are viable and fertile but weigh significantly less than wild type mice (Kwon et al. 2001), which supports the importance of UBR1 in skeletal muscle maintenance. Despite overlapping binding capabilities, functions of UBR1 and UBR2 differ significantly specially in terms of tissue expression levels. In contrast, their simultaneous absence is fatal. UBR1^{-/-}UBR2^{-/-} embryos die at midgestation with defects in neurogenesis and cardiovascular development (An et al. 2006).

The human *UBR1* gene encodes for a 200 kDa protein that contains at least two zinc finger motifs (UBR-box and a RING-H2 domain), an N-domain, a putative internal degron recognition motif (III) and a C-terminal autoinhibitory domain (AIN) (**Figure 1.7**) (Tasaki et al. 2009; Xia, Webster, et al. 2008). Secondary structure predictions show abundant secondary structure between the UBR-box and RING-H2, however, no additional domains have been characterized yet. In 2005, Zenker *et al* identified several disease-associated mutations in the *UBR1* gene as the cause for the congenital disorder known as the Johanson Blizzard syndrome (JBS) (**Table 1.1**).

This autosomal recessive disorder is characterized by nasal wing hypoaplasia and exocrine pancreatic insufficiency (Johanson & Blizzard 1971). JBS includes several other features such as hearing impairment, aplasia of the scalp, dental defects, hypothyroidism, cognitive impairment, short stature and urogenital and anorectal malformations. The birth prevalence of the syndrome in Europe is estimated to be 1 in 250,000 individuals cataloguing it as a rare disease (Zenker et al. 2005). Characterization of *UBR1*^{-/-} mice showed exocrine pancreatic insufficiency, while the pancreas of patients with JBS showed no expression of *UBR1* and destructive pancreatitis. Given the partial functional redundancy of mammalian N-recognins (*UBR1* and *UBR2* in particular), the Arg/N-end rule is still present but at a lower level of activity in JBS patients (Kwon et al. 2001). The *UBR1* protein substrate whose impaired degradation would cause the disorder is not yet known. As JBS affects a variety of functions and processes, it is suggested that the disease does not rise from the impairment of degradation of one substrate but more from a combination of different proteins specific to particular tissues (Atik et al. 2015; Zenker et al. 2006).

Table 1.1. Mutations in the *UBR1* gene causing Johanson Blizzard syndrome.

	Nucleotide alteration	Predicted effect	Previous publications	
UBR-box	c.81+2dupT	r.spl.? p.?	Alkhoury et al. (2008)	
	c.81+5G>C	r.spl.? p.?		
	c.364G>C	p.(Val122Leu)	Hwang et al. (2011)	
	c.380G>T	p.(Cys127Phe)	Zenker et al. (2005); Hwang et al. (2011)	
	c.407A>G	p.(His136Arg)		
N-domain	c.477delT	p.(Gly160Alafs*5)b	Zenker et al. (2005)	
	c.497A>G	p.(His166Arg)		
	c.529-13G>A	<i>p.(Asn177Leufs*10)</i>	Godbole et al. (2013)	
	c.650T>G	p.(Leu217Arg)	Zenker et al. (2005)	
	c.660-2_660-1delAG	r.spl.? p.?		
	c.753_754delTG	p.(Cys251*)	Zenker et al. (2005)	
	c.857T>G	p.(Ile286Arg)	Liu et al. (2011)	
	c.950T>C	p.(Leu317P ro)		
	c.1094-13A>G	<i>p.(Val365Glufs*2)</i>	Zenker et al. (2005)	
	c.1094-12A>G	r.spl.? p.?		
III	c.1166_1177del12	p.(Ala389_Phe392del)	Zenker et al. (2005)	
	c.1280T>G	p.L427R		
	c.1440-1G>A	r.spl.? p.?	Atik et al. (2015) Al-Dosari et al. (2008)	
	c.1507C>T	p.(Arg503*)	Hwang et al. (2011)	
	c.1537C>T	p.(Gln513*)		
	c.1648C>T	p.(Gln550*)	Zenker et al. (2005)	
	c.1688C>A	p.(Ala563Asp)	Zenker et al. (2005)	
	c.1759C>T	p.(Gln587*)		
		c.1886C>G	p.(Ser629*)	Zenker et al. (2005); Schoner et al. (2012)
		c.1911+14C>G	<i>p.(Glu638Valfs*29)</i>	
		c.1979_1981delTTG	p.(Val660del)	Alkhoury et al. (2008)
		c.1993C>T	p.(Arg665*)	

RING	c.2034C>A	p.(Tyr678*)	Almashraki et al. (2011) Zenker et al. (2005); Hwang et al. (2011)
	c.2098T>C	p.(Ser700Pro)	
	c.2254+2T>C	r.spl.? p.?	
	c.2260C>T	p.(Arg754Cys)	Hwang et al. (2011) Zenker et al. (2005)
	c.2261G>A	p.(Arg754His)	
	c.2294_2296delAAG	p.(Glu766del)	
	c.2319dupT	p.(His774Serfs*6)	
	c.2379+1G>C	r.spl.? p.?	
	c.2380-1G>A	r.spl.? p.?	Atik et al. (2015) Zenker et al. (2005) Zenker et al. (2005)
	c.2432+5G>C	r.spl.? p.?	
	c.2546_2547insA	p.(Met849Ilefs*13)c	Elting et al. (2008) Hwang et al. (2011)
	c.2598delA	p.(Pro867Hisfs*12)d	
	c.2608G>T	p.(Glu870*)	Quaio et al. (2014)
	c.2839+5G>A	<i>p.(Arg947Aspfs*7)</i>	
	c.3304C>G	p.(Gln1102Glu)	
c.3328G>T	p.(Glu1110*)		
c.3682C>T	p.(Gln1228*)		
c.3694delC	p.(Leu1232Trpfs*17)		
c.3724A>G	p.(Arg1242Gly)		
c.3745dupA	p.(Arg1249Lysfs*4)		
c.3835G>A	p.(Gly1279Ser)		
c.3998-1G>C	<i>p.(Glu1333_Gly1337del)</i>		
c.4093C>T	p.(Gln1365*)		
c.4188C>A	p.(Cys1396*)		
c.4193delT	p.(Leu1398Argfs*3)		
c.4277C>T	p.(Pro1426Leu)		
c.4280C>T	p.(Ser1427Phe)		
c.4291T>C	p.(Ser1431Pro)		
c.4524T>A	p.(Tyr1508*)		
c.4927G>T	p.(Glu1643*)	Zenker et al. (2005)	
c.4942delG	p.(Glu1648Lysfs*21)	Singh et al. (2014)	
c.4981G>A	p.(Gly1661Arg)		
c.5080G>T	p.(Glu1694*)		
c.5109-3A>G	<i>p.(Arg1704Glyfs*26)</i>		
c.5135_5144del10	p.(Arg1712Leufs*14)		
AIN			

r.spl.?: RNA was not analyzed but the change is expected to affect splicing, e.g. when the splice donor or splice acceptor site is changed. **p.:** protein. **fs*x:** frame shift with a new reading frame length x. Italic letters indicate that the effect of splicing mutations was demonstrated on mRNA level. Highlighted areas indicate mutations present in the UBR-box (blue), N-domain (purple), highly conserved area in UBR1 and UBR2 (brown), RING domain (cyan) and autoinhibitory domain (green). Modified from Sukalo *et al*, 2014.

As regulation of the N-end rule continues to be explored, emerging roles and regulatory mechanisms of the signature UBR1 N-recognin are discovered. One of the most interesting interplays is that of yeast Ubr1p with the HECT type ubiquitin ligase Ufd4, where cooperative interactions facilitate ubiquitylation. For Mgt1, a DNA repair enzyme that demethylates double stranded DNA from O⁶-methylguanine, both ligases were shown to interact and co-target an

internal degron in Mgt1 to induce its proteasomal degradation (Hwang, CS, Shemorry & Varshavsky 2009). The synergy between both ligases increased the yield and processivity of polyubiquitylation. Although Ufd4 is not essential for Ubr1p activity, *ufd4Δ* cells display impaired N-end rule activity. Moreover, Ufd4 is ~10 fold more abundant than Ubr1p in yeast, thus the Ubr-Ufd4 complex is expected to mediate degradation of Arg/N-end rule substrates (Hwang, C-S et al. 2010). It is important to note however, that Ufd4 is not an N-recognin, and it does not recognize N-degrons.

UBR1 has two cognate E2 enzymes: Ube2A/2B (Rad6 in yeast) and USE1. The E1 activating enzyme UBA1 mediates the activation of the vast majority of E2 enzymes and in UBR1 is paired with Ube2A/2B (Xie, Y & Varshavsky 1999). UBA6 is the second E1 enzyme in vertebrates and functions with the cognate E2 USE1. This pair interacts with both, UBR1 and UBR2 (Lee, PC et al. 2011). Notably, these two E2-E3 complexes work in parallel to promote degradation of the G proteins Rgs4/5 (Lee, MJ et al. 2005).

1.2.1.3.1.5 UBR2

In *Saccharomyces cerevisiae* Ubr2p is also an E3 ubiquitin ligase with an UBR-box domain but it does not function as a N-recognin (Wang, L et al. 2004). In mammals UBR2 is an N-recognin (**Figure 1.7**). Cloning and characterization of mouse UBR2 showed similarities in substrate binding properties as well as interactions with the same cognate E2 UBE2A/B (Kwak et al. 2004; Kwon et al. 2003). Interestingly, UBR2-lacking female mice died as embryos while males were born at normal frequency, which defined a gender-dependent role of UBR2 in development. Male mice exhibited infertility and testis degeneration by two months of age due to massive apoptosis in spermatocytes (Kwon et al. 2003). Expression profiling for both N-recognins showed high levels of UBR2 in testes primary in spermatocytes, while UBR1 was only prominent in spermatogonia, suggesting a role in meiosis. An *et al* demonstrated that UBR2 is recruited to chromatin during cell cycle progression to contribute to the induction of ubiquitylation of chromatin-associated proteins. In this process, UBR2 can also be allosterically activated by dipeptides, which promote accelerated UBE2A/B-mediated ubiquitylation of H2A and H2B histones (An et al. 2012; An et al. 2010). Another interesting role of UBR2 in spermatogenesis is its ability to metabolically stabilize Tex19.1 through binding. This interaction

is thought to be independent from the N-end rule, as ubiquitylation does not occur (Yang, F et al. 2010). A recent study on azoospermia (form of male infertility) found numerous rare non-silent variants of *UBR2* and other epigenetic regulators of spermatogenesis, suggesting an association of genetic defects in *UBR2* with infertility in humans (Li, Z et al. 2015).

UBR2 also controls DNA damage response assuring genome integrity upon genotoxic stress. Cells lacking *UBR2* are impaired in homologous recombination repair of double stranded DNA (Ouyang et al. 2006). Moreover, studies on the proteasomal degradation of Sml1 protein, which induces production of dNTPs necessary for DNA repair, showed that *UBR2* and *UBE2A/B* control Sml1 targeting for proteasomal turnover by inducing its ubiquitylation (Andreson et al. 2010).

The interplay between *UBR1* and *UBR2* is essential for degradation of the G-proteins Rgs4 and Rgs5. *UBR1* and *UBR2* function with both E1-E2 complexes *UBA6-USE1* and *UBA1-UBE2A/B*; however, *UBR2* specifically promotes K48 ubiquitin conjugation of Rgs4 (Lee, MJ et al. 2005; Lee, PC et al. 2011). Interactions between *UBR2* and its cognate E2s are mediated through its conserved RING domain. The presence of K48 ubiquitin linkages in the substrate is a specific signal for proteasomal degradation. Specificity towards particular ubiquitin linkages in *UBR1* has not been reported. As most of the substrates known for *UBR1* are targets of the proteasome, it is expected a preference for K48 linkages. Given the high sequence similarity between *UBR1* and *UBR2* in mammals, it is not surprising that their domain architecture is conserved. Reports on the presence of the autoinhibitory domain in mouse *UBR1* (Du et al. 2002) open the door for a similar mechanism in mammalian *UBR2* (**Figure 1.7**).

1.2.1.3.1.6 UBR3

Mouse *UBR3* was first characterized in 2007 as the closest E3 to *UBR1* and *UBR2* in the *UBR* family. Despite the presence of an *UBR*-box domain, *UBR3* does not bind N-degrons (**Figure 1.7**). Full length *UBR3* is a ~200 kDa protein with 22% similarity to *UBR1* and a conserved RING domain that mediates interactions with the cognate E2s *UBE2A/B*. *UBR3*^{-/-} Mouse strains had various phenotypes depending on their genetic background, where some exhibited neonatal lethality and suckling impairment while others died during embryogenesis. Adult female mice

had loss of sense of smell, and expression assays suggest a regulatory role of UBR3 in additional sensory pathways such as vision, touch, hearing and taste (Tasaki et al. 2007). Moreover, in developing *Drosophila* eye discs, loss of UBR3 leads to a delayed differentiation of photoreceptors, while in zebrafish causes a decrease in sonic hedgehog signalling in developing eyes and sensory neurons (Li, T, Fan, et al. 2016). As numerous studies pointed at the essential role of UBR3 in sensory pathways, the group of Hugo Bellen investigated the role of UBR3 in inherited deaf-blind disorders such as the Usher syndrome. Li *et al* found that UBR3 regulates monoubiquitylation of Myosin II, which modulates Myosin II:Myosin VIIa interactions required for normal auditory organ development. Mutations in these two proteins are often associated with inherited deafness in humans. *Drosophila* homologs of Pcdh15 and Sans, proteins involved in the Usher syndrome, were immunoprecipitated by the UBR-box domain of UBR3 (Li, T, Giagtzoglou, et al. 2016). However, the mechanistic details of how these proteins come together to regulate ear development are yet to be elucidated.

Similarly to UBR2, UBR3 regulates cellular levels of proteins involved in DNA repair. *In vitro* studies shown UBR3-dependent ubiquitylation of APE1, a transcription regulator implicated in genome stability (Meisenberg et al. 2012). UBR3 is also involved in hedgehog signalling by mediating ubiquitylation and degradation of *Drosophila* Cos2, a kinesin-related motor protein, and its mammalian homolog Kif7 (Li, T, Fan, et al. 2016). Numerous studies have implicated UBR proteins in different apoptosis mechanisms, either as promoters or negative effectors. UBR3 in particular, was shown to be a positive modulator of apoptosis by regulating the activity of caspase-cleaved DIAP1 during *Drosophila* development. Interestingly, this role is independent of its E3 activity as the RING domain is not required for DIAP1 stabilization. In contrast, the UBR-box domain is essential for the interaction with the caspase-generated fragment of DIAP1, which bears an N-terminal asparagine that is required for binding (Huang, Q et al. 2014).

Probably the most interesting role that has been discovered for UBR3 is the regulation of its ubiquitylation activity by pri peptides in *Drosophila*. In eukaryotes, many non-coding RNAs contain small open reading frames that can often produce peptides. This has been observed in organisms from yeast to plants and humans (Ingolia et al. 2014; Laressergues et al. 2015; Smith

et al. 2014). In *Drosophila*, the polished rice or tarsal-less (pri) RNA contains small ORFs that encode 11- to 32-residue peptides, which are required for development. Zanet *et al* showed that pri peptides control binding of UBR3 to the Shavenbaby (Svb) transcription repressor, and activate its proteasomal degradation (Zanet et al. 2015). It is suggested that the role of the pri peptides on UBR3 is to modify binding selectivity for substrates through a potential conformational change that would reveal the Svb binding site. The authors demonstrated that the isolated UBR-box domain binds Svb independently of pri peptides, proposing that the pri binding site is located outside of the UBR-box. This mechanism not only resembles that seen in UBR1 but also features a function of the UBR-box domain beyond the N-end rule. If direct interaction between the UBR-box and Svb occurs in purified proteins, this would be additional evidence of the substrate-binding capabilities of the UBR-box domain outside the N-end rule pathway. In addition, binding of other UBR3 substrates such as APE1 and DIAP1 was independent of pri peptides, which highlights important differences in the binding mechanisms that modulate UBR3 activity.

1.2.1.3.1.7 UBR4

Also known as p600, UBR4 is a 600 kDa protein localized in both, nucleus and cytoplasm. The *UBR4* gene produces multiple splice variants that presumably have different functions. UBR4 is a N-recognin that binds type 1 and type 2 N-degrons. In the 600 kDa protein there are only two domains identified (**Figure 1.7**). The UBR-box domain, which recognizes type 1 N-degrons and a cysteine-rich domain (CRD), which is unique to UBR4. Even though UBR4 does not have a clear ubiquitin ligase domain such as RING or HECT, studies have shown that the E3 activity depends on the C-terminal fragment of UBR4 that contains the CRD (Morrison et al. 2013).

The initial characterization of p600 suggested a role in the regulation of integrin-mediated signalling and membrane morphogenesis. In particular, UBR4 is associated to detachment-induced apoptosis through interactions with the human papillomavirus (HPV) type 16 E7 oncoprotein and the tumor suppressor retinoblastoma (Nakatani et al. 2005). Later reports showed that UBR4 in fact interacts with all 17 different types of HPV E7 proteins, including those that are not related to cancer, indicating that this interaction might also be important for virus replication. Moreover, the N-terminus of E7 mediates the interaction with UBR4 (White,

Munger & Howley 2016; White et al. 2012). Also, Szalmás *et al* showed that cancer-causing high-risk HPV E7 oncoprotein targets the tumour suppressor protein tyrosine phosphatase PTPN14 for proteasomal degradation by recruiting UBR4 (Szalmas et al. 2017). Other studies have also shown how UBR4 is implicated in other host pathogen interactions. During dengue virus infection, the viral NS5 protein interacts with UBR4 to bridge the host protein STAT2 and induce its degradation. This complex formation allows virus replication and inhibits IFN-I signalling via STAT2 proteolysis (Morrison et al. 2013). During influenza A virus infection, UBR4 is recruited by the viral protein M2 to allow safe passage of viral glycoproteins to the cell membrane, an essential process for virus budding (Tripathi et al. 2015).

UBR4 or p600 can also function in an E3-independent manner. Shim *et al* reported UBR4 as a microtubule-associated protein that binds to the endoplasmic reticulum membranes and is expressed in CNS neurons (Belzil et al. 2014; Shim et al. 2008). Also, UBR4 interacts with calcium-bound calmodulin in the cytoplasm (Nakatani et al. 2005), suggesting a role in calcium-dependent signalling. Similarly, genetic variants of UBR4 are associated with episodic ataxias, rare neurological channelopathies that cause imbalance and lack of coordination (Conroy et al. 2014). Studies on circadian rhythms of behaviour in mice identified UBR4 as a time-of-day dependent and light-inducible protein. However, the physiological targets of UBR4 during the timing mechanisms remain elusive (Ling, HH et al. 2014). Mitochondrial quality control in neurons involves proteasomal degradation through the N-end rule pathway. UBR4 (in addition to UBR1 and UBR2) can target PINK1, a mitochondrial serine/threonine kinase, for degradation by recognizing its N-terminal phenylalanine upon PARL cleavage. When mutated in the autosomal recessive form of Parkinson disease, PINK1 recruits the E3 ubiquitin ligase Parkin in damaged mitochondria to induce its degradation through mitophagy (Yamano & Youle 2013). In summary, different studies have investigated the role of p600 or UBR4 in the brain. Identified functions range from neurogenesis, neuronal signalling and survival to formation of neuronal networks. Further studies on the importance of UBR4 in brain development might shed light onto its implication on neurological diseases (Parsons, Nakatani & Nguyen 2015).

UBR4 is physically associated to the 26S proteasome complex (Besche et al. 2009) and is also involved in bulk degradation in the lysosome (Hong et al. 2015). UBR4 binds cellular cargoes

destined to autophagic vacuoles and degradation by the lysosome. Moreover, UBR4^{-/-} mice died during embryogenesis due to defective vascular development in the yolk sac. This structure supplies amino acids to the embryo in early stages of development, which are generated through lysosomal degradation of endocytosed maternal proteins. These results suggested a dual role of UBR4 in protein degradation through both, autophagy and ubiquitylation. Rad6 or UBE2A/B is associated to UBR4 as its cognate E2, particularly during endosome-lysosome vesicle trafficking (Hong et al. 2015). The mitogen-activated protein kinase (MAPK) signalling controls cell proliferation and is commonly misregulated during tumorigenesis. Recently, UBR4 was found to control MAPK levels in *Drosophila* through the N-end rule pathway, while its action was counteracted by the deubiquitinase USP47 (Ashton-Beaucage et al. 2016).

Ubiquitylation targets of UBR4 vary not only in function but also in tissue and subcellular localization. UBR4 targets the ATP-citrate lyase (ACLY) for ubiquitylation and proteasomal degradation under specific stimuli. Increased cellular concentrations of glucose reduce the interaction of ACLY with UBR4 and therefore its ubiquitylation (Lin et al. 2013). Acetylation of specific lysine residues in ACLY blocks their ubiquitylation preventing proteasomal turnover. These observations expose a crosstalk between the two pathways in the regulation of fatty acid synthesis and cell growth. Similarly, the membrane protein podocin, expressed in the epithelial cells of the kidney, is targeted by UBR4 for ubiquitylation and proteasomal degradation. Ubiquitylation of mislocalized multimers in conserved lysine residues does not only induce degradation but promotes disassembly and unfolding of the monomeric podocin (Rinschen et al. 2016).

1.2.1.3.1.8 UBR5

UBR5, also known as EDD (E3 identified by differential display) or HYD (Hyperplastic discs), was first characterized in humans as an E3 ligase bearing a carboxy-terminal HECT domain (Callaghan et al. 1998). In *Drosophila melanogaster*, UBR5 was first named hyperplastic discs protein (HYD) because of its identification by a temperature-sensitive mutation that causes imaginal discs¹ overgrowth in mutant larvae and infertility (Mansfield, Elizabeth et al. 1994).

¹ An imaginal disc is a sac-like epithelial structure found inside the larva of insects that undergo metamorphosis.

Human UBR5 is ubiquitously expressed, while within the germ cells in rat it is specifically observed in testis and brain and, to a lower level, in lung, liver and kidney (Müller et al. 1992; Oughtred et al. 2002). UBR5 mRNA and protein are also seen frequently overexpressed or mutated in colorectal, breast and ovarian cancers, implying a potential role in cancer development (Clancy et al. 2003; Dompe et al. 2011; Fuja et al. 2004; Mori, Y et al. 2002; O'Brien et al. 2008). UBR5^{-/-} embryos were not viable midgestation, demonstrating a key role of UBR5 in mammalian development. Similar to UBR4-lacking mice, UBR5^{-/-} embryos showed impaired yolk sac vascular development causing nutrient deprivation (Saunders et al. 2004). Point mutations in UBR5 have also been associated with a familial type of epilepsy (Kato et al. 2012).

The Wnt signalling pathway regulates normal development and controls adult stem cell renewal. Hyperactivation of this pathway, usually associated with aberrant stabilization of β -catenin, is commonly found in human cancers. UBR5 was found to stabilize β -catenin by increasing its protein levels and activity through its ubiquitylation. Interestingly, UBR5 does not directly bind β -catenin, instead, it utilizes GSK-3 β as a scaffold to induce ubiquitylation (Hay-Koren et al. 2011). Ubiquitylation is specific to K29 and K11 ubiquitin linkages and does not target β -catenin for degradation. In contrast, ubiquitylation of β -catenin by other E3s down regulates Wnt signalling, an interesting example of complementary yet dissimilar roles of ubiquitylation in the same pathway. Another role of UBR5 in cell cycle progression involves the regulation of the microspherule protein Msp58, a regulator of rRNA gene transcription. UBR5 controls Msp58 through ubiquitylation and proteasomal degradation thus modulating cyclin levels during cell development (Benavides et al. 2013).

In the chromosomes, upon phosphorylation, the telomerase reverse transcriptase (TERT) is targeted for ubiquitylation and proteasomal degradation by a E3 ligase complex formed by UBR5, DDB1 and VprBP proteins, this down regulation leads to inhibition of telomerase activity and subsequent cellular senescence (Jung et al. 2013; Nakagawa, Mondal & Swanson 2013). Interestingly, DDB1 and VprBP are also part of the Cul4A-Roc1 E3 ligase complex, suggesting the ability of UBR5 to act as an E3 ligase under different catalytic complexes (HECT vs. CRL type). The modulator of apoptosis protein 1 (MOAP1) is targeted for proteasomal degradation by

a number of E3 ligases including UBR5. Remarkably, UBR5-induced turnover is cell cycle dependent, suggesting that these different E3 ligases act on the same substrate in response to different stimuli (Matsuura et al. 2017).

The human *UBR5* gene encodes a ~300 kDa HECT-type ubiquitin ligase. There are four domains characterized in UBR5 and two nuclear localization signals. An N-terminal UBA (ubiquitin associated) domain, an UBR-box domain, a MLLE domain (also known as PABC) and a C-terminal HECT domain (**Figure 1.5, Figure 1.7**). UBR5 is a N-recognin because it binds type 1 N-degrons (Tasaki et al. 2009). However, no substrate of the N-end rule pathway has been identified as a UBR5 target. Secondary structure predictions suggest the presence of abundant secondary structure throughout most of the protein, particularly between the UBR-box domain and the MLLE domain. Thus, it is likely that other unidentified domains are found in UBR5 further expanding its functional repertory. Crystal structures of the UBA domain (Kozlov et al. 2007), the MLLE domain (Deo, Sonenberg & Burley 2001) and the C-terminal lobe of the HECT domain (Matta-Camacho et al. 2012) have been determined; however, these domains only account for approximately 15% of the total protein. Reported cognate E2s for UBR5 are UbcH4, UbcH5B and UbcH5C.

UBA domains are common in proteins associated with the UPS, as they are known to mediate ubiquitin interactions. In contrast, the MLLE domain is only present in two proteins in eukaryotic cells: UBR5 and PolyA-binding protein (PABP). MLLE was first characterized in PABP as a protein-protein interaction domain that recognizes PAM2 motifs in effectors of translation initiation, while in UBR5 its role is not fully understood. Studies done on the MLLE domain in PABP and UBR5 highlighted the ability of the latter to bind PAM2 peptides *in vitro* (Lim et al. 2006). The first clue on MLLE function in UBR5 was given by Yoshida *et al*, which showed that Paip2, an inhibitor of translation initiation, was targeted for ubiquitylation and proteasomal degradation by UBR5 (Yoshida et al. 2006). Mammalian $\alpha 4$ phosphoprotein, an essential component of the mTOR pathway, was also found to bind UBR5 and PABP, presumably through the MLLE domain (McDonald et al. 2010). UBR5 also regulates transcription through CDK9, a subunit of positive transcription elongation factor b (P-TEFb).

Polyubiquitylation of CDK9 by UBR5 in association with TFIIIS recruits RNA polymerase II along the γ fibrinogen inducing its transcription (Cojocaru et al. 2011).

DNA damage response is one of the main pathways where UBR5 is actively involved. Henderson *et al*, showed that UBR5 interacts with the calcium- and integrin-binding protein/DNA-dependent protein kinase-interacting protein (CIB) while located in the nucleus (Henderson et al. 2002). This interaction is sensitive to DNA damage. This study also revealed that UBR5 binds importin α 5 and the progesterone receptor PR enhancing its transcriptional activity. Honda *et al* indicated that the DNA topoisomerase II β -binding protein 1 (TopBP1) is ubiquitylated and targeted for proteasomal degradation by UBR5. Upon double strand DNA breaks, TopBP1 is phosphorylated inhibiting its ubiquitylation in cells (Honda et al. 2002). Also, UBR5 interacts with the checkpoint kinase CHK2 to induce its activation in response to DNA double stranded breaks. Down regulation of DNA damage repair proteins is a common promoter of tumorigenesis, which further suggests the involvement of UBR5 in cancer development (Henderson et al. 2006; Munoz et al. 2007). Histone ubiquitylation is a common response to double strand DNA lesions that is triggered by the E3 ligase RNF8. UBR5 along with the HECT type E3 TRIP12 control RNF168 levels by inducing its ubiquitylation and proteasomal degradation to counteract excessive ubiquitylation in intact chromosomes (Gudjonsson et al. 2012; Okamoto et al. 2013). The ATM kinase and its cofactor ATMIN activate cell cycle checkpoints and promote DNA repair. In response to ionizing radiation, UBR5 ubiquitylates ATMIN, which reduces its interaction with ATM, thus facilitating ATM function in double strand breaks repair (Zhang, T et al. 2014). In damaged chromatin, SPT16, part of the histone chaperone complex, is ubiquitylated by UBR5, which possibly leads to inhibition of RNA polymerase II elongation in DNA lesions (Sanchez et al. 2016). TIP60, an acetyltransferase implicated in DNA damage response and apoptosis, is ubiquitylated and targeted for degradation by UBR5. During HPV infection, the E6 viral protein utilizes this interaction to destabilize TIP60 and induce tumour formation (Subbaiah et al. 2016). UBR5 also promotes gastric cancer development by targeting gastrin 1, a stomach-specific protein, for ubiquitylation (Yang, M et al. 2016).

Numerous proteins have been identified as binding partners of UBR5. Such interactions induce a variety of outcomes that control key processes in the cell. Surprisingly, a great number of these associations seem to be E3-independent, that is, do not result in ubiquitylation by UBR5. For instance, MAPKs, also known as extracellular signal-regulated kinases or ERKs, can act as key regulators of various cellular processes including proliferation, differentiation and migration. Identification of ERK substrates revealed UBR5 as a target for *in vivo* and *in vitro* phosphorylation (Eblen et al. 2003). Moreover, later proteomic studies identified 24 phosphorylation sites in human UBR5 (Bethard et al. 2011). Myocardin, a transcription factor that promotes expression of smooth muscle-specific genes, also interacts with UBR5 in an E3-independent manner. This interaction enhanced trans-activation of smooth muscle-specific promoters and prevented myocardin degradation without affecting its mRNA expression (Hu, G et al. 2010). Other E3-independent binding partners of UBR5 are p53 (Ling, S & Lin 2011; Smits 2012), B55 α (Reid et al. 2013), GW182 proteins in the argonaute-miRNA complexes (Su et al. 2011), the HPV type 18 E6/E6AP ligase (Tomaic et al. 2011), RIP1 and cIAP1 during apoptosis (Christofferson et al. 2012), human herpes virus-6 U14 protein during induced cell-cycle arrest (Mori, J et al. 2015), ROR γ t during TGF- β signalling (Rutz et al. 2015) and CDC20, BUB3 and BUBR1 proteins during cellular response to abnormal mitosis (Scialpi, Mellis & Ditzel 2015).

Posttranslational modifications in substrates are a common mechanism to prompt or prevent ubiquitylation by UBR5. During gluconeogenesis, the rate limiting enzyme phosphoenolpyruvate carboxykinase (PEPCK1) is regulated in response to changing levels of glucose. Under high glucose concentrations, PEPCK1 is destabilized by acetylation. This modification induces binding of UBR5 that in turn ubiquitylates and targets PEPCK1 for proteasomal degradation (Jiang, W et al. 2011). The human chemical toxin sensor pregnane X receptor (hPXR) is a ligand induced transcription factor that induces expression of detoxifying enzymes. Phosphorylation of hPXR stimulates interaction with UBR5 and subsequent ubiquitylation and proteasomal degradation (Ong et al. 2014). In contrast, phosphorylation of nuclear myosin 1c (NM1) prevents its polyubiquitylation by UBR5 leading to its stabilization and induction of RNA polymerase I transcription activation (Sarshad et al. 2014).

1.2.1.3.1.9 UBR6

Also known as FBXO11 or VIT1. UBR6 was first observed during differential display experiments in melanocytes of patients with the pigmentary disorder vitiligo, implicating it with the disease (Le Poole et al. 2001; Li, Y et al. 2009). FBXO11 is also associated with otitis media (inflammation of the ear) as well as cleft palate defects in mice (Hardisty-Hughes et al. 2006; Rye, Bhutta, et al. 2011; Rye, Blackwell & Jamieson 2012). Later studies demonstrated the association of inherited chronic otitis media in human with single nucleotide polymorphisms in *FBXO11* (Rye, Wiertsema, et al. 2011; Segade et al. 2006).

UBR6 was also initially identified as a human type II arginine methyltransferase (PRMT9) producing symmetrically double methylated arginine (Cook et al. 2006). As UBR6 is a target of splicing with 6 isoforms identified, Cook *et al* suggest that an isoform that lacks the F-box domain could be a substrate of alternate splicing that activates its methyltransferase function. As a supposedly methyltransferase, its protein sequence is unique amongst other PRMTs. However, no further work supports these observations and another enzyme has been assigned the PRMT9 name (Hadjikyriacou et al. 2015). Moreover, FBXO11 possesses a F-box and an UBR-box domain, strongly suggesting a role as an E3 ubiquitin ligase part of the SCF complex (Fielenbach et al. 2007). Tasaki *et al* showed that despite the presence of the UBR-box domain, UBR6/FBXO11 was unable to bind type 1 or type 2 N-degrons (**Figure 1.7**) (Tasaki et al. 2009).

In the immune response FBXO11 was identified as part of a CDK9/CCNT1 complex that negatively regulates HIV-1 Tat function and viral gene expression. It is suggested that FBXO11 reduces the level of CDK9 and CCNT1 thus limiting their interaction with the viral protein Tat, an activator of RNA polymerase II (Ramakrishnan et al. 2012). In *C. elegans* FBXO11 homolog DRE-1 targets the transcriptional repressor BLMP-1 for proteasomal degradation controlling aging and maturation. This interaction is also conserved in mammals (Horn et al. 2014; Huang, TF et al. 2014).

FBXO11 functions as an E3 ligase part of the SCF complex. FBXO11 promotes p53 neddylation² and not ubiquitylation *in vivo* and *in vitro* leading to its inactivation (Abida et al.

² Conjugation of the ubiquitin-like protein Nedd8

2007; Tateossian et al. 2015; Xue et al. 2016). FBXO11 targets BCL6, a proto-oncoprotein implicated in human B-cell lymphoma, for ubiquitylation and proteasomal degradation. *FBXO11* gene is absent or mutated in different diffuse large B-cell lymphoma cell lines, which was associated with increased BCL6 stability (Duan et al. 2012). Mutations on FBXO11 were specially found in a region of the protein with a putative CASH domain (**Figure 1.7**). These mutations abrogated ubiquitylation of BCL6 by FBXO11 suggesting it functions as a substrate-binding domain for the SCF^{FBXO11} complex. CASH domains bear glycine and hydrophobic residue repeats that fold into right-handed beta-helix structures and are often found in carbohydrate-binding proteins and sugar hydrolases (Cicarelli et al. 2002). However, their function remains elusive. A number of studies have shown a link between mutations on FBXO11 and cancer development. Pancreatic cancer studies on humans and mice (Mann et al. 2012) identified mutations on the *FBXO11* gene that were associated with poor prognosis. Similar studies have identified associations in splenic marginal zone lymphoma (Parry et al. 2013), glioblastoma, melanoma, prostate cancer (Ngollo et al. 2017; Yang, CH et al. 2015), breast cancer (Xue et al. 2016) and leukemia (Nagel et al. 2017).

FBXO11 along with UBR5 are examples of cross-regulation of E3 ligases through the UPS. Cdt2, a substrate recognition subunit in the CRL4 complex of E3 ligases is also a FBXO11 substrate. Its ubiquitylation and proteasomal degradation is important during the response to TGF- β , exit from the cell cycle and cellular migration (Abbas et al. 2013; Rossi et al. 2013). As an exception from the known mechanisms implicated in substrate recognition by F-box proteins, phosphorylation of the substrate can abrogate Cdt2 binding to FBXO11. This is also observed for BCL6. On the other hand, reports on the phosphorylation-dependent degradation of the transcription factor SNAIL by FBXO11 show the highly specific mechanisms that this E3 ligase has to recognize its different substrates (Jin, Y et al. 2015; Zheng et al. 2014). In addition, FBXO11 was shown to regulate voltage-gated sodium channel Nav 1.5 through proteasomal degradation along with UBR3 (Zhao et al. 2015), an interesting interplay between UBR proteins.

1.2.1.3.1.10 UBR7

UBR7 is by far the least studied UBR protein in the family. The *UBR7* gene encodes a ~50 kDa polypeptide that comprises a plant homeodomain (PHD) domain and an UBR-box domain, both

in the N-terminus of the protein. The UBR-box does not bind N-degrons (Tasaki et al. 2009), while the function of the PHD domain remains elusive (**Figure 1.7**). Secondary structure predictions on UBR7 sequence suggest the presence of additional domains due to abundant secondary structure. However, no putative domains can be identified based on structure similarity. The PHD domain is a zinc finger that resembles the structure of RING and FYVE domains. Due to this similarity, UBR7 is suggested to be an E3 ubiquitin ligase although there is no strong evidence as for this or any other function. The only reference to date is that of Zimmerman *et al*, which reported the first study on UBR7 in mammalian spermatozoa. Here UBR7 was found to have E3 ligase activity in mouse and boar testis and spermatozoa. These observations however, suggest that this activity is not essential for sperm function during fertilization (Zimmerman et al. 2014).

1.3 The proteasome

The 26S proteasome is a multi-subunit complex that catalyzes the breakdown of ubiquitylated proteins. In mammals, it controls at least 80% of protein degradation and its function is key for protein homeostasis as it influences the majority of cellular processes. Remarkably, protein degradation is not its only function. Recent studies have shown how the proteasome complex has the ability to select whether an ubiquitylated protein is degraded or remains functional by removing the attached ubiquitin (Collins & Goldberg 2017). The group of Alfred Goldberg first discovered this complex protein machine in 1987 after initial reports on protein degradation by the UPS. Purification of a multi-protein complex from reticulocytes revealed a distinct ATP-dependent protease of 1.5 MDa that only cleaved substrates if they were conjugated to ubiquitin (Waxman, Fagan & Goldberg 1987). Today, numerous other proteins have been identified in the complex, which makes up to over 2.5 MDa with the primary function of ATP-dependent protein breakdown. The proteasome exists in multiple forms but the major assembly contains a 20S 28-subunit core particle, called CP, and a 19S regulatory particle RP, which consists of around 19 subunits (**Figure 1.8**).

The proteasome works in a multi-step mechanism as the rest of the UPS, where the final step catalyzes protein digestion into short peptides averaging 2 to 10 residues in length (Kisselev et

al. 1999). Nearly all products are molecules that are reclaimed by the cell. Interestingly, some of these digestion products can serve as precursors for antigenic peptides displayed on MHC-class I molecules (Collins & Goldberg 2017; Goldberg et al. 2002). The multi-step mechanism starts with the recognition of the ubiquitin-conjugated protein, followed by unfolding, translocation and deubiquitylation of the substrate. All these critical processes are directed by the regulatory particle RP through the action of six distinct ATPases that exert the protein-unfoldase activity (**Figure 1.8**). The proteolytic active site, where peptide bond cleavage occurs, is enclosed in the internal space of the cylindrical hollow core particle CP, whereas the RP unit controls substrate entry through translocation. Recognition of the substrate occurs through its ubiquitin chain by the RP. Next, ubiquitin molecules are cleaved off by proteasomal deubiquitylases (DUBs) and recycled by the cell. Deubiquitylation allows unfolding and subsequent entry of the substrate to the CP.

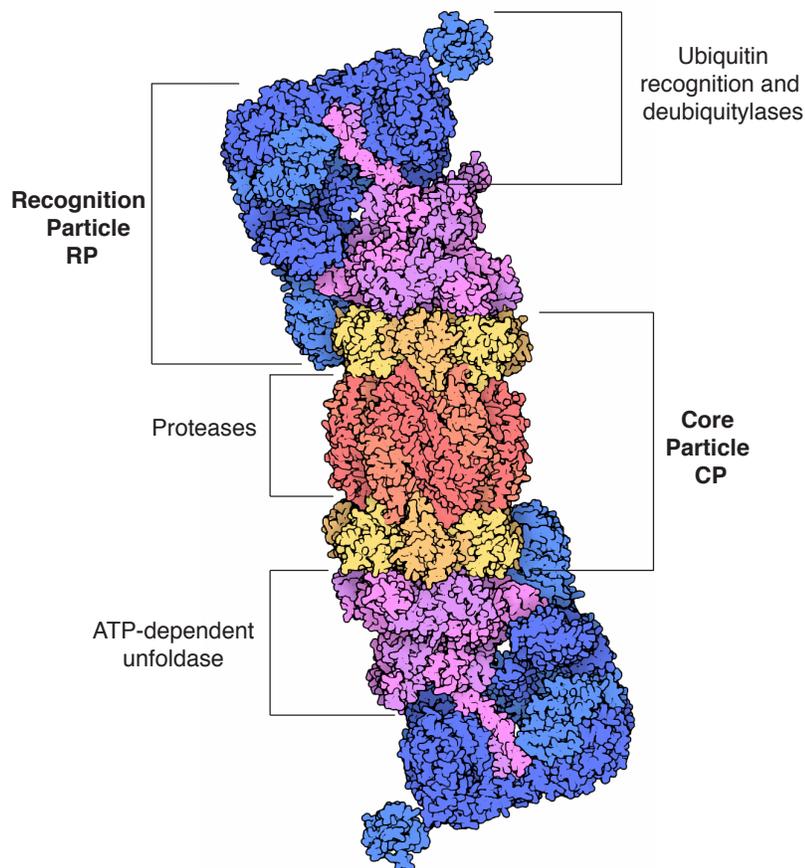


Figure 1.8. The 26S proteasome complex.

Structure of the yeast proteasome determined by hybrid methods. Several crystallographic structures were fit into an electron microscopy reconstruction. The active site is located in the core particle CP shown in yellow and red. Three

types of proteases (red) catalyze the breakdown of the polypeptide chain inside the CP. At both ends of the cylindrical particle, an ATP-dependent unit (magenta) unfolds proteins and induces translocation to the CP. In blue, regulatory subunits such as deubiquitylases and UBA-bearing proteins recycle the ubiquitin moieties detached from the substrates. Modified from “Molecule of the Month” Protein Data Bank. PDB: 4CR2.

One of the distinct characteristics of the proteasome is its ability to specifically and rapidly degrade proteins in the cell. Its function is correlated with environmental conditions that dictate the levels of this multi-enzyme complex during cell proliferation. In *Drosophila*, the 26S proteasome levels increase during growth and decrease in the aging process, whereas in mammals, they are predominantly found in the nuclei during cell proliferation (Tanaka 2009). Therefore, it is not surprising that the proteasome is been studied as a promising target for drug design and development. Inhibition of the UPS through the proteasome impedes cell cycle progression and cell survival. In fact, there are two FDA-approved drugs for the treatment of cancer that inhibit the proteasome: bortezomib and carfilzomib. Despite the advances on understanding the structural and functional features of the system, there is still a lack of information regarding specificity and selectivity of particular proteasome types as well as their relationship with disease and cellular homeostasis. New advances on structure solving methods such as high-resolution cryo-EM have helped advance the knowledge of the system in past few years, including the upstream enzymatic regulators. Hopefully these advances will elucidate detailed mechanisms that serve as platform for therapeutic applications.

1.4 Concluding remarks

Despite the great increase of research on the UPS and in particular the N-end rule pathway, there are still many unresolved questions. As the scope of the N-end rule expands to all natural amino acids, one of the main questions to address is their timely and specific recognition by E3 ligases. As reviewed in this chapter, the vast majority of E3 ligases in the UBR family of proteins are involved in specific maladies. Uncovering the mechanisms underlying substrate recognition and ubiquitin transfer by the E3 ligases is essential not only for understanding the regulation of cellular processes but also will provide a structural framework for the design of specific therapeutic compounds.

In this thesis work I explore some of the substrate recruitment mechanisms in the mammalian UBR family. One particular feature in structures of the UBR-box was the partial occupation of the negative pocket upon binding of arginine (the largest N-degron residue). This observation prompted the question of whether the UBR-box could bind modified N-degrons, small molecules, or even internal degrons. Despite the rigid scaffold by which the UBR-box supports N-degron recognition, there is a great degree of binding plasticity evidenced by the various sizes and shapes of the classical N-degrons. Thus, I investigated the ability of the UBR-box to bind modified N-degrons. This study further clarified the specificity of the Arg/N-end rule for both, N-terminal and second residues, and explained the mechanisms involved in pathological mutations.

CHAPTER TWO: BOUND WATERS MEDIATE BINDING OF DIVERSE SUBSTRATES TO A UBIQUITIN LIGASE

Juliana Muñoz-Escobar, Edna Matta-Camacho, Cordelia Cho, Guennadi Kozlov, Kalle Gehring. *Structure* 25(5): 719-729 (2017).

2.1 Summary

The N-end rule pathway controls the half-life of proteins based on their N-terminal residue. Positively charged type 1 N-degrons are recognized by a negatively charged pocket on the Zn-finger named UBR-box. Here, we show that the UBR-box is rigid but bound water molecules in the pocket provide the structural plasticity required to bind different positively charged amino acids. Ultra-high resolution crystal structures of arginine, histidine and methylated arginine reveal that water molecules mediate binding of N-degron peptides. Using a high-throughput binding assay and isothermal titration calorimetry, we demonstrate that the UBR-box is able to bind methylated arginine and lysine peptides with high affinity and, measure the preference for hydrophobic residues in the second position in the N-degron peptide. Finally, we show that the V122L mutation present in Johansson-Blizzard syndrome patients changes the specificity for the second position due to occlusion of the secondary pocket.

2.2 Introduction

UBR ubiquitin ligases are E3 ligases, which catalyze the last step in the ubiquitylation cascade that modifies proteins for degradation and intracellular signalling (Ciechanover, Aaron & Iwai 2004). The selectivity of ubiquitylation is a result of the creation or exposure of degradation signals (Schrader, Harstad & Matouschek 2009; Varshavsky 1991; Yu et al. 2015). Among the best studied is the N-end rule pathway, where the N-terminal residue, called an N-degron, controls the half-life of proteins (Varshavsky 2011). In the N-end rule, the destabilizing N-terminal residues are grouped in two: type 1 N-degrons, composed of basic residues such as arginine, lysine and histidine, and type 2 N-degrons, formed by bulky hydrophobic residues (Gibbs et al. 2014). The recognition of N-degrons in the N-end rule branch is mediated by the action of a family of ubiquitin ligases, termed the UBR family. The members are defined by the presence of a domain of approximately 70 residues called the UBR-box (Tasaki et al. 2005; Tasaki et al. 2009). In mammals, UBR1, UBR2, UBR4 and UBR5 target type 1 N-degrons (Tasaki et al. 2005). Arginine is the strongest binder followed by N-terminal lysine and histidine.

Crystal structures of yeast and human UBR-box show three zinc atoms organized in two contiguous zinc fingers that stabilize the tertiary structure. A negatively charged pocket is partially occupied by the arginine side chain (Choi et al. 2010; Matta-Camacho et al. 2010), while a secondary pocket interacts with the second position of the N-degron. However, how the binding site is simultaneously optimized to accommodate different N-terminal residues remained elusive.

The study of the molecular determinants responsible for substrate recognition is of particular interest in the development of E3 ubiquitin ligases as drug targets (Bulatov & Ciulli 2015; Hamilton, Lee & Le Roch 2014; Liu et al. 2015; Skaar, Pagan & Pagano 2014). Recent studies highlight the diversity of the physiological roles in which the N-end rule pathway is a regulator (Sriram, SM, Kim & Kwon 2011). UBR1 is the most studied E3 ubiquitin ligase in the pathway. Functional inactivation of the two chromosomal copies of the UBR1 gene causes the congenital disorder called Johanson-Blizzard syndrome (JBS) (Sukalo et al. 2014; Zenker et al. 2006). In mice, double knockout of UBR1 and UBR2 results in early embryonic lethality (An et al. 2006; Kwon et al. 2003; Kwon et al. 2001). Given the stringency of the N-end rule pathway, the UBR-box domain is a potential major target for the development of small-molecule effectors that

mimic the destabilizing character of cognate N-degrons (Agarwalla & Banerjee 2016; Jiang, YXL et al. 2013; Kwon, Levy & Varshavsky 1999; Lee, JH et al. 2015). One of the challenges faced when studying the functional implications of the N-end rule is the lack of specific inhibitors that target each family member independently.

In the present study, we reveal the ability of the UBR-box domain to bind non-canonical N-degrons. We demonstrate that the UBR-box binds N-terminal methylated arginine and lysine peptides with high affinity. We elucidate the mechanism for the plasticity of binding with a complete structural study of six different ligands. The polyvalence of the domain is the result of the strategic binding of water molecules around the negative pocket. Finally, we present a comprehensive study of the molecular determinants for recognition in the second and third positions in the N-degron. We examine the functional and structural implications of the V122L mutation associated with JBS, and show that N-degron specificity is altered and binding affinity decreased in the mutant protein.

2.3 Results

2.3.1 Ultra-high resolution structure of the human UBR-box in complex with arginine N-degron

We determined the crystal structure of the UBR2-box in complex with the tetrapeptide Arg-Leu-Trp-Ser (referred to as RLWS) at 0.79 Å (**Figure 2.1 A**). The UBR-box domain displays a negatively charged pocket responsible for the recognition of the arginine residue. The amino group defines the specificity for N-terminal arginine forming salt bridges with Asp150 and Phe148. The only water molecule involved in binding is located in the proximity of Asp118 and Thr120 bridging them with NH ϵ and NH2 ω ' of arginine. In contrast, Asp153 forms hydrogen bonds with both NH2 ω groups in the arginine side chain (**Figure 2.1 B**). The second residue in the bound peptide, leucine, is accommodated in the hydrophobic groove named secondary pocket. Hydrophobic interactions govern the recognition of the second residue while the third position in the peptide displays more mobility as seen in the alternate conformation of the tryptophan residue. In our previous complex structure with RIFS peptide (Matta-Camacho et al. 2010), the

only water molecule observed in the negative pocket was also found in the proximity of Asp118 and Thr120, while the second residue of the peptide, isoleucine, also binds the secondary pocket in a hydrophobic interaction. The ultra-high resolution of the new structure confirmed the role and location of water molecules in the binding of N-terminal arginine and highlighted the nature of the interactions involved in recognition. The arginine- and histidine-bound structures reveal the modulatory role of polar water molecules in human N-degron recognition.

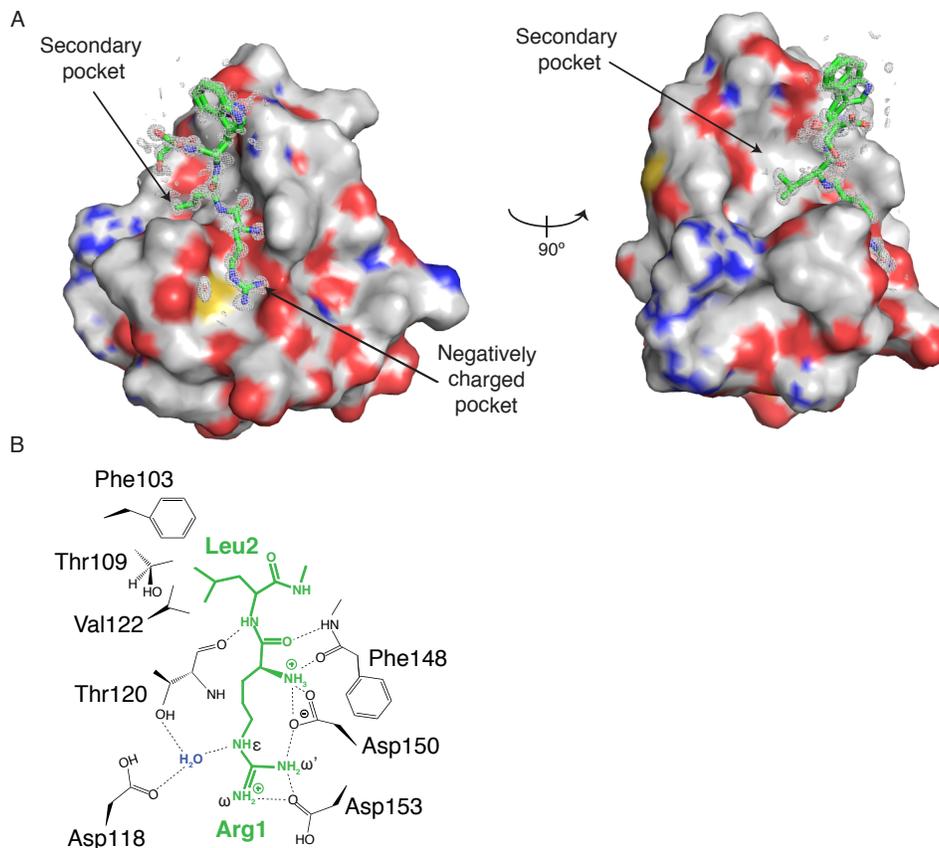


Figure 2.1. Ultra-high resolution structure of the UBR-box domain in complex with destabilizing peptide RLWS at 0.79 Å.

A. Surface representations of the human UBR-box domain from UBR2 in complex with the tetrapeptide RLWS. The 2Fo-Fc electron density map is contoured at 1σ and carved at 2.0 Å around the bound water molecule and peptide. Only Asp118 forms a water bridge with $\text{NH}\epsilon$ from arginine. **B.** Schematic representation of the molecular determinants involved in arginine and leucine recognition.

2.3.2 Crystal structure of the human UBR-box in complex with a histidine N-degron

A high-resolution (1.55 Å) crystal structure was obtained for the complex of human UBR2-box and the tetrapeptide HIFS (**Figure 2.2 A**). The N-terminal histidine is located in the negatively charged groove and establishes a network of organized water molecules that contact the histidine side chain. The interaction of the N-terminal amino group is determinant for binding as in the case of the arginine N-degron (**Figure 2.2 B**). The amino group of histidine forms three hydrogen bonds with Phe148 (NH backbone) and Asp150 (both carboxyls in side chain). As opposed to the arginine-bound structure, the histidine side chain does not interact directly with the pocket. Instead, the N δ 2 in the imidazole group connects to the carboxyl side chains of Asp153 and Glu155 through a network of hydrogen bonds with three organized water molecules in the binding pocket (**Figure 2.2 B, C**). In contrast, N δ 1 locates 3.4 Å away from the closest hydrogen bond donor suggesting that the only important stabilizing contact of the histidine side chain occurs through the N δ 2 and the water bridges. Asp118 and Glu155 undergo conformational changes in comparison to the arginine-bound structure in order to accommodate histidine binding (**Figure 2.2 D**). In the arginine-bound structure, Asp118 mediates hydrogen bonds with water. Upon histidine binding, Asp118 moves away from the binding pocket while Glu155, not involved in arginine binding, approximates to the negative groove. Thr120, involved in arginine binding, is 3.8 Å away from the imidazole group, impeding any electrostatic interaction. We tested the effect of Glu155 on binding of N-terminal histidine by measuring the dissociation constant of E155A UBR2-box with HIFS peptide (**Figure 2.3**). We observed a K_d of 38 μM compared to 34 μM for the wild-type protein, suggesting that most of the stabilizing effect of histidine comes from the interaction with Asp153.

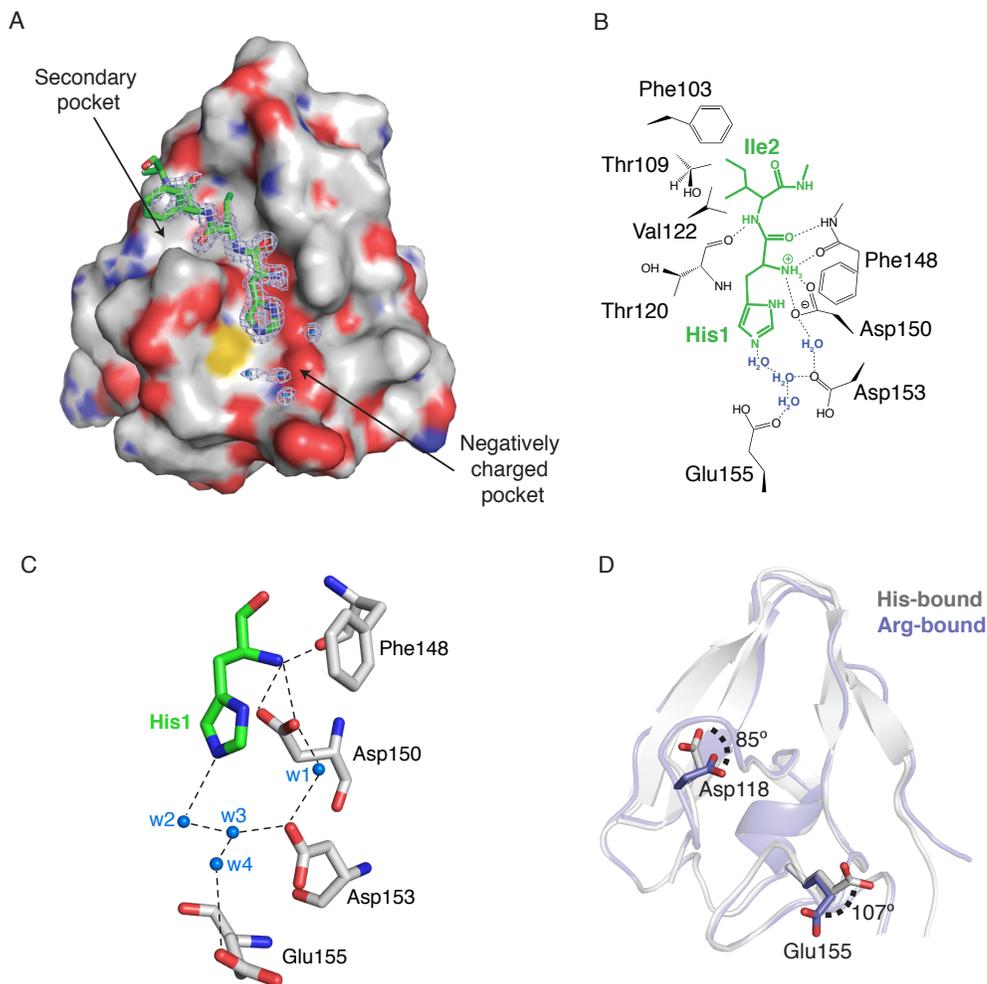


Figure 2.2. Histidine N-degron binds the negative pocket through a highly structured network of water molecules.

Crystal structure of the human UBR2-box in complex with HIFS peptide at 1.5 Å resolution. **A.** Electrostatic potential surface representation of the UBR2-box domain in complex with the tetrapeptide HIFS showing the charge complementarity and three bound water molecules. The 2Fo-Fc electron density map was obtained before any ligand was added to the model. Map was contoured at 1.5 σ and carved at 2.0 Å. **B.** Schematic representation of the interactions involved in histidine recognition. **C.** Water molecules stabilize the contacts between the imidazole group of histidine and the pocket. Water w1 keeps Asp153 in the right orientation to connect with w3. N δ 2 interacts with the pocket through the water network formed by w2 to w4. **D.** Changes in the side chain orientations of Asp118 and Glu155 upon arginine (purple) and histidine peptide binding (grey).

The human UBR2:HIFS complex differs substantially from the yeast UBR1:HIAA structure. In yeast, the N δ 1 in the imidazole group hydrogen bonds directly with Thr144 (Thr120 in human) while no water molecules are present in the negative groove (Choi et al. 2010). This direct interaction of the side chain with the binding pocket explains the higher affinity of histidine N-degron observed for yeast ($K_d = 13 \mu\text{M}$) compared to human UBR-box ($K_d = 34 \mu\text{M}$) (**Figure 2.4**). Human and yeast UBR-box domains have 46% identity, with most interacting residues conserved in both organisms, however, small differences in the global fold change the role of certain residues during binding. It is important to highlight that Choi *et al.* only observed binding of HIAA peptide at pH 6.5. This suggests that the protonation state of histidine influences binding to the UBR-box in yeast. In the human domain, the imidazole group only interacted with water molecules in the pocket, making it difficult to discern the protonation state of histidine.

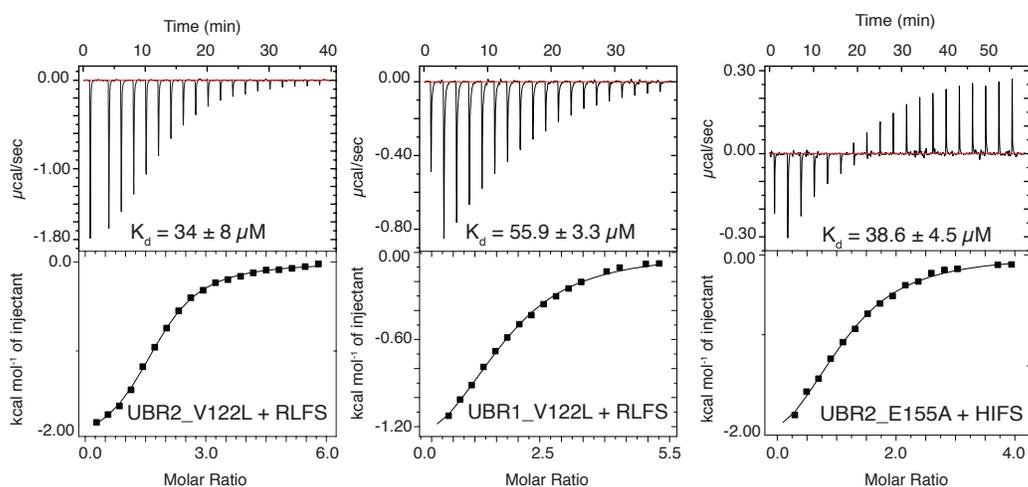


Figure 2.3. ITC affinity measurements for UBR-box domains bearing the V122L and E155A mutations with type 1 N-degron tetrapeptides.

Isothermal titration calorimetry experiments for different UBR-box mutants. V122L decreases N-degron binding in UBR1 and UBR2. E155A mutant in UBR2 does not decrease affinity for HIFS.

The arginine- and histidine-bound structures reveal the modulatory role of water molecules in N-degron binding by the UBR-box. Water molecules are positioned around the negative pocket promoting hydrogen bonds in spaces not occupied by the bound ligand. N-terminal arginine, which binds with the highest affinity, interacts with three residues in the pocket, directly with

Asp150 and Asp153 and through water with Asp118 (**Figure 2.1 B**). In contrast, histidine N-degron needs at least two interconnected water molecules, w2 and w3, to interact with Asp153 and an additional with Glu155 (w4) (**Figure 2.2 C**). W1 bridges Asp150 and Asp153, keeping the latter in optimal position for interaction with w3. The role of w1 in the modulation of N-degron binding seems to be fundamental as our methylated arginine complexes conserve w1 as a stabilizing module (discussed later).

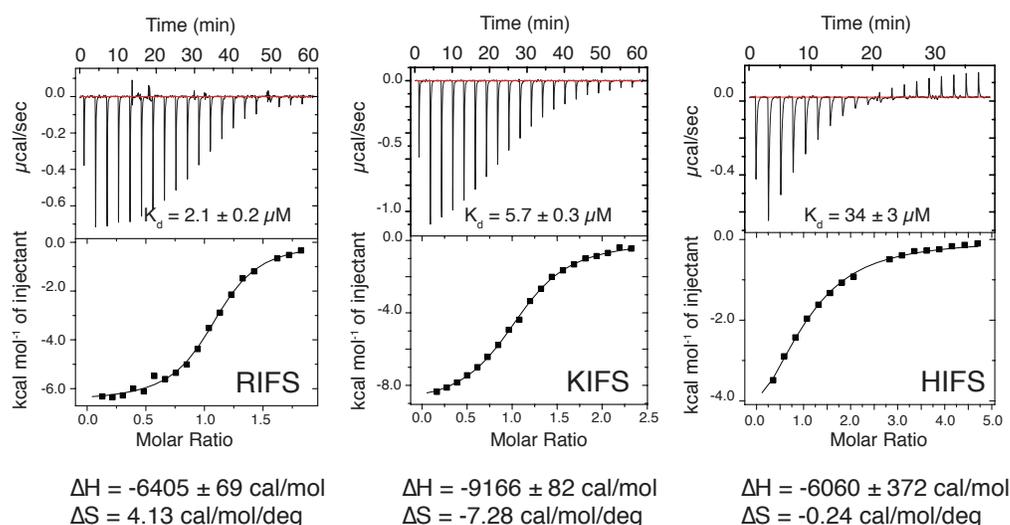


Figure 2.4. ITC affinity measurements of UBR-box domain from UBR2 binding to type 1 N-degrons.

Isothermal titration calorimetry measurements with tetrapeptides bearing N-terminal arginine, lysine and histidine. Arginine is the most destabilizing residue followed by lysine and histidine.

2.3.3 Post-translational modifications and N-degrons

Methylation of the nitrogen atoms of arginine and lysine residues is frequently involved in regulation of gene transcription and molecular recognition. Arginine methylation is catalyzed by different types of methyltransferases, which add one or two methyl groups to generate monomethylated arginine (Han, Kommaddi & Shenoy) or symmetrically or asymmetrically double methylated arginine (SDMA or ADMA) (Beaver & Waters 2016). Similarly, lysine can be mono-, double- or tri- methylated by the action of lysine methyltransferases, frequently

involved in regulation of histones (Boriack-Sjodin & Swinger 2016). We were interested in how these post-translational modifications could affect N-degron recognition. We used thermal shift assays (also called differential scanning fluorimetry) to measure the thermal stability of the UBR2-box:peptide complexes using RIFS and KIFS peptides and compared them with their methylated versions. Increment of the melting temperature upon addition of ligand is characteristic of higher affinity binding. Our results show an increase of $\sim 1^\circ\text{C}$ for asymmetric methylation of RIFS compared to unmodified peptide, while dimethylation of KIFS had the same effect on thermostability as the unmodified peptide (**Figure 2.5 A**). Dissociation constants measured by isothermal titration calorimetry (Soucy et al.) for RIFS and $^{\text{ADMA}}$ RIFS are virtually the same while $^{\text{DM}}$ KIFS affinity is $13\ \mu\text{M}$ compared to $5.7\ \mu\text{M}$ for the unmethylated peptide (**Figure 2.5 B**). These results demonstrate that methylated N-terminal arginine and lysine bind to the UBR-box domain of UBR2 with high affinity.

To understand the mechanism underlying this flexibility in ligand specificity, we determined the crystal structures of the UBR2-box in complex with $^{\text{ADMA}}$ RIFS and the UBR1-box in complex with $^{\text{MMA}}$ RIFS. The resolution of both structures, 1.1 and 1.6 Å, respectively, allowed the identification of key water molecules involved in binding (**Figure 2.5 C, D**). The different states of arginine methylation induce conformational changes in the side chains that highlight the plasticity of the UBR-box domain interactions. As seen in Figure 2.5 C, addition of methyl groups induces a rotation in the arginine side chain towards Asp118 that allows the peptide to hydrogen bond directly without the need of water. At the expense of this turn, Asp153 no longer binds directly the $^{\text{ADMA}}$ Arg side chain and needs a water molecule to mediate the hydrogen bond. In the $^{\text{MMA}}$ Arg- and $^{\text{ADMA}}$ Arg-bound structures, water w1 bridges Asp150 and Asp153 as seen in the histidine complex. In the $^{\text{MMA}}$ Arg, w1 forms an extra contact with the free NH₂ group further stabilizing the complex. All the structures illustrate how water molecules modulate the high affinity interaction with the negative pocket. In particular, w1 keeps Asp150 and Asp153 in the optimal orientation to bind the available NH groups from arginine side chains. In the case of histidine, w1 position is conserved with a similar purpose, preserve side chain positioning for water bridge formation.

Methyl groups in both, ^{ADMA}Arg and ^{MMA}Arg, are stabilized by the aliphatic side chains of Val117, Ala156 and Glu155 (**Figure 2.5 D**). Even though the bound methylated conformer has a different hydrogen bond pattern compared to wild-type arginine, the main interacting residues in the pocket are conserved. The use of water bridges instead of direct interactions often has a negative impact on the affinity. However, conformational restraints imposed by the methylation of the arginine side chain locate hydrogen bond donors closer to Asp118 and Thr120 which enhance binding by direct interactions. This displacement of the ligand over the whole negative surface exposes strategic places for the addition of donor/acceptor groups that could support all hydrogen bond requirements without water bridge formation.

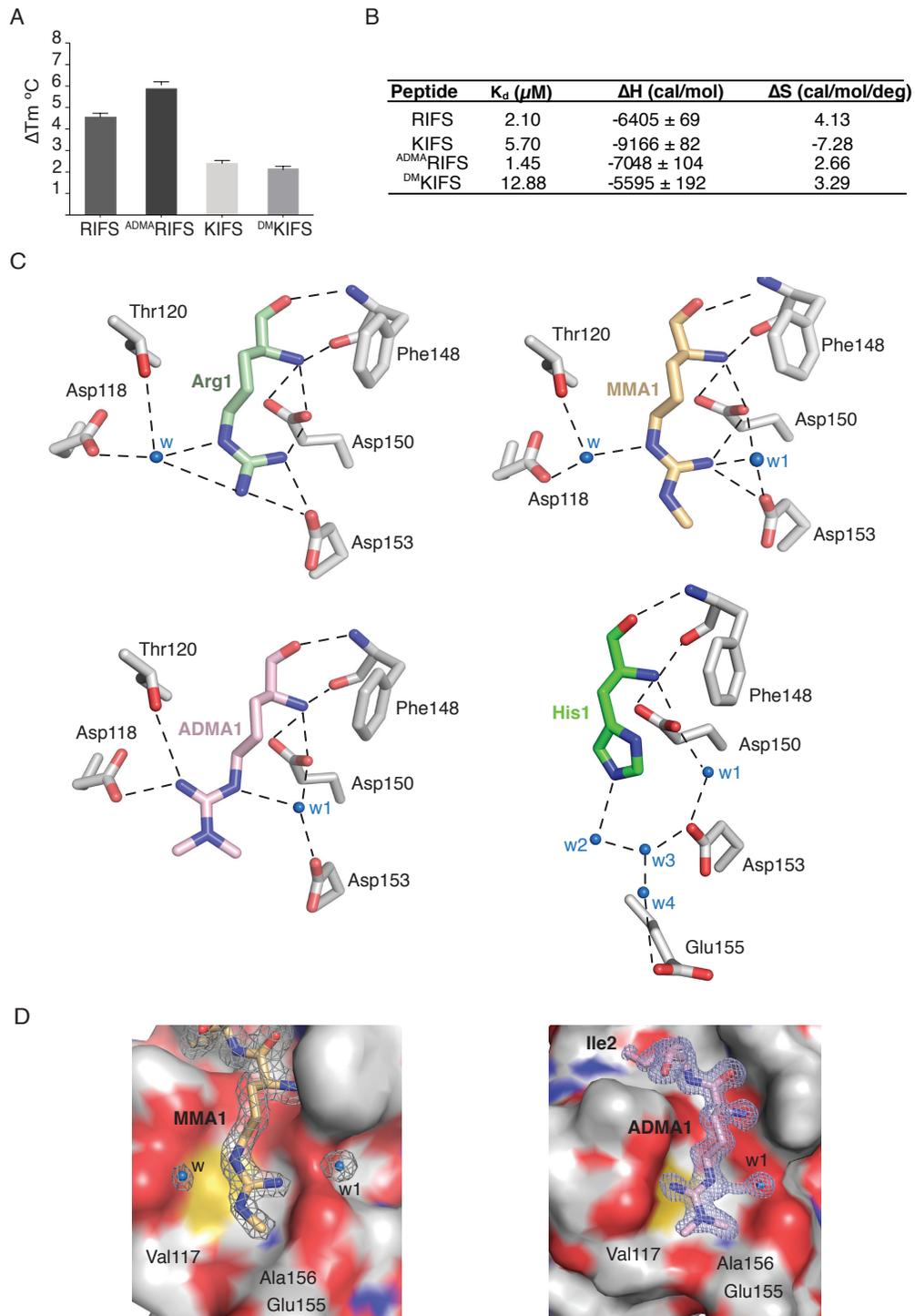


Figure 2.5. Methylated N-terminal arginine and lysine peptides bind to UBR-box domain with high affinity.

A. Melting temperature changes upon binding of methylated and classic N-degron peptides bearing N-terminal arginine and lysine. ΔT_m is the difference between the T_m of the UBR2-box:peptide complex and the T_m of unbound UBR2-box. **B.** Dissociation constants, enthalpy and entropy changes for different peptides with UBR2-box

measured by ITC. **C.** Comparison of the recognition elements involved in arginine, MMA, ADMA and histidine peptide binding to the UBR-box. Hydrogen bonds are represented by dotted lines. Only N-terminal residue of peptides is shown for clarity. Water is a key element in the stabilization of the interactions. W1 adopts the same position for MMA, ADMA and histidine residues. In all three structures, w1 stabilizes Asp153 to optimize contacts with the ligands. **D.** Surface representation of UBR1-box in complex with MMA peptide (left) and UBR2-box in complex with ADMA peptide (right). Electron density maps 2Fo-Fc contoured at 1 σ for MMA and 2 σ for ADMA, carved at 2.0 Å. Maps were calculated before ligand was added to the model. Maps were obtained before any ligand was added to the model.

2.3.4 A rigid scaffold that displays multiplicity of binding

The diverse size and shape of the UBR-box N-degrons prompts the question of how does the domain adapt to bind all different ligands. For all arginine and methylated arginine structures the backbone and side chains of the domain remain in the same conformation as in the unbound structure. Only upon histidine binding, the side chains of Asp118 and Glu155 reorient to form water bridges. We calculated the backbone RMSD per residue of the five crystal structures of the UBR2-box and compared the largest and smallest N-degrons (**Figure 2.6 A**). A comparison of all structures of bound and unbound UBR2-box shows no changes in the overall conformation of the domain (**Figure 2.6 B**). Pairwise comparison of ^{ADMA}RIFS and HIFS complexes showed that the protein remains rigid despite large changes in the size of the N-terminal ligand (**Figure 2.6 A, purple**). To overcome the rigid scaffold and assist binding of different ligands, the domain uses water molecules as flexible adapters. Strategic binding of water molecules supplies specific hydrogen bonds for each ligand.

To further understand the role of bound water molecules in the negative pocket, we analyzed the unbound crystal structure of the UBR-box from UBR2. In the 1.2 Å resolution structure (V122L UBR2-box), two water molecules were observed bound to the negative pocket. The positions of the waters were the same as in the histidine- and methylated arginine-bound proteins: w1 interacts with Asp150 and Asp153 while w3 stabilizes the Asp 153 side chain (**Figure 2.6 C**). This conservation in the positions of bound water molecules highlights their importance for the stability of the protein and ligand binding. The peptide with an N-terminal arginine displaces w1 and w3 by directly interacting with Asp150 and Asp153 side chains.

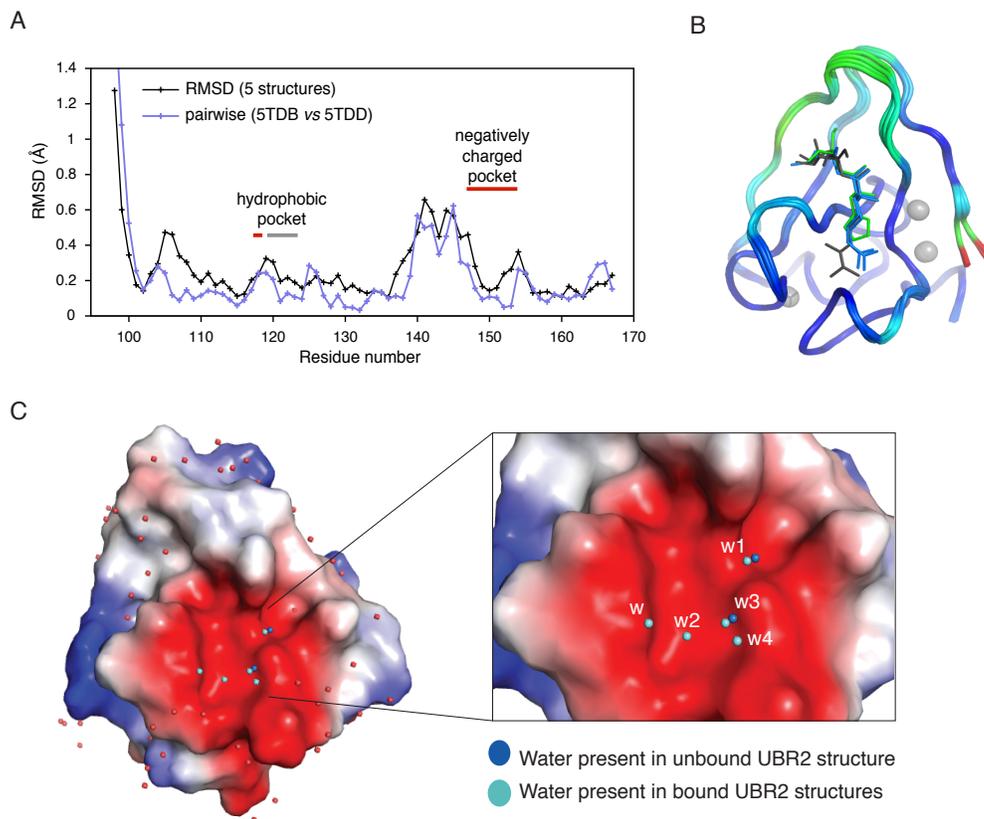


Figure 2.6. The UBR-box of UBR2 forms a rigid scaffold.

A. Backbone RMSD per residue of the five UBR2-box crystal structures and a pairwise comparison between the complexes with the largest and smallest N-degrons. The graph shows an absence of ligand-dependent changes in the N-degron-recognition site. Coordinates used were 3NY2 (ligand-free), 3NY3 (complex with RIFS peptide), 5TDA (RLWS peptide), 5TDB (^{ADMA}RIFS peptide), and 5TDD (HIFS peptide). **B.** Superposition of the five UBR2-box structures color-coded according to the RMSD. Arginine (*blue*), dimethylarginine (*grey*) and histidine (*green*). Grey spheres represent zinc ions. **C.** Surface representation of V122L UBR2-box structure. Red spheres represent bound water molecules around surface. Cyan spheres represent water molecules bound to the negative pocket in bound UBR2-box structures (Arginine, ADMA, MMA and histidine peptides). Blue spheres represent water molecules bound to the negative pocket in unbound UBR2-box structure.

2.3.5 Specificity of the second and third positions of the N-degron

We investigated the dependency of the second and third positions in the N-degron for binding to the UBR box. We selected arginine as the N-terminal residue and evaluated all 19 amino acid substitutions in the second and third positions of the tetrapeptide RXXS. Thermal shift assays were conducted to classify the destabilizing peptides as strong or weak binders. In the assay, high affinity peptides increase the thermal stability of UBR2-box whereas poor binders do not alter its melting temperature. We tested all 20 peptides with the RXFS sequence and observed the highest change in melting temperature (T_m) for large, hydrophobic residues (**Figure 2.7 A**). The maximum increase in stability (ΔT_m) was approximately 7°C for Leu and Tyr. In contrast, the smallest increases were observed for His (~2°C) and proline (~0°C) indicating weak or no binding (**Figure 2. 7A, C, F**). In the UBR2:peptide structures, the backbone amino group of the second residue is a hydrogen bond donor to the hydroxyl group of Thr120. When proline is present in the second position, the absence of a hydrogen atom at the cyclic nitrogen prevents hydrogen bond formation with Thr120 (**Figure 2.1 B, Figure 2.2 B**). In addition, the imino group in proline might be far from the side chain of Val122 reducing stabilizing contacts between the peptide and the secondary hydrophobic pocket. The loss of these contacts is enough to prevent the interaction of the N-degron with the UBR box even in the presence of N-terminal arginine. The ~7°C difference in T_m observed between the best and worst binders in this group highlights the importance of the second position for optimal binding to the UBR box domain.

Next, we evaluated the influence of the third position using peptides bearing the sequence RAXS. The highest T_m change was observed for bulky hydrophobic residues such as tryptophan (~3.5 °C) and phenylalanine (~3.4 °C) whereas the lowest was for histidine (~1.5 °C) (**Figure 2.7 B**). No residue abrogated binding to UBR2 and the T_m difference in this group of peptides was ~2 °C. Our results suggest that the third position has a modest contribution to binding at least in the context of the isolated domain. As seen in the crystal structures, the side chain of the third position does not directly interact with the domain (**Figure 2.1 A, 2.2 A**); it remains to be tested if there is an effect of the third position in the context of the full length protein.

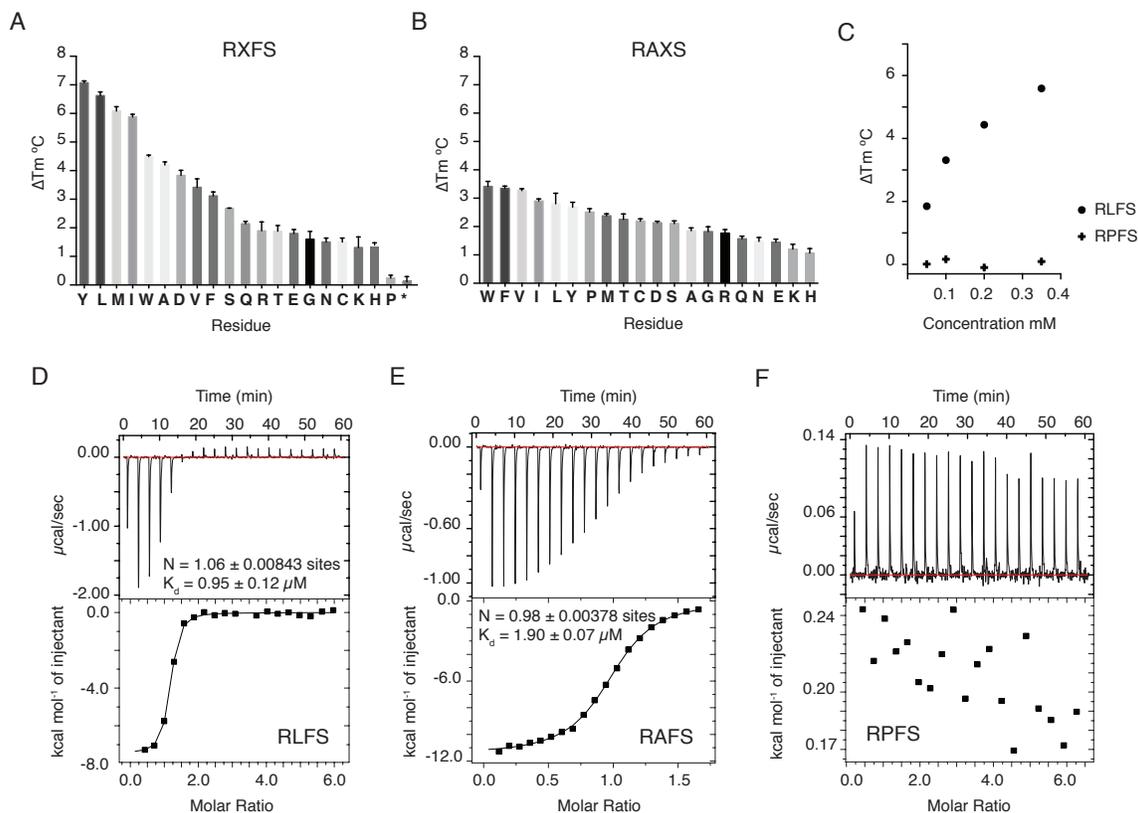


Figure 2.7. The second position of type 1 N-degrons is important for binding to the UBR-box.

A. Melting temperature change of the UBR2-box upon addition of RXFS peptides, where X represents any of the 20 natural amino acids. * Indicates the negative control GRIFS, that bears an N-terminal glycine. Error bars represent standard deviation. **B.** Melting temperature change of UBR2-box upon addition RAXS peptides. **C.** Change of melting temperature versus peptide concentration for UBR2-box with the best and worst binding peptides. **D-F.** Comparison of ITC experiments for selected peptides confirms the selectivity in the second position for hydrophobic residues and against proline.

For select peptides, we measured the affinity by ITC experiments (**Figure 2.7 D-F; Table 2.1, Figure 2.4**). In our previous work, we reported binding affinities for selected tetrapeptides bearing different second positions (Matta-Camacho et al. 2010). In recent experiments, we noticed that the use of 50 mM Tris as a buffer had an impact in the affinities measured previously. We repeated ITC experiments using 20 mM HEPES buffer and observed markedly higher affinities. For the highest affinity peptides, the improvement in affinity was roughly 40-

fold. N-degron binding is highly dependent on the amino group of the N-terminus; competition between Tris and the peptides likely explains the lower affinities observed previously.

Our ITC measurements confirmed that arginine and lysine N-degron have the highest affinity whereas histidine N-degrons bind roughly 15 times less tightly (**Figure 2.4**). Leucine in the second position gave the highest affinity peptide (Table 2.1). We did not observe any binding with proline as the second amino acid (**Figure 2.7 F**). For the third position, tryptophan and phenylalanine (RAWS and RAFS) showed the highest affinity at 2 μ M whereas histidine in the third position (RAHS) only has a 2-fold difference in affinity (**Table 2.1**).

Table 2.1. Dissociation constants, enthalpy and entropy changes for the UBR-box domain from UBR2 binding different tetrapeptides as measured by ITC.

Peptide	K_d (μ M)	ΔH (cal/mol)	ΔS (cal/mol/deg)
RYFS	1.00	-12130 ± 210	-13.9
RLFS	0.95	-7470 ± 96	2.07
RMFS	1.38	-10300 ± 9	-8.31
RAFS	1.90	-11480 ± 62	-13.00
RIFS	2.10	-6405 ± 69	4.13
RWFS	4.46	-10670 ± 285	-11.90
RHFS	4.60	-10280 ± 92	-10.70
RPFS	No binding	-	-
RAWS	2.00	-10090 ± 53	-8.40
RAHS	4.00	-7850 ± 83	-2.08
RLWS	1.20	-6264 ± 59	5.70
KIFS	5.70	-9166 ± 82	-7.28
KLFS	3.70	-7504 ± 49	-0.76
HIFS	34.36	-6060 ± 372	-0.24
^{ADMA} RIFS	1.45	-7048 ± 104	2.66
^{DM} KIFS	12.88	-5595 ± 192	3.29
RLFS (UBR2_V122L)	34.01	-2115 ± 23	13.3
RLFS (UBR1_V122L)	55.86	-1337 ± 83	14.9
HIFS (UBR2_E155A)	38.61	-2075 ± 89	13.1

2.3.6 Johanson Blizzard syndrome (JBS) and the second position in the N-degron

The JBS is a genetic disorder inherited as an autosomal recessive trait, which originates from the loss of UBR1 (Zenker et al. 2005). To date, multiple pathogenic mutations along the *UBR1* gene have been identified. Analysis of some JBS mutations located in the UBR1-box domain (C127F,

H136R, G160AfsX5, H166R) suggests that the disease arises from a loss of Zn²⁺ binding residues and structural stability (Matta-Camacho et al. 2010; Sukalo et al. 2014). In contrast, the V122L mutation does not affect Zn²⁺ binding or residues involved in N-terminal recognition. Previous studies have shown that this mutation inhibits proteasomal degradation of a model substrate in yeast (Hwang, CS et al. 2011). To understand the molecular implication of this mutation, we used thermal shift assays to test the binding properties of all different RXFS peptides. UBR1- and UBR2-box domains share a 77% identity with identical negative pockets. However, during thermal shift assays, UBR1-box did not show a thermal curve preventing the calculation of its melting curve. For this reason, we introduced the V122L mutation in UBR2-box as it shares very similar binding properties to UBR1.

The addition of the peptides to the V122L mutant sample had a small effect on the thermostability of the protein, showing loss of binding. The highest ΔT_m was seen for the RAFS and RMFS peptides (~1.6 and 1.3 °C respectively) while RLFS did not have a significant increase in T_m ($\Delta T_m = 0.4$ °C) (**Figure 2.8 A**). For the wild type UBR2-box, RLFS had an increase of 7°C in the T_m . In contrast, RLFS binding is lost for the V122L mutant. All peptides tested had a similar effect as RLFS where no improved thermal stability was observed in the mutant UBR2-box. ITC experiments also showed a decrease in affinity. Mutant UBR1- and UBR2-boxes dissociation constants were 55 μ M and 36 μ M, respectively, showing a 30-fold decrease in binding affinity (**Figure 2.3 A, B; Table 2.1**). In the structure of UBR2-box in complex with RLWS, the recognition of Leu2 by the secondary pocket is mediated by hydrophobic interactions through residues Phe103, Thr109 and Val122 (**Figure 2.8 B**). The V122L mutation while conservative, introduces a bulkier residue in the secondary pocket. We obtained the crystal structure of the V122L UBR2-box domain. The orientation of the Leu122 side chain suggests that the close proximity of the bound second residue from the N-degron would cause a repulsive effect limiting optimal binding to the pocket (**Figure 2.8 B**). This effect is enough to decrease binding affinity of the arginine peptide by 36-fold in the case of RLFS. These results suggest that this mutation would impact drastically the targeting of substrates with already low affinities such as His-bearing proteins. The hydrophobic effect governing the recognition of the second position in the N-degron is further supported by alanine in the second position. RAFS peptide had the highest increase in thermostability for the mutant UBR2 (1.6 °C),

which shows that a small hydrophobic residue in the second position could still bind although with a clear decrease in affinity.

We also obtained the first structure of the human UBR1-box in complex with a peptide (^{MMA}RIFS). All residues involved in recognition in the negative pocket are conserved as in UBR2; however, the secondary pocket has an additional residue, His141 that is located in proximity to the bound isoleucine from the peptide. In Figure 2.8 C, the surface representation compares the wild type and V122L mutations in the UBR1-box and its effect on the secondary pocket. Leu122 would completely occlude the pocket as His141 also reduces the space available for binding of the second position. Our results explain the fundamental role of the secondary pocket in substrate recognition and contribute to the understanding of the mechanisms involved in the pathogenesis of JBS.

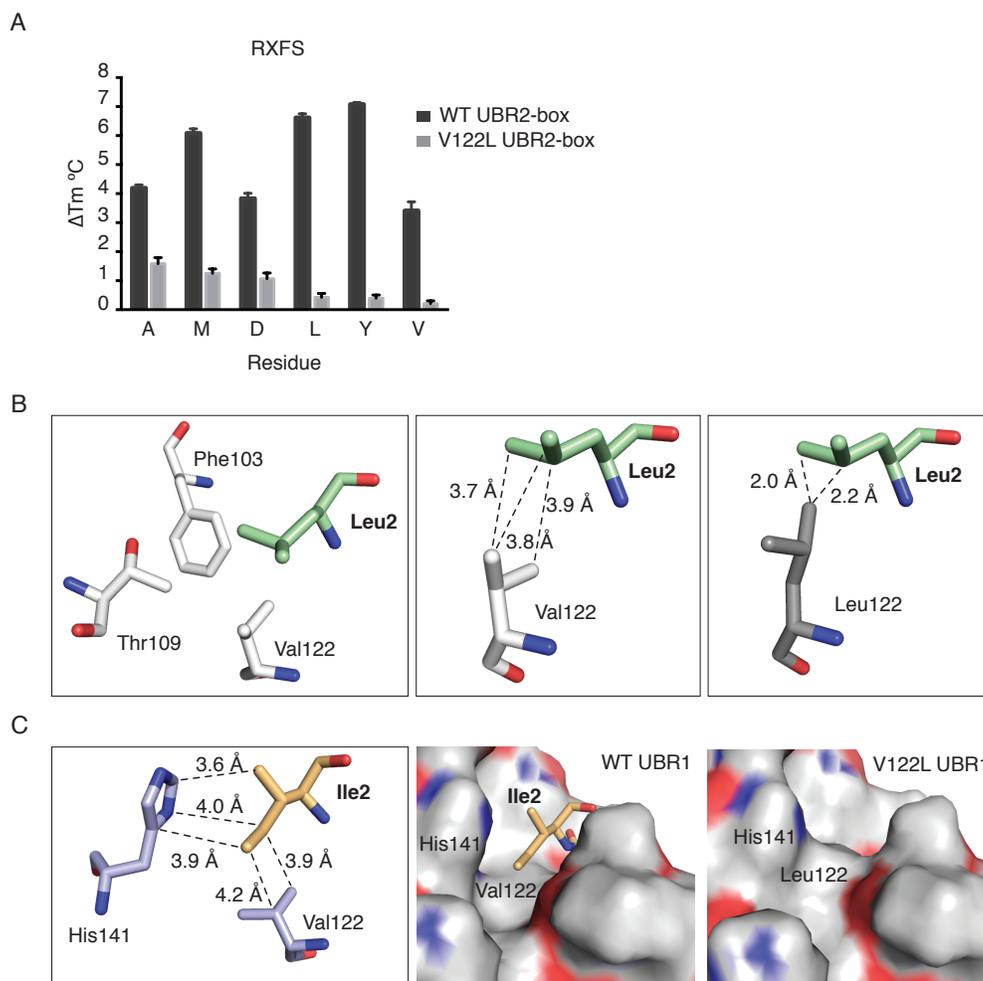


Figure 2.8. The V122L mutation in Johansson Blizzard syndrome prevents binding of N-degrons with bulky residues in the second position.

A. Comparison of the melting temperature change of wild type UBR2-box and V122L UBR2-box upon addition of peptides RXFS, where X is the residue shown. **B.** The hydrophobic pocket in UBR2-box is formed by Phe103, Thr109 and Val122 side chains; with hydrophobic contacts between Leu2 of the N-degron and Val122 of UBR2. Crystal structure of V122L UBR2-box shows how Leu122 occludes the pocket and prevents binding of Leu2 (modelled). **C.** Ile2 from the UBR1 structure in complex with MMA peptide. His141 is part of the secondary pocket in UBR1 along with Val122, Phe103 and Thr109. Surface representation of wild type and V122L UBR1 shows that Leu122 (modelled) occludes the pocket preventing binding of the second residue in the N-degron.

Table 2.2. Collection and refinement statistics for crystallographic data

	UBR2 RLWS	UBR2 ADMA RIFS	UBR1 MMA RIFS	UBR2 HIFS	V122L UBR2
Data collection					
PDB code	5TDA	5TDB	5TDC	5TDD	5UM3
Space group	P 2 ₁ 2 ₁ 2 ₁	P 1 2 ₁ 1	P 2 ₁ 2 ₁ 2 ₁	P 4 ₁	P 2 ₁ 2 ₁ 2 ₁
Cell dimensions					
a, b, c (Å)	28.7, 37.16, 57.47	29.57, 37.17, 29.75	47.27, 49.04, 53.63	29.31, 29.31, 74.15	29.18, 29.37, 66.05
α , β , γ (°)	90, 90, 90	90, 109.50, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	17.03 - 0.79 (0.81 - 0.79)	15.46 - 1.10 (1.14 - 1.10)	26.82 - 1.61 (1.67 - 1.61)	20.73 - 1.55 (1.60 - 1.55)	17.62 - 1.198 (1.241 - 1.198)
R _{merge}	0.048	0.062	0.081	0.1	0.098
I/ σ I	38.11 (1.39)	32.69 (2.03)	17.75 (2.32)	32.12 (4.95)	17.87 (1.61)
Completeness (%)	96.3	99.5	97.4	99.9	98.10
Redundancy	5.4 (2.6)	4.1 (3.9)	4.5 (2.9)	6.4 (5.7)	7.0 (3.1)
Refinement					
Resolution (Å)	0.79	1.1	1.61	1.55	1.19
No. reflections	65534 (4982)	24584 (2376)	16359 (1370)	9083 (902)	126332 (4935)
R _{work} /R _{free}	0.12/0.128	0.125/0.134	0.149/0.179	0.106/0.152	0.137/0.158
No. atoms	1422	1157	1252	650	669
Macromolecules	672	554	1165	595	574
Ligands	3	20	16	11	3
Solvent	107	54	71	44	92
Average B- factors	10.78	20.17	22.17	21.15	17.75
Macromolecules	8.88	18.72	21.76	19.89	16.36
Ligands	5.68	21.39	16.75	35.48	12.27
Solvent	22.87	34.64	30.20	34.49	26.63
R.m.s deviations					
Bond lengths (Å)	0.008	0.007	0.009	0.01	0.006
Bond angles (°)	1.06	0.98	1.05	0.96	0.95
Ramachandran					
Favoured (%)	99	99	99	99	100
Allowed (%)	1.2	1.4	1.4	1.4	0
Outliers (%)	0	0	0	0	0
Rotamer outliers	0	0	0	0	0
Clashscore	3.79	0.9	0	0	0

2.4 Discussion

2.4.1 Water as the sculptor of N-degron specificity

The UBR-box domain mediates N-degron recognition through two interacting pockets. A negatively charged pocket determines specificity towards N-terminal residues while a secondary pocket mediates interactions with the second position of the N-degron. In the first pocket, water plays a key role in allowing different type 1 N-degrons to bind. Recognition of the N-terminus is mediated through hydrogen bond formation and electrostatic interactions (salt bridges) with the N-terminal amino group. This interaction is fundamental and is conserved in all type 1 N-degrons. However, the side chains of different type 1 N-degrons are very different in size and shape. Side chain recognition by the negative pocket involves different hydrogen bond arrangements promoted by water. In the case of arginine, the strongest N-degron, the stabilization of the side chain occurs through two direct hydrogen bonds and one water bridge. For histidine, the weakest binder, the imidazole group forms a water bridge that involves three organized water molecules. The lack of direct hydrogen bond interactions with the pocket and the need for a rather “long” water bridge in the imidazole side chain explains the lower affinity observed for the HIFS peptide. In all known structures (**Figures 2.1 B, Figure 2.2 C and Figure 2.5 C**), water molecules participate in recognition promoting water bridge formation and thus, conferring plasticity to the binding. This was demonstrated by the recognition of arginine and its methylated forms. In all three structures, water is involved in hydrogen bond formation; however, its position changes depending on the bound conformation and location of donor groups in the side chain. Histidine also illustrates this plasticity with the conformational changes that take place in Asp118 and Glu155 compared to the Arg-bound structures (**Figure 2.2 D**). These conformational changes of key residues illustrate the structural mechanisms involved in optimal binding where water bridges act as flexible adapters, matching the hydrogen-bonding requirements of the protein.

Water has been shown to support high selectivity for ligands in several protein-peptide interactions (Ladbury 1996; Levinson & Boxer 2014; Sleight et al. 1999; Urakubo, Ikura & Ito 2008). One of the most important sources of binding free energy comes from the displacement of

water molecules that solvate the protein-binding site (Breiten et al. 2013; Jana & Bandyopadhyay 2012; Snyder et al. 2014). Water molecules involved in hydrogen bonding sites are entropically unfavorable because of the spacial limitations related to forming hydrogen bonds with the protein and solvent simultaneously (Michel, Tirado-Rives & Jorgensen 2009). Our high-resolution structures elucidate for the first time the modulatory role of structured waters in human N-degron binding. From the complex structures, it is suggested that a further displacement of waters from the negative groove would lead to an increase in the dissociation constants. Similarly, the displacement of water molecules in the binding site by strategically modifying the ligand is one of the main approaches to lead optimization in rational drug design. However, a complete thermodynamic analysis of water molecules and the effects of their removal upon ligand modification would be necessary during lead optimization.

2.4.2 Hydrophobicity as the second driving force in N-degron modulation

While the N-terminal side chain is positioned and the amino group is recognized in the negative pocket, the second residue accommodates in a hydrophobic groove called the secondary pocket. Residues Phe103, Thr109, Thr120, Val122 and His141 (UBR1 only) are in close proximity to the bound side chain and backbone of the second position; these interactions are essential to achieve low micromolar affinity. Disruption of the hydrogen bond with the backbone NH group in the second residue is sufficient to abrogate binding (RPFS peptide, **Figure 2.7 A, C, F**). Similarly, the identity of the side chain in the second position is also a driving force in the specificity of binding. In our previous work, ITC experiments showed a preference for acidic residues in the second position (Matta-Camacho et al. 2010; Sriram, SM & Kwon 2010), while yeast UBR-box favoured hydrophobic residues (Choi et al. 2010). However, our recent experiments demonstrated that the use of Tris interfered with the affinity measurements performed in our previous work. Thermal shift assays and ITC revealed the highest affinities for hydrophobic residues in the second position. Our results define a rule for the second position of the N-degron that agrees with what was observed in yeast. Our experiments with the V122L mutant also support the importance of the second position side chain where binding was virtually lost in the case of the best binding peptide, RLFS.

2.4.3 JBS and the second position of the N-degron

Given the high sequence identity between UBR1 and UBR2, the Arg/N-end rule pathway is still present in patients with JBS, although, at a lower level of activity (Zenker et al. 2006). Most of the pathogenic mutations found to date are frameshifts or nonsense that lead to complete loss of UBR1. Recently, the point mutation V122L was described in a JBS patient with mild developmental delay and pancreatic insufficiency (Hwang, CS et al. 2011). Our results demonstrated the need of a stable interaction with the hydrophobic pocket of the UBR-box to achieve high affinity. Despite the lack of a known substrate associated to JBS, it is now clear that the V122L mutant would considerably decrease substrate targeting and therefore its ubiquitylation. Moreover, if the JBS substrate has a histidine N-degron, its recognition would be abrogated affecting its metabolic stability. This negative effect will also depend on the identity of the second residue, as alanine in the second position binds with moderate affinity.

We also reported the first structure of the human UBR1-box in complex with a peptide. In this structure, the imidazole group of His141, not conserved in UBR2, locates in the hydrophobic pocket contributing to the stabilization of Ile2 from the N-degron. In the V122L mutant, the presence of His141 would considerably restrict the positioning of the second residue from the N-degron occluding even more the hydrophobic pocket (**Figure 2.8 C**). It is important to point out that the metabolic effect of this mutation will also depend on the characteristics of the substrate such as cellular levels and the type of the mutation seen in the other chromosome. In particular, the patient bearing the V122L mutation in Hwang *et al.* had a frameshift mutation that introduced a stop codon just before the RING domain. RING domains form the scaffold between the E2 conjugating enzyme and the E3 ubiquitin ligase. Loss of this domain would impair recruitment of the E2 therefore inhibiting ubiquitin transfer. In this particular patient, in one copy of the gene substrate recruitment is impaired whereas in the second copy E2 binding is lost. Additional studies involving the RING domain would contribute to the understanding of the N-end rule pathway and its relationship with JBS pathogenesis.

Given the physiological importance of the N-end rule pathway and its numerous variations recently discovered, there is a great interest in developing more potent inhibitors or effectors. Recent studies have analyzed possible inhibitors for the N-end rule pathway using molecular

dynamics and *in silico* docking (Jiang, YXL et al. 2013). A special interest has been observed for the inhibition of the pathway through binding of small molecules to the N-domain (type 2 recognition unit)(Jiang, Y et al. 2014; Jiang, YXL et al. 2013; Sriram, S et al. 2013) Also, studies have exploited the cooperative binding of type 1 and type 2 residues to develop heterovalent inhibitors that target both, the UBR-box and N-domains (Agarwalla & Banerjee 2016). Cooperative binding showed a higher thermodynamic and kinetic selectivity in protein-ligand interactions compared to monomeric binding (Lee, JH et al. 2015). These prototype inhibitors could be further optimized using the thermodynamic properties of the water networks formed around histidine, arginine and methylated arginine in the pocket. Similarly, incorporation of a hydrophobic group after arginine that occupies the secondary pocket could potentially impact the specificity and affinity of the compound.

2.5 Conclusions

Proteosomal degradation relies entirely on effective protein-protein interactions. Deciphering how substrate binding is optimized for each N-degron will provide a structural framework for the rational design of inhibitors for the N-end rule pathway. Similarly, a greater understanding of N-degron recognition will lead to the identification of novel substrates. We presented a complete and definitive study of the specificity of type 1 N-degron binding and exposed the ability of the UBR-box to bind non-canonical N-degrons. This multiplicity of binding is the result of the strategic binding of water molecules within a rigid domain. Water molecules act as binding adaptors allowing the human UBR-box to bind different N-terminal residues with specificity and high affinity. Ligand binding assisted by water suggests it should be possible to design novel high affinity inhibitors that occupy the entire negative pocket. Moreover, hydrophobic interactions with the secondary pocket potentiate binding of the N-degron and bring the affinity to physiological levels. Disruption of these interactions inhibits the recruitment substrates evidenced by the V122L mutant in JBS.

2.6 Materials and Methods

2.6.1 Cloning, expression and purification

Human UBR1 UBR-box domain (aa 98 – 168) and UBR2 UBR-box domain (aa 98 – 167) were codon optimized for *E. coli* and cloned into the pGEX-6p-1 vector. Both proteins were transformed in BL21 cells and expressed in LB media by induction with 1 mM IPTG at OD_{600nm} of 0.8 for 18 hours at 16°C. At the point of induction, 200 µM ZnCl₂ was added to the culture. Purification was performed using affinity chromatography with Glutathione S-transferase (GST) immobilized in agarose beads followed by removal of the GST tag with 3C protease. The final step of purification consisted of size exclusion chromatography (Superdex 75: GE Healthcare) using 20 mM HEPES pH 7.6, 10 mM NaCl and 2 mM β-mercaptoethanol (buffer A) for UBR2-box or 20 mM Tris pH 7.6, 10 mM NaCl and 2 mM β-mercaptoethanol for UBR1-box. All peptides were synthesized by Fmoc solid-phase peptide synthesis, purified by reverse phase chromatography on a C18 column and verified by ion-spray quadruple mass spectroscopy (BioBasic Canada Inc.). The lyophilized powder was resuspended in buffer A.

2.6.2 Crystallization

For the UBR1:^{MMA}RIFS complex, crystals were obtained by equilibrating 1 µl drop of the protein mixture UBR1:^{MMA}RIFS peptide (1:2 molar ratio) mixed with 1 µl of reservoir solution containing 2.2 M (NH₄)₂SO₄, 0.2 K/Na Tartrate in hanging-drop vapour diffusion system incubated at 20°C. All UBR2:peptide crystals were obtained by equilibrating 0.5 µl of the protein mixture in 1:3 molar ratio mixed with 0.5 µl of reservoir solution in sitting-drop vapour diffusion system and incubated at 20°C. UBR2:^{ADMA}RIFS crystals were obtained in 0.1 M Bis-Tris pH 5.5 and 25% PEG 3350; UBR2:HIFS in 0.1 M MES pH 6.5 and 12% PEG 20000; UBR2:RLWS in 0.1 M MES pH 6.5 and 20% PEG 10000; V122L UBR2 in 0.1 M HEPES pH 7.5 and 25% PEG 4000.

2.6.3 Data collection, structure solution and refinement

Crystals were cryoprotected by soaking in mother liquor supplemented with 15% glycerol, flash cooled in liquid nitrogen, and data collection was performed at 100 K. Data sets were collected on a ADSC Quantum-210 CCD detector (Area Detector Systems Corp.) at CHESS beamline A1 ($\lambda = 0.9769 \text{ \AA}$ and $\lambda = 0.6362 \text{ \AA}$). Crystals of UBR1:^{MMA}RIFS belonged to the orthorhombic system, space group $P2_12_12_1$ with 2 molecules per asymmetric unit corresponding to 28% solvent content. Data were indexed, integrated, scaled and merged using HKL2000 (Otwinowski, Z. & Minor, W. 1997). Molecular replacement using the UBR1-box structure was used to determine the complex structure. The initial model obtained from Phaser (McCoy et al. 2007) was improved by several cycles of refinement, using the programs REFMAC (Murshudov, Vagin & Dodson 1997) from the CCP4 suite (Winn et al. 2011) and Phenix.refine (Adams et al. 2010). Crystals of UBR2:^{ADMA}RIFS belonged to the monoclinic system, space group $P2_1$ with 1 molecule per asymmetric unit corresponding to 32% solvent content. The model was improved by several rounds of refinement using the UBR2:RIFS structure (pdb code 3NY3) without peptide using phenix.refine. Crystals of UBR2:HIFS belonged to the tetragonal system, space group $P4_1$ with 1 molecule per asymmetric unit corresponding to 30% solvent content. Molecular replacement using the UBR2:RIFS model (pdb code 3NY3) without peptide was used to determine the complex structure. The initial model obtained from Phaser was improved by several rounds of refinement using phenix.refine. Crystals of V122L UBR2-box belonged to the orthorhombic system, space group $P2_12_12_1$ with 1 molecule per asymmetric unit corresponding to 26.7% solvent content. Molecular replacement with UBR2:HIFS model (pdb code 5TDD) without peptide was used to determine the structure. The initial model obtained from Phaser was improved by several cycles of refinement using the program Phenix. Crystals of UBR2:RLWS belonged to the orthorhombic system, space group $P2_12_12_1$ with 1 molecule per asymmetric unit corresponding to 27% solvent content. Molecular replacement using UBR2:RIFS model (pdb code 3NY3) without peptide was used to determine the complex structure. The initial model obtained from Phaser was improved by several cycles of refinement, using the program phenix.refine. Refinement parameters included explicit hydrogen in riding position, individual B-factors with anisotropic ADPs and refinement of occupancies. The final round of refinement included optimized X-ray/stereochemistry and X-ray/ADP weights. Extra density corresponding to the peptide bound in each structure was extended manually using the program Coot (Emsley et

al. 2010). Structure figures were made with PyMol (Schrodinger 2015). Data collection and refinement statistics are summarized in Table 2.2. The coordinates and structure factors have been deposited in the RCSB Protein Data Bank (accession numbers 5TDA, 5TDB, 5TDC, 5TDD and 5UM3).

2.6.4 Thermal shift assays

Each reaction contained 20 μ l of solution with 50 μ M UBR2-box, 5 μ l of Protein Thermal Shift™ buffer, 1 \times Protein Thermal Shift™ dye (Life Technologies), buffer A and 0.1 mM to 2.5 mM peptide. Peptide concentrations were estimated based on amount of powder weighted and molecular weight. Samples were heated from 25 °C to 99 °C at a rate of 1 °C per minute and fluorescence signals were monitored by the StepOne Plus quantitative real-time PCR system (Life Technologies). Data were analyzed using Thermal Shift software (Life Technologies). All peptides were also tested alone (no protein controls) and flat lines were observed at all temperatures during each experiment. The maximum change of fluorescence with respect to temperature was used to determine the T_m .

2.6.5 Isothermal titration calorimetry

Experiments were carried out on a MicroCal iTC200 in buffer A at 20°C. The sample cell contained 250 μ l of 60 μ M UBR2 and was titrated with 19 injections of 2 μ l of 0.5 to 4 mM peptide. The binding isotherm was fitted with a binding model employing a single set of independent sites to determine the thermodynamic binding constants and stoichiometry.

2.6.6 Quantification and statistical analysis

Thermal shift assays for each peptide were performed in 8 replicates and standard deviation (SD) was calculated for each ΔT_m measurement and represented by error bars.

2.6.7 Data resources

HKL2000 was used to process the raw diffraction data. Different software found in Phenix and CCP4 suites were used to determine, refine and build the structure model. All software were reported in Method Details and indicated in the Key Resources Table. The accession numbers for the coordinates and structure factors of all structures in this paper have been deposited in the PDB under the codes 5TDA (UBR2:RLWS), 5TDB (UBR2:^{ADMA}RIFS), 5TDC (UBR1:^{MMA}RIFS), 5TDD (UBR2:HIFS) and 5UM3 (UBR2_V122L).

2.7 Acknowledgements

We would like to thank the CCP4/Advanced Photon Source summer school 2016 for the advice on structure refinement. Stipends to J.M.E were provided by the Quebec Network for Research on Protein Function, Engineering, and Applications PROTEO and NSERC CREATE Training Program in Bionanomachines. Data acquisition at the Macromolecular Diffraction (MacCHESS) facility at the Cornell High Energy Synchrotron Source (CHESS) was supported by the National Science Foundation award DMR 0225180 and the National Institutes of Health award RR-01646. This study was funded by the Canadian Institutes of Health Research grant MOP-14219 to K.G.

Studies on the UBR-box domain from N-recognins elucidated the N-end rule binding mechanisms and exposed unique structural features that offer plasticity of binding. In the next chapter my aim was to understand why the UBR-box domains in other members of the family do not bind N-degrons and what alternative functions it could have. Thus, I focused my studies on obtaining the crystal structure of the UBR-box from the SCF ubiquitin ligase UBR6/FBXO11.

CHAPTER THREE: CRYSTAL STRUCTURE OF THE UBR-BOX FROM UBR6/FBXO11 REVEALS A DOMAIN SWAPPING MEDIATED BY ZINC BINDING

Juliana Muñoz-Escobar, Guennadi Kozlov, Kalle Gehring. *Under review.*

3.1 Summary

The UBR-box is a 70-residue zinc finger domain present in the UBR family of E3 ubiquitin ligases that directly binds N-terminal degradation signals in substrate proteins. UBR6, also called FBXO11, is an UBR-box containing E3 ubiquitin ligase that does not bind N-terminal signals. Here, we present the crystal structure of the UBR-box domain from human UBR6. The dimeric crystal structure reveals a unique form of domain swapping mediated by zinc coordination, where three independent protein chains come together to regenerate the topology of the monomeric UBR-box fold. Analysis of the structure reveals novel coordination chemistry for zinc in a protein and suggests that the absence of N-terminal residue binding arises from the lack of an amino acid binding pocket.

3.2 Introduction

E3 ubiquitin ligases catalyze the transfer of ubiquitin to proteins targeted for proteasomal degradation. These enzymes recognize specific degradation signals, termed degrons, located in substrate proteins (Hershko & Ciechanover 1998). The N-end rule pathway is a branch of the ubiquitin proteasome system that relates the *in vivo* half-life of a protein to the identity of its N-terminal residue (Bachmair, Finley & Varshavsky 1986). The UBR-box is a ~70 residue zinc finger domain present in the UBR family of E3 ubiquitin ligases. The mammalian genome encodes seven UBR-containing proteins named UBR1 to UBR7. In mammals, UBR1, UBR2, UBR4 and UBR5 target proteins for proteasomal degradation in the N-end rule pathway by binding destabilizing N-degrons through the UBR-box domain (Tasaki et al. 2005; Tasaki et al. 2009). Remarkably, the presence of the UBR-box does not guarantee the ability to recognize N-degrons. UBR3, UBR6 and UBR7 contain a UBR-box but do not bind destabilizing N-terminal residues (Tasaki et al. 2009). In our previous work, we identified the molecular determinants for N-degron recognition in UBR1 and UBR2 by studying the crystal structure of the UBR-box and its ability to bind N-degron peptides. However, the role of the UBR-box domain in proteins that are not part of the N-end rule is still not understood.

UBR6, also called FBXO11, is an F-box subunit of the Skp1-Cullin-F-box (SCF) ubiquitin ligase complex. As an F-box protein, UBR6 directly binds substrates for ubiquitylation and proteasomal degradation. It comprises an N-terminal F-box, three presumptive substrate-binding CASH domains and a C-terminal UBR-box (Duan et al. 2012). Despite the presence of a UBR-box, UBR6 does not bind N-degrons (Tasaki et al. 2009). To understand the role of the UBR-box in proteins that are not part of the N-end rule pathway, we determined the crystal structure of the UBR-box domain from human UBR6. We analyze the distinct dimerization behavior observed in solution and in the crystal, and propose an explanation for the inability of UBR6 to bind N-degrons.

3.3 Results and Discussion

3.3.1 The UBR-box domain of UBR6 is a dimer in solution

UBR6 eluted in two peaks during size exclusion chromatography, suggesting the protein forms both dimers and monomers in solution (**Figure 3.1 A**). Dimerization was not dependent on time, protein concentration or addition of reducing agents. We used multi-angle light scattering coupled with size exclusion chromatography (SEC-MALS) to confirm the molecular mass of the purified fractions (**Figure 3.1 B**). Dimerization has been observed in other zinc fingers (Aras et al. 2009; McCarty et al. 2003) and F-box proteins (Chew et al. 2007; Zhuang et al. 2009) but the UBR-boxes of UBR1 and UBR2 do not form dimers (Matta-Camacho et al. 2010; Munoz-Escobar et al. 2017). It is unclear if the UBR6 UBR-box is a dimer in the context of the full-length protein and if it affects its function.

3.3.2 Domain swapping in the UBR-box domain from UBR6

Crystallization experiments were set up for both monomer and dimer fractions of UBR6. Crystals grew after approximately six months for both samples. All crystals had the same space group and crystal form regardless of the input sample. Solving the structure by phasing with anomalous diffraction of the zinc atoms showed that the crystals had four identical molecules in the asymmetric unit arranged as two dimers. Each chain consists of three α -helices, two antiparallel β -strands and two long loops (**Figure 3.1 C**). Dimerization is driven by zinc atoms: a central symmetrical tetrahedral zinc coordinated by histidines 848 and 883, and two pairs of zinc fingers with three zinc atoms each (**Figure 3.1 C, E**).

To our knowledge, the tetrahedral topology of the central zinc-binding site composed of only histidine residues is very rare (Banaszak et al. 2012) (**Figure 3.1 C**). The most common coordination topology of tetrahedral zinc sites is Cys₂-Cys/His-Cys/His (with positional variations). Around 30% of structural zinc coordination spheres in the PDB have a Cys₄ topology while only 4% have a His₃-Asp sphere (Laitaoja, Valjakka & Janis 2013). The histidines composing the central zinc-binding site are not conserved in UBR-box domains in other UBR proteins (**Figure 3.2 A**).

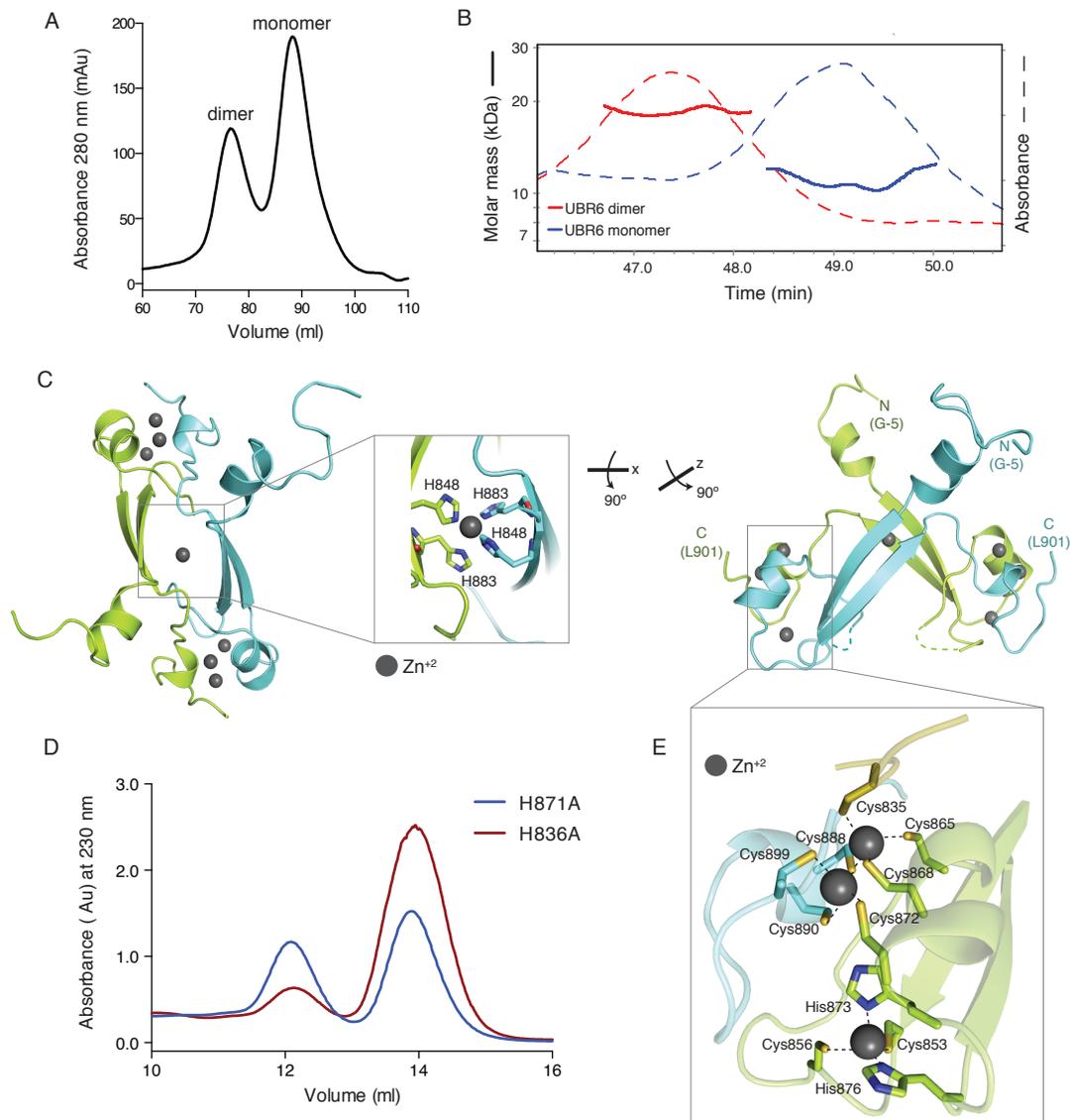


Figure 3.1. UBR-box from UBR6 forms different dimers in solution and in the crystal structure.

A. Size exclusion chromatogram of the UBR-box from UBR6 shows monomer and dimer fractions. **B.** Molecular weight profiles from SEC-MALS for monomer (11 kDa) and dimer (18 kDa) fractions. **C.** Crystal structure of the UBR-box from UBR6. The dimer structure is stabilized by seven zinc atoms. The central zinc binding-site has an unusual topology with four histidine residues in a tetrahedral coordination sphere. Secondary structural elements are α_1 (Ile839-Tyr845), α_2 (Val866-Cys872) and α_3 (Asp889-Ala892), β_1 (Met847-Cys853) and β_2 (Val878-Asp884). **D.** Size exclusion experiment shows that loss of the central zinc coordination sphere alters but does not prevent dimerization in solution. **E.** Intermolecular zinc coordination in the UBR-box structure of UBR6. Zinc binding residues are conserved in the UBR family.

To test the effect of this zinc-binding site on dimerization, we mutated the histidines to alanine and probed the mutant proteins by size exclusion chromatography. Dimer and monomer peaks were observed for both the His848Ala and His883Ala mutants but the His848Ala mutant showed an increase in the fraction of monomers (**Figure 3.1 D**). A His848/883Ala mutant was cloned but the protein did not express. These results suggest that the dimerization observed in solution is not dependent on the central zinc coordination site. We believe this zinc coordination site may be an artifact of crystal packing as zinc was present in the crystallization drop at 10 μ M.

The remaining zinc binding sites offer some insight into the significance of the structure and suggest that it arises from a unique form of domain swapping. The second zinc finger has an unusual topology also observed in UBR1 and UBR2 (Matta-Camacho et al. 2010), consisting of two zinc ions each tetrahedrally coordinated by the shared cysteine 868 (**Figure 3.1 E**). All but one of the zinc-coordinating residues are conserved across the UBR family (**Figure 3.2 A**). In UBR1 and UBR2, a histidine residue forms a Cys₃-His coordination sphere while, in UBR6, this residue is replaced by Cys899 to generate a Cys₄ site. A remaining zinc atom in each zinc finger is bound to a Cys₂-His₂ site found in UBR1/2 structures (**Figure 3.1 E**).

Even though most of the zinc-binding residues are conserved in UBR6 compared to UBR1 and UBR2 (**Figure 3.2 A**), the structure of UBR6 is strikingly different. Analysis of the UBR6 structure suggests that it is the result of domain swapping where three independent protein chains come together to regenerate the topology observed in the monomeric UBR-box fold (**Figure 3.2 B, D**). The first zinc ion is coordinated by Cys868 and Cys872 from chain one (green), and Cys890 and Cys899 from chain two (cyan). The second zinc atom is coordinated by Cys865 and Cys868 from chain one (green), Cys888 from chain two (cyan), and Cys835 from an adjacent chain in the asymmetric unit (yellow) (**Figure 3.1 E, Figure 3.2 B**). This zinc finger is observed twice in the structure as two protein chains coordinate each zinc-binding site in opposite sides of the dimer (**Figure 3.1 C**). The third zinc finger has a typical Cys₂-His₂ topology as observed in UBR1 and UBR2 (**Figure 3.1 E**). This zinc-binding site is the only one formed by one protein chain. In UBR1 and UBR2, these zinc fingers form a rigid scaffold that frames the N-degron binding site. Given the conservation of the zinc coordination spheres, we believe the combined

structure observed in the right panel of Figure 3.2 B is probably a close representation of the physiological structure of the UBR-box from UBR6.

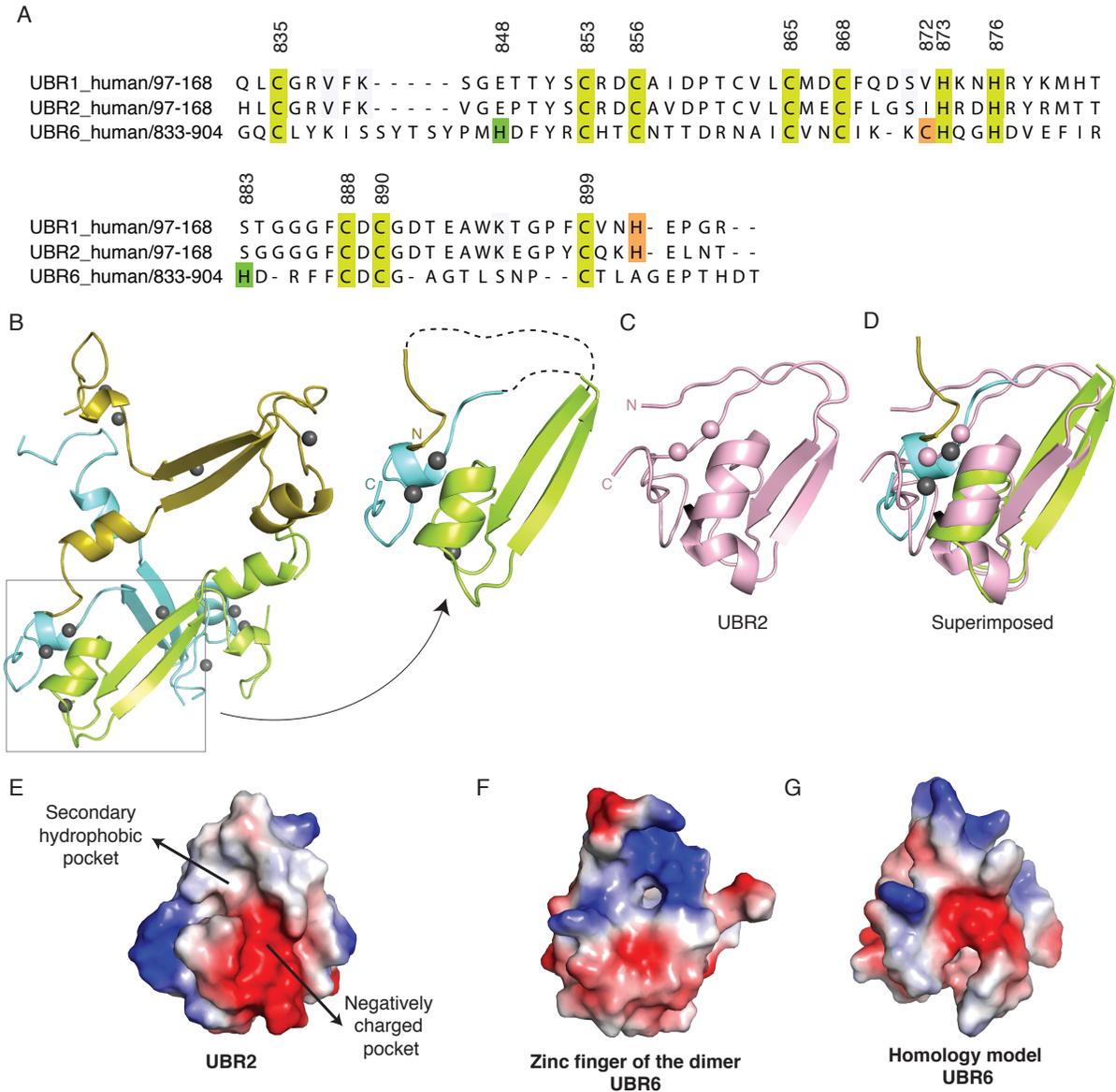


Figure 3.2. Dimer crystal structure emulates the monomeric UBR-box fold.

A. Sequence alignment of the UBR-box domains from human UBR1, UBR2 and UBR6 highlights the conservation of zinc-coordinating residues. The C-terminal histidine residue in UBR1 and UBR2 is not conserved in UBR6. Instead, Cys872 completes a tetrahedral coordination site. The histidine residues coordinating the central zinc ion in UBR6 are not conserved in the UBR family (*green*). **B.** Three protein chains in the asymmetric unit complete the zinc coordination topology that resembles the UBR-box fold. **C.** Crystal structure of the UBR-box from UBR2

(PDB: 3NY3). **D.** Superimposed structures of UBR2 and the zinc finger dimer of UBR6. **E.** Electrostatics potential surface representation of UBR2 (PDB: 3NY3) shows N-degron binding pockets. **F.** Monomer model of UBR6. Zinc finger of the dimer (panel B, right). Absence of pockets could explain the lack of N-degron binding in UBR6. **G.** Homology model shows lack of interacting pockets. Structure model was generated based on human and yeast UBR-box structures using SWISS-MODEL.

One question that remains is why the UBR-box in UBR6 does not bind N-degrons. We were unable to detect binding of peptides with N-terminal arginine residues to the UBR-box of UBR6 by nuclear magnetic resonance. In UBR1 and UBR2, N-degron recognition occurs through a negatively charged pocket that recognizes the positive N-terminal residue, and a secondary hydrophobic pocket, which binds the second residue of the substrate (**Figure 3.2 E**). An electrostatic potential surface representation of the UBR6 fragment containing the conserved zinc fingers lacks any grooves in the area where N-degron binding is expected (**Figure 3.2 F**). A homology model generated by the SWISS-MODEL server based on the available UBR-box structures (human and yeast) (**Figure 3.2 G**) also shows the absence of grooves that would support N-degron binding.

A physiological structure of the UBR-box domain will elucidate the role of the UBR-box that do not bind N-degrons and will clarify the inability of the domain to interact with N-degrons. Amongst the possibilities for UBR-box function are those of zinc finger proteins. Tertiary structured zinc fingers confer specific binding activities to various molecules beyond proteins, such as DNA, RNA and small molecules. It remains to be discovered if the UBR-box is a protein-protein interaction domain or if it serves other macromolecules.

Table 3.1. Collection and refinement statistics for crystallographic data.

UBR6 (833-904)	
Data collection	
Wavelength	1.2822
Resolution range	41.72 - 2.202 (2.281 - 2.202)
Space group	P 21 21 21
Unit cell	67.57 69.2 82.59 90 90 90
Unique reflections	20183 (1952)
Redundancy	4.3 (4.1)
Completeness (%)	99.81 (99)
Mean I/sigma(I)	14.67 (2.88)
Wilson B-factor	36.59
R-meas	0.094 (0.608)
CC1/2	0.851
Refinement	
Reflections used in refinement	20172 (1949)
Reflections used for R-free	1010 (98)
R-work	0.1948 (0.2940)
R-free	0.2474 (0.3414)
Number of non-hydrogen atoms	2377
macromolecules	2210
ligands	50
Protein residues	117
RMS(bonds)	0.005
RMS(angles)	0.69
Ramachandran favored (%)	98.87
Ramachandran allowed (%)	1.13
Ramachandran outliers (%)	0
Rotamer outliers (%)	0.41
Clashscore	5.54
Average B-factor	50.03
macromolecules	49.69
ligands	67.83
solvent	48.78
Number of TLS groups	13

3.4 Materials and Methods

3.4.1 Protein expression and purification

Human UBR6 (833-904) and H848A, H883A and H848/883A mutants were cloned into pGEX-6p-1 vector and expressed in BL21(DE3) *E. coli* cells. Cultures were grown at 37 °C until O.D. at 600 nm reached ~0.7. At this point, 100 µM ZnCl₂ was added and protein was induced with 0.5 mM IPTG. Cultures were grown for 20 h at 16 °C. Lysis of pellets was done in 50 mM HEPES pH 7.5, 200 mM NaCl, 0.1 mM PMSF, 10 µM ZnCl₂, 5% glycerol and 5 mM β-mercaptoethanol buffer. GST-tagged protein was purified using glutathione S-transferase (GST.) sepharose beads and eluted with 20 mM reduced glutathione. GST tag was cleaved using PreScission Protease[®] from GE life sciences. Proteins were further purified by size exclusion chromatography using a HiLoad Superdex 75 16/600 column and 20 mM HEPES pH 7.5, 100 mM NaCl, 2 mM β-mercaptoethanol and 10 µM ZnCl₂ buffer. Gel filtration experiments for UBR6 mutants were performed in a Superdex 75 10/300 GL column.

3.4.2 SEC-MALS

Multi-angle light scattering coupled with size exclusion chromatography experiments were done in with the Wyatt miniDawn TREOS and Optilab rex units using a Superose[™] 6 10/300 GL column. Data analysis was performed with the ASTRA software.

3.4.3 Crystallization, data collection and structure determination

Crystals of 10 mg/ml UBR6 (monomer and dimer fractions) were grown at 22 °C by sitting drop vapor diffusion against 0.2 M NaNO₃, 0.1 M Bis-Tris propane pH 6.5 and 25% PEG 3350. Crystals were cryoprotected with 20% ethylene glycol and flash cooled with liquid nitrogen. X-ray diffraction data sets were collected at λ 1.28 at the Canadian Light Source facility beam line 08B1-1 using a RAYONIX MX300HE detector. Data was processed using HKL2000 (Otwinowski, Z. & Minor, W. 1997). Protein structure was determined using AutoSol program from PHENIX (Adams et al. 2010). Model was finalized manually using COOT (Emsley et al. 2010). Structure refinement was done with CCP4 (Winn et al. 2011) and PHENIX (Adams et al.

2010) suites. Translation/libration/screw (TLS) vibration motion was applied at the last stage of refinement (Painter & Merritt 2006). Refinement statistics are given in Table 1. Figures were produced using PyMol (Schrodinger 2015). The UBR6 homology model was generated by SWISS-MODEL server (Biasini et al. 2014). The coordinates and structure factors of the UBR-box domain from UBR6 have been deposited in the Protein Data Bank (PDB) with accession code 5VMD.

3.5 Acknowledgements

X-ray diffraction data was acquired at the Canadian Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research. This study was funded by Canadian Institutes of Health Research (CIHR) grant MOP-14219.

To understand other substrate recruitment mechanisms in the UBR family I focused on structural studies of UBR5, an N-recognin with N-end rule-independent roles. This E3 ligase did not have identified substrate-binding domains other than the UBR-box domain. However, there are not confirmed N-end rule substrates for UBR5. In this work, I identified MLLE as a substrate-binding domain, which in addition regulates multiple protein-protein interactions within UBR5 and with E3-independent binding partners. This work exposed a modulatory mechanism for ubiquitylation of translation effectors based on an intramolecular interaction involving the HECT and MLLE domains.

CHAPTER FOUR: THE MLLE DOMAIN OF THE UBIQUITIN LIGASE UBR5 BINDS TO ITS CATALYTIC DOMAIN TO REGULATE SUBSTRATE BINDING

Juliana Muñoz-Escobar[‡], Edna Matta-Camacho[‡], Guennadi Kozlov and Kalle Gehring. *J Biol Chem.* 11; 290(37): 22841-50 (2015)

[‡]Equal contributions

4.1 Summary

E3 ubiquitin ligases catalyze the transfer of ubiquitin from an E2 conjugating enzyme to a substrate. UBR5, a HECT-type E3 ligase, mediates the ubiquitination of proteins involved in translation regulation, DNA damage response and gluconeogenesis. In addition, UBR5 functions in a ligase-independent manner by prompting protein-protein interactions without ubiquitination of the binding partner. Despite recent functional studies, the mechanisms involved in substrate recognition and selective ubiquitination of its binding partners remain elusive. The C-terminus of UBR5 harbors the HECT catalytic domain and an adjacent MLLE domain. MLLE domains mediate protein-protein interactions through the binding of a conserved peptide motif, termed PAM2. Here, we characterize the binding properties of UBR5 MLLE domain to PAM2 peptides from Paip1 and GW182. The crystal structure with a Paip1 PAM2 peptide reveals the network of hydrophobic and ionic interactions that drive binding. In addition, we identify a novel interaction of the MLLE domain with the adjacent HECT domain mediated by a PAM2-like sequence. Our results confirm the role of the MLLE domain of UBR5 in substrate recruitment and suggest a potential role in regulating UBR5 ligase activity.

4.2 Introduction

Ubiquitination is one of the most abundant post-translational modifications in eukaryotic cells. Catalyzed by the ubiquitin proteasome system, ubiquitination has two major roles: regulation of protein degradation, essential for normal cellular function and for the removal of harmful, damaged or misfolded proteins, and control of protein activity by regulating protein–protein interactions and subcellular localization (Pickart 2001; Varshavsky 2012). The ubiquitin proteasome system targets proteins through the addition of one or more ubiquitin molecules to specific lysine residues or to the N-terminus. This process is carried out by a complex cascade of reactions catalyzed by activating (E1), conjugating (E2), and ligating enzymes (Ciechanover, A. & Ben-Saadon 2004; Pickart 2001). The E3 ubiquitin ligases mediate the specificity towards substrates and catalyze the final attachment of the 76-residue ubiquitin moiety to the target protein. E3 enzymes fall into two categories based on their catalytic mechanism: RING (Really Interesting New Gene) and U-box ligases promote ubiquitin transfer indirectly while RBR (RING between RING) and HECT (homologous to E6-AP carboxyl terminus) type ubiquitin ligases directly catalyze the transfer of ubiquitin to the substrate. In this latter category, the ubiquitin is transferred from the E2 conjugating enzyme to the substrate in a two-step reaction. In the first step, a catalytic cysteine in the E3 enzyme forms a thioester bond with the ubiquitin from the E2-Ub intermediate. In the final step, ubiquitin is transferred from the thioester bond with the E3 to a lysine residue in the substrate (Scheffner & Kumar 2014).

Ubiquitin protein ligase E3 component N-recognin 5 (UBR5) also known as EDD (E3 isolated by differential display) is a mammalian ortholog of the HYD (hyperplastic discs) protein of *Drosophila melanogaster* (Callaghan et al. 1998; Mansfield, E. et al. 1994). UBR5 belongs to the HECT-type group of E3 ubiquitin ligases. Human UBR5 mediates ubiquitination of several proteins, including β -catenin, TopBP1, TERT, RORyt, Paip2, CDK9, ATMIN, among others, highlighting its role as an important effector in cell cycle progression and DNA damage response (Cojocararu et al. 2011; Hay-Koren et al. 2011; Honda et al. 2002; Jung et al. 2013; Otwinowski, Zbyszek & Minor, Wlodek 1997; Rutz et al. 2015; Wang, X et al. 2013; Yoshida et al. 2006; Zhang, T et al. 2014). UBR5 has also been suggested to be a tumor suppressor. Overexpressed or mutated UBR5 has been found in solid tumors including ovarian, breast, hepatocellular, tongue, gastric and melanoma (Bradley et al. 2014; Fuja et al. 2004; Meissner et al. 2013; Ohshima et al.

2007). In addition, UBR5 exhibits E3-independent activity as a transcriptional cofactor for the progesterone receptor and serves as a binding partner for a diverse subset of proteins such as GW182, p53, CHK2, TFIIIS and DUBA (Cojocaru et al. 2011; Hay-Koren et al. 2011; Henderson et al. 2006; Ling, S & Lin 2011; Rutz et al. 2015; Smits 2012; Su et al. 2011). Despite accumulating knowledge about UBR5 function, the biochemical roles and exact mechanisms of recognition and ubiquitination by UBR5 are yet to be determined.

UBR5 is a large 309 kDa protein and consists of a N-terminal UBA domain followed by two nuclear localization signals, a zinc finger-like UBR-box, a MLLE domain homologous to the C-terminal domain of poly(A)-binding protein (PABP), and a HECT domain at its C-terminus (**Figure 4.1 A**) (Kozlov et al. 2007; Kozlov et al. 2001; Matta-Camacho et al. 2012). Remarkably, only two proteins in eukaryotic cells contain a MLLE domain, PABP and UBR5. In PABP, MLLE is a protein-protein interaction domain that recognizes effectors of translation initiation that display a conserved peptide motif, PAM2 (PABP-interacting motif 2) (Kozlov et al. 2001). The term MLLE comes from a signature motif *MLLEKITG* in the domain and the abbreviation of *Mademoiselle* in French. Solution and crystal structures of MLLE domains from human UBR5 and various PABPs have shown that the domains consist of a bundle of 4 or 5 α -helices (Deo, Sonenberg & Burley 2001; Kozlov et al. 2001). The PAM2 motif was initially identified in three proteins associated with mRNA translation and protein synthesis: Paip1 (PABP-interacting protein 1), Paip2, and eukaryotic Release Factor 3 (Xie, J, Kozlov & Gehring 2014). A bioinformatic survey highlighted the existence of many other PAM2-containing proteins, which include ataxin-2, Tob1/2, USP10, dNF-X1, TPRD/TTC3 and dMAP 205kDa (Albrecht & Lengauer 2004). The NMR solution and crystal structures of the MLLE domain from human PABP in complex with PAM2 peptides revealed that peptides bind to the most conserved helices α_2 , α_3 and α_5 of MLLE (Kozlov et al. 2004; Kozlov, Menade, et al. 2010; Lim et al. 2006). Recently, GW182 was shown to bind to the PABP MLLE surface largely overlapping with the PAM2-binding site (Jinek et al. 2010; Kozlov, Safaei, et al. 2010).

Accumulating evidence supports the model in which competition between UBR5 and PABP for shared binding partners is linked to translation and gene expression regulation. This has been demonstrated for UBR5-mediated proteasomal degradation of Paip2 upon PABP depletion

(Yoshida et al. 2006) and for the recruitment of GW182 and Tob1/2 by UBR5 to Argonaute-miRNA complexes during gene silencing (Su et al. 2011)

The MLLE domain of UBR5 was first shown to bind to a fragment of Paip1 by GST-pull down assays (Deo, Sonenberg & Burley 2001). The peptide binding properties of the UBR5 MLLE domain were later characterized by our lab using NMR chemical shift mapping and isothermal titration calorimetry (Lim et al. 2006). Despite previous studies, there is no atomic structure for UBR5 MLLE bound to a PAM2 peptide. Moreover, the binding of GW182 to UBR5 in miRNA silencing has not been characterized.

A number of substrates for ubiquitination by UBR5 have been described in the last few years. In numerous cases, the C-terminal fragment of UBR5 that includes the MLLE and HECT domains mediates binding. These observations suggest a role of the MLLE domain in the substrate selectivity of UBR5. For instance, Paip2 is targeted for proteasomal degradation by UBR5. However, it is unclear if this interaction is mediated directly by the MLLE domain. A better understanding of PAM2 recognition by UBR5 should help in the identification of novel physiological partners and provide insight into its ability to regulate ubiquitin and E3-independent activity.

In the present study, we determined the crystal structure of the MLLE domain of UBR5 in complex with the PAM2 peptide from Paip1. The structure explains the overlapping binding specificity of the MLLE domains from UBR5 and PABP. We reveal a novel intra-molecular interaction involving the MLLE domain and the HECT domain of UBR5. This interaction is mapped to the N-terminal lobe in the HECT domain and is mediated by a PAM2-like sequence. Our results suggest a regulatory role of the MLLE domain in the catalytic activity of UBR5 beyond binding of PAM2-containing substrates.

4.3 Results

4.3.1 GW182 interacts with the UBR5 MLLE domain

Human GW182, a core component of the miRNA-induced silencing complex, interacts with PABP via its MLLE domain and this interaction is required for miRNA-mediated deadenylation (Jinek et al. 2010; Kozlov, Safaee, et al. 2010). In a similar fashion, UBR5 was recently suggested to be a key component of the miRNA-silencing pathway with the MLLE domain being essential for its silencing function (Su et al. 2011). UBR5 regulated miRNA-mediated gene silencing in an E3 ligase-independent manner by targeting the GW182 family of Argonaute-miRNA complexes. In this study, UBR5 recruited the translation effectors GW182 and Tob1/2 without prompting their proteasomal degradation. Previous studies have characterized the binding properties of several effectors of translation initiation that interact with the PABP MLLE domain through PAM2 motifs. To understand the ability of UBR5 to bind GW182 we performed a titration of the ^{15}N -labeled UBR5 MLLE domain with GW182 (1380-1401). Addition of the peptide produced large chemical shift changes in a number of amides, indicating specific binding (**Figure 4.1 B**). The titration resulted in fast-intermediate exchange that suggests high micromolar binding affinity. A fit of the chemical shift changes measured a K_d of $175 \pm 35 \mu\text{M}$. Although significantly weaker than the interaction with the PABP MLLE domain ($6 \mu\text{M}$), the chemical shift changes upon GW182 binding are similar to those seen upon binding the PAM2 peptide from Paip1 (**Figure 4.1 C**). The largest chemical shift changes upon GW182 peptide binding were leucine, threonine, lysine, glycine and alanine residues in helices α_2 , α_3 and the C-terminus of helix α_5 (**Figure 4.1 D**). This confirms that GW182 binds UBR5 MLLE domain through its PAM2 motif as seen in other PAM2-containing proteins.

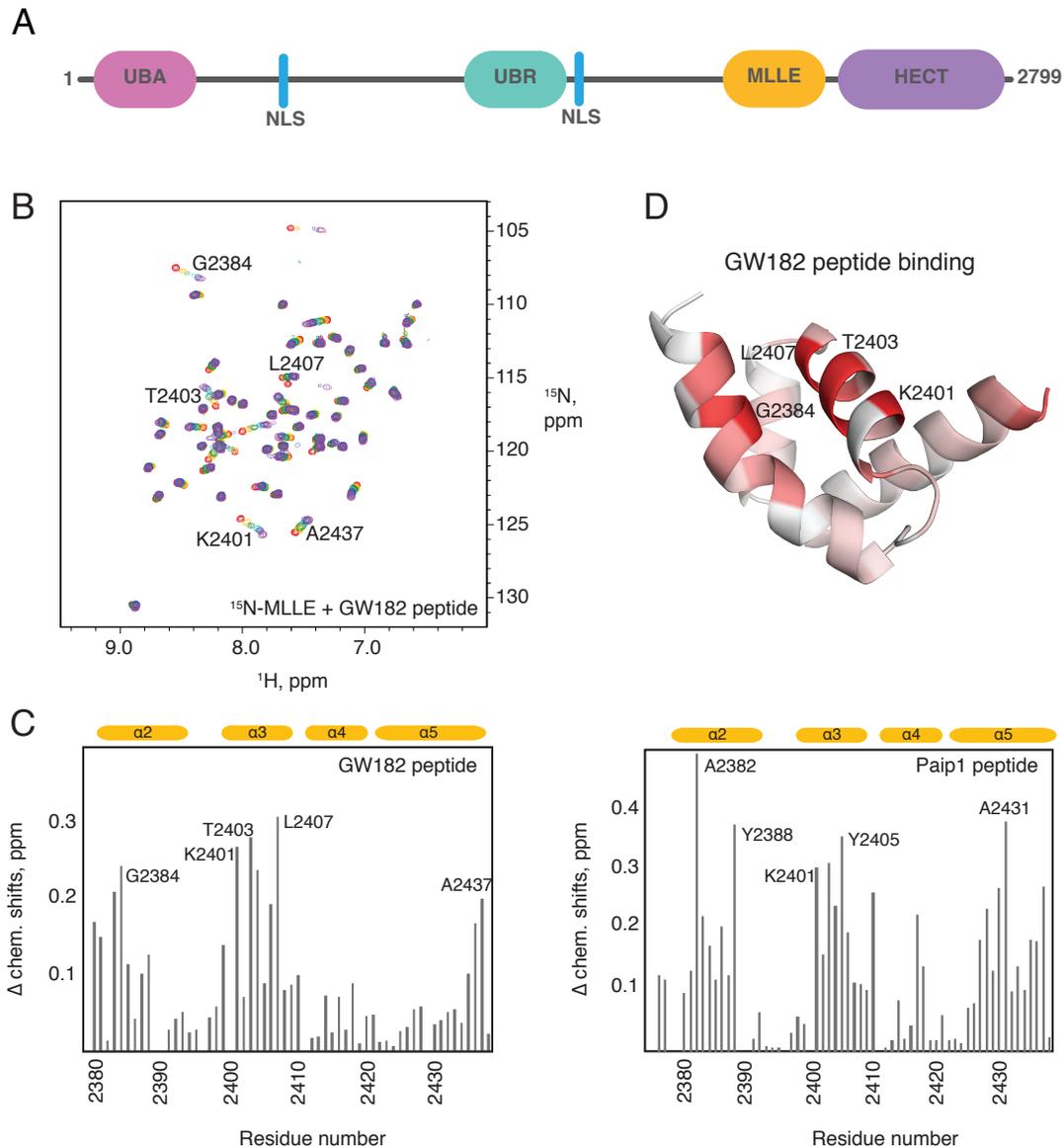


Figure 4.1. The GW182 PAM2-like region interacts with UBR5 MLE.

A. Schematic diagram of the structural domains of UBR5. The catalytic HECT domain is at the C-terminus. UBR5 also contains two nuclear localization signals (NLS) and three protein-protein interaction domains: a ubiquitin-associated domain at its N-terminus, a zinc finger-like UBR domain near the middle of the protein and a domain homologous to the C-terminal domain of poly(A)-binding protein called MLE domain that is adjacent to the HECT domain. **B.** ^{15}N - ^1H NMR correlation spectra of the ^{15}N -labeled UBR5 MLE domain titrated with increasing amounts of the GW182 (1380-1401) peptide, color-coded from red to purple. **C.** Comparison of the chemical shift changes in the ^{15}N -labeled UBR5 MLE domain upon addition of the GW182 peptide (left) and Paip1 peptide (right). Shifts are calculated as a weighted average in ppm as $(\Delta H^2 + (\Delta N/5)^2)^{1/2}$. **D.** Mapping of the NMR chemical shift changes onto a cartoon representation of the unliganded MLE domain (PDB entry 1I2T) upon binding of

GW182 (white, no change; red, maximum change). Helices α_2 , α_3 and α_5 of UBR5 MLLE are involved in peptide binding.

4.3.2 Structure of UBR5 MLLE bound to a PAM2 peptide

To further understand the binding specificity of the MLLE domain from UBR5, we attempted to crystallize the domain in complex with the GW182 peptide. However, no crystals were obtained during crystallization trials. On the other hand, we were able to obtain diffracting crystals for MLLE in complex with the Paip1 peptide. This peptide showed the highest affinity (K_d of 3.4 μM) among those tested in previous isothermal titration calorimetry studies (Lim et al. 2006). The asymmetric unit contains two copies of the MLLE-peptide complex, which are very similar with an RMSD of 0.24 Å over 58 C α atoms. The electron density was missing for 3 and 6 residues at the N- and C-termini of the Paip1 peptide suggesting they are disordered (Table 1).

The structure of the peptide-bound UBR5 MLLE shows a helical bundle with four α -helices folding into a right-handed superhelix. When compared to the structure of unliganded domain from human UBR5 (Deo, Sonenberg & Burley 2001), both structures are very similar, displaying an RMSD of 0.72 Å over residues Gln2381-Ala2437. The only significant difference can be seen in the N-terminal helix, which slightly bends toward the peptide in the complex structure (**Figure 4.2 A**). As the structure of the MLLE domain from PABP contains an additional α -helix at the N-terminus (Kozlov et al. 2001), the helices in the domain from UBR5 are numbered from α_2 to α_5 for easier comparison. In the complex, the Paip1 peptide adopts an extended conformation except for a β -turn at residues Ser129-Ala132 that allows it to wrap around the highly conserved helix α_3 .

Hydrophobic interactions make major contributions to peptide binding to MLLE domains (Kozlov, Safaee, et al. 2010). The side chain of Paip1 Phe135 interacts with C α of Gly2384, the methyl group of Thr2403, and stacks with the side chain of Tyr2388 in a classical “fishbone” stacking arrangement (**Figure 4.2 B**). Next to it, the side chain of Pro137 packs against the aromatic ring of Tyr2388. The side chain of Leu128 inserts into a hydrophobic pocket formed by the side chains of Met2405, Leu2406, Leu2409, Ala2431, Leu2434 and the aliphatic part of Glu2430 (**Figure 4.2 C**). An additional hydrophobic interaction involves Ala132 of Paip1, which

is invariant in PAM2 sequences. The methyl group of Ala132 packs against C α of Met2405, carbonyl of Gly2404 and the C γ of Glu2408 (**Figure 4.2 C**).

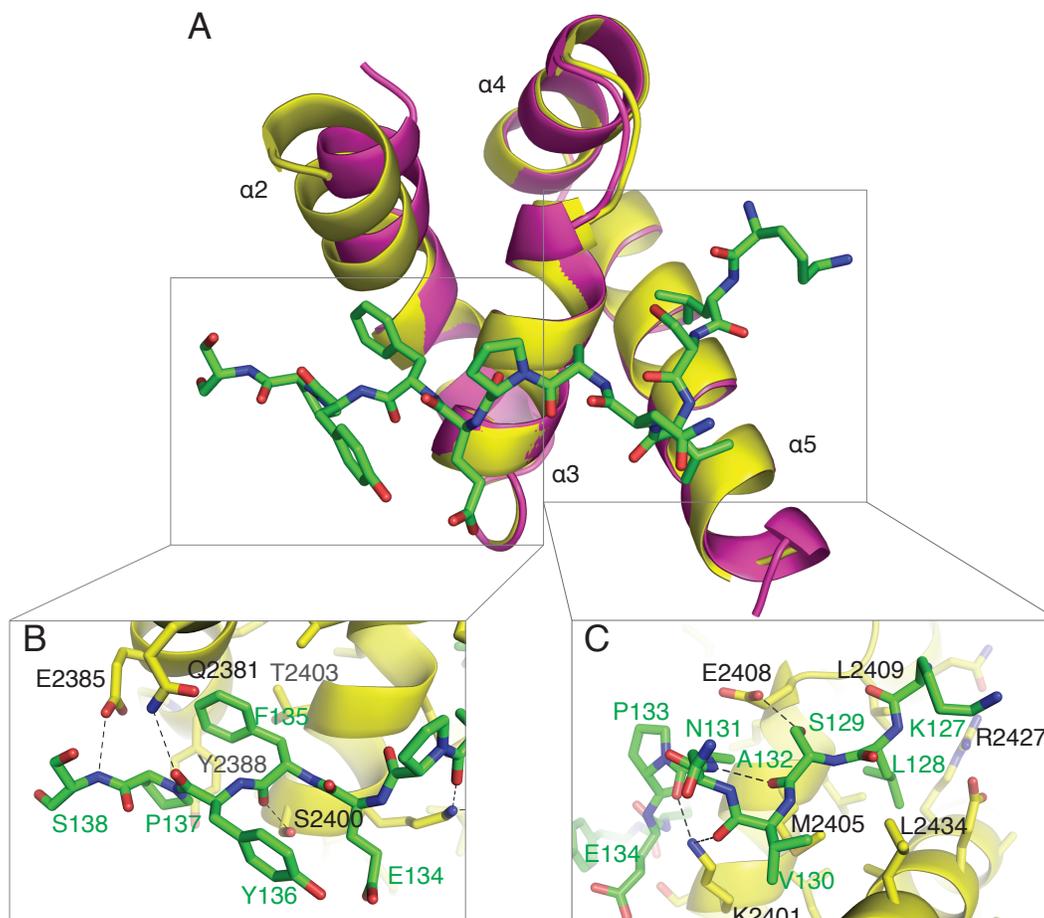


Figure 4.2. Crystal structure of the UBR5 MLE-Paip1 peptide complex.

A. UBR5 MLE domain undergoes minor conformational changes upon binding the PAM2 peptide from Paip1. Ribbon representation of overlaid structures of the liganded (yellow) and unliganded (magenta; PDB entry 1I2T) MLE domain from UBR5. The Paip1 peptide residues are shown as sticks in green. **B.** Close-up of the side chains of Gln2381 and Glu2385 of MLE shows intermolecular hydrogen bonds with carbonyl of Pro137 and amide of Ser138 of Paip1. The aromatic ring of Paip1 Phe135 stacks with the side chain of Tyr2388. **C.** The side chain of Lys2401 of MLE forms intermolecular hydrogen bonds with carbonyls of Val130 and Ala132 of Paip1. The hydrophobic side chain of Leu128 plays a key role binding the MLE domain.

The peptide binding is reinforced by ionic interactions with the UBR5 MLLE domain. The carbonyls of Val130 and Ala132 form hydrogen bonds with the side chain of Lys2401 (**Figure 4.2 C**). The amide of Phe135 forms a hydrogen bond with the carbonyl and the side chain of Ser2400 (**Figure 4.2 B**), which also makes hydrogen bonds with the carbonyl of Phe135. The carbonyl of Tyr136 makes a hydrogen bond with the side chain of Gln2381. The side chain of Glu2385 makes a salt bridge with side chain and amide of Ser138 (**Figure 4.2 B**). The side chain of Ser129 makes a salt bridge with the side chain of Glu2408. Carbonyl of this serine makes an intramolecular hydrogen bond with the amide of Ala132, which stabilizes the bound conformation of the peptide (**Figure 4.2 C**).

4.3.3 UBR5 binds Paip2

Conservation in the binding properties of the MLLE domains from PABP and UBR5 suggests that the E3 ubiquitin ligase activity of UBR5 may play a role in translation. For instance, UBR5 targets the translation inhibitor Paip2 for ubiquitination and proteasomal degradation when PABP is depleted (Yoshida et al. 2006). To confirm this is due to a direct interaction, we tested binding of Paip2 to full-length UBR5 and the UBR5 MLLE-HECT fragment (**Figure 4.1 A**). We performed a series of pull-down assays using GST-fused full-length Paip2 as bait for binding to the full-length UBR5 (**Figure 4.3 A**), and GST-MLLE or GST-MLLE-HECT fragments of UBR5 as bait for Paip2 binding (**Figure 4.3 B**). In all cases binding of Paip2 to either the full-length UBR5 or the MLLE-containing fragments was observed. The presence of a phenylalanine residue in PAM2 motifs is conserved throughout PAM2-containing proteins and required for their interactions with MLLE domains (Kozlov et al. 2004). We tested if the Phe118 of Paip2 was required for the interaction with UBR5 in our binding assays. The Paip2 F118A mutation abrogated binding to both full-length UBR5 and its MLLE domain confirming that the interaction was direct and specific to the MLLE domain of UBR5 (**Figure 4.3 A, B**).

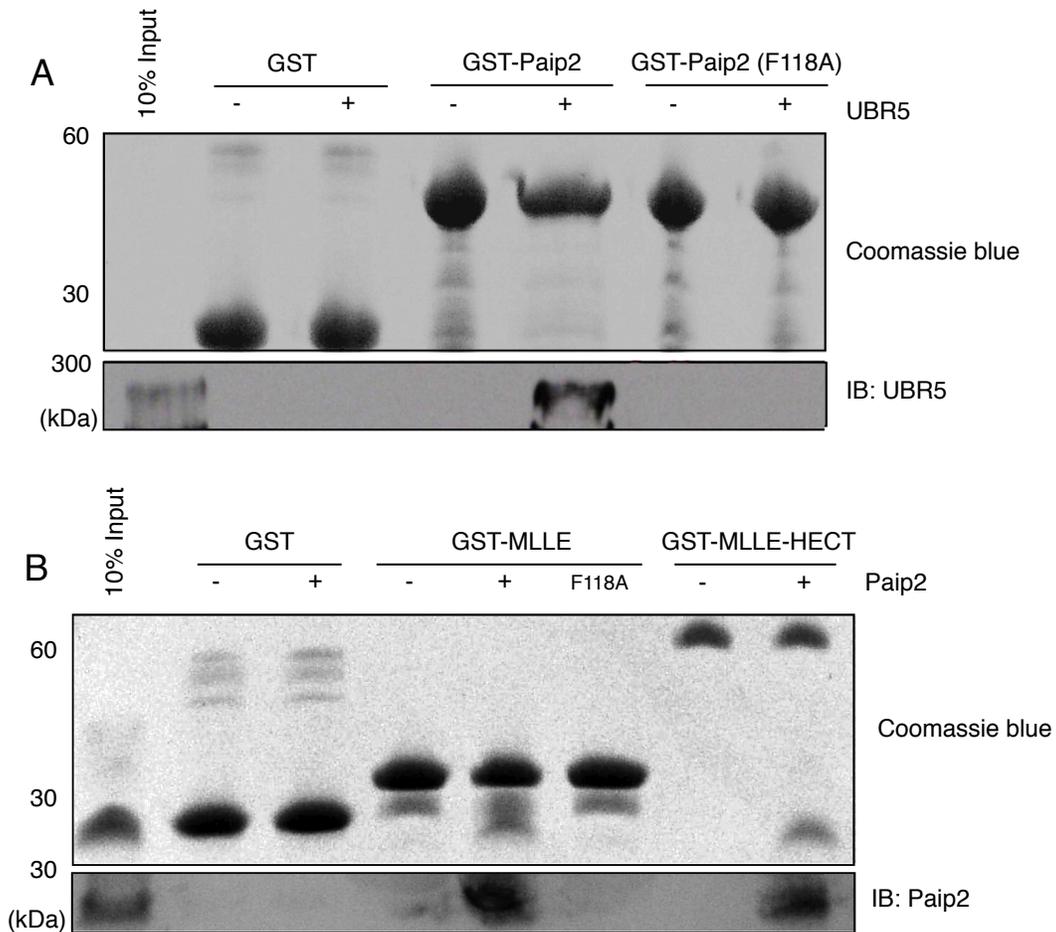


Figure 4.3. Binding of Paip2 to UBR5.

A. Binding of wild-type GST-Paip2 to full-length human UBR5. A mutation of a key phenylalanine residue in Paip2 prevents binding. **B.** Binding of wild-type Paip2 or its F118A mutant to GST-MLLE or GST-MLLE-HECT fragments of UBR5.

4.3.4 MLLE interacts with the HECT domain of UBR5

The ability of UBR5 to regulate its activity throughout the many pathways it is involved in remains elusive. In E3 ligases, sequences or domains located in proximity to the HECT domain often are involved in intra- and/or intermolecular interactions that modulate the catalytic activity (Gallagher et al. 2006; Mari et al. 2014; Scheffner & Kumar 2014; Wiesner et al. 2007). The MLLE domain in UBR5 is located to the N-terminal side of the catalytic HECT domain with a

~50-residue separation. Thus, we asked whether the MLLE domain might interact with the HECT domain. To test this, we performed an NMR titration of ¹⁵N-labeled MLLE (**Figure 4.4 A**) with unlabeled GST-fused HECT domain (residues 2520-2799). Stepwise addition of the GST-HECT domain resulted in severe line broadening and the loss of most of the peaks in the NMR spectrum (**Figure 4.4 B**) suggesting formation of a high molecular weight complex. As controls, titrations of MLLE with GST, with the UBA domain (residues 180-230) or with the UBR-box (residues 1177-1244) of UBR5 showed no spectral changes, indicating no binding (data not shown). An additional control with the MLLE domain of PABP showed that HECT binding was limited to the MLLE domain of UBR5 (data not shown). Together, these data demonstrate that the MLLE domain of UBR5 specifically binds to its HECT domain.

The HECT domain of E3 ligases consists of a bilobal structure with a C-terminal lobe containing the catalytic cysteine residue and an N-terminal lobe that binds the E2 enzyme. The lobes are linked by a flexible region, which presumably facilitates proper positioning of the catalytic cysteine towards the ubiquitin-E2 thioester bond (Scheffner & Kumar 2014). Our next question involved the characterization of the MLLE:HECT interaction and, in particular, how the N- and C-terminal lobes of the HECT domain were involved. We expressed and purified independently the N-lobe (residues 2520-2662) and C-lobe (residues 2687-2799) of UBR5. Addition of the C-lobe fragment to ¹⁵N-labeled MLLE produced no spectral changes indicating no binding (**Figure 4.4 C**). Conversely, the addition of the N-lobe fragment to ¹⁵N-labeled MLLE produced strong line broadening similar to that observed with the HECT domain (**Figure 4.4 D**). Next we tested if the MLLE:N-lobe interaction required the peptide-binding surface of the MLLE domain by adding a PAM2 peptide to the complex of MLLE and N-lobe domains. If the MLLE:N-lobe interaction was dependent on the PAM2-recognizing surface of MLLE then the peptide would compete with the N-lobe for MLLE and the NMR signals would reappear, which is what we observed. Addition of a peptide corresponding to residues 106-127 of Paip2 resulted in reappearance of signals for MLLE at the positions consistent with MLLE binding the PAM2 peptide (**Figure 4.4 E**). Isothermal titration calorimetry measurements of the MLLE:N-lobe interaction measured a K_d of $50 \pm 2.0 \mu\text{M}$ (**Figure 4.4 F**). These results show that the MLLE domain from UBR5 interacts with the HECT domain in a PAM2-dependent manner.

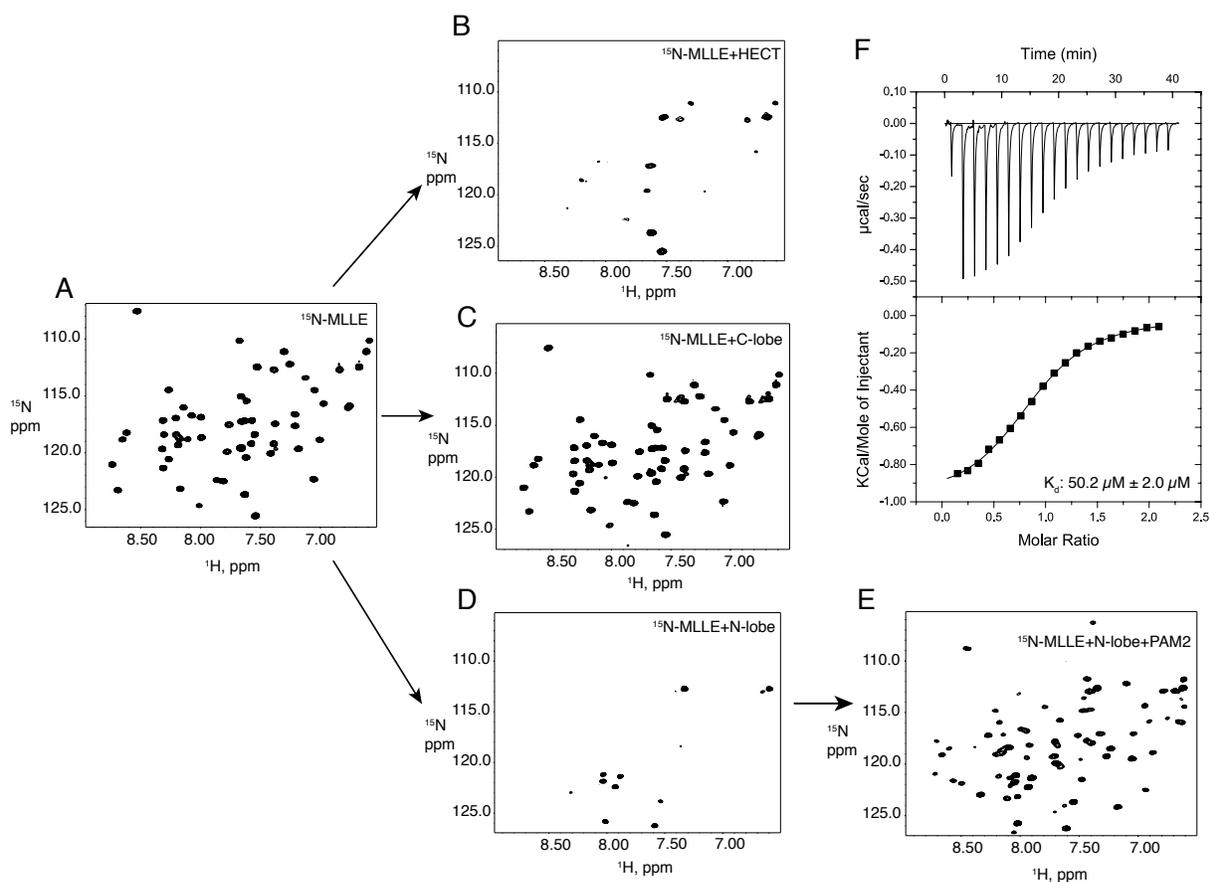


Figure 4.4 The UBR5 MLE domain interacts with the HECT domain.

^1H - ^{15}N correlation NMR spectra of ^{15}N -MLE domain show that the isolated domain forms a higher molecular weight complex in the presence of the HECT domain. **A.** NMR spectrum of 0.15 mM ^{15}N -MLE alone. **B.** NMR spectrum upon addition of 0.4 mM HECT domain. The fast transverse relaxation of magnetization due to the ^{15}N -MLE:HECT interaction leads to loss of most of the MLE NMR signals. **C.** NMR spectrum after addition of 0.5 mM C-lobe domain shows no interaction. **D.** NMR spectrum after addition of 0.5 mM N-lobe shows an interaction. **E.** The addition of 1.1 mM Paip2 (106-127) peptide leads to reappearance of the MLE signals due to displacement of the N-lobe and the lower molecular weight of the MLE:Paip2 peptide complex. **F.** ITC experiment for the binding of the N-terminal lobe of the HECT domain to the MLE domain. The upper curve shows the baseline-corrected thermogram, and the lower graph shows the integrated areas of the heat of binding along with a fit, from which the stoichiometry (N) is 0.900 ± 0.007 sites, the molar association constant (K) is $(1.99 \pm 0.08) \times 10^4 \text{ M}^{-1}$, enthalpy (ΔH) is $-1003 \pm 11 \text{ cal/mol}$ and entropy (ΔS) is 16.3 cal/mol/deg .

4.3.4 The N-terminal lobe of the HECT domain contains a PAM2-like sequence

The alignment of the PAM2 sequences from the proteins known to bind PABP and/or UBR5 to the HECT domain revealed the presence of a sequence in the N-terminal lobe with features of a PAM2 motif (**Figure 4.5 A**). The Phe2505 in the N-lobe of UBR5 can be aligned with the conserved phenylalanine of other PAM2-containing proteins. Conserved asparagine and alanine residues are also present in the N-lobe. We designed a peptide bearing the HECT PAM2-like sequence (residues 2499-2517) and performed 2D NMR titrations using ^{15}N -labeled MLLE from UBR5. Titration of the ^{15}N -labeled MLLE with the HECT peptide produced large amide chemical shift changes in fast-intermediate exchange (**Figure 4.5 B**). The largest chemical shift changes occur in helices $\alpha 2$, $\alpha 3$ and $\alpha 5$ (**Figure 4.5 C**). The changes upon HECT peptide binding are similar to those seen upon binding of GW182 and Paip1 (**Figure 4.1 C**). A fit of the chemical shift changes measured a K_d of $850 \pm 55 \mu\text{M}$. The weak binding affinity of the peptide suggests that additional intra-molecular contacts between HECT N-lobe and MLLE stabilize the association between the intact protein domains. As a control, we tested a second peptide with a mutation in Phe2505. Upon addition of the F2505A peptide, there were almost no changes in the spectrum, indicating abrogation of the binding between MLLE and the mutant peptide (**Figure 4.5 D**).

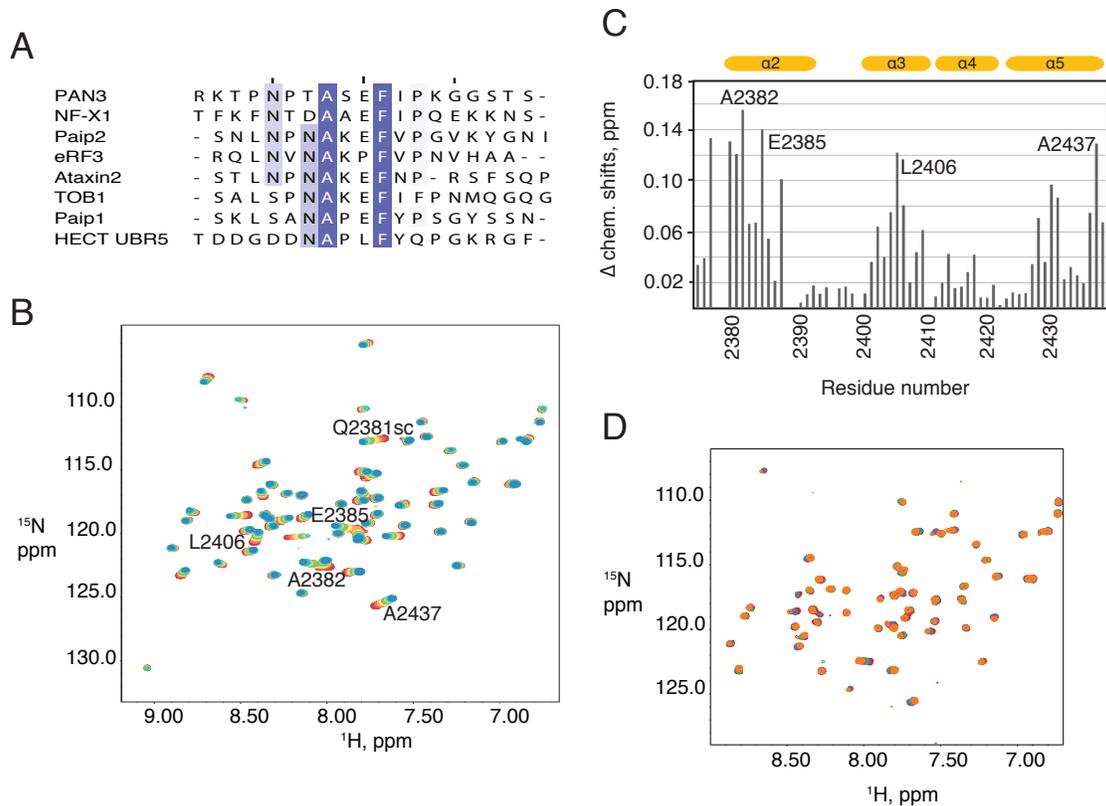


Figure 4.5. The UBR5 MLLE domain recognizes a PAM2-like peptide from the UBR5 HECT domain.

A. Sequence alignment of known PAM2 motifs and a PAM2-like sequence in the N-terminal lobe of the HECT domain. **B.** ^1H - ^{15}N correlation NMR spectra ^{15}N -MLLE domain of UBR5 titrated with increasing concentrations of the HECT peptide. Red is the spectrum of ^{15}N -MLLE alone and blue is the spectrum with the highest concentration of peptide. **C.** Chemical shift changes in the ^{15}N -labeled UBR5 MLLE domain upon addition of HECT peptide. The largest chemical shift changes were by R2380, A2382, E2385 and Y2388 positioned in helix α_2 , L2406 in helix α_3 and A2437 in helix α_5 . **D.** NMR spectra titrated with increasing concentrations of the mutant HECT peptide (F2505A) show no significant changes.

Table 4.1. Collection and refinement statistics for crystallographic data.

UBR5 MLLE-Paip1 (123-144)	
Data collection	
Space group	P6 ₄ 22
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	95.79, 95.79, 82.94
Resolution (Å)	50-2.60 (2.69-2.60)
<i>R</i> _{sym}	0.054 (0.343)
<i>I</i> / <i>sI</i>	24.2 (6.6)
Completeness (%)	99.9 (100.0)
Redundancy	10.1 (8.9)
Refinement	
Resolution (Å)	41.5-2.60
No. reflections	6999
<i>R</i> _{work} / <i>R</i> _{free}	0.227/0.289
No. atoms	1110
MLLE	902
Peptide	200
Water	8
<i>B</i> -factors	
MLLE	16.4
Peptide	24.0
Water	48.3
R.m.s deviations	
Bond lengths (Å)	0.018
Bond angles (°)	2.06
Ramachandran statistics (%)	
Most favored regions	93.4
Additional allowed regions	6.6
WHAT IF structure Z-scores²	
1 st generation packing quality	0.9
2 nd generation packing quality	1.3
Ramachandran plot appearance	0.5
Chi-1/chi-2 rotamer normality	-1.7
Backbone conformation	1.4
WHAT IF RMS Z-scores³	
Bond lengths	0.796
Bond angles	0.954
Omega angle restrains	1.007
Side chain planarity	0.709
Improper dihedral distribution	1.075
B-factor distribution	0.509
Inside/outside distribution	1.008

4.4 Discussion

The specificity of the ubiquitination process relies on the E3 ubiquitin ligases and their ability to directly interact with substrates. Over 600 different E3s have been identified in the human genome and 28 belong to the HECT-type E3 family. In all cases, the HECT domain is located at the carboxyl terminus of the protein and the substrate binding is mediated by various domains located N-terminal to the HECT domain (Scheffner & Kumar 2014). The activity by HECT E3s can be regulated at two levels. In the first level, substrate binding is mediated through protein-protein interactions by domains/motifs located N-terminal to the HECT domain. Some HECT proteins also interact with regulatory/auxiliary proteins that facilitate or interfere with substrate binding (Ichimura et al. 2005; Shea et al. 2012; Shearwin-Whyatt et al. 2006). In the second level, regulation occurs through intra- and/or intermolecular interactions that inhibit Ub-thioester formation or E2 binding (Gallagher et al. 2006; Mari et al. 2014; Pandya et al. 2010; Wiesner et al. 2007). Despite accumulating functional knowledge about the regulatory mechanisms that govern E3 ligases, our structural understanding of the inter- and intra-molecular interactions that modulate the catalytic activity of HECT-type enzymes has lagged. The HECT-type ligases in the Nedd4 family are the most studied to date. In SMURF2, the C2 domain interacts with the HECT domain rendering the full-length protein inactive. The N-terminal lobe of the HECT domain interacts with the C2 domain and with ubiquitin. Both interacting surfaces overlap, affecting transthiolation and non-covalent binding of ubiquitin to the N-lobe (Mari et al. 2014). In the case of Itch, the autoinhibitory mechanism involves an intra-molecular interaction between the WW domains and the HECT domain. Phosphorylation of the PRR regions of Itch causes a conformational change that weakens the WW:HECT interaction increasing its catalytic activity (Gallagher et al. 2006). A similar regulatory mechanism is seen in the non-Nedd4 HECT-type ligase HUWE1. An N-terminal helical element was shown to affect the catalytic activity of the HECT domain in HUWE1. In absence of this N-terminal helix, the isolated HECT domain gained activity relative to the helix-extended counterpart; the authors hypothesize that this could be due to an increase in the inner flexibility of the HECT domain that allows the enzyme to acquire a favourable orientation for ubiquitin transfer or product release (Pandya et al. 2010).

In the case of UBR5, we have identified an intra-molecular interaction between the HECT domain and the adjacent MLLE domain. This interaction has the potential to regulate the

catalytic activity of HECT in a manner similar to that seen in other E3 ligases. We measured the affinity of the interaction between the isolated domains to be 50 μM , which is relatively strong considering that, in the intact UBR5 protein, the two domains are separated by only 50 amino acids (**Figure 4.3 A**). Previous phosphoproteomic studies have reported UBR5 to be heavily phosphorylated (Bethard et al. 2011). It is possible that specific phosphorylation sites in the protein lead to conformational changes that regulate ubiquitin activity. The MLLE-HECT interaction might regulate HECT domain activity by preventing proper E2 binding or positioning of the C-lobe to receive the ubiquitin. In parallel, the MLLE domain also acts as a substrate-binding domain so that substrate binding might be correlated with activation of ligase activity.

UBR5 plays an essential role in cellular process such as DNA damage response, translation initiation and cell cycle progression. However, the mechanistic details of how UBR5 interacts with substrates are poorly understood. To date, the only PAM2-containing protein identified as a substrate for ubiquitination and proteasomal degradation by UBR5 is Paip2. Yoshida, *et al.* proposed a homeostatic mechanism where PABP and UBR5 compete for binding to Paip2, an inhibitor of PABP function. A decrease of PABP levels augments the concentration of Paip2 that is available to interact with UBR5, leading to Paip2 proteasomal degradation. As Paip2 levels decrease, the relative amount of PABP increases and the overall activity of PABP is restored in a positive feedback. In contrast, UBR5 plays an essential role in microRNA-mediated gene silencing independent of its ubiquitin ligase activity (Su et al. 2011). To date, there are two suggested roles of the MLLE domain in miRNA silencing. First, GW182 proteins recruit UBR5 into Ago-miRNA complexes through its MLLE domain. Second, UBR5 MLLE interacts with PAM2-containing proteins in a similar fashion to PABP thus sharing binding partners such as Paip1/2 and Tob1/2. Through protein interactions with these proteins, the extended protein network includes different deadenylase complexes, all of which play key roles in regulating translation and mRNA stability. In the present study, we characterized the binding of the UBR5 MLLE domain to the GW182 PAM2 peptide and determined the crystal structure of the MLLE-Paip1 complex. Comparison of the MLLE domains of UBR5 and PABP shows that the major intermolecular interactions that mediate peptide binding are preserved in both proteins. However, in general, the affinity of the UBR5 MLLE domain for PAM2 peptides is lower than that of the PABP MLLE domain. The complex with the GW182 peptide is no exception. The GW182

PAM2 peptide binds to UBR5 MLLE with approximately 30-fold lower affinity than to the PABP MLLE domain (Kozlov, Safaee, et al. 2010). This likely reflects the unique C-terminal sequence of the GW182 PAM2 motif, which contains a tryptophan residue that inserts between the helices $\alpha 2$ and $\alpha 3$ of PABP MLLE (Jinek et al. 2010; Lim et al. 2006). The biological significance of the wide-range of PAM2 affinities measured *in vitro* is unclear. It would be interesting to investigate the functional significance of the differences in affinity for GW182 in the Ago-miRNA complex formation.

The binding of the PAM2 peptides to UBR5 shows surprising contrasts in function. PAM2 motifs from GW182 and Paip2 have the ability to bind the MLLE domains from both UBR5 and PABP although with a higher affinity for the latter. However, the proteins interact with UBR5 for different purposes. Paip2's fate is to be targeted for proteasomal degradation whereas GW182 promotes gene silencing. In contrast, the interaction of the PAM2 peptide from the HECT domain of UBR5 suggests a role in regulating UBR5 activity. Despite the fact that all of these interactions involve recognition of PAM2-like sequences, each of them seems to have a unique effect in the response of UBR5. It remains to be discovered if the differences in affinity among PAM2 proteins are essential in determining the role of UBR5 or if other events are key in controlling the different activities of UBR5.

In conclusion, we have characterized the PAM2 peptide binding to the MLLE domain of UBR5 by X-ray crystallography and NMR spectroscopy. Future functional and structural studies are required to address the role of the newly discovered MLLE-HECT interaction in the E3 ligase activity of UBR5.

4.5 Materials and Methods

4.5.1 Protein expression, purification and peptide synthesis

Human Paip2 protein and the MLLE, HECT and MLLE-HECT domains of rat UBR5 were cloned into BamHI and XhoI restriction sites of the pGEX-6P-1 vector (Amersham Pharmacia Biotech, Piscataway, NJ), and the construct transformed into the *E. coli* expression host BL21 Gold Magic (DE3) (Stratagene). The proteins were expressed and purified by affinity chromatography to yield a GST-fused domain or an isolated domain with a five-residue (Gly-Pro-Leu-Gly-Ser) N-terminal extension. Prior to crystallization, the MLLE protein was additionally purified using size-exclusion chromatography in gel-filtration buffer (50 mM Tris, 100 mM NaCl, pH 7.5). The final yield of purified protein was ~7 mg per liter of Luria broth culture media.

A plasmid coding for the full-length human UBR5 was kindly donated by Dr. Darren N. Saunders (Garvan Institute of Medical Research) and the protein expressed in HEK293 cells as a His-tag fusion protein.

The Paip1 (123-144), Paip2 (106-127), and GW182 (1380-1401) peptides were synthesized by Fmoc solid-phase peptide synthesis and purified by reverse phase chromatography on a C18 column (Vydac, Hesperia, CA). The composition and purity of the peptides were verified by electrospray ionization mass spectroscopy. The HECT peptide and its F2505A mutant were expressed as GST-fused proteins in *E. coli*, purified with affinity chromatography and cleaved with preScission protease leaving a five-residue (Gly-Pro-Leu-Gly-Ser) N-terminal extension. Peptides were further purified by reverse phase chromatography. Western blot analyses were done using anti-UBR5 and anti-Paip2 antibodies (Sigma-Aldrich).

4.5.2 Isothermal titration calorimetry measurements

Experiments were carried out on a MicroCal iTC200 titration calorimeter in 50 mM Tris-HCl buffer (pH 7.6) and 150 mM NaCl at 20 °C. The reaction cell contained 200 μ l of

0.1 mM HECT-N lobe and was titrated with 19 injections of 2 μ l of 1.0 mM MLLE domain. The binding isotherm was fit with a binding model employing a single set of independent sites to determine the thermodynamic binding constants and stoichiometry.

4.5.3 Crystallization

Crystallization conditions for the UBR5 MLLE/Paip1 (123-144) complex were identified utilizing hanging drop vapor diffusion with the JCSG+ crystallization suite (QIAGEN). The best crystals were obtained by equilibrating a 1.0 μ l drop of MLLE-Paip1 (123-144) mixture in 1:2 ratio (10 mg/ml) in 50 mM Tris-HCl (pH 7.5), 0.1 M NaCl, mixed with 1.0 μ l of reservoir solution containing 1.0 M ammonium sulfate, 0.2 M lithium sulfate, 10% glycerol and 0.1 M Tris-HCl (pH 8.5). Crystals grew in 3-10 days at 20°C. The crystals contain two MLLE and two peptide molecules in the asymmetric unit corresponding to $V_m = 2.89 \text{ \AA}^3 \text{ Da}^{-1}$ and a solvent content of 57.4%. Residue numbers used here and in the PDB deposition are 14 smaller than in UniProt entry Q62671.

4.5.4 Structure solution and refinement

Diffraction data from a single crystal of MLLE-peptide complex were collected on an ADSC Quantum-210 CCD detector (Area Detector Systems Corp.) at beamline A1 at the Cornell High-Energy Synchrotron Source (CHESS) (Table 1). Data processing and scaling were performed with HKL2000 (Otwinowski, Zbyszek & Minor, Wladek 1997). The structure of UBR5 MLLE/Paip1 was determined by molecular replacement with Phaser (McCoy et al. 2007), using the coordinates of MLLE from human UBR5 (PDB entry 1I2T). The initial model obtained from Phaser was completed and adjusted with the program Xfit (McRee 1999) and was improved by several cycles of refinement, using the program REFMAC 5.2 (Murshudov et al. 1999) and model refitting. At the latest stage of refinement, we also applied the translation-libration-screw (TLS) option (Winn, Murshudov & Papiz 2003). The final model has good stereochemistry according to the program PROCHECK (Laskowski et al. 1993). The refinement statistics are given in Table 4.1 (Vriend 1990). The coordinates and structure factors have been deposited in the RCSB Protein Data Bank (accession number 3NTW).

4.5.5 NMR spectroscopy

NMR assignments of the MLLE domain of rat UBR5 were described earlier (Lim et al. 2006). All NMR experiments were recorded at 298 K. NMR titrations were carried by adding unlabeled either protein or peptide into 0.15 mM sample of the ^{15}N -labeled MLLE domain and monitored by ^{15}N - ^1H heteronuclear single quantum correlation spectra. NMR spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed with XEASY (Bartels et al. 1995).

4.5.6 Pull-down assays

Assays were performed in 1X PBS buffer with 0.1% Tween 20 (buffer A). 0.2 mg of GST and GST-fused proteins were incubated in 25 μl of Glutathione Sepharose for 15 min. Supernatant was removed and beads were washed twice with buffer A. 0.05 mg of bait proteins (Paip2 or UBR5) were incubated on the beads for 1h. Supernatant was removed and beads were washed twice with buffer A. 1X loading dye was added to each sample to run SDS-PAGE.

4.6 Acknowledgements

Data acquisition at the Macromolecular Diffraction (MacCHESS) facility at the Cornell High Energy Synchrotron Source (CHESS) was supported by the National Science Foundation award DMR 0225180 and the National Institutes of Health award RR-01646. This study was funded by the Canadian Institutes of Health Research grant MOP-14219 to K.G. and an NSERC CREATE Training Program in Bionanomachines award to J.M. The atomic coordinates and structure factors (code: 3NTW) have been deposited in the Protein Data Bank.

Homology models of the UBR-box from UBR4 and UBR5 (**Figure 5.2**) suggest that N-degron specificity might be significantly different for these UBR family members. In UBR1 and UBR2, the secondary pocket is essential for high affinity binding and specificity for hydrophobic residues. In UBR5, the secondary pocket is predicted to be positively charged and should favour interactions with negatively charged second residues. Speculatively, one might hypothesize that UBR4 and UBR5 might be specialized ligases for secondary N-degrons generated by arginylation of aspartic acid, glutamic acid, or oxidized cysteine N-terminal residues.

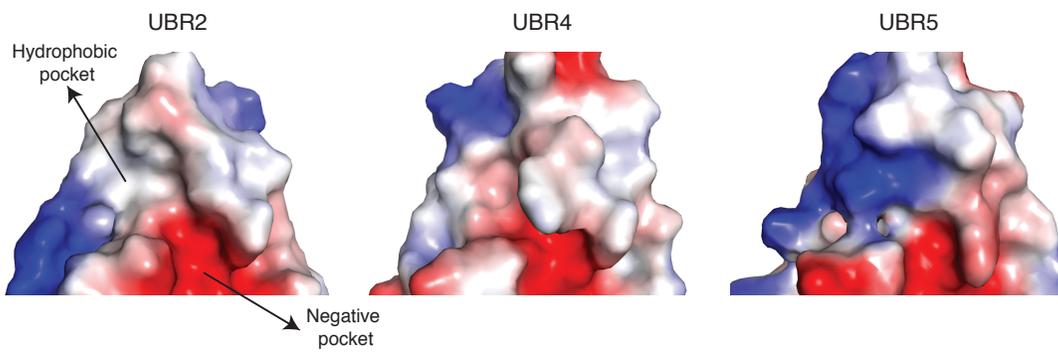


Figure 5.2. Secondary pocket in the UBR-box domains.

Comparison of the UBR-box domain in UBR2 (crystal structure), UBR4 and UBR5 (homology models). The hydrophobic or secondary pockets could be a potential source of specificity towards the second position in the N-degron. The UBR-box in UBR5 could have a positive character compared to the hydrophobic groove in UBR1 and UBR2.

5.2 UBR-box domains that do not bind N-degrons

Initially associated with only type 1 N-degron recognition, the UBR-box domain appears as a malleable scaffold for protein-protein interactions with other potential roles. In UBR family members that do not bind N-degrons, the function of the UBR-box remains unknown. In the case of UBR6, experiments I performed in the group of Dr. Michele Pagano point at UBR6 localization in the nucleus with association to chromatin. There are a great number of possible roles for the UBR-box in UBR6. Zinc-binding domains (such as RING domains in E3 ligases) have the ability to dimerize. Some CRLs use

dimerization as a way to increase processivity or to inhibit catalysis. My experiments in the Pagano lab using a mammalian cell system failed to demonstrate any full-length dimerization mediated by the UBR-box domain. SCF E3 ligases function in a multi-subunit complex that brings together different substrate adaptors for ubiquitylation. In UBR6, the putative CASH domain was shown to serve as a targeting-domain for a few of the substrates (Duan et al. 2012). It is not uncommon for an E3 ligase to display multiple substrate binding domains or to utilize auxiliary proteins to diversify and increase specificity. Given its importance for BCL6 misregulation, structural work on UBR6 activation upon interaction with substrates will likely provide a therapeutic framework for B-cell lymphoma.

An alternative role for the UBR-box is suggested from studies on UBR3. The caspase-generated fragment of DIAP1 that bears N-terminal asparagine binds UBR3 through the UBR-box. Mutation of this N-terminal residue to methionine abrogates binding (Huang, Q et al. 2014). This interaction is E3-independent as the RING domain is dispensable. The UBR3 UBR-box domain also interacts with other proteins such as Svb, Pch15 and Sans. In the case of Svb targeting, small peptides simultaneously bind to UBR3 to activate substrate binding (Zanet et al. 2015). This mechanism is similar to that observed in UBR1 and UBR2 with Cup9 (Xia, Turner, et al. 2008). This interesting interplay further supports a protein-protein interaction role of the UBR-box beyond the N-end rule. Structural and biophysical studies would highlight the molecular determinants for a possible N-terminal and internal degron recognition in this UBR-box domain.

One of the biggest challenges in structural biology research is finding stable protein fragments that are prone to crystallize or that are suitable for *in vitro* biophysical experiments. The UBR family is no exception. Multiple trials to crystallize the UBR-box domains of all members only led to success for UBR1 and UBR2. Given the abundance of secondary structure in almost all UBR proteins according to bioinformatics analysis, a promising approach for future structural work will be to study longer fragments around the UBR-box domains.

5.3 Substrate recognition by the UBR-box

In the N-end rule pathway, the two most studied N-recognins are UBR1 and UBR2. Examination of the binding capabilities of the domain demonstrated that the UBR-box interacts with more than the canonical N-terminal residue. I showed that the domains bind methylated N-terminal arginine and lysine residues with high affinity, which opens the door for identification of additional substrates. Given the prevalence of arginine and lysine methylation in eukaryotic cells, particularly in histones and transcriptional regulation, there is a possibility that methylated N-degrons can be targets of the N-end rule. There are two possible mechanisms for methylated N-degron generation. First, an N-terminal methyltransferase could catalyze methylation of the N-terminal arginine or lysine side chain. Second, an internal methylated arginine or lysine could become exposed upon cleavage by a protease. To date, there are no reports of N-terminal methyltransferases that modify the side chain instead of the α amino group. Similarly, cleavage by proteases at methylated arginine and lysine residues is uncommon, although it does occur. Some bacterial proteases have shown specificity for cleavage at methylation sites but there are no reports for mammalian enzymes with similar specificity (Huesgen et al. 2015). The ability of methylated arginine and lysine to bind the UBR-box also represents an opportunity to explore the structure-activity relationships in the design of more specific inhibitors. The locations of water molecules in the negative pocket provide an interesting framework for rational drug design.

5.3.1 Proline has a dual role in the N-end rule.

An interesting role of proline in the N-end rule emerged with the latest publications that identified it as a destabilizing N-terminal residue. Proline in the second position of a type 1 N-degron abolishes binding to the UBR-box. In contrast, Pro2 in phosphoenolpyruvate carboxykinase (PEPCK1) induces its degradation by the Pro/N-end rule through Gid4 (Chen, SJ et al. 2017). Moreover, PEPCK1 is also a substrate of UBR5. However, it is recognized by an internal degron. To our knowledge, proline is the only residue in the N-end rule that displays a dual role both promoting and blocking E3 binding.

5.3.2 UBR-box and N-domain, a cooperative interaction?

Despite the lack of structural information on the N-domain, multiple studies show interplay with the UBR-box in terms of substrate targeting (Hwang, CS, Shemorry & Varshavsky 2009; Tasaki et al. 2009; Xia, Turner, et al. 2008). The binding of small peptides to the UBR-box is thought to induce a conformational change that exposes the N-domain for type 2 binding. Similarly, this binding also activates internal degron recognition by the third binding site in UBR1 (**Figure 1.7**). This hypothesis is supported by the inhibitory effect of p-chloroamphetamine on type 1 and type 2 recognition (Jiang, Y et al. 2014). Binding of the compound to the N-domain reduces significantly type 1 binding (UBR-box) but the mechanism is poorly understood.

A structure of the UBR-box domain together with the N-domain would elucidate this inhibitory/activation mechanism and greatly aid the design of more potent inhibitors that exploit the proposed conformational change. While the sequence similarity with the Clp domain identifies a ~80-residue region for the N-domain, the presence of abundant, predicted secondary structure in the vicinity suggests a bigger structure. Future structural studies should focus on the ~400-residue fragments of UBR1 and UBR2 that include the UBR-box domain.

5.4 Internal degron recognition.

Given the potential ability of the UBR-box in UBR3 to bind proteins through internal degrons, it remains to be seen if this ability is also present in the UBR-box of other UBR-family proteins. In the case of UBR5, some targets of ubiquitylation depend on the presence of the UBR-box for binding without involving N-terminal recognition (Henderson et al. 2006; Munoz et al. 2007).

5.5 UBR5 as a HECT- and CRL-type ligase.

UBR5 is perhaps one of the most versatile proteins in the family. Numerous and distinct substrate recognition mechanisms have been reported and continue to be explored. Unfortunately, very little is known about its catalytic mechanism and regulation.

Recent reports of UBR5 activity suggest that it associates with DDB1, VprBP and DYRK2 to form an E3 ligase complex (Hossain et al. 2017; Jung et al. 2013). In the complex, VprBP acts as the substrate-binding site, while ubiquitylation occurs through the HECT domain of UBR5. The association of UBR5 with the DDB1 complex would possibly expand its repertoire of substrates and diversify its regulation. The precise mechanism of ubiquitylation by this complex is not understood but it is suggested that phosphorylation by DYRK2 is important for recognition.

As mentioned in Chapter One, UBR5 is a large protein and the structure of only ~15% is known. The studies here provide the first mechanistic view of a regulatory mechanism. The interaction of MLLE with the HECT domain in UBR5 is potentially autoinhibitory as seen in the regulation of other E3 ligases. One of the main questions that remain is how UBR5 is able to regulate its E3-ligase activity in addition to its E3-independent roles (Cojocaru et al. 2011; Ling, S & Lin 2011; McDonald et al. 2010; Su et al. 2011). Given the size of the full length protein (300 kDa), multiple domains and regions are potentially involved in modulating the intricate network of interactions and outcomes that UBR5 seems to control in the cell. A look at its structure and its association with these complexes using electron microscopy could be a valuable approach to understand UBR5 regulation and catalytic activity.

5.6 Concluding remarks

This thesis work provides a comprehensive study of type 1 N-degron recognition by UBR1 and UBR2, expands the scope of binding partners for the UBR-box and proposes a modulatory mechanism for ubiquitylation in a HECT-type ligase. These findings will

hopefully serve as a platform for further investigation on UBR ligases regulation and its relationship with cellular homeostasis and disease.

REFERENCES

- Abbas, T, Mueller, AC, Shibata, E, Keaton, M, Rossi, M & Dutta, A 2013, 'CRL1-FBXO11 promotes Cdt2 ubiquitylation and degradation and regulates Pr-Set7/Set8-mediated cellular migration', *Mol Cell*, vol. 49, no. 6, pp. 1147-58.
- Abida, WM, Nikolaev, A, Zhao, W, Zhang, W & Gu, W 2007, 'FBXO11 promotes the Neddylation of p53 and inhibits its transcriptional activity', *J Biol Chem*, vol. 282, no. 3, pp. 1797-804.
- Acconcia, F, Sigismund, S & Polo, S 2009, 'Ubiquitin in trafficking: the network at work', *Exp Cell Res*, vol. 315, no. 9, pp. 1610-8.
- Adams, PD, Afonine, PV, Bunkoczi, G, Chen, VB, Davis, IW, Echols, N, Headd, JJ, Hung, LW, Kapral, GJ, Grosse-Kunstleve, RW, McCoy, AJ, Moriarty, NW, Oeffner, R, Read, RJ, Richardson, DC, Richardson, JS, Terwilliger, TC & Zwart, PH 2010, 'PHENIX: a comprehensive Python-based system for macromolecular structure solution', *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 2, pp. 213-21.
- Agarwalla, P & Banerjee, R 2016, 'N-end rule pathway inhibition assists colon tumor regression via necroptosis', *Mol Ther Oncolytics*, vol. 3, p. 16020.
- Alagramam, K, Naider, F & Becker, JM 1995, 'A recognition component of the ubiquitin system is required for peptide transport in *Saccharomyces cerevisiae*', *Mol Microbiol*, vol. 15, no. 2, pp. 225-34.
- Albrecht, M & Lengauer, T 2004, 'Survey on the PABC recognition motif PAM2', *Biochem Biophys Res Commun*, vol. 316, no. 1, pp. 129-38.
- An, JY, Kim, E, Zakrzewska, A, Yoo, YD, Jang, JM, Han, DH, Lee, MJ, Seo, JW, Lee, YJ, Kim, TY, de Rooij, DG, Kim, BY & Kwon, YT 2012, 'UBR2 of the N-end rule

pathway is required for chromosome stability via histone ubiquitylation in spermatocytes and somatic cells', *PLoS One*, vol. 7, no. 5, p. e37414.

An, JY, Kim, EA, Jiang, Y, Zakrzewska, A, Kim, DE, Lee, MJ, Mook-Jung, I, Zhang, Y & Kwon, YT 2010, 'UBR2 mediates transcriptional silencing during spermatogenesis via histone ubiquitination', *Proc Natl Acad Sci U S A*, vol. 107, no. 5, pp. 1912-7.

An, JY, Seo, JW, Tasaki, T, Lee, MJ, Varshavsky, A & Kwon, YT 2006, 'Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway', *Proc Natl Acad Sci U S A*, vol. 103, no. 16, pp. 6212-7.

Andreson, BL, Gupta, A, Georgieva, BP & Rothstein, R 2010, 'The ribonucleotide reductase inhibitor, Sml1, is sequentially phosphorylated, ubiquitylated and degraded in response to DNA damage', *Nucleic Acids Res*, vol. 38, no. 19, pp. 6490-501.

Aras, S, Singh, G, Johnston, K, Foster, T & Aiyar, A 2009, 'Zinc coordination is required for and regulates transcription activation by Epstein-Barr nuclear antigen 1', *PLoS Pathog*, vol. 5, no. 6, p. e1000469.

Ashton-Beaucage, D, Lemieux, C, Udell, CM, Sahmi, M, Rochette, S & Therrien, M 2016, 'The Deubiquitinase USP47 Stabilizes MAPK by Counteracting the Function of the N-end Rule ligase POE/UBR4 in *Drosophila*', *PLoS Biol*, vol. 14, no. 8, p. e1002539.

Atik, T, Karakoyun, M, Sukalo, M, Zenker, M, Ozkinay, F & Aydoğdu, S 2015, 'Two novel UBR1 gene mutations in a patient with Johanson Blizzard Syndrome: A mild phenotype without mental retardation', *Gene*, vol. 570, no. 1, pp. 153-5.

Attali, I, Tobelaim, WS, Persaud, A, Motamedchaboki, K, Simpson-Lavy, KJ, Mashahreh, B, Levin-Kravets, O, Keren-Kaplan, T, Pilzer, I, Kupiec, M, Wiener, R,

Wolf, DA, Rotin, D & Prag, G 2017, 'Ubiquitylation-dependent oligomerization regulates activity of Nedd4 ligases', *EMBO J*, vol. 36, no. 4, pp. 425-40.

Bachmair, A, Finley, D & Varshavsky, A 1986, 'In vivo half-life of a protein is a function of its amino-terminal residue', *Science*, vol. 234, no. 4773, pp. 179-86.

Bai, C, Sen, P, Hofmann, K, Ma, L, Goebel, M, Harper, JW & Elledge, SJ 1996, 'SKP1 Connects Cell Cycle Regulators to the Ubiquitin Proteolysis Machinery through a Novel Motif, the F-Box', *Cell*, vol. 86, no. 2, pp. 263-74.

Baker, RT & Varshavsky, A 1991, 'Inhibition of the N-end rule pathway in living cells', *Proc Natl Acad Sci U S A*, vol. 88, no. 4, pp. 1090-4.

Banaszak, K, Martin-Diaconescu, V, Bellucci, M, Zambelli, B, Rypniewski, W, Maroney, MJ & Ciurli, S 2012, 'Crystallographic and X-ray absorption spectroscopic characterization of *Helicobacter pylori* UreE bound to Ni(2)(+) and Zn(2)(+) reveals a role for the disordered C-terminal arm in metal trafficking', *Biochem J*, vol. 441, no. 3, pp. 1017-26.

Barbash, O, Lee, EK & Diehl, JA 2011, 'Phosphorylation-dependent regulation of SCF(Fbx4) dimerization and activity involves a novel component, 14-3-3varepsilon', *Oncogene*, vol. 30, no. 17, pp. 1995-2002.

Bartel, B, Wunning, I & Varshavsky, A 1990, 'The recognition component of the N-end rule pathway', *EMBO J*, vol. 9, no. 10, pp. 3179-89.

Bartels, C, Xia, TH, Billeter, M, Guntert, P & Wuthrich, K 1995, 'The program XEASY for computer-supported NMR spectral analysis of biological macromolecules', *J Biomol NMR*, vol. 6, no. 1, pp. 1-10.

Beaver, JE & Waters, ML 2016, 'Molecular Recognition of Lys and Arg Methylation', *ACS Chem Biol*, vol. 11, no. 3, pp. 643-53.

Behnia, R, Panic, B, Whyte, JR & Munro, S 2004, 'Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p', *Nat Cell Biol*, vol. 6, no. 5, pp. 405-13.

Belzil, C, Ramos, T, Sanada, K, Colicos, MA & Nguyen, MD 2014, 'p600 stabilizes microtubules to prevent the aggregation of CaMKIIalpha during photoconductive stimulation', *Cell Mol Biol Lett*, vol. 19, no. 3, pp. 381-92.

Benavides, M, Chow-Tsang, LF, Zhang, J & Zhong, H 2013, 'The novel interaction between microspherule protein Msp58 and ubiquitin E3 ligase EDD regulates cell cycle progression', *Biochim Biophys Acta*, vol. 1833, no. 1, pp. 21-32.

Besche, HC, Haas, W, Gygi, SP & Goldberg, AL 2009, 'Isolation of mammalian 26S proteasomes and p97/VCP complexes using the ubiquitin-like domain from HHR23B reveals novel proteasome-associated proteins', *Biochemistry*, vol. 48, no. 11, pp. 2538-49.

Bethard, JR, Zheng, H, Roberts, L & Eblen, ST 2011, 'Identification of phosphorylation sites on the E3 ubiquitin ligase UBR5/EDD', *J Proteomics*, vol. 75, no. 2, pp. 603-9.

Bhalla, V, Daidie, D, Li, H, Pao, AC, LaGrange, LP, Wang, J, Vandewalle, A, Stockand, JD, Staub, O & Pearce, D 2005, 'Serum- and glucocorticoid-regulated kinase 1 regulates ubiquitin ligase neural precursor cell-expressed, developmentally down-regulated protein 4-2 by inducing interaction with 14-3-3', *Mol Endocrinol*, vol. 19, no. 12, pp. 3073-84.

Biasini, M, Bienert, S, Waterhouse, A, Arnold, K, Studer, G, Schmidt, T, Kiefer, F, Gallo Cassarino, T, Bertoni, M, Bordoli, L & Schwede, T 2014, 'SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information', *Nucleic Acids Res*, vol. 42, no. Web Server issue, pp. W252-8.

Boriack-Sjodin, PA & Swinger, KK 2016, 'Protein Methyltransferases: A Distinct, Diverse, and Dynamic Family of Enzymes', *Biochemistry*, vol. 55, no. 11, pp. 1557-69.

Bosu, DR & Kipreos, ET 2008, 'Cullin-RING ubiquitin ligases: global regulation and activation cycles', *Cell Div*, vol. 3, p. 7.

Bradley, A, Zheng, H, Ziebarth, A, Sakati, W, Branham-O'Connor, M, Blumer, JB, Liu, Y, Kistner-Griffin, E, Rodriguez-Aguayo, C, Lopez-Berestein, G, Sood, AK, Landen, CN, Jr. & Eblen, ST 2014, 'EDD enhances cell survival and cisplatin resistance and is a therapeutic target for epithelial ovarian cancer', *Carcinogenesis*, vol. 35, no. 5, pp. 1100-9.

Breiten, B, Lockett, MR, Sherman, W, Fujita, S, Al-Sayah, M, Lange, H, Bowers, CM, Heroux, A, Krilov, G & Whitesides, GM 2013, 'Water networks contribute to enthalpy/entropy compensation in protein-ligand binding', *J Am Chem Soc*, vol. 135, no. 41, pp. 15579-84.

Brower, CS, Piatkov, KI & Varshavsky, A 2013, 'Neurodegeneration-associated protein fragments as short-lived substrates of the N-end rule pathway', *Mol Cell*, vol. 50, no. 2, pp. 161-71.

Brzovic, PS, Rajagopal, P, Hoyt, DW, King, MC & Klevit, RE 2001, 'Structure of a BRCA1-BARD1 heterodimeric RING-RING complex', *Nat Struct Biol*, vol. 8, no. 10, pp. 833-7.

Bulatov, E & Ciulli, A 2015, 'Targeting Cullin-RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation', *Biochem J*, vol. 467, no. 3, pp. 365-86.

Byrd, C, Turner, GC & Varshavsky, A 1998, 'The N-end rule pathway controls the import of peptides through degradation of a transcriptional repressor', *EMBO J*, vol. 17, no. 1, pp. 269-77.

Callaghan, MJ, Russell, AJ, Woollatt, E, Sutherland, GR, Sutherland, RL & Watts, CK 1998, 'Identification of a human HECT family protein with homology to the *Drosophila* tumor suppressor gene hyperplastic discs', *Oncogene*, vol. 17, no. 26, pp. 3479-91.

Chen, SJ, Wu, X, Wadas, B, Oh, JH & Varshavsky, A 2017, 'An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes', *Science*, vol. 355, no. 6323.

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*, vol. 66, no. 3, pp. 345-57 e6.

Chew, EH, Poobalasingam, T, Hawkey, CJ & Hagen, T 2007, 'Characterization of cullin-based E3 ubiquitin ligases in intact mammalian cells--evidence for cullin dimerization', *Cell Signal*, vol. 19, no. 5, pp. 1071-80.

Choi, WS, Jeong, BC, Joo, YJ, Lee, MR, Kim, J, Eck, MJ & Song, HK 2010, 'Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases', *Nat Struct Mol Biol*, vol. 17, no. 10, pp. 1175-81.

Christensen, DE, Brzovic, PS & Klevit, RE 2007, 'E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages', *Nat Struct Mol Biol*, vol. 14, no. 10, pp. 941-8.

Christofferson, DE, Li, Y, Hitomi, J, Zhou, W, Upperman, C, Zhu, H, Gerber, SA, Gygi, S & Yuan, J 2012, 'A novel role for RIP1 kinase in mediating TNF α production', *Cell Death Dis*, vol. 3, p. e320.

Ciccarelli, FD, Copley, RR, Doerks, T, Russell, RB & Bork, P 2002, 'CASH--a beta-helix domain widespread among carbohydrate-binding proteins', *Trends Biochem Sci*, vol. 27, no. 2, pp. 59-62.

Ciechanover, A & Ben-Saadon, R 2004, 'N-terminal ubiquitination: more protein substrates join in', *Trends Cell Biol*, vol. 14, no. 3, pp. 103-6.

Ciechanover, A & Iwai, K 2004, 'The Ubiquitin System: From Basic Mechanisms to the Patient Bed', *IUBMB Life*, vol. 56, no. 4, pp. 193-201.

Clancy, JL, Henderson, MJ, Russell, AJ, Anderson, DW, Bova, RJ, Campbell, IG, Choong, DY, Macdonald, GA, Mann, GJ, Nolan, T, Brady, G, Olopade, OI, Woollatt, E, Davies, MJ, Segara, D, Hacker, NF, Henshall, SM, Sutherland, RL & Watts, CK 2003, 'EDD, the human orthologue of the hyperplastic discs tumour suppressor gene, is amplified and overexpressed in cancer', *Oncogene*, vol. 22, no. 32, pp. 5070-81.

Cojocaru, M, Bouchard, A, Cloutier, P, Cooper, JJ, Varzavand, K, Price, DH & Coulombe, B 2011, 'Transcription factor IIS cooperates with the E3 ligase UBR5 to ubiquitinate the CDK9 subunit of the positive transcription elongation factor B', *J Biol Chem*, vol. 286, no. 7, pp. 5012-22.

Collins, GA & Goldberg, AL 2017, 'The Logic of the 26S Proteasome', *Cell*, vol. 169, no. 5, pp. 792-806.

Conroy, J, McGettigan, P, Murphy, R, Webb, D, Murphy, SM, McCoy, B, Albertyn, C, McCreary, D, McDonagh, C, Walsh, O, Lynch, S & Ennis, S 2014, 'A novel locus for episodic ataxia:UBR4 the likely candidate', *Eur J Hum Genet*, vol. 22, no. 4, pp. 505-10.

Cook, JR, Lee, JH, Yang, ZH, Krause, CD, Herth, N, Hoffmann, R & Pestka, S 2006, 'FBXO11/PRMT9, a new protein arginine methyltransferase, symmetrically dimethylates arginine residues', *Biochem Biophys Res Commun*, vol. 342, no. 2, pp. 472-81.

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*, vol. 6, no. 3, pp. 277-93.

Deo, RC, Sonenberg, N & Burley, SK 2001, 'X-ray structure of the human hyperplastic discs protein: an ortholog of the C-terminal domain of poly(A)-binding protein', *Proc Natl Acad Sci U S A*, vol. 98, no. 8, pp. 4414-9.

Ditzel, M, Wilson, R, Tenev, T, Zachariou, A, Paul, A, Deas, E & Meier, P 2003, 'Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis', *Nat Cell Biol*, vol. 5, no. 5, pp. 467-73.

Dompe, N, Rivers, CS, Li, L, Cordes, S, Schwickart, M, Punnoose, EA, Amler, L, Seshagiri, S, Tang, J, Modrusan, Z & Davis, DP 2011, 'A whole-genome RNAi screen identifies an 8q22 gene cluster that inhibits death receptor-mediated apoptosis', *Proc Natl Acad Sci U S A*, vol. 108, no. 43, pp. E943-51.

Dou, H, Buetow, L, Sibbet, GJ, Cameron, K & Huang, DT 2012, 'BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer', *Nat Struct Mol Biol*, vol. 19, no. 9, pp. 876-83.

Dougan, DA, Micevski, D & Truscott, KN 2012, 'The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+proteases', *Biochim Biophys Acta*, vol. 1823, no. 1, pp. 83-91.

Dougan, DA, Reid, BG, Horwich, AL & Bukau, B 2002, 'ClpS, a Substrate Modulator of the ClpAP Machine', *Molecular Cell*, vol. 9, no. 3, pp. 673-83.

Du, F, Navarro-Garcia, F, Xia, Z, Tasaki, T & Varshavsky, A 2002, 'Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain', *Proc Natl Acad Sci U S A*, vol. 99, no. 22, pp. 14110-5.

Duan, S, Cermak, L, Pagan, JK, Rossi, M, Martinengo, C, di Celle, PF, Chapuy, B, Shipp, M, Chiarle, R & Pagano, M 2012, 'FBXO11 targets BCL6 for degradation and is inactivated in diffuse large B-cell lymphomas', *Nature*, vol. 481, no. 7379, pp. 90-3.

Duda, DM, Olszewski, JL, Schuermann, JP, Kurinov, I, Miller, DJ, Nourse, A, Alpi, AF & Schulman, BA 2013, 'Structure of HHARI, a RING-IBR-RING ubiquitin ligase: autoinhibition of an Ariadne-family E3 and insights into ligation mechanism', *Structure*, vol. 21, no. 6, pp. 1030-41.

Eblen, ST, Kumar, NV, Shah, K, Henderson, MJ, Watts, CK, Shokat, KM & Weber, MJ 2003, 'Identification of novel ERK2 substrates through use of an engineered kinase and ATP analogs', *J Biol Chem*, vol. 278, no. 17, pp. 14926-35.

Eisele, F & Wolf, DH 2008, 'Degradation of misfolded protein in the cytoplasm is mediated by the ubiquitin ligase Ubr1', *FEBS Lett*, vol. 582, no. 30, pp. 4143-6.

Eisenhaber, B, Chumak, N, Eisenhaber, F & Hauser, MT 2007, 'The ring between ring fingers (RBR) protein family', *Genome Biol*, vol. 8, no. 3, p. 209.

Emsley, P, Lohkamp, B, Scott, WG & Cowtan, K 2010, 'Features and development of Coot', *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 4, pp. 486-501.

Erbse, A, Schmidt, R, Bornemann, T, Schneider-Mergener, J, Mogk, A, Zahn, R, Dougan, DA & Bukau, B 2006, 'ClpS is an essential component of the N-end rule pathway in Escherichia coli', *Nature*, vol. 439, no. 7077, pp. 753-6.

Etlinger, JD & Goldberg, AL 1977, 'A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes', *Proc Natl Acad Sci U S A*, vol. 74, no. 1, pp. 54-8.

Fallon, L, Belanger, CM, Corera, AT, Kontogianna, M, Regan-Klapisz, E, Moreau, F, Voortman, J, Haber, M, Rouleau, G, Thorarinsdottir, T, Brice, A, van Bergen En Henegouwen, PM & Fon, EA 2006, 'A regulated interaction with the UIM protein Eps15 implicates parkin in EGF receptor trafficking and PI(3)K-Akt signalling', *Nat Cell Biol*, vol. 8, no. 8, pp. 834-42.

Fielenbach, N, Guardavaccaro, D, Neubert, K, Chan, T, Li, D, Feng, Q, Hutter, H, Pagano, M & Antebi, A 2007, 'DRE-1: an evolutionarily conserved F box protein that regulates *C. elegans* developmental age', *Dev Cell*, vol. 12, no. 3, pp. 443-55.

Frescas, D & Pagano, M 2008, 'Deregulated proteolysis by the F-box proteins SKP2 and beta-TrCP: tipping the scales of cancer', *Nat Rev Cancer*, vol. 8, no. 6, pp. 438-49.

Frottin, F, Martinez, A, Peynot, P, Mitra, S, Holz, RC, Giglione, C & Meinnel, T 2006, 'The proteomics of N-terminal methionine cleavage', *Mol Cell Proteomics*, vol. 5, no. 12, pp. 2336-49.

Fuja, TJ, Lin, F, Osann, KE & Bryant, PJ 2004, 'Somatic mutations and altered expression of the candidate tumor suppressors CSNK1 epsilon, DLG1, and EDD/hHYD in mammary ductal carcinoma', *Cancer Res*, vol. 64, no. 3, pp. 942-51.

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*, vol. 103, no. 6, pp. 1717-22.

Ganoth, D, Bornstein, G, Ko, TK, Larsen, B, Tyers, M, Pagano, M & Hershko, A 2001, 'The cell-cycle regulatory protein Cks1 is required for SCF(Skp2)-mediated ubiquitinylation of p27', *Nat Cell Biol*, vol. 3, no. 3, pp. 321-4.

Gautschi, M, Just, S, Mun, A, Ross, S, Rucknagel, P, Dubaquier, Y, Ehrenhofer-Murray, A & Rospert, S 2003, 'The yeast N(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides', *Mol Cell Biol*, vol. 23, no. 20, pp. 7403-14.

Gibbs, DJ, Bacardit, J, Bachmair, A & Holdsworth, MJ 2014, 'The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions', *Trends Cell Biol*, vol. 24, no. 10, pp. 603-11.

Giglione, C, Boularot, A & Meinnel, T 2004, 'Protein N-terminal methionine excision', *Cell Mol Life Sci*, vol. 61, no. 12, pp. 1455-74.

Glenn, KA, Nelson, RF, Wen, HM, Mallinger, AJ & Paulson, HL 2008, 'Diversity in tissue expression, substrate binding, and SCF complex formation for a lectin family of ubiquitin ligases', *J Biol Chem*, vol. 283, no. 19, pp. 12717-29.

Goldberg, AL, Cascio, P, Saric, T & Rock, KL 2002, 'The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides', *Mol Immunol*, vol. 39, no. 3-4, pp. 147-64.

Gonda, DK, Bachmair, A, Wunning, I, Tobias, JW, Lane, WS & Varshavsky, A 1989, 'Universality and structure of the N-end rule', *J Biol Chem*, vol. 264, no. 28, pp. 16700-12.

Gudjonsson, T, Altmeyer, M, Savic, V, Toledo, L, Dinant, C, Grofte, M, Bartkova, J, Poulsen, M, Oka, Y, Bekker-Jensen, S, Mailand, N, Neumann, B, Heriche, JK, Shearer, R, Saunders, D, Bartek, J, Lukas, J & Lukas, C 2012, 'TRIP12 and UBR5 suppress

spreading of chromatin ubiquitylation at damaged chromosomes', *Cell*, vol. 150, no. 4, pp. 697-709.

Hadjikyriacou, A, Yang, Y, Espejo, A, Bedford, MT & Clarke, SG 2015, 'Unique Features of Human Protein Arginine Methyltransferase 9 (PRMT9) and Its Substrate RNA Splicing Factor SF3B2', *J Biol Chem*, vol. 290, no. 27, pp. 16723-43.

Hamilton, MJ, Lee, M & Le Roch, KG 2014, 'The ubiquitin system: an essential component to unlocking the secrets of malaria parasite biology', *Mol Biosyst*, vol. 10, no. 4, pp. 715-23.

Hammerle, M, Bauer, J, Rose, M, Szallies, A, Thumm, M, Dusterhus, S, Mecke, D, Entian, KD & Wolf, DH 1998, 'Proteins of newly isolated mutants and the amino-terminal proline are essential for ubiquitin-proteasome-catalyzed catabolite degradation of fructose-1,6-bisphosphatase of *Saccharomyces cerevisiae*', *J Biol Chem*, vol. 273, no. 39, pp. 25000-5.

Hammond-Martel, I, Yu, H & Affar el, B 2012, 'Roles of ubiquitin signaling in transcription regulation', *Cell Signal*, vol. 24, no. 2, pp. 410-21.

Han, SO, Kommaddi, RP & Shenoy, SK 2013, 'Distinct roles for beta-arrestin2 and arrestin-domain-containing proteins in beta2 adrenergic receptor trafficking', *EMBO Rep*, vol. 14, no. 2, pp. 164-71.

Hao, B, Oehlmann, S, Sowa, ME, Harper, JW & Pavletich, NP 2007, 'Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases', *Mol Cell*, vol. 26, no. 1, pp. 131-43.

Hao, B, Zheng, N, Schulman, BA, Wu, G, Miller, JJ, Pagano, M & Pavletich, NP 2005, 'Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase', *Mol Cell*, vol. 20, no. 1, pp. 9-19.

Hardisty-Hughes, RE, Tateossian, H, Morse, SA, Romero, MR, Middleton, A, Tymowska-Lalanne, Z, Hunter, AJ, Cheeseman, M & Brown, SD 2006, 'A mutation in the F-box gene, Fbxo11, causes otitis media in the Jeff mouse', *Hum Mol Genet*, vol. 15, no. 22, pp. 3273-9.

Hay-Koren, A, Caspi, M, Zilberberg, A & Rosin-Arbesfeld, R 2011, 'The EDD E3 ubiquitin ligase ubiquitinates and up-regulates beta-catenin', *Mol Biol Cell*, vol. 22, no. 3, pp. 399-411.

Helsens, K, Van Damme, P, Degroeve, S, Martens, L, Arnesen, T, Vandekerckhove, J & Gevaert, K 2011, 'Bioinformatics analysis of a *Saccharomyces cerevisiae* N-terminal proteome provides evidence of alternative translation initiation and post-translational N-terminal acetylation', *J Proteome Res*, vol. 10, no. 8, pp. 3578-89.

Henderson, MJ, Munoz, MA, Saunders, DN, Clancy, JL, Russell, AJ, Williams, B, Pappin, D, Khanna, KK, Jackson, SP, Sutherland, RL & Watts, CK 2006, 'EDD mediates DNA damage-induced activation of CHK2', *J Biol Chem*, vol. 281, no. 52, pp. 39990-40000.

Henderson, MJ, Russell, AJ, Hird, S, Munoz, M, Clancy, JL, Lehrbach, GM, Calanni, ST, Jans, DA, Sutherland, RL & Watts, CK 2002, 'EDD, the human hyperplastic discs protein, has a role in progesterone receptor coactivation and potential involvement in DNA damage response', *J Biol Chem*, vol. 277, no. 29, pp. 26468-78.

Hershko, A & Ciechanover, A 1998, 'The ubiquitin system', *Annu Rev Biochem*, vol. 67, pp. 425-79.

Hershko, A, Heller, H, Elias, S & Ciechanover, A 1983, 'Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown', *J Biol Chem*, vol. 258, no. 13, pp. 8206-14.

Hershko, A, Heller, H, Eytan, E, Kaklij, G & Rose, IA 1984, 'Role of the alpha-amino group of protein in ubiquitin-mediated protein breakdown', *Proc Natl Acad Sci U S A*, vol. 81, no. 22, pp. 7021-5.

Honda, Y, Tojo, M, Matsuzaki, K, Anan, T, Matsumoto, M, Ando, M, Saya, H & Nakao, M 2002, 'Cooperation of HECT-domain ubiquitin ligase hHYD and DNA topoisomerase II-binding protein for DNA damage response', *J Biol Chem*, vol. 277, no. 5, pp. 3599-605.

Hong, JH, Kaustov, L, Coyaud, E, Srikumar, T, Wan, J, Arrowsmith, C & Raught, B 2015, 'KCMF1 (potassium channel modulatory factor 1) Links RAD6 to UBR4 (ubiquitin N-recognition domain-containing E3 ligase 4) and lysosome-mediated degradation', *Mol Cell Proteomics*, vol. 14, no. 3, pp. 674-85.

Horn, M, Geisen, C, Cermak, L, Becker, B, Nakamura, S, Klein, C, Pagano, M & Antebi, A 2014, 'DRE-1/FBXO11-dependent degradation of BLMP-1/BLIMP-1 governs *C. elegans* developmental timing and maturation', *Dev Cell*, vol. 28, no. 6, pp. 697-710.

Hossain, D, Javadi Esfehiani, Y, Das, A & Tsang, WY 2017, 'Cep78 controls centrosome homeostasis by inhibiting EDD-DYRK2-DDB1VprBP', *EMBO Rep*, vol. 18, no. 4, pp. 632-44.

Hu, G, Wang, X, Saunders, DN, Henderson, M, Russell, AJ, Herring, BP & Zhou, J 2010, 'Modulation of myocardin function by the ubiquitin E3 ligase UBR5', *J Biol Chem*, vol. 285, no. 16, pp. 11800-9.

Hu, RG, Sheng, J, Qi, X, Xu, Z, Takahashi, TT & Varshavsky, A 2005, 'The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators', *Nature*, vol. 437, no. 7061, pp. 981-6.

Hua, Z & Vierstra, RD 2011, 'The cullin-RING ubiquitin-protein ligases', *Annu Rev Plant Biol*, vol. 62, pp. 299-334.

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*, vol. 286, no. 5443, pp. 1321-6.

Huang, Q, Tang, X, Wang, G, Fan, Y, Ray, L, Bergmann, A, Belenkaya, TY, Ling, X, Yan, D, Lin, Y, Ye, X, Shi, W, Zhou, X, Lu, F, Qu, J & Lin, X 2014, 'Ubr3 E3 ligase regulates apoptosis by controlling the activity of DIAP1 in Drosophila', *Cell Death Differ*, vol. 21, no. 12, pp. 1961-70.

Huang, TF, Cho, CY, Cheng, YT, Huang, JW, Wu, YZ, Yeh, AY, Nishiwaki, K, Chang, SC & Wu, YC 2014, 'BLMP-1/Blimp-1 regulates the spatiotemporal cell migration pattern in *C. elegans*', *PLoS Genet*, vol. 10, no. 6, p. e1004428.

Huesgen, PF, Lange, PF, Rogers, LD, Solis, N, Eckhard, U, Kleifeld, O, Goulas, T, Gomis-Ruth, FX & Overall, CM 2015, 'Lysarginase mirrors trypsin for protein C-terminal and methylation-site identification', *Nat Methods*, vol. 12, no. 1, pp. 55-8.

Huibregtse, JM, Scheffner, M, Beaudenon, S & Howley, PM 1995, 'A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase', *Proc Natl Acad Sci USA*, vol. 92, no. 7, pp. 2563-7.

Hwang, C-S, Shemorry, A, Auerbach, D & Varshavsky, A 2010, 'The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases', *Nat Cell Biol*, vol. 12, no. 12, pp. 1177-85.

Hwang, C-S, Shemorry, A & Varshavsky, A 2010, 'N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals', *Science (New York, N.Y.)*, vol. 327, no. 5968, pp. 973-7.

Hwang, CS, Shemorry, A & Varshavsky, A 2009, 'Two proteolytic pathways regulate DNA repair by cotargeting the Mgt1 alkylguanine transferase', *Proc Natl Acad Sci U S A*, vol. 106, no. 7, pp. 2142-7.

Hwang, CS, Sukalo, M, Batygin, O, Addor, MC, Brunner, H, Aytes, AP, Mayerle, J, Song, HK, Varshavsky, A & Zenker, M 2011, 'Ubiquitin ligases of the N-end rule pathway: assessment of mutations in UBR1 that cause the Johanson-Blizzard syndrome', *PLoS One*, vol. 6, no. 9, p. e24925.

Hwang, CS & Varshavsky, A 2008, 'Regulation of peptide import through phosphorylation of Ubr1, the ubiquitin ligase of the N-end rule pathway', *Proc Natl Acad Sci U S A*, vol. 105, no. 49, pp. 19188-93.

Ichimura, T, Yamamura, H, Sasamoto, K, Tominaga, Y, Taoka, M, Kakiuchi, K, Shinkawa, T, Takahashi, N, Shimada, S & Isobe, T 2005, '14-3-3 proteins modulate the expression of epithelial Na⁺ channels by phosphorylation-dependent interaction with Nedd4-2 ubiquitin ligase', *J Biol Chem*, vol. 280, no. 13, pp. 13187-94.

Ingolia, NT, Brar, GA, Stern-Ginossar, N, Harris, MS, Talhouarne, GJ, Jackson, SE, Wills, MR & Weissman, JS 2014, 'Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes', *Cell Rep*, vol. 8, no. 5, pp. 1365-79.

Jana, M & Bandyopadhyay, S 2012, 'Conformational flexibility of a protein-carbohydrate complex and the structure and ordering of surrounding water', *Phys Chem Chem Phys*, vol. 14, no. 18, pp. 6628-38.

Jiang, W, Wang, S, Xiao, M, Lin, Y, Zhou, L, Lei, Q, Xiong, Y, Guan, KL & Zhao, S 2011, 'Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase', *Mol Cell*, vol. 43, no. 1, pp. 33-44.

Jiang, Y, Choi, WH, Lee, JH, Han, DH, Kim, JH, Chung, YS, Kim, SH & Lee, MJ 2014, 'A neurostimulant para-chloroamphetamine inhibits the arginylation branch of the N-end rule pathway', *Sci Rep*, vol. 4, p. 6344.

Jiang, YXL, Pore, SK, Lee, JH, Sriram, S, Mai, BK, Han, DH, Agarwalla, P, Zakrzewska, A, Kim, Y, Banerjee, R, Lee, SH & Lee, MJ 2013, 'Characterization of mammalian N-degrons and development of heterovalent inhibitors of the N-end rule pathway', *Chemical Science*, vol. 4, no. 8, pp. 3339-46.

Jin, L, Williamson, A, Banerjee, S, Philipp, I & Rape, M 2008, 'Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex', *Cell*, vol. 133, no. 4, pp. 653-65.

Jin, Y, Shenoy, AK, Doernberg, S, Chen, H, Luo, H, Shen, H, Lin, T, Tarrash, M, Cai, Q, Hu, X, Fiske, R, Chen, T, Wu, L, Mohammed, KA, Rottiers, V, Lee, SS & Lu, J 2015, 'FBXO11 promotes ubiquitination of the Snail family of transcription factors in cancer progression and epidermal development', *Cancer Lett*, vol. 362, no. 1, pp. 70-82.

Jinek, M, Fabian, MR, Coyle, SM, Sonenberg, N & Doudna, JA 2010, 'Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation', *Nat Struct Mol Biol*, vol. 17, no. 2, pp. 238-40.

Johanson, A & Blizzard, R 1971, 'A syndrome of congenital aplasia of the alae nasi, deafness, hypothyroidism, dwarfism, absent permanent teeth, and malabsorption', *J Pediatr*, vol. 79, no. 6, pp. 982-7.

Jung, HY, Wang, X, Jun, S & Park, JI 2013, 'Dyrk2-associated EDD-DDB1-VprBP E3 ligase inhibits telomerase by TERT degradation', *J Biol Chem*, vol. 288, no. 10, pp. 7252-62.

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*, vol. 2, p. 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*, vol. 36, no. 6, pp. 1095-102.

Kamura, T, Conrad, MN, Yan, Q, Conaway, RC & Conaway, JW 1999, 'The Rbx1 subunit of SCF and VHL E3 ubiquitin ligase activates Rub1 modification of cullins Cdc53 and Cul2', *Genes Dev*, vol. 13, no. 22, pp. 2928-33.

Kato, T, Tamiya, G, Koyama, S, Nakamura, T, Makino, S, Arawaka, S, Kawanami, T & Tooyama, I 2012, 'UBR5 Gene Mutation Is Associated with Familial Adult Myoclonic Epilepsy in a Japanese Family', *ISRN Neurol*, vol. 2012, p. 508308.

Khmelinskii, A & Knop, M 2014, 'Analysis of protein dynamics with tandem fluorescent protein timers', *Methods Mol Biol*, vol. 1174, pp. 195-210.

Khoury, GA, Baliban, RC & Floudas, CA 2011, 'Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database', *Sci Rep*, vol. 1.

Kim, HC & Huibregtse, JM 2009, 'Polyubiquitination by HECT E3s and the determinants of chain type specificity', *Mol Cell Biol*, vol. 29, no. 12, pp. 3307-18.

Kim, HC, Steffen, AM, Oldham, ML, Chen, J & Huibregtse, JM 2011, 'Structure and function of a HECT domain ubiquitin-binding site', *EMBO Rep*, vol. 12, no. 4, pp. 334-41.

Kim, HK, Kim, RR, Oh, JH, Cho, H, Varshavsky, A & Hwang, CS 2014, 'The N-Terminal Methionine of Cellular Proteins as a Degradation Signal', *Cell*, vol. 156, no. 0, pp. 158-69.

Kim, I, Miller, CR, Young, DL & Fields, S 2013, 'High-throughput analysis of in vivo protein stability', *Mol Cell Proteomics*, vol. 12, no. 11, pp. 3370-8.

Kisselev, AF, Akopian, TN, Woo, KM & Goldberg, AL 1999, 'The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation', *J Biol Chem*, vol. 274, no. 6, pp. 3363-71.

Kleiger, G, Saha, A, Lewis, S, Kuhlman, B & Deshaies, RJ 2009, 'Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates', *Cell*, vol. 139, no. 5, pp. 957-68.

Komander, D & Rape, M 2012, 'The ubiquitin code', *Annu Rev Biochem*, vol. 81, pp. 203-29.

Kozlov, G, De Crescenzo, G, Lim, NS, Siddiqui, N, Fantus, D, Kahvejian, A, Trempe, JF, Elias, D, Ekiel, I, Sonenberg, N, O'Connor-McCourt, M & Gehring, K 2004, 'Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase', *EMBO J*, vol. 23, no. 2, pp. 272-81.

Kozlov, G, Menade, M, Rosenauer, A, Nguyen, L & Gehring, K 2010, 'Molecular determinants of PAM2 recognition by the MLLE domain of poly(A)-binding protein', *J Mol Biol*, vol. 397, no. 2, pp. 397-407.

Kozlov, G, Nguyen, L, Lin, T, De Crescenzo, G, Park, M & Gehring, K 2007, 'Structural basis of ubiquitin recognition by the ubiquitin-associated (UBA) domain of the ubiquitin ligase EDD', *J Biol Chem*, vol. 282, no. 49, pp. 35787-95.

Kozlov, G, Safaee, N, Rosenauer, A & Gehring, K 2010, 'Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mlle domain of poly(A)-binding protein', *J Biol Chem*, vol. 285, no. 18, pp. 13599-606.

Kozlov, G, Trempe, JF, Khaleghpour, K, Kahvejian, A, Ekiel, I & Gehring, K 2001, 'Structure and function of the C-terminal PABC domain of human poly(A)-binding protein', *Proc Natl Acad Sci U S A*, vol. 98, no. 8, pp. 4409-13.

Kwak, KS, Zhou, X, Solomon, V, Baracos, VE, Davis, J, Bannon, AW, Boyle, WJ, Lacey, DL & Han, HQ 2004, 'Regulation of protein catabolism by muscle-specific and cytokine-inducible ubiquitin ligase E3alpha-II during cancer cachexia', *Cancer Res*, vol. 64, no. 22, pp. 8193-8.

Kwon, YT, Kashina, AS, Davydov, IV, Hu, RG, An, JY, Seo, JW, Du, F & Varshavsky, A 2002, 'An essential role of N-terminal arginylation in cardiovascular development', *Science*, vol. 297, no. 5578, pp. 96-9.

Kwon, YT, Levy, F & Varshavsky, A 1999, 'Bivalent inhibitor of the N-end rule pathway', *J Biol Chem*, vol. 274, no. 25, pp. 18135-9.

Kwon, YT, Reiss, Y, Fried, VA, Hershko, A, Yoon, JK, Gonda, DK, Sangan, P, Copeland, NG, Jenkins, NA & Varshavsky, A 1998, 'The mouse and human genes encoding the recognition component of the N-end rule pathway', *Proc Natl Acad Sci U S A*, vol. 95, no. 14, pp. 7898-903.

Kwon, YT, Xia, Z, An, JY, Tasaki, T, Davydov, IV, Seo, JW, Sheng, J, Xie, Y & Varshavsky, A 2003, 'Female lethality and apoptosis of spermatocytes in mice lacking

the UBR2 ubiquitin ligase of the N-end rule pathway', *Mol Cell Biol*, vol. 23, no. 22, pp. 8255-71.

Kwon, YT, Xia, Z, Davydov, IV, Lecker, SH & Varshavsky, A 2001, 'Construction and analysis of mouse strains lacking the ubiquitin ligase UBR1 (E3alpha) of the N-end rule pathway', *Mol Cell Biol*, vol. 21, no. 23, pp. 8007-21.

Ladbury, JE 1996, 'Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design', *Chemistry & Biology*, vol. 3, no. 12, pp. 973-80.

Laitaoja, M, Valjakka, J & Janis, J 2013, 'Zinc coordination spheres in protein structures', *Inorg Chem*, vol. 52, no. 19, pp. 10983-91.

Laskowski, RA, MacArthur, MW, Moss, DS & Thornton, JM 1993, 'PROCHECK: a program to check the stereochemical quality of protein structures', *J. Appl. Crystallogr.*, vol. 26, pp. 283-91.

Lauressergues, D, Couzigou, JM, Clemente, HS, Martinez, Y, Dunand, C, Becard, G & Combier, JP 2015, 'Primary transcripts of microRNAs encode regulatory peptides', *Nature*, vol. 520, no. 7545, pp. 90-3.

Le Poole, IC, Sarangarajan, R, Zhao, Y, Stennett, LS, Brown, TL, Sheth, P, Miki, T & Boissy, RE 2001, 'VIT1', a novel gene associated with vitiligo', *Pigment Cell Res*, vol. 14, no. 6, pp. 475-84.

Lee, JH, Jiang, Y, Kwon, YT & Lee, MJ 2015, 'Pharmacological Modulation of the N-End Rule Pathway and Its Therapeutic Implications', *Trends Pharmacol Sci*, vol. 36, no. 11, pp. 782-97.

Lee, JN, Song, B, DeBose-Boyd, RA & Ye, J 2006, 'Sterol-regulated degradation of Insig-1 mediated by the membrane-bound ubiquitin ligase gp78', *J Biol Chem*, vol. 281, no. 51, pp. 39308-15.

Lee, KE, Heo, JE, Kim, JM & Hwang, CS 2016, 'N-Terminal Acetylation-Targeted N-End Rule Proteolytic System: The Ac/N-End Rule Pathway', *Mol Cells*, vol. 39, no. 3, pp. 169-78.

Lee, MJ, Tasaki, T, Moroi, K, An, JY, Kimura, S, Davydov, IV & Kwon, YT 2005, 'RGS4 and RGS5 are in vivo substrates of the N-end rule pathway', *Proc Natl Acad Sci U S A*, vol. 102, no. 42, pp. 15030-5.

Lee, PC, Sowa, ME, Gygi, SP & Harper, JW 2011, 'Alternative ubiquitin activation/conjugation cascades interact with N-end rule ubiquitin ligases to control degradation of RGS proteins', *Mol Cell*, vol. 43, no. 3, pp. 392-405.

Levinson, NM & Boxer, SG 2014, 'A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity', *Nat Chem Biol*, vol. 10, no. 2, pp. 127-32.

Li, T, Fan, J, Blanco-Sanchez, B, Giagtzoglou, N, Lin, G, Yamamoto, S, Jaiswal, M, Chen, K, Zhang, J, Wei, W, Lewis, MT, Groves, AK, Westerfield, M, Jia, J & Bellen, HJ 2016, 'Ubr3, a Novel Modulator of Hh Signaling Affects the Degradation of Costal-2 and Kif7 through Poly-ubiquitination', *PLoS Genet*, vol. 12, no. 5, p. e1006054.

Li, T, Giagtzoglou, N, Eberl, DF, Jaiswal, SN, Cai, T, Godt, D, Groves, AK & Bellen, HJ 2016, 'The E3 ligase Ubr3 regulates Usher syndrome and MYH9 disorder proteins in the auditory organs of *Drosophila* and mammals', *Elife*, vol. 5.

Li, W, Bengtson, MH, Ulbrich, A, Matsuda, A, Reddy, VA, Orth, A, Chanda, SK, Batalov, S & Joazeiro, CA 2008, 'Genome-wide and functional annotation of human E3

ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling', *PLoS One*, vol. 3, no. 1, p. e1487.

Li, Y, Chen, F, Lin, F, Guan, C, Wei, X, Wan, Y & Xu, A 2009, 'VIT1/FBXO11 knockdown induces morphological alterations and apoptosis in B10BR mouse melanocytes', *Int J Mol Med*, vol. 23, no. 5, pp. 673-8.

Li, Z, Huang, Y, Li, H, Hu, J, Liu, X, Jiang, T, Sun, G, Tang, A, Sun, X, Qian, W, Zeng, Y, Xie, J, Zhao, W, Xu, Y, He, T, Dong, C, Liu, Q, Mou, L, Lu, J, Lin, Z, Wu, S, Gao, S, Guo, G, Feng, Q, Li, Y, Zhang, X, Wang, J, Yang, H, Wang, J, Xiong, C, Cai, Z & Gui, Y 2015, 'Excess of rare variants in genes that are key epigenetic regulators of spermatogenesis in the patients with non-obstructive azoospermia', *Sci Rep*, vol. 5, p. 8785.

Liew, CW, Sun, H, Hunter, T & Day, CL 2010, 'RING domain dimerization is essential for RNF4 function', *Biochem J*, vol. 431, no. 1, pp. 23-9.

Lim, NS, Kozlov, G, Chang, TC, Groover, O, Siddiqui, N, Volpon, L, De Crescenzo, G, Shyu, AB & Gehring, K 2006, 'Comparative peptide binding studies of the PABC domains from the ubiquitin-protein isopeptide ligase HYD and poly(A)-binding protein. Implications for HYD function', *J Biol Chem*, vol. 281, no. 20, pp. 14376-82.

Lin, R, Tao, R, Gao, X, Li, T, Zhou, X, Guan, KL, Xiong, Y & Lei, QY 2013, 'Acetylation stabilizes ATP-citrate lyase to promote lipid biosynthesis and tumor growth', *Mol Cell*, vol. 51, no. 4, pp. 506-18.

Ling, HH, Beaulé, C, Chiang, CK, Tian, R, Figeys, D & Cheng, HY 2014, 'Time-of-day- and light-dependent expression of ubiquitin protein ligase E3 component N-recognin 4 (UBR4) in the suprachiasmatic nucleus circadian clock', *PLoS One*, vol. 9, no. 8, p. e103103.

Ling, S & Lin, WC 2011, 'EDD inhibits ATM-mediated phosphorylation of p53', *J Biol Chem*, vol. 286, no. 17, pp. 14972-82.

Lipkowitz, S & Weissman, AM 2011, 'RINGS of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis', *Nat Rev Cancer*, vol. 11, no. 9, pp. 629-43.

Liu, J, Shaik, S, Dai, X, Wu, Q, Zhou, X, Wang, Z & Wei, W 2015, 'Targeting the ubiquitin pathway for cancer treatment', *Biochim Biophys Acta*, vol. 1855, no. 1, pp. 50-60.

Mann, KM, Ward, JM, Yew, CC, Kovochich, A, Dawson, DW, Black, MA, Brett, BT, Sheetz, TE, Dupuy, AJ, Australian Pancreatic Cancer Genome, I, Chang, DK, Biankin, AV, Waddell, N, Kassahn, KS, Grimmond, SM, Rust, AG, Adams, DJ, Jenkins, NA & Copeland, NG 2012, 'Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma', *Proc Natl Acad Sci U S A*, vol. 109, no. 16, pp. 5934-41.

Mansfield, E, Hersperger, E, Biggs, J & Shearn, A 1994, 'Genetic and Molecular Analysis of hyperplastic discs, a Gene Whose Product Is Required for Regulation of Cell Proliferation in *Drosophila melanogaster* Imaginal Discs and Germ Cells', *Developmental Biology*, vol. 165, no. 2, pp. 507-26.

Mansfield, E, Hersperger, E, Biggs, J & Shearn, A 1994, 'Genetic and molecular analysis of hyperplastic discs, a gene whose product is required for regulation of cell proliferation in *Drosophila melanogaster* imaginal discs and germ cells', *Dev Biol*, vol. 165, no. 2, pp. 507-26.

Mari, S, Ruetalo, N, Maspero, E, Stoffregen, MC, Pasqualato, S, Polo, S & Wiesner, S 2014, 'Structural and functional framework for the autoinhibition of Nedd4-family ubiquitin ligases', *Structure*, vol. 22, no. 11, pp. 1639-49.

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*, vol. 12, no. 4, pp. 342-9.

Matsuura, K, Huang, NJ, Cocce, K, Zhang, L & Kornbluth, S 2017, 'Downregulation of the proapoptotic protein MOAP-1 by the UBR5 ubiquitin ligase and its role in ovarian cancer resistance to cisplatin', *Oncogene*, vol. 36, no. 12, pp. 1698-706.

Matta-Camacho, E, Kozlov, G, Li, FF & Gehring, K 2010, 'Structural basis of substrate recognition and specificity in the N-end rule pathway', *Nat Struct Mol Biol*, vol. 17, no. 10, pp. 1182-7.

Matta-Camacho, E, Kozlov, G, Menade, M & Gehring, K 2012, 'Structure of the HECT C-lobe of the UBR5 E3 ubiquitin ligase', *Acta Crystallogr Sect F Struct Biol Cryst Commun*, vol. 68, no. Pt 10, pp. 1158-63.

McCarty, AS, Kleiger, G, Eisenberg, D & Smale, ST 2003, 'Selective dimerization of a C2H2 zinc finger subfamily', *Mol Cell*, vol. 11, no. 2, pp. 459-70.

McCoy, AJ, Grosse-Kunstleve, RW, Adams, PD, Winn, MD, Storoni, LC & Read, RJ 2007, 'Phaser crystallographic software', *J Appl Crystallogr*, vol. 40, no. Pt 4, pp. 658-74.

McDonald, WJ, Sangster, SM, Moffat, LD, Henderson, MJ & Too, CK 2010, 'alpha4 phosphoprotein interacts with EDD E3 ubiquitin ligase and poly(A)-binding protein', *J Cell Biochem*, vol. 110, no. 5, pp. 1123-9.

McRee, DE 1999, 'XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density', *J Struct Biol*, vol. 125, no. 2-3, pp. 156-65.

Meisenberg, C, Tait, PS, Dianova, II, Wright, K, Edelman, MJ, Ternette, N, Tasaki, T, Kessler, BM, Parsons, JL, Kwon, YT & Dianov, GL 2012, 'Ubiquitin ligase UBR3 regulates cellular levels of the essential DNA repair protein APE1 and is required for genome stability', *Nucleic Acids Res*, vol. 40, no. 2, pp. 701-11.

Meissner, B, Kridel, R, Lim, RS, Rogic, S, Tse, K, Scott, DW, Moore, R, Mungall, AJ, Marra, MA, Connors, JM, Steidl, C & Gascoyne, RD 2013, 'The E3 ubiquitin ligase UBR5 is recurrently mutated in mantle cell lymphoma', *Blood*, vol. 121, no. 16, pp. 3161-4.

Metzger, MB, Pruneda, JN, Klevit, RE & Weissman, AM 2014, 'RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination', *Biochim Biophys Acta*, vol. 1843, no. 1, pp. 47-60.

Michel, J, Tirado-Rives, J & Jorgensen, WL 2009, 'Energetics of displacing water molecules from protein binding sites: consequences for ligand optimization', *J Am Chem Soc*, vol. 131, no. 42, pp. 15403-11.

Michelle, C, Vourc'h, P, Mignon, L & Andres, CR 2009, 'What was the set of ubiquitin and ubiquitin-like conjugating enzymes in the eukaryote common ancestor?', *J Mol Evol*, vol. 68, no. 6, pp. 616-28.

Morett, E & Bork, P 1999, 'A novel transactivation domain in parkin', *Trends Biochem Sci*, vol. 24, no. 6, pp. 229-31.

Mori, J, Kawabata, A, Tang, H, Tadagaki, K, Mizuguchi, H, Kuroda, K & Mori, Y 2015, 'Human Herpesvirus-6 U14 Induces Cell-Cycle Arrest in G2/M Phase by Associating with a Cellular Protein, EDD', *PLoS One*, vol. 10, no. 9, p. e0137420.

Mori, Y, Sato, F, Selaru, FM, Olaru, A, Perry, K, Kimos, MC, Tamura, G, Matsubara, N, Wang, S, Xu, Y, Yin, J, Zou, TT, Leggett, B, Young, J, Nukiwa, T, Stine, OC, Abraham,

JM, Shibata, D & Meltzer, SJ 2002, 'Instabilotyping reveals unique mutational spectra in microsatellite-unstable gastric cancers', *Cancer Res*, vol. 62, no. 13, pp. 3641-5.

Morrison, J, Laurent-Rolle, M, Maestre, AM, Rajsbaum, R, Pisanelli, G, Simon, V, Mulder, LC, Fernandez-Sesma, A & Garcia-Sastre, A 2013, 'Dengue virus co-opts UBR4 to degrade STAT2 and antagonize type I interferon signaling', *PLoS Pathog*, vol. 9, no. 3, p. e1003265.

Müller, D, Rehbein, M, Baumeister, H & Richter, D 1992, 'Molecular characterization of a novel rat protein structurally related to poly(A) binding proteins and the 70K protein of the U1 small nuclear ribonucleoprotein particle (snRNP)', *Nucleic Acids Research*, vol. 20, no. 7, pp. 1471-5.

Munoz, MA, Saunders, DN, Henderson, MJ, Clancy, JL, Russell, AJ, Lehrbach, G, Musgrove, EA, Watts, CK & Sutherland, RL 2007, 'The E3 ubiquitin ligase EDD regulates S-phase and G(2)/M DNA damage checkpoints', *Cell Cycle*, vol. 6, no. 24, pp. 3070-7.

Munoz-Escobar, J, Matta-Camacho, E, Cho, C, Kozlov, G & Gehring, K 2017, 'Bound Waters Mediate Binding of Diverse Substrates to a Ubiquitin Ligase', *Structure*.

Murshudov, GN, Vagin, AA & Dodson, EJ 1997, 'Refinement of macromolecular structures by the maximum-likelihood method', *Acta Crystallogr D Biol Crystallogr*, vol. 53, no. Pt 3, pp. 240-55.

Murshudov, GN, Vagin, AA, Lebedev, A, Wilson, KS & Dodson, EJ 1999, 'Efficient anisotropic refinement of macromolecular structures using FFT', *Acta Crystallogr D Biol Crystallogr*, vol. 55, no. Pt 1, pp. 247-55.

Nagel, S, Pommerenke, C, Meyer, C, Kaufmann, M, MacLeod, RA & Drexler, HG 2017, 'Identification of a tumor suppressor network in T-cell leukemia', *Leuk Lymphoma*, pp. 1-15.

Nakagawa, T, Mondal, K & Swanson, PC 2013, 'VprBP (DCAF1): a promiscuous substrate recognition subunit that incorporates into both RING-family CRL4 and HECT-family EDD/UBR5 E3 ubiquitin ligases', *BMC Mol Biol*, vol. 14, p. 22.

Nakatani, Y, Konishi, H, Vassilev, A, Kurooka, H, Ishiguro, K, Sawada, J, Ikura, T, Korsmeyer, SJ, Qin, J & Herlitz, AM 2005, 'p600, a unique protein required for membrane morphogenesis and cell survival', *Proc Natl Acad Sci U S A*, vol. 102, no. 42, pp. 15093-8.

Narita, K 1958, 'Isolation of acetylpeptide from enzymic digests of TMV-protein', *Biochim Biophys Acta*, vol. 28, no. 1, pp. 184-91.

Ngollo, M, Lebert, A, Daures, M, Judes, G, Rifai, K, Dubois, L, Kemeny, JL, Penault-Llorca, F, Bignon, YJ, Guy, L & Bernard-Gallon, D 2017, 'Global analysis of H3K27me3 as an epigenetic marker in prostate cancer progression', *BMC Cancer*, vol. 17, no. 1, p. 261.

Noyes, NC, Hampton, B, Migliorini, M & Strickland, DK 2016, 'Regulation of Itch and Nedd4 E3 Ligase Activity and Degradation by LRAD3', *Biochemistry*, vol. 55, no. 8, pp. 1204-13.

O'Brien, PM, Davies, MJ, Scurry, JP, Smith, AN, Barton, CA, Henderson, MJ, Saunders, DN, Gloss, BS, Patterson, KI, Clancy, JL, Heinzelmann-Schwarz, VA, Murali, R, Scolyer, RA, Zeng, Y, Williams, ED, Scurr, L, Defazio, A, Quinn, DI, Watts, CK, Hacker, NF, Henshall, SM & Sutherland, RL 2008, 'The E3 ubiquitin ligase EDD is an adverse prognostic factor for serous epithelial ovarian cancer and modulates cisplatin resistance in vitro', *Br J Cancer*, vol. 98, no. 6, pp. 1085-93.

Oh, JH, Hyun, JY & Varshavsky, A 2017, 'Control of Hsp90 chaperone and its clients by N-terminal acetylation and the N-end rule pathway', *Proc Natl Acad Sci U S A*.

Ohshima, R, Ohta, T, Wu, W, Koike, A, Iwatani, T, Henderson, M, Watts, CK & Otsubo, T 2007, 'Putative tumor suppressor EDD interacts with and up-regulates APC', *Genes Cells*, vol. 12, no. 12, pp. 1339-45.

Okamoto, K, Bartocci, C, Ouzounov, I, Diedrich, JK, Yates, JR, 3rd & Denchi, EL 2013, 'A two-step mechanism for TRF2-mediated chromosome-end protection', *Nature*, vol. 494, no. 7438, pp. 502-5.

Ong, SS, Goktug, AN, Elias, A, Wu, J, Saunders, D & Chen, T 2014, 'Stability of the human pregnane X receptor is regulated by E3 ligase UBR5 and serine/threonine kinase DYRK2', *Biochem J*, vol. 459, no. 1, pp. 193-203.

Otwinowski, Z & Minor, W 1997, 'Processing of X-ray diffraction data collected in oscillation mode', in Charles W. Carter, Jr. (ed.), *Methods in Enzymology*, Academic Press, vol. Volume 276, pp. 307-26.

Otwinowski, Z & Minor, W 1997, 'Processing of X-ray diffraction data collected in oscillation mode', *Methods Enzymol*, vol. 276, pp. 307-26.

Oughtred, R, Bedard, N, Adegoke, OA, Morales, CR, Trasler, J, Rajapurohitam, V & Wing, SS 2002, 'Characterization of rat100, a 300-kilodalton ubiquitin-protein ligase induced in germ cells of the rat testis and similar to the *Drosophila* hyperplastic discs gene', *Endocrinology*, vol. 143, no. 10, pp. 3740-7.

Ouyang, Y, Kwon, YT, An, JY, Eller, D, Tsai, SC, Diaz-Perez, S, Troke, JJ, Teitell, MA & Marahrens, Y 2006, 'Loss of Ubr2, an E3 ubiquitin ligase, leads to chromosome

fragility and impaired homologous recombinational repair', *Mutat Res*, vol. 596, no. 1-2, pp. 64-75.

Painter, J & Merritt, EA 2006, 'Optimal description of a protein structure in terms of multiple groups undergoing TLS motion', *Acta Crystallogr D Biol Crystallogr*, vol. 62, no. Pt 4, pp. 439-50.

Pandya, RK, Partridge, JR, Love, KR, Schwartz, TU & Ploegh, HL 2010, 'A structural element within the HUWE1 HECT domain modulates self-ubiquitination and substrate ubiquitination activities', *J Biol Chem*, vol. 285, no. 8, pp. 5664-73.

Park, SE, Kim, JM, Seok, OH, Cho, H, Wadas, B, Kim, SY, Varshavsky, A & Hwang, CS 2015, 'Control of mammalian G protein signaling by N-terminal acetylation and the N-end rule pathway', *Science*, vol. 347, no. 6227, pp. 1249-52.

Parry, M, Rose-Zerilli, MJ, Gibson, J, Ennis, S, Walewska, R, Forster, J, Parker, H, Davis, Z, Gardiner, A, Collins, A, Oscier, DG & Strefford, JC 2013, 'Whole exome sequencing identifies novel recurrently mutated genes in patients with splenic marginal zone lymphoma', *PLoS One*, vol. 8, no. 12, p. e83244.

Parsons, K, Nakatani, Y & Nguyen, MD 2015, 'p600/UBR4 in the central nervous system', *Cell Mol Life Sci*, vol. 72, no. 6, pp. 1149-60.

Peng, J, Schwartz, D, Elias, JE, Thoreen, CC, Cheng, D, Marsischky, G, Roelofs, J, Finley, D & Gygi, SP 2003, 'A proteomics approach to understanding protein ubiquitination', *Nat Biotechnol*, vol. 21, no. 8, pp. 921-6.

Piatkov, KI, Brower, CS & Varshavsky, A 2012, 'The N-end rule pathway counteracts cell death by destroying proapoptotic protein fragments', *Proc Natl Acad Sci U S A*, vol. 109, no. 27, pp. E1839-47.

Pickart, CM 2001, 'Mechanisms underlying ubiquitination', *Annu Rev Biochem*, vol. 70, pp. 503-33.

Poyurovsky, MV, Priest, C, Kentsis, A, Borden, KL, Pan, ZQ, Pavletich, N & Prives, C 2007, 'The Mdm2 RING domain C-terminus is required for supramolecular assembly and ubiquitin ligase activity', *EMBO J*, vol. 26, no. 1, pp. 90-101.

Rahighi, S, Ikeda, F, Kawasaki, M, Akutsu, M, Suzuki, N, Kato, R, Kensche, T, Uejima, T, Bloor, S, Komander, D, Randow, F, Wakatsuki, S & Dikic, I 2009, 'Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation', *Cell*, vol. 136, no. 6, pp. 1098-109.

Ramakrishnan, R, Liu, H, Donahue, H, Malovannaya, A, Qin, J & Rice, AP 2012, 'Identification of novel CDK9 and Cyclin T1-associated protein complexes (CCAPs) whose siRNA depletion enhances HIV-1 Tat function', *Retrovirology*, vol. 9, p. 90.

Rao, H, Uhlmann, F, Nasmyth, K & Varshavsky, A 2001, 'Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability', *Nature*, vol. 410, no. 6831, pp. 955-9.

Reid, MA, Wang, WI, Rosales, KR, Welliver, MX, Pan, M & Kong, M 2013, 'The B55alpha subunit of PP2A drives a p53-dependent metabolic adaptation to glutamine deprivation', *Mol Cell*, vol. 50, no. 2, pp. 200-11.

Reiss, Y, Kaim, D & Hershko, A 1988, 'Specificity of binding of NH2-terminal residue of proteins to ubiquitin-protein ligase. Use of amino acid derivatives to characterize specific binding sites', *J Biol Chem*, vol. 263, no. 6, pp. 2693-8.

Riley, BE, Loughheed, JC, Callaway, K, Velasquez, M, Brecht, E, Nguyen, L, Shaler, T, Walker, D, Yang, Y, Regnstrom, K, Diep, L, Zhang, Z, Chiou, S, Bova, M, Artis, DR, Yao, N, Baker, J, Yednock, T & Johnston, JA 2013, 'Structure and function of Parkin E3

ubiquitin ligase reveals aspects of RING and HECT ligases', *Nat Commun*, vol. 4, p. 1982.

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*, vol. 290, no. 39, pp. 23875-87.

Rinschen, MM, Bharill, P, Wu, X, Kohli, P, Reinert, MJ, Kretz, O, Saez, I, Schermer, B, Hohne, M, Bartram, MP, Aravamudhan, S, Brooks, BR, Vilchez, D, Huber, TB, Muller, RU, Kruger, M & Benzing, T 2016, 'The ubiquitin ligase Ubr4 controls stability of podocin/MEC-2 supercomplexes', *Hum Mol Genet*, vol. 25, no. 7, pp. 1328-44.

Rodrigo-Brenni, MC & Morgan, DO 2007, 'Sequential E2s drive polyubiquitin chain assembly on APC targets', *Cell*, vol. 130, no. 1, pp. 127-39.

Rope, AF, Wang, K, Evjenth, R, Xing, J, Johnston, JJ, Swensen, JJ, Johnson, WE, Moore, B, Huff, CD, Bird, LM, Carey, JC, Opitz, JM, Stevens, CA, Jiang, T, Schank, C, Fain, HD, Robison, R, Dalley, B, Chin, S, South, ST, Pysher, TJ, Jorde, LB, Hakonarson, H, Lillehaug, JR, Biesecker, LG, Yandell, M, Arnesen, T & Lyon, GJ 2011, 'Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency', *Am J Hum Genet*, vol. 89, no. 1, pp. 28-43.

Rossi, M, Duan, S, Jeong, YT, Horn, M, Saraf, A, Florens, L, Washburn, MP, Antebi, A & Pagano, M 2013, 'Regulation of the CRL4(Cdt2) ubiquitin ligase and cell-cycle exit by the SCF(Fbxo11) ubiquitin ligase', *Mol Cell*, vol. 49, no. 6, pp. 1159-66.

Rutz, S, Kayagaki, N, Phung, QT, Eidenschenk, C, Noubade, R, Wang, X, Lesch, J, Lu, R, Newton, K, Huang, OW, Cochran, AG, Vasser, M, Fauber, BP, DeVoss, J, Webster, J, Diehl, L, Modrusan, Z, Kirkpatrick, DS, Lill, JR, Ouyang, W & Dixit, VM 2015,

'Deubiquitinase DUBA is a post-translational brake on interleukin-17 production in T cells', *Nature*, vol. 518, no. 7539, pp. 417-21.

Rye, MS, Bhutta, MF, Cheeseman, MT, Burgner, D, Blackwell, JM, Brown, SD & Jamieson, SE 2011, 'Unraveling the genetics of otitis media: from mouse to human and back again', *Mamm Genome*, vol. 22, no. 1-2, pp. 66-82.

Rye, MS, Blackwell, JM & Jamieson, SE 2012, 'Genetic susceptibility to otitis media in childhood', *Laryngoscope*, vol. 122, no. 3, pp. 665-75.

Rye, MS, Wiertsema, SP, Scaman, ES, Oommen, J, Sun, W, Francis, RW, Ang, W, Pennell, CE, Burgner, D, Richmond, P, Vijayasekaran, S, Coates, HL, Brown, SD, Blackwell, JM & Jamieson, SE 2011, 'FBXO11, a regulator of the TGFbeta pathway, is associated with severe otitis media in Western Australian children', *Genes Immun*, vol. 12, no. 5, pp. 352-9.

Sanchez, A, De Vivo, A, Uprety, N, Kim, J, Stevens, SM, Jr. & Kee, Y 2016, 'BMI1-UBR5 axis regulates transcriptional repression at damaged chromatin', *Proc Natl Acad Sci USA*, vol. 113, no. 40, pp. 11243-8.

Sander, B, Xu, W, Eilers, M, Popov, N & Lorenz, S 2017, 'A conformational switch regulates the ubiquitin ligase HUWE1', *Elife*, vol. 6.

Sarshad, AA, Corcoran, M, Al-Muzzaini, B, Borgonovo-Brandter, L, Von Euler, A, Lamont, D, Visa, N & Percipalle, P 2014, 'Glycogen synthase kinase (GSK) 3beta phosphorylates and protects nuclear myosin 1c from proteasome-mediated degradation to activate rDNA transcription in early G1 cells', *PLoS Genet*, vol. 10, no. 6, p. e1004390.

Saunders, DN, Hird, SL, Withington, SL, Dunwoodie, SL, Henderson, MJ, Biben, C, Sutherland, RL, Ormandy, CJ & Watts, CK 2004, 'Edd, the murine hyperplastic disc

gene, is essential for yolk sac vascularization and chorioallantoic fusion', *Mol Cell Biol*, vol. 24, no. 16, pp. 7225-34.

Scheffner, M & Kumar, S 2014, 'Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects', *Biochim Biophys Acta*, vol. 1843, no. 1, pp. 61-74.

Scheffner, M, Nuber, U & Huibregtse, JM 1995, 'Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade', *Nature*, vol. 373, no. 6509, pp. 81-3.

Schlesinger, DH, Goldstein, G & Niall, HD 1975, 'The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells', *Biochemistry*, vol. 14, no. 10, pp. 2214-8.

Schnell, JD & Hicke, L 2003, 'Non-traditional functions of ubiquitin and ubiquitin-binding proteins', *J Biol Chem*, vol. 278, no. 38, pp. 35857-60.

Schrader, EK, Harstad, KG & Matouschek, A 2009, 'Targeting proteins for degradation', *Nat Chem Biol*, vol. 5, no. 11, pp. 815-22.

Schrodinger 2015, *The PyMOL Molecular Graphics System, Version~1.8*, Version~1.8 edn, Schrodinger LLC.

Schulman, BA & Harper, JW 2009, 'Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways', *Nat Rev Mol Cell Biol*, vol. 10, no. 5, pp. 319-31.

Scialpi, F, Mellis, D & Ditzel, M 2015, 'EDD, a ubiquitin-protein ligase of the N-end rule pathway, associates with spindle assembly checkpoint components and regulates the mitotic response to nocodazole', *J Biol Chem*, vol. 290, no. 20, pp. 12585-94.

Scott, DC, Monda, JK, Bennett, EJ, Harper, JW & Schulman, BA 2011, 'N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex', *Science*, vol. 334, no. 6056, pp. 674-8.

Segade, F, Daly, KA, Allred, D, Hicks, PJ, Cox, M, Brown, M, Hardisty-Hughes, RE, Brown, SD, Rich, SS & Bowden, DW 2006, 'Association of the FBXO11 gene with chronic otitis media with effusion and recurrent otitis media: the Minnesota COME/ROM Family Study', *Arch Otolaryngol Head Neck Surg*, vol. 132, no. 7, pp. 729-33.

Setty, SR, Stochlic, TI, Tong, AH, Boone, C & Burd, CG 2004, 'Golgi targeting of ARF-like GTPase Arl3p requires its Nalpha-acetylation and the integral membrane protein Sys1p', *Nat Cell Biol*, vol. 6, no. 5, pp. 414-9.

Shea, FF, Rowell, JL, Li, Y, Chang, TH & Alvarez, CE 2012, 'Mammalian alpha arrestins link activated seven transmembrane receptors to Nedd4 family e3 ubiquitin ligases and interact with beta arrestins', *PLoS One*, vol. 7, no. 12, p. e50557.

Shearwin-Whyatt, L, Dalton, HE, Foot, N & Kumar, S 2006, 'Regulation of functional diversity within the Nedd4 family by accessory and adaptor proteins', *Bioessays*, vol. 28, no. 6, pp. 617-28.

Shemorry, A, Hwang, CS & Varshavsky, A 2013, 'Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway', *Mol Cell*, vol. 50, no. 4, pp. 540-51.

Shim, SY, Wang, J, Asada, N, Neumayer, G, Tran, HC, Ishiguro, K, Sanada, K, Nakatani, Y & Nguyen, MD 2008, 'Protein 600 is a microtubule/endoplasmic reticulum-associated protein in CNS neurons', *J Neurosci*, vol. 28, no. 14, pp. 3604-14.

Skaar, JR, Pagan, JK & Pagano, M 2013, 'Mechanisms and function of substrate recruitment by F-box proteins', *Nat Rev Mol Cell Biol*, vol. 14, no. 6, pp. 369-81.

— 2014, 'SCF ubiquitin ligase-targeted therapies', *Nat Rev Drug Discov*, vol. 13, no. 12, pp. 889-903.

Skowyra, D, Craig, KL, Tyers, M, Elledge, SJ & Harper, JW 1997, 'F-Box Proteins Are Receptors that Recruit Phosphorylated Substrates to the SCF Ubiquitin-Ligase Complex', *Cell*, vol. 91, no. 2, pp. 209-19.

Sleigh, SH, Seavers, PR, Wilkinson, AJ, Ladbury, JE & Tame, JR 1999, 'Crystallographic and calorimetric analysis of peptide binding to OppA protein', *J Mol Biol*, vol. 291, no. 2, pp. 393-415.

Smit, JJ, Monteferrario, D, Noordermeer, SM, van Dijk, WJ, van der Reijden, BA & Sixma, TK 2012, 'The E3 ligase HOIP specifies linear ubiquitin chain assembly through its RING-IBR-RING domain and the unique LDD extension', *EMBO J*, vol. 31, no. 19, pp. 3833-44.

Smith, JE, Alvarez-Dominguez, JR, Kline, N, Huynh, NJ, Geisler, S, Hu, W, Coller, J & Baker, KE 2014, 'Translation of small open reading frames within unannotated RNA transcripts in *Saccharomyces cerevisiae*', *Cell Rep*, vol. 7, no. 6, pp. 1858-66.

Smits, VA 2012, 'EDD induces cell cycle arrest by increasing p53 levels', *Cell Cycle*, vol. 11, no. 4, pp. 715-20.

Snyder, PW, Lockett, MR, Moustakas, DT & Whitesides, GM 2014, 'Is it the shape of the cavity, or the shape of the water in the cavity?', *Eur Phys J-Spec Top*, vol. 223, no. 5, pp. 853-91.

Soucy, TA, Smith, PG, Milhollen, MA, Berger, AJ, Gavin, JM, Adhikari, S, Brownell, JE, Burke, KE, Cardin, DP, Critchley, S, Cullis, CA, Doucette, A, Garnsey, JJ, Gaulin, JL, Gershman, RE, Lublinsky, AR, McDonald, A, Mizutani, H, Narayanan, U, Olhava,

EJ, Peluso, S, Rezaei, M, Sintchak, MD, Talreja, T, Thomas, MP, Traore, T, Vyskocil, S, Weatherhead, GS, Yu, J, Zhang, J, Dick, LR, Claiborne, CF, Rolfe, M, Bolen, JB & Langston, SP 2009, 'An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer', *Nature*, vol. 458, no. 7239, pp. 732-6.

Spratt, DE, Walden, H & Shaw, GS 2014, 'RBR E3 ubiquitin ligases: new structures, new insights, new questions', *Biochem J*, vol. 458, no. 3, pp. 421-37.

Sriram, S, Lee, JH, Mai, BK, Jiang, Y, Kim, Y, Yoo, YD, Banerjee, R, Lee, SH & Lee, MJ 2013, 'Development and characterization of monomeric N-end rule inhibitors through in vitro model substrates', *J Med Chem*, vol. 56, no. 6, pp. 2540-6.

Sriram, SM, Kim, BY & Kwon, YT 2011, 'The N-end rule pathway: emerging functions and molecular principles of substrate recognition', *Nat Rev Mol Cell Biol*, vol. 12, no. 11, pp. 735-47.

Sriram, SM & Kwon, YT 2010, 'The molecular principles of N-end rule recognition', *Nat Struct Mol Biol*, vol. 17, no. 10, pp. 1164-5.

Starheim, KK, Gevaert, K & Arnesen, T 2012, 'Protein N-terminal acetyltransferases: when the start matters', *Trends in Biochemical Sciences*, vol. 37, no. 4, pp. 152-61.

Stolz, A, Besser, S, Hottmann, H & Wolf, DH 2013, 'Previously unknown role for the ubiquitin ligase Ubr1 in endoplasmic reticulum-associated protein degradation', *Proc Natl Acad Sci U S A*, vol. 110, no. 38, pp. 15271-6.

Su, H, Meng, S, Lu, Y, Trombly, MI, Chen, J, Lin, C, Turk, A & Wang, X 2011, 'Mammalian hyperplastic discs homolog EDD regulates miRNA-mediated gene silencing', *Mol Cell*, vol. 43, no. 1, pp. 97-109.

Subbaiah, VK, Zhang, Y, Rajagopalan, D, Abdullah, LN, Yeo-Teh, NS, Tomaic, V, Banks, L, Myers, MP, Chow, EK & Jha, S 2016, 'E3 ligase EDD1/UBR5 is utilized by the HPV E6 oncogene to destabilize tumor suppressor TIP60', *Oncogene*, vol. 35, no. 16, pp. 2062-74.

Sukalo, M, Fiedler, A, Guzman, C, Spranger, S, Addor, MC, McHeik, JN, Oltra Benavent, M, Cobben, JM, Gillis, LA, Shealy, AG, Deshpande, C, Bozorgmehr, B, Everman, DB, Stattin, EL, Liebelt, J, Keller, KM, Bertola, DR, van Karnebeek, CD, Bergmann, C, Liu, Z, Duker, G, Rezaei, N, Alkuraya, FS, Ogur, G, Alrajoudi, A, Venegas-Vega, CA, Verbeek, NE, Richmond, EJ, Kirbiyik, O, Ranganath, P, Singh, A, Godbole, K, Ali, FA, Alves, C, Mayerle, J, Lerch, MM, Witt, H & Zenker, M 2014, 'Mutations in the human UBR1 gene and the associated phenotypic spectrum', *Hum Mutat*, vol. 35, no. 5, pp. 521-31.

Szalmas, A, Tomaic, V, Basukala, O, Massimi, P, Mittal, S, Konya, J & Banks, L 2017, 'The PTPN14 Tumor Suppressor Is a Degradation Target of Human Papillomavirus E7', *J Virol*, vol. 91, no. 7.

Tanaka, K 2009, 'The proteasome: overview of structure and functions', *Proc Jpn Acad Ser B Phys Biol Sci*, vol. 85, no. 1, pp. 12-36.

Tang, X, Orlicky, S, Lin, Z, Willems, A, Neculai, D, Ceccarelli, D, Mercurio, F, Shilton, BH, Sicheri, F & Tyers, M 2007, 'Suprafacial orientation of the SCFCdc4 dimer accommodates multiple geometries for substrate ubiquitination', *Cell*, vol. 129, no. 6, pp. 1165-76.

Tasaki, T & Kwon, YT 2007, 'The mammalian N-end rule pathway: new insights into its components and physiological roles', *Trends Biochem Sci*, vol. 32, no. 11, pp. 520-8.

Tasaki, T, Mulder, LC, Iwamatsu, A, Lee, MJ, Davydov, IV, Varshavsky, A, Muesing, M & Kwon, YT 2005, 'A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons', *Mol Cell Biol*, vol. 25, no. 16, pp. 7120-36.

Tasaki, T, Sohr, R, Xia, Z, Hellweg, R, Hortnagl, H, Varshavsky, A & Kwon, YT 2007, 'Biochemical and genetic studies of UBR3, a ubiquitin ligase with a function in olfactory and other sensory systems', *J Biol Chem*, vol. 282, no. 25, pp. 18510-20.

Tasaki, T, Zakrzewska, A, Dudgeon, DD, Jiang, Y, Lazo, JS & Kwon, YT 2009, 'The substrate recognition domains of the N-end rule pathway', *J Biol Chem*, vol. 284, no. 3, pp. 1884-95.

Tateossian, H, Morse, S, Simon, MM, Dean, CH & Brown, SD 2015, 'Interactions between the otitis media gene, Fbxo11, and p53 in the mouse embryonic lung', *Dis Model Mech*, vol. 8, no. 12, pp. 1531-42.

Theodoraki, MA, Nillegoda, NB, Saini, J & Caplan, AJ 2012, 'A network of ubiquitin ligases is important for the dynamics of misfolded protein aggregates in yeast', *J Biol Chem*, vol. 287, no. 28, pp. 23911-22.

Tomaic, V, Pim, D, Thomas, M, Massimi, P, Myers, MP & Banks, L 2011, 'Regulation of the human papillomavirus type 18 E6/E6AP ubiquitin ligase complex by the HECT domain-containing protein EDD', *J Virol*, vol. 85, no. 7, pp. 3120-7.

Trempe, JF, Sauve, V, Grenier, K, Seirafi, M, Tang, MY, Menade, M, Al-Abdul-Wahid, S, Krett, J, Wong, K, Kozlov, G, Nagar, B, Fon, EA & Gehring, K 2013, 'Structure of parkin reveals mechanisms for ubiquitin ligase activation', *Science*, vol. 340, no. 6139, pp. 1451-5.

Tripathi, S, Pohl, MO, Zhou, Y, Rodriguez-Frandsen, A, Wang, G, Stein, DA, Moulton, HM, DeJesus, P, Che, J, Mulder, LC, Yanguez, E, Andenmatten, D, Pache, L,

Manicassamy, B, Albrecht, RA, Gonzalez, MG, Nguyen, Q, Brass, A, Elledge, S, White, M, Shapira, S, Hacohen, N, Karlas, A, Meyer, TF, Shales, M, Gatorano, A, Johnson, JR, Jang, G, Johnson, T, Verschueren, E, Sanders, D, Krogan, N, Shaw, M, Konig, R, Stertz, S, Garcia-Sastre, A & Chanda, SK 2015, 'Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding', *Cell Host Microbe*, vol. 18, no. 6, pp. 723-35.

Tsai, YC, Fishman, PS, Thakor, NV & Oyler, GA 2003, 'Parkin facilitates the elimination of expanded polyglutamine proteins and leads to preservation of proteasome function', *J Biol Chem*, vol. 278, no. 24, pp. 22044-55.

Turner, GC, Du, F & Varshavsky, A 2000, 'Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway', *Nature*, vol. 405, no. 6786, pp. 579-83.

Turner, GC & Varshavsky, A 2000, 'Detecting and measuring cotranslational protein degradation in vivo', *Science*, vol. 289, no. 5487, pp. 2117-20.

Uldrijan, S, Pannekoek, WJ & Vousden, KH 2007, 'An essential function of the extreme C-terminus of MDM2 can be provided by MDMX', *EMBO J*, vol. 26, no. 1, pp. 102-12.

Urakubo, Y, Ikura, T & Ito, N 2008, 'Crystal structural analysis of protein-protein interactions drastically destabilized by a single mutation', *Protein Sci*, vol. 17, no. 6, pp. 1055-65.

Van Damme, P, Lasa, M, Plevoda, B, Gazquez, C, Elosegui-Artola, A, Kim, DS, De Juan-Pardo, E, Demeyer, K, Hole, K, Larrea, E, Timmerman, E, Prieto, J, Arnesen, T, Sherman, F, Gevaert, K & Aldabe, R 2012, 'N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB', *Proc Natl Acad Sci U S A*, vol. 109, no. 31, pp. 12449-54.

van der Reijden, BA, Erpelinck-Verschueren, CA, Lowenberg, B & Jansen, JH 1999, 'TRIADs: a new class of proteins with a novel cysteine-rich signature', *Protein Sci*, vol. 8, no. 7, pp. 1557-61.

van der Veen, AG & Ploegh, HL 2012, 'Ubiquitin-like proteins', *Annu Rev Biochem*, vol. 81, pp. 323-57.

van Drogen, F, Sangfelt, O, Malyukova, A, Matskova, L, Yeh, E, Means, AR & Reed, SI 2006, 'Ubiquitylation of cyclin E requires the sequential function of SCF complexes containing distinct hCdc4 isoforms', *Mol Cell*, vol. 23, no. 1, pp. 37-48.

Varshavsky, A 1991, 'Naming a targeting signal', *Cell*, vol. 64, no. 1, pp. 13-5.

—— 1996, 'The N-end rule: functions, mysteries, uses', *Proc Natl Acad Sci U S A*, vol. 93, no. 22, pp. 12142-9.

—— 2011, 'The N-end rule pathway and regulation by proteolysis', *Protein Sci*, vol. 20, no. 8, pp. 1298-345.

—— 2012, 'The ubiquitin system, an immense realm', *Annu Rev Biochem*, vol. 81, pp. 167-76.

—— 2017, 'The Ubiquitin System, Autophagy, and Regulated Protein Degradation', *Annu Rev Biochem*, vol. 86, pp. 123-8.

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*, vol. 11, no. 1, pp. 249-59.

Verma, R, Oania, RS, Kolawa, NJ & Deshaies, RJ 2013, 'Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome', *Elife*, vol. 2, p. e00308.

Vriend, G 1990, 'WHAT IF: a molecular modeling and drug design program', *J Mol Graph*, vol. 8, no. 1, pp. 52-6, 29.

Wang, KH, Roman-Hernandez, G, Grant, RA, Sauer, RT & Baker, TA 2008, 'The molecular basis of N-end rule recognition', *Mol Cell*, vol. 32, no. 3, pp. 406-14.

Wang, L, Mao, X, Ju, D & Xie, Y 2004, 'Rpn4 is a physiological substrate of the Ubr2 ubiquitin ligase', *J Biol Chem*, vol. 279, no. 53, pp. 55218-23.

Wang, X, Singh, S, Jung, HY, Yang, G, Jun, S, Sastry, KJ & Park, JI 2013, 'HIV-1 Vpr protein inhibits telomerase activity via the EDD-DDB1-VPRBP E3 ligase complex', *J Biol Chem*, vol. 288, no. 22, pp. 15474-80.

Wang, Z, Liu, P, Inuzuka, H & Wei, W 2014, 'Roles of F-box proteins in cancer', *Nat Rev Cancer*, vol. 14, no. 4, pp. 233-47.

Wauer, T & Komander, D 2013, 'Structure of the human Parkin ligase domain in an autoinhibited state', *EMBO J*, vol. 32, no. 15, pp. 2099-112.

Waxman, L, Fagan, JM & Goldberg, AL 1987, 'Demonstration of two distinct high molecular weight proteases in rabbit reticulocytes, one of which degrades ubiquitin conjugates', *J Biol Chem*, vol. 262, no. 6, pp. 2451-7.

Welcker, M, Singer, J, Loeb, KR, Grim, J, Bloecher, A, Gurien-West, M, Clurman, BE & Roberts, JM 2003, 'Multisite phosphorylation by Cdk2 and GSK3 controls cyclin E degradation', *Mol Cell*, vol. 12, no. 2, pp. 381-92.

Wenzel, DM, Lissounov, A, Brzovic, PS & Klevit, RE 2011, 'UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids', *Nature*, vol. 474, no. 7349, pp. 105-8.

Wenzel, DM, Stoll, KE & Klevit, RE 2011, 'E2s: structurally economical and functionally replete', *Biochem J*, vol. 433, no. 1, pp. 31-42.

Wertz, IE, Kusam, S, Lam, C, Okamoto, T, Sandoval, W, Anderson, DJ, Helgason, E, Ernst, JA, Eby, M, Liu, J, Belmont, LD, Kaminker, JS, O'Rourke, KM, Pujara, K, Kohli, PB, Johnson, AR, Chiu, ML, Lill, JR, Jackson, PK, Fairbrother, WJ, Seshagiri, S, Ludlam, MJ, Leong, KG, Dueber, EC, Maecker, H, Huang, DC & Dixit, VM 2011, 'Sensitivity to antitubulin chemotherapeutics is regulated by MCL1 and FBW7', *Nature*, vol. 471, no. 7336, pp. 110-4.

White, EA, Munger, K & Howley, PM 2016, 'High-Risk Human Papillomavirus E7 Proteins Target PTPN14 for Degradation', *MBio*, vol. 7, no. 5.

White, EA, Sowa, ME, Tan, MJ, Jeudy, S, Hayes, SD, Santha, S, Munger, K, Harper, JW & Howley, PM 2012, 'Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses', *Proc Natl Acad Sci U S A*, vol. 109, no. 5, pp. E260-7.

Wiesner, S, Ogunjimi, AA, Wang, HR, Rotin, D, Sicheri, F, Wrana, JL & Forman-Kay, JD 2007, 'Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain', *Cell*, vol. 130, no. 4, pp. 651-62.

Winn, MD, Ballard, CC, Cowtan, KD, Dodson, EJ, Emsley, P, Evans, PR, Keegan, RM, Krissinel, EB, Leslie, AG, McCoy, A, McNicholas, SJ, Murshudov, GN, Pannu, NS, Potterton, EA, Powell, HR, Read, RJ, Vagin, A & Wilson, KS 2011, 'Overview of the CCP4 suite and current developments', *Acta Crystallogr D Biol Crystallogr*, vol. 67, no. Pt 4, pp. 235-42.

Winn, MD, Murshudov, GN & Papiz, MZ 2003, 'Macromolecular TLS refinement in REFMAC at moderate resolutions', *Methods Enzymol*, vol. 374, pp. 300-21.

Wu, K, Chen, A & Pan, ZQ 2000, 'Conjugation of Nedd8 to CUL1 enhances the ability of the ROC1-CUL1 complex to promote ubiquitin polymerization', *J Biol Chem*, vol. 275, no. 41, pp. 32317-24.

Xia, Z, Turner, GC, Hwang, CS, Byrd, C & Varshavsky, A 2008, 'Amino acids induce peptide uptake via accelerated degradation of CUP9, the transcriptional repressor of the PTR2 peptide transporter', *J Biol Chem*, vol. 283, no. 43, pp. 28958-68.

Xia, Z, Webster, A, Du, F, Piatkov, K, Ghislain, M & Varshavsky, A 2008, 'Substrate-binding sites of UBR1, the ubiquitin ligase of the N-end rule pathway', *J Biol Chem*, vol. 283, no. 35, pp. 24011-28.

Xie, J, Kozlov, G & Gehring, K 2014, 'The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins', *Biochim Biophys Acta*, vol. 1839, no. 11, pp. 1062-8.

Xie, Y & Varshavsky, A 1999, 'The E2-E3 interaction in the N-end rule pathway: the RING-H2 finger of E3 is required for the synthesis of multiubiquitin chain', *EMBO J*, vol. 18, no. 23, pp. 6832-44.

Xu, P, Duong, DM, Seyfried, NT, Cheng, D, Xie, Y, Robert, J, Rush, J, Hochstrasser, M, Finley, D & Peng, J 2009, 'Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation', *Cell*, vol. 137, no. 1, pp. 133-45.

Xue, J, Chi, Y, Chen, Y, Huang, S, Ye, X, Niu, J, Wang, W, Pfeffer, LM, Shao, ZM, Wu, ZH & Wu, J 2016, 'MiRNA-621 sensitizes breast cancer to chemotherapy by suppressing FBXO11 and enhancing p53 activity', *Oncogene*, vol. 35, no. 4, pp. 448-58.

Yamano, K & Youle, RJ 2013, 'PINK1 is degraded through the N-end rule pathway', *Autophagy*, vol. 9, no. 11, pp. 1758-69.

Yang, CH, Pfeffer, SR, Sims, M, Yue, J, Wang, Y, Linga, VG, Paulus, E, Davidoff, AM & Pfeffer, LM 2015, 'The oncogenic microRNA-21 inhibits the tumor suppressive activity of FBXO11 to promote tumorigenesis', *J Biol Chem*, vol. 290, no. 10, pp. 6037-46.

Yang, F, Cheng, Y, An, JY, Kwon, YT, Eckardt, S, Leu, NA, McLaughlin, KJ & Wang, PJ 2010, 'The ubiquitin ligase Ubr2, a recognition E3 component of the N-end rule pathway, stabilizes Tex19.1 during spermatogenesis', *PLoS One*, vol. 5, no. 11, p. e14017.

Yang, M, Jiang, N, Cao, QW, Ma, MQ & Sun, Q 2016, 'The E3 ligase UBR5 regulates gastric cancer cell growth by destabilizing the tumor suppressor GKN1', *Biochem Biophys Res Commun*, vol. 478, no. 4, pp. 1624-9.

Yoshida, M, Yoshida, K, Kozlov, G, Lim, NS, De Crescenzo, G, Pang, Z, Berlanga, JJ, Kahvejian, A, Gehring, K, Wing, SS & Sonenberg, N 2006, 'Poly(A) binding protein (PABP) homeostasis is mediated by the stability of its inhibitor, Paip2', *EMBO J*, vol. 25, no. 9, pp. 1934-44.

Yu, G, Rosenberg, JN, Betenbaugh, MJ & Oyler, GA 2015, 'Pac-Man for biotechnology: co-opting degrons for targeted protein degradation to control and alter cell function', *Curr Opin Biotechnol*, vol. 36, pp. 199-204.

Zanet, J, Benrabah, E, Li, T, Pelissier-Monier, A, Chanut-Delalande, H, Ronsin, B, Bellen, HJ, Payre, F & Plaza, S 2015, 'Pri sORF peptides induce selective proteasome-mediated protein processing', *Science*, vol. 349, no. 6254, pp. 1356-8.

Zenker, M, Mayerle, J, Lerch, MM, Tagariello, A, Zerres, K, Durie, PR, Beier, M, Hulskamp, G, Guzman, C, Rehder, H, Beemer, FA, Hamel, B, Vanlieferinghen, P, Gershoni-Baruch, R, Vieira, MW, Dunic, M, Auslender, R, Gil-da-Silva-Lopes, VL, Steinlicht, S, Rauh, M, Shalev, SA, Thiel, C, Ekici, AB, Winterpacht, A, Kwon, YT, Varshavsky, A & Reis, A 2005, 'Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome)', *Nat Genet*, vol. 37, no. 12, pp. 1345-50.

Zenker, M, Mayerle, J, Reis, A & Lerch, MM 2006, 'Genetic basis and pancreatic biology of Johanson-Blizzard syndrome', *Endocrinol Metab Clin North Am*, vol. 35, no. 2, pp. 243-53, vii-viii.

Zhang, L, Fairall, L, Goult, BT, Calkin, AC, Hong, C, Millard, CJ, Tontonoz, P & Schwabe, JW 2011, 'The IDOL-UBE2D complex mediates sterol-dependent degradation of the LDL receptor', *Genes Dev*, vol. 25, no. 12, pp. 1262-74.

Zhang, T, Cronshaw, J, Kanu, N, Snijders, AP & Behrens, A 2014, 'UBR5-mediated ubiquitination of ATMIN is required for ionizing radiation-induced ATM signaling and function', *Proc Natl Acad Sci U S A*, vol. 111, no. 33, pp. 12091-6.

Zhao, C, Wang, L, Ma, X, Zhu, W, Yao, L, Cui, Y, Liu, Y, Li, J, Liang, X, Sun, Y, Li, L & Chen, YH 2015, 'Cardiac Nav 1.5 is modulated by ubiquitin protein ligase E3 component n-recognin UBR3 and 6', *J Cell Mol Med*, vol. 19, no. 9, pp. 2143-52.

Zheng, H, Shen, M, Zha, YL, Li, W, Wei, Y, Blanco, MA, Ren, G, Zhou, T, Storz, P, Wang, HY & Kang, Y 2014, 'PKD1 phosphorylation-dependent degradation of SNAIL

by SCF-FBXO11 regulates epithelial-mesenchymal transition and metastasis', *Cancer Cell*, vol. 26, no. 3, pp. 358-73.

Zhuang, M, Calabrese, MF, Liu, J, Waddell, MB, Nourse, A, Hammel, M, Miller, DJ, Walden, H, Duda, DM, Seyedin, SN, Hoggard, T, Harper, JW, White, KP & Schulman, BA 2009, 'Structures of SPOP-substrate complexes: insights into molecular architectures of BTB-Cul3 ubiquitin ligases', *Mol Cell*, vol. 36, no. 1, pp. 39-50.

Zimmerman, SW, Yi, YJ, Sutovsky, M, van Leeuwen, FW, Conant, G & Sutovsky, P 2014, 'Identification and characterization of RING-finger ubiquitin ligase UBR7 in mammalian spermatozoa', *Cell Tissue Res*, vol. 356, no. 1, pp. 261-78.