Elucidating ubiquitin recognition by the HECT-type ubiquitin ligase HUWE1

Studien zur Ubiquitinerkennung durch die HECT-Typus Ubiquitinligase HUWE1

Doctoral thesis

for a doctoral degree at the Graduate School of Life Sciences,

Julius-Maximilians-Universität Würzburg,

Section Biomedicine

submitted by

Rahul Mony Nair

From Kerala, India

Würzburg 2020

Submitted on: ……..

Office stamp

Members of the Thesis Committee

Chairperson: Prof. Christian Janzen

Primary Supervisor: Dr. Sonja Lorenz

Supervisor (Second): Prof. Antje Gohla

Supervisor (Third): Prof. Nikita Popov

Date of Public Defence: ……

Date of Receipt of Certificates: …….

Dedicated to my INDIA

"We are what our thoughts have made us so take care about what you think. Words are secondary, thoughts live; they travel far....."

Swami Vivekananda

Summary

The small protein modifier ubiquitin is at the heart of an immensely versatile posttranslational modification system that orchestrates countless physiological and disease-associated cellular processes. Key to this versatility are the manifold modifications that can be assembled from ubiquitin “building blocks” and are associated with specific functional outcomes for the modified substrates. In particular, ubiquitin molecules can form polymeric chains of distinct lengths and linkage types that give rise to distinct chain conformations, thereby providing recognition sites for specific signaling receptors/effectors. The class of E3 enzymes (ubiquitin ligases) provides critical specificity determinants in ubiquitin linkage formation; it is therefore crucial to unravel precisely how E3 enzymes operate in order to understand the structural basis of ubiquitin signaling and exploit these insights for therapeutic benefit.

Overexpression and deregulation of the HECT-type ubiquitin ligase HUWE1 is implicated in several different cancer types and neurodegenerative disorders. It is largely unknown which factors control the ubiquitin modifications formed by HUWE1, how the catalytic HECT domain interacts with functionally distinct ubiquitin molecules (donor, acceptor and regulatory ubiquitin molecules) and which conformational transitions enable these interactions during ubiquitin chain formation.

One aim of this study was to structurally elucidate the recognition of donor ubiquitin by the HECT domain of HUWE1. To this end I utilized a ubiquitin activity-based probe to reconstitute a proxy for a donor ubiquitin-linked conjugate of the HECT domain of HUWE1 and determined its structure by X-ray crystallography. This structure reveals that the donor ubiquitin binds to the C-lobe of HUWE1 in the same way as NEDD4-type ligases, corroborating the idea that HECT ligases utilize a conserved mode of donor ubiquitin recognition. independent of their linkage and substrate specificities. With the help of biochemical analyses, I also validated specific features of the structure, in particular the positioning of the C-terminal tail of the ligase, which was known to be critical for activity. In the newly determined structure, which reflects an “L-shaped”, active state of the HECT domain, this tail is fully resolved and coordinated at the N-lobe-C-lobe interface. I defined residues that are critical for this coordination and showed that they are also essential for

the activity of HUWE1, including auto-ubiquitination, free ubiquitin chain formation, and substrate ubiquitination.

Furthermore, I discovered that the N-lobe of HUWE1 harbors a ubiquitin-binding exosite similar to NEDD4-type ligases and E6AP. My *in-vitro* activity and binding assays show that HUWE1 uses the exosite for isopeptide bond formation, but that it is dispensable for thioester bond formation. The binding assays further show that the donor ubiquitin loaded HECT domain binds an additional ubiquitin molecule at the exosite more tightly than the *apo* HECT domain, which possibly suggests allosteric communication between the two sites.

Finally, I showed that the ubiquitin activity-based probe (ubiquitin-propargylamine) can label the catalytic cysteine of HUWE1 and NEDD4-type with close to quantitative turn-over, while it does not react with the HECT domain of the evolutionarily more divergent E6AP. The determinants underlying these differential reactivities remain to be explored.

Taken, together my results significantly enhance our mechanistic understanding of the catalytic domain of HUWE1 and pinpoint linchpins for therapeutic interventions with the activity of this disease-relevant enzyme.

Zusammenfassung

Der kleine Proteinmodifikator Ubiquitin ist das Herzstück eines immens vielseitigen posttranslationalen Modifikationssystems, das unzählige physiologische und krankheitsassoziierte zelluläre Prozesse orchestriert. Der Schlüssel zu dieser Vielseitigkeit sind die vielfältigen Modifikationen, die sich aus Ubiquitin-"Bausteinen" zusammensetzen lassen und mit spezifischen funktionellen Ergebnissen für die modifizierten Substrate verbunden sind. Insbesondere können Ubiquitin-Moleküle Ketten unterschiedlicher Länge und Verknüpfungstypen bilden, die zu unterschiedlichen Kettenkonformationen führen und dadurch Erkennungsstellen für spezifische Signalrezeptoren/-effektoren bieten. Die Klasse der E3-Enzyme (Ubiquitin-Ligasen) liefert kritische Spezifitätsdeterminanten für die Bildung von Ubiquitin-Bindungen; daher ist es entscheidend, die genaue Funktionsweise der E3-Enzyme zu entschlüsseln, um die strukturelle Grundlage der Ubiquitin-Signalisierung zu verstehen und diese Erkenntnisse für therapeutische Anwendungen zu nutzen.

Die Überexpression und Deregulierung der Ubiquitin-Ligase HUWE1 aus der Klasse der HECT-E3-Ligasen ist an mehreren verschiedenen Krebsarten und neurodegenerativen Erkrankungen beteiligt. Es ist weitgehend unbekannt, welche Faktoren durch die von HUWE1 gebildeten Ubiquitin-Modifikationen kontrolliert werden, wie die katalytische HECT-Domäne mit funktionell unterschiedlichen Ubiquitin-Molekülen (Donor-, Akzeptor- und regulatorische Ubiquitin-Moleküle) interagiert und welche Konformationsübergänge diese Interaktionen während der Ubiquitin-Kettenbildung ermöglichen.

Ein Ziel dieser Studie war es, die Erkennung des Donor-Ubiquitin-Moleküls durch die HECT-Domäne von HUWE1 strukturell aufzuklären. Zu diesem Zweck verwendete ich eine *´ubiquitin activity-based probe´*, um ein Konjugat der HUWE1-HECT-Domäne mit einem Donor-Ubiquitin-Molekül zu rekonstitutieren und die Struktur mittels Röntgenkristallographie zu bestimmen. Diese Struktur zeigte, dass das Donor-Ubiquitin-Molekül auf die gleiche Weise an den *C-lobe* von HUWE1 bindet wie die Klasse der NEDD4-Ligasen, was die Idee bestätigt, dass HECT-Ligasen einen vergleichbaren Mechanismus bei der Donor-Ubiquitin-Erkennung verwenden, unabhängig von ihrer Bindung und Substratspezifität. Mit Hilfe biochemischer Analysen validierte ich auch spezifische Merkmale der Struktur, insbesondere die Positionierung des *C-terminal tail*

der Ligase, der entscheidend für die Aktivität ist. In der neu bestimmten Struktur, die einen "L-förmigen", aktiven Zustand der HECT-Domäne widerspiegelt, ist der *C-terminal tail* an der Grenzfläche von *N-lobe* und *C-lobe* vollständig aufgelöst und koordiniert. Ich konnte Seitenketten festmachen, die für diese Koordination kritisch sind, und habe gezeigt, dass sie auch für die Aktivität von HUWE1 wesentlich sind, einschließlich der Auto-Ubiquitinierung, der freien Ubiquitin-Kettenbildung und der Substrat-Ubiquitinierung.

Darüber hinaus entdeckte ich, dass der *N-lobe* von HUWE1 eine Ubiquitin-bindende *exosite* aufweist, ähnlich wie für die Klasse der NEDD4-Ligasen und E6AP. Meine *in vitro* Aktivitäts- und Bindungstests ergaben, dass HUWE1 die *exosite* für die Bildung von Isopeptidbindungen verwendet, diese aber für die Bildung von Thioesterbindungen entbehrlich ist. Die Bindungstests zeigten ferner, dass die Donor-Ubiquitin-beladene HECT-Domäne ein zusätzliches Ubiquitin-Molekül an der *exosite* stärker bindet als die Apo-HECT-Domäne, was möglicherweise auf eine allosterische Kommunikation zwischen den beiden Ubiquitin-Bindestellen hindeutet.

Schließlich zeigte ich, dass die *´ubiquitin activity-based probe´* (Ubiquitin-Propargylamin) das katalytische Cystein von HUWE1 und NEDD4 mit nahezu quantitativem Umsatz markieren kann, während es nicht mit der HECT-Domäne des evolutionär stärker divergierenden E6AP reagiert. Die Faktoren, die diesen unterschiedlichen Reaktivitäten zugrunde liegen, müssen noch erforscht werden.

Zusammengenommen verbessern meine Ergebnisse unser mechanistisches Verständnis der katalytischen Domäne von HUWE1 und geben uns die Dreh- und Angelpunkte für therapeutische Interventionen mit der Aktivität dieses krankheitsrelevanten Enzyms an die Hand.

Table of contents

Summary iv

Zusammenfassung vi

1 Introduction 1

1.1 Post-translational protein modifications 1

1.2 The ubiquitin system 2

1.2.1 Role in physiology and disease 2

1.2.2 Enzymatic machinery 7

1.3 Homologous to E6AP C-terminus-type (HECT) ligases 11

1.3.1 Domain architecture and classification 11

1.3.2 Structural mechanism of the HECT domain 12

1.3.3 HUWE1 16

1.3.3.1 Pathophysiological roles 16

1.3.3.2 Structural mechanism 19

1.3.4 Manipulating HECT-type ligases 20

1.3.4.1 Small-molecule ligands 20

1.3.4.2 Ubiquitin-activated probes 21

1.4 Aims of the thesis 23

2 Materials 24

2.1 Primers 24

2.2 Bacterial strains and expression constructs 25

2.3 Bioreagents, kits and enzymes 27

2.4 Chemicals 27

2.5 Crystallization screens 29

2.6 Specialized consumables 30

A list of specialized concumables is provided in Table 9. 30

2.7 Relevant scientific equipment 30

A list of relevant scientific equipment is provided in Table 10 ..... *30*

2.8 Software, servers and databases *31*

3 Methods 33

3.1 Protein production *33*

3.1.1 Molecular biology 33

3.1.1.1 Preparation plasmid DNA and transformation into E. coli 33

3.1.1.2 Cloning techniques 34

3.1.1.2.1 Site-directed mutagenesis 34

3.1.1.2.2 Sub-cloning 34

3.1.2 Protein Expression and purification 34

3.1.2.1 General culturing routine 34

3.1.2.2 Protein purification 35

3.1.2.2.1 HUWE1 HECT domain and variants 35

3.1.2.2.2 E6AP HECT domain and variants 36

3.1.2.2.3 NEDD4 HECT domain and variants 37

3.1.2.2.4 Ubiquitin and variants 37

3.1.2.2.5 Ubiquitin-PA 38

3.1.2.2.6 HUWE1 HECT domain-Ub-PA conjugate 39

3.1.2.2.7 K48 linked-di ubiquitin HUWE1 HECT complex 39

3.1.2.2.8 UBE2L3 40

3.1.2.2.9 Accessory proteins 40

3.1.3 Protein concentration determination 40

3.2 Biochemical and biophysical methods *41*

3.2.1 Electrophoretic methods 41

3.2.1.1 SDS-PAGE 41

3.2.1.2 Immunoblotting 41

3.2.2 Thermoflour assay 42

3.2.3 Circular dichroism 43

3.2.4 Fluorescence polarization 43

3.2.5 In vitro activity assays 44

3.2.5.1 Isopeptide bond formation 44

3.2.5.2 Thioester bond formation 45

3.2.6 Ubiquitin-PA labelling assay 45
3.2.7 Mass spectrometry 45
3.3 X-ray crystallography 46
3.3.1 Protein crystallization and data collection 46
3.3.2 Structure determination and refinement 46
4 Results and Discussion 48
4.1 Ubiquitin-PA as a tool to study ubiquitin recognition by the HECT domain of HUWE1 48
4.1.1 Preparation of ubiquitin-PA 48
4.1.2 Preparation of the HECT domains of HUWE1, NEDD4 and E6AP 51
4.1.3 HECT ligase profiling with Ubiquitin-PA 54
4.2 Donor ubiquitin recognition by the HECT domain of HUWE1 58
4.2.1 Crystal structure of HUWE1 HECT domain-Ub-PA conjugate 60
4.2.2 Structure-based design of mutations to study the coordination of the C-tail 64
4.3 Using ubiquitin-PA to illuminate the mechanistic details of HUWE1 69
4.3.1 Ub-PA labelling of structure-guided HUWE1 HECT domain variants 69
4.4 Interrogating additional ubiquitin binding sites in the HECT domain of HUWE1 71
4.4.1 Conformational dynamics of the HECT domain of HUWE1 affect ubiquitin binding 71
4.4.2 Characterization of allosteric ubiquitin binding sites in the HECT domain of HUWE1 74
4.5 Acceptor ubiquitin: The key to linkage specificity 77
5 Conclusions 79
5.1 Recognition of donor ubiquitin by the HECT domain of HUWE1 79
5.2 Acceptor ubiquitin: Key to understand linkage specificity 80
5.3 The non-covalent ubiquitin binding exosite in HUWE1 82
5.4 Differential reactivity of HECTs towards activity-based probes 87
5.5 Activity-based probes as precursors of HECT ligase-directed drugs? 88

6 References 89

7 Appendix 100

7.1 Abbreviations *100*

7.2 List of tables *105*

7.3 List of figures *106*

7.4 Curriculum vitae ..... ***Error! Bookmark not defined.***

7.5 Acknowledgements *109*

7.6 Affidavit *111*

1 INTRODUCTION

1.1 Post-translational protein modifications

Proteins are critical mediators of the homeostasis of eukaryotic cells through a diverse number of processes, which need to be precisely regulated in space and time. Post-translational modifications (PTMs) of proteins - the covalent attachment of small chemical groups or protein modifiers to target proteins - constitute a versatile and highly dynamic platform for the regulation of protein functions. PTMs influence the properties of target proteins at different levels, including protein concentration/turnover/stability, localization, structure/conformational dynamics, interactions and enzymatic activities [1]. There are an estimated 200 types of PTMs [2], including phosphorylation, ubiquitination, methylation, acetylation and proline hydroxylation, which also crosstalk with one another [1].

In contrast to modifying chemical groups, ubiquitin and ubiquitin-like proteins (Ubls) are small protein modifiers and open a particularly large range of signaling functions through the formation of protein-protein interactions. Moreover, ubiquitin and Ubls can modify target proteins in different ways, including the modification of a single site (mono-ubiquitination), several individual sites (multi-mono-ubiquitination), or in the form of chains (polyubiquitination), which can in turn have many several different linkage types and topologies [3]. Remarkably, ubiquitin can itself be post-translationally modified by acetylation and phosphorylation at Ser, Thr or Tyr residues, respectively [4]. The phosphorylation of ubiquitin at Ser 65 by PTEN-induced protein kinase 1 (PINK1) [5] recruits parkin-an E3 ligase mutated in the context of autosomal recessive juvenile Parkinsonism-and autophagy adaptors to initiate mitophagy [4].

A detailed functional, mechanistic and structural understanding of the catalytic machineries and specificity determinants of PTMs is essential to uncover and therapeutically target the molecular bases of human pathologies, such as cancer, inflammatory and neurodegenerative diseases. In particular, my thesis work deals with the structural mechanisms of the ubiquitin system and the overall question of how catalytic efficiency and specificity is achieved in this complex signaling system.

1.2 The ubiquitin system

1.2.1 Role in physiology and disease

Ubiquitin is a small (8.5 kDa), evolutionary conserved post-translational modifier that controls a vast number of biological processes, such as protein degradation, DNA repair, endocytosis, autophagy, transcription, immunity and inflammation [6]. The conjugation of ubiquitin to a target protein typically occurs at a primary amino group of the target (Lys or the N-terminus), but it has been recently been shown that additional residues, such as Thr, Ser and Tyr can be modified, as well [4]. The C-terminal carboxyl group of ubiquitin is linked to the aforementioned residues by way of an isopeptide, peptide, or ester bond, respectively. Additionally, ubiquitin can form chains through one of its eight primary amino groups - Lys 6, Lys 11, Lys 27, Lys 29, Lys 33, Lys 48 and Lys 63 or the N-terminus, [6] (Figure 1).

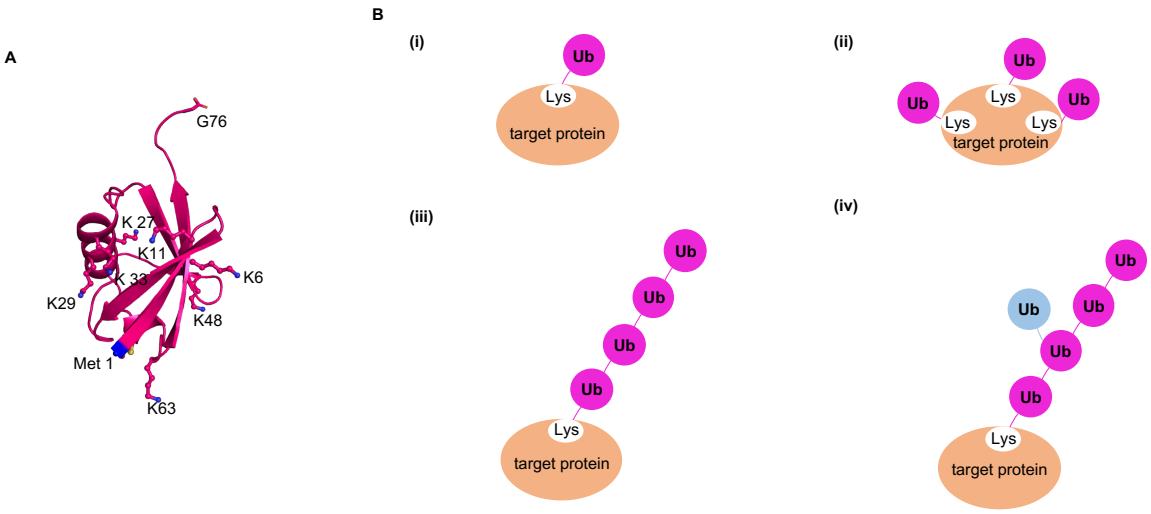


Figure 1: Structural features of ubiquitin and modes of ubiquitination

(A) Crystal structure of ubiquitin shown in cartoon representation (PDB: 1UBQ). The side chains of lysine residues and Met 1 are shown as ball-and-stick models. The C-terminal G76 is shown in stick representation. **(B) (i)** Single-site conjugation of a ubiquitin (Ub) molecule on a target **(ii)** multi-monoubiquitination of a target **(iii)** homotypic and **(iv)** heterotypic, branched ubiquitin chains on a target protein.

The linkage type determines which surfaces are exposed in a given ubiquitin chain and thereby determines the interaction with downstream effectors. Besides the different linkage types, ubiquitin chains can contain more than one type (mixed chains) or branches, thus increasing the complexity of chain topologies and associated signaling functions. The cell, therefore, uses the entire repertoire of linkage types and topological/structural diversity to generate functional diversity [4].

Ubiquitin was first discovered by Gideon Goldstein in his search for thymopoietin [7]. However, the physiological role of ubiquitin was not known until the 1980s when Avram Hershko's, Aaron Ciechanover's and Irwin Rose's work identified two stable proteins in reticulocyte fractions. One was a small heat-stable protein, then named "ATP-dependent proteolysis factor 1" (now known as ubiquitin) and the second a high-molecular weight protein, then named "ATP-dependent proteolysis factor 2" (now known as the 26S proteasome) [8,9]. This and the following seminal work led to the discovery of the ubiquitin-proteasome system (UPS), which was recognized with the Nobel Prize in Chemistry in 2004. Further groundbreaking work from the laboratories of Alexander Varshavsky and Cecile Pickart showed that Lys 48-linked ubiquitin chains, the most prominent chain type in higher eukaryotes [10], can direct targets to proteasomal degradation [11,12].

To date, much detailed insight into the molecular basis and functional consequences of individual ubiquitin modifications has been generated: Distinct ligases, such as the SKP, Cullin, F-box containing complex (SCF), glycoprotein 78 (GP78), or E6-associated protein (E6AP) were identified as the enzymes responsible for building Lys 48-linked chains in the cell [6]. Lys 63-linked chains, the second most abundant chain type in eukaryotic cells [4] is implicated in non-degradative functions, e.g., the endocytosis of membrane receptors, such as G-protein coupled receptors (GPCRs) [6] and nuclear factor -κappa B signaling (NFκB) [13,14], lysosomal degradation [15] and DNA repair [16] Notably, the latter is also regulated by mono-ubiquitination, as illustrated by the FA-core complex substrates FANCD2 and FANCI, whose modification promotes the recruitment of the FAN1 nuclease in DNA inter-strand cross-link repair [17]. Lys 27-linked chains, as catalyzed by the ubiquitin ligase RNF168, promote the formation of histone 2A, which is crucial for the recruitment of DNA damage response mediators in DNA double-strand breaks [18]. Lys 11-linked chains, assembled by the anaphase-promoting complex/cyclosome (APC/C) during mitosis drive target proteins to proteasomal degradation, thereby promoting mitotic progression [19]. Lys

29-linked chains, assembled by the ubiquitin ligase SMAD ubiquitylation regulatory factor 1 (SMURF1), negatively regulates WNT-signaling by modifying the WNT/β-catenin signaling adaptor protein axin [20]. The attachment of Lys 33-linked chains to the F-actin regulatory protein coronin 7 by the ubiquitin ligase CUL3–KLHL20 inhibits the depolymerization of F-actin during post-Golgi trafficking [21]. LUBAC, a multi-subunit ubiquitin ligase, utilizes the N-terminal methionine (Met-1) to assemble 'linear' ubiquitin chains during NF-κB signaling [22]. The cellular function of Lys 6-linked chains is not fully understood, yet, a recent study showed that the ubiquitin ligase HUWE1 promotes Lys 6 chains on mitofusin 2 (MFN2), a protein important for mitochondrial biogenesis [23].

Given the importance of ubiquitination for cellular and tissue homeostasis, it is not surprising that dysregulation of the ubiquitin system is associated with a plethora of pathophysiological conditions. For example, human papilloma virus-induced cervical cancer is caused by the interaction of the viral E6 protein with the ubiquitin ligase E6AP, which is thus reprogrammed to target P53 for proteasomal degradation by Lys 48-linked ubiquitin chains [24]. Xeroderma pigmentosum E, an autosomal recessive disorder which can lead to tumorigenesis, is caused by UV-damage to a DNA-binding ubiquitin ligase complex (UV-DDB), which under normal circumstances recruits nucleotide excision repair factors to damaged DNA through mono-ubiquitination of histone H2A [25]. Spontaneous mutations in the *Sharpin* gene, which encodes a component of the LUBAC ubiquitin ligase complex, a main regulator of NEMO, was shown to cause chronic proliferative dermatitis, inflammation and defects in lymphoid organ development in mouse models [26]. Mutations in PTEN-induced putative phosphatase 1 (*PINK1*), which encodes a well-characterized binding partner of the ubiquitin ligase parkin, causes early-onset familial Parkinson's disease [27]. De-ubiquitinating enzymes (DUBs), which reverse the function of ubiquitin ligase, are also associated with numerous human diseases. One such example is CYLD, the inactivation or downregulation of which results in increased invasion and proliferation of B-cell lymphoma through alteration of the localization pattern of NF-κB [28].

These and multiple other examples suggest that the ubiquitin system provides an interesting target for therapeutic efforts. Despite considerable effort in academia and pharmaceutical industry to target components of the ubiquitin system with small-molecule effectors, however, progress has been rather slow [29]. Yet, there are several impressive examples that highlight the potential that manipulating the ubiquitin-proteasome system holds for

cancer therapy: Most prominently, proteasome-inhibitors have revolutionized the treatment of hematological disorders, including multiple myeloma and mantle cell lymphoma [29]. The dipeptide boronic acid compound known initially as PS341, later Bortezomib (Velcade), was the first reversible proteasome inhibitor to gain FDA-approval [30]. With an EC_{50}-value of 0.6 nM, Bortezomib shows remarkable regression in patients suffering from relapsed multiple myeloma and mantle cell lymphoma [29]. A second-generation compound, Carfilzomib (PR-171, Kyprolis), inhibits the proteasome irreversibly and with higher efficiency than Bortezomib [31]; additional proteasome inhibitors with improved bioavailability have recently been generated, such as MLN9708 and CEP-18770 (Delanzomib) [29]. Collectively, these inhibitors can sensitize cancer cells and induce cancer-cell specific apoptosis [32]. Despite the remarkable success of these inhibitors, some of them are associated with dose-limiting toxicity, which required a reduction of the dose or discontinuation of the drug [33].

Efforts have also been invested into targeting components of the ubiquitination and deubiquitination machinery upstream of the proteasome. In particular, DUBs and the ubiquitin-activating (E1) enzyme have proven amenable to manipulation by small molecules due to their defined active sites [29]. Yet no such compounds have been approved for clinical use. The large group of ubiquitin ligases, which make interesting therapeutic targets due to their signaling specificity have turned out to be particularly challenging to access therapeutically. One reason is that many ligases lack a defined active site (section 1.3.1.) and instead act through the formation of protein-protein interfaces, which are notoriously difficult to target [29]. Nevertheless, some ligase-directed compounds reached clinical trials. For example, cis-imidazoline analogs known as nutlins inhibit the interaction of the ubiquitin ligase MDM2 with a key substrate, the tumor suppressor P53, thus causing a strong accumulation of wild-type P53; unfortunately, however, they do not work for mutated P53, as required for cancer therapy [29].

A very interesting, recent development is the field of protein-targeting chimeric molecules (PROTACs) [34]. These are hetero-bifunctional small molecules that contain a ligand for a ubiquitin ligase at one end and a ligand for a specific protein of interest on the other, connected by a short linker (Figure 2).

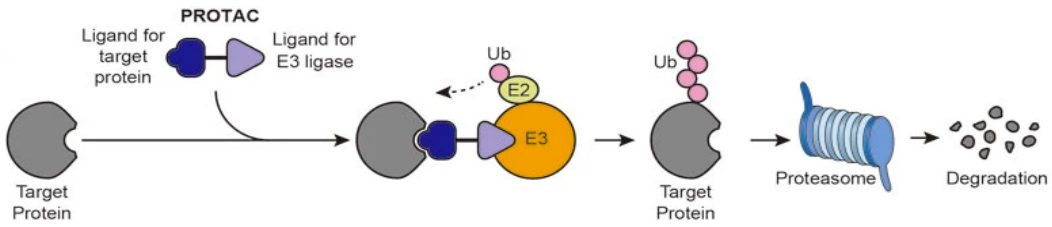


Figure 2: PROTAC-mediated target protein degradation

The bifunctional molecule PROTAC simultaneously binds to its target protein and a specific E3. The target is then ubiquitinated by the E3 and directed to degradation by the proteasome (figure taken from [29]).

The first generation PROTAC, PROTAC-1, was developed by Craig Crews and colleagues. As a proof-of-concept, they used IκB phospho-peptide, a ligand that can recognize the ubiquitin protein ligase $SCF^{\beta\text{-}TRCP}$ at one end and a covalent binder of the target protein methionine aminopeptidase-2, ovalicin on the other end was generated. The degradation of methionine aminopeptidase-2 in a PROTAC-1 dependent manner was shown [29]. The potential of a PROTAC-based strategy lies in making thus-far considered 'undruggable' oncoproteins accessible to therapeutic use. Notably, a ligand for the pathogenic protein of interest still needs to be found, but it is sufficient to identify a binder rather than a compound that interferes with the protein activity. Given the size of the bifunctional PROTAC bioavailability and pharmacokinetics may be harder to optimize than for conventional 'mono-functional' drugs and the linker connecting the two parts of the molecule needs to be carefully optimized. The PROTAC strategy also requires the ubiquitin ligase and the target protein to be expressed in the same tissue at the same time, which provides an additional challenge [29].

To accelerate progress in exploiting and harnessing the ubiquitin system for therapeutic use it is of utmost importance to understand the enzymatic mechanisms of ubiquitination and associated specificity factors and regulation principles, which is the overall motivation of my thesis research.

1.2.2 Enzymatic machinery

The ubiquitin cascade usually involves three steps (Figure 3), driven by distinct classes of enzymes. Initially, a ubiquitin-activating enzyme (E1) utilizes ATP for the activation of ubiquitin by covalently attaching the C-terminus of ubiquitin to the active-site cysteine of the E1. Next the E1 transfers ubiquitin to the active-site cysteine of a ubiquitin-conjugating enzyme (E2) in a trans-thioesterification reaction. Finally, a ubiquitin ligase (E3) mediates the final transfer of ubiquitin from the E2 to a primary amino group of a target protein or another ubiquitin molecule (during ubiquitin chain formation) [35]. The modification on the target protein or chain assembly can be reversed by the action of DUBs, contributing to the maintenance of free ubiquitin within the cell [36].

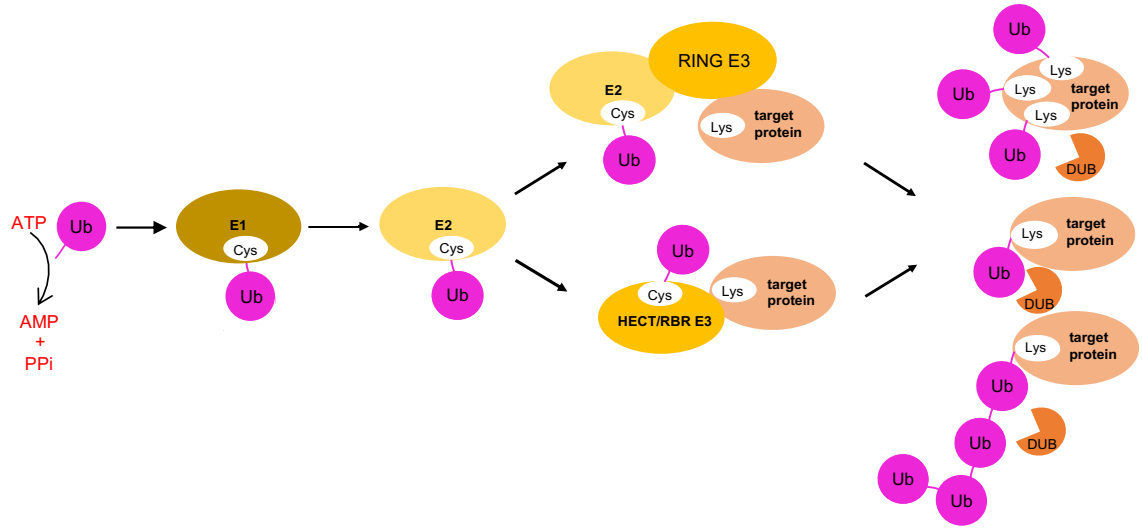


Figure 3: Ubiquitin conjugation machinery

Schematic representation of ubiquitin conjugation. The E1 activates the C-terminus of ubiquitin in an ATP-dependent manner and forms a thioester between E1 and ubiquitin. The activated ubiquitin is transferred to the catalytic cysteine of the E2 and the E3 facilitates the transfer of ubiquitin from the E2 to a target protein either directly for RING ligases (Really Interesting New Gene) or through an E3-ubiquitin intermediate for HECT (Homologous to E6AP C-terminus) or RBR (Ring-between-Ring) ubiquitin protein ligases.

There are two ubiquitin-specific E1s, UBA1 and UBA6 of 110 and 118 kDa [37]. These enzymes contain two heterodimeric Rossmann-type folds, a domain containing the catalytic cysteine (the 'cysteine domain'), and a ubiquitin-fold domain A catalytic cysteine-bearing half-domain and a four-helix bundle are inserted between the adenylation domains; (UFD), which recruits the E2 [37]. In the first step of the catalytic cycle, the E1 binds to ATP-Mg^{2+} and ubiquitin and catalyzes a C-terminal acyl adenylation of ubiquitin. In the next step, the catalytic cysteine on the E1 attacks the ubiquitin-adenosine monophosphate (AMP) complex, yielding a thioester bond between the E1 and ubiquitin. In the last step, the E1 transfers ubiquitin to an E2 through a transthioesterification reaction [37], forming an E2~Ub conjugate.

The E2 family is composed of nearly 40 members with a molecular weight ranging from 14 to 35 kDa that share a conserved catalytic (UBC) domain of ~ 150 residues [35]. Class-1 E2s are composed of a UBC domain only, while class-2 E2s are characterized by an N-terminal region and class-3 E2s by a C-terminal extension; class-4 E2s have extensions on either end of the UBC domain. E1s and E3s interact with E2s through the conserved UBC domain [38-40]. During the catalytic cycle, the activated ubiquitin that is thioester-linked at the active site of the E2 is termed the "donor" ubiquitin. The ubiquitin molecule that performs a nucleophilic attack on the activated donor during ubiquitin chain formation is known as the "acceptor" ubiquitin [41].

E3s mediate target protein recognition in the ubiquitin system and are categorized into three subfamilies based on their architecture and catalytic mechanism: Really Interesting New Gene-type ligases (RING), Homologous to E6AP C-terminus-type ligases (HECT) (see section 1.3) and RING-between RING-type ligases (RBR) (Figure 4). There are nearly 600 annotated human RING-type ligases, a significantly higher number than the E2 enzymes with which they interact [42]. This hierarchical arrangement of the ubiquitin cascade allows the E2 enzymes to specifically transfer ubiquitin to appropriate targets via their interaction with specific RING partners. The eponymous RING domain of RING-type ligases is characterized by cysteine and histidine-rich cross-brace structure, which are coordinated by two zinc ions [43] and provides a platform for protein-protein interactions. Effective ubiquitylation by RING-type ligases requires the E2~Ub conjugate to be orientated in a conformation where the C-terminal Gly 76 of Ub at the E2 active site is primed for

nucleophilic attack by a lysine residue, commonly referred to as a “closed conformation” [41,44-47].

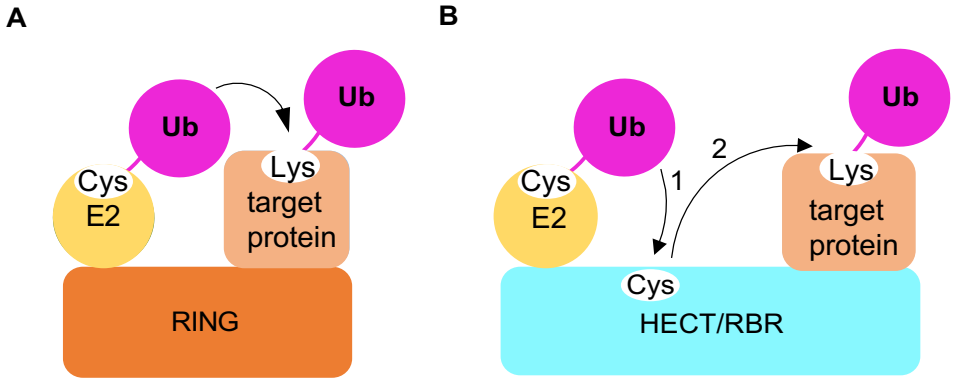


Figure 4: Schematic overview of E3 mechanisms

(**A**) RING-type ligases facilitates the direct transfer of ubiquitin from an E2 to the lysine of a target protein. **(B)** HECT and RBR-type ligases follow a two-step mechanism of ubiquitin transfer by forming a thioester intermediate with ubiquitin before attaching it to a target protein (figure redrawn [48]).

E2~Ub conjugates are generally conformationally malleable in solution, depending on the type of E3 and the absence/presence of an E3 [41,44]. RING domain binding typically promotes a population shift in the dynamic E2~Ub conjugate towards closed conformations which primes the active site for ubiquitin transfer [44,45,49].In contrast, HECT and RBR-type ligases interact with E2-conjugates in an open conformation [50,51], It is interesting to note that E2~Ub conjugates that populate closed conformations in the absence of an E3, such as Ubc13 [47], Ubc1 [52], Ube2S [41], and Cdc34 [46], shows an E3-independent ubiquitin transfer to the substrate. Ubiquitin has also been shown to bind to the backside of several E2s, including RAD6 [53], UBCH5 [54], UBCH6 [55], and UBE2G2 [56], which activates ubiquitin transfer allosterically [57]. It is also noteworthy that not all RING-type ligases have an intrinsic ability to promote ubiquitin transfer but are highly regulated by a diverse range of mechanisms, one example being dimerization with a secondary RING E3 for ligase activity [43]. The largest subset of RING-type E3s are the cullin-RING ligases (CRLs), which exist as multi-subunit protein complexes [58]. Structural studies of a

chemically trapped complex of $CRL1^{\beta\text{-}TRCP}$ with NEDD8 recently showed the mode of ubiquitin transfer to a substrate, as activate by the Neural precursor cell expressed developmentally down-regulated protein 8 (NEDD8). Conformational dynamics within the CRL domains and local structural remodeling induced by NEDD8, primes the neddylated $CRL1^{\beta\text{-}TRCP}$ for the transfer of ubiquitin from the E2 to its phosphorylated IκBα target [59]. The reversible attachment of NEDD8 to CRLs and the sequestration of cullins by CAND1 are some of the ways by which the activity of CRLs can be regulated [60]. As explained above, the linkage specificity of ubiquitin chain formation is a major determinant of the biological outcome of ubiquitination. Notably, in the case of RING-type ligases, the associated E2 typically determines linkage specificity. In contrast, HECT and RBR ligases can inherently determine linkage specificity.

The 28 human HECT-type ligases possess a conserved C-terminal catalytic HECT domain and an extended N-terminal region of varying size [50]. Their architecture, conformational dynamics and catalytic mechanism will be discussed in more detail below (section 1.3).

The 14 RBR-type ligases annotated in the human proteome contain features of both RING and HECT ligases. Similar to HECTs, RBRs follow a two-step mechanism by forming a thioester intermediate with ubiquitin in the first step [61].Like RING ligases, RBR Es contain a domain known as RING1, which resembles a classical RING domain and functions in the transfer of ubiquitin from the E2 to the cysteine in a so-called RING2 domain, which constitutes the catalytic core of the RBR. Furthermore, RBR E3s contain an In-Between-RING (IBR) domain, which provides a platform for the E2-ubiquitin complex to dock onto the RING1 domain.

I am particularly interested in understanding HECT-type ligases, since they have important pathophysiological roles and are poorly understood at a structural and mechanistic level. My research thus focussed on particular aspects in the catalytic cycle of the catalytic HECT domain, which we need to understand at the atomic level to eventually exploit or target HECT ligases for therapeutic benefit.

1.3 Homologous to E6AP C-terminus-type (HECT) ligases

1.3.1 Domain architecture and classification

The first HECT-type ligase identified was the human papilloma virus (HPV) E6-associated protein (E6AP) [62]. The C-terminal HECT domain comprises ~350 residues folded into two lobes, C-lobe and the N-lobe, and a short inter-lobe linker. The HECT domain is typically sufficient to promote ligase activity (auto-ubiquitination and/or free ubiquitin chain formation), but insufficient to recruit substrates. The latter are recruited by the extended N-terminal regions, which also fulfill regulatory functions. For example, in the case of NEDD4-type ligases, an N-terminal C2 domain binds to phospholipids and targets the ligases to the plasma membrane, endosomes and multivesicular bodies [63]. Additional WW domains in the N-terminal extension mediate substrate recognition and regulate the HECT domain. A disruption of WW-HECT domain interactions by phosphorylation in the N-terminal extension of ITCH stabilizes this ligase in an active conformation [64]. Thus, intra- and intermolecular interactions of control HECT ligase activities.

The 28 human HECT-type ligases are divided into subfamilies based on the domain composition of their N-terminal extensions [50] (Figure 5). The best-characterized subfamily are the 9 NEDD4-type ligases, which contain a calcium-dependent lipid binding domain (C2) [65,66], two to four WW domains which recognize proline-rich (PY) motifs, besides the C-terminal HECT domain. NEDD4-type ligases are implicated in a series of human pathologies like cancer, immune response, growth and development, and neurodegeneration [64]. For example, NEDD4 ubiquitinates RAS-related protein 2A (RAP2A), a negative regulator of dendritogenesis, by forming a trimeric complex with TRAF2 and NCK-interacting kinase (TINK) and RAP2A, which then promotes dendrite growth and arborization [67].

The subfamily of HERC-type ligases has six members, which are divided on the basis of their molecular masses, into large HERCs (HERC1 and HERC2) and small HERCs (HERC3, HERC4, HERC5 and HERC6). All HERC ligases are characterized by at least one regulator of chromosome condenstaion1-like (RCC1) domain [64]. This domain interacts with histones and guanine-nucleotide exchange factors [68].

The remaining human HECT ligases show diverse domain compositions of their N-terminal regions are this category as “other HECTs”. This latter family contains the large ligase HUWE1, which is the main focus of this work and will be discussed in detail below (section 1.3.3)

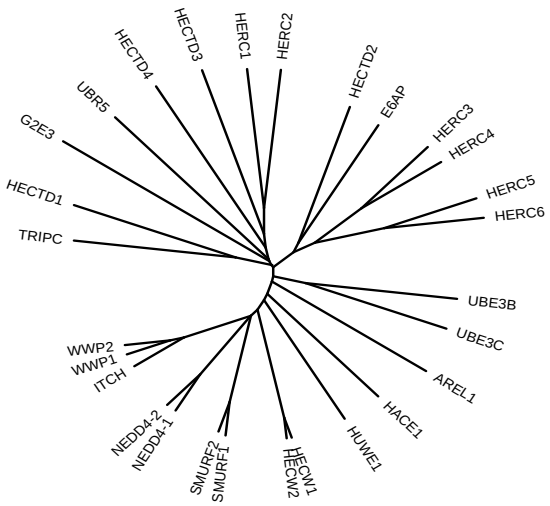


Figure 5: Phylogenetic analysis of the 28 human HECT ligases

The HECT domain sequences of the 28 human HECT ligases were aligned using Clustal Omega [69] and converted to a phylogenetic tree using Interactive Tree of Life (iTOL) v3 server [70]. The figure was generated by Dr. Sonja Lorenz [71].

1.3.2 Structural mechanism of the HECT domain

The first structure of a HECT domain was determined for E6AP [72], showing the characteristic two lobes joint by a hinge region (Figure 6) [73], which was later found to be flexible [74], thus enabling the different relative arrangement of the two lobes required during catalysis. The C-lobe harbors the catalytic cysteine and the N-lobe interacts with the cognate E2s (UBCH7 is widely assumed to be the main E2 that functions with HECT E3s, but UBCH5, 6 and 8 have also been reported) and, at least in some cases, a regulatory ubiquitin molecule [75-82]. The HECT domain enables a two-step reaction [72,83-86]. In the first step, an E2-donor ubiquitin conjugate interacts with the N-lobe and the ubiquitin is transferred to the catalytic cysteine on the C-lobe in a trans-thioesterification reaction. In

the second step, the activated C-terminal carbonyl group of donor ubiquitin is nucleophilically attacked by a primary amino group of an acceptor ubiquitin or a target protein (Figure 7) [50]. During the process of chain formation, the ubiquitin acts as a target protein.

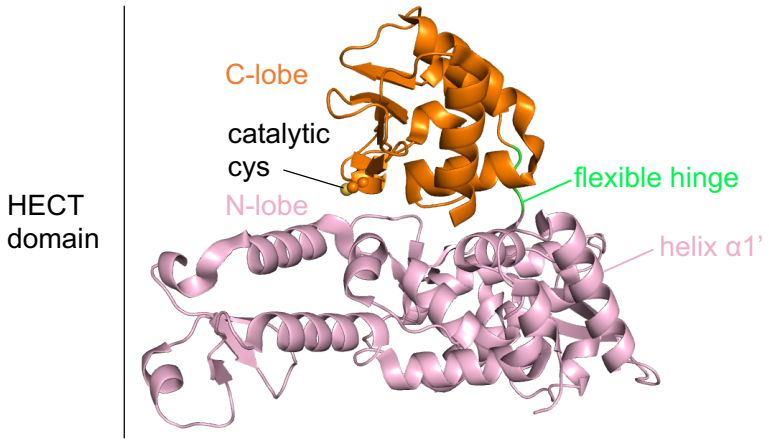


Figure 6: Architecture of the HECT domain

Crystal structure of the HECT domain of HUWE1 is shown in cartoon representation (PDB: 3H1D) [73]. The flexible hinge region is shown in green and the N and C lobes are labelled. The side chain of the catalytic cysteine is shown as a ball-and-stick model.

To achieve linkage specificity in ubiquitin chain formation, the HECT domain ought to orient the acceptor ubiquitin in such a way that it presents a particular primary amino group to the donor-bound active-site cysteine. How this occurs structurally has remained unclear. Another open question is whether binding partners of HECT E3s in the cell influence the interactions of the E3 with target proteins, ubiquitin and E2s, respectively. For example, the affinities of HECT domains with UBCH7 fall into the low micromolar range *in-vitro* [50], suggesting that additional factors or co-localization mechanisms may contribute to HECT ligase activities in the cell [50]. More research is also needed to unravel which E2-E3 pairs are relevant in the context of the cell.

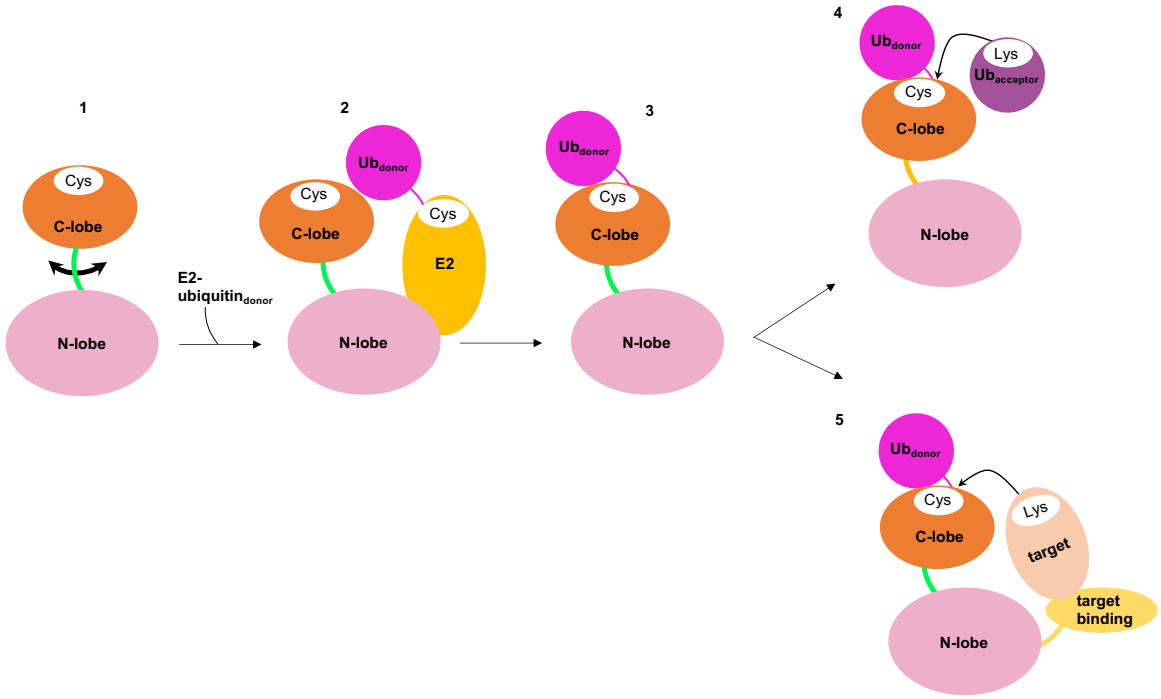


Figure 7: Schematic of the reaction mechanism of the HECT domain

(**1**) The *apo* state of the HECT domain in which the N- and C-lobe are supposed to be flexible with respect to each other; (**2**) the trans-thioesterification step with the E2 and resulting product (**3**). Thereafter there are two possibilities: (**4**) the nucleophilic attack of an acceptor ubiquitin molecule or (**5**) target protein on the activated C-terminal carbonyl group of the donor ubiquitin, resulting in ubiquitin chain formation and target protein modification, respectively. For the recognition of target proteins, regions outside of the HECT domain are required.

The flexibility of the hinge region was shown to be indispensable for catalysis and induce conformational rearrangements of the two lobes [72,74,81]. That such flexibility is necessary was predicted based on the crystal structure of an E6AP-UBCH7 complex [72] where the catalytic cysteine of the E3 is at a distance of 41 Å from the catalytic cysteine of the E2 (Figure 8A). Many conformations of the HECT domain have been observed in crystal structures and have been categorized as an “inverted T-shape” and an “L-shape” [72] (Figure 8A). The first step is necessary during trans-thioesterifiction, based on a crystal structure of a complex containing the HECT domain of NEDD4L, E2 and donor ubiquitin.

The latter state is important during the second catalytic step of isopeptide bond formation, based on a crystal structure of a complex containing the HECT domain of Rsp5, donor

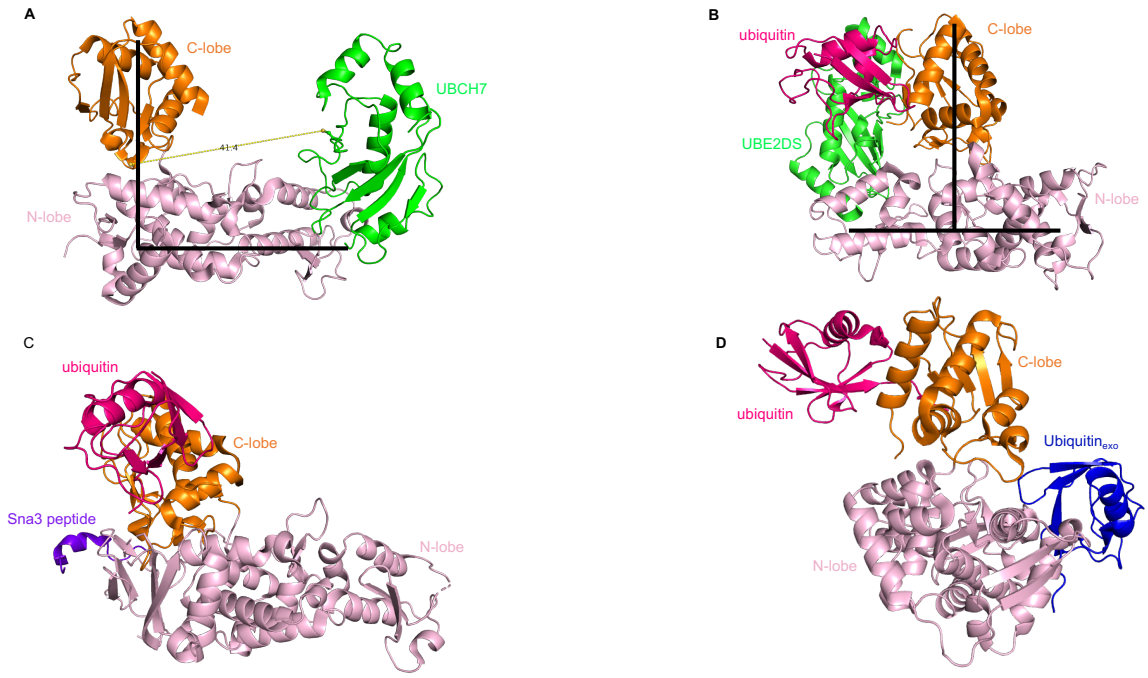


Figure 8: Crystal structures of HECT domains in different conformations

(**A**) Crystal structure of the E6AP HECT domain-UBCH7 complex in the L-shape (extracted from PDB: 1C4Z) [72], in an "L-shape" (black line). The side chains of the catalytic cysteines of E6AP and UBCH7 are shown in ball-and-stick representation. The distance of 41.4 Å between the two sulphur atoms is indicated. (**B**) Crystal structure of a NEDD4L HECT domain-UBCH5-ubiquitin complex (extracted from PDB: 3JVZ) [84] in an "inverted T-shape" (black line); (**C**) Crystal structure of Rsp5 HECT domain-ubiquitin-Sna3 complex (extracted from PDB: 4LCD) [87] (**D**) Crystal structure of a NEDD4 HECT domain complex with donor ubiquitin and an additional ubiquitin molecule bound to the exosite on the N-lobe ("ubiquitin $_{exo}$") (extracted from PDB: 4BBN) [80].

ubiquitin and a SNA3-derived substrate peptide (Figure 8C) stabilized by a three-way crosslinker [87]. In contrast, the crystal structure of a NEDD4L HECT domain-UBE2D2-ubiquitin complex (Figure 8B) shows an inverted-T shape, as required for trans-thioesterification [87]. The structure indicates that the C-terminus of the donor is

sandwiched between the catalytic centers of the UBE2D2 and NEDD4L and the C-lobe of NEDD4L engaging in hydrophobic interactions with the donor ubiquitin. Note that in this structure, the catalytic cysteine of the E2 was replaced with a serine, which forms an oxyester with the C-terminus of ubiquitin in lieu of the native thioester, as the latter is susceptible to hydrolysis. In all the available HECT domain structures that contain an active site-linked ubiquitin molecule, the donor forms a conserved hydrophobic interface (comprising Ile 36, Leu 71 and Leu 73 of ubiquitin) with the C-lobe of HECT domain (e.g. comprising Asp 4333, Arg 4334 and Leu 4335 in the case of HUWE1 [88](Figure 24)). The interface is maintained throughout the trans-thioesterification reaction [80,87].

In NEDD4- type ligases and E6AP, the presence of a non-covalent ubiquitin binding site, called “exosite”, on the N-lobe was shown to promote ubiquitin chain elongation [50]. The underlying mechanism is still unclear, but was suggested to involve the stabilization of a growing ubiquitin chain (Figure 8D) [79,80]. Consistently, mutations in the exosite of Rsp5, NEDD4, NEDD4L and WWP1 impair ubiquitin chain elongation, but not chain initiation on the target protein [79]. A small-molecule, exosite-directed inhibitor of NEDD4-1 showed that blocking the exosite can switch the enzyme from being processive to distributive [89].

To understand the specificity and processivity of HECT-type ligases and illuminate the factors that control chain length, processivity and linkage specificity, we need to understand on a structural level how HECT domains interaction with ubiquitin and how the different ubiquitin molecules (donor, acceptor and exosite-bound ubiquitin) impact each other and the conformational dynamics of the HECT domain. My thesis work aims to address these questions, at least in part, using HUWE1 as a model system.

1.3.3 HUWE1

1.3.3.1 Pathophysiological roles

The HECT-type ligase HUWE1 (also known as HECTH9, ARF-BP1, URE-B1, MULE and LASU1) is associated with proliferation/differentiation, apoptosis, DNA repair, stress responses and overexpressed in many different types of cancer [63]. However, its various roles in the cell are incompletely understood. HUWE1 can monoubiquitinate target proteins

and also decorate them with Lys 6, Lys 48, Lys 63-linked chains [90]. Which factors influence the modification type formed by HUWE1 is unknown. In the following section, a selection of actions of HUWE1 on particular disease-associated substrates are outlined, illustrating the complexity of its biological functions.

MYC, a transcription factor and hallmark proto-oncogene is identified as one of the known substrates of HUWE1 [91]. MYC can form a binary complex with MAX that binds to E-boxes on DNA (CACGTG), thereby activating the transcription of pro-growth genes [91]. Upon formation of a ternary complex with the MYC-interacting zinc finger protein 1 (MIZ1), the interaction of MIZ1 and the P300 histone acetyltransferase is impeded, thus repressing transcription and promoting the recruitment of the DNA methyltransferase, DNMT3a [92,93]. HUWE1 was reported to form Lys 63-linked chains on C-MYC, which is required for the recruitment of P300 [91]. Knockdown studies of HUWE1 in HeLa cells showed a reduction in MYC-associated functions and co-depletion of MNT, a MYC transactivation antagonist, pointing towards an oncogenic role of HUWE1 in MYC-associated tumors [94]. HUWE1 also interacts with N-MYC, which is expressed in neural stem cells and neuroectodermal progenitors [95]. Biochemical and proteomic studies identified N-MYC as a substrate for HUWE1 in neural cells, where HUWE1 makes Lys 48-linked chains on N-MYC and directs it for proteasomal degradation [96]. In the same study, N-MYC knockdown in mouse embryonic cells decreased the level of Cyclin D2, a downstream target of N-MYC [96]. It was also shown that C-MYC, just like N-MYC, can be modified with K48-linked chains by HUWE1, leading to a context and time-dependent regulation of MYC [96].

HUWE1 was also reported to form Lys 48-linked chains on MIZ1 upon tumor necrosis factor alpha (TNFα) stimulation, thereby directing MIZ1 for proteasomal degradation [97]. In the absence of HUWE1, MIZ1 suppresses the TNFα-induced JNK activation and cell death [97]. Studies in mouse embryonic fibroblasts showed that knockdown of HUWE1 tightly regulated TNFα-induced JNK activation and cell death, but the effect was abolished upon knockout of MIZ1 [97]. A reduction in HUWE1 levels resulted in the accumulation of C-MYC/MIZ1 complex however, knockout of C-MYC rescued the HUWE1 phenotype [98]. Thus, a finely tuned balance and regulation of MIZ1 and MYC by HUWE1 controls tumorigenesis [99].

The P53 tumor suppressor protein is another well-studied substrate of HUWE1 with diverse physiological functions ranging from stress responses, oncogene activation and hypoxia to the DNA damage response [100]. HUWE1 polyubiquitinates P53, which directs it for protein degradation and suppresses p53-dependent apoptosis [101]. In MYC-driven B cell lymphomas, the P53 protein levels are elevated due to a suppression of HUWE1, leading to growth arrest and apoptosis [102].

Another HUWE1 substrate is the anti-apoptotic factor MCL1, which is frequently upregulated in cancers to promote cell survival [63]. MCL1 associates with the proapoptotic BAK protein to maintain the latter in an inactive state under normal conditions [103]. HUWE1 interacts with MCL-1 via its BH3 domain and promotes the degradation of MCL1, thereby enhancing apoptosis in response to DNA damage [104]. The HUWE1-MCL1 interplay also has a role in the regulation of obesity, where depletion of HUWE1 was shown to stabilize MCL1 via the inactivation of GSK3β [105].

HUWE1 also associates with the BRCA1-MERIT40/RAP80 complex and together mediate the polyubiquitination and degradation of BRCA1, a key regulator of genomic stability [106]. HUWE1 mediated degradation of BRCA1 is related to decreased BRCA1 protein levels in sporadic breast cancer and IR induced DNA damage and repair response [107]. HUWE1 also functions in protein quality-control, by targeting unassembled, soluble protein subunits for degradation [108]. The knowledge about DUBs associated with HUWE1 is limited. A recent study showed that HUWE1 assembles Lys 63-linked chains on USP7 to increase its deubiquitinase activity and thus generates a platform for the recruitment of HIF-1α and other regulatory proteins, influencing tumor progression [109].

How these various, structurally distinct substrates are recruited by HUWE1, oriented towards the catalytic center and decorated with specific ubiquitin modifications, is unknown. It will therefore be important to reconstitute selected HUWE1-substrate complexes together with associated cellular factors to elucidate the structural basis of substrate and linkage specificity of this crucial ligase.

1.3.3.2 Structural mechanism

HUWE1 is one of the largest HECT-type ligases with a molecular weight of 482 kDa. The domain architecture of HUWE1 (Figure 9) shows a signature HECT domain at the C-terminus. The N-terminal region contains a an Armadillo repeat domain, ubiquitin-associated domain (UBA) of yet unknown functions, a WWE domain, a BH3 domain, which interacts with MCL1 [104] and a UBM domain of unknown function. Several crystal structures of the HECT domain of HUWE1 are available [73,88,110]. Its N-terminal helix α1 was found to be important for the stability and function of the HECT domain in line with studies on other HECT ligases [50,73]. A crystal structure of the isolated C-lobe of HECT domain of HUWE1 linked to the donor ubiquitin showed a canonical, NEDD4-type interface [88], in line with HUWE1 being the closest phylogenetic neighbor of NEDD4-type ligases (Figure 5).

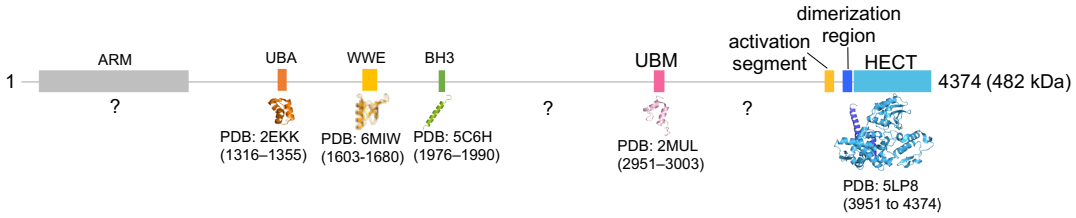


Figure 9: Domain organization of HUWE1

Predicted domains of HUWE1 are indicated, including domain boundaries and structural information, if available. UBA domain [111], WWE domain [112], BH3 domain [113], UBM domain [114], HECT domain with dimerization region [110]. This figure was prepared by Dr. Sonja Lorenz.

A region flanking the HECT domain N-terminally was structurally characterized and shown to mediate dimerization [110]. Dimerization mediates the autoinhibition of HUWE1 and is regulated by both intra- and intermolecular interactions [110]. Engagement of the dimerization region with a region known as the “activation segment”, located ~ 50 residues N-terminally, counteracts dimerization and activates ligase activity. This interaction of the dimerization region is competing with the intermolecular engagement of the tumor

suppressor p14ARF by the activation segment, which stabilizes the ligase in the inactive, dimeric state [110]. It will be critical to investigate how the extended N-terminal region of HUWE1 impacts the conformational dynamics and function of the catalytic domain. Furthermore, open questions remain regarding specific properties of the catalytic domain, such as whether it binds ubiquitin through an exosite (like NEDD4-type ligases and E6AP) or how the acceptor ubiquitin is oriented to encode linkage specificity in ubiquitin chain formation.

1.3.4 Manipulating HECT-type ligases

1.3.4.1 Small-molecule ligands

To date, only a few small molecules have been developed to target the HECT-type ligases. For example, a derivative of 1H-indole-3-yl-carbinol (I3C) named HECLIN, selectively inhibits the activity of several NEDD4-type ligases *in-vitro* and *in-vivo*, through an unclear mechanism, which was suggested to involve oxidation of the active-site cysteine [115]. Clomipramine, an FDA-approved antidepressant, can block ITCH auto-ubiquitination and P73 ubiquitination [116,117]. The first crystal structure of a HECT domain in complex with a small molecule was provided by Statsyuk and co-workers. They used a 100 fragment-based electrophilic library screen to identify an indole-based compound that binds adjacent to the exosite cysteine (Cys 627) of NEDD4-1 [117]. No labelling of E6AP and WWP1 was observed as these ligases are lacking a cysteine residue at the homologous position [117]. Another class of I3C-derived molecules identified as NEDD4 inhibitors inhibit the ubiquitination of PTEN by NEDD4-1 [117]. An N-aryl benzimidazole (NAB) derivative, was found to modulate Rsp5 activity in neurodegenerative processes, rescuing yeast cells expressing toxic levels of α-synuclein by restoring vesicular trafficking pathways. This makes NAB derivatives attractive leads for drug development against neurodegenerative disorders [117]. A high throughput *in-vitro* screen against HUWE1 identified two hits, BI8622 and BI8626, which specifically inhibit HUWE1 with low micromolar IC_{50}-values [99,117]. N-acetylphenylalanine was identified as a potential small molecule inhibitor of E6AP [117]. No HECT-directed inhibitors are in clinical studies. To develop high-quality chemical probes, structural knowledge of the yet uncharacterized steps in the catalytic cycle of HECT ligases will be indispensable.

1.3.4.2 Ubiquitin-activated probes

Activity-based probes (ABPs) have been developed for different families of enzymes, including serine hydrolases, cysteine protease and others [118-120]. ABPs usually contains three components: a reactive group (also known as “warhead”), a recognition element and a reporter tag [121]. The choice of warhead will affect both the reactivity and selectivity of the probe. The selectivity is also a function of the recognition element, which can be a small molecule, a short peptide, or a protein. The reporter tag is used to detect or isolate ABP-labeled targets [122]. In the ubiquitin system, DUBs have proven suitable targets for electrophilic ubiquitin-based ABPs and have been successfully used to identify new DUBs, characterize enzymatic features *in-vitro* and in cells, determine the structures of DUB-ubiquitin complexes and explore the accessibility of DUBs as therapeutic targets [123]. Various warheads have been attached to ubiquitin and characterized in terms of their reactivity with DUBs, including vinyl methyl sulfone [124], aldehyde and nitrile groups [125]. Presumably, the most widely used ubiquitin-based ABPs are ubiquitin-propargylamide (Ub-PA) or ubiquitin-vinyl methyl ester (Ub-VME) [126-128]. Internal and terminal diubiquitin-based probe [129-133], ubiquitin photoaffinity probes [134,135], ubiquitin interactor affinity probes [136,137], ubiquitin-substrate protein probes [138,139] and a reactive-site-centric DUB probes [140] are also available.

Interestingly, ubiquitin-based ABPs were found to target certain other members of the ubiquitination machinery despite DUBs: For example, HA-ubiquitin vinyl methyl ester (HA-VME) labels the HECT domain of HUWE1 [141]. HUWE1 was also found to react with Ub-PA in cell lysate [127]. Additionally, Ub-VME and Ub-PA were found to react with the catalytic cysteines on HECT ligases *in-vitro* [126,141-143], however, with rather low turnover [142]. A list of available ubiquitin based ABPs that were shown to modify HECT ligases is provided below;

Table 1: Table of activity-based ubiquitin probes for HECT-type ubiquitin ligases

HECT	ABPs
NEDD4	His-UBCH7-Ub [142]
	His-UBCH7-Ub [142]
	His-UBCH5B-Ub [142]
	Ub-VME [142]

	biotin-Ub-PA [142]
	Ub-Dha [143]
UBE3C	His-UBCH7-Ub [142]
	His-UBCH7-Ub [142]
	His-UBCH5B-Ub [142]
	Ub-VME [142]
	biotin-Ub-PA [142]
	Ub-Dha [143]
HUWE1	Ub-VME [141,142]
	biotin-Ub-PA [142]
	Ub-PA [126]
HECTD1	His-UBCH5B-Ub [142]
	Ub-VME [142]
	Ub-Dha [143]
UBR5	Ub-VME [142]
E6AP	Ub-VME [141]

1.4 Aims of the thesis

The overall aim of my thesis work was to elucidate mechanisms of ubiquitin recognition by HECT-type ligases and understand whether they are conserved or enzyme-specific. As a model system, I chose HUWE1 since it is an important cancer-therapeutic target, yet its structural mechanism is poorly understood.

In particular, I aimed to address the following specific biological questions:

- How is the donor ubiquitin oriented in the context of the HECT domain of HUWE1?
- What conformational changes does donor binding induce in the HECT domain of HUWE1?
- Does HUWE1 use a ubiquitin-binding exosite, like NEDD4-type ligases and E6AP?
- Is there a structural communication between donor binding and the putative binding of ubiquitin to the exosite?
- How is the acceptor ubiquitin recognized by the HECT domain of HUWE1 to enable linkage-specific chain formation?

On a more technical level, my work also addressed the question of whether ubiquitin based ABPs offer tools to access yet uncharacterized structural features of HECT ligases.

To answer these questions, I combined biochemical with biophysical and high-resolution structural analyses and validated structure-guided predictions *in-vitro*.

2 MATERIALS

2.1 Primers

Oligonucleotides for site-directed mutagenesis, restriction free (RF) cloning and sequencing were purchased from Sigma-Aldrich in high-purity salt-free (HPSF) quality and lyophilized form and dissolved in reagent-grade water.

Table 2: Oligonucleotide sequences, used in polymerase chain reactions (PCRs)

Sl no	Primer Name		Primer Sequence
1	HUWE1 HECT C4184A	F	GTCCAAGAGTTTGGAGTTGCGGAAGTTCGTGACCTCAAAC
		R	GTTTGAGGTCACGAACTTCCGCAACTCCAAACTCTTGGAC
2	HUWE1 HECT C4099A	F	CAATCCATCTTCCCACGCGAACCCCAACCACCTC
		R	GAGGTGGTTGGGGTTCGCGTGGGAAGATGGATTG
3	HUWE1 HECT C4367A	F	GTTGGCTATCCAGGAGGCGTCTGAAGGCTTTGGG
		R	CCCAAAGCCTTCAGACGCCTCCTGGATAGCCAAC
4	HUWE1 HECT C4341A	F	TGCCTTCAGCTCACACAGCGTTTAATCAGCTGGATCTG
		R	CAGATCCAGCTGATTAAACGCTGTGTGAGCTGAAGGCA
5	HUWE1 HECT L4061A	F	GATGCTGGTGGGGCCCTGCGGGAGTGGTAT
		R	ATACCACTCCCGCAGGGCCCCACCAGCATC
6	HUWE1 HECT L4061D	F	GATGCTGGTGGGGATCTGCGGGAGTGGTAT
		R	ATACCACTCCCGCAGATCCCCACCAGCATC
7	HUWE1 HECT E4054A	F	ATAGTATTTGAAGGAGAAGCAGGGCAGGATGCT
		R	AGCATCCTGCCCTGCTTCTCCTTCAAATACTAT
8	HUWE1 HECT E4064A	F	GGTGGGCTCCTGCGGGCCTGGTATATGATC
		R	GATCATATACCAGGCCCGCAGGAGCCCACC
9	HUWE1 HECT C lobe	F	AAGTTCTGTTCCAGGGGCCCCTGCCCACCATTGACATCGATGATCTGAAA

		R	TCCACCAGTCATGCTAGCCATATGTTATTAGGCCAGCC CAAAGCCTTC
10	HUWE1 HECT Y4078A	F	ATGATCATCTCTCGAGAGATGTTTAACCCTATGGCCGC CTTGTTCCGTACCTCACCTGGTG
		R	CACCAGGTGAGGTACGGAACAAGGCGGCCATAGGGT TAAACATCTCTCGAGAGATGATCAT
11	HUWE1 HECT F4181A	F	CTCACCTTCAGCACTGAGGTCCAAGAGGCCGGAGTTT GTGAAGTTCGTGACCTC
		R	GAGGTCACGAACTTCACAAACTCCGGCCTCTTGGACC TCAGTGCTGAAGGTGAG

Additional primers were designed by Dr. Sonja Lorenz, while working at the University of California, Berkeley, CA/USA.

2.2 Bacterial strains and expression constructs

Table 3: Bacterial strains for cloning and protein expression

Organism	Strain	Genotype	Supplier
E. coli	Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG	Invitrogen
E. coli	DH5α	F- φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rk -, mk +) phoA supE44 thi-1 gyrA96 relA1 λ-	Invitrogen
E. coli	BL21 (DE3)	F- ompT hsdS(rB – mB -) dcm+ ga (DE3)	Invitrogen

Table 4: Vectors for protein expression in bacteria

Sl . No	Vector Name	Host	Affinity tag	Cleavage site	Resistance	Supplier
1	pSKB2	bacterial	N-terminal His6-tag	Precision protease	Kan	
2	pTXB1	bacterial	CBD		Amp	Prof. David komander
3	pM41	bacterial	N-terminal His6-tag	Prescision protease	Kan	Barbara Orth

Table 5: Expression constructs. The protein sequences are from *homo sapiens*, except for the proteases

Vector	Insert	Residue number	Provided by
pSKB2	HUWE1 HECT domain	3993-4374	Dr. Sonja Lorenz
pSKB2	HUWE1 HECT domain-triple cys-ala mutant (TM)	3993-4374	
pSKB2	HUWE1 HECT L4061A	3993-4374	
pSKB2	HUWE1 HECT L4061D	3993-4374	
pSKB2	HUWE1 HECT E4054A	3993-4374	
pSKB2	HUWE1 HECT E4064A	3993-4374	
pSKB2	HUWE1 HECT L4061A	3843-4374	
pSKB2	HUWE1 HECT L4061D	3843-4374	
pSKB2	HUWE1 HECT E4054A	3843-4374	
pSKB2	HUWE1 HECT E4064A	3843-4374	
pM41	MIZ1	1-282	Barbara Orth
pET-30a	ubiquitin	1-76	Dr. Sonja Lorenz
pTXB1	ubiquitin	1-76	Prof. David Komander
pET-28a	E6AP HECT domain	495-852	Dr. Lena Ries
pET-28a	E6AP HECT C-lobe	495-740	Dr. Lena Ries
pET-28a	NEDD4 HECT domain	514-900	Sino Biological Inc.
pET-28a	NEDD4 HECT C-lobe	782-900	Dr. Sonja Lorenz
pET-2d	UBE2L3	1-154	Prof. Michael Rape
pET-28a	3C protease	1-401	Dr. Florian Sauer
pRK793	TEV protease	1-303	Prof. John Kuriyan

2.3 Bioreagents, kits and enzymes

Table 6: Bioreagents, enzymes and kits

Designation	Supplier
5x Q5 reaction buffer	New England Biolabs
5x Q5 High GC Enhancer	New England Biolabs
10x Standard Taq Reaction Buffer	New England Biolabs
Albumin Fraktion V (BSA)	Roth
DNaseI	Invitrogen
DpnI	New England Biolabs
GC buffer (PCR)	New England Biolabs
Lysogeny broth (LB) medium	Carl Roth
NucleoSpin Gel and PCR cleanup kit	Macherey & Nagel
NucleoSpin Plasmid kit	Macherey & Nagel
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific
Q5 high fidelity DNA Polymerase	New England Biolabs
SignalFire ECL Reagent	Cell Signalling Technology
Taq DNA Polymerase	New England Biolabs
Terrific Broth (TB) medium	Carl Roth

2.4 Chemicals

Table 7: Chemicals

Substance	Supplier
2-Deoxyadenosine 5-triphosphate (dATP) sodium salt solution	New England Biolabs
2-Deoxycytidine 5-triphosphate (dCTP) sodium salt solution	New England Biolabs
2-Deoxyguanosine 5-triphosphate (dGTP) sodium salt solution	New England Biolabs
2-Deoxythymidine 5-triphosphate (dTTP) sodium salt solution	New England Biolabs
2-Propanol (Isopropanol)	Carl Roth
3-Morpholinopropane-1-sulfonic acid (MOPS)	Sigma-Aldrich
4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Carl Roth
Acetic acid	Carl Roth
Ammonium persulfate (APS)	Carl Roth
Ammonium acetate (NH_4COOH)	Carl Roth
Ampicillin sodium salt	Carl Roth

Benzamidin hydrochloride monohydrate	Carl Roth
Beta-mercapthoethanol (β-ME)	Sigma-Aldrich
Bis-Acrylamid 29:1	Fisher Bioreagents
Bromphenol blue	Carl Roth
Calcium chloride dihydrate ($CaCl_2$)	Carl Roth
Chloramphenicol	Carl Roth
Coomassie Brilliant Blue G-250	Carl Roth
Coomassie Brilliant Blue R-250 Carl Roth	Carl Roth
Dimethylsulfoxide (DMSO)	Carl Roth
Disodium hydrogen phosphate (Na_2HPO_4)	Carl Roth
Dithiothreitol (DTT)	Carl Roth
Ethanol (EtOH)	Carl Roth
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth
Ethylene glycol	Sigma-Aldrich
Glucose	Carl Roth
Glycerol	Carl Roth
Glycine	Carl Roth
Hydrochloric acid (HCl)	Carl Roth
Imidazole	Carl Roth
Iron(II) sulfate heptahydrate ($FeSO_4$)	Sigma-Aldrich
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Carl Roth
Kanamycin sulfate	Carl Roth
Magnesium chloride hexahydrate	Carl Roth
Magensium sulfate heptahydrate (Mg_2SO_4)	Sigma-Aldrich
Manganese(II) chloride tetrahydrate ($MnCl_2 \cdot 4\ H_2O$)	Sigma-Aldrich
2-Mercaptoethanesulfonic acid sodium salt (MESNa)	Sigma-Aldrich
Methanol	Carl Roth
Nickel(II) sulfate hexahydrate	Carl Roth
Perchloric acid	Sigma-Aldrich
Polyethylenglycol (PEG) 3350	Sigma-Aldrich
Polyethylenglycol (PEG) 4000	Sigma-Aldrich
Potassium dihydrogen phosphate	Carl Roth
Potassium formate (KOOH)	Carl Roth
Propargylamine	Sigma-Aldrich
Sodium acetate	Carl Roth
Sodium cacodylate trihydrate	Sigma-Aldrich
Sodium chloride (NaCl)	Carl Roth
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich
Sodium dihydrogenphosphate dihydrate (NaH_2PO_4)	Carl Roth

Sodium hydroxide (NaOH)	Carl Roth
Tetramethylethylenediamin (TEMED)	Carl Roth
Thiamine hydrochloride	Sigma-Aldrich
Tris-(2-carboxyethyl)-phosphine (TCEP)	Carl Roth
Tris-(hydroxymethyl)-aminoethane (Tris)	Carl Roth
Triton X-100	Sigma-Aldrich
Tween 20	Sigma-Aldrich
Yeast nitrogen base	Sigma-Aldrich
Zinc chloride ($ZnCl_2$)	FlukaBioChemica

2.5 Crystallization screens

All crystallization screens (Table 8) were prepared by the in-house facility according to the composition of the corresponding commercial screens.

Table 8: Commercial crystallization screens used as templates for in-house screens

Screen	Supplier
Additive Screen	Hampton Research
Crystal Screen, Crystal Screen 2	Hampton Research
Index Screen HT	Hampton Research
JCSG+	Molecular Dimensions
Nextal PEG Suite	Qiagen
Nucleix Suite	Qiagen
PEGs Suite, PEGs II Suite	Qiagen
Protein Complex Suite	Qiagen
Silver Bullets Bio	Hampton Research
Wizard 1+2, Wizard 3+4	Emerald BioSystems

2.6 Specialized consumables

A list of specialized concumables is provided in Table 9.

Table 9: Specialized consumables

Type	Model	Supplier
24-well hanging-drop crystallization plates	Crystalgen Super Clear Plate	Jena Bioscience
96-well sitting-drop crystallization plates	Crystal quick 1 square well, flat bottom, low profile	Greiner Bio-One
Microplate 96 well	96-well half area microplates	Greiner Bio-One
Cover slides	22 mm, siliconized	Jena Bioscience
Dialysis membranes	Spectra/Por	Spectrum Laboratories
Nickel-beads	Ni-NTA agarose	Macherey & Nagel
PVDF membrane	Roti -PVDF	Carl Roth
SDS gels	Novex 10-20% Tris-Glycine Mini Gels, Wedge Well	Thermo Fisher Scientific
ultrafiltration units	Amicon MWCO 3-30 kDa, 0.5-20 ml	Merck Millipore

2.7 Relevant scientific equipment

A list of relevant scientific equipment is provided in Table 10.

Table 10: Scientific equipment

Device	Model	Company
Affinity chromatography	HisTrap HP 1 ml/5 ml	GE Healthcare
Anion exchange chromatography column	Mono Q 4.6/100 PE	GE Healthcare
Cation exchange chromatography column	HiTrap SP 5 ml	GE Healthcare
CD cuvette	QS-110	Hellma
CD spectropolarimeter	J-810	Jasco
Chemiluminescence	FluorchemQ Multi image	Alpha Innotech
Crystallization loops	CryoLoop	Hampton Research
Crystallization robot	Analytic Honey Bee 963	Digilab
SEC columns	HiLoad 16/600 Superdex (SD) 200 pg	GE Healthcare
	HiLoad 16/600 Superdex 75 pg	

	Superdex 75 increase 10/300 GL	
FPLC system	AKTA pure 25	GE Healthcare
Gel electrophoresis chamber	Mini-ProteanR 3-cell	Bio-Rad
	Xcell SureLock Mini-Cell	Invitrogen
Liquid handling robot	LISSY 2002	Zinsser Analytik
Microplate Reader	CLARIOstar	BMG LABTECH
Robotic sealing unit for microplates	RoboSeal	H.J. BIOANALYTIC
Scanner	Odyssey	LI-COR
Sonicator	LabsonicR B.	Braun Biotech International
Spectrophotometer	Bio-Photometer Plus	Eppendorf
Nanodrop	ND 2000c	Thermo Fisher Scientific
Thermo block	Rotilabo-Block Heater250	Carl Roth
	Thermomixer Comfort	Eppendorf
uv imaging system	Gel Doc XR System	Bio-Rad
Western blot	Trans-Blot Turbo Transfer- System	Bio-Rad

2.8 Software, servers and databases

The software as well as server-based tools and databases (Table 11) were used in the latest version published at the time.

Table 11: Software, server-based tools and databases

Program	Description	Supplier / Reference
AIMLESS	Scaling and merging of diffraction data	[144]
AlphaView	Capturing images of the FluorchemQ system	Alpha Innotech
blastp/ blastn	Sequence search	[145]
CCP4	Software suite for macromolecular X-ray crystallography	[146]
COOT	Model-building software	[147]
Clustal Omega	Sequence alignments	[148-150]
CrystalClear	X-ray data collection and basic processing	Rigaku
ExPASy ProtParam	Computation of physical and chemical properties of proteins	[151]
ExPASy Translate	Translation tool of nucleotide sequences to protein sequences	[151]

ImageJ	image processing program, 1.48v	[152]
Interactive Tree of Life V3 server	Generation of phylogenetic tree	[70]
Image Studio Software	image processing program	LI-COR
MARS	Clariostar data analysis software	BMG LABTECH
Microsoft Excel	Spreadsheet software	Microsoft Corporation
Microsoft Power	Generation of figures	Microsoft Corporation
MolProbity	Structure validation for macromolecular crystallography	[153]
MXCube2	X-ray data collection GUI	[154]
ODYSSEY	Infra-red imaging software	LI-COR
OriginPro	Graphics and data analysis software	OriginLab
PDB	Protein Data Bank	[155]
PHASER	Phasing software	[156]
Phenix	Software suite for macromolecular X-ray crystallography	[157]
Phyre^2	Biosequence analysis; protein 3D structure prediction	[158]
PISA	Bioinformatic characterization of interfaces	[159]
PrimerX	Automated design of mutagenic primers for site-directed mutagenesis	Lapid, 2003
PyMOL	3-dimensional visualization and graphical illustration software	DeLano Scientific LLC
Pubmed	Literature search	[160]
RF-cloning.org	automated primer design process for RF cloning	[161]
Quantity One	UV imaging system control; UV image recording and analysis	Bio-Rad
Spectra Manager	CD data acquisition and analysis	Jasco
UNICORN	FPLC instrument control; recording, analysis and management of chromatograms	GE Healthcare
XDS	Indexing, and integration of diffraction images	[162]

3 METHODS

3.1 Protein production

3.1.1 Molecular biology

3.1.1.1 Preparation plasmid DNA and transformation into *E. coli*

Chemical competent cells were prepared by the Lorenz lab technical assistant, Julia Haubenreisser. To this end, 5 ml LB medium supplied with the appropriate antibiotic(s) was inoculated with the desired *E. coli* strain and incubated shaking at 37°C overnight. 1% of the overnight culture was then transferred into 100 ml of LB medium, supplied with the appropriate antibiotic(s) and kept shaking at 37oC until the $O.D_{600}$ reached a value of 0.4 to 0.6. The culture was then cooled on ice for 5 minutes and harvested by centrifugation at 4000 rpm and 4oC. The cells were then resuspended in 30 ml of ice-cold TFB1 (30 mM KCOOH pH 5.8, 75 mM CaCl2, 10 mM RbCl, 15 % glycerol), incubated on ice for 90 minutes and pelleted again by centrifugation at 4000 rpm and 4oC. The cells were then resuspended gently in 4 ml of ice-cold TFB2 (10 mM MOPS pH 6.5, 75 mM CaC12, 10 mM RbCl, 15 % glycerol), incubated on ice for 15 minutes, frozen in liquid nitrogen and stored at -80°C for future use.

Single bacterial colonies carrying the plasmid of interest were inoculated into 5 ml of LB medium supplied with the appropriate antibiotic(s) and incubated shaking at 37°C overnight. The bacterial cells were then pelleted by centrifugation at 6000 rpm at 4°C for 8 minutes and the plasmid DNA was isolated with the help of a NucleoSpin Plasmid Kit (Macherey & Nagel), following the manufacturer's protocol.

In order to express a protein of interest, the corresponding plasmid was transformed into chemical competent E. coli cells. To this end, 50 µl of frozen cells were thawed on ice, 10 to 100 ng of plasmid DNA was added and incubated on ice for 30 minutes. The cells were then subjected to a heat shock at 42°C for 45 seconds in a water bath, incubated on ice for 5 minutes, supplied with 700 µl of LB medium and incubated shaking at 37°C for 1 hour. Subsequently, the cells were pelleted by centrifugation at 2000 rpm and room temperature for 3 minutes, resuspended in 100 µl of LB medium, plated on an agar plate supplemented with the appropriate antibiotic(s) and incubated at 37°C overnight.

3.1.1.2 Cloning techniques

3.1.1.2.1 Site-directed mutagenesis

Site-specific mutations or deletions were introduced by the QuickChange (QC) procedure. The oligonucleotides were designed with the help of the server PrimerX [163] in such a way that the mutation site was in the middle, with at least 14 to 16 bp overhangs on the 5' and 3' end. For PCRs Q5®, high fidelity DNA polymerase (New England Biolabs) was used and the amplification cycle adjusted to the size of the respective plasmid (20-30 seconds per kb). Dpn1 (New England Biolabs) digests were performed at 37°C overnight to remove the methylated template DNA. The remaining DNA was then transformed into chemical competent *E. coli* TOP10 cells (ThermoFischer) (3.1.1.1), the plasmid DNA isolated (3.1.1.1) and the sequence verified (Eurofins Genomics or Microsynth Seqlab).

3.1.1.2.2 Sub-cloning

Generally, sub-cloning was conducted by restriction free (RF) methods [164]. The corresponding primers were designed using the online platform RF-cloning.org [161] In the first step the gene of interest is amplified with the designed primers, giving rise to complementary to the target plasmid on either end. This PCR product then served as a megaprimer in the following step, i.e. the insertion of the gene into the target plasmid. The PCR protocols for the 2 steps were adapted from www.rf-cloning.org.

3.1.2 Protein Expression and purification

All the buffers were prepared from filtrated water using a GenPure system (TKA). Unless specified all, the chemicals used were purchased from Carl Roth and Sigma-Aldrich. The buffers were filtered and degassed wherever needed and typically stored at 4°C.

3.1.2.1 General culturing routine

DNA constructs were transformed into the respective *E. coli* expression strains, as described in 3.1.1.1. Typically, 10-15 colonies were inoculated into a 50 or 100 ml of LB

medium and incubated at 37° C in an incubator-shaker overnight. This starting culture was then used to inoculate 2 l of TB medium in a way that the OD_{600}-value of the 2l-culture was ~ 0.05. The culture was incubated shaking at 37°C until the OD_{600} reached values of 1.0- to 1.2 and thereafter supplemented with 0.5 mM IPTG to induce the recombinant protein expression. Expression was generally performed overnight at a temperature suitable for each protein (see section 3.1.2.2). The cultures were then harvested by centrifugation at 5000 rpm and 4° C for 15 minutes and the pellets are stored at -80° C.

In the case of E3 and E2 expression, cell pellets were resuspended in the respective lysis buffer (see section 3.1.2.2.1) and the cells disrupted using a Microfluidizer® bench-top processor M-110P-20 (Microfluidics) at 1500 bar (3 passages or until the lysate was clear). The lysate was then spun at 25000 rpm and 4°C for 1 hour using a Beckmann JA 25.50 rotor.

In the case of ubiquitin and variants expression, the cell pellets were resuspended in the respective lysis buffer (see section 3.1.2.2.4) and lysed by sonication using a Labsonic® device (B. Braun biotech International). To avoid heating of the cell lysate 1-minute ultrasonic frequency (>20 kHz) with a minimum break of 1 minute after each cycle was used. The resulting lysate was cleared by centrifugation at 25000 rpm and 4°C in a Beckmann JA 25.20 rotor for 40 minutes.

3.1.2.2 Protein purification

3.1.2.2.1 HUWE1 HECT domain and variants

The HUWE1 HECT domain WT and variants thereof were typically purified from a 2 l-cell culture in TB medium [110]. To this end, the respective plasmid encoding the N-terminally His_6-tagged protein was transformed into LOBSTR RIL cells (Kerafast, Boston, MA), the main culture was induced with 0.5 mM IPTG and grown at 15° C overnight. The harvested cells were resuspended in buffer A containing 80 mM HEPES pH 8.0, 500 mM NaCl, 20 mM imidazole and 10% glycerol), supplemented with cOmplete EDTA free protease inhibitor cocktail (ROCHE) (typically half a tablet is dissolved in 1 ml of lysis buffer) and 5 mM β-mercaptoethanol. The cleared lysate was passed through a 0.2 μm filter and

subjected to Nickel-affinity chromatography. To this end, the sample was applied to a 5 mL HisTrap HP column (GE Healthcare), washed with buffer A containing 5 mM β-mercaptoethanol and the bound protein eluted by a gradient of buffer B containing 80 mM HEPES pH 8.0, 500 mM NaCl, 1 M imidazole, 10 % glycerol, 5 mM β-mercaptoethanol. The gradient was set from 2.5% to 100% of buffer A with 1ml/min of flow rate. The His_6-tag protein eluted from 2.5% of 10.2 % of buffer B. The His_6-tag was cleaved from the protein with the addition of His_6-tagged precision protease (3C protease) in a molar ratio of 1:50 during dialysis at 4°C into a buffer containing 20 mM HEPES pH 8.0, 150 mM NaCl, 20 mM imidazole and 3 mM β-mercaptoethanol. The cleaved protein was separated from uncleaved protein and the protease by another round of Nickel-affinity chromatography, followed by a gel filtration using a HiLoad 16/600 SD 75 pg column (GE Healthcare) in 20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA and 5 mM DTT. Pure protein fractions were pooled, concentrated and stored in -80°C after flash freezing for future use.

3.1.2.2.2 E6AP HECT domain and variants

The E6AP HECT domain WT and variants thereof, were typically purified from a 2 l-cell culture in TB medium [71]. The respective plasmids encoding the N-terminal His_6-tag was transformed into *E. coli* BL21(DE3) cells, 2 l cultures grown in TB-medium and protein expression induced with 0.5 mM IPTG. After overnight incubation at 18° C cells were harvested by centrifugation and resuspended in a buffer containing 50 mM Tris/HCL pH 8.0, 500 mM NaCl, 5 mM benzamidine, 25 mM imidazole, half a tablet of cOmplete EDTA free protease inhibitor cocktail (ROCHE) , 3 % glycerol, 0.4 % Triton X-100 and 8 mM β-mercaptoethanol. The supernatant was passed through a 0.2 μm filter and subjected to Nickel-affinity chromatography. To this end a 5 ml HisTrap HP column was equilibrated in buffer C containing 50 mM Tris/HCL pH 8.0, 400 mM NaCl, 8 mM β-mercaptoethanol, 25 mM imidazole and bound protein eluted by applying a linear gradient of buffer D containing 50 mM Tris/HCL pH 8.0, 400 mM NaCl, 8 mM β-mercaptoethanol, 500 mM imidazole from 0% to 100% buffer C. The protein elution started at 8.6% of buffer D containing 50 mM Tris/HCL pH 8.0, 400 mM NaCl, 8 mM β-mercaptoethanol, 500 mM imidazole and lasted until 34.7%. The column was cleaned with an additional 5 CV of buffer with 500 mM imidazole, after protein elution. The His_6-tag was cleaved from the target protein by His_6-

tagged precision protease (3C protease) in a molar ratio of 1:50, during dialysis at 4°C into a buffer containing 50 mM Tris/HCL pH 7.5, 150 mM NaCl and 2 mM DTT. The cleaved protein for future experiments was separated from uncleaved fraction and the His_6-tagged protease by another round of Nickel affinity chromatography, followed by gel filtration using a HiLoad 16/600 SD 75 pg column (GE Healthcare) in 50 mM Tris/HCL pH 7.5, 150 mM NaCl and 2 mM DTT. Pure protein fractions were pooled, concentrated and stored in -80°C after flash freezing for future use.

3.1.2.2.3 NEDD4 HECT domain and variants

The protein expression and purification protocol for the NEDD4 HECT domain and C-lobe is the same as described for E6AP and its variants in section 3.1.2.2.2**.** In these cases, however, the dialysis and gel filtration were performed in a buffer containing 100 mM HEPES pH 7.5 and 300 mM NaCl.

3.1.2.2.4 Ubiquitin and variants

Ubiquitin WT and variants were purified following a protocol previously described [164], typically from a 2 l culture in TB medium The plasmid pET30a encoding the protein without tag, was transformed into *E. coli* BL21(DE3), expression induced with 0.5 mM IPTG and the cell cultured at 20° C overnight. The harvested cells were resuspended in a buffer containing 50 mM Tris/HCL pH 7.5 and 50 mM NaCl and lysed by sonication. The lysate was cleared by centrifugation at 25000 rpm for 35 minutes at 4°C. The supernatant was transferred to a small beaker and 0.5 ml of 75% perchloric acid was slowly added while stirring on ice under the hood. The formed precipitate was separated from the soluble (ubiquitin-containing) fraction by centrifugation at 8000 rpm and 4°C for 30 minutes. The supernatant was then dialyzed overnight into 50 mM NH_4COOH, pH 4.5 at 4°C, passed through a 0.2 μM filter and subjected to cation exchange chromatography. To this end a 5 ml HiTrap HP-SP column (GE Healthcare) was equilibrated with buffer D containing 50 mM NH_4COOH, pH 4.5 and ubiquitin was eluted by a linear gradient of buffer E containing 50 mM NH_4COOH, pH 4.5, 500 mM NaCl. The column was washed for 5 CV with buffer D and eluted by using a gradient from 0 to 44% buffer E in 4 CVs. Ubiquitin eluted at 45% of buffer E in (6 CV). Ubiquitin-containing fractions were then buffer-exchanged through

dialysis into 50 mM Tris/HCL pH 7.5 and 100 mM NaCl, concentrated and subjected to gel filtration using a HiLoad 16/600 SD 75 pg column (GE Healthcare) in the same buffer. The fractions containing pure ubiquitin were pooled, concentrated and stored at -80°C after flash freezing for future use. Cysteine-containing variants of ubiquitin were purified according to the same protocol, but with the addition of 2 mM DTT in each buffer used.

3.1.2.2.5 Ubiquitin-PA

The expression plasmid (pTXB1-Ub) was transformed into *E. coli* BL21(DE3). Pre-cultures were grown in LB medium and main cultures in 4 l TB medium. Protein expression was induced with 0.5 mM IPTG at A_{600} ~ 0.8 and conducted at 20°C overnight. The cells were harvested by centrifugation and resuspended in a buffer containing 50 mM HEPES pH 8.0 and 100 mM NaCl. Cells were harvested, lysed and cleared, as described above. The lysate was then passed through a 0.2 μm filter and transferred to a gravity-flow chromatography column filled with 60 ml of NEB chitin resin, equilibrated in 50 mM HEPES pH 8.0 and 100 mM NaCl. This binding step was performed at 4° C on a rocker for 5 hours. The beads were washed with 200 ml of 50 mM HEPES pH 8.0 and 100 mM NaCl, before overnight cleavage. The beads with bound ubiquitin were then subjected to overnight cleavage (~20 hours) at RT on a rocker after addition of 300 ml of 150 mM sodium 2-mercaptoethanesulfonate (MESNa, Sigma-Aldrich) in 50 mM HEPES pH 8.0, 100 mM NaCl. The cleaved product, Ub~MESNa thioester, was then concentrated to a volume of 30 to 60 ml and dialyzed at 4°C overnight to get rid of the excess MESNa. The dialyzed Ub~MESNa thioester product was then incubated with 150 mM propargylamine (Sigma-Aldrich) at RT under the hood for 5 hours. The sample was then dialyzed at 4° C overnight to remove excess propargylamine, concentrated and subjected to gel filtration with SD 75 16/600 GL column (GE Healthcare) in 50 mM HEPES pH 8.0 and 100 mM NaCl. Pure fractions, as determined by SDS-PAGE, were pooled, concentrated and stored at -80° C for future use.

3.1.2.2.6 HUWE1 HECT domain-Ub-PA conjugate

The preparation of the HUWE1 HECT domain-Ub-PA conjugate was performed based on the individual proteins, described in the above sections, 3.1.2.3.1 and 3.1.2.3.8. The proteins mixed at a molar ratio of 1:5 and incubated at 30^{o} C overnight and subjected to gel filtration using a SD 75 16/600 GL column (GE Healthcare) equilibrated in 20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM EDTA and 5 mM DTT. The fractions were analyzed by SDS-PAGE, pooled and concentrated to 10mg/ml. For crystallography, the pure protein was used fresh without freezing. The left-over protein was flash frozen and stored at -80$^{\Omega}$.

3.1.2.2.7 K48 linked-di ubiquitin HUWE1 HECT complex

A ternary complex containing K48-linked-di ubiquitin and a HUWE1 HECT domain variant (HUWE1 HECT act-Cys variant) was prepared by Han Zhou, a graduate student in Prof. Jun Yin's laboratory at Georgia State University, GA, USA. and sent to me for structural studies. To this end, additional purification of the shipped material was required, including gel filtration using a SD75 10/300 GL column (GE Healthcare) in a buffer containing 25 mM Tris pH 7.5, 100 mM NaCl. Further purification was done by anion exchange chromatography using a Mono Q 4.6/100 PE column (GE Healthcare). The latter column was equilibrated in a buffer F containing 20 mM Tris/HCl pH 8.0 and 20 mM NaCl and bound protein eluted with a an increasing gradient of buffer G 20 mM Tris/HCL pH 8.0, 1 M NaCl from 0 to 100% over 12 CV in buffer F. Additional purification attempts included Nickel-affinity chromatography using a HisTrap™HP 1 ml column (GE Healthcare) and a gradient of 0 to 100% buffer G over 12 CVs as specified below:

- Buffer A: 20 mM Tris pH 7.5, 100 mM NaCl and 5 mM imidazole
- Buffer B: 20 mM Trsi pH 7.5, 100 mM NaCl, 100 mM imidazole. The gradient was set from an increasing 0-100% buffer B over 12 CV

3.1.2.2.8 UBE2L3

The expression and purification procedure were similar to that of E6AP as described above 3.1.2.3.3**.** However, the buffers used for the purification were different, as specified in the following:

- lysis buffer: 50 mM Tris/HCl pH 8.0, 500 mM NaCl, 5 mM benzamidine, 8 mM β-mercaptoethanol, 0.4% Triton X-100
- Ni-affinity buffer A: 50 mM Tris/HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 8 mM β-mercaptoethanol
- Ni-affinity buffer B: 50 mM Tris/HCl pH 8.0, 500 mM NaCl, 250 mM imidazole, 8 mM β-mercaptoethanol
- dialysis buffer: 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 2 mM DTT
- gel filtration buffer: 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 2 mM DTT

3.1.2.2.9 Accessory proteins

UBA1 was prepared by Dr. Sonja Lorenz, as described [41]. The precision protease and TEV protease used for this study were prepared by Julia Haubenreisser, following standard protocols.

3.1.3 Protein concentration determination

Protein concentrations were determined by absorbance at 280 nm, using a Nanodrop ND 2000c spectrophotometer (Peqlab). In total 3 measurements were taken and averaged, the concentration calculated based on the Lambert-Beer law (see below) and the respective molar extinction coefficient [151].

$$c = \frac{A_{280}}{\varepsilon * \mathrm{d}}$$

where c is the concentration (mol l^{-1}); A_{280} the absorbance; ε the molar extinction coefficient (l $mol^{-1}cm^{-1}$); and d the path length of the light

3.2 Biochemical and biophysical methods

3.2.1 Electrophoretic methods

3.2.1.1 SDS-PAGE

12% gels were best suited for separating proteins in the desired size range of 10 to 70 kDa. The resolving gel contained 12% (w/v) bis-acrylamide 29:1, 375 mM Tris/HCl pH 8.8, 0.1% (w/v) SDS, 0.1 % APS, 0.025 % TEMED. The stacking gel contained 4%(w/v) bis-acrylamide, 29:1, 125 mM Tris/HCl pH 6.8, 0.1%(w/v) SDS, 0.1% APS, 0.025% TEMED. The samples were mixed with 4X SDS loading dye (62.5 mM Tris/HCl pH 7.0, 40 mM EDTA, 15% (w/v) SDS, 48% (w/v) glycerol, 0.04% (w/v) bromophenol blue, $\pm$ 120 mM β-mercaptoethanol, heated at 95^{o}C for 3 minutes, spun down and then loaded alongside a protein ladder (PageRuler Plus, Thermo Fisher Scientific). The electrophoresis was carried out at RT and 180 volt for 45-50 minutes in a buffer containing 25 mM Tris, 192 mM glycine, 0.1% SDS. For Coomassie staining the gels were briefly boiled in 0.1% (w/v) Coomassie G-250, 25% (v/v) isopropanol, 10% (v/v) acetic acid, incubated for 10 minutes on a gel- rocker and then destained in 10% acetic acid, after brief heating. Images were acquired using a LI-COR odyssey scanner. For publication purposes, precast gels (Novex™ 12%, 4-12% Tris-Glycine Mini gels, Thermo Fisher Scientific) and MES running buffer (Thermo Fisher Scientific) were used.

Immunoblotting of in-house prepared gels was done in transfer buffer containing 25 mM Tris/HCl pH 7.5, 192 mM glycine, 20 % methanol. For the precast gels Novex™ 12% and 4-12% Tris-glycine gels, the transfer buffer was commercial MOPS (Thermo Fisher Scientific), with 20% methanol added to it.

3.2.1.2 Immunoblotting

Once the proteins were separated by SDS-PAGE, the gel was washed three times in pre-cooled transfer buffer (1X MOPS buffer (Thermo Fisher Scientific), 20% (v/v) methanol) for 5 minutes. A PVDF membrane (Transfer membrane ROTI®PVDF 0.45) was activated by incubation with 100% methanol for 20 seconds, washed twice with water for 5 seconds and then immersed in the transfer buffer (see section 3.2.1.2, above) until use. Whatman

filter papers and sponges were also soaked in transfer buffer before assembling a blotting 'sandwich', to be used a Mini Trans-Blot® (BioRad) cell for wet blotting, Figure 10.

After assembling a 'sandwich' as shown above (Figure 1), the chamber was filled with transfer buffer and blotting conducted in the cold room (4oC) at 100 volts with constant current for 1 hour. The membrane was then incubated in a blocking solution of blotting grade miilk powder (CarlRoth) diluted in TBST buffer (20 mM Tris/HCl pH 7.5, 150 mM NaCl and 0.1% Tween®20) at RT temperature with gentle shaking for one hours. The membrane was then probed with a primary antibody added to the blocking buffer (1:1000 dilution in blocking buffer (v/v)) at 4oC over-night. The membrane was washed three times in TBST for 5 minutes buffer and incubated at RT with a secondary-HRP conjugated antibody (diluted in the blocking buffer at 1:10000 dilution (v/v)). for 1 hour. The membrane was then washed three times with TBST for 5 minutes and imaged using a SignalFire™ ECL reagent (Cell Signaling) with a FluorchemQ Multi image system (Alpha Innotech).

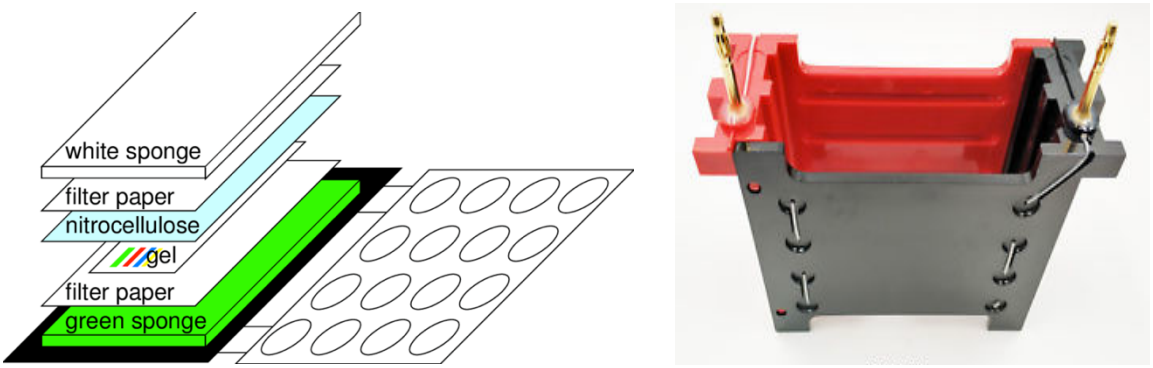


Figure 10: Western blotting

The images were taken from [165], which shows how a wet transfer sandwich is prepared prior to blotting. The black side shown in the left panel figure will face the anode side of the apparatus (black color) and the white side will face the cathode side (red color) of the apparatus shown in the right panel.

3.2.2 Thermoflour assay

The assay was carried out to assess the stability of the K48-linked-diubiquitin-HUWE1 HECT complex in different buffer conditions. 0.5 µl of 2.5% SYPRO orange fluorophore

dye (Thermo Fischer Scientific) was mixed with 0.5 µl protein (at a final concentration of 5.7 mg/ml) and the final volume adjusted to 25 µl with buffer specified in Table-1. The samples were prepared in a 96-well PP-PCR plate (Greiner Bio-One) and heated from 25°C to 95°C in 1°C-per-minute increments, using a real-time PCR cycler (Stratagene Mx3005P). The excitation and emission wavelengths were set to 492 nm and 610 nm, respectively. The data were analyzed using a routine provided by the Structural Genomic Consortium Oxford.

3.2.3 Circular dichroism

The structural integrity of the HECT domain of HUWE1 and its variants were analyzed by far-UV circular dichroism in the range of 190 to 260 nm using a JASCO J-810 spectropolarimeter. The protein concentration was 2.5 µM in 50 mM potassium phosphate and 50 mM sodium phosphate at pH 7.5. In order to improve the signal-to-noise ratio 15 spectra were recorded and the signal averaged, corrected for the buffer signal and normalized to molar ellipticity, according to the equation

$$(\theta)=\frac{\theta * 100}{c*d*N_{AS}}$$

where (θ) is the molar ellipticity in (deg cm^2 $dmol^{-1}$); θ is the measured ellipticity in mdeg; c is the protein concentration in mM; d is the thickness of the cuvette in cm; and N_{AS} is the number of amino acids of the protein.

3.2.4 Fluorescence polarization

The protocol used for the fluorescence polarization (FP) assays were adapted from published protocols [79]1mg of the thiol-reactive fluorescent probe, BODIPY® TMR C5-malemide (Thermo Fisher Scientific) was dissolved in 25 µl of 100% DMSO, vortexed and centrifuged. Ubiquitin G76C was diluted to 1mM target concentration in 50 mM Tris/HCl pH 7.4, 100 mM NaCl and 1 mM TCEP and later dialyzed in the same buffer, but without TCEP, to remove the reducing reagent. Ubiquitin G76C was then mixed with a 10-fold molar excess of the dye. The reaction was wrapped with aluminum foil to protect it from light and incubated at 4°C overnight. Excess dye was removed by desalting with a HiPrep 26/10 column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.4 and 100 mM NaCl. Thereafter, the column was cleaned extensively with 1 M NaOH and water.

FP assays were performed at 30^{o} C in FP buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 5 mM DTT and 0.01% Triton-X), in a 384-well flat-bottom microplate (Greiner Bio-One) using a Clariostar microplate reader (BMG-LabtecH), with an excitation wavelength of 540 nm and emission at 590 nm. Ubiquitin G76C-dye conjugate was used at 50 nM and the different HECT variants concentration varied from 0 to 500 µM. Average binding curves from three independent experiments were fitted according to single using the formula (see below); in OriginPro 9.1 [166]

Fluorescence polarization = 1000 * (parallel - perpendicular) / (parallel + perpendicular)

Where; parallel – represents the parallel measurement channel and perpendicular – represents the perpendicular measurement channel.

3.2.5 *In vitro* activity assays

3.2.5.1 Isopeptide bond formation

In vitro auto-ubiquitination and substrate ubiquitination assays were performed in a buffer containing 25 mM Tris (pH 7.4), 8 mM MgCl2, 100 mM NaCl 1 mM DTT, 3 mM ATP at 30°C with HECT domain of HUWE1 and its variants and for substrate ubiquitination assay, a slightly extended construct of HECT domain of HUWE1 was used (named as HUWE1-C) and its variants. The reaction was carried out in the presence of 0.2 µM E1 (UBA1), 5 µM E2 (UBE2L3), 5 µM HECT domain of HUWE1 and variants and HECT domain of HUWE1 (3843-4374) and variants, 100 µM Ub and 30 µM MIZ1 for substrate ubiquitination assays. The reactions were stopped at the indicated time points with SDS-reducing dye. The samples were loaded on 4-12% Bis-Tris (NuPAGE) gels for isopeptide bond formation assays and for MIZ1 assays. The appropriate reaction products were quantified with Image J [152] and normalized against the input enzyme (minus ATP lane). The mean and standard deviation from at least three independent experiments were plotted (see results section 4.2.4, for more details).

3.2.5.2 Thioester bond formation

For monitor the single-turnover pulse chase assays, UBE2L3~Ub thioester was generated by incubating 1 µM UBA1, 5 µM UBE2L3 and 100 µM ubiquitin WT in 25 mM HEPES, pH 7.4, 100 mM NaCl, 1 mM DTT, 8 mM MgCl2 and 2 mM ATP at 30 °C and incubated for 30 minutes; the reactions were quenched by diluting it with 50 mM EDTA and incubated on ice for another 30 minutes. The UBE2L3-ubiquitin conjugate was then incubated with 10 µM of HECT domain of HUWE1 and its variants for the indicated reaction times and the reactions were stopped by adding SDS-reducing and non-reducing dye. The samples were loaded on 12% Bis-Tris (NuPAGE) gels. The appropriate reaction products were quantified with Image J [152]. The mean and standard deviation from three independent experiments were plotted.

3.2.6 Ubiquitin-PA labelling assay

The labelling of HECT domain of HUWE1 and its variants with ubiquitin-PA (Ub-PA) was performed in a buffer containing 50 mM HEPES pH 8.0 and 100 mM NaCl at a molar ratio of 1:10 at 30^{o} C. Each protein used for labelling was diluted to a target concentration of 10 µM in a buffer containing 50 mM HEPES pH 8.0, 100 mM NaCl. Ub-PA was added directly into the reaction mix to achieve a target concentration of 100 µM. Reactions were quenched at the specified time points (10, 60 and 120 minutes) are quenched with SDS-reducing dye. Samples were analyzed with 12% SDS-PAGE gels and the protein bands were detected by Coomassie staining. The appropriate reaction and products were quantified with Image J [152]. The mean and standard deviation from three independent experiments were plotted. Unless specified, specially all the labelling is done as described above.

3.2.7 Mass spectrometry

Mass spectrometry analyses were performed in Prof. Dr. Henning Urlaub´s laboratory (Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry,Göttingen, Germany) and analyzed by Dr. Uwe Pleßmann.

The HECT domain of HUWE1 and one of its variants (Cys 4099 Ala, Cys 4367 Ala and Cys 4184 Ala) respectively were reacted with Ub-PA as discussed in section 3.2.6; the reactions were quenched with SDS-reducing dye and subjected to SDS-PAGE. The resulting bands at the size of the conjugate from both species were excised, digested with trypsin. The samples were acidified and analyzed by ESI-MS (electrospray ionization mass spectrometry). Unfortunately, the expected mass addition to the active site containing HUWE1 peptide was not observed, likely due to the susceptibility of the conjugates to hydrolysis under acidic conditions.

3.3 X-ray crystallography

3.3.1 Protein crystallization and data collection

Crystals of HUWE1(3993–4374) conjugated with ubiquitin-PA initially grew at 20°C in sitting drops in 100 mM HEPES pH 7.0, 800 mM sodium phosphate monobasic and 800 mM potassium phosphate dibasic. Crystallization was performed at three different protein concentrations of 5, 10 and 15 mg/ml using a HoneyBee 963 crystallization robot (Digilab). The crystals were improved for diffraction quality by streak seeding. The crystals finally used for data collection had grown in sitting drops containing 100 mM HEPES pH 7.0, 600 mM sodium phosphate monobasic and 600 mM potassium phosphate dibasic with streak seeding and were cryo-protected in the same solution, including 10% glycerol.

3.3.2 Structure determination and refinement

Diffraction data were collected at the European Molecular Biology Laboratory (EMBL),Hamburg, Germany, beamline P14 to 2.3 Å resolution and were processed with XDS[162]Molecular replacement was performed with Phaser [156]as implemented in the python-based hierarchical environment for integrated crystallography (PHENIX) [157] The structure of the HUWE1 HECT domain (PDB ID: 3H1D) was split into its two individual lobes as search models: C-lobe (residues 3993-4256) and N-lobe (residues 4257-4374). Refinement was performed using Phenix.refine with individual B-factors, rigid body and torsion angle NCS (non-crystallographic symmetry) restraints. Manual model building was performed in Coot [147]The electron density illustrations were created with phenix.maps

tool and the structural images of crystal structures were created with PyMOL (open source, V1.7.6; DeLano Scientific LLC). Dr. Dan Chen kindly performed the data collection and also helped with data processing and refinement.

4 RESULTS AND DISCUSSION

4.1 Ubiquitin-PA as a tool to study ubiquitin recognition by the HECT domain of HUWE1

The preparation of conjugates of HECT-type ligases with donor ubiquitin for structural studies is challenging, since the thioester linkage between the catalytic cysteine in the HECT domain and the C-terminus of the donor ubiquitin is highly susceptible to hydrolysis. One way to generate such a complex is to replace the thioester linkage by a disulfide bond [80,88], such chemical mimics are also shown for E2 enzymes [167,168]. Oxyester and isopeptide-bonded complexes have also been prepared. As discussed in section 1.3.4.2, a recent study highlighted that ubiquitin C-terminal electrophiles may be used to generated complexes with HECT-type ligases [142], offering opportunities for structural analyses, analogously to DUBs [126]. We thus sought to use ubiquitin-PA (from here referred to as “Ub-PA”) as a tool to study ubiquitin recognition by the HECT domain of HUWE1. Ub-PA can be easily expressed in *E. coli* with relatively high yields (20 mg per liter of bacterial culture) - a pre-requisite for structural studies. I aimed at using Ub-PA to unravel mechanistic features of HUWE1 HECT domain to:

- compare the labeling of selected HECT domains with Ub-PA systematically to understand the basis of reactivity
- generate a HUWE1 HECT domain-Ub-PA complex for structural studies
- determine its structure by X-ray crystallography
- validate structural features by structure-guided mutagenesis in combination with biochemical assays
- use the complex to reconstitute larger complexes with additional ubiquitin molecules and/or substrates (ongoing) and elucidate mechanistic features of the HECT domain of HUWE1

4.1.1 Preparation of ubiquitin-PA

The first step towards using Ub-PA as a tool to study HECT E3 mechanism was to prepare Ub-PA based on a construct kindly provided by Prof. David Komander [169]. This construct (see section 2.1.2.2.5) encodes ubiquitin residues 1-75 with a C-terminal intein-

chitin binding domain (CBD) affinity tag. The intein has a cysteine residue, which attacks the carbonyl oxygen of the peptide bond formed between Arg 74 and Gly 75 of ubiquitin in an intramolecular N-S acyl shift (Figure 11). The resulting intermediate can be attacked by a thiol-containing reagent, such as sodium 2-mercaptoethanesulfonate (MESNa). A transthioesterification reaction yields a ubiquitin~MESNa thioester intermediate, which then reacts with propargylamine (Figure 11) to generate the desired product, ubiquitin-PA. Ubiquitin-PA can then be used to generate a stable vinyl-thioether linkage with the catalytic cysteine in the HECT domain of HUWE1.

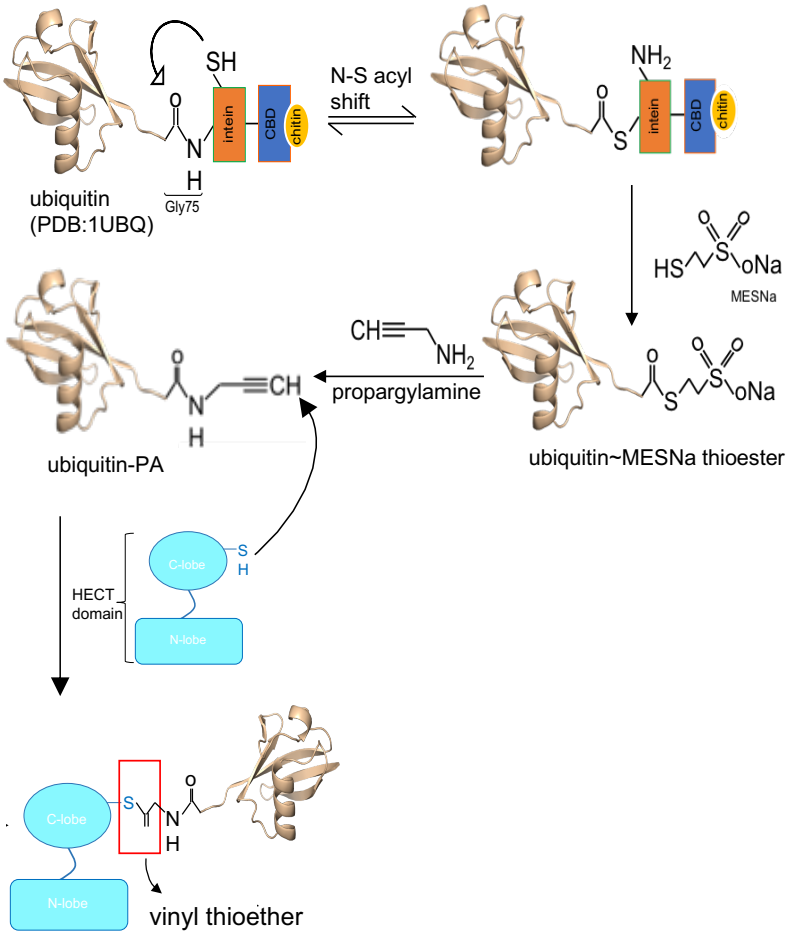


Figure 11: Schematic of the generation of HECT domain-linked Ub-PA

Diagram detailing the reaction steps to generate a stable vinyl-thioether linkage between ubiquitin and the catalytic cysteine of a HECT domain by intein-based chemical ligation.

Ub1-75-intein-CBD was expressed and purified as outlined in section 3.1.2.3.8 and the progress of the purification followed by SDS-PAGE (Figure 12): Lane 2 represents the flow-through of the initial chitin column and shows a significant fraction of unbound tagged protein, indicating that more resin should have been used. Lane 3 represents the washing step, and lane 4 the sample after overnight MESNa-induced cleavage from the chitin resin (thioester product MESNa~Ub), the sample shows a faint band running at an expected molecular weight of 8.5 kDa. The samples in lane 5,6,7 and 8 are fractions after gel filtration (Figure 12(B), which shows a mix of un-cleaved ubiquitin-CBD (30 kDa) and cleaved CBD-tag (27 kDa) which was eluted from the chitin column following MESNa cleavage. The purity of Ub-PA can be assed from section 4.1.3. The yield of pure Ub-PA amounted to 27 mg from 4L of bacterial culture. The protein was next used for comparative labeling assays with HECT domains of different ligases and for the generation of a stable donor-like ubiquitin conjugate of the HECT domain of HUWE1 for structural and mechanistic analyses.

(A) SDS-PAGE **(B) SEC**

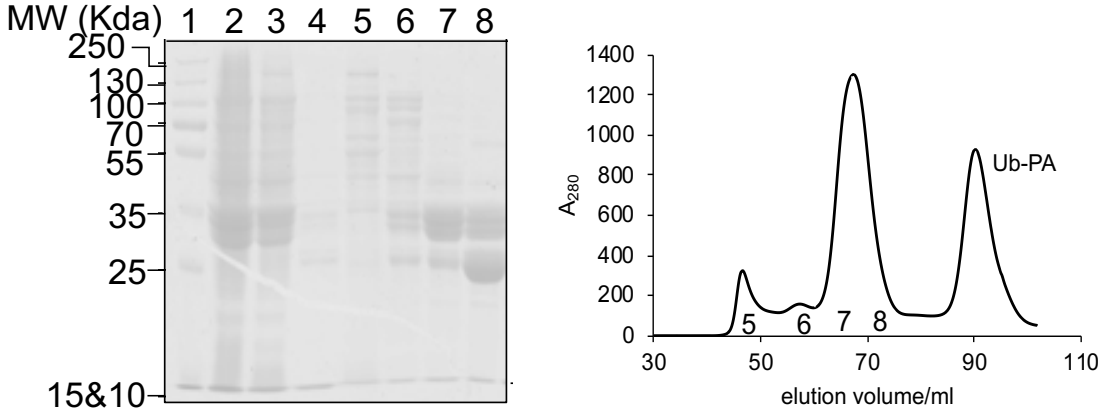


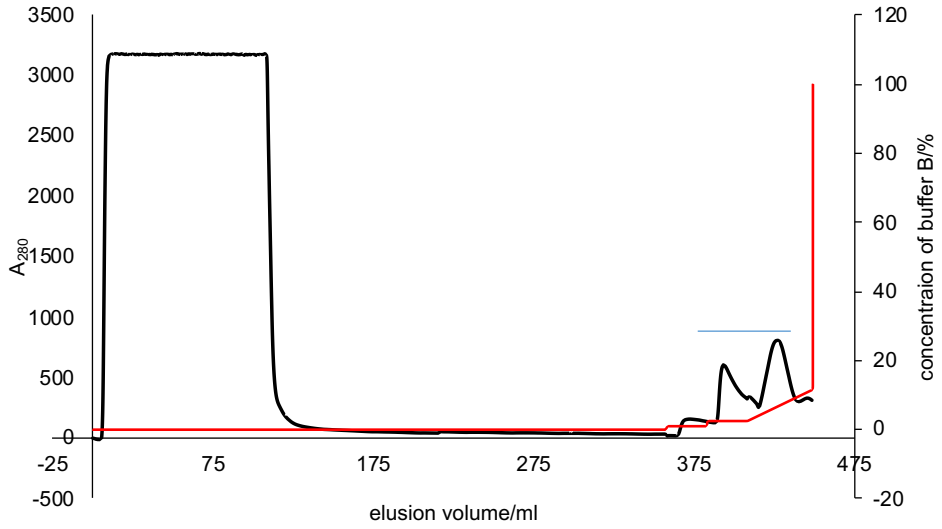
Figure 12: SDS-PAGE and SEC in the preparation of Ub-PA

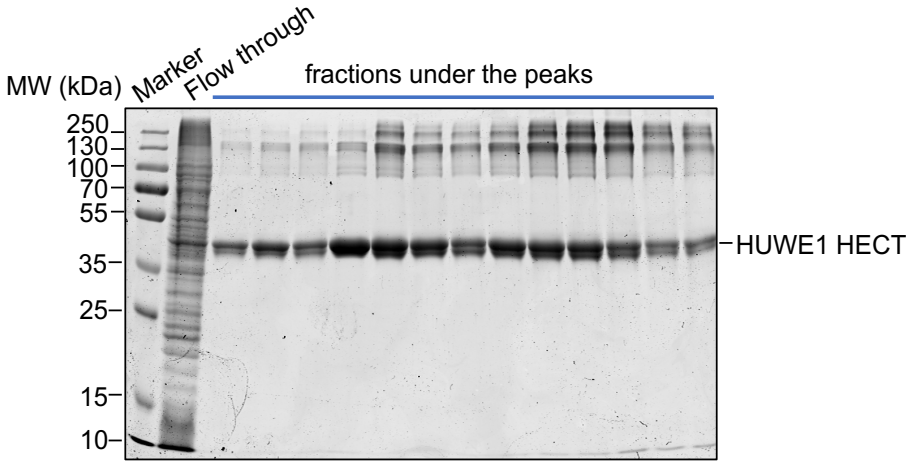
(A) SDS-PAGE showing the progress of the purification of Ub-PA: 1 – marker; 2,3,4-samples taken during the chitin column chromatography; 5,6,7,8 – samples taken after the SEC **(B)** SEC of Ub-PA with a HiLoad SD 16/600 75 pg: 5,6,7,8 – corresponds to samples loaded on the SDS-PAGE

4.1.2 Preparation of the HECT domains of HUWE1, NEDD4 and E6AP

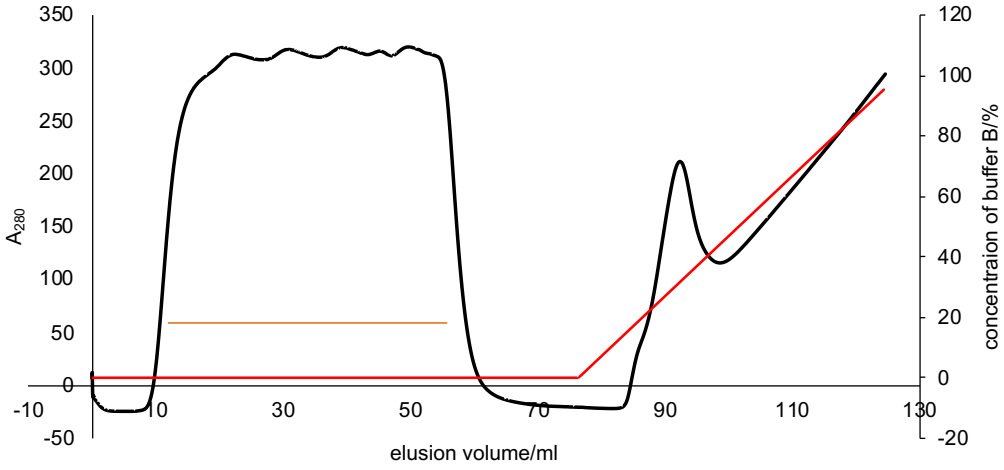
This section describes the preparation of the HECT domains of selected ligases, HUWE1, NEDD4 and E6AP to be tested in labeling reactions with Ub-PA. After harvesting *E. coli* cells recombinantly expressing the individual HECT domains, the proteins were purified according to section 2.1.2.2. Figure 13, shows a representative SDS PAGE of the initial Nickel-affinity chromatography (A) and the corresponding SDS-PAGE for the preparation of the HUWE1 HECT domain; (B) a second round of Nickel-affinity chromatography

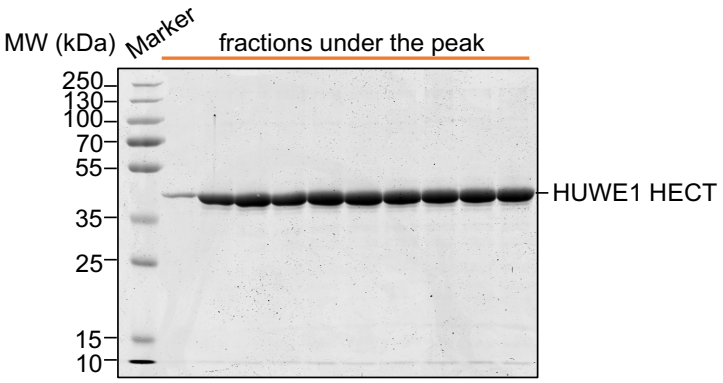
(A) Ni-affinity chromatography



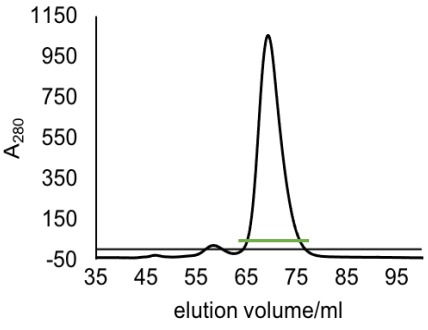


(B) Reverse Ni-affinity chromatography





(C) Size-exclusion chromatograhy



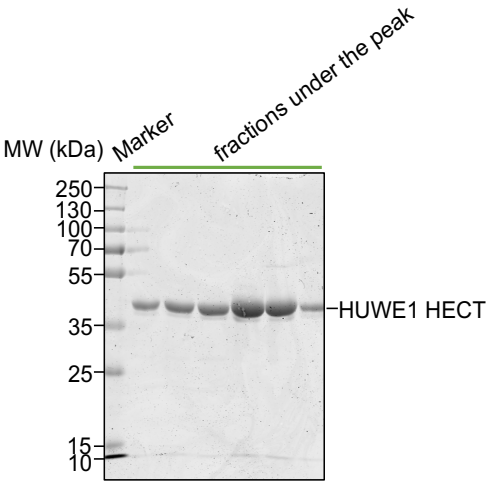


Figure 13: Purification of the HUWE1 HECT domain

(A) Ni-NTA affinity chromatography profile (left); black line: absorption at 280 nm; red line: concentration of buffer B in percent; corresponding SDS-PAGE (right); the blue line correlates the samples loaded with the chromatogram. **(B)** inverse-NTA affinity chromatography (left); black line: absorption at 280 nm; red line: concentration of buffer B in with the chromatogram. **(C)** SEC profile (HiLoad 16/600 75 pg) (left); corresponding SDS-PAGE (right); the green line correlates the loaded samples with the chromatogram.

(inverse) and the corresponding SDS-PAGE; (C) the subsequent size-exclusion chromatography and the corresponding SDS-PAGE. The purification yielded a pure, homogenous amount (24 mg per liter bacterial culture) of the protein suitable for my subsequent experiments. Analogous size-exclusion chromatographic profiles for the HECT domain of NEDD4 and E6AP is shown in Figure 14. The purity of the latter two proteins can be assessed from the subsequent labelling assays (Figure 15).

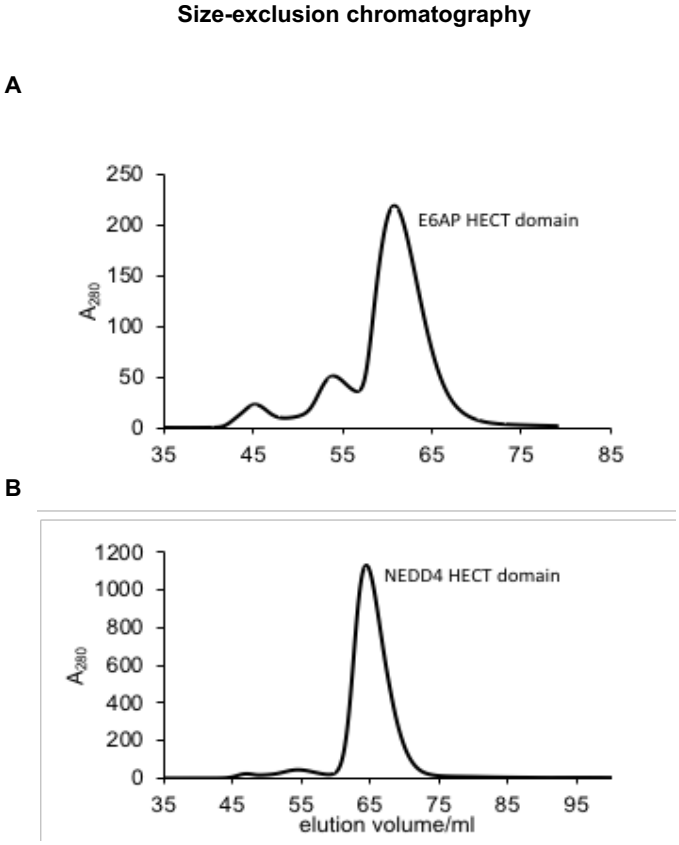


Figure 14: Size-exclusion chromatography of the E6AP and NEDD4 HECT domains

(A) Size-exclusion chromatography of E6AP with HiLoad 16/600 75 pg. **(B)** Size-exclusion chromatography of NEDD4 with HiLoad 16/600 75 pg.

Following the successful purification of the individual HECT domains of HUWE1, NEDD4 and E6AP, I subjected them to labeling reactions with Ub-PA.

4.1.3 HECT ligase profiling with Ubiquitin-PA

Previous studies [142] using Ub-PA to modify HECT ligases reported rather low turnover. However, these studies used purified GST-tagged HECT domains, and the consequences of GST-dimerization on the reactivity of the HECT domain are not known. To understand the limiting factors in this reaction, I set up comparative labeling assays with the HECT domains of HUWE1, NEDD4 and E6AP in their untagged, monomeric form (see section 3.1.2). As a negative control, I used a catalytically dead variant of the HUWE1 HECT domain, in which the catalytic cysteine, Cys 4341, had been replaced by Ala.

The HECT domains of HUWE1 and NEDD4 show almost complete labelling (Figure 15) within 2 hours of incubation at 30°C, whereas E6AP does not. Additionally, the assay indicates that the active-site cysteine of HUWE1 is labeled exclusively, since (i) the C4341A variant shows no modification and (ii) no higher-order products besides the mono-ubiquitin conjugate are observed. This is remarkable because the HECT domains of HUWE1 and NEDD4 have each more than one cysteine residue. For example, the HECT domain of HUWE1 has six cysteines, four of which are surface-exposed, including the active site. Nevertheless, Ub-PA specifically labelled the active-site cysteine. This indicates that the active-site (as expected) is either significantly more reactive towards Ub-PA than other cysteines and/or the presence of a ubiquitin binding site near the catalytic center (donor binding site) is required for the reaction with Ub-PA.

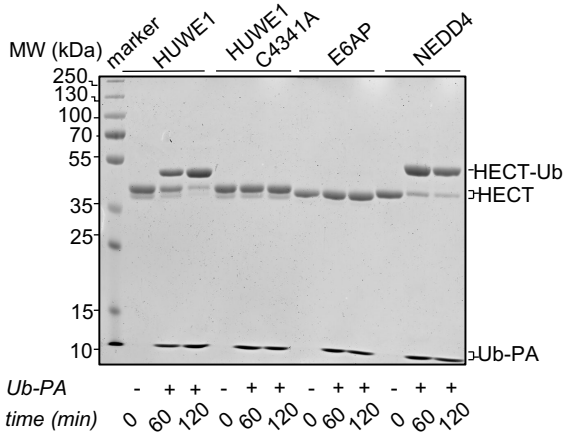


Figure 15: HECT ligase reactivity profiling

Ub-PA labelling reaction monitoring protein turnover at 60 and 120 and visualized by SDS-PAGE; minus Ub-PA is the negative control. Reactions were carried out at 30°C.

I next aimed at analyzing the differences in reactivity between E6AP and HUWE1/NEDD4 towards Ub-PA. First of all, it is interesting to note that HUWE1 is the closest phylogenetic neighbor of the NEDD4-subfamily, while E6AP is more divergent (Figure 16).

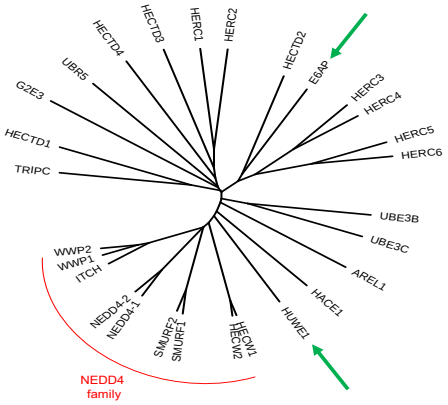


Figure 16: Phylogenetic analysis of the 28 human HECT ligases

The C-lobe sequences all 28 human HECT ligases were aligned using Clustal Omega and converted to a phylogenetic tree using the Interactive Tree of Life (iTOL) v3 server. Green arrows point to HUWE1 and E6AP (tree generated by Dr. Sonja Lorenz [71]).

I noticed a loop region in the C-lobe of HUWE1 and NEDD4 that is not found in E6AP (Figure 17). A structural superposition of the C-lobes of HUWE1, NEDD4 and E6AP highlighting this loop region is shown (Figure 18). I hypothesized that this non-conserved loop might enhance the reactivity of HUWE1/NEDD4 compared to E6AP toward Ub-PA. I thus engineered the non-conserved loop from the HECT domain of NEDD4 to E6AP (the chimeric protein is referred to as 'E6APloopinsert') and deleted the loop in the HECT domain of HUWE1 (referred to as 'HUWE1loopdelete'). The labelling assay with the purified, chimeric proteins did not result in an inversion of the labeling behavior (Figure 19); instead, the HUWE1 variant is labeled efficiently as the WT protein.

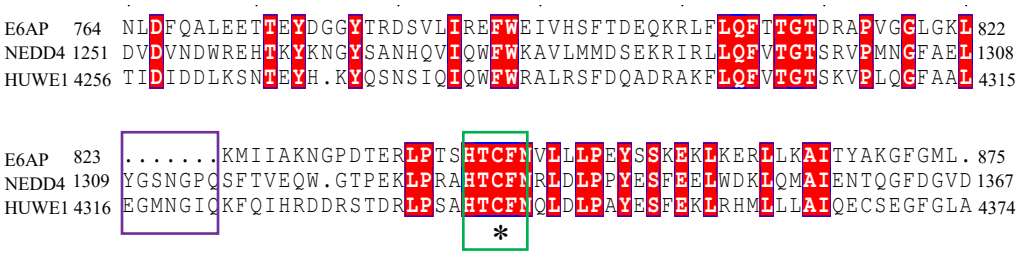


Figure 17: Amino acid sequence alignment for the C-lobes of human E6AP, NEDD4 and HUWE1

The sequence alignment was done with Clustal Omega [69] and visualized with ESPript [170] following a coloring scheme of % equivalence. The violet box marks the non-conserved loop residues, the green box marks the conserved sequence around the active-site cysteine (asterisk).

and the engineered E6AP variant still does not react with Ub-PA. I conclude that the non-conserved loop region does not determine the different labeling efficiencies of HUWE1 and E6AP with Ub-PA. While it remains, unclear what determines the reactivity of the tested HECT domains, it is possible that the N-lobe contributes to the observed differences. I next focused on the HECT domain of HUWE1 due to its high reactivity to prepare a Ub-PA-linked ubiquitin conjugate for structural and mechanistic analyses.

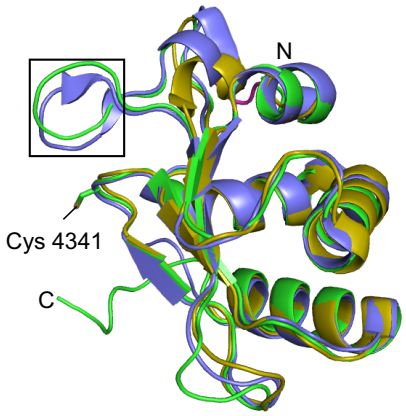


Figure 18: Superposition of crystal structures of the C-lobes of HUWE1, E6AP and NEDD4

C-lobe of HUWE1 in green (extracted from PDB ID 6XZ1), NEDD4 in blue (extracted from PDB: 4BBN [80]) and E6AP in yellow (extracted from PDB: 1D5F [72]). The structures are shown in ribbon representation. The side chain of the catalytic cysteine of HUWE1 is shown as a stick model. The non-conserved loop of NEDD4 and HUWE1 is boxed and the N- and C-terminus are marked.

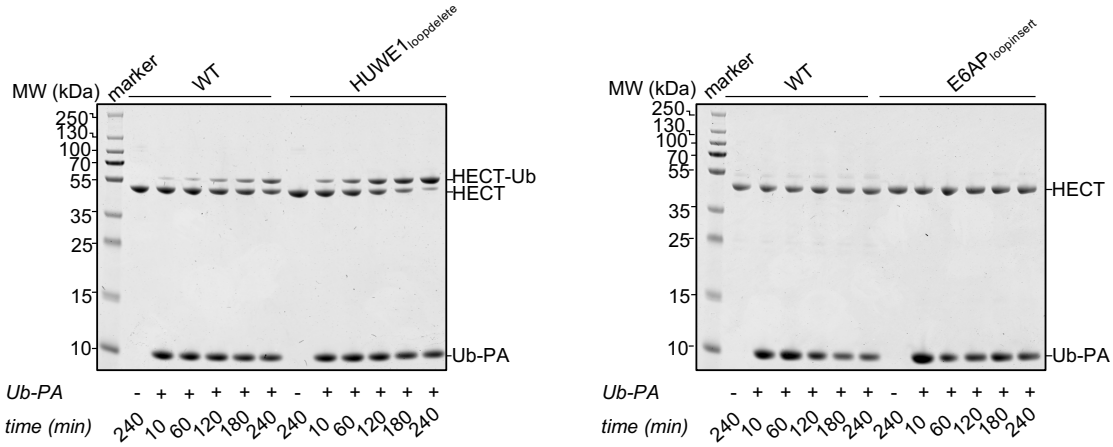


Figure 19: Role of the non-conserved loop region within the C-lobe for Ub-PA labeling

Ub-PA labelling assay monitoring the protein turnover at 10, 60, 120, 180 and 240 minutes and visualized by SDS-PAGE; minus Ub-PA is the negative control. Reactions were carried out at 30°C.

4.2 Donor ubiquitin recognition by the HECT domain of HUWE1

In order to generate the HUWE1 HECT domain-Ub-PA conjugate in preparative amounts, I had initially been concerned about the presence of several surface-exposed cysteines in the HECT domain and a potential heterogeneity in labeling with Ub-PA (chronologically, these experiments were performed before I had demonstrated the specificity for C4341 (Figure 15, see above).

So, I decided to replace the three, surface-exposed, non-catalytic cysteines of the HUWE1 HECT domain (Cys 4099, 4184 and 4367) by alanine in order to reduce the risk of heterogeneous labeling by Ub-PA (Figure 20). The corresponding triple-mutant variant of the HUWE1 HECT domain (from here on referred to as 'TM') was then subjected to an isopeptide bond formation assay to compare its ability to form free ubiquitin chains and to auto-ubiquitinate.

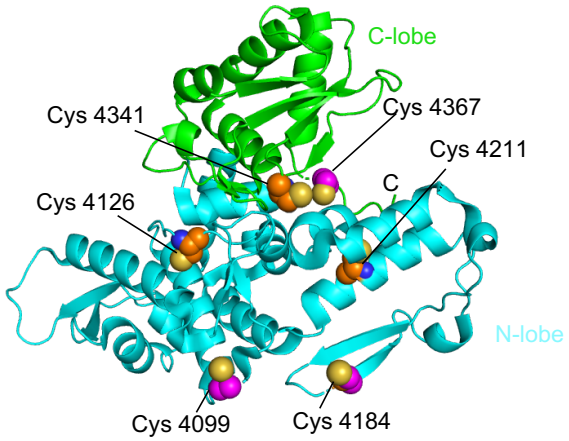


Figure 20: Crystal structure of the HUWE1 HECT domain

The crystal structure of the human HUWE1 HECT domain (extracted from PDB: 5LP8 [110], chain B). All cysteine side chains are represented by spheres. Cysteines that I chose to mutate in this work are colored magenta and the non-mutated cysteines are colored orange; nitrogen atoms are shown in blue and sulfur in light orange. C-termini is labeled.

This assay shows that the TM variant is catalytically impaired with respect to the WT (Figure 21). The reduced activity of the TM variant indicates that the cysteine-to-alanine mutations interfere with the macromolecular interfaces and/or the fold or conformational dynamics of the HECT domain required for catalysis. Therefore, the TM variant does not provide a meaningful target protein for structural studies. Nevertheless, I included the TM variant in labeling experiments with Ub-PA. To this end, 55 µM of the WT and TM variants were mixed with Ub-PA in a molar ratio of 1:10 in a buffer containing 50 mM HEPES pH 8.0 and 150 mM NaCl and the reactions stopped at four different time points (10, 30, 60 and 240 minutes). At each time point, 10 µl of the reaction was quenched by the addition of 4X SDS-loading dye and analyzed by SDS PAGE.

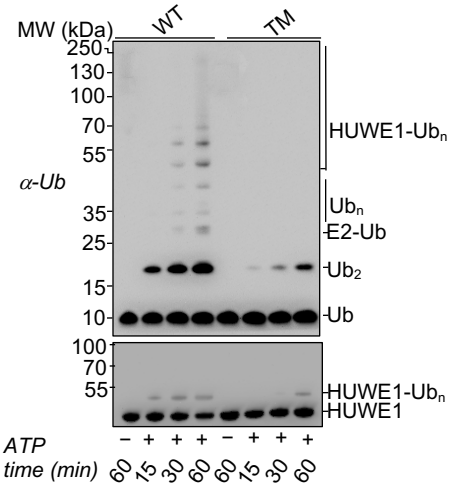


Figure 21: Role of cysteine residues in the HECT domain of HUWE1 in catalytic activity

Comparison of the *in-vitro* activities of the HECT domain of HUWE1 WT and TM, by anti-ubiquitin and anti-HUWE1 Western blotting. The assay was performed in triplicates and stopped at three different time points (15, 30 and 60 minutes), minus ATP is the negative control.

As shown in Figure 22, both the WT and the TM showed good labeling efficiencies. This observation implies that the requirements for the enzymatic and chemical reactivity of the HECT domain of HUWE1 are not the same and chemical labelling assay can still occur for a protein variant that is catalytically compromised. In the following, I focused on the

WT HECT domain of HUWE1 rather than the TM variant, since I wanted to work with a catalytically intact protein.

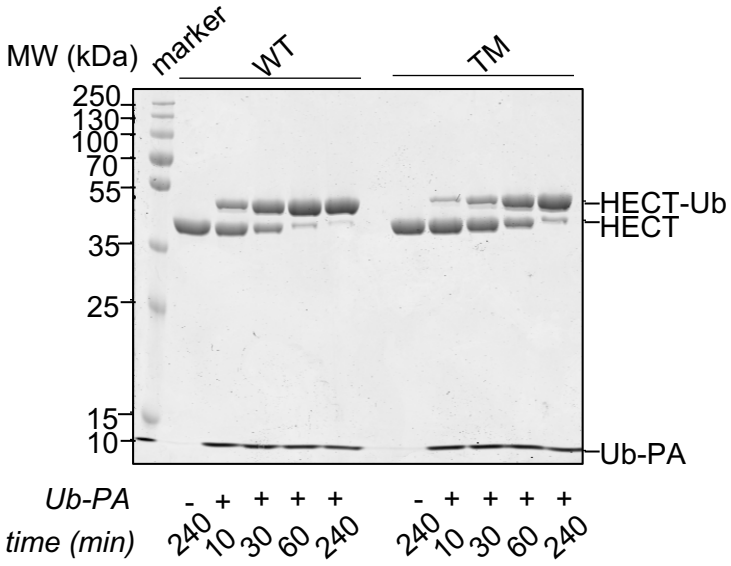


Figure 22: Role of cysteine residues in the HECT domain of HUWE1 in reactivity towards Ub-PA

Ub-PA labelling assay monitoring the protein turnover, at 10, 30, 60 and 240 minutes and visualized with SDS-PAGE. WT: HUWE1 HECT domain, TM: HUWE1 HECT triple mutant; minus Ub-PA is used as a negative control. Reactions were carried out at 30°C.

4.2.1 Crystal structure of HUWE1 HECT domain-Ub-PA conjugate

The *in-vitro* chemical labeling assays showed that Ub-PA reacts specifically with the catalytic cysteine of the HUWE1 HECT domain (Figure 15), thus presumably mimicking the donor ubiquitin. To determine the structure of this conjugate, I employed X-ray crystallography. Crystals of the HUWE1 HECT domain-Ub PA conjugate initially grew at 20 °C in sitting drops from a screening condition containing 100 mM HEPES pH 7.0, 800 mM sodium phosphate monobasic, 800 mM potassium phosphate dibasic. Crystals were further optimized to improve diffraction quality by streak seeding. The crystals used for data collection grew in sitting drops containing 100 mM HEPES pH 7.0, 600 mM sodium

phosphate monobasic, 600 mM potassium phosphate dibasic with streak seeding and were cryo-protected in the same buffer solution, including 10% glycerol.

Together with Dr. Dan Chen, I solved the crystal structure of the conjugate at a resolution of 2.8 Å by molecular replacement. Molecular replacement was performed with Phaser, as implemented in the python-based hierarchical environment for integrated crystallography (PHENIX) [157]. The structure of the HUWE1 HECT domain (PDB ID: 3H1D) was split into two sub-domains as search models: C-lobe (residues 3993-4256) and N-lobe (residues 4257-4374), see Figure 23 and Table 12. The structure shows the typical β-grasp fold of ubiquitin, with four β sheets and one helix. The HUWE1 HECT domain adopts an L-shaped conformation, as it was also seen for NEDD4-type ligases [78,79,82,87] and E6AP [72]. Other crystal structures of HECT domains, however, contained a so-called 'inverted T'- conformation [50].

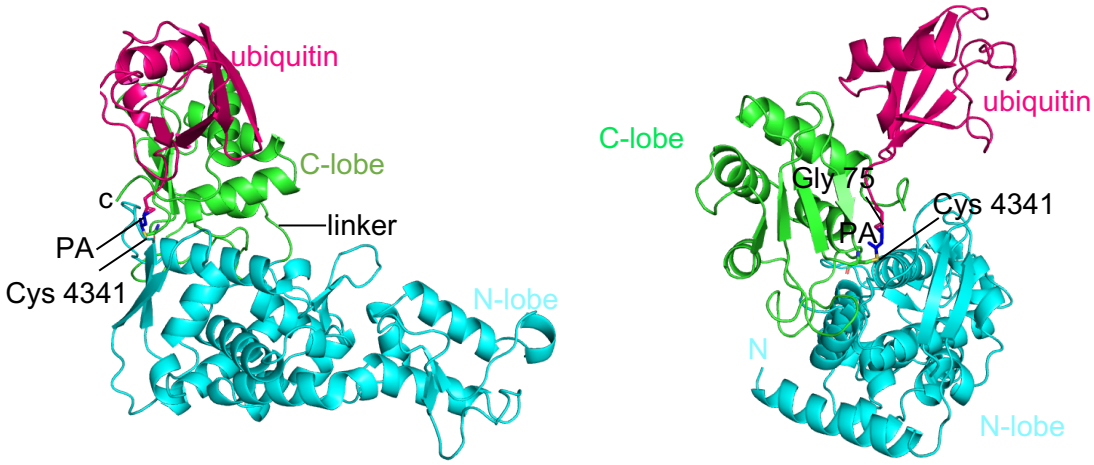


Figure 23: Crystal structure of the HUWE1 HECT domain-Ub-PA conjugate

The cartoon representation shows the HUWE1 HECT N-lobe in cyan, C-lobe in green and the bound ubiquitin in magenta. The PA linker and the catalytic cysteine are represented as ball-and-stick models.

Table 12: X-ray crystallographic data collection and refinement statistics for the structure of the HUWE1 HECT domain-Ub-PA conjugate

	HUWE1 HECT-Ub-PA conjugate	Refinement	
wavelength	0.9763	resolution (Å)	49.4 – 2.8 (2.9 – 2.8)
space group	C121	R_{work} / R_{free}	23.5 / 25.7
unit cell parameters		no. of atoms	7540
a, b, c (Å)	137.5 142.3 102.5	macromolecules	7456
α, β, γ (°)	90 129.2 90	water	84
total reflections	140314 (14709)	average B-factors	69.52
unique reflections	37296 (3729)	macromolecules	69.31
R_{pim}	5.49 (46.93)	water	65.12
completeness (%)	98.8 (99.2)	RMSD from ideality	
I/σ(I)	9.43 (1.48)	bond lengths (Å)	0.006
redundancy	3.8 (3.9)	bond angles (°)	0.71
Wilson B factor	61.9	Ramachandran statistics	
CC ½	0.998 (0.774)	favored (%)	95.38
		disallowed (%)	0.00
		MolProbity clash score	14.93
		MolProbity overall score	2.29

The C-lobe of the HECT domain makes contacts with the ubiquitin using a hydrophobic patch, which includes residues Ile 36, Leu 71 and Leu 73, from ubiquitin and Asp 4333, Leu 4335 and Arg 4336 from HUWE1 (Figure 24). This interface has been observed in other crystal structures mimicking C-lobe or HECT-domain complexes with the donor ubiquitin [80,84,87,88].

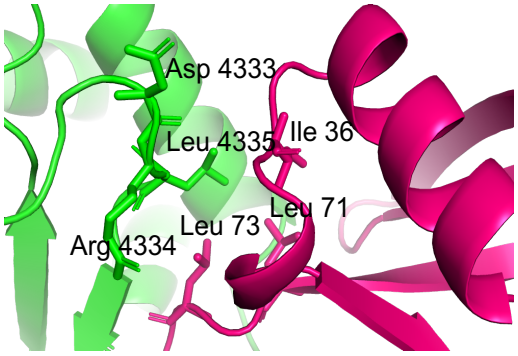


Figure 24: Detailed view of the HUWE1 C-lobe-ubiquitin interface

The side chains of hydrophobic residues (Ile 36, Leu 71 and Leu 73) of ubiquitin that contact HUWE1 are represented as sticks. The contacting residues on the HUWE1 C-lobe (Asp 4333, Arg 4334 and Leu 4335) are also represented as stick models. Ubiquitin is shown in magenta and HUWE1 C-lobe in green.

Thus, the Ub-PA labeling strategy captured ubiquitin in a donor-like orientation, mimicking the product of the first enzymatic step of ubiquitin-E3 thioester formation. The electron density map shows complete density around all the atoms of Ub-PA linked to the catalytic cysteine (Cys 4341) of HUWE1 HECT domain, confirming the vinyl-thioether linkage (Figure 25).

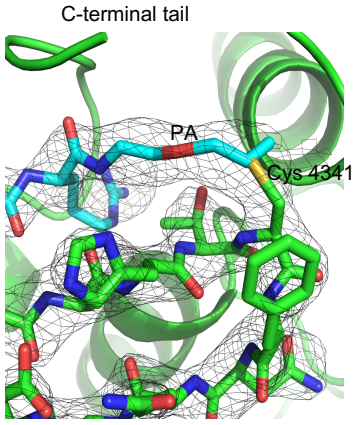


Figure 25: Detailed view on the catalytic center in the HUWE1 HECT domain-Ub-PA conjugate

The electron density (2Fo-Fc) around the linker Ub-PA, the catalytic cysteine and neighboring residues are shown as a gray mesh. The side chain of the catalytic cysteine

(sulphur atom in yellow) and PA are shown in stick representation. The C-terminal tail of the HUWE1 HECT Ub-PA conjugate is shown in stick representation.

4.2.2 Structure-based design of mutations to study the coordination of the C-tail

Our crystal structure of HECT domain of HUWE1 shows a fully resolved C-terminal tail (Figure 26). This is remarkable because previous structures of HECT domains, with the exception of an auto-inhibited, dimeric state of HUWE1, had this catalytically critical region disordered [110]. The C-terminal tail adopts a conformation in our structure that is distinct from the auto-inhibited state (Figure 27).

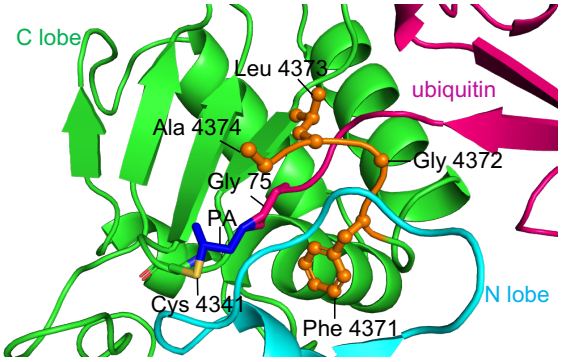


Figure 26: Detailed view of the position of C-terminal tail in the crystal structure of the HUWE1 HECT domain- Ub-PA conjugate

The side chains of the C-terminal four residues (Phe 4371, Gly 4372, Leu 4373 and Ala 4374) of the C-lobe are represented as ball-and-stick models. The PA linker (blue), the catalytic cysteine (Cys 4341, sulfur atom in yellow) and Gly 75 (magenta) are also represented as stick models.

The C-terminal tail is indispensable for isopeptide formation activity in several HECT ligases [71,80,87,171]. In our structure, the tail makes intramolecular contacts with both N-lobe and C-lobe residues, particularly Glu 4054, Glu 4064 and Leu 4061 on the N-lobe (Figure 28). To further validate the coordination of the C-terminal tail observed in the

crystal structure, these three residues were individually replaced by alanine by site-directed mutagenesis. The structural integrity of individual variants was analyzed by CD spectroscopy. As shown in Figure 29 none of the protein variants disrupted the protein fold.

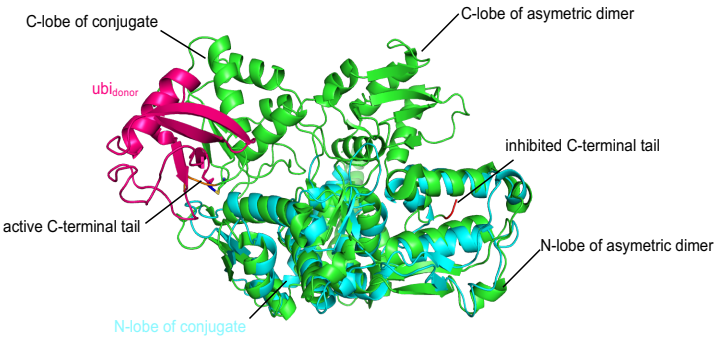


Figure 27: Superposition of the crystal structure of the HUWE1 HECT domain asymmetric dimer and HUWE1 HECT-Ub-PA conjugate

Asymmetric dimer formed of molecule B (PDB: 5LP8) [110] and the HUWE1 HECT -Ub-PA (PDB:6XZ1) conjugate crystal structure is shown in cartoon representation. The inhibited C-terminal tail of molecule B is highlighted in red and the active C-terminal tail of HUWE1 HECT-Ub-PA conjugate is highlighted in orange.

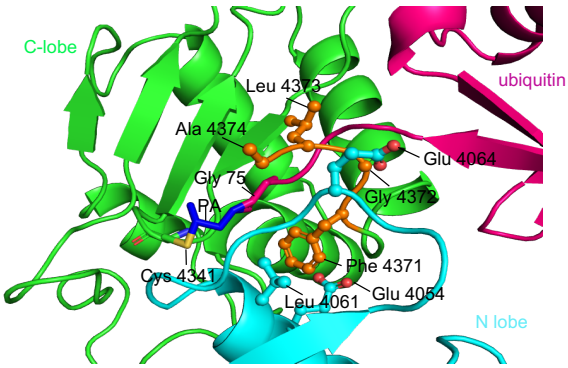


Figure 28: Structural view on the C-terminal tail-coordinating residues of the HUWE1 HECT domain-Ub-PA conjugate

The side chains of the C-terminal four residues (Phe 4371, Gly 4372, Leu 4373 and Ala 4374) of the HUWE1 HECT domain-Ub-PA conjugate are shown as ball-and-stick models.

The PA linker (blue), the catalytic cysteine (Cys 4341, sulfur atom in yellow) and Gly 75 of ubiquitin (magenta) are also represented as stick models.

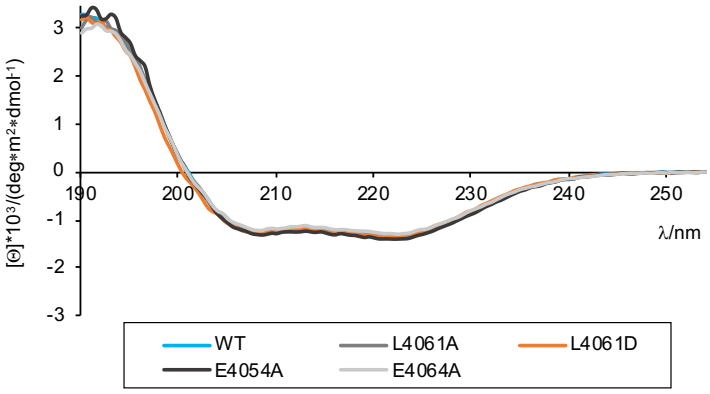


Figure 29: CD spectra of HUWE1 HECT domain variants with substitutions of the C-terminal tail-coordinating residues

Superposition of individual spectra of the indicated HUWE1 HECT domain variants. A total number of 15 spectra were measured and averaged.

After confirming the structural integrity of the variants, I checked their ability to promote isopeptide bond formation by monitoring free ubiquitin chains and to auto-ubiquitinate, compared to WT. A HUWE1 HECT domain variant, deficient of the C-terminal tail (referred to as '∆4') was used as a negative control. The ∆4 variant of the HECT domain of HUWE1 had compromised activity compared to the full-length HECT domain [141]. The assays show that the L4061A/D mutation has the most severe effects on the activity of the HECT domain of HUWE1, **(**Figure 30); the other two mutants also show a reduction in activity compared to WT, but not as strong as L4061A/D. Taken together, these data are in line with the notion that the L-conformation observed in our crystal structure and the associated coordination of the C-tail at the N-lobe-C-lobe interface is important for catalysis.

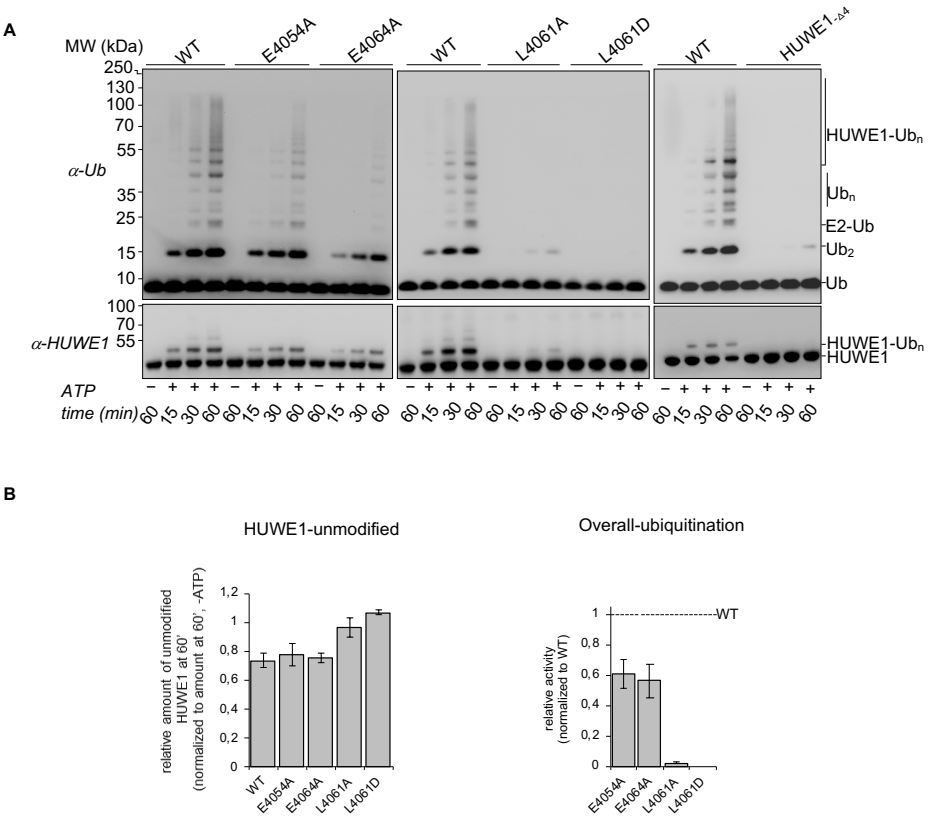


Figure 30: Significance of C-terminal tail-coordinating residues for the catalytic activity of the HECT domain of HUWE1

(A) *In-vitro* activity assay comparing the activities of the HECT domain of HUWE1 WT, E4054A, E4064A, L4061A/D and Δ4, by anti-ubiquitin and anti-HUWE1 Western blotting. **(B)** Quantification of the amounts of unmodified HUWE1 (from anti-HUWE1 blot) and overall ubiquitination (from anti-ubiquitin blot). The assays were stopped at three different time points (15, 30 and 60 minutes); the quantification is based on 3 independent experiments; the mean and standard deviation were plotted for the 60-minute time point.

Next, I tested whether these C-terminal tail-coordinating variants also play a role in substrate ubiquitination, using the physiological substrate MYC interacting zinc-finger protein 1 (MIZ1) [91] as a model. In the *in-vitro* assays, I used HA-tagged MIZ1 in order to monitor activity by anti-HA Western blotting and an extended construct of HUWE1 (3843-4374), since Dr. Bodo Sander and Barbara Orth (Lorenz lab) had shown that this construct

recognizes MIZ1. The assays show that the L4061A mutation has the most severe effects (~90% impairment) on the ubiquitination of MIZ1 by the HECT domain of HUWE1 (Figure 31). The other two variants (E4054A and E4064A) also show a reduction in MIZ1 ubiquitination compared to the WT, but not as strong as the L4061A variant. Thus, the C-terminal tail-coordinating residues in the L-conformation are important for substrate ubiquitination, as it was seen for auto-ubiquitination and free chain formation.

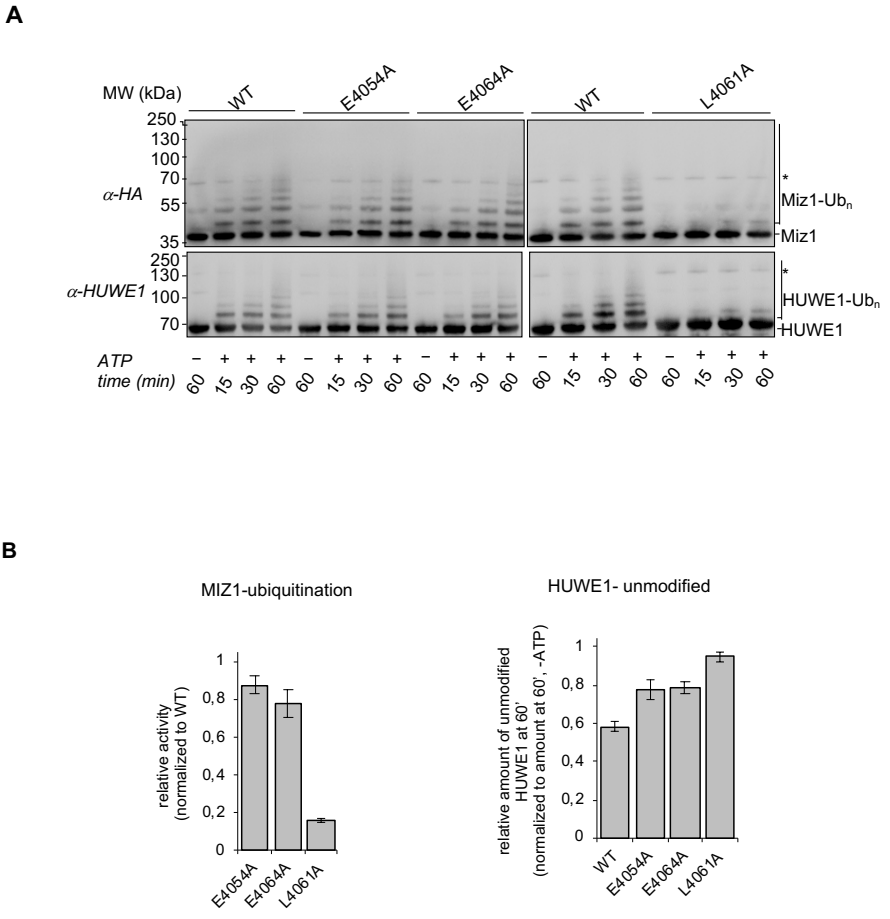


Figure 31: Significance of C-terminal tail-coordinating residues for substrate ubiquitination

(A) *In-vitro* activity assay comparing the activities of HUWE1 (3843-4374) WT, E4054A, E4064A, L4061A by anti-HA and anti-HUWE1 Western blotting. **(B)** Quantification of the amounts of unmodified HUWE1 (from anti-HUWE1 blot) and MIZ1 ubiquitination (from anti-HA blot). The assays were stopped at three different time points (15, 30 and 60

minutes); the quantification is based on 3 independent experiments; the mean and standard deviation were plotted for the 60-minute time point.

Next, I sought to investigate the role of the C-terminal tail-coordinating residues in the cell. For this purpose, I teamed up with Barbara Orth (AG Lorenz) and performed cell-based activity assays using HeLa cell lysate, monitoring the steady state level of endogenous MIZ1 by anti-MIZ1 Western blotting upon overexpression of HUWE1 (2474-4374) WT and mutants. Unfortunately, we did not observe reproducible, significant effects in this context, possibly due to the presence of endogenous HUWE1. Cellular assays may require more advanced strategies (HUWE1 knock-out), which remain to be established.

4.3 Using ubiquitin-PA to illuminate the mechanistic details of HUWE1

4.3.1 Ub-PA labelling of structure-guided HUWE1 HECT domain variants

I next explored whether the reactivity towards Ub-PA can be used to illuminate yet unknown mechanistic aspects in HUWE1. To this end, I initially monitored the reactivity of the C-terminal tail-coordinating variants (E4054A, E4064A and L4061A), towards Ub-PA. The relative labelling efficiencies of the HUWE1 HECT domain variants show a similar trend as the *in-vitro* isopeptide bond formation and substrate ubiquitination assays (Figure 30 and Figure 31 see above). The strongest reduction in labelling is seen for the L4061A and Δ4 variants, both of which show no detectable labelling anymore (Figure 32). This suggests that the C-terminal tail and its coordination by Leu 4061 on the N-lobe is also indispensable for the reactivity of the HECT domain towards Ub-PA. Taken together, the data from the enzymatic assays and labelling assays, it is clear that the coordination of the C-tail at the N-lobe-C-lobe interface is important for activity of the enzyme.

Consequently, I tested if the C-lobe alone reacts with Ub-PA. To this end, I purified the C-lobes of HUWE1, E6AP and NEDD4 and subjected them to labelling reactions with Ub-PA. These assays showed that the isolated C-lobe is insufficient to promote labelling for any of the HECT-type ligases used (Figure 33). Hence, the N-lobe is required for the reaction of the HECT domain with Ub-PA.

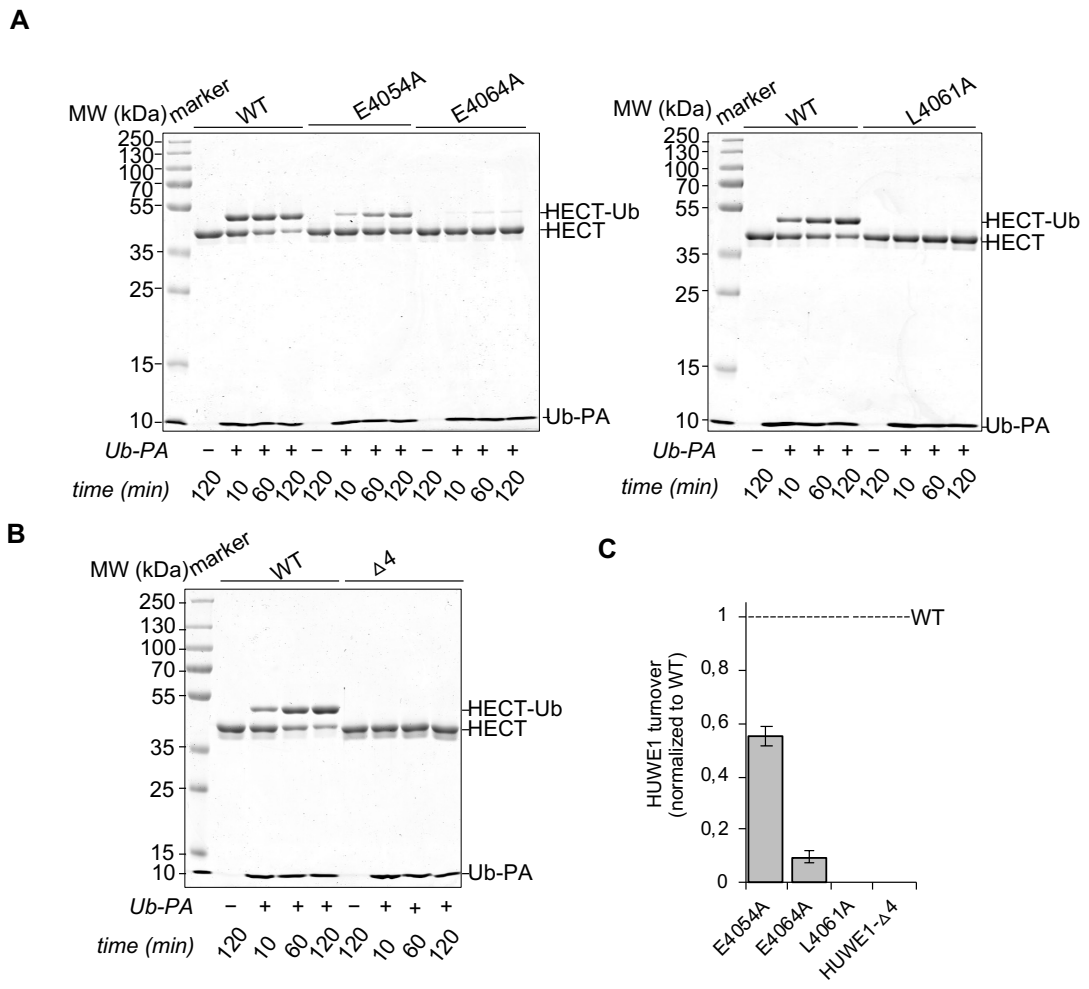


Figure 32: Reactivity of the C-terminal tail-coordinating and Δ4 variants of the HUWE1 HECT domain towards Ub-PA

(A) The labelling efficiency of C-terminal tail coordinating variants towards Ub-PA is analyzed by SDS-PAGE. **(B)** The labelling efficiency of the Δ4 variant towards Ub-PA analyzed by SDS-PAGE. All the reactions were run for 10, 60 and 120 minutes at 30°C. **(C)** Quantification of the unmodified HUWE1 HECT domain after 120', normalized to minus Ub-PA conditions (negative control).

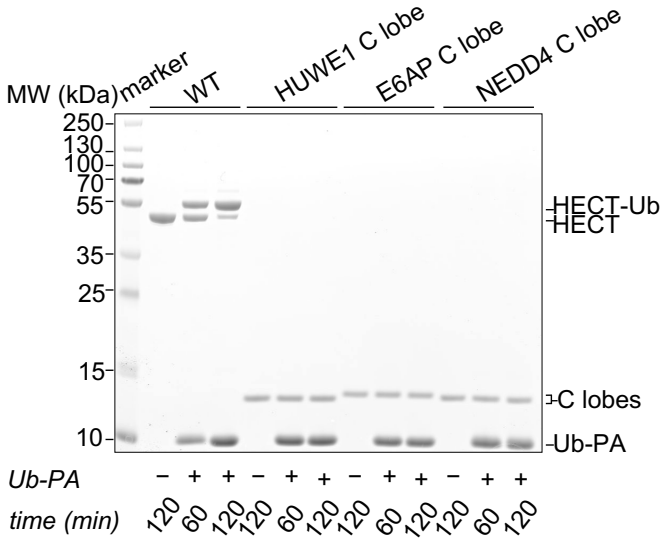


Figure 33: Ub-PA labeling reaction with the C-lobes of HUWE1, E6AP and NEDD4

The labelling efficiency of the isolated C-lobes of HUWE1, E6AP and NEDD4 with Ub-PA, analyzed by SDS-PAGE. The reactions were carried out for 60 and 120 minutes at 30°C; minus Ub-PA conditions are a negative control.

4.4 Interrogating additional ubiquitin binding sites in the HECT domain of HUWE1

4.4.1 Conformational dynamics of the HECT domain of HUWE1 affect ubiquitin binding

E6AP and many NEDD4-type ligases have an allosteric ubiquitin binding site on the N-lobe, known as “exosite” [71,79]. I interrogated the presence of such an exosite in the HECT domain of HUWE1, which had so far not been studied. I initially performed a sequence alignment of the N-lobe of HUWE1 with the N-lobes of all NEDD4- ligases and E6AP (Figure 34) and found that the exosite region is rather conserved in HUWE1. The non-covalent ubiquitin binding exosite in the case of NEDD4-type of ligases, has been shown to be important for polyubiquitin chain elongation but not for the transfer of ubiquitin from E2 to E3 or for the conjugation of the first ubiquitin to a substrate. The mutations of exosite residues equivalent in NEDD4-type enzymes (Tyr 605, Phe 707), Rsp5 (Tyr 516, Phe 618) shows a drastic reduction in binding with KD for NEDD4 F707A:

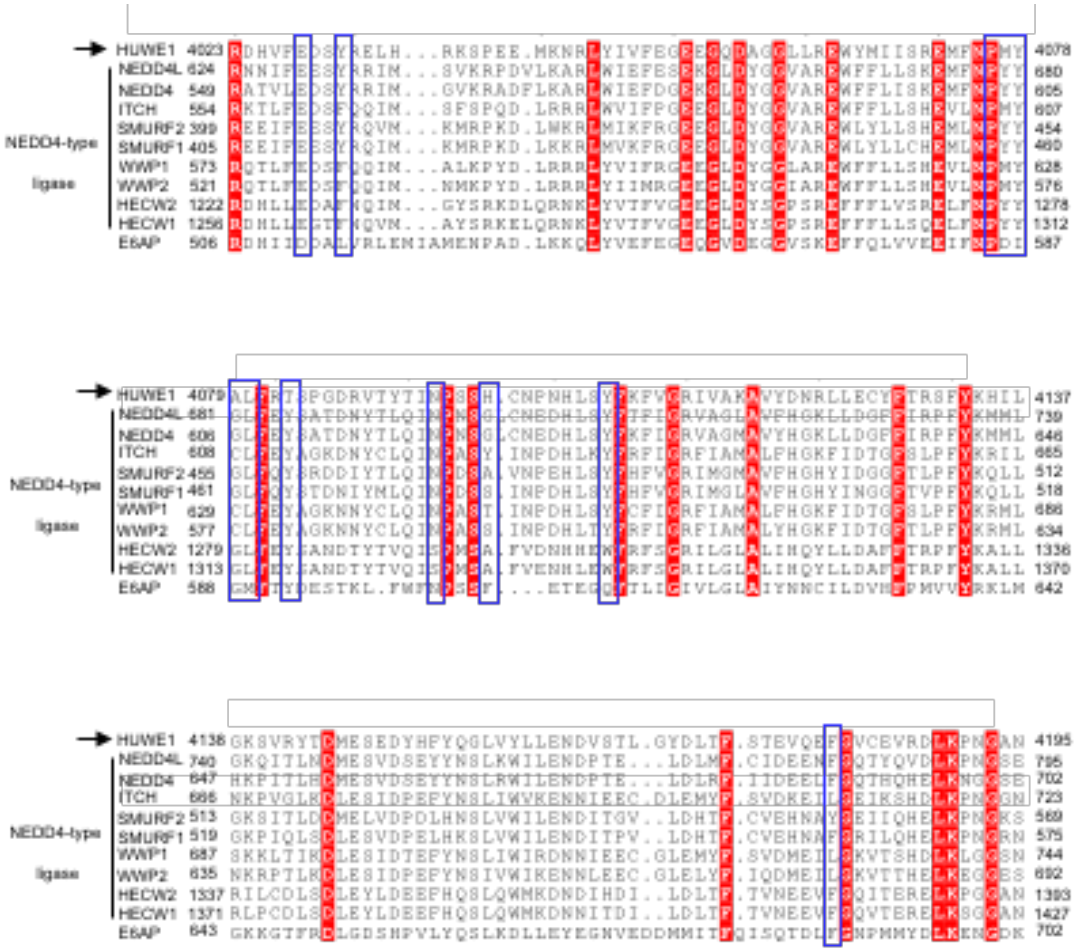


Figure 34: Amino acid sequence alignment of the N-lobe of selected human HECT ligases

Amino acid sequence alignment for the N-lobes of HUWE1, NEDD4-type ligases and E6AP, (NEDD4 isoform 4, rest isoform 1), as output by Clustal Omega [69], the alignment is illustrated by ESPript 3.0 alignment program [170] and coloring is based on PAM 250 [172,173]. Boxes refer to the position homologous to residues Tyr 4078 and Phe 4181 of HUWE1 in NEDD4-type ligases and E6AP. Residues in the exosite were defined based on the crystal structure of the ubiquitin-bound HECT domain of NEDD4 (PDB: 4BBN) [80] and a minimum surface burial upon complex formation of at least 50% determined with the PISA server [159].

340 μM; Y605A: 87 μM; Rsp5 K_D could not be determined and for E6AP (Ile 564, Phe 665), it shows a moderately weak effect with a K_D of 120 μM for I564D; 110 μM for F665D. However the tested mutations of exosite in E6AP markedly reduce the isopeptide bond formation, also in this case not affecting the transthiolation step [71,78,79].

To test whether HUWE1 binds to ubiquitin via its exosite (Tyr 4078, Phe 4181) , I performed fluorescence-polarization (FP) experiments with HECT domain variants and fluorophore-labelled ubiquitin (see section 2.2.4) (Figure 35).

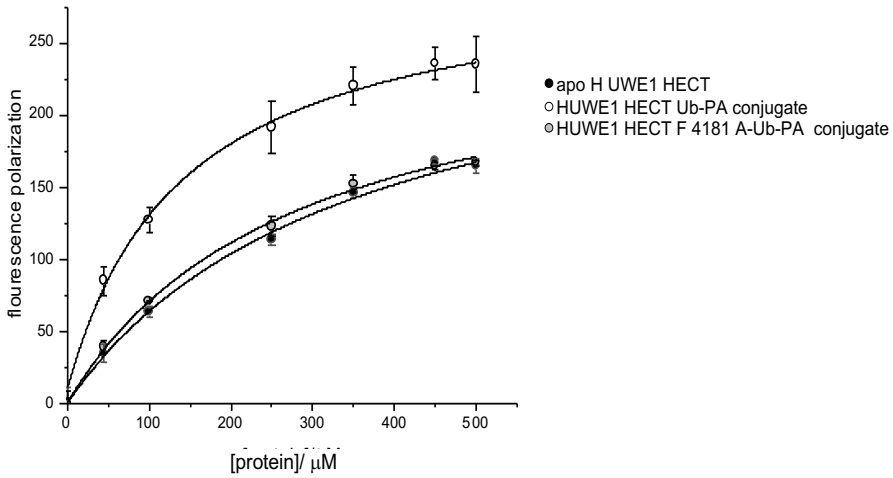


Figure 35: FP analysis of HUWE1 HECT domain-ubiquitin interactions

FP measurements of HUWE1 binding to Ub-fluorophor in the presence of different concentrations of the indicated proteins. The data points and errors reflect the mean and standard deviation from three independent experiments.

The measurements were performed as triplicates and plotted as a function of concentration for *apo* HECT domain of HUWE1, HUWE1 HECT domain-Ub-PA conjugate and an exosite variant (F4181A), respectively. The FP measurements shows that ubiquitin binding is weak in the context of the *apo* HECT domain, with a K_D of 346 +/- 59 μM. I next tested the binding of ubiquitin to the Ub-PA-linked ubiquitin-conjugate. Interestingly, the binding was strengthened to a K_D of 135 +/- 20 μM, in this context. To determine whether the interaction with ubiquitin depends on the same contacts as the exosite of NEDD4 ligases, I replaced a conserved phenylalanine (Phe 4181) in the HECT domain of HUWE1

by alanine. The homologous phenylalanine in NEDD4 and Rsp5 was shown to be critical for isopeptide bond formation activity and ubiquitin binding [79,174]. The homologous mutation in HUWE1 indeed weakens the affinity of the HECT-ubiquitin conjugate for ubiquitin to a K_D of 277 +/- 44 μM, indicating that the observed binding relies on the same key contact as that observed in NEDD4-type ligases and E6AP. Taken together, these data indicate that the HECT domain of HUWE1 binds to ubiquitin in *trans* and a key residue in the exosite (Phe 4181) in HUWE1 is important for this interaction. Moreover, interestingly, the interaction with ubiquitin is enhanced when the donor ubiquitin is loaded on the HECT domain. This observation may suggest that the exosite in HUWE1 is more accessible in the context of the donor-conjugate than in the unloaded HECT domain, that the donor and exosite-bound ubiquitin contact each other, or that the donor ubiquitin influences the properties of the exosite allosterically.

4.4.2 Characterization of allosteric ubiquitin binding sites in the HECT domain of HUWE1

To test the functional consequences of ubiquitin binding to the exosite region of the HECT domain of HUWE1, I performed *in-vitro* activity assays upon introducing exosite-mutations in HUWE1 (Y4078A), F4181A), analogous to previous studies on NEDD4-type enzymes and E6AP [71,78-80] (Figure 36). The purified exosite variants were first analyzed by CD spectroscopy to confirm their structural integrity. As shown in Figure 37, neither of the mutations perturbed the native protein fold. I next tested their ability to promote isopeptide bond formation by monitoring free ubiquitin chains and auto-ubiquitination, compared to WT. The data from the *in-vitro* assay show that the exosite mutations reduce the activity of HECT domain in both types of read-out.

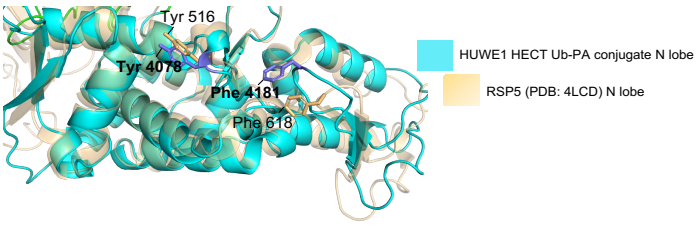


Figure 36: Structural superposition of exosite residues of two human HECT ligases

Structural superposition of the N-lobe of the HUWE1 HECT domain-Ub-PA conjugate (PDB ID 6XZ) and an Rsp5 (extracted from PDB: 4LCD) [87].The proteins are shown in

cartoon representation, the side chains of two critical exosite residues, whose effect was tested by mutagenesis are displayed as stick models (HUWE1 bold).

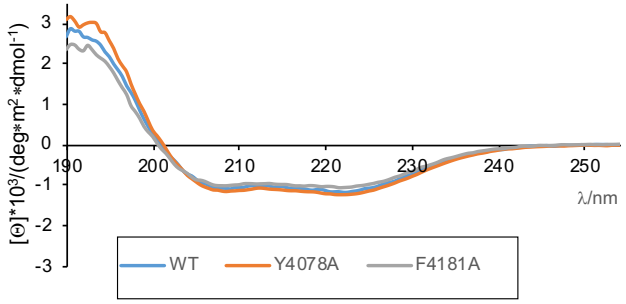


Figure 37: CD spectra of HUWE1 HECT domain exosite variants

Superposition of individual spectra of the indicated HUWE1 HECT domain variants. A total of 15 spectra were averaged.

Next, I interrogated the significance of these exosite residues for the ability of the HUWE1 HECT domain to receive donor ubiquitin from UBCH7. For this purpose, I produced the exosite variants in the context of the Δ4 construct [73], since the full-length HECT domain transfers the donor ubiquitin too rapidly to internal lysine residues to be monitored in manual-mixing experiments. The data shows that neither mutation impairs thioester transfer (Figure 39), which implies that the mutated exosite residues do not take part in the first step of the enzymatic reaction. In line with this finding, previous studies on NEDD4-type of ligases [78-80] and E6AP [71] showed that ubiquitin binding to the exosite is exclusively required for ubiquitin chain elongation. Therefore, I hypothesized that the HUWE1 exosite residues might also not be required for the reactivity of the HECT domain towards Ub-PA, which involves only one ubiquitin (the equivalent of the donor). To test this, I performed Ub-PA labelling experiments with the exosite variants of HECT domain of HUWE1 WT, Y4078A and F4181A (Figure 40**)**. These experiments show similar labeling efficiency of Y4078A and F4181A towards Ub-PA, as the WT. Hence, HUWE1 uses the same key residues in the exosite as NEDD4-type ligases and E6AP for isopeptide bond formation, but those are not required for thioester formation nor reactivity towards Ub-PA.

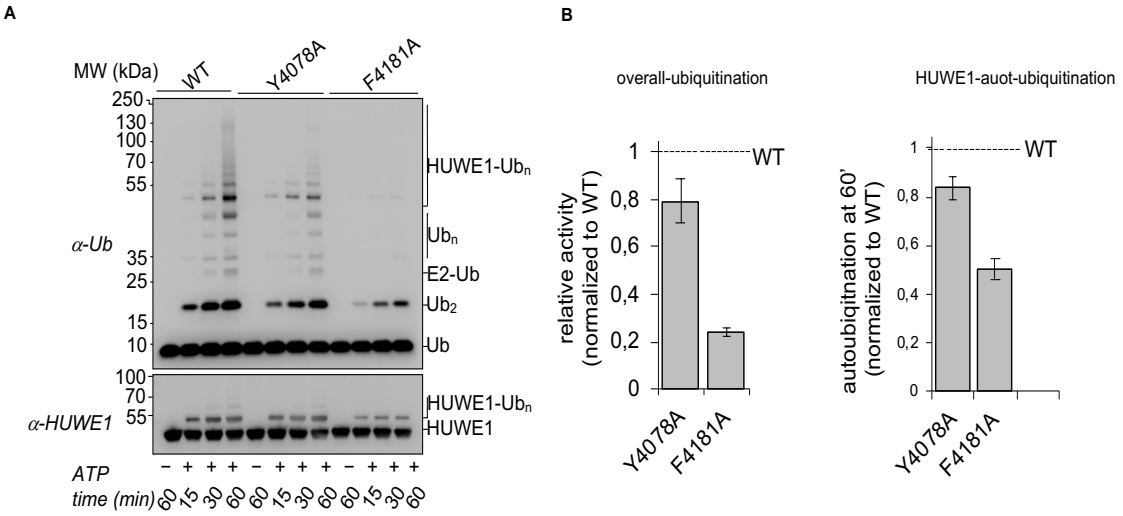


Figure 38: Role of the exosite in the HUWE1 HECT domain for activity

(A) *In-vitro* activity assay comparing the activities of the HECT domain of HUWE1 WT with two exosite variants Y4078A and F4181A by anti-ubiquitin and anti-HUWE1 Western blotting. **(B)** Quantification of the overall ubiquitination (from anti-ubiquitin blot) and autoubiquitination of HUWE1(from anti-HUWE1 blot). The assays were performed in triplicates and stopped at three time points (15, 30 and 60 minutes); the quantification is based on 3 independent experiments, the mean and standard deviation were plotted for the 60-minute time point.

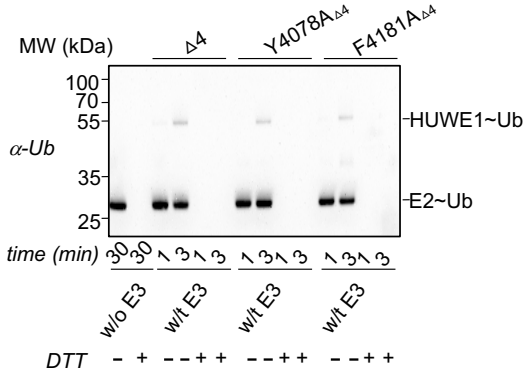


Figure 39: Thioester transfer of ubiquitin from UBCH7 to HUWE1 exosite variants

Thioester transfer of ubiquitin from UBCH7 to the Δ4 variants of HECT domain of HUWE1, monitored by single-turnover pulse-chase assays at two time points, as indicated, and monitored by anti-ubiquitin Western blotting.

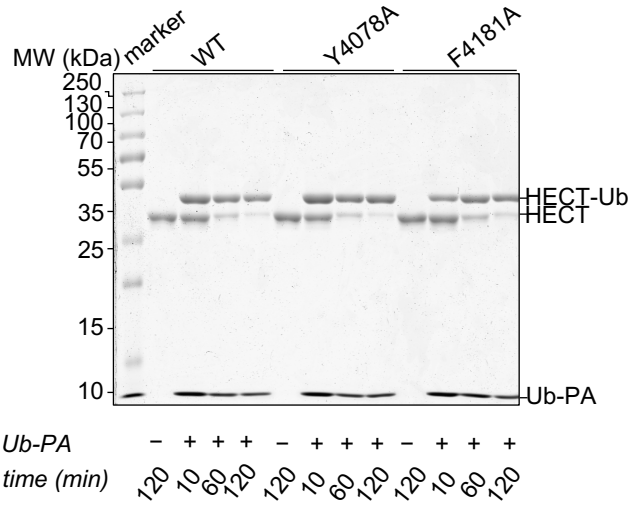


Figure 40: **Ub-PA labelling reactions of the HUWE1 HECT domain and exosite variants**

The labelling efficiency of HUWE1 WT, Y4078A and F4181A with Ub-PA is analyzed by SDS-PAGE. The reactions were carried out for 10, 60 and 120 minutes at 30°C; minus Ub-PA conditions represent a negative control.

4.5 Acceptor ubiquitin: The key to linkage specificity

To understand the position of acceptor ubiquitin, a long-standing, open question in the field of HECT ligases, I teamed up with Prof. Jun Yin and Han Zhou (graduate student in Prof. Yin's lab), Georgia State University, GA, USA, to generate a three-way cross-linking strategy to tether the acceptor at the E3-donor complex. This is necessary, since the binding of the acceptor ubiquitin is otherwise too weak for structural analyses. Our collaborators established a strategy to crosslink Lys48-linked diubiquitin to the catalytic cysteine of HUWE1 by an engineered disulphide bond, thus closely mimicking the intermediate during isopeptide bond formation. For this strategy to be applied, we generated a single-cysteine HECT domain variant, which retains only the active-site cysteine (section 2.1.2.2.7). The ternary protein complex produced by the Yin lab, was sent to me and subjected to size-exclusion chromatography (Figure 41). However, the SDS-PAGE analysis showed that the eluted protein was not sufficiently pure and

contained unconjugated HUWE1 HECT and additional purification attempts failed and-or were prohibited by lack of sufficient material. Due to aggregation problems during protein concentration, a maximum concentration of 5 mg/ml (unpure) complex was reached, which was used for 3 crystallization screens (Index and Wizard 1+2). No crystals have been obtained so far. Additional joint efforts will be needed for more extensive purification and crystallization trials. At the same time, alternative chemical biology strategies may be employed to reconstitute acceptor complexes with donor-HECT conjugates and decipher the structural underpinnings of linkage specificity in these ligases.

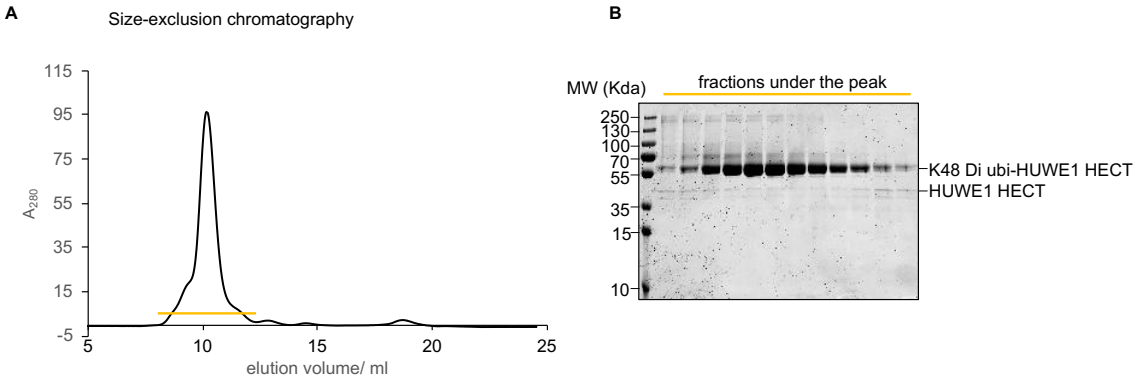


Figure 41: Purification of K48-linked di-ubiquitin in complex with the HUWE1 HECT domain

(**A**) Size-exclusion chromatography using a SD 75 increase 10/300 GL column. (**B**) SDS-PAGE loaded with fractions after the SEC.

5 CONCLUSIONS

5.1 Recognition of donor ubiquitin by the HECT domain of HUWE1

The low affinities of ubiquitin for the catalytic HECT domain makes it challenging to capture functionally relevant complexes of the HECT domain with the donor and acceptor ubiquitin – an essential prerequisite to understanding the structural underpinnings of linkage specificity in ubiquitin chain formation. In order to gain insight into ubiquitin recognition by HUWE1 I determined the crystal structure of a donor ubiquitin-linked conjugate of the HECT domain of HUWE1. To provide access to sufficient quantities of this conjugate, I applied a ubiquitin-based ABP. In line with previous biochemical studies, I found the probe (Ub-PA) to react with HUWE1 and NEDD4; however, I observed high turnover (of 95% in 120 min), which reflects a significant improvement compared to those [142]. I assume that this improvement is due to the difference in protein constructs that I have used in my study, I cleaved the affinity tags off the proteins. The buffer conditions may be another reason for the difference in the protein turnover that I observed. Our structure is the first of the HECT domain of HUWE1 bound to ubiquitin, as opposed to the ubiquitin-bound isolated C-lobe of HUWE1 [88]. In our structure the donor ubiquitin binds to the HECT C-lobe in the same way as seen for the isolated C-lobe and for C-lobes and HECT domains of NEDD4-type ligases, corroborating the idea that HECT ligases may utilize a conserved mode of donor ubiquitin recognition that is independent of linkage and substrate specificities [80,84,87,88]. However, it remains to be tested whether the numerous yet uncharacterized ligases outside of the NEDD4 family indeed utilize the same mode of donor recognition.

Notably, our structure has the C-terminal tail resolved in an active conformation. The only other structure that included the C-tail was a dimeric state of an extended HECT domain construct of HUWE1, in which the tail is engaged in inhibitory interactions at the dimer interface. In our structure, the C-tail adopts a distinct conformation, in which it is coordinated intramolecularly by residues of both the N-lobe and the C-lobe. This coordination was mutationally validated, demonstrating that the conformation that was trapped in the crystal is indeed important for isopeptide bond formation. It was previously shown that the determinants of ubiquitin chain specificity are located within the last ~ 60 of E6AP or HUWE1 into RSP5, changed its chain type specificity to the either Lys 48 or

HUWE1 into RSP5, changed its specificity to Lys 48 and Lys 63, respectively [174]. Similarly, the last six residues of E6AP (KGFGML), in which mostly the Gly^{848} and Phe^{849} are conserved among HECT-ligases (Figure 42), A variant of E6AP F849A, failed to promote substrate ubiquitination, but showed no defects in the transfer of ubiquitin from E2 to E6AP. Terminating E6AP at Ala^{846}, also showed defects in substrate ubiquitination; similar results were obtained when the C-terminal tail of E6AP was extended with the insertion of several random amino acids or with an epitope tag (reference missing). Similar results were obtained for Rsp5 variants [171]. The substitution of Asp^{900} to lysine in NEDD4 abrogated ubiquitin chain formation, showing the importance of a single acidic residue of the C-terminal tail [80]. The substitution of the last three residues of E6AP into the sequence of NEDD4 resulted in a reduction in the formation of Lys 48 chains (reference missing).

Another study showed that the last three amino acids of HECT- type ligases might be important for determining the chain type specificity together with other HECT domain determinants [80]. The crystal structure of a chemically trapped proxy of Rsp5, ubiquitin, and a target peptide $SNA3^{C}$ during isopeptide bond formation suggests that Phe^{806} contacts the N-lobe and deletion of this residue severely affects substrate ubiquitination [87]. Thus, the C-terminal tail is important to anchor both the lobes of the HECT domain for effective substrate ubiquitination and may also play yet undefined role in positioning the donor or acceptor ubiquitin.

5.2 Acceptor ubiquitin: Key to understand linkage specificity

The missing key to understand the linkage specificity of HECT-type ligases is the positioning of the acceptor ubiquitin, which nucleophilically attacks the activated C-terminus of the E3-bound donor with a specific primary amino group. To gain structural insight into this mechanism, we employed a chemical biology strategy with Han Zhou and Prof. Jun Yin (Georgia State University, GA, USA) to produce a ternary complex of the

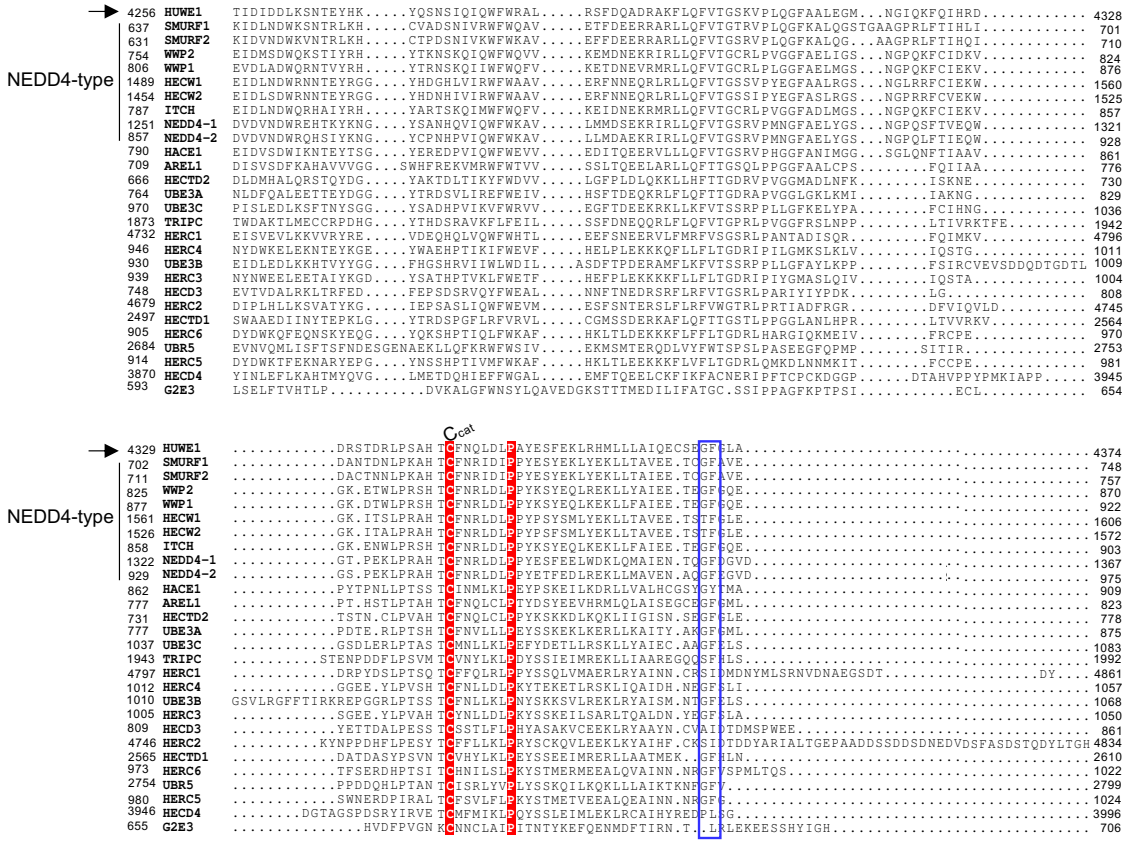


Figure 42: Multiple amino acid sequence alignment of C-lobe of different human HECT ligases

Amino acid sequence alignment for the C-lobes of different human HECT- type ligases, as output by Clustal Omega [69], the alignment is illustrated by ESPript 3.0 alignment program [170] and coloring is based on PAM 250 [172,173]. The Glycine and phenylalanine residues in all HECTs are boxed. The catalytic cysteine is labeled with C_{cat}.

E3, donor and acceptor ubiquitin. This strategy employs unnatural amino acid incorporation to form a three-way cross-link between a Lys 48-linked diubiquitin and the catalytic cysteine of HUWE1, thus closely mimicking the native intermediate of linkage formation. Following this strategy, a ternary-complex could be produced by Han Zhou, establishing a proof of principle. However, the material proved insufficient in purity and homogeneity for successful crystallization. While we will continue to pursue this

collaborative approach, we will also pursue alternative strategies to generate the desired ternary complex [132,175].

It will be important to understand the structural basis of linkage specificity for HUWE1, because it can modify its cellular substrates with three different ubiquitin chain types in a context-dependent manner. It has been shown that, HUWE1 can regulate the stability of MYC through Lys 48-linked ubiquitination and the knockdown of HUWE1 leads to increased levels of MYC and promotes colon cancer [176]. Studies on other cancer types shows, HUWE1 can activate the transcription of MYC through Lys 63-linked ubiquitination and HUWE1 silencing or inhibition prevents MYC signaling and reduces the tumor cell proliferation [91,99]. These observations suggest that HUWE1 may play a cell-type and context-dependent role in MYC associated tumors. HUWE1 was also reported to form Lys 6 chins on MFN2 [23]. Lys 6 chains were significantly less abundant in HUWE1−/− cell lines, showing a decrease of ~75%. HUWE1 is involved in pathological conditions, like DDR or mitophagy, where Lys 6-linked chains could serve important adaptor functions [23]. Therefore, a detailed structural view of how HUWE1 can assemble these different chain types is required. If our chemical biology crosslinking strategies work, we can apply them to additional linkage-specific HECT ligases, such as the Lys 48-specific E6AP [71] and the Lys 63-specific NEDD4 [80], and any cysteine-containing ubiquitination enzyme (HECTs, RBRs, and E2s).

5.3 The non-covalent ubiquitin binding exosite in HUWE1

My studies suggest that the N-lobe of HUWE1 harbors a ubiquitin-binding exosite similar to NEDD4-type [76-82] and E6AP [71]. In line with the function of the exosite in other ligases [76,78,79,177], my *in-vitro* activity and binding studies show that HUWE1 uses this exosite for isopeptide bond formation, but that it is dispensable for thioester bond formation bond formation (Figure 35 and Figure 39). Moreover, previous studies showed that the exosite-mediated ubiquitin interaction is needed for the processivity of the enzyme: A small-molecule inhibitor of NEDD4-1 that blocks the exosite was shown to switch the NEDD4-1 from a processive to a distributive state [89].

The modulation of enzymatic activities through non-covalent interactions with other proteins is a general mechanism found in the regulation of E2 and E3 enzymes. For example, E2 enzymes have a site, known as 'backside', of the UBC domain (Figure 43). Structural and biochemical studies on UBE2D3 and several other E2 enzymes has shown that even though this non-covalent interaction falls into a high-micromolar affinity range, it can stimulate the processivity of the enzyme to build polyubiquitin chains [57,178-181].

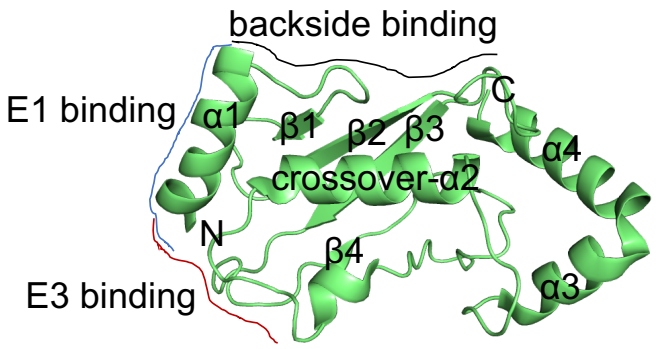


Figure 43: The solution NMR structure of the catalytic core domain of UBE2D3

The solution NMR structure of the UBC domain of UBE2D3 (PDB: 2FUH) [54] is shown in cartoon representation. The important structural features are labelled; the E1, E3 and the non-covalent ubiquitin binding side are labelled. The α/β-fold which consists of four α-helices and four stranded β-sheets are also labelled.This figure was adapted from [51].

A superposition of the β1/β2-loop conformations of NEDD8 and ubiquitin shows that, the non-covalent binding of NEDD8 or ubiquitin to UBE2D3 can only happen if the loops are in a loop-out conformation, and also only a loop-out conformation is compatible for the CUL1-linked NEDD8 to bind the catalytic module of the enzyme for its reactivity [59]. The backside binding of ubiquitin in the case of UBE2B, helps the enzyme to build Lys 11-linked polyubiquitin chains, in an E3 independent manner [53]. The intrinsic chain-building ability of RAD6, is enhanced by the backside ubiquitin binding [182] and the ubiquitin-like protein SUMO binds to the backside of UBE2I, which initiates chain building by SUMO [183]. The SUMO E3 ligase, RANBP2, which lacks the RING domain has a natively unfolded region known as internal repeat region 1 (IR1) that exhibits SUMO E3 ligase activity which makes critical contacts with the canonical E3-binding surface of UBE2I, the backside of UBE2I, and by interactions with SUMO binds to the backside of UBE2I [184-

186]. Thus, the non-covalent binding of a Ub/Ubl to the distal backside surface of E2 enzymes has a positive effect on E2 processivity [57].

The non-covalent ubiquitin binding in NEDD4-type of ligases has been shown to be important for ubiquitin chain elongation but not for the transfer of ubiquitin from E2 to E3 or for the conjugation of the first ubiquitin to a substrate [78,79]. That the ubiquitin-binding exosite is distal from the E3~ubiquitin thioester suggests it does not serve as a docking site for the acceptor ubiquitin, but might promote processive ubiquitin chain formation by tethering a growing chain [78]. The exosite interacts with ubiquitin utilizing a series of predominantly hydrophobic residues to contact the canonical Ile 44 hydrophobic patch [50] (Figure 44). It has been shown that the UEV domain present in the ESCRT-I subunit VPS23 compete with the exosite in RSP5 for the target protein-bound ubiquitin, thus inhibiting ubiquitin chain elongation and favoring mono-ubiquitination [187].

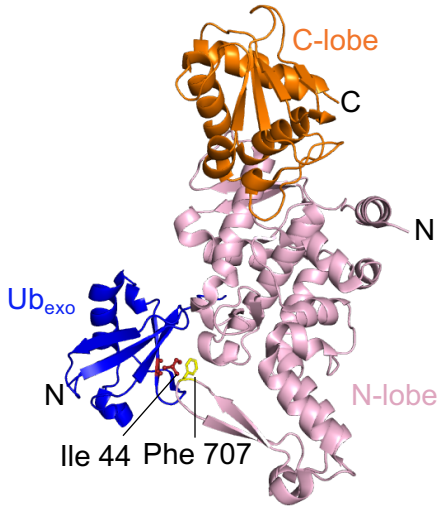


Figure 44: Top view of the crystal structure of the HECT domain of NEDD4 bound to exosite ubiquitin

The crystal structure of the HECT domain of NEDD4-exosite ubiquitin is shown in cartoon representation (PDB extracted from: 2XBB) [79]. A key residue on the exosite ubiquitin (Ile 44) which interacts with the N-lobe residue on NEDD4 (Phe 707) ia shown in ball-and-stick representation. The two lobes and the N and the C-terminus of the protein are also labelled.

In the case of SMURF2, the ubiquitin binding exosite lies close to the C2-domain binding region and to the catalytic cysteine of the enzyme. Thus, the binding of C2 domain to the HECT domain of SMURF2 inhibits the processivity of the SMURF2 and also the transfer of ubiquitin to the E3's active site [65,66]. A crystal structure of a WW-2-(2,3-linker)-HECT fusion construct, which is one of the regulatory elements in NEDD4-type ligases, reveals extensive interactions between the 2,3-linker region, the N-lobe and the ubiquitin-binding exosite [188]. The mutational analysis of the ubiquitin binding exosite in Rsp5 and the Ile 44 patch of ubiquitin has found to interfere with the oligomerization of the enzyme [75].

Detailed biochemical analysis and HECT-ubiquitin variant (UbV) co-crystallization experiments revealed that UbV binding at the N-lobe exosite can modulate E3 ligase activity through a variety of mechanisms (reference missing). Binding of the UbV to the NEDD4-2 and Rsp5 exosite promoted the transfer of ubiquitin from E2 to E3, whereas other UbV interactions promoted ubiquitin transfer from E3 to the substrate. Another two UbVs were shown to modulate NEDD4-2 activity by decreasing processivity and increasing distributive multi mono-ubiquitination. As the engineered UbVs are highly specific for one HECT E3 by way of their design and selection process, their specificity may be useful to target and modulate a specific HECT E3 ligase without affecting other HECT E3s in the cell [82]. From my study, the binding assays (Figure 35) shows that a donor ubiquitin loaded HECT domain binds an additional ubiquitin molecule at the exosite better than the apo HECT domain. This may suggest

1) direct contact between donor ubiquitin and the ubiquitin in the exosite.

- This is indeed unlikely to occur due to steric reasons.

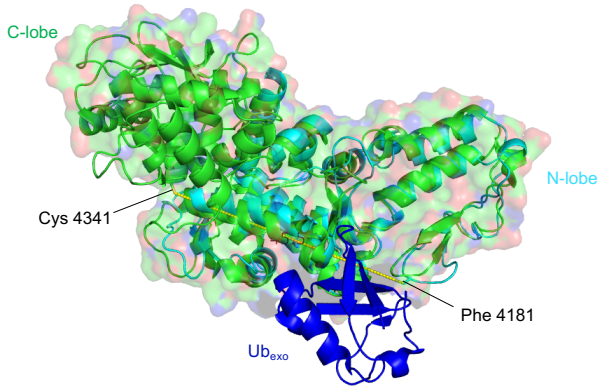


Figure 45: Modeling of exosite ubiquitin into HUWE HECT domain-Ub-PA conjugate

The crystal structure of HUWE1 HECT domain-Ub-PA conjugate (PDB: 6XZ1) is aligned with the crystal structure of NEDD4-Ub_{exo} (PDB: 2XBB) [79] in pymol. The surface representation of the structures is shown with a transparency set to 0.5 in pymol. The distance between the catalytic cysteine 4341 and Phe 4181 are measured. The N-lobe, C-lobe and the Ub_{exo} are labelled.

2) that the exosite is buried in the *apo* HECT domain of HUWE1

- This may be possible due to the flexibility of the inter-lobe linker; yet there are no structures available in favor of this model.

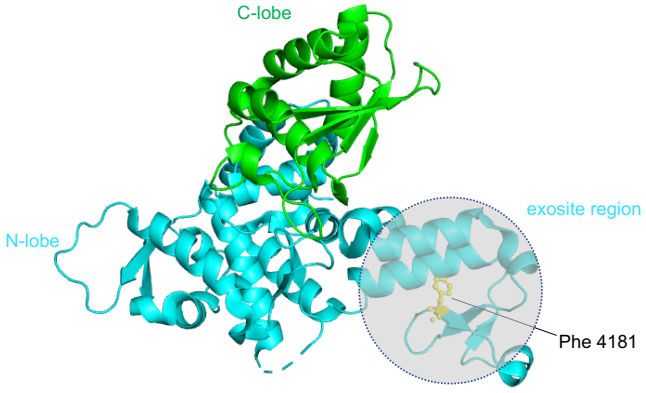


Figure 46: Crystal structure of the *apo* HECT domain of HUWE1

The crystal structure of the apo HECT domain of HUWE1 (PDB: 3H1D) [73] is represented in cartoon representation with the exosite region marked on the N-lobe.

3) that an allosteric link between the L-shape and exosite ubiquitin binding

- This effect is currently studied with the help of MD simulations together with Stephan Boehler and Prof. Christoph Sotriffer (University of Wuerzburg).

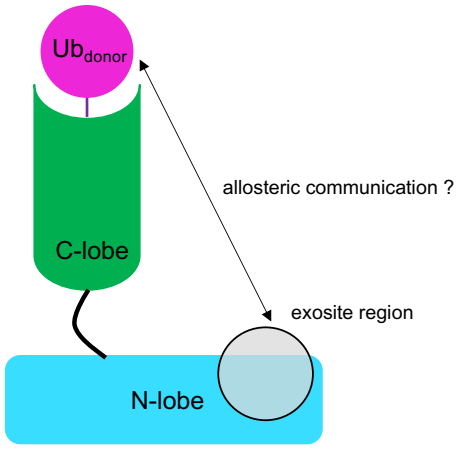


Figure 47: Cartoon representation of the allosteric communication

The cartoon here shows the possible allosteric communication between the donor ubiquitin loaded C-lobe of the HECT domain of HUWE1 and the exosite on the N-lobe.

5.4 Differential reactivity of HECTs towards activity-based probes

The expansion of the ABPs has contributed to a rapid increase in our understanding of deubiquitinases. However, recent studies suggest that C-terminal electrophilic probes can also label Cys residues in HECT ligase [126,141,142]. In line with these results, I showed Ub-PA can label the catalytic cysteine of HUWE1 and NEDD4-type with close to quantitative turn-over, while E6AP did not show any labelling. This is in line with the evolutionary relationship of these ligases (HUWE1 and NEDD4 being similar, while E6AP is more divergent) (Figure 5), but the exact factors that impact reactivity are unclear. My observation that the C-lobes of HUWE1, NEDD4 and E6AP did not show any labelling suggests that the N-lobe contributes to the reactivity. This could occur through an impact of the N-lobe on the chemical environment of the active site, influencing the pK_a of the thiol group of the active site cysteine or through the different affinities of the HECT domains for ubiquitin, thus leading to differences in the recruitment of Ub-PA. Furthermore, it could be that the C-terminal tail has a potential role in reactivity of the HECT domain towards Ub-PA.

5.5 Activity-based probes as precursors of HECT ligase-directed drugs?

Chemical probes that react with the active site in an enzyme are termed ‘activity-based probes’ (ABPs). ABPs rely on some aspect of the catalytic mechanism of the enzyme to ensure that reaction occurs only for catalytically-active species [189]. Propargylamine derivatives are well-known mechanism-based covalent monoamine oxidase (MAO)-B inhibitors in the treatment of Parkinson’s disease and terminal alkynes have found application as activity-based probes for CYP enzymes [190]. Biologically inert terminal alkynes have recently emerged as a tool to covalently label proteins that has a catalytic cysteine. Ubiquitin-PA was shown to label DUBs [126] as well as HECT-type ligases [126,142], which makes it an interesting precursor for active site-directed inhibitors of these enzyme classes, e.g. by replacing the ubiquitin moiety in Ub-PA with a small-molecule binder. New binders could be identified by high-throughput screening. Another possibility is selecting existing weak binders of HECT-type ligases and attaching them to PA, with which one can possibly switch the molecule to a stronger, covalent binder.

Manipulating and expanding the knowledge on activity-based probes towards drug discovery is in principle feasible though it comes with some obstacles, for example selection of a reactive warhead. Computational reactivity prediction for certain electrophiles is possible, but precise and comparable rating of these nucleophiles still remains as a major challenge [189]. Another challenge is when the target residue is poorly reactive, and difficult to access, or incompatible with the spatial and geometric requirements of the nucleophilic headgroup. Tissue-specificity and redundancy are other limiting factors. However, HECT-type ligases are very attractive targets due to their key roles in various diseases and their diverse domain structures, once structurally understood, may facilitate specific manipulations.

6 REFERENCES

1. Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. Nature Structural & Molecular Biology 17, 666-672.
2. Kho, Y., Kim, S.C., Jiang, C., Barma, D., Kwon, S.W., Cheng, J., Jaunbergs, J., Weinbaum, C., Tamanoi, F., Falck, J.*, et al.* (2004). A tagging-via-substrate technology for detection and proteomics of farnesylated proteins. Proc Natl Acad Sci U S A 101, 12479-12484.
3. Ikeda, F., and Dikic, I. (2008). Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. EMBO Rep 9, 536-542.
4. Swatek, K.N., and Komander, D. (2016). Ubiquitin modifications. Cell Research 26, 399-422.
5. Koyano, F., Okatsu, K., Kosako, H., Tamura, Y., Go, E., Kimura, M., Kimura, Y., Tsuchiya, H., Yoshihara, H., Hirokawa, T.*, et al.* (2014). Ubiquitin is phosphorylated by PINK1 to activate parkin. Nature 510, 162-166.
6. Komander, D., and Rape, M. (2012). The Ubiquitin Code. Annual Review of Biochemistry 81, 203-229.
7. Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D.H., Niall, H.D., and Boyse, E.A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proc Natl Acad Sci U S A 72, 11-15.
8. Ciechanover, A., Heller, H., Elias, S., Haas, A.L., and Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. Proc Natl Acad Sci U S A 77, 1365-1368.
9. Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., and Rose, I.A. (1980). Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. Proc Natl Acad Sci U S A 77, 1783-1786.
10. Peng, J., Schwartz, D., Elias, J.E., Thoreen, C.C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S.P. (2003). A proteomics approach to understanding protein ubiquitination. Nature Biotechnology 21, 921-926.
11. Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science 243, 1576-1583.
12. Chen, Z., and Pickart, C.M. (1990). A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. J Biol Chem 265, 21835-21842.
13. Talreja, J., and Samavati, L. (2018). K63-Linked Polyubiquitination on TRAF6 Regulates LPS-Mediated MAPK Activation, Cytokine Production, and Bacterial Clearance in Toll-Like Receptor 7/8 Primed Murine Macrophages. Front Immunol 9, 279-279.
14. Chen, J., and Chen, Z.J. (2013). Regulation of NF-κB by ubiquitination. Curr Opin Immunol 25, 4-12.
15. Kirkin, V., McEwan, D.G., Novak, I., and Dikic, I. (2009). A role for ubiquitin in selective autophagy. Mol Cell 34, 259-269.
16. Liu, P., Gan, W., Su, S., Hauenstein, A.V., Fu, T.-m., Brasher, B., Schwerdtfeger, C., Liang, A.C., Xu, M., and Wei, W. (2018). K63-linked polyubiquitin chains bind to DNA to facilitate DNA damage repair. Science Signaling 11, eaar8133.
17. Liu, T., Ghosal, G., Yuan, J., Chen, J., and Huang, J. (2010). FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. Science 329, 693-696.

18. Gatti, M., Pinato, S., Maiolica, A., Rocchio, F., Prato, M.G., Aebersold, R., and Penengo, L. (2015). RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. Cell Rep 10, 226-238.
19. Matsumoto, M.L., Wickliffe, K.E., Dong, K.C., Yu, C., Bosanac, I., Bustos, D., Phu, L., Kirkpatrick, D.S., Hymowitz, S.G., Rape, M.*, et al.* (2010). K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. Mol Cell 39, 477-484.
20. Fei, C., Li, Z., Li, C., Chen, Y., Chen, Z., He, X., Mao, L., Wang, X., Zeng, R., and Li, L. (2013). Smurf1-mediated Lys29-linked nonproteolytic polyubiquitination of axin negatively regulates Wnt/β-catenin signaling. Mol Cell Biol 33, 4095-4105.
21. Yuan, W.C., Lee, Y.R., Lin, S.Y., Chang, L.Y., Tan, Y.P., Hung, C.C., Kuo, J.C., Liu, C.H., Lin, M.Y., Xu, M.*, et al.* (2014). K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking. Mol Cell 54, 586-600.
22. Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., Sano, S., Tokunaga, F., Tanaka, K., and Iwai, K. (2006). A ubiquitin ligase complex assembles linear polyubiquitin chains. EMBO J 25, 4877-4887.
23. Michel, M.A., Swatek, K.N., Hospenthal, M.K., and Komander, D. (2017). Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. Molecular cell 68, 233-246.e235.
24. Scheffner, M., Huibregtse, J.M., Vierstra, R.D., and Howley, P.M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell 75, 495-505.
25. Kapetanaki, M.G., Guerrero-Santoro, J., Bisi, D.C., Hsieh, C.L., Rapić-Otrin, V., and Levine, A.S. (2006). The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. Proc Natl Acad Sci U S A 103, 2588-2593.
26. Popovic, D., Vucic, D., and Dikic, I. (2014). Ubiquitination in disease pathogenesis and treatment. Nature Medicine 20, 1242-1253.
27. Karbowski, M., and Youle, R.J. (2011). Regulating mitochondrial outer membrane proteins by ubiquitination and proteasomal degradation. Curr Opin Cell Biol 23, 476-482.
28. Bonnet, M., and Courtois, G. (2011). [CYLD deubiquitinase as a recurrent target in oncogenic processes]. Med Sci (Paris) 27, 626-631.
29. Huang, X., and Dixit, V.M. (2016). Drugging the undruggables: exploring the ubiquitin system for drug development. Cell Research 26, 484-498.
30. Adams, J. (2001). Proteasome inhibition in cancer: Development of PS-341. Seminars in Oncology 28, 613-619.
31. Kortuem, K.M., and Stewart, A.K. (2013). Carfilzomib. Blood 121, 893-897.
32. Buac, D., Shen, M., Schmitt, S., Kona, F.R., Deshmukh, R., Zhang, Z., Neslund-Dudas, C., Mitra, B., and Dou, Q.P. (2013). From bortezomib to other inhibitors of the proteasome and beyond. Curr Pharm Des 19, 4025-4038.
33. Goldberg, A.L. (2012). Development of proteasome inhibitors as research tools and cancer drugs. J Cell Biol 199, 583-588.
34. Sakamoto, K.M., Kim, K.B., Kumagai, A., Mercurio, F., Crews, C.M., and Deshaies, R.J. (2001). Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. Proceedings of the National Academy of Sciences 98, 8554.
35. Pickart, C.M., and Eddins, M.J. (2004). Ubiquitin: structures, functions, mechanisms. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1695, 55-72.
36. Wilkinson, K.D. (1997). Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. The FASEB Journal 11, 1245-1256.

37. Schulman, B.A., and Harper, J.W. (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nature reviews Molecular cell biology 10, 319-331.
38. Gundogdu, M., and Walden, H. (2019). Structural basis of generic versus specific E2–RING E3 interactions in protein ubiquitination. Protein Science 28, 1758-1770.
39. Eletr, Z.M., Huang, D.T., Duda, D.M., Schulman, B.A., and Kuhlman, B. (2005). E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. Nat Struct Mol Biol 12, 933-934.
40. Huang, D.T., Paydar, A., Zhuang, M., Waddell, M.B., Holton, J.M., and Schulman, B.A. (2005). Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1. Molecular cell 17, 341-350.
41. Wickliffe, K.E., Lorenz, S., Wemmer, D.E., Kuriyan, J., and Rape, M. (2011). The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. Cell 144, 769-781.
42. Metzger, M.B., Hristova, V.A., and Weissman, A.M. (2012). HECT and RING finger families of E3 ubiquitin ligases at a glance. Journal of Cell Science 125, 531.
43. Zheng, N., and Shabek, N. (2017). Ubiquitin Ligases: Structure, Function, and Regulation. Annual Review of Biochemistry 86, 129-157.
44. Pruneda, J.N., Littlefield, P.J., Soss, S.E., Nordquist, K.A., Chazin, W.J., Brzovic, P.S., and Klevit, R.E. (2012). Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. Mol Cell 47, 933-942.
45. Dou, H., Buetow, L., Sibbet, G.J., Cameron, K., and Huang, D.T. (2012). BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. Nat Struct Mol Biol 19, 876-883.
46. Saha, A., Lewis, S., Kleiger, G., Kuhlman, B., and Deshaies, R.J. (2011). Essential role for ubiquitin-ubiquitin-conjugating enzyme interaction in ubiquitin discharge from Cdc34 to substrate. Mol Cell 42, 75-83.
47. Pruneda, J.N., Stoll, K.E., Bolton, L.J., Brzovic, P.S., and Klevit, R.E. (2011). Ubiquitin in motion: structural studies of the ubiquitin-conjugating enzyme~ubiquitin conjugate. Biochemistry 50, 1624-1633.
48. Morreale, F.E., and Walden, H. (2016). Types of Ubiquitin Ligases. Cell 165, 248-248.e241.
49. Plechanovová, A., Jaffray, E.G., McMahon, S.A., Johnson, K.A., Navrátilová, I., Naismith, J.H., and Hay, R.T. (2011). Mechanism of ubiquitylation by dimeric RING ligase RNF4. Nat Struct Mol Biol 18, 1052-1059.
50. Sonja, L. (2018). Structural mechanisms of HECT-type ubiquitin ligases. Biological Chemistry 399, 127-145.
51. Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Research 26, 423-440.
52. Hamilton, K.S., Ellison, M.J., Barber, K.R., Williams, R.S., Huzil, J.T., McKenna, S., Ptak, C., Glover, M., and Shaw, G.S. (2001). Structure of a conjugating enzyme-ubiquitin thiolester intermediate reveals a novel role for the ubiquitin tail. Structure 9, 897-904.
53. Hibbert, R.G., Huang, A., Boelens, R., and Sixma, T.K. (2011). E3 ligase Rad18 promotes monoubiquitination rather than ubiquitin chain formation by E2 enzyme Rad6. Proceedings of the National Academy of Sciences 108, 5590.
54. Brzovic, P.S., Lissounov, A., Christensen, D.E., Hoyt, D.W., and Klevit, R.E. (2006). A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. Mol Cell 21, 873-880.
55. Miura, T., Klaus, W., Gsell, B., Miyamoto, C., and Senn, H. (1999). Characterization of the binding interface between ubiquitin and class I human ubiquitin-conjugating enzyme

2b by multidimensional heteronuclear NMR spectroscopy in solution. J Mol Biol 290, 213-228.
56. Bocik, W.E., Sircar, A., Gray, J.J., and Tolman, J.R. (2011). Mechanism of polyubiquitin chain recognition by the human ubiquitin conjugating enzyme Ube2g2. J Biol Chem 286, 3981-3991.
57. Buetow, L., Gabrielsen, M., Anthony, Nahoum G., Dou, H., Patel, A., Aitkenhead, H., Sibbet, Gary J., Smith, Brian O., and Huang, Danny T. (2015). Activation of a Primed RING E3-E2–Ubiquitin Complex by Non-Covalent Ubiquitin. Molecular Cell 58, 297-310.
58. Hua, Z., and Vierstra, R.D. (2011). The cullin-RING ubiquitin-protein ligases. Annu Rev Plant Biol 62, 299-334.
59. Baek, K., Krist, D.T., Prabu, J.R., Hill, S., Klügel, M., Neumaier, L.-M., von Gronau, S., Kleiger, G., and Schulman, B.A. (2020). NEDD8 nucleates a multivalent cullin–RING–UBE2D ubiquitin ligation assembly. Nature 578, 461-466.
60. Petroski, M.D., and Deshaies, R.J. (2005). Function and regulation of cullin–RING ubiquitin ligases. Nature Reviews Molecular Cell Biology 6, 9-20.
61. Smit, J.J., and Sixma, T.K. (2014). RBR E3-ligases at work. EMBO Rep 15, 142-154.
62. Huibregtse, J.M., Scheffner, M., Beaudenon, S., and Howley, P.M. (1995). A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc Natl Acad Sci U S A 92, 2563-2567.
63. Rotin, D., and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. Nature Reviews Molecular Cell Biology 10, 398-409.
64. Scheffner, M., and Kumar, S. (2014). Mammalian HECT ubiquitin-protein ligases: Biological and pathophysiological aspects. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1843, 61-74.
65. Wiesner, S., Ogunjimi, A.A., Wang, H.-R., Rotin, D., Sicheri, F., Wrana, J.L., and Forman-Kay, J.D. (2007). Autoinhibition of the HECT-Type Ubiquitin Ligase Smurf2 through Its C2 Domain. Cell 130, 651-662.
66. Mari, S., Ruetalo, N., Maspero, E., Stoffregen, Mira C., Pasqualato, S., Polo, S., and Wiesner, S. (2014). Structural and Functional Framework for the Autoinhibition of Nedd4-Family Ubiquitin Ligases. Structure 22, 1639-1649.
67. Kawabe, H., Neeb, A., Dimova, K., Young, S.M., Jr., Takeda, M., Katsurabayashi, S., Mitkovski, M., Malakhova, O.A., Zhang, D.-E., Umikawa, M.*, et al.* (2010). Regulation of Rap2A by the ubiquitin ligase Nedd4-1 controls neurite development. Neuron 65, 358-372.
68. Rosa, J.L., Casaroli-Marano, R.P., Buckler, A.J., Vilaró, S., and Barbacid, M. (1996). p619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and Rab proteins. EMBO J 15, 4262-4273.
69. Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D.*, et al.* (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47, W636-W641.
70. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44, W242-245.
71. Ries, L.K., Sander, B., Deol, K.K., Letzelter, M.-A., Strieter, E.R., and Lorenz, S. (2019). Analysis of ubiquitin recognition by the HECT ligase E6AP provides insight into its linkage specificity. Journal of Biological Chemistry.
72. Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P.M., Huibregtse, J.M., and Pavletich, N.P. (1999). Structure of an E6AP-UbcH7 Complex: Insights into Ubiquitination by the E2-E3 Enzyme Cascade. Science 286, 1321.

73. Pandya, R.K., Partridge, J.R., Love, K.R., Schwartz, T.U., and Ploegh, H.L. (2010). A structural element within the HUWE1 HECT domain modulates self-ubiquitination and substrate ubiquitination activities. J Biol Chem 285, 5664-5673.
74. Verdecia, M.A., Joazeiro, C.A., Wells, N.J., Ferrer, J.L., Bowman, M.E., Hunter, T., and Noel, J.P. (2003). Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. Mol Cell 11, 249-259.
75. Attali, I., Tobelaim, W.S., Persaud, A., Motamedchaboki, K., Simpson-Lavy, K.J., Mashahreh, B., Levin-Kravets, O., Keren-Kaplan, T., Pilzer, I., Kupiec, M.*, et al.* (2017). Ubiquitylation-dependent oligomerization regulates activity of Nedd4 ligases. EMBO J 36, 425-440.
76. French, M.E., Klosowiak, J.L., Aslanian, A., Reed, S.I., Yates, J.R., 3rd, and Hunter, T. (2017). Mechanism of ubiquitin chain synthesis employed by a HECT domain ubiquitin ligase. J Biol Chem 292, 10398-10413.
77. French, M.E., Kretzmann, B.R., and Hicke, L. (2009). Regulation of the RSP5 ubiquitin ligase by an intrinsic ubiquitin-binding site. J Biol Chem 284, 12071-12079.
78. Kim, H.C., Steffen, A.M., Oldham, M.L., Chen, J., and Huibregtse, J.M. (2011). Structure and function of a HECT domain ubiquitin-binding site. EMBO Rep 12, 334-341.
79. Maspero, E., Mari, S., Valentini, E., Musacchio, A., Fish, A., Pasqualato, S., and Polo, S. (2011). Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation. EMBO Rep 12, 342-349.
80. Maspero, E., Valentini, E., Mari, S., Cecatiello, V., Soffientini, P., Pasqualato, S., and Polo, S. (2013). Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. Nature Structural & Molecular Biology 20, 696-701.
81. Ogunjimi, A.A., Wiesner, S., Briant, D.J., Varelas, X., Sicheri, F., Forman-Kay, J., and Wrana, J.L. (2010). The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. J Biol Chem 285, 6308-6315.
82. Zhang, W., Wu, K.-P., Sartori, M.A., Kamadurai, H.B., Ordureau, A., Jiang, C., Mercredi, P.Y., Murchie, R., Hu, J., Persaud, A.*, et al.* (2016). System-Wide Modulation of HECT E3 Ligases with Selective Ubiquitin Variant Probes. Molecular cell 62, 121-136.
83. Eletr, Z.M., and Kuhlman, B. (2007). Sequence determinants of E2-E6AP binding affinity and specificity. Journal of molecular biology 369, 419-428.
84. Kamadurai, H.B., Souphron, J., Scott, D.C., Duda, D.M., Miller, D.J., Stringer, D., Piper, R.C., and Schulman, B.A. (2009). Insights into Ubiquitin Transfer Cascades from a Structure of a UbcH5B∼Ubiquitin-HECTNEDD4L Complex. Molecular Cell 36, 1095-1102.
85. Ronchi, V.P., Klein, J.M., and Haas, A.L. (2013). E6AP/UBE3A ubiquitin ligase harbors two E2~ubiquitin binding sites. J Biol Chem 288, 10349-10360.
86. Schwarz, S.E., Rosa, J.L., and Scheffner, M. (1998). Characterization of human hect domain family members and their interaction with UbcH5 and UbcH7. J Biol Chem 273, 12148-12154.
87. Kamadurai, H.B., Qiu, Y., Deng, A., Harrison, J.S., MacDonald, C., Actis, M., Rodrigues, P., Miller, D.J., Souphron, J., Lewis, S.M.*, et al.* (2013). Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. eLife 2, e00828.
88. Jäckl, M., Stollmaier, C., Strohäker, T., Hyz, K., Maspero, E., Polo, S., and Wiesner, S. (2018). β-Sheet Augmentation Is a Conserved Mechanism of Priming HECT E3 Ligases for Ubiquitin Ligation. Journal of Molecular Biology 430, 3218-3233.
89. Kathman, S.G., Span, I., Smith, A.T., Xu, Z., Zhan, J., Rosenzweig, A.C., and Statsyuk, A.V. (2015). A Small Molecule That Switches a Ubiquitin Ligase From a Processive to a Distributive Enzymatic Mechanism. J Am Chem Soc 137, 12442-12445.
90. Kao, S.-H., Wu, H.-T., and Wu, K.-J. (2018). Ubiquitination by HUWE1 in tumorigenesis and beyond. Journal of Biomedical Science 25, 67.

91. Adhikary, S., Marinoni, F., Hock, A., Hulleman, E., Popov, N., Beier, R., Bernard, S., Quarto, M., Capra, M., Goettig, S*., et al.* (2005). The Ubiquitin Ligase HectH9 Regulates Transcriptional Activation by Myc and Is Essential for Tumor Cell Proliferation. Cell 123, 409-421.
92. Vervoorts, J., Lüscher-Firzlaff, J.M., Rottmann, S., Lilischkis, R., Walsemann, G., Dohmann, K., Austen, M., and Lüscher, B. (2003). Stimulation of c-MYC transcriptional activity and acetylation by recruitment of the cofactor CBP. EMBO Rep 4, 484-490.
93. Brenner, C., Deplus, R., Didelot, C., Loriot, A., Viré, E., De Smet, C., Gutierrez, A., Danovi, D., Bernard, D., Boon, T*., et al.* (2005). Myc represses transcription through recruitment of DNA methyltransferase corepressor. EMBO J 24, 336-346.
94. Walker, W., Zhou, Z.-Q., Ota, S., Wynshaw-Boris, A., and Hurlin, P.J. (2005). Mnt-Max to Myc-Max complex switching regulates cell cycle entry. The Journal of cell biology 169, 405-413.
95. Stanton, B.R., Perkins, A.S., Tessarollo, L., Sassoon, D.A., and Parada, L.F. (1992). Loss of N-myc function results in embryonic lethality and failure of the epithelial component of the embryo to develop. Genes Dev 6, 2235-2247.
96. Zhao, X., Heng, J.I.-T., Guardavaccaro, D., Jiang, R., Pagano, M., Guillemot, F., Iavarone, A., and Lasorella, A. (2008). The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. Nat Cell Biol 10, 643-653.
97. Yang, Y., Do, H., Tian, X., Zhang, C., Liu, X., Dada, L.A., Sznajder, J.I., and Liu, J. (2010). E3 ubiquitin ligase Mule ubiquitinates Miz1 and is required for TNFalpha-induced JNK activation. Proc Natl Acad Sci U S A 107, 13444-13449.
98. Inoue, S., Hao, Z., Elia, A.J., Cescon, D., Zhou, L., Silvester, J., Snow, B., Harris, I.S., Sasaki, M., Li, W.Y*., et al.* (2013). Mule/Huwe1/Arf-BP1 suppresses Ras-driven tumorigenesis by preventing c-Myc/Miz1-mediated down-regulation of p21 and p15. Genes Dev 27, 1101-1114.
99. Peter, S., Bultinck, J., Myant, K., Jaenicke, L.A., Walz, S., Müller, J., Gmachl, M., Treu, M., Boehmelt, G., Ade, C.P*., et al.* (2014). Tumor cell-specific inhibition of MYC function using small molecule inhibitors of the HUWE1 ubiquitin ligase. EMBO Mol Med 6, 1525-1541.
100. Bykov, V.J.N., Eriksson, S.E., Bianchi, J., and Wiman, K.G. (2018). Targeting mutant p53 for efficient cancer therapy. Nat Rev Cancer 18, 89-102.
101. Chen, D., Kon, N., Li, M., Zhang, W., Qin, J., and Gu, W. (2005). ARF-BP1/Mule is a critical mediator of the ARF tumor suppressor. Cell 121, 1071-1083.
102. Qi, C.-F., Kim, Y.-S., Xiang, S., Abdullaev, Z., Torrey, T.A., Janz, S., Kovalchuk, A.L., Sun, J., Chen, D., Cho, W.C*., et al.* (2012). Characterization of ARF-BP1/HUWE1 interactions with CTCF, MYC, ARF and p53 in MYC-driven B cell neoplasms. Int J Mol Sci 13, 6204-6219.
103. Willis, S.N., Chen, L., Dewson, G., Wei, A., Naik, E., Fletcher, J.I., Adams, J.M., and Huang, D.C.S. (2005). Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. Genes & development 19, 1294-1305.
104. Zhong, Q., Gao, W., Du, F., and Wang, X. (2005). Mule/ARF-BP1, a BH3-only E3 ubiquitin ligase, catalyzes the polyubiquitination of Mcl-1 and regulates apoptosis. Cell 121, 1085-1095.
105. Gruber, S., Straub, B.K., Ackermann, P.J., Wunderlich, C.M., Mauer, J., Seeger, J.M., Büning, H., Heukamp, L., Kashkar, H., Schirmacher, P*., et al.* (2013). Obesity promotes liver carcinogenesis via Mcl-1 stabilization independent of IL-6Rα signaling. Cell Rep 4, 669-680.

106. Shao, G., Patterson-Fortin, J., Messick, T.E., Feng, D., Shanbhag, N., Wang, Y., and Greenberg, R.A. (2009). MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. Genes Dev 23, 740-754.
107. Wang, X., Lu, G., Li, L., Yi, J., Yan, K., Wang, Y., Zhu, B., Kuang, J., Lin, M., Zhang, S.*, et al.* (2014). HUWE1 interacts with BRCA1 and promotes its degradation in the ubiquitin–proteasome pathway. Biochem Biophys Res Commun 444, 290-295.
108. Xu, Y., Anderson, D.E., and Ye, Y. (2016). The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. Cell Discov 2, 16040.
109. Wu, H.T., Kuo, Y.C., Hung, J.J., Huang, C.H., Chen, W.Y., Chou, T.Y., Chen, Y., Chen, Y.J., Chen, Y.J., Cheng, W.C.*, et al.* (2016). K63-polyubiquitinated HAUSP deubiquitinates HIF-1α and dictates H3K56 acetylation promoting hypoxia-induced tumour progression. Nat Commun 7, 13644.
110. Sander, B., Xu, W., Eilers, M., Popov, N., and Lorenz, S. (2017). A conformational switch regulates the ubiquitin ligase HUWE1. eLife 6, e21036.
111. Solution structure of RUH-074, a human UBA domain FAU - Kitasaka, S., Ruhul Momen, A.Z.M., Hirota, H., Muto, Y., Yokoyama, S. CRDT - 2007/03/23 12:00 AID - 10.2210/pdb2ekk/pdb [doi].
112. WWE domain of human HUWE1 FAU - Halabelian, L., Loppnau, P., Tempel, W., Wong, F., Bountra, C., Arrowsmith, C.H., Edwards, A.M., Tong, Y. CRDT - 2018/09/20 12:00 AID - 10.2210/pdb6miw/pdb [doi].
113. Structure of Mcl-1 complexed with Mule at 2.05 Angstroms resolution FAU - Song, T., Wang, Z., Ji, F., Chai, G., Liu, Y., Li, X., Li, Z., Fan, Y., Zhang, Z. CRDT - 2015/06/23 12:00 AID - 10.2210/pdb5c6h/pdb [doi].
114. Solution structure of Ubiquitin Binding Motif of human Arf-bp1 FAU - Khatun, R., Sheng, Y. CRDT - 2014/09/12 12:00 AID - 10.2210/pdb2mul/pdb [doi].
115. Mund, T., Lewis, M.J., Maslen, S., and Pelham, H.R. (2014). Peptide and small molecule inhibitors of HECT-type ubiquitin ligases. Proceedings of the National Academy of Sciences 111, 16736.
116. Rossi, M., Rotblat, B., Ansell, K., Amelio, I., Caraglia, M., Misso, G., Bernassola, F., Cavasotto, C.N., Knight, R.A., Ciechanover, A.*, et al.* (2014). High throughput screening for inhibitors of the HECT ubiquitin E3 ligase ITCH identifies antidepressant drugs as regulators of autophagy. Cell Death Dis 5, e1203.
117. Chen, D., Gehringer, M., and Lorenz, S. (2018). Developing Small-Molecule Inhibitors of HECT-Type Ubiquitin Ligases for Therapeutic Applications: Challenges and Opportunities. Chembiochem 19, 2123-2135.
118. Jessani, N., and Cravatt, B.F. (2004). The development and application of methods for activity-based protein profiling. Curr Opin Chem Biol 8, 54-59.
119. Liu, Y., Patricelli, M.P., and Cravatt, B.F. (1999). Activity-based protein profiling: The serine hydrolases. Proceedings of the National Academy of Sciences 96, 14694.
120. Kato, D., Boatright, K.M., Berger, A.B., Nazif, T., Blum, G., Ryan, C., Chehade, K.A., Salvesen, G.S., and Bogyo, M. (2005). Activity-based probes that target diverse cysteine protease families. Nat Chem Biol 1, 33-38.
121. Cravatt, B.F., Wright, A.T., and Kozarich, J.W. (2008). Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. Annu Rev Biochem 77, 383-414.
122. Martell, J., and Weerapana, E. (2014). Applications of copper-catalyzed click chemistry in activity-based protein profiling. Molecules 19, 1378-1393.
123. Gopinath, P., Ohayon, S., Nawatha, M., and Brik, A. (2016). Chemical and semisynthetic approaches to study and target deubiquitinases. Chemical Society Reviews 45, 4171-4198.

124. Borodovsky, A., Kessler, B.M., Casagrande, R., Overkleeft, H.S., Wilkinson, K.D., and Ploegh, H.L. (2001). A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. EMBO J 20, 5187-5196.
125. Lam, Y.A., Xu, W., DeMartino, G.N., and Cohen, R.E. (1997). Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome. Nature 385, 737-740.
126. Ekkebus, R., van Kasteren, S.I., Kulathu, Y., Scholten, A., Berlin, I., Geurink, P.P., de Jong, A., Goerdayal, S., Neefjes, J., Heck, A.J.R*., et al.* (2013). On terminal alkynes that can react with active-site cysteine nucleophiles in proteases. J Am Chem Soc 135, 2867-2870.
127. Sommer, S., Weikart, N.D., Linne, U., and Mootz, H.D. (2013). Covalent inhibition of SUMO and ubiquitin-specific cysteine proteases by an in situ thiol–alkyne addition. Bioorganic & Medicinal Chemistry 21, 2511-2517.
128. Misaghi, S., Galardy, P.J., Meester, W.J., Ovaa, H., Ploegh, H.L., and Gaudet, R. (2005). Structure of the ubiquitin hydrolase UCH-L3 complexed with a suicide substrate. J Biol Chem 280, 1512-1520.
129. Mulder, M.P., El Oualid, F., ter Beek, J., and Ovaa, H. (2014). A native chemical ligation handle that enables the synthesis of advanced activity-based probes: diubiquitin as a case study. Chembiochem 15, 946-949.
130. Haj-Yahya, N., Hemantha, H.P., Meledin, R., Bondalapati, S., Seenaiah, M., and Brik, A. (2014). Dehydroalanine-Based Diubiquitin Activity Probes. Organic Letters 16, 540-543.
131. McGouran, J.F., Gaertner, S.R., Altun, M., Kramer, H.B., and Kessler, B.M. (2013). Deubiquitinating enzyme specificity for ubiquitin chain topology profiled by di-ubiquitin activity probes. Chem Biol 20, 1447-1455.
132. Li, G., Liang, Q., Gong, P., Tencer, A.H., and Zhuang, Z. (2014). Activity-based diubiquitin probes for elucidating the linkage specificity of deubiquitinating enzymes. Chem Commun (Camb) 50, 216-218.
133. Flierman, D., van der Heden van Noort, G.J., Ekkebus, R., Geurink, P.P., Mevissen, T.E., Hospenthal, M.K., Komander, D., and Ovaa, H. (2016). Non-hydrolyzable Diubiquitin Probes Reveal Linkage-Specific Reactivity of Deubiquitylating Enzymes Mediated by S2 Pockets. Cell Chem Biol 23, 472-482.
134. Tan, X.D., Pan, M., Gao, S., Zheng, Y., Shi, J., and Li, Y.M. (2017). A diubiquitin-based photoaffinity probe for profiling K27-linkage targeting deubiquitinases. Chem Commun (Camb) 53, 10208-10211.
135. Liang, J., Zhang, L., Tan, X.L., Qi, Y.K., Feng, S., Deng, H., Yan, Y., Zheng, J.S., Liu, L., and Tian, C.L. (2017). Chemical Synthesis of Diubiquitin-Based Photoaffinity Probes for Selectively Profiling Ubiquitin-Binding Proteins. Angew Chem Int Ed Engl 56, 2744-2748.
136. Zhang, X., Smits, A.H., van Tilburg, G.B., Jansen, P.W., Makowski, M.M., Ovaa, H., and Vermeulen, M. (2017). An Interaction Landscape of Ubiquitin Signaling. Mol Cell 65, 941-955.e948.
137. Zhang, X., Smits, A.H., van Tilburg, G.B., Ovaa, H., Huber, W., and Vermeulen, M. (2018). Proteome-wide identification of ubiquitin interactions using UbIA-MS. Nat Protoc 13, 530-550.
138. Gong, P., Davidson, G.A., Gui, W., Yang, K., Bozza, W.P., and Zhuang, Z. (2018). Activity-based ubiquitin-protein probes reveal target protein specificity of deubiquitinating enzymes. Chem Sci 9, 7859-7865.
139. Meledin, R., Mali, S.M., Kleifeld, O., and Brik, A. (2018). Activity-Based Probes Developed by Applying a Sequential Dehydroalanine Formation Strategy to Expressed Proteins Reveal a Potential α-Globin-Modulating Deubiquitinase. Angew Chem Int Ed Engl 57, 5645-5649.

140. Hewings, D.S., Heideker, J., Ma, T.P., AhYoung, A.P., El Oualid, F., Amore, A., Costakes, G.T., Kirchhofer, D., Brasher, B., Pillow, T.*, et al.* (2018). Reactive-site-centric chemoproteomics identifies a distinct class of deubiquitinase enzymes. Nat Commun 9, 1162.
141. Love, K.R., Pandya, R.K., Spooner, E., and Ploegh, H.L. (2009). Ubiquitin C-terminal electrophiles are activity-based probes for identification and mechanistic study of ubiquitin conjugating machinery. ACS Chem Biol 4, 275-287.
142. Byrne, R., Mund, T., and Licchesi, J.D.F. (2017). Activity-Based Probes for HECT E3 Ubiquitin Ligases. Chembiochem 18, 1415-1427.
143. Xu, L., Fan, J., Wang, Y., Zhang, Z., Fu, Y., Li, Y.-M., and Shi, J. (2019). An activity-based probe developed by a sequential dehydroalanine formation strategy targets HECT E3 ubiquitin ligases. Chemical Communications 55, 7109-7112.
144. Evans, P. (2011). An introduction to data reduction: Space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr 67, 282-292.
145. Shiryev, S.A., Papadopoulos, J.S., Schäffer, A.A., and Agarwala, R. (2007). Improved BLAST searches using longer words for protein seeding. Bioinformatics 23, 2949-2951.
146. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67, 235-242.
147. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-2132.
148. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., and Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res 43, W580-584.
149. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Res 41, W597-600.
150. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7, 539.
151. Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E.*, et al.* (2012). ExPASy: SIB bioinformatics resource portal. Nucleic acids research 40, W597-W603.
152. Abramoff, M., Magalhães, P., and Ram, S.J. (2003). Image Processing with ImageJ. Biophotonics International 11, 36-42.
153. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66, 12-21.
154. Gabadinho, J., Beteva, A., Guijarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M., Brockhauser, S., Flot, D., Gordon, E., Hall, D.*, et al.* (2010). MxCuBE: A synchrotron beamline control environment customized for macromolecular crystallography experiments. Journal of synchrotron radiation 17, 700-707.
155. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res 28, 235-242.
156. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr 40, 658-674.
157. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a

comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221.
158. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845-858.
159. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol 372, 774-797.
160. Lu, Z. (2011). PubMed and beyond: a survey of web tools for searching biomedical literature. Database (Oxford) 2011, baq036.
161. Bond, S.R., and Naus, C.C. (2012). RF-Cloning.org: an online tool for the design of restriction-free cloning projects. Nucleic Acids Res 40, W209-W213.
162. Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr 66, 125-132.
163. David, M.P.C., Lapid, C.M., and Daria, V.R.M. (2008). An efficient visualization tool for the analysis of protein mutation matrices. BMC Bioinformatics 9, 218.
164. van den Ent, F., and Lowe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods 67, 67-74.
165. Ross, J. (2020). Biological physics studies of microtubules, taxol, and the microtubules-associated protein, tau.
166. Seifert, E. (2014). OriginPro 9.1: Scientific Data Analysis and Graphing Software—Software Review. Journal of Chemical Information and Modeling 54, 1552-1552.
167. Lorenz, S., Bhattacharyya, M., Feiler, C., Rape, M., and Kuriyan, J. (2016). Crystal Structure of a Ube2S-Ubiquitin Conjugate. PLoS One 11, e0147550-e0147550.
168. Serniwka, S.A., and Shaw, G.S. (2009). The structure of the UbcH8-ubiquitin complex shows a unique ubiquitin interaction site. Biochemistry 48, 12169-12179.
169. Swatek, K.N., Aumayr, M., Pruneda, J.N., Visser, L.J., Berryman, S., Kueck, A.F., Geurink, P.P., Ovaa, H., van Kuppeveld, F.J.M., Tuthill, T.J.*, et al.* (2018). Irreversible inactivation of ISG15 by a viral leader protease enables alternative infection detection strategies. Proc Natl Acad Sci U S A 115, 2371-2376.
170. Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42, W320-W324.
171. Salvat, C., Wang, G., Dastur, A., Lyon, N., and Huibregtse, J.M. (2004). The -4 Phenylalanine Is Required for Substrate Ubiquitination Catalyzed by HECT Ubiquitin Ligases. Journal of Biological Chemistry 279, 18935-18943.
172. Hersh, R.T. (1967). Atlas of Protein Sequence and Structure, 1966. Systematic Biology 16, 262-263.
173. States, D.J., Gish, W., and Altschul, S.F. (1991). Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. Methods 3, 66-70.
174. Kim, H.C., and Huibregtse, J.M. (2009). Polyubiquitination by HECT E3s and the determinants of chain type specificity. Mol Cell Biol 29, 3307-3318.
175. Streich Jr, F.C., and Lima, C.D. (2016). Capturing a substrate in an activated RING E3/E2–SUMO complex. Nature 536, 304-308.
176. Myant, K.B., Cammareri, P., Hodder, M.C., Wills, J., Von Kriegsheim, A., Győrffy, B., Rashid, M., Polo, S., Maspero, E., Vaughan, L.*, et al.* (2017). HUWE1 is a critical colonic tumour suppressor gene that prevents MYC signalling, DNA damage accumulation and tumour initiation. EMBO molecular medicine 9, 181-197.
177. Ogunjimi, A.A., Briant, D.J., Pece-Barbara, N., Le Roy, C., Di Guglielmo, G.M., Kavsak, P., Rasmussen, R.K., Seet, B.T., Sicheri, F., and Wrana, J.L. (2005). Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. Mol Cell 19, 297-308.
178. Sakata, E., Satoh, T., Yamamoto, S., Yamaguchi, Y., Yagi-Utsumi, M., Kurimoto, E., Tanaka, K., Wakatsuki, S., and Kato, K. (2010). Crystal structure of UbcH5b~ubiquitin

intermediate: insight into the formation of the self-assembled E2~Ub conjugates. Structure 18, 138-147.
179. Page, R.C., Pruneda, J.N., Amick, J., Klevit, R.E., and Misra, S. (2012). Structural insights into the conformation and oligomerization of E2~ubiquitin conjugates. Biochemistry 51, 4175-4187.
180. Ranaweera, R.S., and Yang, X. (2013). Auto-ubiquitination of Mdm2 enhances its substrate ubiquitin ligase activity. J Biol Chem 288, 18939-18946.
181. Brzovic, P.S., and Klevit, R.E. (2006). Ubiquitin Transfer from the E2 Perspective: Why is UbcH5 So Promiscuous? Cell Cycle 5, 2867-2873.
182. Kumar, P., Magala, P., Geiger-Schuller, K.R., Majumdar, A., Tolman, J.R., and Wolberger, C. (2015). Role of a non-canonical surface of Rad6 in ubiquitin conjugating activity. Nucleic Acids Res 43, 9039-9050.
183. Knipscheer, P., van Dijk, W.J., Olsen, J.V., Mann, M., and Sixma, T.K. (2007). Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. EMBO J 26, 2797-2807.
184. Pichler, A., and Melchior, F. (2002). Ubiquitin-related modifier SUMO1 and nucleocytoplasmic transport. Traffic 3, 381-387.
185. Pichler, A., Knipscheer, P., Saitoh, H., Sixma, T.K., and Melchior, F. (2004). The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. Nat Struct Mol Biol 11, 984-991.
186. Reverter, D., and Lima, C.D. (2005). Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. Nature 435, 687-692.
187. Herrador, A., Leon, S., Haguenauer-Tsapis, R., and Vincent, O. (2013). A Mechanism for Protein Monoubiquitination Dependent on a trans-Acting Ubiquitin-binding Domain. J Biol Chem 288.
188. Chen, Z., Jiang, H., Xu, W., Li, X., Dempsey, D.R., Zhang, X., Devreotes, P., Wolberger, C., Amzel, L.M., Gabelli, S.B.*, et al.* (2017). A Tunable Brake for HECT Ubiquitin Ligases. Molecular Cell 66, 345-357.e346.
189. Wang, S., Tian, Y., Wang, M., Wang, M., Sun, G.-B., and Sun, X.-B. (2018). Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. Front Pharmacol 9, 353-353.
190. Youdim, M.B., Gross, A., and Finberg, J.P. (2001). Rasagiline [N-propargyl-1R(+)-aminoindan], a selective and potent inhibitor of mitochondrial monoamine oxidase B. Br J Pharmacol 132, 500-506.

7 APPENDIX

7.1 Abbreviations

Prefixes

µ	micro
m	milli
k	kilo
Units	
°	degree
°C	degree Celsius
A	ampere
Å	Angström
Da	Dalton
g	gram
x g	gravitational force
h	hour
K	Kelvin
l	liter
m	meter
min	minute
M	molar (mol/l)
OD	optical density
rpm	revolutions per minute
s	second
v/v	volume per volume
w/v	weight per volume
α	anti
A	alanine, Ala
A_{280}	measured absorbance at 280 nm

AKT1	serine/threonine-protein kinase
AMP	adenosine-5´-monophosphate
Amp	ampicillin
APC/C	anaphase-promoting complex/cyclosome
APS	ammonium persulfate
ATP	adenosine-5´-triphosphate
BESSY	Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung
bp	base pairs
BSA	bovine serum albumin
[c]	concentration
C	cysteine, Cys
cat	catalytic
CC1/2	correlation coefficient 1/2
CCP4	Collaborative Computational Project, Number 4, 1994
CD	circular dichroism
CV	column volume
D	aspartic acid, Asp
ddH2O	bi-distilled water
DNA	desoxyribonucleic acid
DNase	deoxyribunucelase
dNTP	desoxyribonucleoside-5´-triphosphate (dATP, dCTP, dGTP, dTTP)
DSB	double strand break
DTT	dithiothreitol
DUB	deubiquitinating enzyme
E	glutamic acid, Glu
E1	ubiquitin-activating enzyme
E2	ubiquitin-conjugating enzyme
E3	ubiquitin ligase

E. coli	*Escherichia coli*
ECL	enhanced chemiluminescence
EDTA	ethylendiamintetraacetate desoxyribonucleic acid
ESI-MS	electrospray ionization mass spectrometry
EtOH	ethanol
F	phenylalanine, Phe
F	forward
FA	Fanconi anemia
FAN	Fanconi-associated nuclease
FANCD2	Fanconi anemia protein D
FANCI	Fanconi anemia group 1 protein
FL	full-length
FPLC	fast protein liquid chromatography
G	glycine, Gly
H	histidine, His
HA	human influenza hemagglutinin
HF	high fidelity
HIF-1α	hypoxia-inducible factor 1-alpha
His_6	hexa-histidine
HECT	homologous to E6AP C-Terminus
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HRP	horseradish peroxidase
I	isoleucine, Ile
IMAC	immobilized metal affinity chromatography
IMiD	immune modulators
IPTG	isopropyl β-D-1-thiogalactopyranoside
K	lysine, Lys
Kan	kanamycin
KD	dissociation constant

L	leucine, Leu
LB	lysogeny broth
B-ME	β-mercapthoethanol
MeOH	methanol
MES	2-(N-morpholino)ethanesulfonic acid
MESNa	2-mercaptoethanesulfonic acid sodium salt,
MOPS	3-(N-morpholino)propanesulfonic acid
MR	molecular replacement
MS	mass spectrometry
MW	molecular weight
MWCO	molecular weight cut-off
NEB	New England Biolabs
OD600	optical density (= absorbance) measured at a wavelength of 600 nm
PA	propargylamine
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
pI	isoelectric point
PINK1	serine/threonine-protein kinase PINK1
PTEN	phosphatase and tensin homolog
PVDF	polyvinylidene difluoride
R	reverse
RBR	RING-between-RING
RF	restriction-free
RING	Really Interesting New Gene
RMSD	root-mean-square deviation
RNA	ribonucleic acid
RNF4	RING-type ubiquitin ligase 4

PROTAC	proteolysis targeting chimeras
RT	room temperature
SD	standard deviation
SD	Superdex
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
SGC	Structural Genomics Consortium
SUMO	Small ubiquitin-related modifier
TB	Terrific Broth
TBS-T	Tris-buffered saline with tween 20
TCEP	Tris-(2-carboxymethyl)-phosphine
TEMED	N,N,N´,N´-tetramethylethylendiamine
TEV	Tobacco Etch Virus
TFB	transformation buffer
TM	melting temperature
TRIM21	ubiquitin ligase TRIM21
Tris	tris-(hydroxymethyl)-aminomethan
Ub	ubiquitin
UBC	catalytic core domain of ubiquitin-conjugating enzymes
UBC2	ubiquitin-conjugating enzyme 2 from yeast
UBE2	ubiquitin-conjugating enzyme, E2
ULP1	ubiquitin-like-specific protease 1
V	valine, Val
W	tryptophane, Trp
Wnt	wingless-type MMTV integration site family member
WT	wild-type
XRCC4	X-ray repair Cross-Complementing protein 4
Y	tyrosine, Tyr
YNB	yeast nitrogen base

7.2 List of tables

Table 1: Table of activity-based ubiquitin probes for HECT-type ubiquitin ligases 21

Table 2: Oligonucleotide sequences, used in polymerase chain reactions (PCRs)..... 24

Table 3: Bacterial strains for cloning and protein expression 25

Table 4: Vectors for protein expression in bacteria..... 26

Table 5: Expression constructs. The protein sequences are from *homo sapiens*, except for the proteases 26

Table 6: Bioreagents, enzymes and kits 27

Table 7: Chemicals 27

Table 8: Commercial crystallization screens used as templates for in-house screens... 29

Table 9: Specialized consumables 30

Table 10: Scientific equipment..... 30

Table 11: Software, server-based tools and databases..... 31

Table 12: X-ray crystallographic data collection and refinement statistics for the structure of the HUWE1 HECT domain-Ub-PA conjugate 62

7.3 List of figures

Figure 1: Structural features of ubiquitin and modes of ubiquitination 2

Figure 2: PROTAC-mediated target protein degradation 6

Figure 3: Ubiquitin conjugation machinery 7

Figure 4: Schematic overview of E3 mechanisms 9

Figure 5: Phylogenetic analysis of the 28 human HECT ligases 12

Figure 6: Architecture of the HECT domain 13

Figure 7: Schematic of the reaction mechanism of the HECT domain 14

Figure 8: Crystal structures of HECT domains in different conformations 15

Figure 9: Domain organization of HUWE1 19

Figure 10: Western blotting 42

Figure 11: Schematic of the generation of HECT domain-linked Ub-PA 49

Figure 12: SDS-PAGE and SEC in the preparation of Ub-PA 50

Figure 13: Purification of the HUWE1 HECT domain 53

Figure 14: Size-exclusion chromatography of the E6AP and NEDD4 HECT domains 54

Figure 15: HECT ligase reactivity profiling 55

Figure 16: Phylogenetic analysis of the 28 human HECT ligases 55

Figure 17: Amino acid sequence alignment for the C-lobes of human E6AP, NEDD4 and HUWE1 56

Figure 18: Superposition of crystal structures of the C-lobes of HUWE1, E6AP and NEDD4 57

Figure 19: Role of the non-conserved loop region within the C-lobe for Ub-PA labeling 57

Figure 20: Crystal structure of the HUWE1 HECT domain 58

Figure 21: Role of cysteine residues in the HECT domain of HUWE1 in catalytic activity 59

Figure 22: Role of cysteine residues in the HECT domain of HUWE1 in reactivity towards Ub-PA 60

Figure 23: Crystal structure of the HUWE1 HECT domain-Ub-PA conjugate..... 61

Figure 24: Detailed view of the HUWE1 C-lobe-ubiquitin interface 63

Figure 25: Detailed view on the catalytic center in the HUWE1 HECT domain-Ub-PA conjugate 63

Figure 26: Detailed view of the position of C-terminal tail in the crystal structure of the HUWE1 HECT domain- Ub-PA conjugate 64

Figure 27: Superposition of the crystal structure of the HUWE1 HECT domain asymmetric dimer and HUWE1 HECT-Ub-PA conjugate 65

Figure 28: Structural view on the C-terminal tail-coordinating residues of the HUWE1 HECT domain-Ub-PA conjugate 65

Figure 29: CD spectra of HUWE1 HECT domain variants with substitutions of the C-terminal tail-coordinating residues 66

Figure 30: Significance of C-terminal tail-coordinating residues for the catalytic activity of the HECT domain of HUWE1 67

Figure 31: Significance of C-terminal tail-coordinating residues for substrate ubiquitination 68

Figure 32: Reactivity of the C-terminal tail-coordinating and Δ4 variants of the HUWE1 HECT domain towards Ub-PA 70

Figure 33: Ub-PA labeling reaction with the C-lobes of HUWE1, E6AP and NEDD4 71

Figure 34: Amino acid sequence alignment of the N-lobe of selected human HECT ligases 72

Figure 35: FP analysis of HUWE1 HECT domain-ubiquitin interactions..... 73

Figure 36: Structural superposition of exosite residues of two human HECT ligases 74

Figure 37: CD spectra of HUWE1 HECT domain exosite variants 75

Figure 38: Role of the exosite in the HUWE1 HECT domain for activity 76

Figure 39: Thioester transfer of ubiquitin from UBCH7 to HUWE1 exosite variants 76

Figure 40: Ub-PA labelling reactions of the HUWE1 HECT domain and exosite variants 77

Figure 41: Purification of K48-linked di-ubiquitin in complex with the HUWE1 HECT domain 78

Figure 42: Multiple amino acid sequence alignment of C-lobe of different human HECT ligases 81

Figure 43: The solution NMR structure of the catalytic core domain of UBE2D3 83

Figure 44: Top view of the crystal structure of the HECT domain of NEDD4 bound to exosite ubiquitin 84

Figure 45: Modeling of exosite ubiquitin into HUWE HECT domain-Ub-PA conjugate... 86

Figure 46: Crystal structure of the *apo* HECT domain of HUWE1 86

Figure 47: Cartoon representation of the allosteric communication..... 87

7.4 Acknowledgements

In the beginning, I am greatly indebted to my advisor, Dr. Sonja Lorenz, for her continuous support during my years at the University of Würzburg. Her patience, guidance, and enthusiasm helped me in all the time of research and writing this dissertation. She is the busiest and the most available person at the same time. She was always there to listen to my issues and give timely advice at every stage of my research. I am grateful for her valuable advice, constructive criticism, positive appreciation, and counsel throughout the research tenure. Thank you for believing in me and teaching me that 'everything will fall into place.'

Next, I also have to thank Prof. Antje Gohla, who was always kind to answer all of my questions and have many insightful discussions during my thesis committee meetings. I appreciate her willingness to meet at short any notice and guiding me throughout my research. I also have to thank Prof. Nikita Popov for his helpful advice and guidance in general for my research. No research is possible without requisite materials and resources, for this, I extend my sincere gratitude towards all my lab mates who helped me with the data analysis and other research oriented help throughout my stay in the lab.

Very special thanks to the all the members of GSLS and especially to Dr. Gabriele Blum-Oehler for being so kind and helpful from the start till the end of my PhD. I am also grateful for the way she extended her help during the COVID-19 crisis. I would also like to thank my funding body GRK 2243 for their financial support, without which this research wouldn't have been possible. I would also take this opportunity to thank all the students and faculties of GRK2243 for their valuable inputs and fruitful discussions during yearly progress report presentations. A heartfelt thanks to all my colleagues especially Nasir bhai, Ngoc, Ashwin, Susobhan bhai, Aparna, Radhika, Umair, Barbara, Lena, Anna, Bodo, Dan, Ayshwarya, Bing, Florian, Julia.H, Monika and everyone at the Rudolf Virchow Zentrum who has befriended, guided, supported, or annoyed me over the past three years. I am the product of our relationships !! My family away from home, four people who made my life here in Würzburg the best; Mohindar, Mano, Ravi and Abhishek. I also want

to thank my near and dear buddies, Vishnu, Sarath and Raj for being there for me and cracking me up with the worst jokes one can imagine!

This journey would not have been possible without the support and love of my family; My father, Mr. Mony T.R, My mother Miss. Radhamani K.G, my sister, my brother-in-law and my two sweet twin nephews. Thank you papa and amma for giving me the liberty to follow my instincts and pursue my dream. I would like to acknowledge the most special person in my life – my fiancé Shraddha Sagar, for all her thoughts, wellwishes, phone calls, texts, and being there whenever I was feeling low. I would also like to extend my love and thanks to her parents.

Last but not the least, I want to thank Dr. Aravind Penmatsa, who gave me the very first opportunity to work in his laboratory at the Indian Institute of Science, Bangalore, India. He also taught me that being patient is what you should learn if you want to be a good crystallographer. My all-time buddy and brother Ashu who taught me how to do simple calculations for dilutions by making me stand in the 4-degree cold room for several minutes. He always boosted my confidence and supported me well in my research studies and always tapped my shoulder and gave me positive feeling. I want to THANK Pratibha who fed me for over a year from her own student mess account, when I was physically weak because of health issues. I want to thank all my other collogues of AP lab for all the fun and the joy time they had given me during my stay there.

7.5 Affidavit

I hereby confirm that my thesis entitled “Elucidating ubiquitin recognition by the HECT-type ubiquitin ligase HUWE1” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Studien zur Ubiquitinerkennung durch die HECT-Typus Ubiquitinligase HUWE1“ eigenständig, d.h. insbesondere selbstständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift