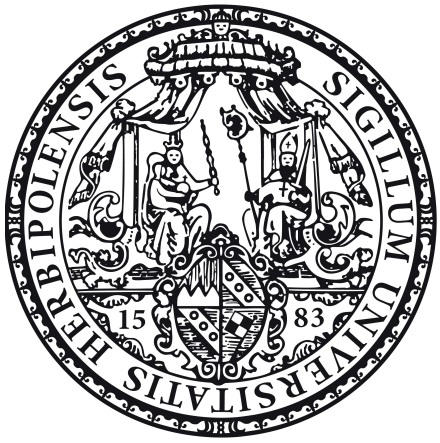


"Knowing then defeating: The ubiquitin activating enzyme, a promising target for cancer therapy."

"Erst verstehen, dann besiegen: Das Ubiquitin-aktivierende Enzym, ein vielversprechender Kandidat in der Krebstherapie"

Doctoral thesis

for a doctoral degree at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,

Section Biomedicine

submitted by

Mohit Misra

from

Lucknow, India

Würzburg 2017

Reverse page

Submitted on: ……

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Thomas Müller

Primary Supervisor: Prof. Dr. Hermann Schindelin

Supervisor (Second): Prof. Dr. Alexander Buchberger

Supervisor (Third): Dr. Sonja Lorenz

Date of Public Defence: ……....

Date of Receipt of Certificates: ……..

Dedicated to my uncle late Kamla Kant Misra

TABLE OF CONTENTS

SUMMARY.....1-2

ZUSAMMENFASSUNG.....3-4

1. MAIN INTRODUCTION

1.1 Ubiquitin and ubiquitin-like post-translational modification.....5-6

1.1.1 Ubiquitin……...6-7

1.1.2 NEDD8……..7-8

1.1.3 SUMO……8-9

1.1.4 FAT10……9

1.1.5 ISG15……..10

1.1.6 ATGs……..10

1.1.7 UFM1……10-11

1.1.8 URM1……11

1.1.9 MoaD and ThiS……11-12

1.1.10 FUB1……..12

1.1.11 HUB1……12-13

1.2 The Ubiquitylation cascade.....14-15

1.2.1 **E1 enzymes**……15

1.2.1.1 Canonical E1s……..15

1.2.1.1.1 UBA1.....17

1.2.1.1.2 SAE1/SAE2.....17

1.2.1.1.3 APPBP1/UBA3.....17

1.2.1.1.4 UBA6.....17-18

1.2.1.1.5 UBA7.....18

1.2.1.2 Noncanonical E1s……18-19

1.2.1.2.1 ATG7.....19

1.2.1.2.2 UBA5.....20

1.2.1.2.3 UBA4.....20-21

1.2.2 **E2 enzymes**……21-22

1.2.2.1 Canonical E2s……..23

1.2.2.2 Noncanonical E2s……...24

1.2.3 **E3 enzymes**……24

1.2.3.1 RING/ U-box ligases……24-25

1.2.3.2 HECT ligases……25-26

1.2.3.3 RBR ligases……..26

1.2.4 **Deubiquitinases**……26

1.2.4.1 Thiol Proteases.....…….....27-28

1.2.4.2 JAMM/MPN+ Mettaloproteases……28

1.2.5 **Specificity Determinants**……29-32

1.2.6 **Regulation and Cross-talks**……32-34

1.3 Ubiquitin activating enzyme……34

1.3.1 Discovery……..35

1.3.2 Function……35-37

1.3.3 Structural Characterization……...37-40

1.3.4 Splice variants and localization……..……....40

1.3.5 Tissue Expression……..41

1.3.6 Post-translational modifications……....41-42

1.3.7 Stability/ half-life……...42

1.3.8 Interaction partners……....43

1.3.9 Implication in cancer and neurodegenerative diseases…….....43-45

1.4 Aim of Work……..46

2. CHAPTER 2: Structural and biochemical insights into the catalytic cycle of Uba1

2.1 Abstract……47-48

2.2 Introduction……49-51

2.3 Materials and methods.....52

2.3.1 Plasmids and constructs.....52

2.3.2 Mutants.....52-53

2.3.3 Protein expression and purification.....54-55

2.3.4 Circular Dichroism spectroscopy.....55

2.3.5 Crystallization and data collection.....55-56

2.3.6 Structure determination and refinement.....57

2.3.7 Biomol Green Assay.....57-59

2.3.8 SDS-PAGE analysis of E1-E2 complex formation.....59

2.3.9 Author contributions.....59

2.4 Results and Discussions……60

2.4.1 Uba1 purification optimization……..60-62

2.4.2 Uba1 purification.....62

2.4.3 Crystallization and structure determination.....62-64

2.4.4 Mg-ATP bound Uba1 crystal structure……....64-67

2.4.5 Biochemical characterization of the ATPase activity…..….....67-71

2.4.6 Uba1-ubiquitin complex……..71-73

2.4.7 Active site remodeling.....73-75

2.4.8 The cross-over loop.....……..75-78

2.4.9 E1-E2 transthioesterification…….....78-80

2.4.10 The Uba1-Ubc13 complex.....……...80-84

2.4.11 ATP induced conformational changes allow E1-E2 disulfide bridging.....84-86

2.4.12 Complex formation does not require ATP hydrolysis.....87-88

2.4.13 ATP binding mutants are deficient in complex formation.....88

2.4.14 R21 involved in a three-way cross-talk.....88-90

2.4.15 UFD domain motion…….....90-93

2.4.16 A second E2-binding site on E1.....93-95

2.5 Conclusion.....96-98

3. Chapter 3: Targeting the ubiquitin activating enzyme for cancer therapy

3.1 Abstract.....99-100

3.2 Introduction.....101-106

3.3 Materials and methods

3.3.1 Constructs and expression.....107

3.3.2 Human UBA1 purification.....107

3.3.3 Crystallization and structure determination.....107-108

3.3.4 Inhibition assays.....108

3.3.5 MD simulations and LIE calculations.....109-110

3.3.6 Contributions to the work.....110

3.4 Results and Discussion

3.4.1 Crystal structures of Uba1-ubiquitin.inhibitor covalent adducts.....111-113

3.4.2 Closer look at the inhibitor-Uba1 interfaces.....114-115

3.4.3 Similarities between inhibitor•ubiquitin adduct and Ub•adenylate.....115-116

3.4.4 Comparative analysis with NEDD8 E1 bound NEDD8•MLN4924 structure.....116-118

3.4.5 UBA1~UB thioester inhibition assays.....118-120

3.4.6 MD simulations and LIE calculations.....120-122

3.4.7 A548T mutation negatively effects MLN7243-binding to UBA1.....122-123

3.4.8 Unique features of the ATP-binding pocket of E1 enzymes can serve as specificity regulators.....124-126

3.5 CONCLUSION.....127-129

4. CONCLUDING DISCUSSION.....130-131

5. FUTURE PERSPECTIVE.....132

5.1 Tetrahedral transition state for Uba1.....132

5.2 Fragment based screening.....132-133

5.2.1 Thermal Shift Assays.....134

5.2.2 X-ray structures of fragment bound complexes.....134-136

5.2.3 Activity assays.....136-137

5.2.4 Fragment evolution.....137-138

5.3 Targeting UBA1 of eukaryotic pathogens.....138-140

6. LIST OF PDB ENTRIES.....141

7. LIST OF PUBLICATIONS.....142

8. LIST OF EXPRESSION CONSTRUCTS.....143-144

9. ACKNOWLEDGMENT.....145-146

10. ABBREVIATIONS.....147-149

11. AFFIDAVIT.....150

12. CURRICULUM VITAE.....151-152

13. REFERENCES.....153-168

SUMMARY

Ubiquitin is a 76 amino acid long polypeptide, which is present throughout eukaryotes in a highly conserved fashion. Ubiquitin can modify proteins by becoming covalently attached to them. Eukaryotic cells employ ubiquitin to maintain and regulate fundamental cellular processes like protein degradation, the immune response and transcriptional and translational regulation. Transfer of ubiquitin to the substrate is achieved by the catalysis of three classes of enzymes namely E1, E2 and E3. Together these enzymes form a pyramidal hierarchy, where E1 stands at the apex and E3 enzymes form the base of the pathway.

The ubiquitin activating enzyme 1 (UBA1) plays a major role in ubiquitylation being the ubiquitin-dedicated E1 enzyme. In addition, it is the only enzyme in this pathway to use ATP as an energy source to catalyze two important reactions. The products of these reactions, ubiquitin adenylate and ubiquitin thioester, are the essential intermediate states of ubiquitin, for being conjugated to the target protein. With the help of X-ray crystallography and biochemical approaches, snapshots of multiple catalytic states of UBA1, where it is bound to Mg-ATP, ubiquitin and the E2 Ubc13 as substrates could be captured. With the help of these high-resolution crystal structures, deeper insights into the enzymatic mechanism of UBA1 could be attained. The resulting insights into the catalytic cycle were further validated by biochemical assays. It could be shown that ATP acts as a molecular switch to induce the enzyme's open conformation. Ubiquitin-binding to the enzyme leads to domain rotations, which facilitate the recruitment of a cognate E2 enzyme. The interdomain communication as well as the cross-talk with the substrates and the products fuel the enzymatic cycle of UBA1.

Due to the proven efficacy of proteasome inhibitors for cancer treatment, which block degradation of proteins labeled with ubiquitin, enzymes participating in the ubiquitylation cascade have been targeted by researchers for the development of novel anti-cancer therapeutics. UBA1 inhibition has been shown to preferentially induce cell death in malignant cells, and it can also be used as a strategy to overcome resistance against proteasome inhibitors. MLN7243, an adenosyl sulfamate inhibitor developed by Millenium Pharmaceutical to specifically target UBA1, is currently in Phase-I clinical trials for the treatment of solid tumors. UBA1 could be crystallized in complex with three adenosyl sulfamate inhibitors covalently linked to ubiquitin, which are promising drug candidates for cancer therapy. The inhibitors employed, MLN7243, MLN4924 and ABPA3, show distinct specificities towards different E1 enzymes. With the help of crystal structures the specificity determinants of these inhibitors could be deciphered, which were further confirmed by inhibition assays as well as molecular dynamics

simulations. Together these crystal structures provide a starting point for developing E1-specific inhibitors, which, besides their potential for medicinal purposes, are important tools to better understand the function of the ubiquitin system as well as the action of ubiquitin-like proteins.

ZUSAMMENFASSUNG

Ubiquitin ist ein 76 Aminosäuren langes Polypeptid, das in allen Eukaryoten vorkommt und hoch konserviert ist. Ubiquitin kann Proteine modifizieren indem es mittels einer kovalente Bindung an diese angeheftet wird. Eukaryotische Zellen nutzen Ubiquitin, um fundamentale zelluläre Prozesse wie den Proteinabbau, die Immunantwort sowie die Regulation der Transkription und Translation aufrecht zu erhalten. Der Transfer von Ubiquitin auf das Substrat wird durch die Katalyse von drei Enzymklassen E1, E2 und E3 erreicht. Zusammen bilden diese Enzyme eine pyramidale Hierarchie in der das E1 an der Spitze steht und die E3 Enzyme die Basis bilden.

Das Ubiquitin-aktivierende Enzym 1 (UBA1) ist das E1 Enzym für Ubiquitin und spielt somit eine Hauptrolle in der Ubiquitinierung. Weiterhin ist es das einzige Enzym dieses Stoffwechselweges, das ATP als Energie nutzt, um zwei wichtige Reaktionen zu katalysieren. Die Produkte dieser Reaktionen, Ubiquitin-Adenylat und thioesterifiziertes Ubiquitin, sind die essentiellen Ubiquitinintermediate für die Konjugation an das Zielprotein. Mit Hilfe der Röntgenstrukturanalyse und biochemischer Ansätze konnten für UBA1 Momentaufnahmen multipler katalytischer Zustände erfasst werden, in denen es an die Substrate Mg-ATP, Ubiquitin und dem E2 Enzym Ubc13 gebunden vorliegt. Mit Hilfe dieser hochaufgelösten Kristallstrukturen konnten tiefere Einblicke in den enzymatischen Mechanismus des Enzyms erzielt werden. Die gesammelten Erkenntnisse zum katalytischen Zyklus wurden mittels biochemischer Methoden validiert. Es konnte gezeigt werden, dass ATP als molekularer Schalter fungiert, um das Enzym in seine offene Konformation zu überführen. Die Bindung von Bindung an das Enzym führt zur Rotation einzelner Domänen welche die Rekrutierung des E2 Enzymes erleichtern. Die Interdomänen-Interaktion sowie molekulare Wechselwirkungen mit den Substraten und Produkten treiben den enzymatischen Zyklus von UBA1 an.

Die Einsatz von Proteasominhibitoren, die den Abbau von Ubiquitin-markierten Proteinen blockieren, in der Krebstherapie weckte das Interesse von Forschern Enzyme, die an der Ubiquitinierungs-Kaskade beteiligt sind, als neue therapeutische Ziele zur Bekämpfung von Krebserkrankungen zu erschließen. Es konnte gezeigt werden, dass die Inhibierung von UBA1 bevorzugt den Tod maligner Zellen induziert und als Strategie für die Überwindung der Resistenz von Proteasominhibitoren genutzt werden kann. MLN7243, ein Adenosyl-Sulfamat Inhibitor, der von Millenium Pharmaceuticals entwickelt wurde und spezifisch UBA1 angreift, befindet sich gegenwärtig in klinischen Studien der Phase I mit dem Ziel einer Behandlung von soliden Tumoren. UBA1 konnte im Komplex mit drei an Ubiquitin gekoppelten Adenosyl-Sulfamat Inhibitoren, die vielversprechende Wirkstoffe in der Krebstherapie

sind, kristallisiert werden. Die Inhibitoren MLN7243, MLN4924 und ABPA3 besitzen unterschiedliche Spezifitäten für verschiedene E1 Enzyme. Mit Hilfe von Kristallstrukturen konnten Spezifitätsfaktoren dieser Inhibitoren entschlüsselt werden, die im Weiteren mittels Inhibierungstests und molekulardynamischer Simulationen bestätigt werden konnten. Diese Kristallstrukturen lieferten ein klareres Bild für die Entwicklung E1-spezifischer Inhibitoren mit deren Hilfe, neben ihrer potentiellen medizinischen Anwendung, ein besseres Verständnis des Ubiquitinsystems und Ubiquitin-ähnlicher kovalenter Verknüpfungen gewonnen werden kann.

1. MAIN INTRODUCTION

1.1 Ubiquitin and ubiquitin-like post-translational modification

Post-translational modification (PTM) is a common mode of regulation of proteins which is present in abundance in eukaryotes. Usually a post-translational modification refers to the covalent attachment of a functional group on a polypeptide. Some of the most well studied types of post-translational modifications are phosphorylation, acetylation, ubiquitylation, glycosylation, S-nitrosylation and methylation. Post-translational modifications are dynamic and reversible, allowing the cell to exploit them as regulatory switches. Prokaryotes have rather few reports of post-translational modifications in the form of phosphorylation and acetylation, indicating that these modifications have been favored by evolution from simpler prokaryotic organisms to complex higher eukaryotes. Ubiquitylation is one of the highly utilized and complex forms of post-translational modifications present throughout eukaryotes. This modification is represented by the covalent attachment of ubiquitin, a 76 amino acid long polypeptide, predominantly on lysine residues of the target protein. Ubiquitin is regarded as one of the highly conserved proteins in eukaryotes and its C-terminal carboxylate can form an isopeptide bond with the ε-amino group of lysine (Hershko and Ciechanover, 1982; Varshavsky, 1997). Although proteasomal degradation is the best studied function of ubiquitin, where cells tag the target protein with ubiquitin for proteasomal degradation (Glickman and Ciechanover, 2002; Rock et al., 1994), over the years rigorous research has shown that ubiquitylation effects a myriad of cellular processes including antigen processing, DNA repair, stress responses, cell cycle progression, vesicular transport and apoptosis to name a few (Swatek and Komander, 2016). Thus, ubiquitin justifies its name very well in terms of its presence throughout eukaryotes, its omnipresence within the cell from the nucleus to the cytoplasm to various cellular compartments and its widespread influence over cellular functions.

Ever since the discovery of ubiquitin, many other ubiquitin-like modifiers have been discovered that closely resemble the structure of ubiquitin and possess the ability to modify target lysine residues like ubiquitin (Cajee et al., 2012; Kerscher et al., 2006; Welchman et al., 2005). So far 17 ubiquitin-like proteins have been reported which contain two features: 1) A globular domain called β-grasp or ubiquitin fold featuring a five-stranded mixed β-sheet surrounding a central α-helix and 2) a conformationally flexible C-terminal tail terminating in a diglycine motif. Even though many of these modifiers share very low sequence identity (as low as 14%) they have closely related 3D structures. Another interesting feature of some proteins of this family is that they are expressed in translation cassettes which possess additional amino acid residues at the

C-terminus and until specific proteases called deubiquitinating enzymes or the Ubl-specific proteases (ULPs) cleave these additional amino acids (represented as green bars in Figure 1) to free the terminal carboxy group of the diglycine motif, these precursor proteins stay in their inactive forms. A brief description of some of these ubiquitin-like proteins is provided below:

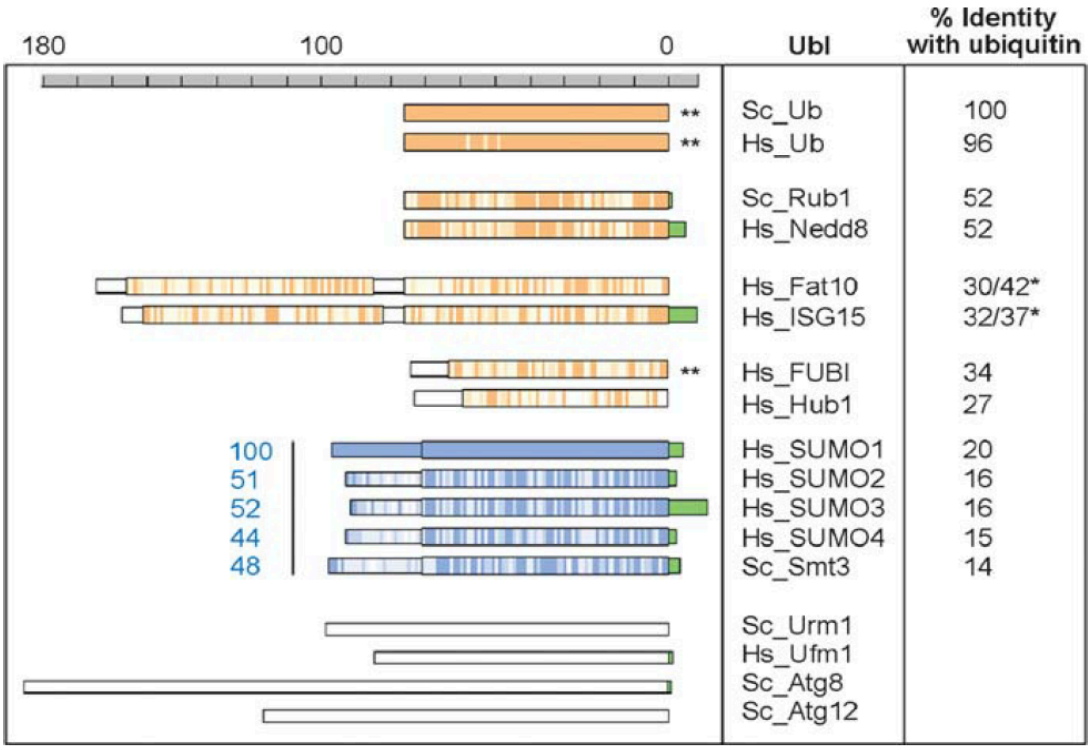


Figure 1: Sequence based comparison of ubiquitin-like modifiers with ubiquitin.

1.1.1 Ubiquitin:

Ubiquitin, first discovered in 1975, is the founder member of the ubiquitin-like proteins. It consists of two α-helices and five β-strands in a ββαββαβ arrangement (Vijaykumar et al., 1987). Other than the isopeptide linkage with the side chain of lysine residue, ubiquitin is also capable of forming covalent bonds with a cysteine residue via thioester bond, a serine or threonine residue through ester bonds and the amino group of N-terminal residue via a peptide bond (Ikeda and Dikic, 2008). Since ubiquitin itself contains seven lysine residues, it can modify itself at these positions as well as at the N-terminal amino group. Therefore, eight types of ubiquitin linkages can be imagined, which all exist in eukaryotes and are referred to as M1, K6, K11, K27, K29, K33, K48 and K63 linkage types (Swatek and Komander, 2016). With the help of these linkage forms, ubiquitin can form chains of one or mixed linkages which in turn contain various signals for regulating fate, function or localization of the target protein. For example, K11 and K48 linked polyubiquitin chains are mostly signals for proteasomal degradation

of modified protein, thereby regulating cell cycle regulation, antigen presentation and endoplasmic reticulum associated degradation (ERAD), whereas K63 linked ubiquitin chains are involved in the immune response, DNA repair, endocytosis and signal transduction (Haglund and Dikic, 2005). Although the other linkage types of ubiquitin are rather underexplored, research has shown Lys-6 linked chains to be involved in DNA repair, Lys-29 chains to play a role in lysosomal degradation, Lys-33 linked ubiquitin chains have been implicated in modifying kinases and M-1 linked linear ubiquitin chains activate nuclear factor κB (NF-κB). Moreover, K48-K63 mixed branched chains have been shown to regulate NF-κB signaling. In humans, four genes are responsible for the production of ubiquitin: UBB, UBC, UBA52 and RPS27A. UBA52 and RPS27A genes encode a single copy of ubiquitin fused to the ribosomal proteins L40 and S27a, respectively, whereas the UBB and UBC genes code for polyubiquitin precursor proteins (Finley et al., 1989; Redman and Rechsteiner, 1989). In total, three residues differ in human ubiquitin from yeast ubiquitin (S19P, D24E and S28A), corresponding to 96% sequence identity.

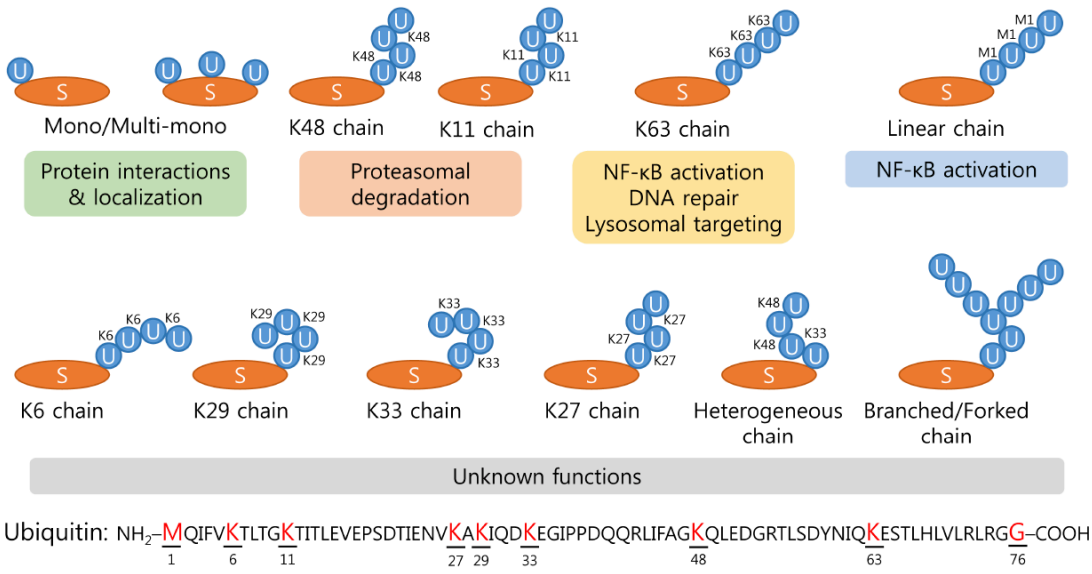


Figure 2: Types and modes of ubiquitin chain synthesis with biological relevance.

1.1.2 NEDD8:

Neural precursor cell-Expressed Developmentally Down-regulated 8 (NEDD8) is the closest relative of ubiquitin with a sequence identity of 52% (Kumar et al., 1993). The yeast homolog of this protein is called Rub1. NEDD8 modification indirectly regulates the ubiquitylation pathway as it activates nearly all members of the cullin ubiquitin ligase family (except the anaphase promoting complex-2), a major subclass of ubiquitin E3 enzymes which participate in modifying a variety of ubiquitin target substrates. Although

chain formation on substrates has not been reported for NEDD8, NEDD8 chains can be linked to the catalytic cysteine residue of Ubc12 (the E2 enzyme for NEDDylation pathway). As NEDD8 modification influences ubiquitylation, it plays an important function in regulating a subset of ubiquitin function including NF-κB signaling, cell growth, signal transduction, transcriptional control, genomic integrity and tumor suppression. In addition to cullins, other proteins like p53, EGFR, pVHL, BRCA and ribosomal proteins have also been shown to undergo NEDDylation, which also impacts cell proliferation, viability and development (Hori et al., 1999; Xirodimas et al., 2004). Interestingly, the downregulation of the deneddylase DEN1, an enzyme that cleaves NEDD8 from target proteins, was found to suppress apoptosis indicating that NEDDylation prevents apoptosis (Broemer et al., 2010).

1.1.3 SUMO:

After ubiquitin, one of the best studied ubiquitin-like modifiers is the Small Ubiquitin-like MOdifier (SUMO) and the attachment of SUMO to target protein is termed as SUMOylation. It was discovered in 1996 when it was observed to stabilize RanGAP1 (Dohmen, 2004; Matunis et al., 1996). Four isoforms of this protein have been reported in humans (SUMO1-4) whereas in yeast there is only one homolog named Smt3 (Suppressor of MIF2 mutations-3)(Su and Li, 2002). It is involved in a variety of cellular activities like cell cycle control, nuclear transport, DNA repair, response to viral infections to name a few (Johnson, 2004a; Meulmeester and Melchior, 2008). The identity of SUMO isoforms to ubiquitin ranges from 15%-20% but in the structural alignment to ubiquitin it represents the ubiquitin fold like other ubiquitin-like modifiers. SUMO-2 and 3 are practically identical sharing 86% sequence identity and are distinct from SUMO-1, which has around 44% sequence identity to SUMO-2-3. SUMO-4 shares 85% identity to SUMO-2 and is expressed mainly in the kidney (Bohren et al., 2004). SUMO isoforms demonstrate long N- and C-terminal extensions in comparison to ubiquitin. Also, due to the presence of a SUMOylation motif in the N-terminal extension, both SUMO2 and 3 can form SUMO chains at K11 whereas SUMO-1 lacks this ability (Vertegaal, 2010). The biological role of the N-terminal extensions is not well understood. SUMO-1 target proteins are involved in transcription, maintaining chromatic structure and DNA repair. Strikingly, several cases have been reported where SUMOylation competes with ubiquitylation for the same site, thereby blocking ubiquitiylation of the target proteins, *e.g*. IκB and PCNA (Wilson and Heaton, 2008). SUMO-4 appears to have a different functional mechanism as in its mature form it contains a proline residue in its C-terminus instead of a glycine which restricts its ability to covalently modify proteins. Instead this protein has been proposed to function in

conjugating DNA repair proteins involved in base excision repair non-covalently (Owerbach et al., 2005).

1.1.4 FAT10:

Human HLA-F Adjacent Transcript 10 (FAT10) is an 18 kDa protein which resembles a diubiquitin separated by a flexible linker. The N-terminal and C-terminal ubiquitin like domain of FAT10 share 29% and 36% sequence identity with ubiquitin, respectively. FAT10 was first identified amongst genes in the major histocompatibilty complex class I locus and was found to be expressed in lymphoid cell lines. (Fan et al., 1996) Remarkably, this ubiquitin-like modification is present in vertebrates only and the protein is expressed by stimuli such as interferon-γ and TNFα (tumor necrosis factor α) (Lukasiak et al., 2008). This is the only other modifier other than ubiquitin that has been implied to modify target proteins for proteasomal degradation (Schmidtke et al., 2014). FAT10 appears to mediate activation of NF-κB, a key regulator of the innate immunity, and FAT10 knockout in mice results in lymphocytes that display spontaneous cell death when compared to wild type. Furthermore, cancer cells have been shown to express elevated levels of FAT10, suggesting a role of FAT10 in tumor cell progression (Aichem and Groettrup, 2016).

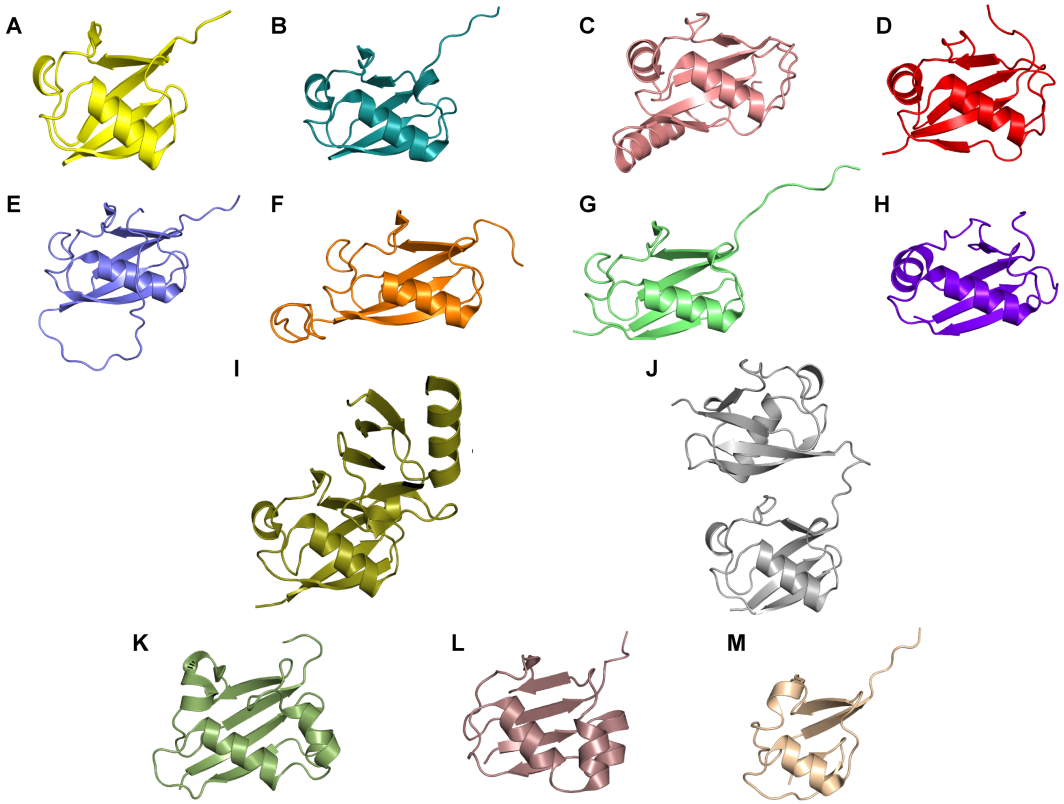


Figure 3: Three dimensional structures of ubiquitin and members of the ubiquitin family. (A) Ubiquitin, (B) NEDD8, (C) ATG8, (D) ATG12, (E) SUMO-1, (F) SUMO-2,

(G) SUMO-3, (H) UFM1, (I) ISG15, (J) homology model of FAT10 (derived from a linear diubiquitin structure), (K) yeast Urm1, (L) prokaryotic MoaD, and (M) ThiS.

1.1.5 ISG15:

ISG15 (Interferon Stimulated Gene 15) was the second Ubl to be discovered after ubiquitin and it also resembles a diubiquitin in its structure (Martensen and Justesen, 2004). The 17 kDa protein is encoded by an anti-viral response gene whose expression is induced by type I IFN (interferon) and LPS (lipopolysaccharide). Like FAT10, ISG15 is also reported in higher eukaryotes only and is absent from insects implying its co-evolution with the IFN signaling pathway (Skaug and Chen, 2010). Interestingly, ISG15 is the only other Ubl, which has the last six residues identical to ubiquitin (LRLRGG). Overexpression of ISG15 in cell culture blocks replication of a wide range of viruses including HIV, Ebola VP40 and Herpes (Lenschow et al., 2007; Okumura et al., 2006; Okumura et al., 2008). Consequently, viruses have evolved proteases that can deconjugate ISG15 from viral proteins, thus counteracting anti-viral response. Notably, ISG15 has also been reported to form chains, although their physiological relevance is not yet clear (Giannakopoulos et al., 2005).

1.1.6 ATGs:

Autophagy is a well-characterized system for degrading bulky cellular components with help of the lysosomal pathway. Two ubiquitin-like modifiers, ATG8 and ATG12, have been shown to play crucial role in this process (Mizushima et al., 1998b). Both proteins possess a ubiquitin-like fold and employ ubiquitin-like mechanisms to modify substrates. Remarkably, ATG8 is the only ubiquitin-like modifier known which modifies lipids, specifically PE (phosphatidylethanolamine) instead of a protein substrate (Ichimura et al., 2000). In this way, ATG8 is anchored to membranes. ATG8 has numerous homologues in humans, for example LC3, GABARAP and GATE-16. It has been shown that there is a linkage between the ATG12 and ATG8 pathways in such a manner that a defect in one pathway negatively affects the other (Mizushima et al., 1999; van der Veen and Ploegh, 2012).

1.1.7 UFM1:

Ubiquitin-Fold Modifier 1 (Ufm1) is a 9.1 kDa protein that shares 16% identity with ubiquitin. Exceptionally, ATG8, ATG12 and UFM1 are the only Ubls that have single glycine residue at their C-terminus (Komatsu et al., 2004). UFM1 is one of the latest ubiquitin-like modifiers to be discovered and recent studies revealed that it participates in fatty acid metabolism, endoplasmic reticulum stress and erythroid development.

Ufm1 is present in plants and animals but is absent in fungi (Tatsumi et al., 2010; Tatsumi et al., 2011).

1.1.8 URM1:

Ubiquitin-related modifier 1 is one of the least understood ubiquitin-like modifications. It shares very limited sequence identity with ubiquitin, however, it is closely related to the prokaryotic sulfur-carrier proteins ThiS and MoaD, which are involved in thiamine and molybdenum cofactor (Moco) biosynthesis, respectively (Furukawa et al., 2000; Schmitz et al., 2008a; Xu et al., 2006). URM1, therefore is believed to be the evolutionary link between these bacterial proteins and ubiquitin. Unusually, Urm1 contains a thiocarboxylate group at the C-terminus and is involved in the transfer of a sulfur atom thereby modifying certain species of tRNA (Huang et al., 2008a; Nakai et al., 2008). URM1 has also been shown to modify the peroxiredoxin Ahp1 (alkyl hydroperoxide reductase-1) indicating its involvement in oxidative stress responses (Goehring et al., 2003a). Moreover, Δurm1 mutants were found to be sensitive to rapamycin, which suggests that urmylation takes part in TOR (target of rapamycin) signaling (Chan et al., 2000; Goehring et al., 2003b).

1.1.9 MoaD and ThiS:

In prokaryotes, synthesis of Moco and thiamine requires a sulfur atom to be inserted into their precursor (Rajagopalan and Johnson, 1992; Taylor et al., 1998). The sulfur-carrier proteins MoaD and ThiS perform this function for molybdenum cofactor and thiamine synthesis, respectively, where they transfer the sulfur atom a thiocarboxylate group located at the C-terminus of these proteins. Both MoaD and ThiS share the ubiquitin fold (β-grasp fold) and terminate with the diglycine motif conserved in ubiquitin and the ubiquitin-like modifiers family. Furthermore, both proteins share the same mechanism of activation via adenylation of the C-terminus in an ATP dependent manner by homologous activating (E1-like) enzymes (MoeB for MoaD and ThiF for ThiS) (Rudolph et al., 2001; Wang et al., 2001).

Although MoaD and ThiS provide the evolutionary link of ubiquitin-like systems in prokaryotes still these proteins differ from ubiquitin in terms of their function. Whereas ubiquitin is covalently attached to the target proteins, MoaD and ThiS act as sulfur donor proteins. Although there have been several reports of bacterial proteins hijacking the host's ubiquitin proteasome system (UPS), it is mostly in the context of pathogenicity (Ashida et al., 2014). Therefore, in prokaryotes there is still no system that is directly comparable to ubiquitin- like modifications in eukaryotes. However, in *Mycobacterium tuberculosis* a 64 residue protein called Pup (prokaryotic ubiquitin-like

protein) has been shown to be covalently attached to the lysine residues of specific proteins and, strikingly, to target them for degradation by the mycobacterial proteasome. Pup, however, does not possess a β-grasp fold and does not contain a C-terminal glycine. Instead, it terminates with a glutamine residue which is later converted to a glutamate (Maupin-Furlow, 2014). Recently, Lehmann et al. reported UBact (Ubiquitin actinobacteria), a family of ubiquitin-like proteins in gram-negative bacteria based on bioinformatics analysis of newly sequenced bacterial genomes. They propose that due to gene association, size and a conserved C-terminal G/E/Q motif, these proteins can serve as conjugatable tags in bacteria (Scheel et al., 2005).

1.1.10 FUB1:

Fau ubiquitin-like protein 1 (FUB1) is a subunit of MNSF (Monoclonal Nonspecific Suppressor Factor) that inhibits several immunological responses (Nakamura et al., 1995). FUB1 contains a conjugatable C-terminal glycine, which associates with the MNSFα subunit (Nagata et al., 1998). The active MNSF complex interacts with interleukin-11-like receptors on appropriate cell types and induces phosphorylation of p31 and p65 (Nakamura and Tanigawa, 2000). Interestingly, in a mass-spectroscopy based study, FUB1 has been shown to be conjugated to Bcl-G (B-cell lymphoma-G), a member of the BCL2 family which induces apoptosis (Nakamura and Tanigawa, 2003).

1.1.11 HUB1:

Hub1 stands for homologous to ubiquitin-1 and its mammalian homolog is called UBL5 (Ubiquitin-like protein-5) (Friedmann et al., 2001). This protein has a dityrosine motif at its C-terminus instead of a diglycine. However, Hub1 possesses the ubiquitin fold, but differs in its electrostatic surface and its C-terminus is rather structured in comparison to ubiquitin (McNally et al., 2003; Ramelot et al., 2003). Interestingly, the C-terminal tyrosine is not essential for the conjugation of Hub1 to other proteins (Luders et al., 2003). Hub1 has been shown to play a essential role in cell cycle control in *S. pombe*. It has also been shown to play a role in pre-mRNA splicing (Mishra et al., 2011). A recent study has shown that Hub1 associates with the helicase Prp5, a key regulator of early spliceosome assembly, and stimulates its ATPase activity (Karaduman et al., 2017).

Other than the aforementioned ubiquitin-like modifiers there are several other known and putative ubiquitin-like proteins that also possess the ability to reversibly modify target proteins for example DWNN (Domain With No Name), BUBL1/2 (Bacterial intein-like UBiquitin-Like), SF3A120 and oligoadenylate synthetase (Dassa et al., 2004; Pugh et al., 2006). These proteins generally require three classes of enzymes that take-up

the free pool of these modifiers and ultimately deposit them on the substrate proteins. These three classes of enzymes are named as E1, E2 and E3 class which comprise the respective activating, conjugating and ligating enzymes. On top of that, since these modifications are reversible, deconjugating enzymes can decouple these modifications, thus maintaining the free pool of modifiers available in the cell.

Starting with a very simple question of how protein degradation is an ATP dependent process, Avram Hershko, together with his graduate student Aaron Ciechanover and collaborator Irvin Rose deciphered the entire cascade of ubiquitin dependent proteolysis system, for which they were awarded the Nobel prize in Chemistry in 2004 (Wilkinson, 2005). Over the last few decades, research has established ubiquitin as a universal player of eukaryotic cell biology, which modulates virtually every aspect of basic cellular needs. In this light, it is not surprising that the list of diseases implicating misregulation of the ubiquitin system is growing steadily, and currently includes many types of cancer as well as severe types of neurodegenerative disorders such as Parkinson's disease, Huntington's disease and Alzheimer's disease and type 2 diabetes.

1.2 The ubiquitylation cascade

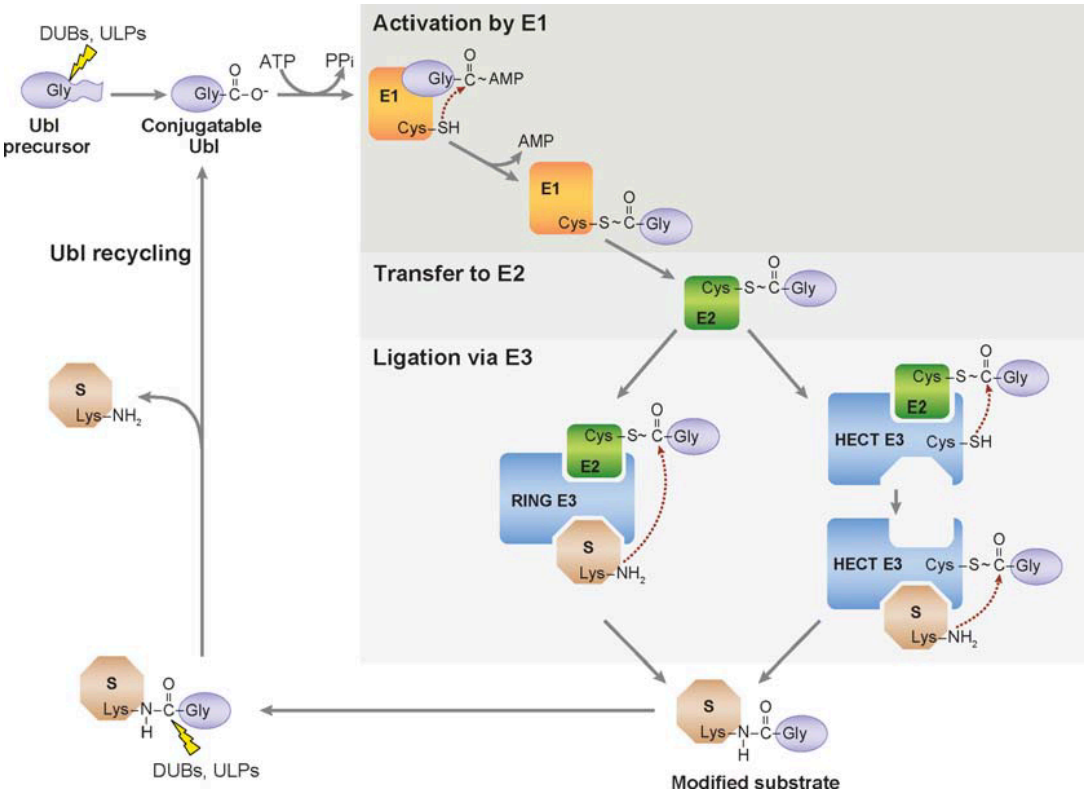


Figure 4: The ubiquitin or ubiquitin-like conjugation system.

The ubiquitylation cascade in humans is comprised of ubiquitin, two known E1 enzymes for ubiquitin, around 35 E2 enzymes and more than six hundred E3 enzymes, which together and in a specific manner covalently modify substrate proteins with a ubiquitin tag. As the numbers suggest, the ubiquitylation pathway adopts a pyramidal architecture, where E1 enzymes reside atop and the E3 enzymes form the base. The huge number of E2 and E3 enzymes determines the directionality and selectivity for the target proteins to be tagged by ubiquitin. This array of E1-E2-E3 enzyme system imparts ubiquitylation specificity, modularity and fidelity (Figure 4). Moreover, this cascade includes more than one hundred deubiquitinating enzymes which play two essential roles. First, they cleave the C-terminus of ubiquitin precursors to produce mature ubiquitin that can be taken up by E1 enzymes and second they cleave ubiquitin tags from the substrate proteins, thus keeping the process reversible and maintaining the free pool of cellular ubiquitin.

Ubl	E1	E2	E3	Biological function
Ubiquitin	UBA1/UBA6	Around fifty, (also UBCH8)	More than six hundred (HECT/RING/RBR)	Protein degradation, immune response, transcriptional regulation,
NEDD8	APP-BP1/UBA3	Ubc12/UBE2F	RING E3 ligases eg. Rbx1, Rbx2	Regulation of Cullin E3 ligases
SUMO1-4	SAE1/SAE2 (UBA2)	Ubc9	SIZ/PIAS	Nuclear localization, transcriptional regulation, DNA repair, antagonizing ubiquitylation
ISG15	UBA7 (UBE1L)	UBCH8	HERC5/Epf1	Viral immune response
FAT10	UBA6	UBE2Z	BIRC6, CHIP, RBBP6, TRIM25,TRIM51 and UHRF1	Ub/ FAT10 dependent protein degradation, cell cycle control, apoptosis, immune response
UFM1	UBA5	UFC1	UFL1	Unfolded protein response
URM1	UBA4	NI	NI	tRNA thiolation, oxidative stress response
ATG8	ATG7	ATG3	ATG5-ATG12	autophagy
ATG12	ATG7	ATG10	NI	autophagy

Table 1: Ubls with their respective E1, E2, E3 enzymes and relevant biological functions.

1.2.1 E1 enzymes

E1 enzymes directly bind the mature ubiquitin or ubiquitin-like modifiers and activate them in an ATP-dependent manner. Therefore, this class is also referred as the activating enzymes. For example, the E1 enzyme for ubiquitin is called the ubiquitin activating enzyme. Each Ubl has its own, mostly unique E1 enzyme.

So far eight activating enzymes have been reported that can activate the seventeen known ubiquitin or UBLs. According to the domain architecture and the oligomerization state, they have been classified into two sub-classes: Canonical and non-canonical E1 enzymes (Schulman and Harper, 2009). Canonical E1 enzymes either exist as monomers or heterodimers to catalyze the adenylation and thioesterification of their respective UBLs. Non-canonical E1 enzymes resemble the prokaryotic MoeB and ThiF enzymes, which homodimerize to catalyze the adenylation of the MoaD and ThiS proteins, respectively (Duda et al., 2005; Rudolph et al., 2001).

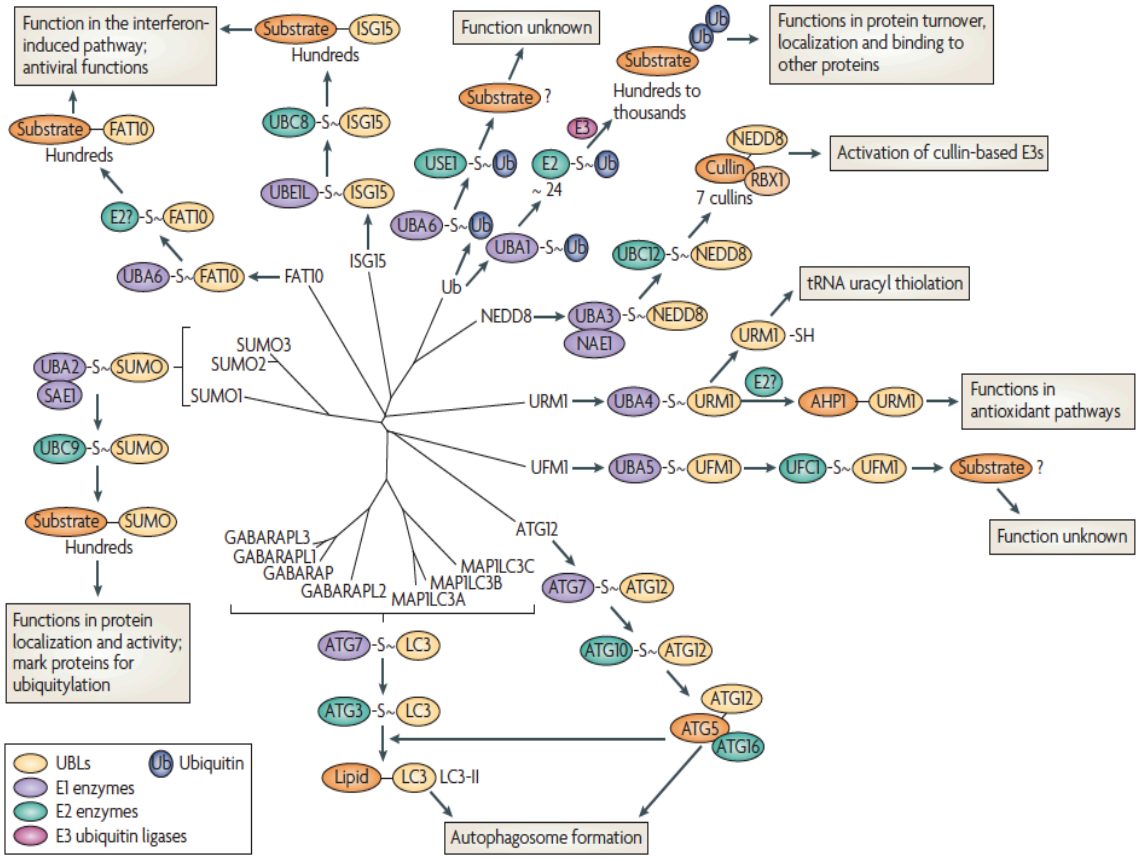


Figure 5: Evolutionary tree of ubiquitin and ubiquitin-like post translational modification with the corresponding E1-E2-E3 enzymes and substrates.

1.2.1.1 Canonical E1s

Of the eight known E1 enzymes, five belong to the canonical subclass. These enzymes exist either as monomers or heterodimers. Canonical E1s have two MoeB- and ThiF-homologous repeats, either in a single polypeptide or as two subunits which then correspond to the N- and C-terminal halves of homodimeric E1s. The adenylation domains of canonical E1s are pseudosymmetrical but functionally different in a way that one MoeB/ThiF repeat binds to Mg^{2+}-ATP and UBL and the other primarily provides structural stability (Schulman and Harper, 2009). Therefore, canonical E1s can only activate one Ubl molecule at a time, which can be seen as a compromise for the processing of downstream reactions. Additionally, they possess two additional domains for Ubl transfer to E2 enzymes, where one domain, known as the ubiquitin fold domain (UFD), is involved in the recruitment of the E2 and the second domain contains the catalytic cysteine residue required for Ubl thioester formation.

1.2.1.1.1 UBA1:

Ubiquitin activating enzyme, is the only dedicated E1 enzyme for ubiquitin, however, UBA6 can also activate ubiquitin in higher eukaryotes. It is the best characterized E1 enzyme with which the general catalytic mechanism of E1 enzymes was deciphered. As this thesis is dedicated to the structural and biochemical characterization of this enzyme, it will be discussed in more detail later.

1.2.1.1.2 SAE1/SAE2 (UBA2):

Unlike UBA1, canonical E1 enzymes also exist as heterodimeric complexes. SUMO activating enzyme is one example of these heterodimeric E1s. In this heterodimeric complex, one subunit (SAE1) acts as a structural platform and the other subunit (SAE2) contains the active sites for both adenylation as well as thioesterification (Johnson et al., 1997). The catalytic subunit is also known as UBA2. The C-terminus of UBA2 has a nuclear localization sequence which supports high level of SUMO conjugation in the nucleus. The SUMO E1 is capable of activating all four isoforms of SUMO (SUMO1-4) (Johnson, 2004b). SUMO E1 also contains a zinc binding site at the interface between the adenylation domain and the UFD domain, however, this coordination is not critical for the catalytic activity of the enzyme (Lois and Lima, 2005). Also, SUMO E1 has been shown to undergo disulfide bridge formation with its E2 (UBC9) under oxidative stress conditions, resulting in lower levels of sumoylation (Bossis and Melchior, 2006).

1.2.1.1.3 APPBP1/UBA3:

The closest relative of ubiquitin, NEDD8 has its own activating enzyme, which also exists as a heterodimeric complex (Liakopoulos et al., 1998). The NEDD8 E1 (APPBP1/UBA3) was the first E1 to be structurally characterized (Walden et al., 2003c). The APPBP1 subunit supports the catalytic activity of the enzyme whereas the UBA3 subunit contains the catalytic centers. Interestingly, unlike SAE1, APPBP1 does contact its Ubl NEDD8 through a 200 amino acid insertion that is missing in SAE1 (Walden et al., 2003b). APPBP1/UBA3 also possesses a Zn^{2+} coordination site similar to SAE1/SAE2.

1.2.1.1.4 UBA6 (UBE1L2):

The only canonical E1 enzyme that shows dual UBL-specificity is UBA6 (Jin et al., 2007a; Pelzer et al., 2007). It can activate both FAT10 and ubiquitin, however, it exhibits a higher affinity towards FAT10 (Chiu et al., 2007). UBA1 and UBA6 are distantly related with approximately 40% sequence identity. Interestingly, UBA1 is evolutionarily more closely related to UBA7 (the E1 for ISG15). UBA6 is present from in

organisms ranging from sea urchins, over zebra fish to humans but is absent in yeast, worms and flies, which indicates its selective role in deuterostomes. The deletion of the mouse *uba6* gene leads to embryonic lethality (Chiu et al., 2007). Surprisingly, mice that lack FAT10 survive, indicating that the essential function of UBA6 is not linked to FAT10 activation (Chen et al., 2014b). Where UBA1 together with its E2 and E3 enzymes adopts a pyramidal hierarchy, UBA6, with only one E2 enzyme and several predicted E3 enzymes forms an "inverted T" like architecture (Liu et al., 2017b).

1.2.1.1.5 UBA7 (UBE1L):

The activation of ISG15 is dependent on the catalytic activity of monomeric UBA7. Although structurally and mechanistically underexplored, UBA7 is supposed to possess a similar domain architecture as UBA1 as it is the closest relative of UBA1 sharing 46% sequence identity. The expression of ISG15, UBA7 and their E2 enzyme UbcH8 is induced at the transcriptional level by type-1 interferons (IFN-α/β) (Yuan and Krug, 2001). The C-terminal ubiquitin fold domain (UFD) of UBA7 has been shown to be essential for the transfer of ISG15 to UbcH8, thus suggesting that this domain discriminates the E2 enzyme for ISGylation from the E2 enzymes of other ubiquitin or ubiquitin like pathways (Durfee et al., 2008).

E1	UBA1	SAE1/ SAE2	APPBP1/ UBA3	UBA4	UBA5	UBA6	UBA7	ATG7
UBA1	**100**	**29**	**25**	**22**	**19**	**42**	**46**	**16**
SAE1/SAE2	**29**	**100**	**29**	**22**	**21**	**27**	**28**	**18**
APPBP1/UBA3	**25**	**29**	**100**	**18**	**20**	**25**	**25**	**19**
UBA4	**22**	**22**	**18**	**100**	**22**	**21**	**23**	**19**
UBA5	**19**	**21**	**20**	**22**	**100**	**19**	**19**	**17**
UBA6	**42**	**27**	**25**	**21**	**19**	**100**	**37**	**19**
UBA7	**46**	**28**	**25**	**23**	**19**	**37**	**100**	**18**
ATG7	**16**	**18**	**19**	**19**	**17**	**19**	**18**	**100**

Table 2: Percent sequence identity among E1 enzymes.

1.2.1.2 Noncanonical E1s

UBA4, UBA5 and ATG7 are the three reported non-canonical E1 enzymes. These enzymes act as homodimers and activate URM1, UFM1 and ATG8/12, respectively. The non-canonical E1 enzymes are symmetrical homodimers in the sense that each monomer of the homodimer complex can bind to its ubiquitin-like modifier. Another

general feature of this class of enzymes is that it does not possess a specialized domain containing the catalytic cysteine residue unlike the canonical E1 enzymes which undergo huge conformational alterations during thioesterification. Instead, they have the catalytic cysteine lying close to the C-terminus of ubiquitin on a loop. Noncanonical E1 enzymes share close resemblance to the prokaryotic MoeB and ThiF enzymes, both structurally and mechanistically. MoeB and ThiF are also homodimeric proteins containing two symmetric catalytic centers. Each monomer contains a four-stranded β-sheet that binds to a hydrophobic surface area of MoaD or ThiS which corresponds to the hydrophobic patch of ubiquitin. The UBL-binding site of MoaD or ThiS is stabilized by homodimerization of MoeB and ThiF. The C-terminal tail of MoaD or ThiS extends towards the bound ATP. A conserved ATP-binding arginine finger comes from the distal monomer reaching into the nucleotide binding pocket and coordinating to the γ-phosphate of ATP, while a conserved aspartate from the proximal subunit coordinates Mg^{2+}.

1.2.1.2.1 ATG7:

Autophagy or self-eating is an essential process which is initiates by the formation of a phagosome which ultimately fuses with the lysosome to form the autophagosome (Lamb et al., 2013). This process is involved in mitochondrial clearance, proteasome degradation, clearance of ubiquitin-positive aggregates and host immunity against pathogens (Kaur and Debnath, 2015). At least 18 distinct so called “Atg” proteins are key components for autophagosomal membrane formation (Farre and Subramani, 2016). As mentioned earlier, two ubiquitin-like modifiers, ATG8 and ATG12, play a critical role in this process (Mizushima et al., 1998a). Both of these ubiquitin-like proteins are activated by a single E1 enzyme, referred to as ATG7 (Tanida et al., 1999). ATG7’s C-terminal domain (CTD) homodimerizes whereas the N-terminal domains (NTD) are involved in interactions with the E2 enzymes, thus mimicking the UFD domains of canonical E1 enzymes (Taherbhoy et al., 2011). Together, the CTD and NTD form a bird-like structure where the CTD resembles the body and the NTDs appear like wings (Kaiser et al., 2012). The C-terminal helical region following the adenylation domain provides additional contacts to the ATG8 binding and plays a crucial role in activation (Noda et al., 2011). Interestingly, much like the zinc binding motif present in the SUMO and NEDD8 E1s, ATG7 also contains a zinc coordination site located between the adenylation domain and the C-terminal helical region, which seems important for its interaction with ATG8 (Hong et al., 2011).

1.2.1.2.2 UBA5:

Ufmylation is a rather recently discovered post-translational modification (Komatsu et al., 2004). However, we know the E1-E2-E3 involved in this cascade namely UBA5-UFC1-UFL1. UBA5 is another non-canonical E1 enzyme that exists as homodimer in solution and represents a minimalistic E1 composed of a single domain (Bacik et al., 2010). Recent structural and biochemical studies have shown that UBA5 activates UFM1 through a trans-binding mechanism where, other than the adenylation anchoring point for UFM1, UBA5 possesses an additional interaction surface for UFM1 at the C-terminus of the adenylation domain which is provided by the neighboring monomer (Oweis et al., 2016). This interaction surface contains a UFM1 interaction sequence (UIS) constituted of 13 amino acids (Padala et al., 2017; Xie, 2014). Furthermore, the C-terminus of the UIS is proposed to bind to UFC1, the E2 enzyme of this pathway, therefore the E2 transthioesterification also takes place via a trans-binding mechanism, similar to ATG7-ATG8/ATG12 system.

1.2.1.2.3 UBA4:

UBA4, the least understood of all the E1 enzymes, has been proposed to be the evolutionary link between the prokaryotic MoeB and ThiF proteins and eukaryotic E1s (Furukawa et al., 2000; Leimkuhler et al., 2017). This is the only E1 enzyme that is involved in sulfur transfer much like its prokaryotic orthologs and is involved in the thiolation of wobble uridine, in addition to post-translational modifications in the form of urmylation. For the tRNA thiolation function, Urm1 provides sulfur to thiolase (Ncs2-Ncs6) which together with the elongator pathway forms 5-methoxy-carbonyl-methyl-2-thio-uridine (mcm^5s^2U) at tRNA anticodon wobble positions (Chavarria et al., 2014; Dauden et al., 2017; Klassen et al., 2015; Leidel et al., 2009; Schlieker et al., 2008).

The sulfur atom required for the C-terminal thiocarboxylation of Urm1 is provided by the mobilization of the sulfur from cysteine, either directly by the desulfurase Nfs1 in an S-transfer onto Uba4 or via the sulfur transferase Tum1 (Marelja et al., 2008; Noma et al., 2009). Uba4 contains a unique Rhodanese Homology Domain (RHD), which has been shown to participate in the sulfur transfer activity via its thiol-active cysteine residue. The Urm1 thiocarboxylate has been reported to be important for both functions of Uba4, i.e. thiolation and urmylation (Judes et al., 2016).

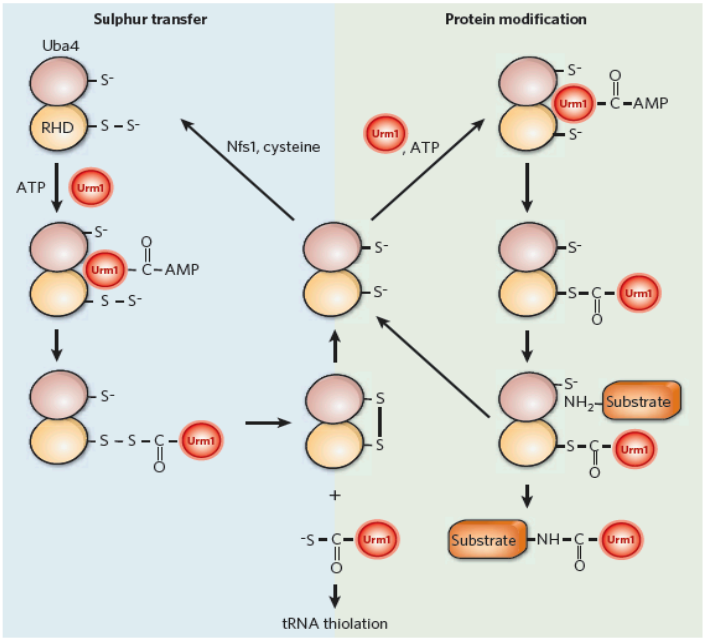


Figure 6: Dual function of the ubiquitin-like modifier Urm1 in wobble uridine thiolation as well as post-translation modification in the form of urmylation. Due to this dual role, Urm1 is considered as ubiquitin-like fossil at the crossroads of S-transfer and protein conjugation.

Interestingly, both active site cysteines present on the MoeB-like domain and the RHD domain of Uba4 seem to be synergistically involved in the uridine thiolation activity. Whereas it was shown that C397 (active site cysteine of the RHD domain) can be persulfurated, which enables the formation of an acyl-disulfide with Urm1 and finally production of Urm1-COSH, the role of C225, the catalytic cysteine of the MoeB like domain is unclear (Mueller, 2006; Schmitz et al., 2008b). C225 was proposed to have a role in the reductive cleavage of the acyl-disulfide bond between Uba4 and Urm1. Furthermore, Uba4 itself undergoes urmylation, however, the physiological relevance of this modification is not known (Judes et al., 2016).

1.2.2 E2 enzymes

E2 enzymes act as mediators between E1 and E3 enzymes. Whereas E1 enzymes act as ubiquitin or Ubl filters, E3 enzymes provide substrate specificity. Therefore, E2

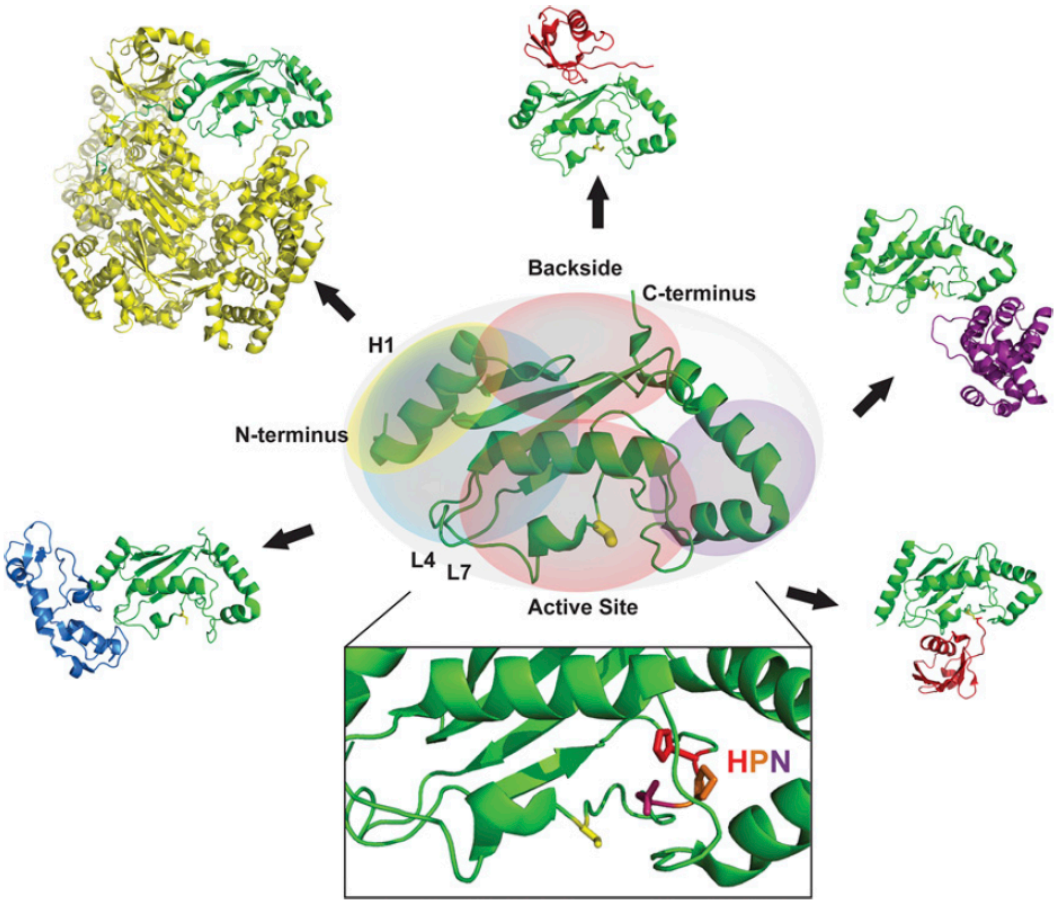


Figure 7: Core Ubc domain structure of an E2 enzyme and its multiple binding surfaces for the interaction partners. The central panel shows Ubc13 in cartoon structure (PDB: 2GMI), with the interaction surfaces described in the text colored. The E2 in each interaction pair is shown in green. Clockwise from the lower left corner: the E1/E3- binding surface as seen in the complex of Ubc13 with the RING E3 TRAF6 (blue) (PDB: 3HCT) and in Ubc12 in complex with NEDD8 E1 (yellow) (PDB: 2NVU); the backside binding surface as seen in the E2-Ub complex UbcH5c-Ub (red) (PDB: 2FUH); the substrate binding surface as observed in the SUMO E2 Ubc9 in complex with its substrate RanGAP1 (purple) (PDB: 1Z5S); the activated Ub/Ubl surface as shown in the UbcH5b~Ub complex (PDB: 3A33). The inset at the bottom presents the highly conserved HPN motif at the active site, with the active site cysteine shown in yellow.

enzymes ensure that the correct ubiquitin-like modifier is transferred from E1 to E3 or directly to the target protein depending on the class of E3 enzyme it is dealing with. Whereas baker's yeast and fission yeast possess thirteen and sixteen E2 enzymes for ubiquitylation, respectively, humans have more than 35 E2 enzymes for the same general activity, thus indicating the evolution of E2 enzymes for achieving diverse functions without compromising specificity (van Wijk and Timmers, 2010).

1.2.2.1 Canonical E2s

The E2s that can mediate transfer of ubiquitin, SUMO, NEDD8, ISG15 and FAT10 are classified as canonical E2 enzymes. These E2 enzymes possess a characteristic Ubc catalytic domain which is ~150 amino acids long in length. This domain adopts an α/β-fold, typically with four α-helices and a four stranded β-sheet. Furthermore, many E2s contain additional regions aside of the catalytic core domain. Based on this, canonical E2s have been classified into four classes: Class I includes E2s with only the Ubc domain, classes II and III have N- or C-terminal extensions, respectively, and class IV E2s have extensions at both ends. The Ubc domain is characterized by an N-terminal helix (α1), a four-stranded β-sheet (β1-4), a short 3_{10}-helix that leads into the central cross-over helix (α2), followed by two C-terminal helices (α3-4). The E2 catalytic cysteine is present in a shallow groove with the conserved HPN motif preceding the 3_{10}-helix. The histidine of the HPN motif is proposed to play a structural role in maintaining the E2 active site, while the asparagine is critical for mediating the catalysis of an isopeptide bond between ubiquitin and the substrate lysine residue by stabilizing the oxyanion reaction intermediate (Pickart et al., 2003).

Multiple surfaces on the Ubc domain are exploited to interact with a diverse range of binding partners. The N-terminal helix H1 has been shown to interact both with the UFD domain of E1 enzymes as well as to bind to the E3 enzymes (Huang et al., 2007). Due to this observation, it is clear that binding of E2 to E1 and E3 is mutually exclusive, meaning that the E2 must disengage from the E1 to relay the ubiquitin or ubiquitin-like modifier to E3 enzyme or substrate (Eletr et al., 2005; Huang et al., 2008b). Furthermore, the phenylalanine residue present on the loop L4 is required for E2 to bind to the HECT type E3 ligases, whereas hydrophobic residues lying in the L7 loop contribute to the interactions with both HECT and RING type E3 ligases (Stewart et al., 2016).

The backside surface of the E2s located on the β-meander has been shown to bind to ubiquitin or Ubls as well as to certain E3 enzymes, interactions required either for linkage specificity of ubiquitin chains or to achieve E2-E3 specificity. Cases reported for backside ubiquitin binding suggest that it can either act as a throttle or a brake for chain building. The C-terminal helices have been shown to bind to the catalytic domain of E1 enzymes as well as to the substrate to achieve specificity. Other than that, there are numerous examples of Ub/Ubl tethered to the active site via their C-terminal tail. These complexes result mostly from the exchange of the catalytic cysteine with serine which leads to the formation of an oxyester bond (Wenzel et al., 2011).

1.2.2.2 Noncanonical E2s

Noncanonical E2s comprise those E2s that carry ATG8, ATG12 and UFM1 as thioester adducts. They were considered to be too diverge to be included in the canonical group. While they do reflect canonical E2s in terms of their topology, they exhibit a lack of several structural feature, for example, the last two α-helices that are part of the Ubc fold. The noncanonical E2s are in turn charged by noncanonical E1s, i.e. ATG7 and UBA5.

E2s can further be classified into three classes: 1) The so called priming or monoubiquitylating E2s that transfer Ub/Ubl onto target protein residues (lysine/cysteine/serine/threonine), 2) the chain-building E2s that transfer Ub/Ubl onto another Ub/Ubl (Li et al., 2007) and 3) promiscuous E2 that can do either. Notably, the E2~Ub X-ray and NMR structures represent an ensemble of conformations ranging from compact “closed” to flexible “open” conformations and it has been shown that the RING E3s restrict the E2~Ub to the closed conformation to ensure efficient ubiquitin discharge (Petroski and Deshaies, 2005; Rodrigo-Brenni and Morgan, 2007).

1.2.3 E3 enzymes

E3 enzymes perform the last step of the ubiquitylation reaction, where they either directly or indirectly enable isopeptide bond formation between the ε-amino group of the substrate lysine residue and the C-terminal carboxylate of ubiquitin, which originates from ubiquitin being thioesterified to the active site cysteine of the E2 enzyme. In addition to this, the E3 enzyme also enables ubiquitin chain building on the ubiquitin for which one of the seven lysines on the donor ubiquitin itself performs the nucleophilic attack on the C-terminus of the acceptor ubiquitin (Berndsen and Wolberger, 2014). The ubiquitin ligases represent one of the largest families of enzymes in humans, comprising more than six hundred members. These enzymes maintain the fidelity of the ubiquitin system so that the right protein is given the right tag to serve the right purpose. The E3 enzymes are subdivided into three major sub-classes, which are briefly discussed below:

1.2.3.1 RING/ U-box ligases

Really Interesting New Gene (RING) ligases constitute the largest number of enzymes within the E3 enzyme family accounting to roughly eighty percent of its members (Deshaies and Joazeiro, 2009; Lydeard et al., 2013). The RING ligases are mediators and do not possess an active site, therefore the E2 enzymes which associate with the RING ligases directly transfer ubiquitin onto the target protein. RING domains originally described in 1991, have the signature of three β strands, an α-helical domain and two

free loops that are structured by two groups of Zn^{2+}- stabilizing cysteine residues (Berndsen and Wolberger, 2014). There are multiple variations to the zinc coordinating motifs as well as primary sequence of the RING domains but they all adopt a similar overall fold at the tertiary structure level. One variation of RING domains does not even have zinc coordinating residues and they are classified as U-box proteins (Patterson, 2002). Both the RING domain and U-box proteins serve as a platform for E2 enzymes thioesterified with

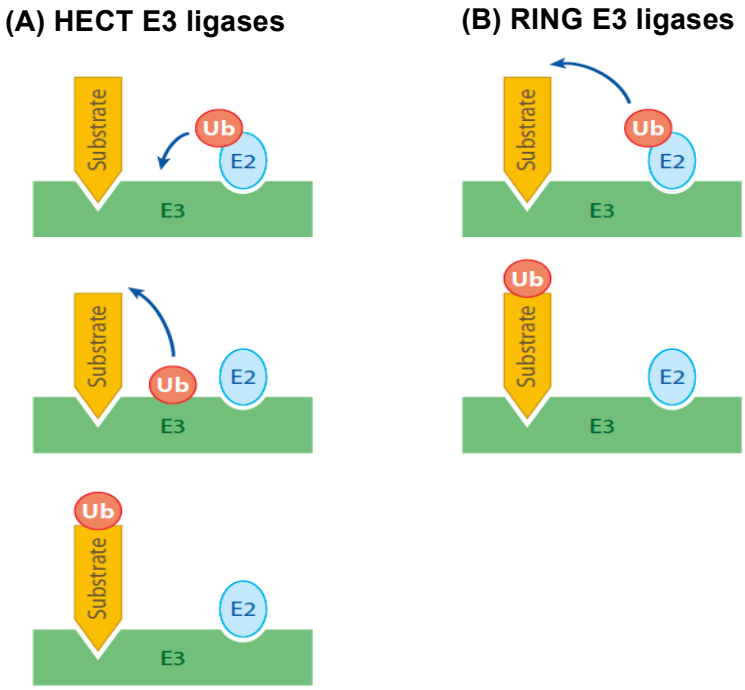


Figure 8: Mechanism of HECT and RING E3 ligases.

ubiquitin which contacts both the E2 as well as ubiquitin. This binding serves as a catalyst for the discharge of ubiquitin from the E2.

1.2.3.2 HECT ligases

Homologous to E6AP C-Terminus (HECT) is another class of E3 ligases that, instead of mediating the transfer of ubiquitin from E2 to substrate, pass the ubiquitin to the substrate lysine through another thioester intermediate formed between its own catalytic cysteine and the ubiquitin acquired from the E2 enzyme (Zheng, 2003). E6AP is a representative enzyme for this class as reflected in the name of this E3 family. HECT ligases have N-terminal substrate binding domains and at their C-termini they carry two domains, namely the N-lobe and the C-lobe, which together form the HECT domain. The C-lobe is present at the very C-terminus of the enzyme and contains the catalytic cysteine residue required for transthioesterification. The N-lobe provides the

platform for the binding of the E2~Ub species. As ubiquitin transfer is direct in case of the HECT ligases, it is believed that these enzyme regulate the positioning of the ubiquitin modification as well as the linkage specificity, in contrast to RING ligases where E2 enzymes determine linkage specificity (Lorenz et al., 2013; Rotin and Kumar, 2009).

1.2.3.3 RBR ligases

RING Betweeen RING (RBR) E3 ligases comprise both RING and HECT like properties and represent the most recently discovered class of E3 enzymes (Buetow and Huang, 2016). Interestingly, these enzymes are autoinhibited by accessory domains in the native state. The RBR domain was originally identified in fruit fly Ariadne-1 and the Parkinson associated E3 ligase Parkin whose mutations are reported in the early onset of the disease. These enzymes possess two RING domains where one RING domain contains the catalytic cysteine to mediate a HECT-like mechanism and the other RING domain serves as the binding interface for charged E2s, as reported for RING ligases (Smit and Sixma, 2014). Parkin contains an N-terminal ubiquitin-like domain (UBL), which keeps the enzyme in an auto-inhibited state by blocking the E2 binding site, whereas in Ariadne a C-terminal domain performs the analogous function by masking the active site cysteine. Remarkably, post-translational modifications, in particular phosphorylation, play critical roles in activating these enzymes, for example, phosphorylation of Ser65 of ubiquitin as well as the ubiquitin-like domain present in the enzyme has been shown to trigger the activity of Parkin (Berndsen and Wolberger, 2014).

1.2.4 Deubiquitinases

Like other post-translational modifications ubiquitylation is also a reversible process. Deubiquitinases (DUBs) play three crucial functions: 1) As mentioned before, ubiquitin and many ubiquitin-like modifiers are expressed with additional amino acids at their C-termini. Deubiquitinases remove these additional amino acids through their protease activity and free the C-terminus of these post-translational modifiers to enable their activation. 2) DUBs counteract the process of ubiquitylation by removing the ubiquitin tags from the substrate proteins. This can either rescue them from being degraded by the proteasome or can change the meaning of the tag to alter its purpose. 3) These enzymes are also critical to maintain the free pool of ubiquitin in the cell. In other words, they recycle ubiquitin to be reused and retargeted (Komander et al., 2009). There have been around one hundred different DUB enzymes reported in humans, which can be classified into two major groups:

1.2.4.1 Thiol Proteases

As suggested by the name, thiol proteases possess an active site cysteine, which is accompanied by a histidine and an aspartate residue to form a catalytic triad. In some cases, the aspartate is missing and the catalytic dyad of a histidine and a cysteine residue is sufficient for the protease activity (Mevissen and Komander, 2017; Wolberger, 2014).

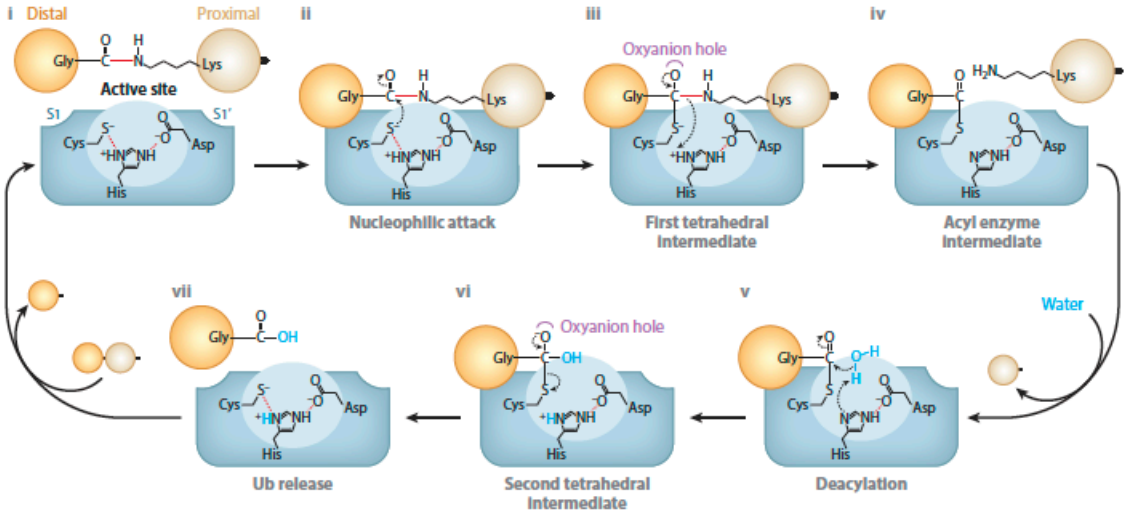


Figure 9: Reaction mechanism of thiol proteases.

Figure 9 shows the catalytic mechanism of the cysteine proteases, where in the first step the catalytic cysteine carries out a nucleophilic attack on the C-terminal carbon from the distal ubiquitin (i.e. the ubiquitin molecule whose C-terminus is involved in the isopeptide bond) which results in an acyl intermediate. The negatively charged transition state is stabilized by the hydrogen donating histidine residue. In the subsequent step, a water molecule hydrolyses the acyl intermediate, thus cleaving the isopeptide bond. The aspartate holds the histidine in the conformation apt for catalysis.

Thiol proteases have been classified into five sub-classes based on their structural folds. **1) Ubiquitin Specific Proteases (USPs)** represent the biggest family of deubiquitinases and they can be discriminated by their finger-palm-thumb subdomain architecture of the catalytic domain. USPs are generally unspecific deubiquitinases, meaning they are able to cleave several types of ubiquitin linkage types, however, there are some exceptions. **2) OTU (Ovarian Tumour) proteases** are structurally distinct from USPs and usually show linkage specific cleavage that is why there are proposed to specifically detect ubiquitin linkages on substrates. (Mevissen et al., 2013) **3) Ubiquitin C-terminus Hydrolases (UCHs)** were the first family to be structurally characterized and they feature a prominent loop covering the active site. This

configuration limits the ability of UCHs to accommodate large, folded ubiquitin conjugates, therefore it was proposed that these enzymes can only act on small peptide conjugates emerging as a result of proteasomal or lysosomal activity. **4) Josephins** represent yet another fold for cysteine DUBs and consists of an extended helical arm that appears to regulate access to active site. The extension of a poly-glutamine stretch of the representative protein Ataxin-3 leads to the neurodegenerative disorder Machado-Joseph disease and that is how this family obtained its name. **5) MINDY (Motif Interacting with ubiquitin containing novel DUB family)** is the most recently discovered class of thiol proteases which does not share any homology to known DUBs and is highly selective for K48 linked ubiquitin chains. Moreover, this class of enzymes acts on the distal end of the ubiquitin chain, thereby behaving like an exo-DUB (Rehman et al., 2016).

1.2.4.2 JAMM/MPN+ Metalloproteases

JAMM (JAB1/MPN/MOV34) family metalloproteases have two zinc coordination sites and feature a glutamate, an aspartate, two histidines and a serine residue at their active site. One of the zinc ions activates a water molecule for peptide bond hydrolysis whereas the other active site residues not involved in metal coordination position the substrate and reaction intermediates through direct interactions (Mevissen and Komander, 2017).

Interestingly, these metalloproteases are usually found associated with large protein complexes like the 26 S proteasome-associated Rpn11, COP9 signalosome subunit 5 and the ESCRT machinery associated AMSH. Most JAMM/MPN+ DUBs cleave Lys63 ubiquitin chains and often display high specificity. AMSH bound Lys63 linked diubiquitin was the first DUB structure with the substrate bound across its active site.

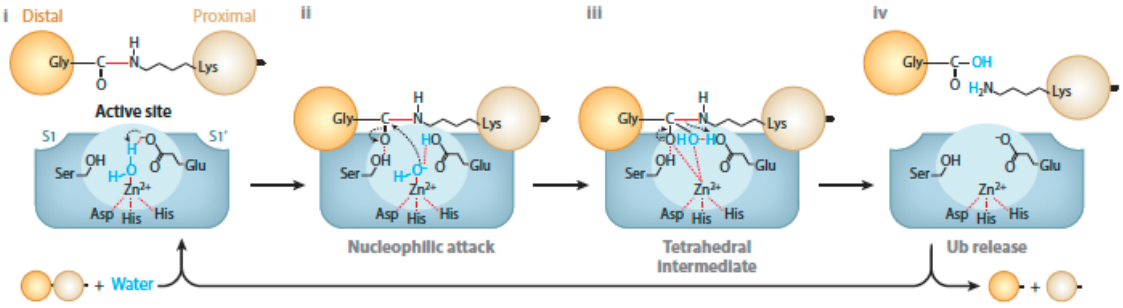


Figure 10: Reaction mechanism of JAMM/MPN+ proteases.

Together, the E1-E2-E3 cascade and deubiquitinases keep ubiquitylation as a dynamic modification, which is kept under tight control, both temporally and spatially.

1.2.5 Specificity determinants

Due to more than eight hundred enzymes and proteins involved in ubiquitin or ubiquitin-like post-translational modifications belonging to the different classes discussed above as well as the thousands of substrates they are believed to act upon, specificity is a major challenge in order to understand how a eukaryotic cell utilizes these post-translational modifications to achieve the required specific function. Starting with Ub/Ubls this problem comes down to, how do they recognize a particular E1 enzyme, then how do the E1 enzymes pair with the specific E2 enzymes, how do E2 enzymes achieve specificity towards their cognate E3s, how do E3s differentiate between their substrates and non-substrates and, last but not the least, how do DUBs identify their substrates and specific ubiquitin linkage types? These are central questions in the efforts targeting ubiquitin and ubiquitin-like modifiers. After decades of research, some of these questions have been answered while some remain to be explored. Some of the well-understood general concepts of specificity are discussed below:

First of all, ubiquitin and ubiquitin-like modifiers have unique signature sequences so that they can be recognized by their specific E1 enzymes and other ubiquitin binding elements. Foremost all ubiquitin like modifiers have unique C-terminal sequences which contain specificity determinant sequences (Bohnsack and Haas, 2003; Whitby et al., 1998). In ubiquitin, Arg72 is the primary specificity determinant, and this residue binds to a negatively charged pocket on the E1 surface, where it engages in a network of hydrogen bonds and ion pairs (Lee and Schindelin, 2008). The corresponding residue in SUMO is Glu93, which interacts with Arg119 and Tyr159 of the SUMO E1 SAE1 subunit, and in NEDD8 Ala72 interacts with Leu206 and Tyr207 of the NEDD8 E1 UBA3 subunit (Lois and Lima, 2005; Walden et al., 2003a). Similarly, each Ubl has its own signature sequence at its C-terminus which allows the E1 enzyme to identify its respective substrate.

Ubiquitin-like modifiers	Carboxy-terminal sequences
Ubiquitin	LRLRGG
ISG15	LRLRGG
NEDD8	LALRGG
SUMO1	QEQTGG
SUMO2/3	QQQTGG
FAT10	CYCIGG
URM1	STLHGG
UFM1	PRDRVG

ATG12	KSQAWG
MAP1LC3-A, -B, -C	SQETPG
GABARAP, -L1	SESVYG
GABARAP,-L2	GENTWG
GABARAP-L3	NESVYG

Table 3: C-terminal sequences for ubiquitin and ubiquitin-like modifiers.

The UFD domain of the canonical E1 enzymes and the UFD-like elements of the non-canonical E1 enzymes play crucial roles in the recruitment of the E2 enzyme by interacting with its N-terminal helix (Wang et al., 2010a). Moroever, other regions on the E2 enzyme have been shown to bind to the catalytic cysteine domain of canonical E1 enzymes. Many E2s have specific extension, either at the N-terminus or the C-terminus of the UBC fold, that act as E1-specificity elements. For example, UBC12 has an N-terminal helical extension which unfolds to interact specifically with the adenylation domain of UBA3 (Huang et al., 2004; Huang et al., 2005).

Due to the larger combinatorial landscape, E2-E3 interactions follow multiple mechanisms to achieve specificity. These include, to name a few, backside binding of E3 or Ubl, E2-RING interactions in case of RING or RBR E3 ligases and E2-N-lobe interactions in case of HECT E3 ligases. Furthermore C-terminal signature sequences and the surface properties of Ubl itself help to match the correct the E2~Ub complex with the right E3 (Berndsen and Wolberger, 2014; Stewart et al., 2016).

Other than the Ubl-E1-E2-E3 members, the proper reception of the tagged substrates is also a critical step for accomplishing ubiquitylation. For this several ubiquitin binding modules have been reported that not only recognize the Ubl but also can distinguish between ubiquitin linkage types (Dikic et al., 2009). Ubiquitin binding modules are present in the enzymes catalyzing ubiquitylation, for example in the Ubc domain of E2s as well as in deubiquitinating enzymes. Moreover, these modules are often seen on ubiquitin receptors, for example in the Endoplasmic Reticulum Associated Degradation (ERAD) or proteasome shuttle factors as well as on the proteasome regulatory particle. Some examples of ubiquitin binding domains (UBDs) are the Ubiquitin Interacting Motif (UIM), the inverted UIM, Ubiquitin Binding Zinc fingers (UBZ) and Pleckstrin Homology Receptors for Ubiquitin domain (PRU). Usually they are 20-50 amino acids in length and bind non-covalently to ubiquitin. Ubl recognition is also well studied for SUMO where the short SUMO Interacting Motif (SIM) has been shown to bind tightly to these modifiers. UBDs that adopt α-helical structure bind to the hydrophobic patch on the central β strand of ubiquitin, whereas zinc fingers can recognize three surfaces of

ubiquitin: 1) The I44 hydrophobic patch, 2) the D58 polar surface or 3) the C-terminal residues. The Pleckstrin Homology (PH) domain of Rpn13, the proteasomal ubiquitin receptor is consists of two antiparallel β sheets which create a platform for binding of the β-strand surface of ubiquitin. Moreover, several ubiquitin receptors can have multiple UBDs to decipher different ubiquitin linkage types for recognition.

On top of the UBDs, ubiquitin receptors and shuttle factors also contain Ubiquitin-like modules (UBLs) which, despite being different from ubiquitin in sequence, share an overall similar fold with ubiquitin. These modules can bind to UBDs through a hydrophobic patch closely related to the I44 patch of ubiquitin. Examples are the recruitment of proteasomal shuttle factors (HR23 and PLIC/Dsk2) to the UIM domains in the proteasomal subunit S5a, the intramolecular interaction between the ULD (Ubiquitin-like domain) and UBA2 domains in HR23 (Rad23 in yeast), as well as the UBA domain mediated recruitment of FAT10 by NUB1L (Schmidtke et al., 2006; Walters et al., 2003). The heterodimeric ERAD shuttling factor UFD1/NPL4 contains the Ubiquitin regulatory X (UBX) domain, another variant of UBLs (Park et al., 2005). ULD in PARKIN keeps the enzyme in autoinhibited state (Buetow and Huang, 2016). Thus, UBDs and UBLs together act as the readers and modulators of the ubiquitin code. A forth class of ubiquitylating enzymes, referred to as E4 enzymes have been shown to maintain, elongate and edit ubiquitin chains to generate specific kinds of homotypic and heterotypic ubiquitin chains (Hoppe, 2005).

Specificity for DUBs is also quite well understood. First of all, there are Ub and Ubl specific proteases that can recognize these modifiers through their surface properties as well as the C-terminal flexible tails (Mevissen and Komander, 2017). As ubiquitin and NEDD8 are quite similar, several DUBs have been reported that can cleave both ubiquitin and NEDD8. Although rather underexplored, ISG15 modifications may also be cleaved by some of these DUBs as it represent diubiquitin and has an identical C-terminal sequence. Moreover, some DUBs act at the end of the ubiquitin chains (Exo DUBs), whereas others bind in the middle of the ubiquitin chain for cleavage (Endo DUBs). To remove monoubiquitin or ubiquitin chains of a certain linkage type on the modified substrates certain DUBs have substrate recognition domains to achieve specificity. Together, writers (E1-E2-E3), readers (UBDs), editors (ULDs, E4) and erasers (DUBs) tightly control the specificity of the ubiquitin system.

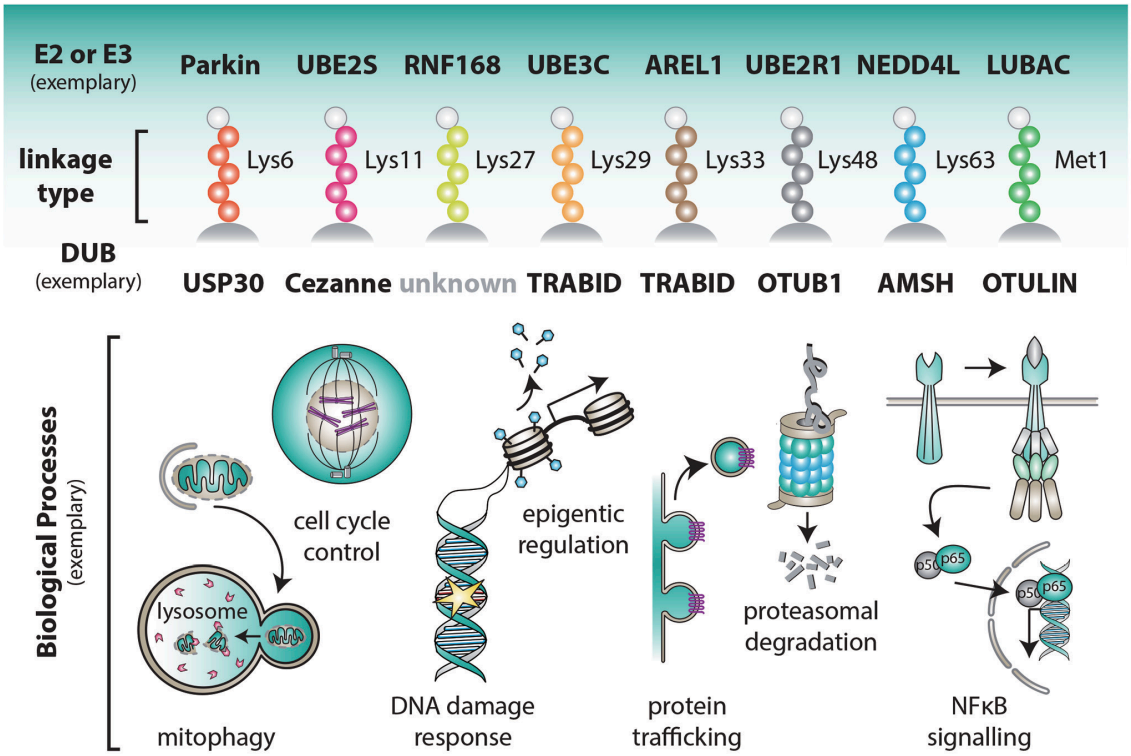


Figure 11: Linkage specific E2/E3 enzymes and their counteracting deubiquitinating enzymes together specify unique signals for various biological functions.

1.2.6 Regulation and cross-talk

As the ubiquitin and ubiquitin-like post-translational modifiers are themselves proteins they can be further subjected to other post-translational modifications in the form of phosphorylation, acetylation, methylation or lipidation. Moreover, ubiquitin chains can be mixed with SUMO2/3 or NEDD8 polypeptides. This way, another layer of tuning these modifications is established (Swatek and Komander, 2016). Ubiquitin contains eleven serine, threonine and tyrosine residues and mass spectrometry studies have shown that each of these residues can be phosphorylated. As mentioned for the activation of PARKIN, Ser65 of ubiquitin undergoes phosphorylation by PTEN-induced protein kinase 1 (PINK1) which is thus linked to mitophagy (the bulk degradation of mitochondria). Ser65 phosphorylation was shown to induce two conformations in ubiquitin that are in dynamic equilibrium. The predominant major conformation resembles the common ubiquitin structure observed more than 300 times in the Protein Data Bank (PDB). The previously unobserved ‘minor’ conformation has the last β-strand and the C-terminal tail of ubiquitin withdrawn into the molecule, altering many surface properties including the Ile44 patch and restricting the C-terminus for chain assembly (Wauer et al., 2015). Other than being ubiquitylated, the lysine residues on ubiquitin can also be acetylated. Mining of available datasets suggest that other than

Lys29, all other lysines on ubiquitin undergo acetylation. These modifications can alter the charge and surface properties of ubiquitin. Lysine acetylation can directly influence ubiquitin linkage by depriving ubiquitin of possible modification sites. Artificially acetylated Lys6 and Lys48 and phospho-Ser65 modifications on ubiquitin do not interfere with E1-mediated ubiquitin charging, however, the discharge of ubiquitin onto substrates was inhibited in a variety of E2/E3 assembly reactions. Importantly, Lys6- and Lys48-AcUb has been identified attached to histones in cells. An 'acetyl-mimetic' ubiquitin mutation (K6Q) was found to stabilize monoubiquitinated histone H2B.

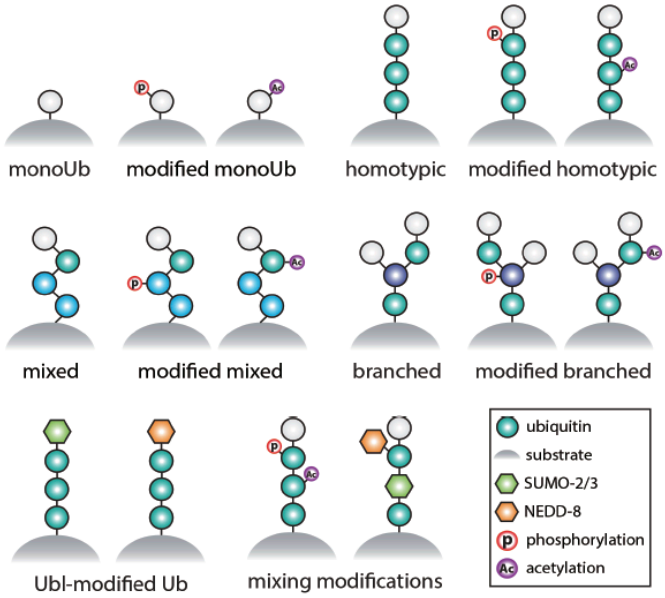


Figure 12: Increasing the complexity: Post-translational modifications on ubiquitin and ubiquitin chains.

SUMO is the only other ubiquitin-like modifier that has been shown to form chains and SUMO-targeted ubiquitin ligases ('StUbLs') target SUMO chains for ubiquitylation. In addition to ubiquitylated SUMO chains, recent findings show that ubiquitin can be SUMOylated. By overexpressing NEDD8, a NEDDylated ubiquitin chain could be observed. Similarly, it is possible that other Ubls like ISG15 can also modify ubiquitin chains. However, the physiological role of such mixed chains is still unclear.

Another mode of regulation of the Ub/Ubl system is observed under oxidative stress. The SUMO E1 and E2 were shown to form a disulfide cross-link via their catalytic cysteine residues upon oxidation, leading to a downregulation of global SUMOylation levels (Bossis and Melchior, 2006; Stankovic-Valentin et al., 2016). Yeast Uba1 and Ubc3 were also shown to form similar cross-links, resulting in stabilization of their substrates (Doris et al., 2012). The reactive cysteines of DUBs are also susceptible to oxidation. Consequently, oxidation by reactive oxygen species (ROS) has been shown to inactivate members of the OUT, USP and UCH families *in vitro* and *in vivo* (Kulathu

et al., 2013; Lee et al., 2013; Silva et al., 2015). Importantly, oxidation events may be reversible and thus upon relief these enzymes can return back to function normally. It is believed that cysteine-based DUBs serve as ROS sensors and that reversible inhibition is crucial for fine-tuning cellular stress responses.

Ub and Ubl modifications do not just come together to produce mixed chains. Instead, sometimes one modification is essential to generate another modification in the substrate. Such is the relationship between NEDDylation and ubiquitylation. NEDDylation of the Cullin subunit of the Cullin-RING ligases is critical for their activity and this modifier induces an active conformation of the Cullin subunit, which is required for the transfer of ubiquitin from the E2 to the target protein (Merlet et al., 2009).

In contrast to cooperating, these post-translational modifiers also compete with each other. The best studied example of this phenomenon is the Proliferating Cell Nuclear Antigen (PCNA) which is required for processive DNA replication as well as DNA repair processes like nucleotide excision repair, homologous recombination and mismatch repair. The repertoire of PCNA function is dramatically amplified by its post-translational modifications (Choe and Moldovan, 2017). K164 on PCNA undergoes both monoubiquitylation and SUMOylation. Because both these modifications occur at the same lysine residue, they are mutually exclusive and, in addition, result in diverging signals. While monoubiquitination on K164 promotes translesion synthesis, a process to bypass DNA lesions during replication by employing low-fidelity DNA polymerases, SUMOylation at this site inhibits homologous recombination. Interestingly, the same lysine residue can also be modified by K63 linked ubiquitin chain, which recruits the translocase ZRANB3. Furthermore, SUMOylation of E2s has been shown to inhibit ubiquitylation.

Thus, the cross-talk between ubiquitin and other ubiquitin-like modifiers results in a wide range of signals needed for a myriad of cellular processes. To render things even more complex, these modifiers can themselves be modified by kinases, acetyltransferases to generate a mind-boggling combination of modifications. The interplay of these post-translational modifications is a highly dynamic, yet specific process and understanding this diversity and its significance is a big challenge in the field.

1.3 UBIQUITIN ACTIVATING ENZYME

After a bird eye's view over the ubiquitylation cascade, I would like to come back to the main topic of this thesis, which focuses on the apex of the ubiquitylation machinery where the E1 enzyme activates ubiquitin and passes it onto its E2 enzymes. There is only one E1 enzyme (UBA1) for ubiquitin in yeast and *C. elegans* and in deuterostomes

(from sea urchin onward) there is again only one “ubiquitin dedicated” E1 enzyme, however, UBA6 which is only present in the organisms can also activate ubiquitin (Jin et al., 2007b). Not surprisingly, being atop of the ubiquitylation cascade UBA1 was one of the first pathway enzymes to be discovered. The human gene for UBA1 is located on the short arm of the X chromosome (Xp11.23) and does not escape X chromosome inactivation (Zacksenhaus and Sheinin, 1990). Another gene for a protein homologous to E1 is located on the Y chromosome of rodents and marsupials, however, it is missing on the human Y chromosome. The respective mouse gene, Sby, exhibits a testis-specific expression and is a candidate for a spermatogenesis gene required for survival and proliferation of spermatogonia (Mitchell et al., 1998).

1.3.1 Discovery

By the year 1980, Hershko et al. had proposed the mechanism of protein breakdown in a soluble ATP-dependent proteolytic system from reticulocytes and had found that it is composed of several essential components (Hershko and Ciechanover, 1982; Hershko et al., 1980). The discovery of the heat-stable polypeptide APF-1 (ATP dependent Proteolysis Factor-1), which was later found out to be ubiquitin, led to the idea that this protein serves as a tag for targeting substrates for proteasomal degradation as it did not possess proteolytic activity by itself but is required for protein breakdown (Ciechanover et al., 1981; Hershko et al., 1980). Ubiquitin had been shown to form isopeptide linkages with substrate proteins in reticulocyte fractions as well as with the histone H2A (Goldknopf and Busch, 1977). Since this process was ATP dependent, Ciechanover et al. imagined this process to be analogous to the activation of free amino acids for amide bond formation by aminoacyl tRNAs or to the nonribosomal synthesis of peptide antibiotics (Ciechanover et al., 1981). Both these processes required an acyl-adenylate intermediate and result in the release of pyrophosphate. Ciechanover et al. could purify the enzyme catalyzing this step and could demonstrate a ubiquitin and Mg^{2+} dependent ATP-PP_i exchange activity (Ciechanover et al., 1982; Haas et al., 1981). Furthermore, the ubiquitin-dependent ATP-AMP exchange reaction indicated that the activated group is transferred to an acceptor with the liberation of AMP resulting in a high energy covalent linkage between a thiol group in the enzyme and ubiquitin (Haas and Rose, 1982; Haas et al., 1982; Haas et al., 1983). This enzyme was later named the ubiquitin activating enzyme or the E1 enzyme for ubiquitin.

1.3.2 Function

UBA1 performs three important catalytic steps in the ubiquitin conjugation system. UBA1 is made up of a single polypeptide containing two active sites, namely the

adenylation active site and the catalytic cysteine domain. First, UBA1 can bind non-covalently to both ubiquitin and Mg-ATP in a pseudo-substrate addition manner (Tokgoz et al., 2006). In the presence of both substrates, UBA1 catalyzes the adenylation of the C-terminus of ubiquitin. The carboxylate group of the C-terminal glycine residue attacks the α-phosphate of ATP, cleaving the α-β phosphoanhydride bond of ATP, resulting in the formation of a mixed acyl-phosphate anhydride between the C-terminal glycine of ubiquitin and AMP (from here on denoted as Ub•AMP). In this reaction pyrophosphate (PP_i) is released. In the second step, the active site cysteine of UBA1 attacks the acyl carbon to release AMP and form a thioester bond. This step leaves the adenylation active site empty and thus a second ubiquitin undergoes adenylation occupying the vacant site. In this way two ubiquitin molecules are accommodated on the ubiquitin activating enzyme where one ubiquitin is bound noncovalently as Ub•AMP adduct and the other ubiquitin is attached covalently at the catalytic cysteine. The third important role of UBA1 is the transfer of thioesterified ubiquitin to the active site cysteine of the E2 enzyme in a transthioesterfication reaction. For this to occur the E2 enzyme has to bind UBA1 and the catalytic cysteine residues of both these enzymes have to be juxtaposed.

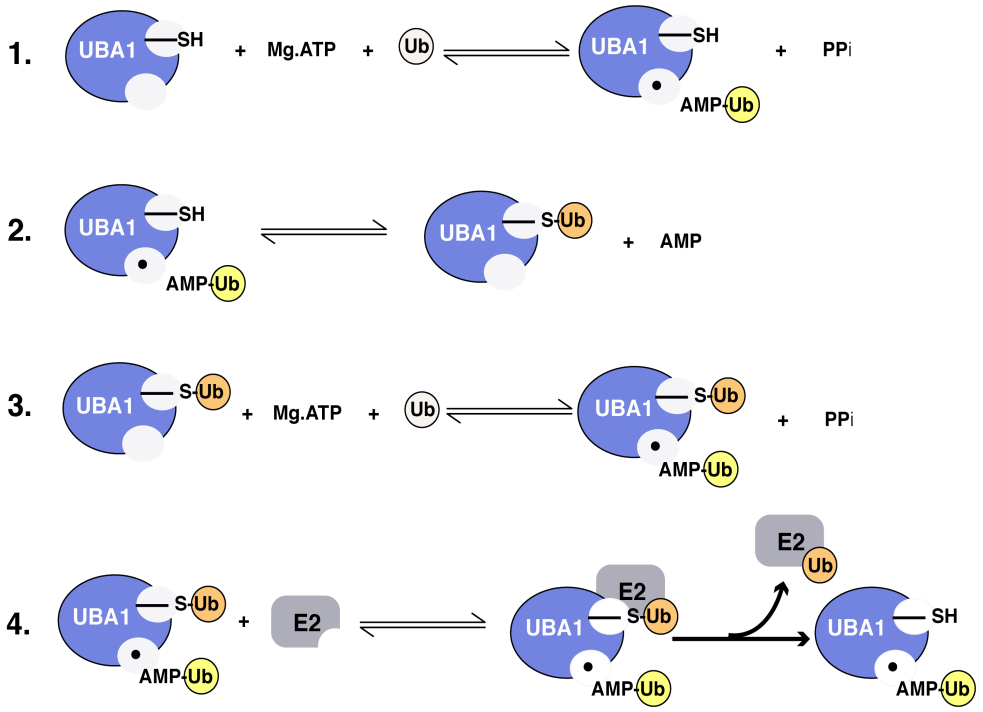


Figure 13: Schematic representation of the enzymatic activity of the ubiquitin activating enzyme.

Although UBA1 is very specific for the activation of ubiquitin, it can activate NEDD8 under stress conditions like heat shock or oxidative stress (Leidecker et al., 2012). Uba1, the homologous enzyme of UBA1 in yeast, plays an essential role in cell

proliferation and sporulation and deletion of yeast *uba1* gene is lethal (Mcgrath et al., 1991). Ubiquitin activating enzyme corresponding to yeast and human organisms will be depicted as Uba1 and UBA1 respectively in the rest of the thesis.

1.3.3 Structural characterization

UBA1 is a 117 kDa modular polypeptide with multiple distinct domains carrying out their specific functions. In addition to the MoeB/ ThiF homology domains that are associated with the adenylation catalytic activity, new domains have been incorporated in the enzyme for specific functions. Despite the presence of two MoeB repeats in UBA1, only the one present in the C-terminal half acts as the functional active site for adenylation compared to two for the MoeB homodimer (Lake et al., 2001). Hence, for UBA1 the N-terminal MoeB homology region, which lacks the ATP binding motif is referred to as the 'inactive' adenylation domain (IAD) and the C-terminal MoeB homology region as the "active" adenylation domain (AAD) (Lee and Schindelin, 2008). Furthermore, UBA1 like SAE2 and UBA3 has large insertions in both the IAD and AAD domain, which have evolved to carry out E1 specific functions. These insertions in the IAD and AAD are referred to as first catalytic cysteine half-domain (FCCH) and second catalytic cysteine half-domain (SCCH), respectively. Interestingly, the FCCH forms a dimer interface in all *Saccharomyces cerevisiae* Uba1 crystal structure, however, in solution Uba1 is present as a monomer (Lee and Schindelin, 2008; Schafer et al., 2014). The SCCH has evolved to carry out the second catalytic half-reaction of the UBA1 and contains the catalytic cysteine residue. In the same region, MoeB possesses the aforementioned 8-residue mobile loop, which also contains a cysteine residue; however, it does not appear to be essential for Moco biosynthesis (Leimkuhler et al., 2001). Uba1 also possesses a short four-helix bundle (4HB) domain which originates from an insertion in the IAD, following the FCCH or equivalent regions in APPBP1 and SAE1 (Lois and Lima, 2005; Walden et al., 2003a). Another E1 specific domain lies at the very C-terminus of UBA1, which is known as the ubiquitin-fold domain (UFD) due to its structural similarity to ubiquitin and other Ubls.

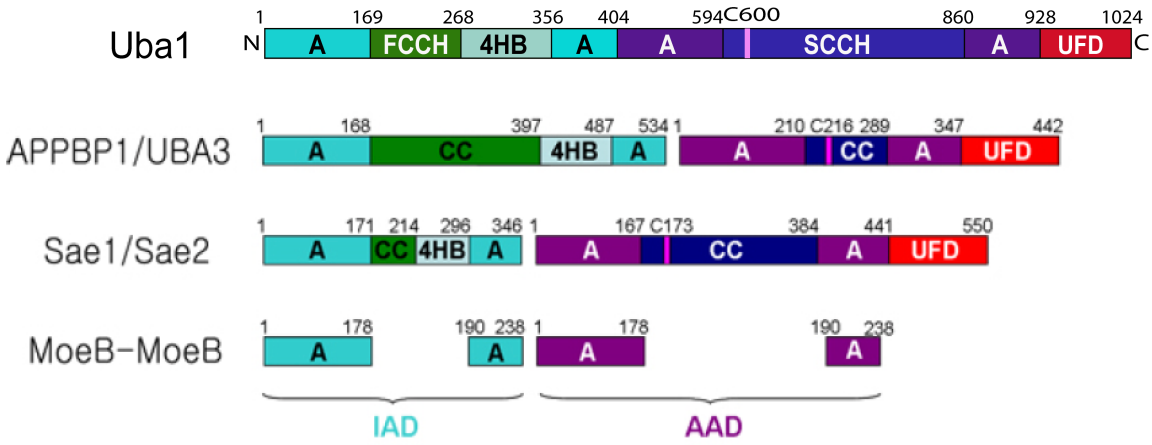


Figure 14: Domain architecture of yeast Uba1, the APPBP1/UBA3 heterodimer (NEDD8 E1), the Sae1/Sae2 heterodimer (SUMO E1) and the prokaryotic MoeB dimer.

Uba1-Ub complex structure:

Although crystal structures of the heterodimeric NEDD8 E1 and SUMO E1 were reported earlier, the first glimpse of the ubiquitin activating enzyme was presented in 2006 by Lee and Schindelin at 2.7 Å resolution, where a Δ9 yeast Uba1 construct was crystallized in the presence of ubiquitin (Lee and Schindelin, 2008). This crystal structure revealed the overall domain architecture of Uba1 and, due to a level of 50% sequence identity, the human UBA1 is assumed to adopt a similar domain structure. This structure presented the relative positions of the six structural domains referred to as IAD, AAD, FCCH, SCCH, 4HB and UFD with overall dimensions of 85 Å x 90 Å x 60 Å. Three Uba1 domains, UFD, FCCH and SCCH, are linked with the respective adjacent domains by flexible linkers. The cysteine catalytic domain (SCCH) is connected to the AAD through a long flexible crossover loop at its N-terminus and a shorter reentry loop at its C-terminus. Moreover, the structure revealed how ubiquitin is noncovalently bound to Uba1 at the AAD. Ubiquitin's globular domain is bound at the opposite end of the highly conserved ATP binding site with its C-terminal flexible tail extending in a tunnel under the crossover loop to reach into the ATP-binding site. The multidomain architecture featuring flexible connecting linkers give the indication that Uba1 can undergo large conformational changes during its catalytic cycle which seems essential as the catalytic cysteine residue is 35 Å apart form the C-terminus of ubiquitin at the AAD. A comparison of the two copies in the asymmetric unit revealed the rotation of ~10° of the UFD domain which is known to recruit E2 enzymes to E1.

Uba1 doubly loaded with ubiquitin:

The second structure of *S. cerevisiae* Uba1 was reported by Schäfer et al., where Uba1 is bound to the products of the first and second steps of catalysis and is ready to bind to the E2 enzyme to transfer the thioesterified ubiquitin to E2 (Schafer et al., 2014). The structure presented for the first time the acyladenylate of ubiquitin bound to Uba1 at the AAD as well as a second ubiquitin covalently bound to the catalytic cysteine at the SCCH domain. Regarding the formation of an E1-Ubl acyladenylate multiple E1 structures exist which contain Mg^{2+}-ATP prior to hydrolysis, but the formation of a Ubl-AMP conjugate bound to an E1 enzyme was never observed before. Furthermore, the second ubiquitin molecule was only visible in one copy of Uba1 of the two present in the asymmetric unit, presumably owing to crystal contacts mediated by a Uba1 symmetry mate which favourably interacts with the thioesterified ubiquitin whereas in the second copy there is not enough space to allow the second ubiquitin to accommodate the space due to crystal contacts.

Uba1-Ubc4-Ub-ATP complex:

Another structure of *Saccharomyces pombe* Uba1 was crystallized by Olsen and Lima in the presence of Mg-ATP, ubiquitin (K6R, K11R, K27R, S28A, K29R, K33R, K48R, S57A, K63R mutant) and Ubc4 (C21S/C107S) where they cross-linked the E1 and E2 via a disulfide bridge using 2,2'-dipyridyldisulfide (Olsen and Lima, 2013). This structure informed about three binding sites of Ubc4 on Uba1. The first binding site arises between the N-terminal helix of E2 and the acidic patch on the UFD domain of Uba1, the second binding interface lies at the hydrophobic patch located near the cysteine cap of SCCH and the helical and loop regions at the C-terminus of the Ubc domain of E2 situated opposite to the N-terminal helix. The third tripartite interaction site is located between the Uba1 crossover loop, ubiquitin and loops surrounding the active site of E2. Another interesting feature highlighted in the Ubc4 bound Uba1 complex is that the UFD domain is in its closed conformation (proximal to the SCCH), further demonstrating the importance of UFD flexibility during catalysis of the ubiquitin activating enzyme.

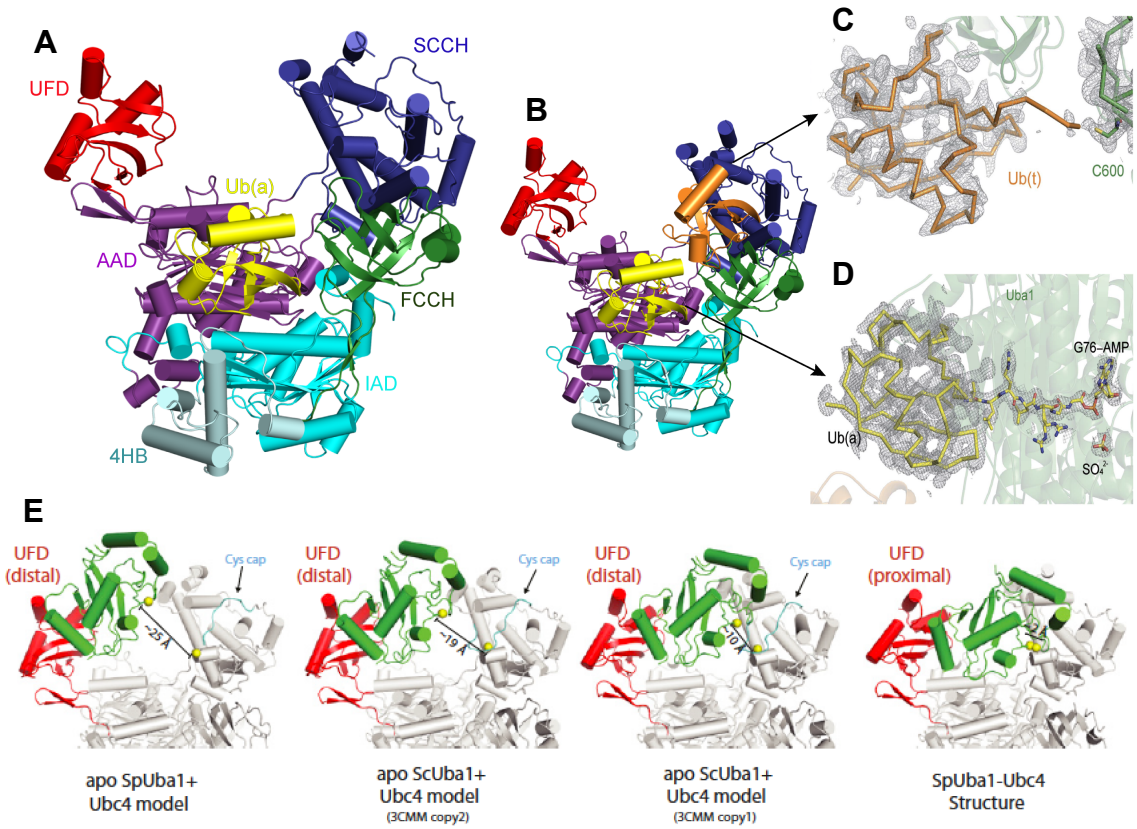


Figure 15: (A) Overall domain architecture of *S. cerevisiae* Uba1 (as colored in figure 14) bound to ubiquitin (yellow) PDB: 3CMM. (B) Doubly loaded Uba1 with thioesterified ubiquitin in orange (panel C) and acyl adenylated ubiquitin in yellow (panel D). (E) Modeled structures of E2 enzyme Ubc4 displaying the recruitment of Ubc4 by UFD in its distal conformations in various crystal structrures as well as the *Sp*Uba1-*Sp*Ubc4 crystal structure where the UFD is in the proximal conformation.

1.3.4 Splice variants and localization

The *uba1* gene comprises 27 exons and contains one alternative splice site. The two isoforms of UBA1 correspond to either 117 kDa (UBA1a) or 110 kDa (UBA1b) (Cook and Chock, 1992). UBA1a and UBA1b both exist in the cytosol and nucleus, however, UBA1a is predominantly nuclear (Cook and Chock, 1991; Sugaya et al., 2015). Approximately, 20% of total cellular E1 is localized in the nucleus, and 80% is localized in the cytosol. However the subcellular localization of UBA1 changes depending on the cell cycle state (Grenfell et al., 1994). During G1 and G2 phase, UBA1 is almost exclusively nuclear. It has been shown that the N-terminal region of the longer isoform contains the nuclear localization signal ($_{5}$PLSKKRR$_{11}$) (Stephen et al., 1995). In another study, the M256I mutation has been reported to cause an instability of UBA1 in the nuclei of CHO-1 cells (Shang et al., 2000).

1.3.5 Tissue Expression

In yeast cells, Uba1 has been accounted to be in the range of around twenty thousand molecules per cell (Ghaemmaghami et al., 2003). In humans, UBA1 is ubiquitously expressed like ubiquitin. The expression profile of UBA1 in human tissue is presented in Figure 16 obtained from the Human Protein Atlas database showing universal presence of UBA1 in all tissue types, however, it is expressed in heart and skeletal muscles and the hippocampus in relatively low amounts. UBA6 is far less abundant in comparison to UBA1 reflecting its role in smaller subset of protein ubiquitylation events (Yang et al., 2013).

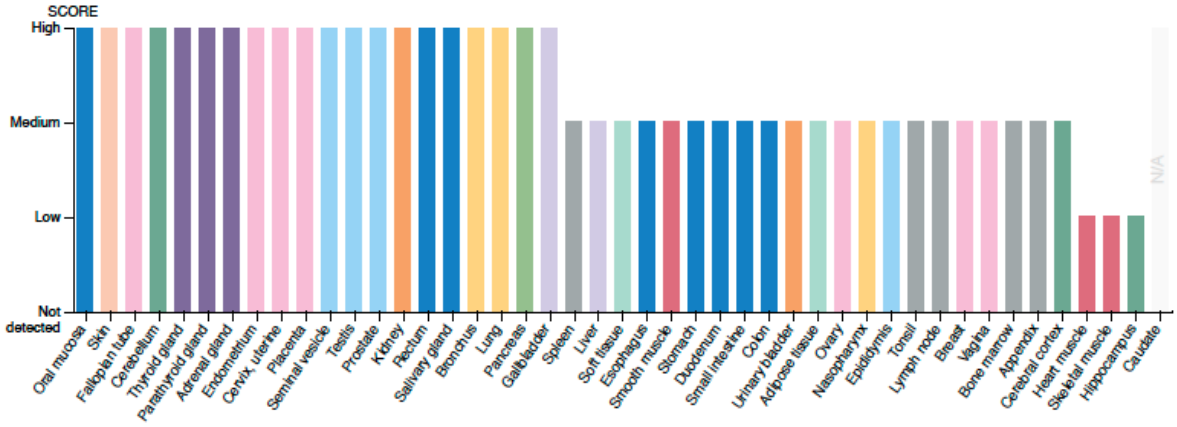


Figure 16: Expression profile of UBA1 based on immunohistological staining methods obtained by the Human Protein Atlas (www.proteinatlas.org).

1.3.6 Post-translational modifications

UBA1 can be phosphorylated at several serine residues, and these modifications are closely linked to the cell cycle status. It has been shown that the longer isoform of UBA1 is phosphorylated at three sites present in the isoform specific N-terminal region during the G2 phase of the cell cycle, which increases its import or retention in the nucleus. Reduced phosphorylation of UBA1a in macrophages was shown to attenuate nucleotide excision repair deficiencies in terminally differentiated macrophages. Table 4 depicts the reported post-translational modification sites for yeast and human UBA1. Interestingly, UBA1 has also been shown to undergo ISGylation as well as FAT10ylation (Giannakopoulos et al., 2005; Rani et al., 2012).

Post-translational modification	Residue	Organism
Ubiquitylation	K28, K53, K381, K408, K494, K561, K595, K608	*Saccharomyces cerevisiae*
	K68, K185, K296, K304, K322, K411, K416, K465, K526, K528, K593, K604, K627, K635, K657, K671, K746, K838, K889, K923, K980	*Homo sapiens*
Phosphorylation	S187, S265, S473, S914	*Saccharomyces cerevisiae*
	S4, S13, S21, S24, S46, Y55, Y560, T800, S810, S816, S820, S835, Y873	*Homo sapiens*
Acetylation	S2, K671, K980, A2 (isoform UBA1b)	*Homo sapiens*
Succinylation	K250	*Saccharomyces cerevisiae*
	K528	*Homo sapiens*

Table 4: Post-translational modification sites on yeast and human UBA1. (Sources: Saccharomyces Genome Database, Uniprot and Phosphosite)

1.3.7 Stability/half-life

The half-life of yeast Uba1 has been reported to be 9.2 hours (Christiano et al., 2014). From 1980 to 1990, many temperature sensitive (ts) mutants of UBA1 were isolated from several cell lines: ts85 of FM3A, ts20 of CHO, ts131b of FM3A, ts20 of Balb/c 3T3, tsBN75 of BHK21, tsFS20 of FM3A and tsFT5 of FM3A (Finley et al., 1984; Ghaboosi and Deshaies, 2007; Tsuji et al., 1990). Figure 17 shows multiple mutations reported on UBA1 responsible for its thermolabile nature in ts cell lines and several disease-associated mutations. The phenotypes of these ts cell lines under non-permissive temperatures established a mutifaceted significance of ubiquitylation for the cell. Other than being responsible for the degradation of cellular proteins, UBA1 is itself degraded by the proteasome system. However, it is still unclear whether the proteasomal degradation of UBA1 is ubiquitin dependent or independent like shown for the degradation of p53, p27 and HIF-1α under certain conditions. Recently, FATylation of UBA1 has been shown to target it for proteasomal degradation (Bialas et al., 2015).

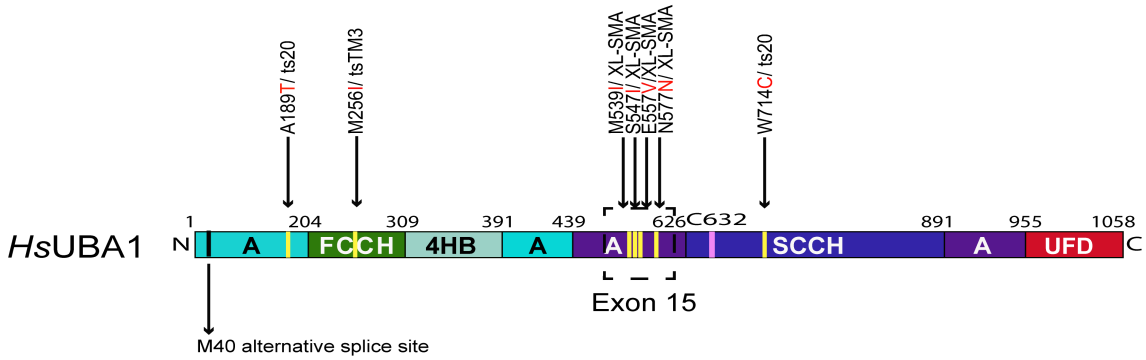


Figure 17: Temperature sensitive (ts) and X-linked Spinal Muscular Atrophy related mutants mapped on the human UBA1 protein. All SMA related mutants are located in the Exon 15 region.

1.3.8 Interaction partners

UBA1 has its own set of E2 enzymes that number to around fifty in humans, however, UBA6 has only one unique E2 enzyme, UBE2Z. In addition to the ubiquitin E2s, UBA1 can also charge UBCH8 the E2 for ISGylation with ubiquitin (Durfee et al., 2008). Due to various post-translational modifications reported for UBA1, several kinases, acetyl-transferases as well as ubiquitin ligases are assumed to interact with UBA1, however, as these interactions are rather transient, they have been hard to capture. Interestingly, in a recent report, UBA1 has been shown to physically interact with the Spinal Motor Neuron (SMN) protein in co-immunoprecipitation experiments (Wishart et al., 2014).

1.3.9 Implication in cancer and neurodegenerative diseases

The degradation of the majority of intracellular proteins takes place either by the proteasome or the lysosomal pathway. Either of these proteolytic pathways has some characteristic features in terms of its substrate preference. The proteasome mainly targets short-lived regulatory proteins, including damaged or partially unfolded proteins. These proteins contain internal degradation signal or degrons that can be a unique motif or a single amino acid residue at the N-terminus of the protein. The lysosomal pathway mainly covers degradation of membrane-bound or organelle-associated proteins. Recent studies suggest that both these pathways work in a coordinated manner and have connecting links between them (Cohen-Kaplan et al., 2016). Ubiquitylation has been shown to play a role in both of these pathways, hence underlining its essential role in the cellular degradative machinery. This impacts the study of cell and developmental biology, in diseases such as cancer, and on the study of protein folding and stability in Alzheimer's disease and other diseases of protein aggregation and misfolding. Being the master regulator of the ubiquitylation machinery, UBA1 critically contributes to maintaining protein homeostasis. Disruption of normal

protein quality control and degradation leads to an accumulation of toxic proteins and neurodegenerative disorders. Temperature-sensitive mutations revealed how loss of UBA1 function leads to an overall reduction in the levels of ubiquitylated proteins and protein degradation, causing cell cycle arrest. Importantly, several lines of evidence suggest that the timing and severity of UBA1 perturbations are likely to dictate the resulting phenotype. For example, studies in Drosophila have shown that partial loss of UBA1 leads to defects in apoptosis, whereas complete loss of UBA1 leads to cell cycle arrest (Lee et al., 2008). In *Caenorhabditis elegans*, loss of UBA1 function at different developmental stages leads to a range of phenotypes including embryonic or larval lethality, decreased fertility in adult stages, and late-onset paralysis (Kulkarni and Smith, 2008). UBA1 has been shown to be associated with Spinal Muscular Atrophy (SMA) and Huntington's disease (HD). Mutations in UBA1 causes a rare form of SMA known as X-linked SMA (XL-SMA), a disease that is clinically similar to SMA but not caused by the homozygous deletion of the SMN1 gene (Groen and Gillingwater, 2015; Jedrzejowska et al., 2015). The clinical characteristics of XL-SMA are muscle weakness associated with anterior horn motor neuron loss, hypotonia and areflexia (Dlamini et al., 2013). In Huntington's disease, decreased expression of UBA1 leads to increased accumulation of the toxic form of the huntingtin protein containing polyglutamine stretches.

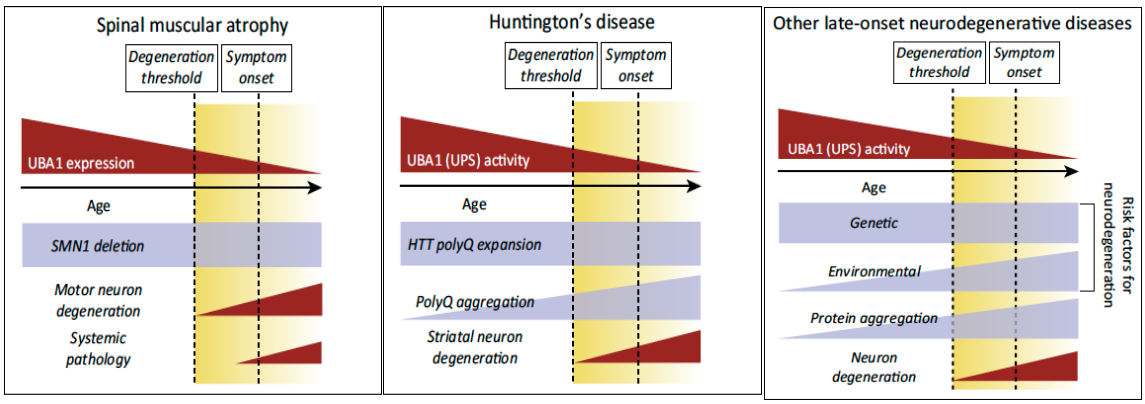


Figure 18: Role of UBA1 in neurodegenerative diseases. Notably, in case of SMA, the pathological symptoms are not the result of an accumulation of aggregated proteins as seen in the case of Huntington's disease or other neurodegenerative disorders.

The reported links between the altered levels or activity of UBA1 with neurodegenerative diseases places UBA1 at the center of a molecular 'hub' capable of modulating neurodegenerative pathways in the nervous system triggered by a diverse range of genetic defects and/or environmental factors. Interestingly, the Ubiquitin

Proteasomal System (UPS) appears to be forming an axis for the see-saw where on one end the downregulation of UBA1 activity leads to neurodegenerative diseases, while, on the other end, upregulation of UBA1 activity is associated with several forms of cancers. However, it is clear that identifying compounds that can modulate UBA1 levels (via transcription, translation or protein stability) or activity is of great interest for overcoming untreatable human health conditions.

1.4 AIM OF WORK

Due to the prevalent existence and significance in the ubiquitylation pathway and its therapeutic potential in the treatment of cancer and neurodegenerative diseases, UBA1 holds a great interest for basic research as well as its translational applications. In terms of basic research, understanding the molecular mechanism of the catalytic cycle of UBA1 will shed light on the first and essential step of the ubiquitin pathway where ubiquitin gets activated by the E1 enzyme in an ATP-dependent manner and is further relayed from the catalytic cysteine of UBA1 to the active site cysteines of the E2 enzymes. For this purpose, structural snapshots of UBA1 at each of the enzymatic steps will help to "connect the dots" and enable visualization of the catalytic cycle of UBA1. After deciphering the enzymatic mechanism of UBA1 at the molecular level, it is intriguing to apply this knowledge to exploit the therapeutic potential of the enzyme by identifying and characterizing modulators of UBA1 enzyme activity. With this motivation, I can categorize the work presented in this thesis into two goals:

I) **To attain structural insights into the catalytic cycle of UBA1 using X-ray crystallography and biochemical tools as discussed in Chapter 2.**

II) **To identify and characterize chemical compounds that modulate UBA1 activity and unravel the structural basis for their mode of action as described in Chapters 3.**

Chapter 2: Structural and biochemical insights into the catalytic cycle of Uba1

2.1 ABSTRACT

Albeit reports of several crystal structures and in depth analysis of enzyme kinetics for UBA1 exist, the complete picture of how the enzyme efficiently achieves its catalytic function is structurally intriguing. The modular architecture of UBA1 shows several domains connected to each other by flexible linkers and as is evident from the crystal structures reported for other E1 enzymes, these domains show conformational flexibility in order to support the catalytic mechanism of UBA1 (Lee and Schindelin, 2008). The C-terminal UFD that is responsible for the recruitment of the E2 enzyme has been shown to adopt closed and open conformations in two different chains located in the same asymmetric unit, thus further substantiating this idea. However, in most of the crystal structures the rest of the protein adopts similar conformations, which is either due to crystallization favoring the protein in these particular states or due to a lack of appropriate substrates inducing alternate conformations that may shed light on the transition states of the enzyme. After the formation of the ubiquitin-acyladenylate the catalytic cysteine residue present in the SCCH domain carries out a nucleophilic attack on the C-terminus of ubiquitin to form a thioester bond. From the structural point of view, this action would necessitate huge conformational changes as the distance between the catalytic cysteine residue and the C-terminus of ubiquitin is approximately 35 Å and simple rigid body domain rotation will result in severe clashes between the SCCH and the AAD. In this chapter, description of how the interdomain communication is achieved upon substrate-binding by describing various unpublished Uba1 structures is presented. Furthermore, by comparison of published ubiquitin and SUMO E1 structures as well as molecular details of key amino acids residues that are required for the enzymatic activity, an attempt is made to understand molecular basis of catalysis for UBA1.

The ubiquitin activating enzyme has three substrates, ubiquitin, ATP and the E2 enzyme. How the binding of these substrates triggers the catalysis and domain rotations is an important question that needs to be answered to entirely understand the enzyme's catalytic properties. Moreover, as E1 enzymes and, in particular, canonical E1 enzymes like the SUMO E1 and NEDD8 E1 feature highly similar domain architecture, such an exploration can further aid in developing a generalized mechanism of the entire E1 enzyme family. UBA7 and UBA6 are the E1 enzymes most closely related to UBA1 and they are present only in higher eukaryotes and activate unique ubiquitin-like modifiers, ISG15 and FAT10, respectively. Both of these enzymes

are monomeric canonical E1 enzymes and, therefore, understanding UBA1 in greater detail will be of importance to comprehend how UBA6 and UBA7 function as enzymes catalyzing similar reactions but with a ubiquitin-like modifier that resembles diubiquitin.

2.2 INTRODUCTION

In the early 1980s, when ubiquitylation was a newly discovered phenomenon, the search for the enzymes responsible for this process led to the discovery of the E1 enzyme that catalyzed the formation of the ubiquitin-adenylate and ubiquitin-thioester (Ciechanover et al., 1981; Haas and Rose, 1982). The reaction scheme and activity of the enzyme were established mostly utilizing the isotope exchange assays where ATP:PP_i/PP_i:ATP and AMP:ATP/ATP:AMP exchange assays defined the two catalytic half reactions. In 1982, Haas et al. showed that both substrate binding and product release are strictly ordered reactions where ATP acts as a leading substrate over ubiquitin while PP_i release precedes AMP release. The kinetic studies have shown that formation of the enzyme-bound ubiquitin adenylate is rapid and the rate-limiting step for the formation of the ternary complex is the transfer of activated ubiquitin to the catalytic cysteine resulting in thioesterified ubiquitin. However, while most of these findings remain true, the ordered substrate addition does not appear to be an essential feature of the enzyme mechanism. In 2006, Tokgöz et al. corrected this notion by experiments performed with the recombinantly purified human UBA1 and suggested it to be a pseudo-ordered substrate addition mechanism where either ATP or ubiquitin bind the enzyme in an independent manner (Tokgoz et al., 2006). As shown vividly by the ubiquitin-bound Uba1 structure by Lee and Schindelin, binding of ATP is not a prerequisite for binding of ubiquitin. Therefore, the binding of either ubiquitin or ATP as a leading substrate depends on the surrounding concentration of each of the substrates as well as their individual binding affinities towards Uba1. The cellular concentration of free ubiquitin and ATP were measured to be approximately 10-20 µM (HEK293 cells) and 2-4 mM, respectively (Gribble et al., 2000; Kaiser et al., 2011). Furthermore, the K_m values reported for ubiquitin and ATP are ~0.5 µM and ~5 µM, respectively, which suggests that both ubiquitin and ATP are present in saturating concentrations (Haas and Rose, 1982; Tokgoz et al., 2006).

The MoeB or ThiF like activity of UBA1 is contained within the active adenylation domain (Duda et al., 2005; Lake et al., 2001; Lehmann et al., 2006). Whereas MoeB or ThiF form homodimers, presenting two adenylation sites in a symmetric fashion, only one of the adenylation repeats in UBA1 serves as an active site. The inactive and active adenylation domain repeats present in Uba1 nevertheless each feature a Rossmann-like fold, with the Active Adenylation Domain (AAD) forming the catalytic center for adenylation and providing binding sites for both ubiquitin and ATP. The AAD consists of eight β-strands, with the strand order 87654123 that form a continuous β-sheet surrounded by eight α-helices. In the N-terminal half of the domain, four parallel β-strands show a variation of the Rossmann fold (Burroughs et al., 2009). Two 3_{10}

helices (η7 and η8) are inserted between the second β-strand (β18) and the fourth α-helix (α15), breaking the continuity of the classical βαβαβ-topology (Figure 19). The first of these 3_{10} helices (η7) contains five residues with the sequence $_{477}$SNLNR$_{481}$ that are highly conserved in the E1 family enzymes and MoeB. The loop between β17 and α14 contains a highly conserved glycine rich motif with the sequence GXGXXG (where X denotes any amino acid), which is reminiscent of the P-loop typically found in ATP and GTP-binding proteins. The C-terminal half of the domain contains an antiparallel β-sheet (β21- β24), which is critical for ubiquitin-binding.

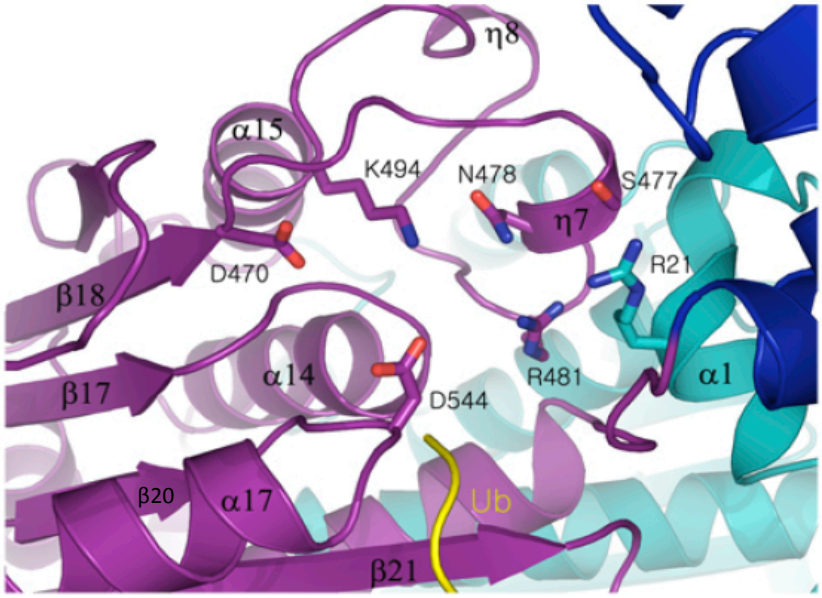


Figure 19: The Active Adenylation Domain (AAD) represented in magenta with the conserved residues for ATP coordination shown in sticks. The arginine finger coming form the N-terminal region, although present in the Inactive Adenylation Domain (IAD), contributes to the ATP binding site of the enzyme. Both AAD and IAD show variations of the Rossmann fold.

The IAD exhibits an identical set of structural features as the AAD (and MoeB), except that it lacks the GXGXXG ATP binding motif. In addition, a four-helix bundle (4HB) formed by residues 269 to 356, packs against the Rossmann-like fold in the IAD, thus blocking access of ubiquitin to this domain (Lee and Schindelin, 2008). However, the IAD contributes the conserved residue R21 to the adenylation active site. A conserved aspartate situated at the loop between β20 and α17 of AAD is involved in coordinating Mg^{2+}, analogous to the aspartate from the Walker B motif of the P-loop NTPase fold. Within the Rossmann fold, E1/E1-like proteins are most closely related to NAD(P)/FAD-dependent dehydrogenases and S-AdoMet-dependent methyltransferases. The phosphotransfer activity of the Rossmanoid domain of E1 enzymes represents an “invention” that occurred independently in an ancestral version of the Rossmann fold

resembling the nucleotide-binding version in NAD(P)/FAD-dependent dehydrogenases and S-AdoMet-dependent methyltransferases (Burroughs et al., 2009).

The prokaryotic E1 like versions (ThiF and MoeB) function as homodimers, whereas eukaryotic canonical E1 proteins function as a heterodimer. This appears to have emerged concomitant with a certain “division of labor” between the two subunits of the dimer. Members of the canonical E1 family, UBA1-N terminal half, SAE1 and APPBP1, are by themselves catalytically inactive but supply the arginine finger to the active site. Conversely, the UBA1-C terminal half, SAE2 and UBA3 lack an arginine finger, but constitute the rest of the active site. The resulting asymmetry in the location of the active site with respect to the dimer interface appears to be critical for positioning the E2 during transthioesterification, as will be described later.

UBA1 appears to have a customized architecture where individual domains are serving a purpose in the catalytic cycle of the enzyme, however, the mechanism of interdomain communication to accomplish enzyme function remains unclear in the lack of structural data. One possible method of achieving such multidomain cooperation is that the binding of the substrates i.e. either ubiquitin, ATP or E2 itself induces conformational changes that trigger catalysis. The formation of the ubiquitin-adenylate, however, does not require drastic conformational changes as both the α-phosphate of ATP and the ubiquitin C-terminus are in close spatial proximity facilitating formation of the first product and pyrophosphate release. However, two important conformational changes are essential to achieve the rest of the catalytic activity of E1 enzymes. The first one is rotation of the SCCH to allow the nucleophilic attack of the catalytic cysteine on the activated ubiquitin species and the second is the UFD domain movement which accommodates the E2 enzymes. Based on the orientation of the catalytic cysteine domain, E1 enzymes have been shown to adopt either open or closed conformations. In the open conformation this domain is distant form the adenylation domain and in the closed state the active site cysteine gains intimacy to the adenylation active site. To find out how either ubiquitin or ATP binding regulate these movements will only be possible when we have access to the whole array of structural snapshots that represent each step of the catalytic cycle, from individual substrate-bound structures as well as product-bound states. Placement of several crystal structures of Uba1 representing different steps of the catalytic cycle in a way that correlate with the reported step by step mechanism of the enzyme, can enable us imagine how the enzyme works. The developed hypothesis for enzymatic mechanism can be further conformed by biochemical methods. This approach has been followed in this chapter to successfully determine several key features of the catalytic cycle of UBA1.

2.3 MATERIALS AND METHODS

2.3.1 Plasmids and constructs

For the expression of N-terminally His_6-tagged Uba1 from *S. cerevisiae* the previously described pET28a-*UBA1* construct encoding for residues 10-1024 was used (Lee and Schindelin, 2008). This gene was generated using the NheI and EcoRI restriction sites of the vector. The coding sequence for *S. cerevisiae* ubiquitin was introduced into pET30a enabling expression without any fusion tag. All thirteen known Ubcs from *S. cerevisiae* were cloned into pETM11 (EMBL, Heidelberg, Germany). The pET30a vector encoding human ubiquitin was a generous gift from Dr. Sonja Lorenz (Rudolf Virchow Center of Experimental Biomedicine, University of Würzburg, Germany).

2.3.2 Mutants

UBA1 mutants were generated in the pET28a-UBA1 construct using site-directed mutagenesis. The list of mutants is provided in Table 5:

	Mutant	**Vector**	**Antibiotic resistance**
1.	D472A	Δ9 *Sc*Uba1-pET 28a	Kan^R
2.	D544A	Δ9 *Sc*Uba1-pET 28a	Kan^R
3.	D474/ 544A	Δ9 *Sc*Uba1-pET 28a	Kan^R
4.	R21A	Δ9 *Sc*Uba1-pET 28a	Kan^R
5.	R481A	Δ9 *Sc*Uba1-pET 28a	Kan^R
6.	R21/ 481A	Δ9 *Sc*Uba1-pET 28a	Kan^R
7.	D472E	Δ9 *Sc*Uba1-pET 28a	Kan^R
8.	D544E	Δ9 *Sc*Uba1-pET 28a	Kan^R
9.	K494E	Δ9 *Sc*Uba1-pET 28a	Kan^R
10.	C600A	Δ9 *Sc*Uba1-pET 28a	Kan^R
11.	D824A	Δ9 *Sc*Uba1-pET 28a	Kan^R
12.	D824R	Δ9 *Sc*Uba1-pET 28a	Kan^R
13.	K494A	Δ9 *Sc*Uba1-pET 28a	Kan^R
14.	Y586A	Δ9 *Sc*Uba1-pET 28a	Kan^R
15.	R861A	Δ9 *Sc*Uba1-pET 28a	Kan^R
16.	E594A	Δ9 *Sc*Uba1-pET 28a	Kan^R
17.	D544C	Δ9 *Sc*Uba1-pET 28a	Kan^R
18.	C600S	Δ9 *Sc*Uba1-pET 28a	Kan^R
19.	C600K	Δ9 *Sc*Uba1-pET 28a	Kan^R

20.	E910A	Δ9 *Sc*Uba1-pET 28a	Kan^R
21.	E910R	Δ9 *Sc*Uba1-pET 28a	Kan^R
22.	C87A	*Sc*Ubc13-pET M11	Kan^R
23.	C97A	*Sc*Ubc3-pET M11	Kan^R

Table 5: Summary of Uba1 mutants (1-21) and active site mutants of Ubc3 and Ubc13 (22 and 23).

For each desired mutation, 50 ng of template DNA was mixed with forward and reverse primers (0.5 µM each) carrying the desired variation of the sequence individually in two reactions. Additionally, 5 µl of 10x reaction buffer and 1 µl of dNTP mix were added and the total volume of each reaction was adjusted to 24 µl with RNAse free ddH_2O. Finally, 1 µl of Phusion DNA polymerase (New England Biolabs) was pipetted into the reaction. This mixture was subjected to 8 rounds of the polymerase chain reaction (PCR) using the following cycling parameters: 95°C denaturation for 30 seconds, 55 °C annealing for 1 minute, and 68°C elongation for 9 minutes. Thereafter, both reactions were mixed and the PCR was repeated for another 35 cycles with the same parameters. Next, 2 µl of the DpnI enzyme (10 U/ µl) were added to the reaction mixture and incubated at 37 °C overnight to digest the parental DNA template. Finally, the resultant mixture was transformed into 100 µl of chemically competent DH5α cells and plated onto LB agar plates containing the appropriate antibiotic. DNA sequences from isolated plasmids of the resulting colonies were verified by automated DNA sequencing.

The ATP-binding mutants will be discussed in section 2.3.3. Three mutations of the catalytic cysteine (C600) of yeast Uba1 were generated including the catalytically inactive C600A mutation discussed in the results section 2.3.3. The C600S and C600K mutants were designed to purify Uba1 that is capable of forming either an oxyester or isopeptide bond with ubiquitin, respectively. Mutations in the crossover loop were created in the form of the Y586A and E594A variants. In addition an Ala mutation of R861 was generated that coordinates with E594 in the crossover loop in the ubiquitin-bound state of the enzyme. The E910 mutations (E910A and E910R) were produced to test the role of this residue in the UFD motion upon binding of ubiquitin at the adenylation site as proposed in section 2.3.8. The catalytic cysteine to alanine mutations of both Ubc13 and Ubc3 were used to probe their contribution during E1-E2 complex formation. The D544C variant was generated to test whether the catalytic cysteine can be captured via a disulfide bridge with this mutant to lock the enzyme in the closed state.

2.3.3 Protein expression and purification

Ubiquitin and all His_6-Ubcs from yeast and all mutated versions were expressed in *E. coli* BL21(DE3) or BL21(DE3) RIL cells (Novagen) by induction with 0.1 mM IPTG at an OD_{600} of 0.6-1 followed by overnight growth at 16°C. For expression of yeast His_6-TEV-Uba1 *E. coli* BL21(DE3)-RIL cells were induced with 0.1 mM IPTG in TB media (containing 0.5% glucose) at an OD_{600} of 1 followed by overnight growth at 25°C.

An average amount of 32 l culture was used for Uba1 expression. All procedures for the purification of proteins were performed at 4 °C. The cells were harvested by centrifugation (20 min at 12,000 x g) and then flash frozen in liquid nitrogen followed by storage at -80 °C. On the day of purification, the cells were thawed in cold water and then resuspended in buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 5 mM β-mercaptoethanol, 25 mM imidazole, complete protease inhibitor (EDTA free, Roche), 5% glycerol). Cell walls were ruptured by passing the lysate twice through a microfluidizer (M-110P Microfluidics) at a pressure of 1.3 MPa and the lysate was further centrifuged for one hour at 75,000 x g to remove cell debris. The supernatant was loaded onto a column containing 25 ml of Ni-NTA beads (Qiagen). The column was thoroughly washed by the addition of 100 ml wash buffer (50 mM TRIS, pH 7.5, 500 mM NaCl, 25 mM imidazole) prior to elution with 250 mM imidazole in the same buffer. The eluted fractions were checked on 12% SDS-PAGE for protein identity and quality. The fractions containing the protein were pooled and solid ammonium sulfate was slowly added under gentle stirring conditions up to a final concentration of 1 M was reached. The resulting solution was centrifuged (30 min at 4000 xg) and the supernatant was loaded onto a HiLoad phenyl sepharose column (GE, Healthcare) equilibrated in 50 mM Tris, pH 7.5, 1 M ammonium sulfate and 10 mM β-mercaptoethanol prior to elution with a descending linear gradient to 100 mM ammonium sulfate in the same buffer. The pooled fractions from the hydrophobic interaction chromatography were concentrated to 5 ml using a Centricon plus-20 (50 kDa cutoff; Millipore) concentrator and then loaded on a HiLoad Superdex 26/60 200pg column (GE, Healthcare) equilibrated with 25 mM Tris, pH 7.5, 200 mM NaCl and 5 mM β-mercaptoethanol). The fractions containing the desired protein were concentrated to approximately 20 mg/ml as determined with a Nanodrop ND 1000 (Peqlab) spectrophotometer using a calculated extinction coefficient of 70,600 $M^{-1}cm^{-1}$ at 280 nm. 20 µl protein aliquots were prepared and flash-frozen in liquid nitrogen and stored at -80 °C. The yield of the protein was typically 1-2 mg per liter of culture.

Yeast E2 proteins were purified by a two-step purification including Ni-NTA affinity chromatography followed by size exclusion chromatography and stored in buffer

containing 25 mM Tris, pH 7.5, 150 mM NaCl and 5 mM β-mercaptoethanol. A similar procedure was followed for the purification of the mutants.

To purify tag free ubiquitin, acid precipitation was followed where after lysis and centrifugation steps, the supernatant was subjected to slow addition of 60% perchloric acid (0.4 ml/ 50 ml lysate) while stirring on ice. The resulting solution was centrifuged (75,000 x g for 30 min) and the supernatant was dialyzed with 50 mM ammonium acetate buffer, pH 4.5, overnight. After filtration of the dialyzed solution, cation exchange chromatography was performed using a HiTrap SP-HP column (GE, Healthcare). A gradient of 0-0.5 M NaCl was used for this step. The elution of ubiquitin was checked by 18% SDS-PAGE. The pH of the pooled protein fractions was adjusted to 7.5 and then concentrated to 2 ml for size exclusion chromatography. A HiLoad 16/60 Superdex 75pg column (GE, Healthcare) was used for this step and the protein was eluted with 25 mM TRIS, pH 7.5, 200 mM NaCl buffer. The protein was flash frozen in liquid nitrogen and stored at -80 °C in this buffer. A typical yield of tag-free ubiquitin was around 40 mg per liter culture.

2.3.4 Circular dichroism spectroscopy

To confirm the correct folding of the mutants circular dichroism (CD) spectroscopy was used. These experiments were conducted with a JASCO J-810 spectropolarimeter at room temperature. Far UV spectra from 260 nm to 190 nm were recorded at a scanning speed of 20 nm/min with a response time of 1 s and a bandwidth of 2 nm. Certain ions and reducing agent present in the buffer impeded the collection of data at wavelength below 200 nm, therefore the buffer of the protein samples was exchanged to 50 mM potassium phosphate with the aid of ultrafiltration units (Sartorius Vivaspin 500). Ten spectra for each sample were accumulated to optimize the signal to noise ratio. For analysis, the buffer spectrum was subtracted as reference from the protein spectra.

2.3.5 Crystallization and data collection

The crystals were obtained using the liquid handling robot Honeybee 963 (Genomic Solutions) in sitting drop vapor diffusion experiments with drops containing 0.5 µl of protein solution mixed with 0.5 µl of mother liquor and a reservoir of 40 µl mother liquor in a 96-well crystallization plate sealed with adhesive sealing film with the help of a Roboseal unit (HJ-BIOANALYTIC). All crystals were grown at 20 °C. The crystallization conditions for ubiquitin, Mg-ATP or E2-bound Uba1 crystals are summarized in Table 6.

Complex	Crystallization condition	Space group
Uba1-Mg-ATP	0.2 M ammonium sulfate, 0.1 M HEPES pH 7.5, 25% PEG 3350	C 2 2 2_1
C600A Uba1-Mg-ATP	0.2 M ammonium sulfate, 0.1 M HEPES pH 7.5, 25% PEG 3350, 5 mM VLRLRGG peptide	C 2 2 2_1
Uba1-Ub	0.2 M ammonium formate 20% PEG 3350	P 2 2_1 2_1
Uba1-Ubc13 (1:1)	0.8 M ammonium nitrate, 0.1 M CHES pH 9.5, 18% PEG 4000	C 1 2 1
Uba1-Ubc13 (1:2)	0.2 M ammonium sulfate, 0.1 M HEPES pH 7.5, 25 % PEG 3350	C 1 2 1
Uba1-Mg-ATP	0.2 M ammonium nitrate, 0.1 M HEPES pH 7.5, 25 % peg 3350	P 1 2_1 1

Table 6: Crystallization conditions for various complexes and their indexed space groups.

The concentration of His-tagged *Sc*Uba1 used in the protein drop of the crystallization set-up varied from 12 mg/ml to 15 mg/ ml. For the Mg-ATP bound ATP and $MgCl_2$ were present at final concentrations of 2.5 mM and 5 mM $MgCl_2$, respectively, with a final Uba1 concentration of 12 mg/ml. The same concentration of Uba1 was used to get complex crystals of ubiquitin-bound Uba1 where ubiquitin was added in a 1:2 molar ratio. For the Uba1-Ubc13 crystals either a 1:1 or a 1:2 molar ratio for His_6-Uba1: His_6-Ubc13 was utilized in the presence of 2.5 mM ATP and 5 mM $MgCl_2$. The final Uba1 concentration was kept at 15 mg/ml for these crystals. In addition, 1 mM H_2O_2 was added for the 1:1 mixture to generate oxidizing conditions.

All crystals were cryoprotected by soaking them in mother liquor supplemented with 15 %- 20% glycerol before being flash-frozen in liquid nitrogen. Beamlines used at either the BESSY (HZB, Berlin, Germany) or ESRF (Grenoble, France) synchrotrons are summarized in Table 7. All data collections were performed at 100 K.

2.3.6 Structure determination and refinement

The collected data were indexed and integrated using XDS (Kabsch, 2010). For scaling and merging Aimless from the CCP4 suite was used. For Uba1-Ubc13 complexes the anisotropic diffraction was corrected by subjecting the dataset to the UCLA MBI anisotropy server (Strong et al., 2006). For molecular replacement PhaserMR (McCoy, 2007) from the CCP4 suite was used (Winn et al., 2011). As search models, Uba1 and ubiquitin from PDB entry 4NNJ and for yeast Ubc13 PDB entry 1JBB was utilized. For initial rigid body refinements refmac5 (Murshudov et al., 1997) from the CCP4 suite was used. Structural refinement employing NCS and reference-model restraints was carried out using PHENIX version 1.8-1092 in combination with Coot (Adams et al., 2010; Emsley et al., 2010). For the graphical representation of protein structures the PyMOL software (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) was used. Protein-protein interfaces were calculated with the help of PDBePISA (Krissinel and Henrick, 2007).

2.3.7 BIOMOL Green Assay

BIOMOL Green reagent (Enzo Life Sciences) provides a simple and convenient method for a colorimetric phosphate quantification. Unlike other molybdate/ malachite green-based assays, it does not require multiple solutions or reagents prepared fresh on the day of the assay. To measure the relative activity of the mutants against the wild type as mentioned in section 2.3.3, a step by step procedure of performing the assay is shown in Figure 20. After initial optimizations, I was able to determine the right concentration of ubiquitin and enzyme units of inorganic pyrophosphatase (SIGMA, I5907) that were kept constant for all measurements at 80 µM and 40 U, respectively. The molar concentration of wild type and mutant Uba1 proteins was varied between 0 µM to 6 µM and eight different concentrations were used for each assay. The assay was initiated by mixing Uba1 (wt or mutant), ubiquitin and pyrophosphatase in the assay buffer (20 mM Tris, pH 7.5, 50 mM NaCl) and the reaction was started by the addition of ATP and $MgCl_2$ at final concentrations of 1 mM and 2 mM, respectively. The catalytic reaction was allowed to proceed at 30 °C for half an hour leading to the production of pyrophosphate. The released pyrophosphate is cleaved into phosphate by

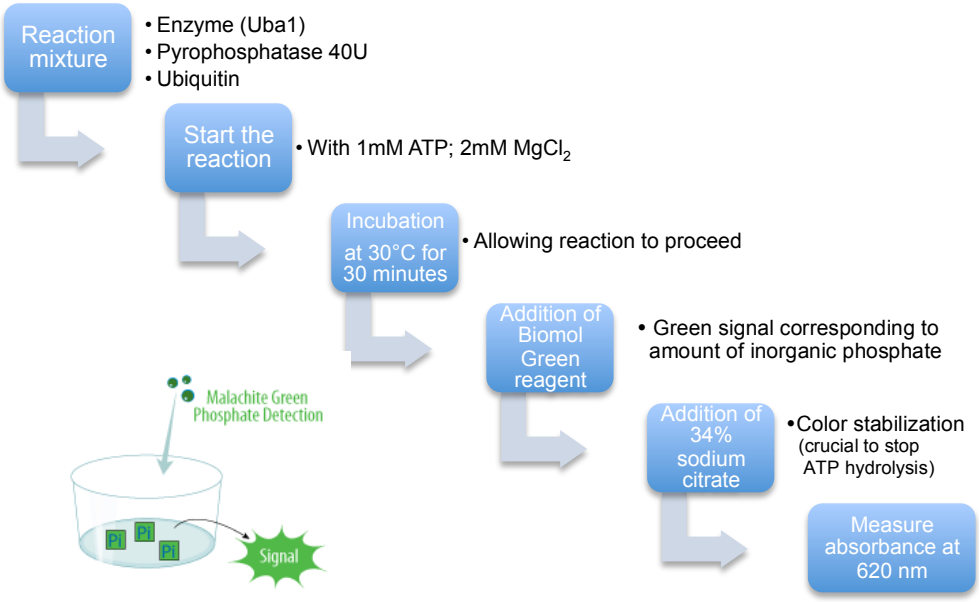


Figure 20: Workflow of the BIOMOL green assay.

pyrophosphatase. After 30 minutes, 100 µl BIOMOL Green reagent were added to 50 µl of reaction. The solution turns green due to the reaction of the Biomol green reagent with phosphate. 5 minutes after addition of the reagent, 15 µl of 34% sodium citrate (w/v) was added (Enzo life sciences, AK111-Addendum datasheet for ATP usage) to the reaction to buffer the solution and stop hydrolysis of ATP under acidic conditions. All reactions were carried out in 96-NUNC microwell plates (Thermo Fischer Scientific) and after addition of sodium citrate the absorbance measurement was performed at 620 nm using a CLARIO star plate reader (BMG, Labtech, Germany). All reactions were performed in triplicates and the data were plotted with absorbance on the y-axis and enzyme concentration on the x-axis. The slope of the fitted curve showed a linear correlation and reflected the specific activity of either the wild-type or mutant. The mean and standard deviations were derived from triplicate measurements.

For testing the possibility of ATP hydrolysis during Uba1-Ubc13 complex formation, 3.5 µM Uba1, 20 µM ubiquitin and 8 µM Ubc13, each in assay buffer, were incubated in various combinations as indicated in Figure 38B in the presence of 2 mM ATP and 4 mM $MgCl_2$ for three hours at room temperature. After this time, 25 U of inorganic pyrophosphatase were added for 30 min to each mixture. Prior to addition of the BIOMOL green reagent, the proteins were separated from the samples using a 10K Amicon® Ultra centrifugal filter (Merck/Millipore, Cork, Ireland). 100 µl of BIOMOL green reagent was pipetted into 50 ul of each sample and after five minutes 15 µl of 34% sodium citrate were added. After a one hour incubation, the absorbance was measured

as described before. Each sample was measured in duplicates in five different experiments. The data were normalized against a control containing 2 mM ATP and 25 U pyrophosphatase in assay buffer.

2.3.8 SDS-PAGE analysis of E1-E2 complex formation

The analysis of E1-E2 complex formation was carried using 4-20% gradient gels purchased from BioRad. The reactions were prepared as indicated by the labels of the individual lanes. The final concentrations of wtUba1/mutant Uba1 and E2 were kept at 10 µM. For ubiquitin, a two-fold molar concentration of Uba1 was used. The concentration of nucleotides (ATP/GTP/CTP/TTP/ATPαS) was constant at a final value of 2.5 mM, and 5 mM $MgCl_2$ was present whenever nucleotides were added in the reaction. The final reaction volume was adjusted to 15 µl with the help of assay buffer (50 mM Tris, pH 7.5, 50 mM NaCl). After overnight incubation at 4 °C, 5 µl of samples were loaded on the SDS-PAGE. The gels were stained with Coomassie brilliant blue G-250 via a standard protocol. After proper destaining, the gels were scanned using an Odyssey scanner (LICOR). The band intensities corresponding to the E1-E2 complexes were quantified for duplicate experiments and mean value were derived using those quantifications. The strength of complex formation for individual E1-E2 pair was determined from the mean values and is depicted qualitatively (Table 8).

2.3.9 Contributions to the work

The *Ct*Uba1-Ubc4 complex and the Uba1-Ubc13 equimolar complex were crystallized by Dr. Antje Schäfer. Monika Kuhn had carried out the SDS-PAGE analysis of E1-E2 complex formation.

2.4 RESULTS AND DISCUSSION

2.4.1 Uba1 purification optimization

Overexpression of large proteins in bacterial hosts often results in low expression levels. Not surprisingly, this is the case for Uba1 as well. Moreover, Uba1 and other E1 proteins have been shown to be prone to degradation, which results in a loss of the active full-length protein. To improve the protein yields for crystallization experiments as well as biochemical assays, I tested several conditions and checked the expression results on 12% SDS-PAGE gels. The Δ9 *Sc*Uba1 construct cloned in pET28a vector with N-terminal His tag and TEV cleavage site between the His tag and the construct was used to further optimize the expression. This construct will be referred to as Uba1 during the rest of the thesis. As the biomass extracted from the bacterial expression cultures, reflected lower yields of cell pellets compared to other well expressing proteins, it gave rise to the notion that higher Uba1 expression might be toxic for the cells as reflected by a lower number of cells after a given expression time. To evaluate this hypothesis, I tested four variables for the expression conditions: 1) Expression host strains, 2) IPTG concentration, 3) induction temperature and 4) culture media ingredients.

Expression host strains:

The previously reported expression of Uba1 utilized the *E. coli* BL21 RIL expression host. This expression host has the advantage of higher abundances of tRNA for the rare codons arginine, isoleucine and leucine. We used this expression host as control. Additionally, we tested two other *E. coli* strains, Rosetta 2 pLys and Rosetta blue lacI, to check whether these strains can overcome the toxic effects of Uba1 expression. Rosetta 2 pLys cells have three special properties: The deletion of lac permease (lacY) enabling the uniform entry of IPTG into cells, additional tRNA for the rare codons arginine, glycine, isoleucine, leucine and proline and expression of T7-lysozyme at low level inhibiting the basal levels of T7-RNAP, thereby reducing the leaky expression of the gene of interest. Rosetta blue lacI strain carries a gene for the expression of the lacI gene, which regulates the lac operon more tightly. The addition of 0.5% glucose in the media was also tested to keep the lac operon inactive and avoid leaky expression of Uba1. Figure 21A shows that, although there were fewer impurities present when Rosetta blue placI cells were used, the overall expression was not increased compared to the BL21 RIL cells used as control. However, the optical density (O.D.) of the cells was found to be higher in the culture where 0.5% glucose was added. Therefore, glucose was added in future expressions.

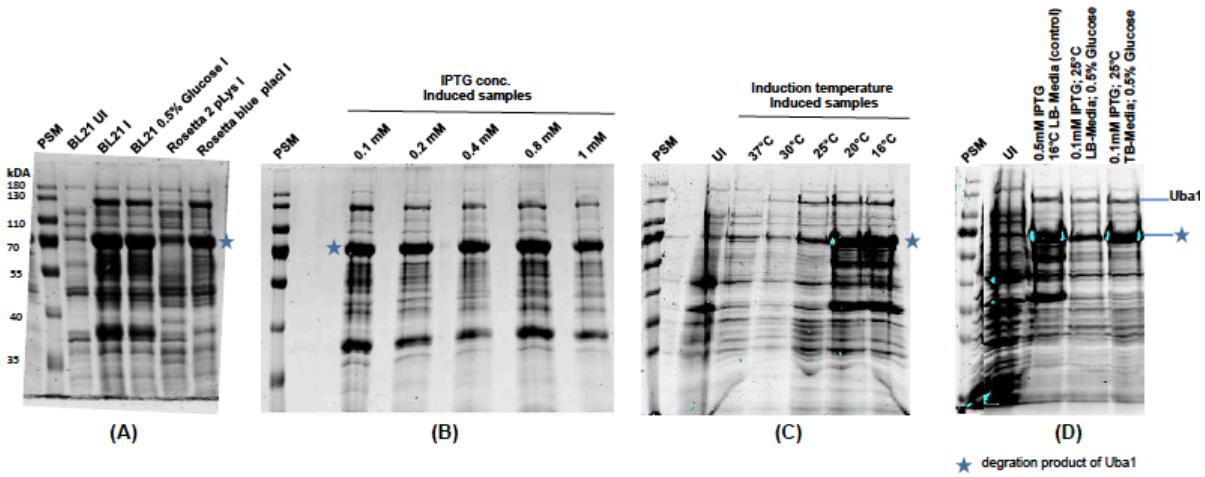


Figure 21: Optimization of Uba1 expression by testing (A) E. coli expression host strains, (B) IPTG concentrations, (C) induction temperatures and (D) culture media

IPTG concentrations:

Finding the optimal IPTG concentration to get the best expression as well as avoiding the toxic effects of overexpression was the second test in the course of expression optimization. 0.1 mM, 0.2 mM, 0.4 mM, 0.8 mM and 1 mM final IPTG concentrations were tested to induce the expression of Uba1, however, higher concentrations of IPTG did not appear to express higher amounts of Uba1. Therefore, 0.1 mM IPTG concentration was chosen for all future expression experiments.

Temperature variation during induction:

With induction temperature being a critical factor during expression it was checked next and 37 °C, 30 °C, 25 °C, 20 °C and 16 °C were tested. As shown in panel C of Figure 21, induction at 25 °C showed the least amount of degradation of the protein (indicated by the dominant band around 70 kDa) while producing similar expression levels as seen for 20 °C and 16 °C. In addition, cells grew to a much higher O.D. at this temperature compared to other temperatures mentioned. Therefore, we changed the induction temperature to 25 °C for Uba1 expression.

Culture media:

We further checked the effect of using either Luria Broth (LB) or Terrific Broth (TB) as culture media to find where the cells grew to higher O.D. as well as the quality of expressed protein. Although the use of LB media ensured less degradation, the cells grew to lower amounts in these cultures. 25 °C again could be seen as a better induction temperature compared to the control. Even though the degradation of Uba1 was found to be slightly higher when using TB media, the cells could be grown to much

higher O.D. We chose to use TB media with 0.5% glucose and an IPTG concentration of 0.1 mM for induction and 25 °C for induction temperature to purify Uba1 in subsequent experiments.

2.4.2 Uba1 purification

A typical purification of yeast Uba1 as described in the materials and methods section 2.5.3 is presented in Figure 22. The degradation product of Uba1 seen around 70 kDa in the (Figure 22A) could be largely separated with the help of hydrophobic interaction chromatography (Figure 22B). Mass-spectrometry confirmed degradation of the C-terminus of the protein (data not shown).

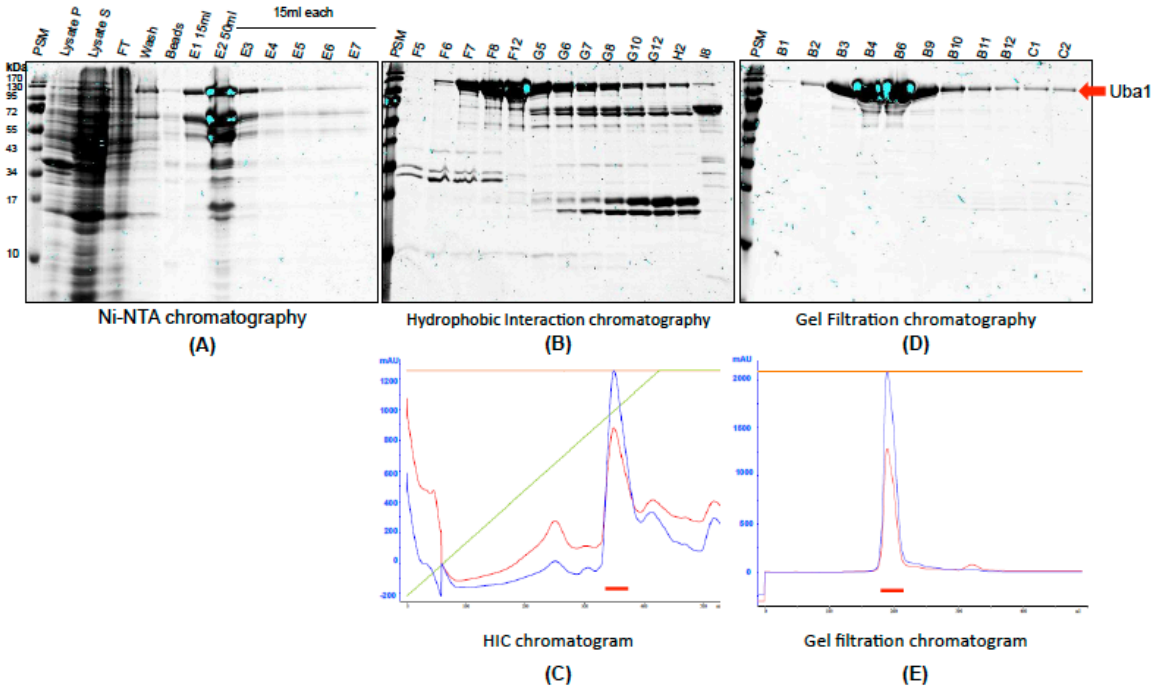


Figure 22: Three-step purification of yeast His_6-TEV-Δ9 *Sc*Uba1. The corresponding 12% SDS-PAGE gels for each step of purification are shown (A-C). The chromatogram for (C) hydrophobic interaction chromatopgraphy (HIC) and (E) size exclusion chromatpgraphy are presented. The significant peaks correspond to protein of interest. The red and blue colored peaks represent absorbance at 230 nm and 280 nm respectively.

2.4.3 Crystallization and structure determination:

Freshly purified Uba1 protein was used for crystallization set-ups as described in section 2.3.5. The crystals appeared between 3-4 days of crystallization (Figure 23). Whereas Uba1-Ub crystals were stable for several weeks, Uba1-ATP/ Uba1-Ubc13 crystals started degrading after 2-3 days of appearance.

(A) (B) (C)

Figure 23: Crystals of the complexes of yeast Uba1. (A) Mg-ATP-bound Uba1 crystals. (B) Uba1-Ub complex crystals. (C) Uba1-Ubc13 crystals obtained in the presence of Mg-ATP.

The space group for each crystallized complex is shown in Table 7. ATP-bound crystals belonged to the orthorhombic space group $C222_1$ where one molecule of Uba1 bound to the ligand was present in the asymmetric unit. However, this complex was also crystallized in the closely related space group $P2_1$, which had two molecules in the asymmetric unit. The Uba1-Ub complex was crystallized in another orthorhombic space group $P2_12_12$ with two molecules in the asymmetric unit. Both complex crystals for Uba1-Ubc13 belonged to the monoclinic space group C2 and possessed one molecule in the asymmetric unit. The data collection and refinement statistics of the final models are presented in Table 7.

Parameters	**Structure**					
	Uba1-Mg-ATP	**C600A-Mg-ATP**	**Uba1-Ub**	***Ct*Uba1-*Ct*Ubc4**	***Sc*Uba1-*Sc*Ubc13**	***Sc*Uba1-*Sc*Ubc13**
Beamline	BESSY BL14.1	BESSY BL14.1	BESSY BL14.1	ESRF ID 23-2	ESRF ID23-2	ESRF ID 29
Wavelength (Å)	0.9184	0.9184	0.9184	0.873	0.873	0.9762
Space group	C2 2 2_1	C2 2 2_1	P2 2_1 2_1	P3_2 2 1	C1 2 1	C1 2 1
Unit cell Parameters (Å)	106.91, 118.31, 196.39	110.013, 117.564, 196.063	77.04, 173.57, 215.23	124.22, 124.22, 219.35	181.39, 58.27, 137.46	177.71, 72.9, 139.52
R_{sym}[a]	0.09 (1.43)	0.09 (0.8)	0.18 (1.06)	0.14 (0.81)	0.16 (1.56)	0.23 (1.7)
R_{pim}[b]	0.05 (0.94)	0.06 (0.58)	0.123 (0.73)	0.06 (0.38)	0.1 (1.001)	0.17 (1.24)
$CC_{1/2}$	0.999 (0.53)	0.998 (0.7)	0.99 (0.52)	0.99 (0.56)	0.997 (0.36)	0.97 (0.41)
Redundancy	6.7	4.6	5.6	5.3	6.8	5.2

Unique reflections	102,936	72,519	104,172	12,984	60,597	63,720
Completeness	0.99 (0.97)	0.95 (0.95)	0.99 (0.97)	1 (1)	1 (1)	0.99 (0.98)
$\langle I/\sigma I \rangle^c$	13 (1.3)	10.1 (1.6)	8.4 (1.6)	7.1 (2)	7.5 (1.0)	5.8 (1.2)
Resolution limits (Å)	46.95-1.87	43.7-2.08	46.27-2.47	76.8-4.4	45.1-2.3	25.0-2.35
R^d /R_{free}^e	0.1631/ 0.2006	0.1897/ 0.2385	0.1887/ 0.2252	0.2837/ 0.3387	0.2110/ 0.2581	0.2270/ 0.2916
Bond distances (Å)	0.005	0.003	0.003	0.008	0.003	0.007
Bond angles (°)	0.957	0.775	0.720	1.208	0.758	1.165
Planar groups (Å)	0.004	0.003	0.003	0.006	0.003	0.004
Chiral centers ($Å^3$)	0.037	0.029	0.027	0.045	0.03	0.063
Ramachandran statistics (%) (Preferred/ Allowed/ Outliers)	95.97/ 3.93/ 0.10	95.89/ 3.52/ 0.59	97.18/ 2.72/ 0.09	91.02/ 6.93/ 2.04	96.12/ 3.7/ 0.18	95.5/ 4.0/ 0.5
Overall average B factor (Chiral centers $Å^2$)	50.4	55.6	45.0	270.7	40.1	34.5

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i |I_i - \langle I \rangle| / \Sigma_{hkl}\Sigma_i I_i$ where I_i is the i^{th} measurement and $\langle I \rangle$ is the weighted mean of all measurements of I.

[b] $R_{pim} = \Sigma_{hkl}\ 1/(N-1)^{½}\ \Sigma_i |I_i - \langle I \rangle| / \Sigma_{hkl}\Sigma_i I_i$, where N is the redundancy of the data and $I(hkl)$ the average intensity.

[c] $\langle I/\sigma I \rangle$ indicates the average of the intensity divided by its standard deviation.

[d] $R = \Sigma_{hkl} ||F_o| - |F_c|| / \Sigma_{hkl}|F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes.

[e] R_{free} same as R for 5% of the data randomly omitted from the refinement. The number of reflections includes the R_{free} subset.

Table 7: Data collection and refinement statistics.

2.4.4 Mg-ATP bound Uba1 crystal structure

Although several Mg-ATP bound crystal structures have been reported for other E1s as well as *Sp*Uba1, which could serve to position Mg-ATP with fairly high confidence, a structure where Uba1 is crystallized with this nucleotide in the absence of ubiquitin was missing. Such a structure would reflect the conformation of protein that is not affected by the binding of ubiquitin and will represent the impact of Mg-ATP binding on the Uba1 structure. To achieve this, Uba1 in the presence of Mg-ATP was crystallized. These crystals had Uba1 packed in the space group C 222_1. Surprisingly, these crystals diffracted to fairly high resolution, up to 1.9 Å, which is the best resolution achieved for

any canonical E1 enzyme. The overall domain architecture of Uba1 appeared to be in the open conformation when compared to previously reported crystal structures.

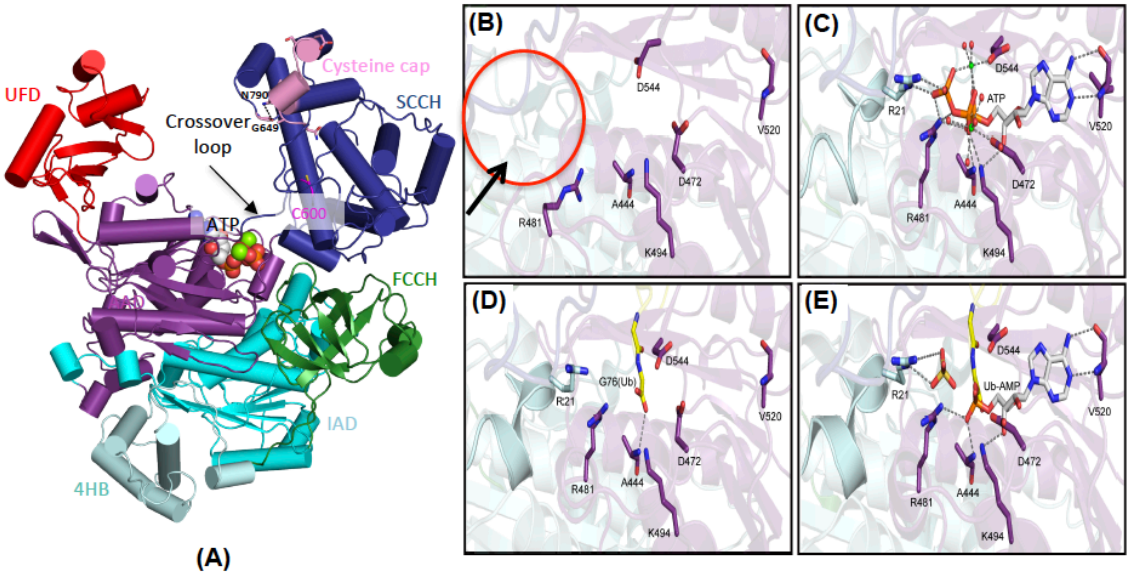


Figure 24: (A) Mg-ATP bound crystal structure of *Sc*Uba1 at 1.9 Å resolution. Enlarged view of the adenylation site in four different crystal structures representing the apo form (B), bound to Mg-ATP (C), in the presence of ubiquitin (D) [PDB code: 3CMM] and the AMP•Ub bound state (E) [PDB code: 4NNJ]. The red circle and black arrow represent in panel (A) highlights the disordered N-terminal in this structure and missing R21 interaction site.

The high resolution reveals features in the structure that were not yet seen in any other Uba1 structure deposited in the PDB. A loop region masking the cysteine domain termed the cysteine cap could never be visualized in the Uba1 structure due to disorder. It is believed to either mask the catalytic cysteine to prevent it from unspecific reactions as well as to protect it from oxidative and chemical damage. This region corresponds to residues 775 to 792 in *Sc*Uba1. Although some parts of this loop were visible in previously reported structures, the region colored in pink in Figure 24A did not show sufficient electron density to be modeled before. The cysteine cap is mostly unstructured and comprises of two β-turns. The side chain of residue N790 appears to play an important role in bringing order to this loop as it is involved in two hydrogen bonds with the main chain of residue 649 and 651 (Figure 24A). Interestingly, residues 646-649 are also disordered in several structures suggesting correlated structural changes in this region.

At the ATP binding pocket of Uba1, ATP is coordinated by several key residues that not only participate in interactions with the ligand but also contribute to catalysis by stabilizing the transition states for acyl adenylate and thioester formation. R21, the arginine finger, stabilizes the γ-phosphate by compensating its negative charge.

Similarly, R481 interacts with the γ-phosphate as well as the phosphoanhydride link between the β and γ-phosphates. The side chains of N478 and Q482 and the main chain of A444 also coordinate with the phosphate backbone of ATP. Another key residue, K494, interacts with the β-phosphate on the one hand and the 3' hydroxyl group of the ribose on the other. Aspartate 470 is coordinating both hydroxyl groups of the ribose. The adenine base is coordinated by the the main chain carbonyl and amino groups of V520 as well as the side chain of N545. Another unique feature of the ATP binding pocket of Uba1 is the coordination of two magnesium ions with the help of D544 and D472 to ATP. Although D544 is highly conserved and is shown to be crucial for catalysis, coordination of the second magnesium ion by D472 was not visualized in any other E1 enzyme structure as this residue, although being conserved, is spatially away from ATP there. Previous reports of second magnesium ion in a similar position in the *Sp*Uba1 structure further supports the presence of this feature in Ub E1 (Olsen and Lima, 2013). With the help of the highly conserved set of residues, the entire ATP molecule is held in place in preparation for catalysis. These residues are present in a similar arrangement from the prokaryotic MoeB to the eukaryotic E1 enzymes reflecting their high sequence conservation (Figure 25).

In another structure, discussed later in section 2.3.9, we could crystallize the E2 enzyme Ubc13 in complex with Uba1. In this structure, neither ATP nor ubiquitin are bound to the complex. Therefore, we can consider this structure as an apo-structure for the adenylation site of the enzyme, which has not been captured before. In this apo-form, the entire N-terminal region up to residue 23 is found to be disordered, including the arginine finger (Figure 24B). Moreover, the ATP-coordinating residues are not positioned to interact with ATP, indicating movement of the amino acid side chains to accommodate ATP. Interestingly, in the ubiquitin-bound structure these residues appear to be poised to interact with ATP, even the arginine finger from the N-terminus is positioned as in the ATP-bound state (Figure 24D). Although being present in this orientationm R21 interacts with the side chain of D824, which is located at the C-terminus of the catalytic cysteine domain, and this interaction seems to order the N-terminal region of Uba1 in the absence of ATP. Upon superimposition of the ubiquitin-bound and Ubc13-bound Uba1 structures, we realized that in the E2-bound complex the catalytic cysteine domain has slightly moved away from the adenylation site, thus making it impossible for R21 to interact with D824, resulting in the disorder of this region. Furthermore, in the AMP•ubiquitin adduct-bound structure, R21 is coordinating a sulfate ion present in the crystallization solution which is mimicking the γ-phosphate of ATP or the released pyrophosphate after formation of the acyl adenylate (Figure 24E). Therefore, the ATP-binding pockets represented in Figure 24B-E, cover all catalytic

states at this site, the apo-enzyme and the ATP, ubiquitin and AMP-ubiquitin adduct bound states, thus providing general insights into binding and hydrolysis of ATP by E1 enzymes as this pocket is highly conserved among all E1 enzymes. Other than the cysteine cap and the ATP-binding pocket, additional regions of Uba1 structure adopt different conformation including the crossover loop and the UFD compared with the ubiquitin-bound structure highlighting ubiquitin induced conformational rearrangements. This will be discussed in more detail in section 2.4.8 and 2.4.15.

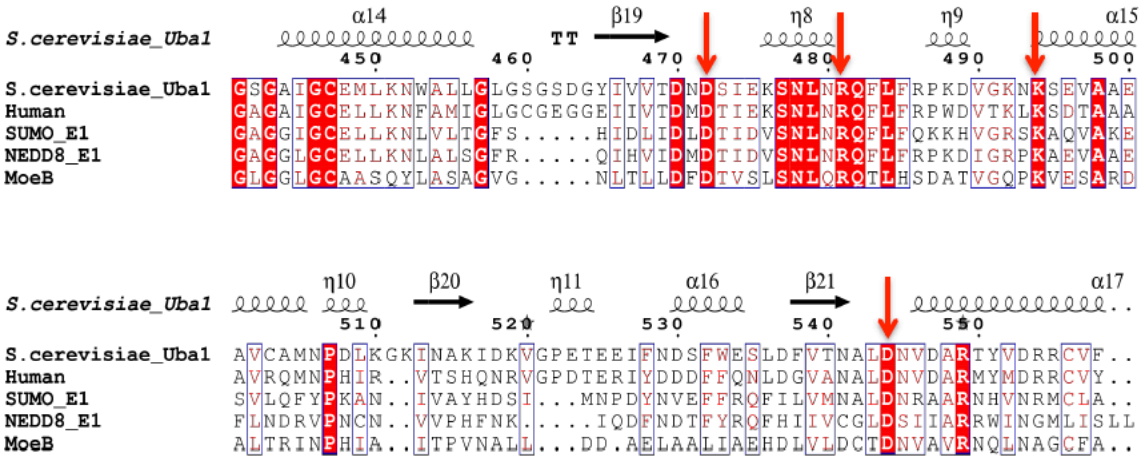


Figure 25: Multiple sequence alignment of the adenylation site of the yeast and human UBA1, SUMO and NEDD8 E1 and prokaryotic MoeB. Conserved residues are highlighted in red.

2.4.5 Biochemical characterization of the ATPase activity

To characterize the contribution of individual residues coordinating with ATP for the catalytic activity of the enzyme, the BIOMOL green assay was utilized which relies on a malachite green based reagent that turns green when reacting with phosphate. As the first catalytic half reaction yields Ub•AMP and pyrophosphate, pyrophosphatases can be used to break PP_i into two inorganic phosphates, which can be quantified using the BIOMOL green reagent. Using this assay, we tested several mutants as shown in Figure 26A as well as the catalytic cysteine mutant. The D472/544A double mutant could not be expressed as it was heavily degraded, therefore, it could not be tested in this assay.

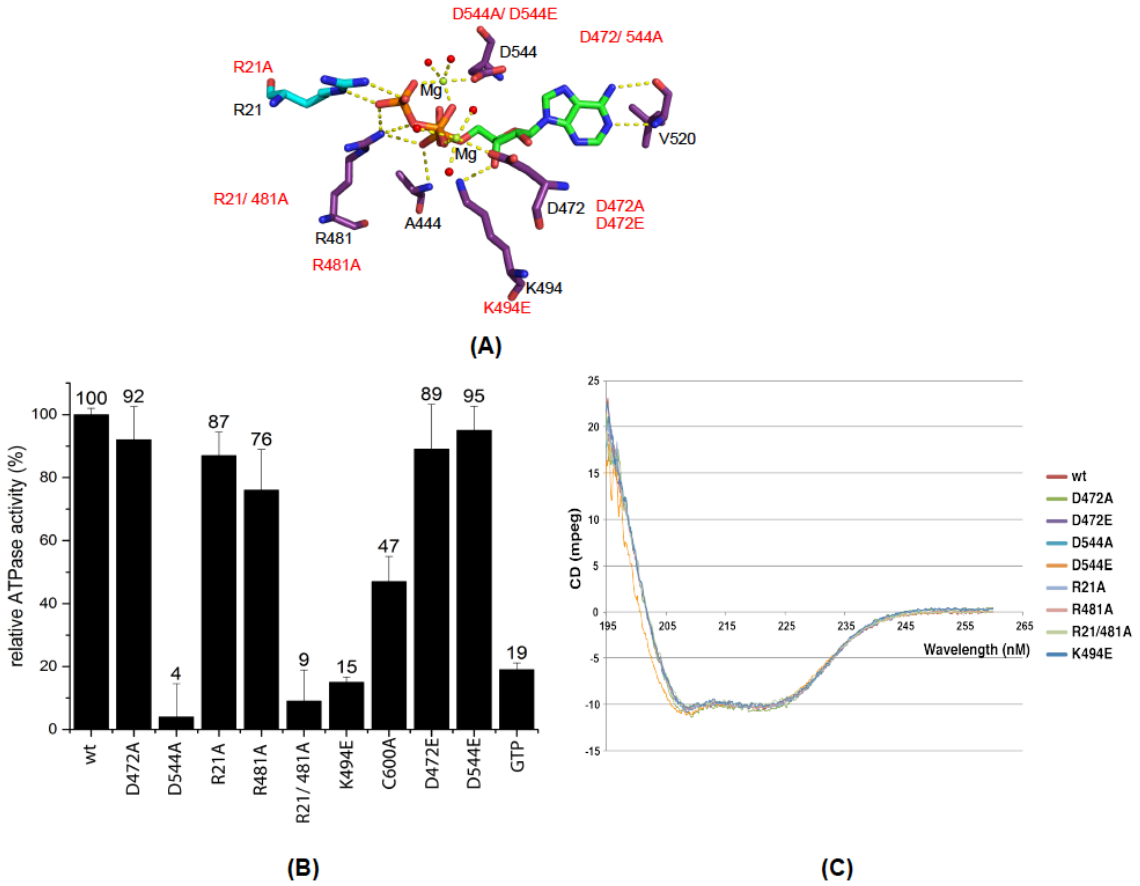


Figure 26: (A) ATP coordinating residues at the adenylation site of Uba1. The respective mutants of these residues are indicated in red. (B) Results of the BIOMOL green assay depicted as a bar graph representing the relative ATPase activity of the mutants compared to the wild-type. (C) CD spectra overlay of the wt and Uba1 mutant proteins.

The second aspartate coordination unique to Uba1 was interesting in terms of its role for the ATPase activity, therefore I mutated residue D472 to an alanine. When compared to the wild-type, the mutant did not show a significant difference in the relative ATPase activity, thus ruling out a role of the second magnesium coordination for acyl adenylate formation. The analogous mutation in the SUMO E1, D50A, abrogated thioester bond formation but had no detectable effect on adenylation, suggesting a role in stabilizing the closed conformation required for thioesterification (Olsen et al., 2010). As previously published reports demonstrated, the mutation of D544 to alanine drastically affected the catalytic activity. The two arginine residue coordinating with the phosphate backbone were mutated individually and results showed that R481 had a somewhat stronger negative effect on the ATPase activity, yet both mutants still retained a fairly high activity. However, a double mutant of both these arginine to alanine turned the enzyme nearly inactive. The K494A mutation was previously reported to dramatically affect the Ub adenylate turnover and K_m value for

ATP (Tokgoz et al., 2006). We tested how inverting the charge of the side chain would affect the ATPase activity by mutating this lysine residue to glutamate. As expected this mutant showed a severe reduction in relative activity to 15%. We also checked the mutation of the catalytic cysteine to alanine to investigate how a lack in thioesterification activity affects acyladenylate formation and detected a twofold lower activity. Therefore, the catalytic cysteine residue, which plays crucial role in vacating the adenylation site by attacking the ubiquitin C-terminus, plays an important role in the rate of the first catalytic half reaction. Furthermore, we mutated the two aspartates coordinating the two magnesium ions as visualized in the crystal structure to glutamate to check how the addition of one carbon atom in the side chain of aspartate impacts the geometry of ATP coordination and thus the rate of catalysis. Surprisingly, these mutations did not affect the ATPase activity to a significant extent, thus indicating that the side chains can rearrange themselves to keep the Mg ion in position for catalysis, in particular this was true for the D544E mutation. All purified mutants were checked via Circular Dichroism (CD) Spectroscopy to confirm whether they folded properly. As shown by Figure 26C the mutants show identical profiles and hence are properly folded.

We also checked whether Uba1 can generate a Ub•GMP adenylate when GTP is introduced instead of ATP. The enzyme displayed GTPase activity, however, it was only 20% of its ATPase activity. This is consistent with results presented by Ciechanover et al. when they used GTP to covalently bind UBA1 with a ubiquitin affinity column instead of ATP (Ciechanover et al., 1982). The presence of GTP could associate UBA1 with the column, however, the amount of protein bound to the column was fourfold higher when ATP was used. CTP could also bind UBA1 in this experiment but to a slightly lower extent than GTP.

The omit map for Mg-ATP bound to Uba1 in the aforementioned crystal structure is displayed in Figure 27A. Interestingly, we could solve another crystal structure of Mg-ATP bound Uba1 at 2.1 Å where a ubiquitin C-terminal mimicking heptapeptide was added during crystallization. Although we could not visualize sufficient electron density to model the complete peptide in this structure, remarkably, we found the phosphate backbone of ATP in a different orientation in this crystal structure (Figure 27B). Strikingly, the phosphoanhydride bond between the α-β phosphate adopts alternate conformation in this structure supported by the omit map electron density (Figure 27C). In addition, we could not locate the second magnesium ion in this structure due to the oxygen from the β-phosphate not being in the same position as in the other structure. As presented in Figure 27E, upon superimposition of the two conformations of Mg-ATP in these structures, we can see that where in one structure K494 is coordinating the β-phosphate and the 3’ hydroxyl group of the ribose, in the other structure the

coordination is shifted to the α-β phosphoanhydride link while the hydrogen bond with the ribose stays the same. The α-β phosphoanhydride bond is cleaved during the enzymes catalysis to release pyrophosphate and the reorientation of this bond stabilized by K494 suggests it to be involved in transition state stabilization. Similarly, coordination of ATP via the magnesium ion which in turn is held in place by D544 is altered in the two structures. Mutations of D576 or K528 in human UBA1 have previously been shown to impair binding of Mg-ATP and shift the rate-limiting step to ubiquitin adenylate formation, thus suggesting that these residue might be critical for transition state stabilization (Tokgoz et al., 2006). These structures support these data, confirming critical role of these residues for the ATP to AMP hydrolysis during the catalytic cycle of the enzyme.

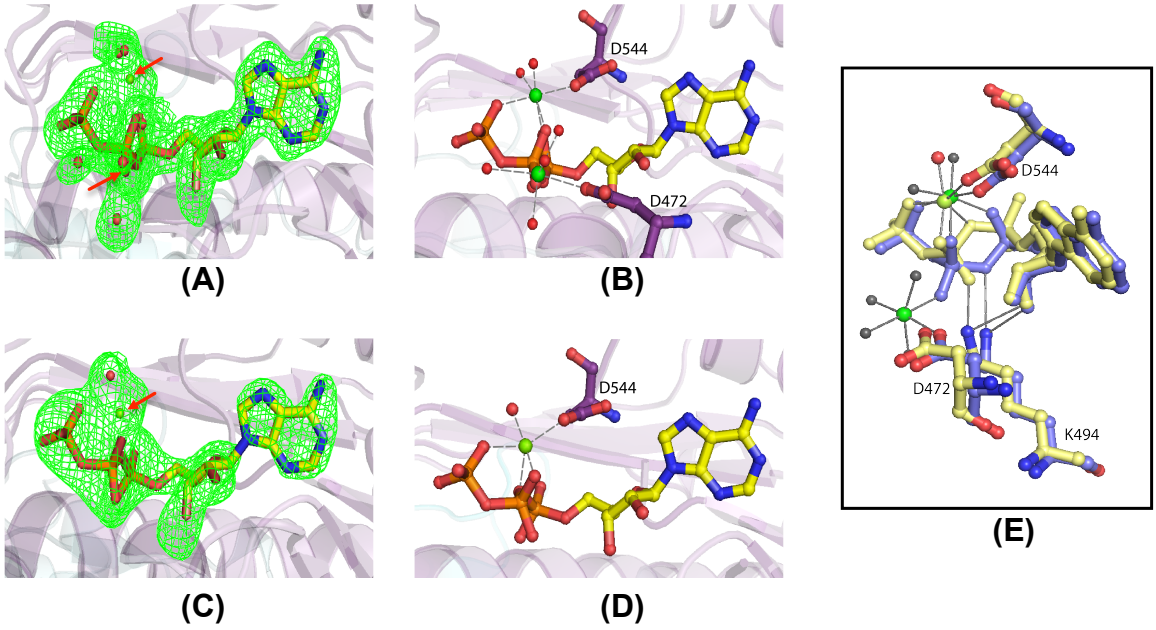


Figure 27: (A, B) Omit map for Mg-ATP contoured at an rmsd of 3 in Uba1 structures where in one structure two magnesium ions are coordinating with ATP and in the other only one magnesium ion is visible. (C, D) Details of the Mg-ion coordination spheres (E) Superposition of the two ATP conformations depicting differences in the α-β phosphoanhydride link, K494 and Mg-coordination residues.

Based on the crystal structures and supporting biochemical data, we can model the first catalytic half reaction of Uba1 where in the presence of ATP in the designated pocket, the C-terminal carboxylate of ubiquitin carries out a nucleophilic attack on the α-phosphate of ATP. This attack results in a pentacoordinate transition state, which is unstable and results in α-β phosphate scission, thus releasing pyrophospahate and magnesium. This transition state is stabilized by magnesium coordination as well as the

side chain of K494. This model takes into account the significance of D544 and K494 during catalysis as demonstrated by the activity assay.

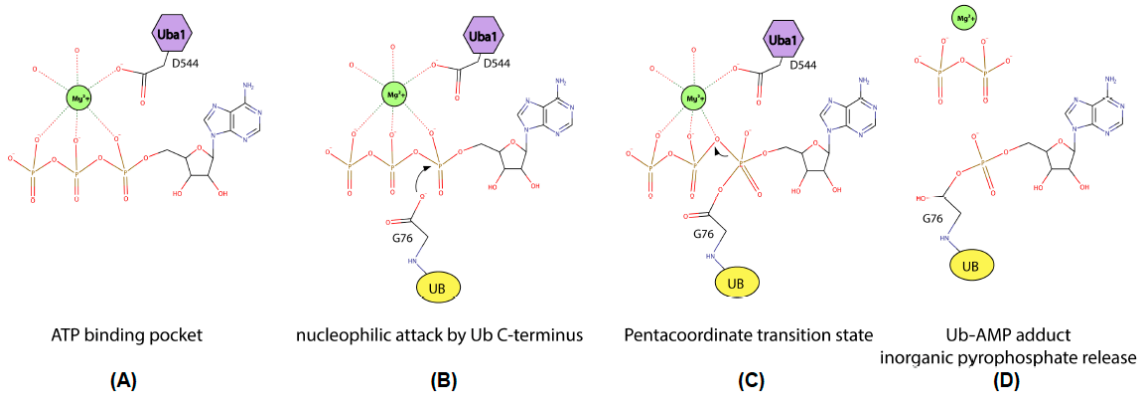


Figure 28: Model of Ub•AMP adduct formation starting with ATP and ubiquitin coupled to the release of PP_i.

It is important to note that UBA1 is not just able to catalyze the forward reaction where ubiquitin adenylate and ubiquitin thioester are formed and PP_i and AMP are released but it can also promote the backward reaction if both AMP and PP_i are present in excess. This was shown by Ciechanover et al. by adding AMP and PP_i to the ubiquitin affinity column as one of the possibilities to elute the covalently bound UBA1 (Ciechanover et al., 1982). This implies that both ubiquitin forms are interchangeable, ie. thioesterified ubiquitin can be converted back to adenylated ubiquitin and visa versa. Moreover, adenylated ubiquitin can be modified back to ubiquitin in the presence of an excess of PP_i. This might be the explanation for the structures of Uba1, SUMO E1 and NEDD8 E1 where both ubiquitin or the ubiquitin-like modifier and ATP were visualized at their respective positions but no electron density was observed corresponding to the Ub/Ubl•AMP adducts.

2.4.6 Uba1-ubiquitin complex

Although the structure of Uba1 noncovalently bound to ubiquitin was already reported by Lee and Schindelin, I could improve the resolution of this structure from 2.7 Å to 2.5 Å. These crystals emerged in a new crystallization condition consisting of 0.2 ammonium formate and 20% peg 3350. The overall organization of this structure appeared to be very similar when compared to the previously reported structure (Figure 29A). However, due to the higher resolution we could model alternate conformations of the side chains for some residues at the Uba1-ubiquitin interface, which appear to be involved in interactions and were not reported before.

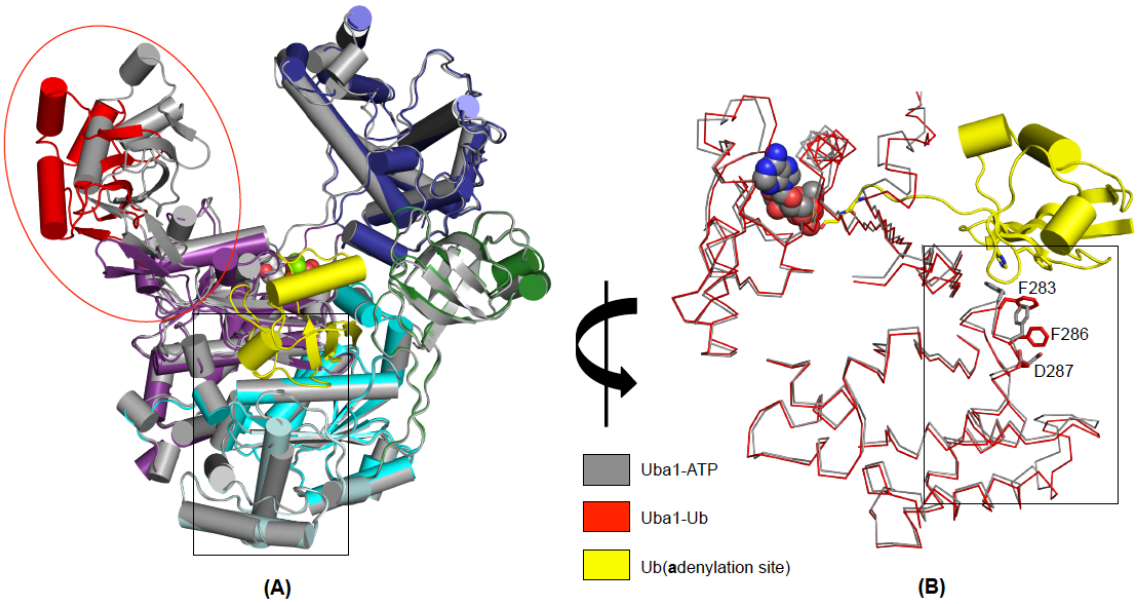


Figure 29: (A) Superposition of the Mg-ATP bound Uba1 structure (in grey) with the ubiquitin-bound Uba1 structure (colored according to domain architecture). (B) Ribbon representation of the altered Cα backbone of the 4HB region (indicated by black rectangle) upon ubiquitin-binding. Residues showing variable conformations are depicted as sticks.

These interactions include the interface of ubiquitin with the crossover loop of Uba1, as will be described in detail in section 2.4.8. When we superimpose the Mg-ATP and ubiquitin-bound Uba1 structures, three regions in particular stand out for their altered conformations upon ubiquitin binding. These regions include the UFD, the crossover loop and the four-helix bundle. The UFD and crossover loop regions will both be discussed in separate sections of the thesis, therefore only the changes in the 4HB domain will be discussed here. The hydrophobic interface between ubiquitin and Uba1 involves both the AAD and 4HB from Uba1 and the canonical hydrophobic patch on ubiquitin that is comprised of residues L8, I44, H68 and V70. Uba1 presents four surface exposed phenylalanines at this interface. The side chain of F898 has to move away upon ubiquitin binding to accommodate both L8 and V70. F905 participates in van der Waals interactions with H68 of ubiquitin and undergoes a slight change in its side chain conformation. Furthermore, F283 has to move away from H68 to avoid clashes and this disrupts the π-π stacking between F283 and F286, also resulting in a rotation of its side chain (Figure 29B). This sequence of movements forces D287 to also adopt a different orientation. As in the absence of a Uba1 structure crystallized without ubiquitin these alterations could not be monitored, this further underscores the value of the Mg-ATP-bound Uba1 structure. Due to the binding of ubiquitin a slight bent (rmsd

0.339) is observed in helices 2 and 3 of the 4HB, as a result of tight hydrophobic interactions in this region.

Upon comparison of ubiquitin forms in the three Uba1-bound states, non-covalently bound ubiquitin, ubiquitin adenylate and thioesterfied ubiquitin, we observe the first two forms to be structurally more closely related. The RMS deviation for ubiquitin and ubiquitin adenylate is 0.29 Å while it is 0.69 Å between ubiquitin/ubiquitin adenylate and thioesterified ubiquitin. This is mainly due to the rearrangements in the flexible C-terminal tail and alterations in the loop region between β-strands 3 and 4 involved in interactions with the FCCH domain at the transthioesterification ubiquitin binding site. Interestingly, the interaction interface of ubiquitin at both the adenylation site and the transthioesterification site utilizes similar regions that include the hydrophobic patch discussed earlier as well as residue R42. However, R72 which is the specificity determining residue for ubiquitin for Uba1 recognition appears to be disordered when ubiquitin is covalently bound.

2.4.7 Active site remodeling

The structure of SUMO E1 where it has been captured in a state that mimics the tetrahedral intermediate using a chemical trick lays the foundation to understand how the catalytic cysteine domain fulfills its role in the thioesterification of ubiquitin at the adenylation catalytic center (Olsen et al., 2010). This was the first instance where an E1 enzyme was crystallized in a 'closed' conformation. The authors used two semisynthestic protein inhibitors especially designed to trap the enzyme in a SUMO-adenylate mimicking state and a transition state where the catalytic cysteine makes a nucleophilic attack on the C-terminus of ubiquitin to generate a tetrahedral intermediate state (Lu et al., 2010). The SUMO adenylate mimicking probe was generated by chemical ligation of a CGG-AMSN(5'-(sulphamoylaminodeoxy)adenosine) tripeptide to a SUMO1$^{1-94}$-ΔC construct missing the three C-terminal residues of the mature SUMO. This inhibitor had a sulfamide as a non-hydrolyzable analogue of the α-phosphate of AMP. For capturing the tetrahedral intermediate state, a SUMO-AVSN (5'-(vinylsulphonylaminodeoxy)adenosine) probe was synthesized in a similar fashion, except it contained a vinyl sulfonamide electrophile to trap the incoming cysteine nucleophile in the second half reaction. The SUMO1-AMSN bound structure resembled the open conformation of the enzyme (figure 30B) as seen in the Mg-ATP bound structure (Figure 30A) while the SUMO1-AVSN presented a closed conformation.

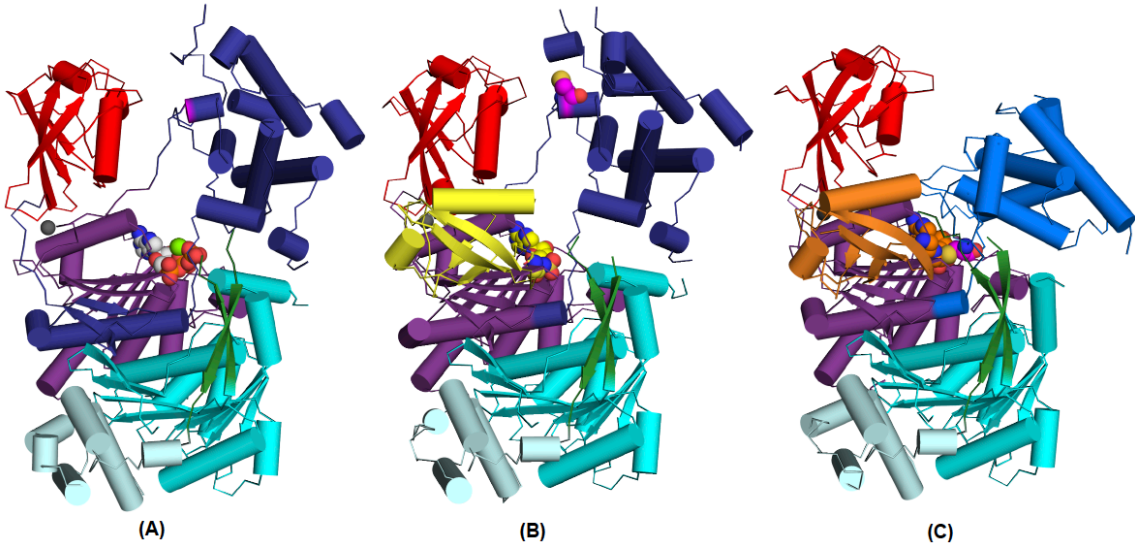


Figure 30: Various catalytic states of SUMO E1. (A) Mg-ATP bound structure, (B) SUMO1-AMSN mimicking the SUMO adenylate bound to SUMO E1, and (C) SUMO1-AVSN bound form mimicking the tetrahedral intermediate state of the enzyme. SUMO1-AMSN is colored in yellow while the SUMO1-AVSN is colored in orange. The SCCH (light blue in panel C, dark blue in A and B) undergoes dramatic conformations changes. The catalytic cysteine residue is colored in magenta in all three structures.

The closed conformation showed that the catalytic cysteine domain undergoes a 130° rotation compared to the open state and this is supported by remodeling of several components of the adenylation site as well as the two linkers, the crossover loop and the reentry loop, which connect the catalytic cysteine domain with the active adenylation domain. The remodeling in the adenylation site included H1 and H2 of SAE1, which become disordered in the closed state, and H3 of UBA2 that forms one side of the adenylation active site in the open conformation and unfolds into a loop in the closed state. The unfolding of the latter shifts the corresponding residues into the space, which is occupied by helix H1 of SAE1 in the open state. Displacement of the residues required for adenylation facilitates the movement of the catalytic cysteine towards the adenylation site as well as the release of pyrophosphate produced in the first catalytic half reaction. In addition, both the crossover loop and the reentry loop act as hinges for the rotation of the catalytic cysteine domain and, finally, the helix in which the catalytic cysteine resides melts to enable the thioesterification reaction at the adenylation active site.

Recently Lv et al., reported the first structure of *Sp*Uba1 in the absence of any of the substrates (either ATP, ubiquitin or E2 enzyme) representing its apo-form at both the adenylation site and the transthioesterification site (Figure 31C) (Lv et al., 2017b). This structure reflects the closed state of Ub E1 seen for the first time for this enzyme and

shows several similarities to the closed state of SUMO E1. The remodeling of the SCCH and AAD is not just required to bring the catalytic cysteine residue into close proximity of the adenylation site, but it also favors pyrophosphate release, thereby further pushing the reaction in the forward direction and abrogating product inhibition of the enzyme. Whereas in the closed state of SUMO E1 the catalytic cysteine is trapped in the tetrahedral intermediate state after travelling a distance of 36 Å, in the *Sp*Uba1 structure, despite a rotation of 106° of the SCCH domain, the catalytic cysteine is still 23 Å away from the modeled position of the ubiquitin C-terminus bound at the adenylation site (Figure 31D).

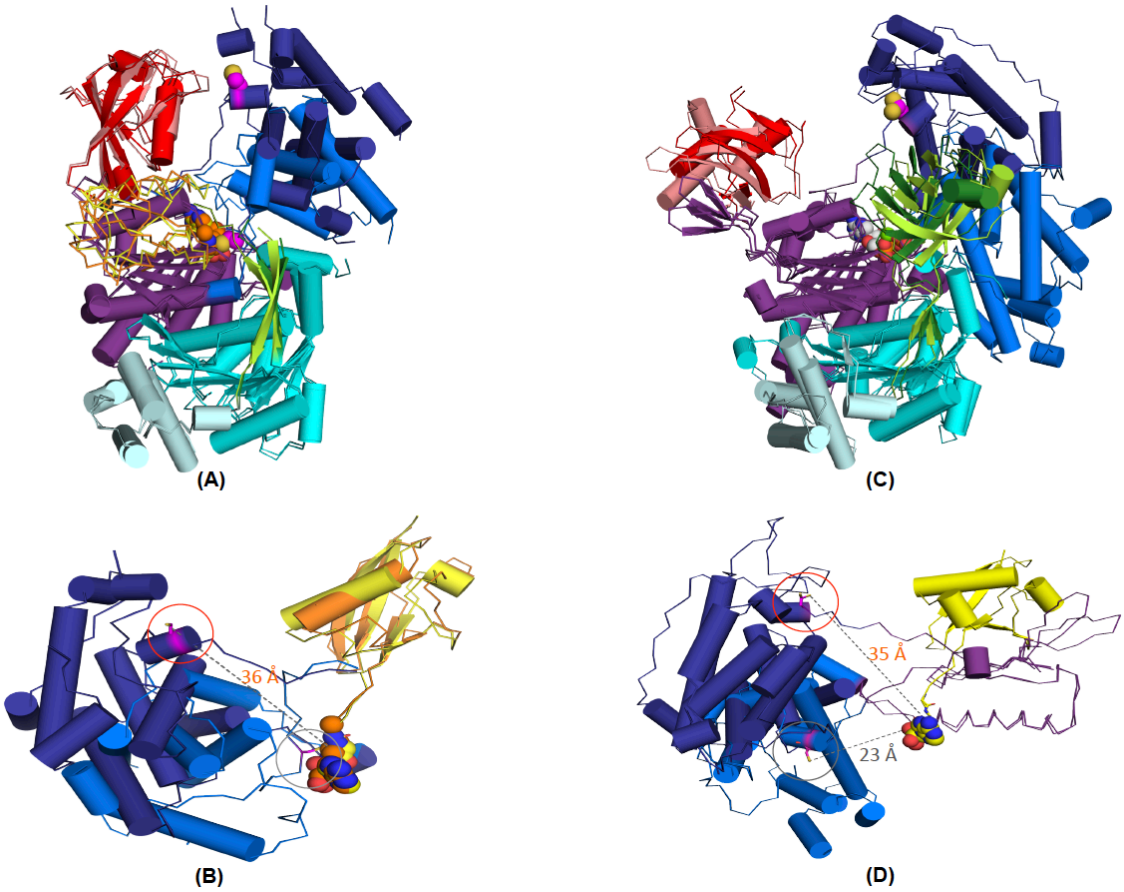


Figure 31: Panels A and C represent the superimpositions of the open and closed states of SUMO E1 and Ub E1, respectively. Panels B and D depict the altered catalytic cysteine domains in the open conformation (dark blue) and closed conformation (light blue). The distance of the catalytic cysteine residue from the C-terminus of SUMO1 or ubiquitin are shown. Adenylated forms of Ub/Ubl are shown in yellow while the thioester intermediate is colored in orange.

2.4.8 The cross-over loop

The crossover loop is a long loop connecting the catalytic cysteine domain to the adenylation domain and appears to be playing a crucial role in bringing the two active sites into close spatial proximity. Interestingly, from a structural perspective it lies close

to both the ubiquitin binding site as well as the ATP binding pocket and has been shown to be part of the polar interface for the binding of ubiquitin to the adenylation site. Upon overlaying this region in the Mg-ATP complex and the ubiquitin bound complex, we observe altered conformations of multiple residues (Figure 32). Four residues from ubiquitin, R42, R72, R74 and D39, participate in interactions with the crossover loop. The three arginine residues have been shown to be critical for adenylation site binding of ubiquitin by site directed mutagenesis (Burch and Haas, 1994).

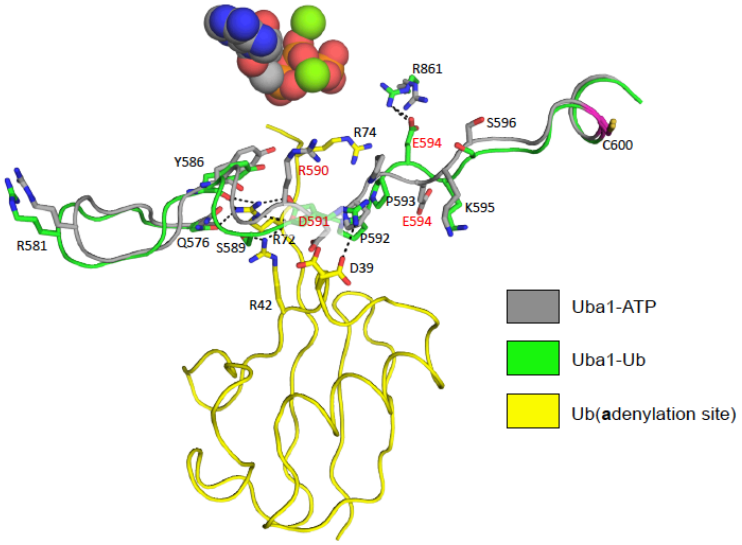


Figure 32: Superposition of the crossover loop in the ATP bound (grey) and ubiquitin bound (green) structures of Uba1. Ub bound at the adenylation site is shown in yellow and Mg-ATP is represented as spheres.

The D39 residue, which was not resolved in the low resolution structure, shows binding to the side chain of R590 in an alternate conformation of this side chain, which is visible in the higher resolution structure of the ubiquitin-bound complex. This interaction rotates R590 by more than 90° away from its orientation seen in Mg-ATP complex. The other two residues showing significant conformational changes on ubiquitin binding are D591 and E594. D591 is involved in several hydrogen bonds with R42 and R72 of ubiquitin , whereas E594 is interacting with R861 of the SCCH domain upon ubiquitin binding. The ratio of the changes in the main chain dihedral angles (Δphi and Δpsi) of individual residues shows dramatic rearrangement on ubiquitin binding as well as in the closed state of the enzyme (Lv et al., 2017b). In Figure 33, the close-up view of the crossover loop in the presence of either of the substrates, ubiquitin adenylate and apo enzyme in closed conformation reported for *Sp*Uba1 is illustrated.

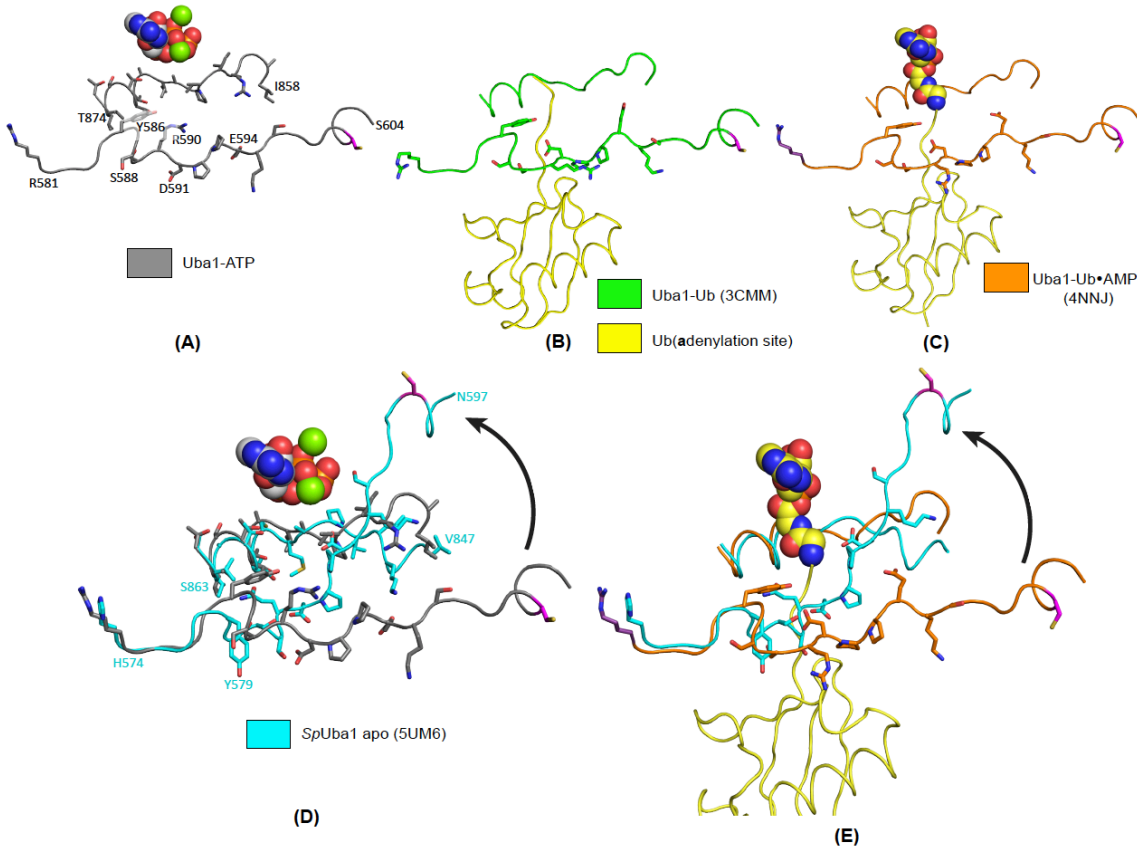


Figure 33: Snapshots of the crossover loop (residues 581-604) and reentry loop (residues 858-874) in *Sc*Uba1 in various catalytic states of the enzyme. The catalytic states involve (A) the Mg-ATP-bound form, (B) the complex with ubiquitin and (C) the ubiquitin-adenylate bound Uba1. (D) Superimposition of closed state of *Sp*Uba1 (apo-state) and the open state of *Sc*Uba1 (in complex with Mg-ATP). In (E) the latter is replaced with the *Sc*Uba1 ubiquitin adenylate complex.

The arrangement of residues in the crossover loop are quite similar in the ubiquitin and ubiquitin adenylate bound forms, except for residue R590, presumably this reflects an intrinsic flexibility of this residues. The overlay of open and closed states with respect to the crossover loop reveals the path of rotation of the loop to bring the catalytic cysteine closer to the adenylation site, similar to that seen for the SUMO E1. However, as the cysteine is still 23 Å away from the C-terminus of ubiquitin, further rearrangements are required including the unfolding of the helix containing the catalytic cysteine. Notably, upon superimposition of the closed apo-state to the open ubiquitin-bound state, steric clashes of the crossover loop region with the C-terminus of ubiquitin are observed (Figure 62), reflecting that the reported structure is not catalytically competent for thioesterification. Bending of the crossover loop starts at residue Y579 (corresponding to Y586 in *Sc*Uba1), which is highly conserved in the Ub E1 from yeast to humans. Interestingly, the crossover loop region is not highly conserved among canonical E1

enzymes even though they follow the overall similar mechanism which requires critical rearrangements in this loop. This may be in part due to the fact that both SAE2 and UBA3 possess a zinc coordination site preceding the crossover loop, thus restricting conformational flexibility of the loop (Wang and Chen, 2010). The two cysteines for the zinc coordination arise from the crossover loop region whereas the other two cysteines are present in the UFD domain. The Ub E1 is bereft of this zinc binding site where the coordinating C-X-X-C motif at the crossover loop is replaced by S-X-X-S. Due to coordination of Zn^{2+} by these residues the bending of the SUMO E1 crossover loop is shifted by five residues in the C-terminal direction, in comparison to Ub E1. Mutations of residues in the crossover loop in case of the SUMO E1 did not significantly alter either the adenylation or thioesterification activity of the enzymes, suggesting that conformational changes over several residues could compensate for the effects of single point mutations. It can also be that drastic rearrangements of the main chain in this region devalue the effect of point mutations. It would be interesting to test whether the crossover loop of Uba1 follows similar trend as seen for SUMO E1 or the amino acid side chains of crossover loop in its case contribute to proper catalysis of the enzyme. In contrast, mutations in the reentry loop region abrogate the thioester formation activity as shown for the K850P mutation in *Sp*Uba1 (K850 corresponds to R861 in *Sc*Uba1). The comparison of the closed and open states of E1 enzymes suggests that the crossover and reentry loop regions are far apart in the open state, whereas they interact with each other as shown in the SUMO E1 closed state where they form a parallel β-sheet (Olsen et al., 2010).

2.4.9 E1-E2 transthioesterification

The doubly loaded Ub E1 structure with adenylated ubiquitin and thioesterified ubiquitin unravels the positioning of the two ubiquitin molecules on the E1 enzyme, which suggests that after capturing activated ubiquitin in the closed conformation accompanied by the release of AMP, the enzyme again adopts the open conformation (Schafer et al., 2014). Why the enzyme needs to get back into the open conformation is clear by the structure reported by Olsen and Lima where *Sp*Uba1 is cross-linked to the E2 enzyme Ubc4, since the enzyme can provide the binding interface to the incoming E2, which is required for the transthioesterification reaction to occur, only in its open conformation (Figure 34A) (Olsen and Lima, 2013). In this structure *Sp*Ubc4 is sandwiched between the UFD and SCCH domain and is involved in a tripartite interaction with these two domains and, via a third smaller interface, with the crossover loop. Another structure of *Sp*Ubc15 in complex with *Sp*Uba1 was recently reported by Lv et al., which reflects plasticity in Ub E1-E2 interactions (Lv et al., 2017a). Ub E1 is

able to transfer ubiquitin in a thioester transfer reaction to a number of E2 enzymes, ranging from 10-35 depending on the organism (van Wijk and Timmers, 2010). In contrast, NEDD8 E1 and SUMO E1 work with two (Ubc12 and Ube2F) and one (Ubc9) E2 enzymes, respectively. The E1 enzymes have evolved to recognize their related E2 enzymes and reject the E2s that belong to other Ubl-modifier systems. The complexes for Uba1-Ubc4 and Uba1-Ubc15 both were formed by cross-linking the E1 and E2 with 2,2'-dipyridylsulfide, thus resulting in a E1-E2 disulfide bridge which is not a catalytically pertinent state, however, a mutational analysis of the interface confirmed the binding interface required for thioester transfer. Although the three interfaces between the E1 and E2 are common between both structures, noticeably the unique features of Ubc15 were accommodated differently, when compared with the Ubc4 bound complex.

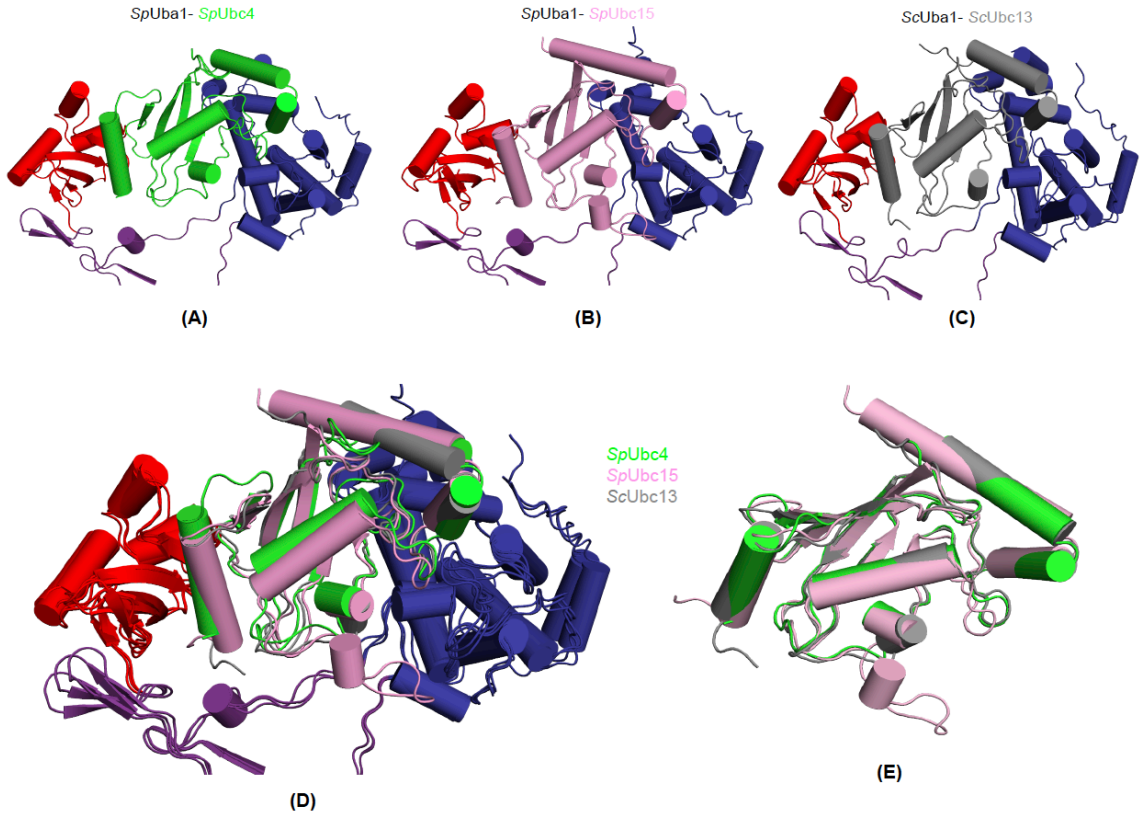


Figure 34: Representation of the E1-E2 interfaces in (A) the *Sp*Uba1-*Sp*Ubc4 [PDB code: 4II2], (B) *Sp*Uba1-*Sp*Ubc15 [PDB code: 5KNL] and (C) *Sc*Uba1-*Sc*Ubc13 [unpublished] structures. The UFD, SCCH and AAD are depicted in red, dark blue and magenta, respectively. In panel (D) all three complexes are superimposed and panel (E) shows only the isolated E2 enzymes from the complexes.

The *Sp*Ubc15 has two unique features on top of the highly conserved UBC domain of E2 enzymes (Figure 34B and E). First, it has a longer N-terminal region, which is docked on the UFD domain by shifting the N-terminal helix by 1.7 Å (Figure 34D). When Ubc15 is modeled to bind in a similar fashion as Ubc4, clashes with the UFD are

observed. Second, Ubc15 has a 13-residue long acidic loop insertion, close to its active site, which provides an additional interface for E1 binding. As shown in the structure, (Figure 34D) the loop is wedged in between the SCCH and FCCH domain, shifting them both by a few angstroms away form each other. Deletion of this acidic loop in Ubc15 and other E2s with the corresponding insertion has been shown to increase the thioester transfer activity. Therefore, it is unclear whether these loop are negative regulators of E1-E2 interaction or they reserve additional purposes. Due to the longer N-terminal sequence Ubc15 engages in more extensive contacts with the crossover loop in addition to the acidic loop insertion, whereas in the Ubc4 complex the latter contacts are absent. Ubc15 also possesses a longer C-terminal helix, however, it is devoid of contacts with the SCCH (Figure 34D-E). A comparison of the reported E1-E2 complex structures suggests that differences in the primary sequences of E2 enzymes are correlated with an altered E1-E2 topology, which facilitates the generation of a matching E1-E2 interface.

We could crystallize another E2, Ubc13 in complex with *Sc*Uba1, at a resolution of 2.3 Å (Figure 34C). These crystals belonged to the monoclinic space group C2. Remarkably, we did not have to follow the cross-linking strategy to obtain this complex. Instead, complex formation was promoted in the presence of Mg-ATP and 1 mM H_2O_2 as oxidizing agent. resulting in a disulfide bridge between the catalytic cysteine residues of *Sc*Uba1 and *Sc*Ubc13.

2.4.10 The Uba1-Ubc13 complex

The Uba1-Ubc13 complex is the first example of an E1-E2 pairing via a disulfide link obtained in the absence of any cross-linking agent. Figure 35A shows the overall topology of the Uba1-Ubc13 complex. The tripartite E1-E2 interaction surfaces as seen in the *Saccharomyces pombe* Uba1-E2 complexes are also present in the Uba1-Ubc13 complex. Structural alignment of the SCCH domain obtained from the Mg-ATP-bound Uba1 structure with the Ubc13 derived from the E1-E2 complex indicate a clash imposed by the cysteine cap region of the SCCH domain onto Ubc13 (figure 35C-D). In all E1-E2 complexes reported for both *S. cerevisiae* and *S. pombe*, this region is disordered, suggesting it opens into a more flexible conformation in these structures.

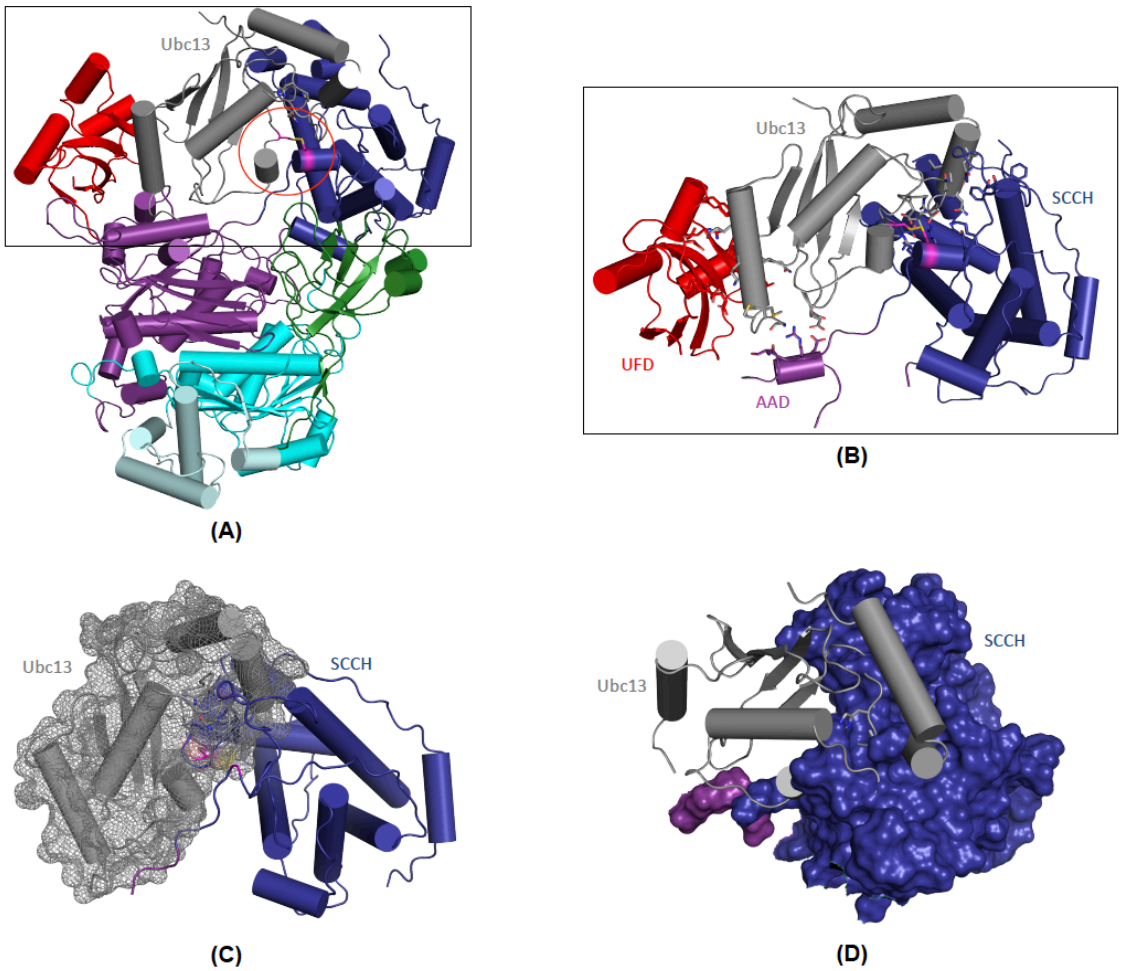


Figure 35: (A) Crystal structure of Uba1 in complex with Ubc13 (grey) at 2.3 Å resolution. The catalytic cysteine residues of both Uba1 and Ubc13 are colored in magenta showing the disulfide bridge in stick representation. (B) Zoomed-in view of the tripartite interaction of Ubc13 with UFD (red), SCCH (dark blue) and the crossover loop (magenta). The interacting residues are depicted as sticks. (C) The steric clash observed when Ubc13 (mesh surface) is superimposed onto the SCCH domain of the Mg-ATP-bound complex and the same in top view with SCCH now shown as solid surface.

The interaction of the UFD of the Ub E1 and the E2 is achieved between the N-terminal helix Ubc13 and β-strands 28-30 and the β29-H32 loop of the UFD. Moreover, the β1-β2 loop of Ubc13 contacts the β28-β29 loop and H33 of the UFD. A combination of hydrophobic and electrostatic interaction appears to stabilize this interface. The hydrophobic interactions are centered around residues I9, V16 and L32 of Ubc13 whereas residues S3, K6, K10, K13 and D29 are involved in the formation of hydrogen bonds. The interactions of the N-terminal helix of Ubc13 are presented in Figure 36A while the residues at the β1-β2 loop of Ubc13 are shown with their interacting partners in Figure 36B.

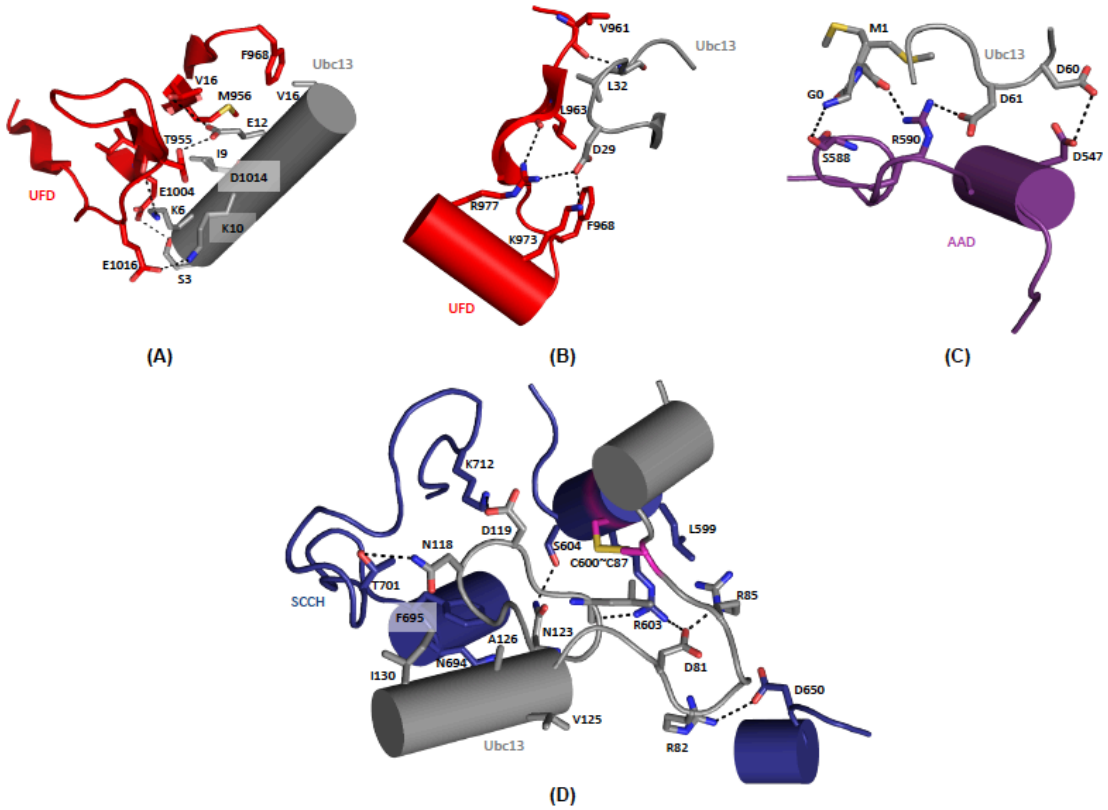


Figure 36: Interactions of Uba1 residues corresponding to the UFD (red), AAD (magenta) and SCCH domain (dark blue) with Ubc13. The participating residues are depicted as sticks and hydrogen bonds are shown as dashed lines.

At the interface with the AAD, Ubc13 engages in fewer contacts with Uba1 mainly involving R590 of AAD and the main chain carbonyl group of the N-terminal Met and the side chain of D61 present at the β3-β4 loop of the UBC fold of Ubc13. The N-terminal Gly of Ubc13 in the structure is a non-native residue derived from the protease cleavage site between the engineered His-tag and Ubc13. Two aspartates, D60 of Ubc13 and D547 of Uba1, are in close proximity in the structure, and these residues may form hydrogen bonds assuming that one of these residues has an unusually high pK_a-value (Figure 36C).

The third interface provided by the catalytic cysteine domain is rather extensive. Other than the disulfide bridge connecting the catalytic cysteines, residues R82, N118, D119 and N123 pair with complementary residues D650, K712 main chain and side chain and S604 in the SCCH via electrostatic interactions (Figure 36D). The hydrophobic interactions are provided by P120, L121, A126 and I130 of Ubc13 contacting L599, F605 and F695 of Uba1. Notably, residues 777-794 corresponding to the cysteine cap loop are disordered in this structure.

A comparison of the E1-E2 complexes discussed here unravels the determinants of the Uba1-E2 interactions at the primary sequence level, however, the overall topology of

these complexes appear to be very similar with slight adaptations. Therefore, this binding mode can be generalized for the E1-E2 interactions, especially for class I E2s possessing only the typical UBC fold. Whereas Ubc13 is more closely related to Ubc4, the N- terminal helix of Ubc13 is accommodated like that seen in the Ubc15 structure indicating malleability of the UFD-E2 interface. Other than that the Ubc13-Uba1 structure adopts similar topology of E1-E2 interaction as observed for the *Sp*Uba1-Ubc4 structure. It lacks the acidic loop insertion and a longer N-terminal sequence like Ubc15. The *Sp*Uba1-Ubc15 complex provides insight into the first example of how insertions in the UBC fold affect the E1-E2 interaction. Furthermore, the acidic patch on the UFD of Uba1 has been shown to be quite sensitive to acidic substitutions at the N-terminal helix of the E2 enzymes (Lv et al., 2017a). The presence of several Ser and Thr residues in this region in a majority of E2s, indicates possible phosphorylation events that negatively regulate E1-E2 or E2-E3 interactions as reported for the human E2, UBE2T. The reported E1-E2 structures provide an initial platform to evaluate the affinity of individual E1-E2 pairs derived from biochemical assays in terms of structural compatibility.

For the transthioesterification reaction to occur, a juxtaposition of the active sites of the E1 and E2 is required so that the active site cysteine of the E2 can attack the C-terminus of ubiquitin, which is covalently linked to the active site cysteine of the E1. As seen in the doubly loaded Uba1 structure (PDB code: 4NNJ), the thioesterified ubiquitin is positioned in front of the Uba1 structure above the FCCH domain (see Figure 15B). Upon superimposing the E1-E2 complex onto this structure it is clear how the modeled quartenary complex is poised for the thioester transfer. In addition, as both the Uba1-Ubc4 and Uba1-Ubc15 complexes have ubiquitin present at the adenylation site in their structures, few contacts between the E1 and the ubiquitin near the crossover loop are observed in these structures suggesting the possibility of a sensing mechanism by the E2 for a doubly loaded E1 (Olsen and Lima, 2013). It has been shown that the doubly loaded E1 transfers ubiquitin more efficiently to an E2 enzyme in comparison to the complex with only thioesterified ubiquitin (Haas et al., 1988). This could be explained by the steric clash of the adenylated ubiquitin with the thioesterified ubiquitin keeping the E1 in an open state which facilitates its binding to an E2 enzyme. Upon thioesterification, the E2~Ub conjugate is forced to leave since superimpositions of several E2~Ub structures adopting the open conformation onto the reported E1-E2 complexes present serious collisions of ubiquitin with E1 (Wenzel et al., 2011). This observation supports the mutually exclusive nature of E2 binding to either E1 or E3 (Eletr et al., 2005).

Strikingly, the crystallized Uba1-Ubc13 complex was stabilized by a disulfide bridge between the catalytic cysteine residues of Uba1 (C600) and Ubc13 (C87) even in the lack of any cross-linking agent. Formation of a disulfide-linked E1-E2 complex under conditions of oxidative stress has been previously reported for the SUMO system and the Uba1-Ubc3 complex, as a mechanism of resisting oxidative damage (Bossis and Melchior, 2006; Doris et al., 2012). Other than the E1-E2 disulfide bridging, several components of the UPS are modulated under oxidative stress including the proteasome and several deubiquitinases (Lee et al., 2013; Pajares et al., 2015). Stress conditions like UV stress or heat shock can induce reactive oxygen species (ROS) leading to damage and/or modifications of proteins, lipids and DNA. To avoid this, cellular machinery has developed several scavenger systems including enzymes (catalases, superoxide dismutase, peroxiredoxin) and antioxidants (vitamins, glutathione and thioeredoxin). Additionally, enzymes containing reactive cysteine residues have been shown to undergo reversible modification in order to bypass cysteine oxidation. Since the intermolecular disulfide bridge between Uba1 and Ubc13 was obtained in the presence of Mg-ATP and hydrogen peroxide (oxidixizing agent), it led us to evaluate the possibility of either oxidative stress or nucleotide induced E1-E2 covalent linking.

2.4.11 ATP-induced conformational changes allow E1-E2 disulfide bridging

We obtained the Uba1-Ubc13 complex in the presence of Mg-ATP which appeared to promote complex formation, however, the absence of the nucleotide in the structure, despite being present at a concentration of 2.5 mM, was quite surprising. As oxidative stress has been shown to induce E1-E2 disulfide-linked complexes, we speculated that this complex could be the result of the oxidizing conditions in form of 1 mM H_2O_2. To check whether complex-formation could be induced by either ATP or oxidizing conditions, we checked the formation of the disulfide-linked Uba1-Ubc13 complex by SDS-PAGE. As shown in Figure 37A, the amount of complex formed in the presence of 2.5 mM ATP turns out to be higher compared to the amount formed in the presence or absence of 0.5 mM H_2O_2 as oxidizing agent. This outcome suggests that complex formation is preferentially induced by the presence of the nucleotide. Furthermore, in the presence of catalytically inactive forms of either of the enzymes where the catalytic cysteine was mutated to alanine or a reducing agent like β-mercaptoethanol was added, complex formation was abrogated owing to its dependence on the disulfide bridge. A previously published report showed that a disulfide-linked Uba1-Ubc3 forms upon oxidative stress, however, when several other E2 enzymes were tested in the same study they did not form a corresponding complex (Doris et al., 2012). Therefore, to test whether Ubc3 shows a higher sensitivity for disulfide bond formation in the

presence of oxidizing agent, we repeated the same experiment except this time with Ubc3 as the E2 enzyme. The result showed higher intensity of the E1-E2 complex band in the presence of ATP, however, even in this case, the presence of H_2O_2 did not promote disulfide bond formation as dramatically as seen with the nucleotide (Figure 37B). This outcome is somewhat contradictory to the previously published report and shows limited ability of E1-E2 disulfide bridging in response to oxidative stress (0.5 mM H_2O_2) *in vitro*.

We tested all thirteen E2 enzymes present in *Saccharomyces cerevisiae* to investigated influence of both ATP or H_2O_2 individually and together in promoting Uba1-E2 complex formation using similar SDS-PAGE analysis. The experiment was performed in duplicates and band intensities of the resulting complexes were quantified by densitometric analyses.

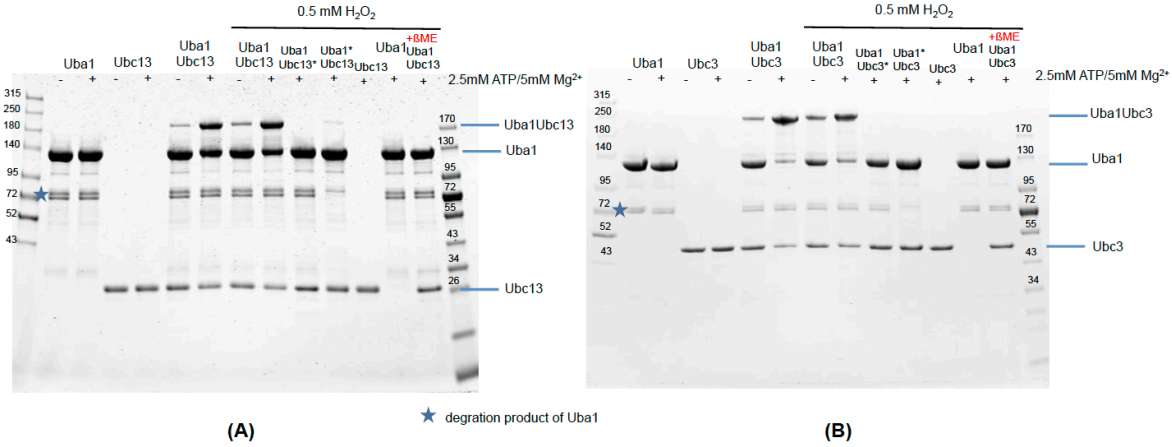


Figure 37: SDS-PAGE analysis of E1-E2 complex formation in the presence of either Mg-ATP or H_2O_2. Catalytically inactive mutant of the enzyme where the catalytic cysteine is replaced with an alanine are indicated by . "*" The blue lines on the right side show the expected height on the gel for either the individual proteins or the respective complexes.

The averaged intensities for individual E1-E2 pairs are presented qualitatively in Table 8. The presence of ATP always yielded the highest amount of the respective complex and the presence of both ATP and H_2O_2 did not increase the yield of the complex. The only exception was Ubc7 where together ATP and H_2O_2 resulted in a slightly higher level of complex formation compared to ATP alone. The exclusive effect of oxidizing conditions were tested for Ubc3, Ubc6 and Ubc13 and none of them showed more complex formation in comparison to the amount induced by ATP. The best result for the complex induced by ATP was seen for the Uba1-Ubc3 pair, followed by similar extents of complex formation seen for the Uba1-Ubc2, Uba1-Ubc10 and Uba1-Ubc13 pairs. Both Ubc9 and Ubc12 were checked as controls as these E2 enzymes belong to the SUMO and NEDD8 system, respectively. The complete loss of complex formation for

these E2 enzymes confirms no cross-reactivity between the Ub E1 and the E2s for SUMO and NEDD8 conjugation. Although the nucleotide was not visualized in the structure of the Uba1-Ubc13 complex, these experiments confirmed that the presence of Mg-ATP drastically induces E1-E2 disulfide bridging for certain pairs. The induction of complex formation by ATP can be appreciated even for the pairs where the shift of Uba1 to Uba1-E2 is not as strong since in those cases the amount of complex is even lower in the absence of Mg-ATP. This observations indicates clear correlation between presence of ATP and more E1-E2 complex formation.

UBA1-E2 complex	**-ATP/ -H_2O_2**	**+ATP**	**+ATP/+ H_2O_2**	**+ H_2O_2**
Ubc1	**–**	**+**	**+**	**N.D.**
Ubc2	**+**	**++**	**++**	**N.D.**
Ubc3	**++**	**+++**	**+++**	**++**
Ubc4	**–**	**+**	**+**	**N.D.**
Ubc5	**–**	**+**	**+**	**N.D.**
Ubc6	**+**	**+**	**+**	**+**
Ubc7	**(+)**	**+**	**++**	**N.D.**
Ubc8	**–**	**(+)**	**+**	**N.D.**
$Ubc9^{SUMO}$	**–**	**–**	**–**	**N.D.**
Ubc10	**+**	**++**	**++**	**N.D.**
Ubc11	**–**	**+**	**+**	**N.D.**
$Ubc12^{NEDD8}$	**–**	**–**	**–**	**N.D.**
Ubc13	**+**	**++**	**++**	**+**

Table 8: Qualitative representation of the effect of ATP, H_2O_2 or their combination on Uba1-E2 disulfide bond formation. The lowest intensity of complex is represented by – (minus sign). +, (+) and ++ show higher yields of complex in this order and +++ shows the maximum amount of E1-E2 complex formation. N.D. refers to levels not determined for certain E2s.

2.4.12 Complex formation does not require ATP hydrolysis

To investigate the impact of the nucleotide on complex formation, we tested four purine and pyrimidine nucleoside triphosphates and checked which nucleotide enforced the Uba1-Ubc13 disulfide the most. As seen in Figure 38A, with ATP, the natural substrate of the enzyme, a significant shift of Uba1 into the Uba1-Ubc13 complex occurred. We also checked whether ATPαS, a non-hydrolysable analog of ATP, which prevents ATP to AMP hydrolysis, results in complex formation as observed for ATP. Surprisingly, ATPαS acts as an even better inducer of this complex as reflected by the higher intensity of the Uba1-Ubc13 band in this case. This means that the mechanism behind E1-E2 complex formation does not require hydrolysis of ATP, instead its presence seems to induce a conformation that promotes complex formation. To further confirm this observation, we checked for ATP hydrolysis using the BIOMOL green assay via the release of phosphate in the reaction. When incubated with either ubiquitin or ubiquitin and Ubc13, Uba1 catalyzed the reaction resulting in the release of pyrophosphate, which could be detected in this assay by the action of pyrophosphatases, however, when Uba1 and Ubc13 were incubated together with Mg-ATP resulting in complex formation no phosphate production was detected and hence no pyrophosphate was generated during the reaction (Figure 38B). This result confirms that nucleotide hydrolysis is not required for E1-E2 complex formation.

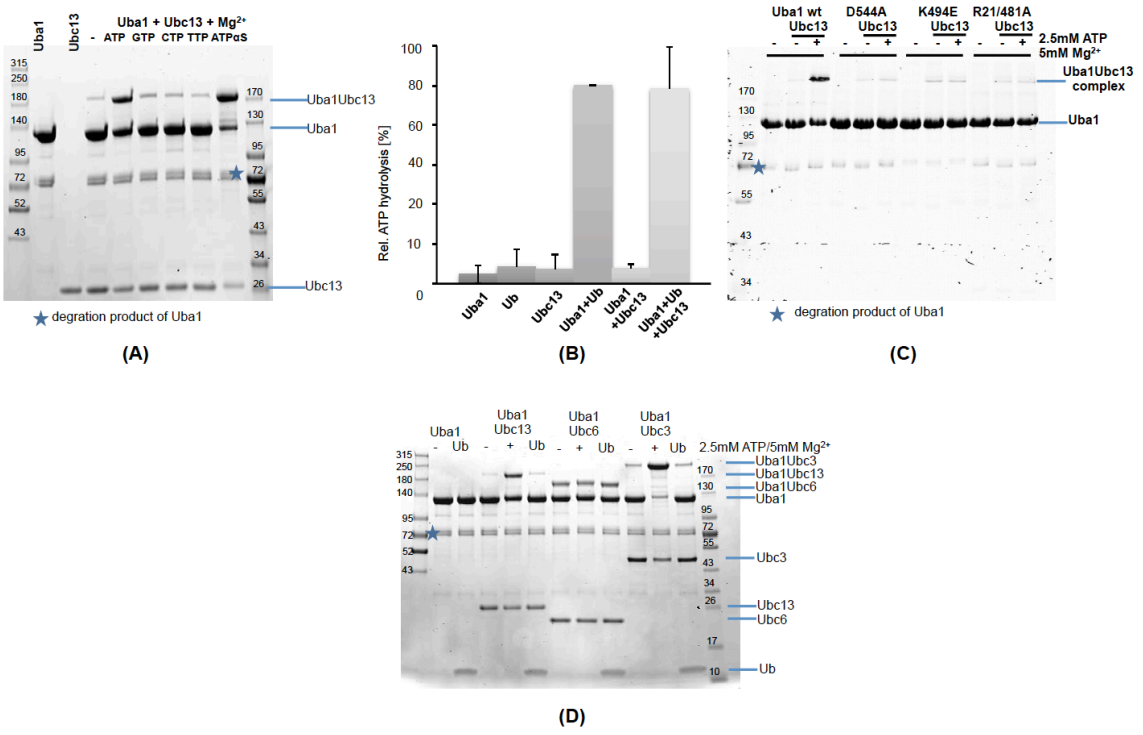


Figure 38: ATP-binding in the relevant binding pocket is required for E1-E2 complex formation. (A) Effect of various nucleotides on Uba1-Ubc13 complex formation. (B) Results of the BIOMOL green assay to check whether ATP

hydrolysis is a requirement for Uba1-Ubc13 complex formation. (C) ATP-binding mutants were tested to confirm the binding of ATP to the enzyme as a prerequisite for complex formation. (D) Influence of either ATP or ubiquitin on Uba1-Ubc13 complex formation.

2.4.13 ATP-binding mutants are deficient in complex formation

Given the positive influence of ATP on complex formation it is still puzzling why we did not detect the nucleotide at the ATP-binding pocket of the enzyme in the crystal structure. To test whether binding of ATP to the nucleotide binding pocket of Uba1 is in fact a prerequisite for complex formation, we utilized those ATP-binding mutants that showed an almost complete loss of the ATPase activity of the enzyme as described earlier in section 2.4.5. These mutants were D544A, K494E and the R21/ 481A double mutant. Figure 38C shows that complex formation is impaired for these mutants, even in the presence of ATP, indicating that complex induction is indeed a result of ATP binding to the corresponding binding site in the enzyme.

Finally, we tested whether ubiquitin binding at the adenylation site could also enforce formation of the complex as seen in the presence of Mg-ATP. Figure 38D suggests that the presence of only ubiquitin is not as effective as the addition of Mg-ATP for inducing the assembly of the Uba1-Ubc13, Uba1-Ubc6 or Uba1-Ubc3 complexes.

2.4.14 R21 is involved in a three-way cross-talk

All the aforementioned experiments refer to the fact that ATP-binding induces conformational changes in Uba1 that in turn facilitate the formation of the E1-E2 disulfide bridged complex and the role of ATP in this process does not include its hydrolysis. As reported for active site remodeling of both SUMO E1 and *Sp*Uba1 (Lv et al., 2017b; Olsen and Lima, 2013) (Figure 30-31), in the closed conformation the N-terminal helix carrying the arginine finger R21 is disordered and as seen in the *Cht*Uba1 structure (see Fig. 42B) this rearrangement moves the N-terminal region away from the ATP-binding site to avoid clashes with the incoming catalytic cysteine domain. Furthermore, the loop lining one face of the ATP binding pocket that contains R481, another residue critical for ATP-binding is in a different orientation when compared to open state of the enzyme. These observations suggest that ATP-binding to Uba1 most likely favors its open conformation as in the closed conformation both R21 and R481 are away from the ATP binding site and from the BIOMOL green assay we know that the the R21A/R481A double mutant is impaired in ATP-hydrolysis. The binding of an E2 to Uba1 requires the latter to be in its open conformation as this state positions the SCCH domain to be able to interact with the E2 enzyme. All crystal structures reported so far for the yeast E1-E2 complexes represent the open conformation of Uba1 where

the E2 is buried in the central canyon between the UFD and SCCH in each structure. Combining these observations suggests that ATP-binding results E1-E2 complex formation, since it promotes the open conformation of Uba1, which favors E2 binding.

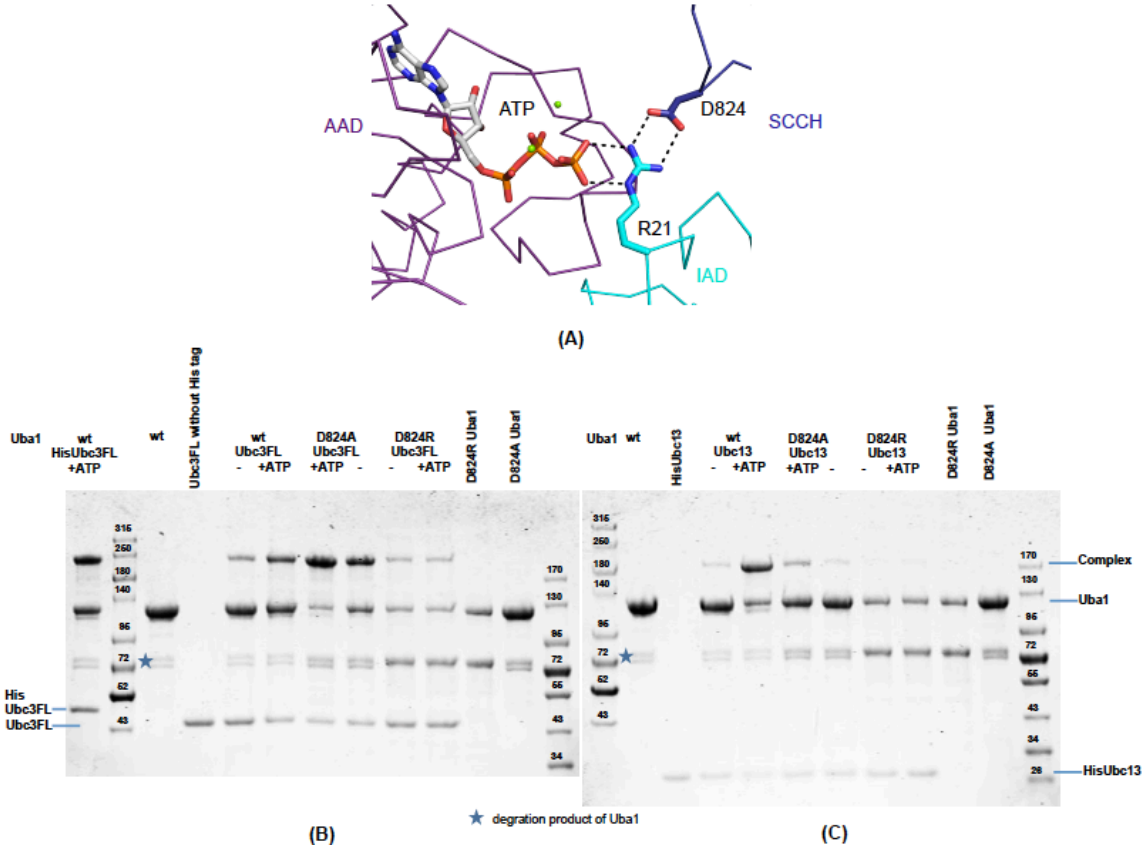


Figure 39: The coordination of R21 by ATP and D824 in the Mg-ATP-bound Uba1 structure. D824A and D824R mutants were tested for their ability to form Uba1-E2 complexes for (B) Ubc3 and (C) Ubc13.

Upon closer inspection of the Mg-ATP-bound Uba1 structure for clues that may explain the ATP-induced complex formation, we focused on residue D824, which is involved in a salt bridge with R21 in this structure. Figure 39A represents the three-way communication involving the N-terminal residue R21 originating from the inactive adenylation domain, the ATP bound at the active adenylation domain and residue D824 located near the C-terminus of SCCH domain. We suspected that the interaction of R21 and D824 is a critical contact keeping the catalytic cysteine domain in the open conformation. To check this we mutated this residue to both an alanine and an arginine and evaluated the ability of these mutants in the formation of the Uba1-Ubc3 or Uba1-Ubc13 complexes. D824R mutant, being a more disruptive mutation, showed impaired complex formation for both E2s. While the D824A mutation did not effect Uba1-Ubc3 complex formation, Uba1-Ubc13 complex formation was reduced to the level of

complex formation seen in the absence of ATP. The fact that the same Uba1 mutant behaves differently with two E2 enzymes suggests a variation in the binding modes of these two E2 enzymes. Ubc3, also referred to as Cdc34 belongs to the class IV of E2 enzymes which posses both an N-terminal and a C-terminal extension. In addition, this E2 enzyme possesses a Ubc15-like acidic loop insertion near its active site cysteine. These additional features of Ubc3 may be involved in generating a more extensive binding interface with Uba1 as reflected in the strongest level of Uba1-Ubc3 complex formation in the presence of ATP in comparison to any other E1-E2 pair.

In an earlier report it was shown that binding of an E2 enhances ubiquitin adenylate formation (Tokgoz et al., 2006). The above explanation of ATP-induced complex formation also supports this observation as binding of E2 to Uba1 keeps the latter in open conformation, thus keeping the adenylation active site intact to catalyze formation of the ubiquitin-adenylate.

2.4.15 UFD domain motion

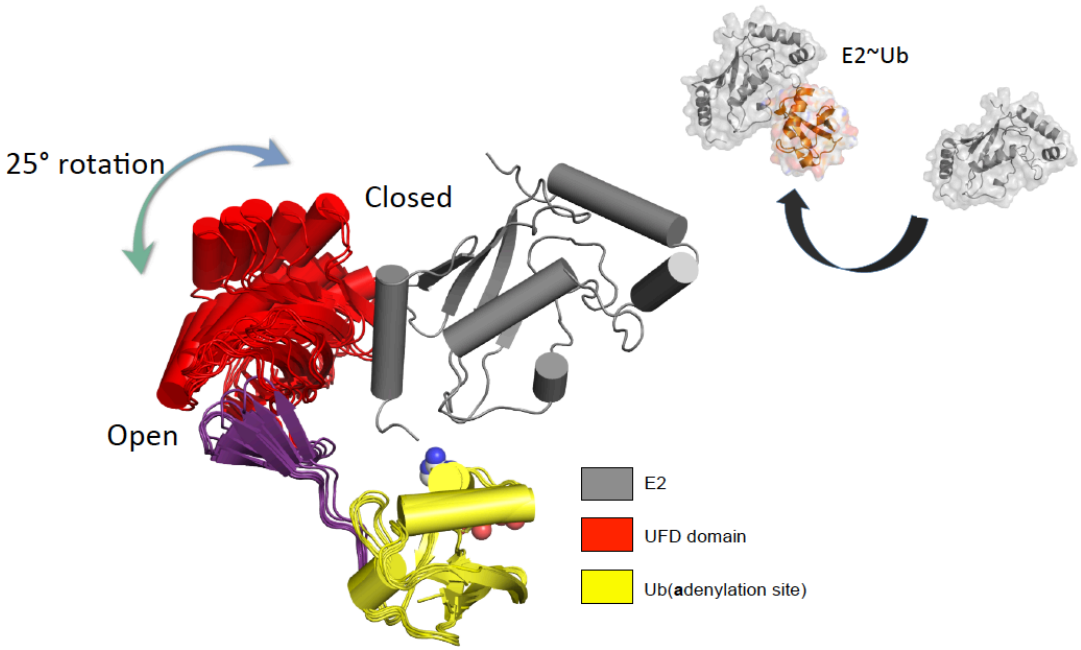


Figure 40: Rotation of the UFD as seen in several Uba1 structures ranging from closed (or proximal) to distal (or open) conformation. The flexibility of the UFD is important for E2 recruitment and transthioesteridication.

As mentioned earlier, the UFD has been observed to show conformational flexibility in the reported Uba1 structures. The two extreme positions of UFD visualized so far are termed proximal (or closed) and distal (or open) conformations. In the proximal conformation the UFD is close to the SCCH domain and in the distal conformation it is remote from it, leaving a wide space between the two domains. All reported Uba1

structures fall onto the trajectory between these two states. Interestingly, both the Mg-ATP-bound and E2-bound structures show the UFD in its closed conformation while the ubiquitin-bound structures present the UFD in several distal conformations, characterized by rotations varying from 10-25° degrees away from the proximal conformation (Figure 41). The consistency of this feature makes it tempting to speculate that ubiquitin-binding to the E1 enzyme widens the canyon between the UFD and the SCCH domain to accommodate the E2 enzyme. Upon superposition of Mg-ATP bound, ubiquitin-bound and Ubc13-bound structures, we realized that the hinge point for the rotation resides at a ubiquitin interaction site on the AAD and this rotation translocates the UFD, thus forcing it into the distal conformations (Figure 41B). This interaction involves the K48 side chain and G47 main chain of ubiquitin and E912 of Uba1 (Figure 41C).

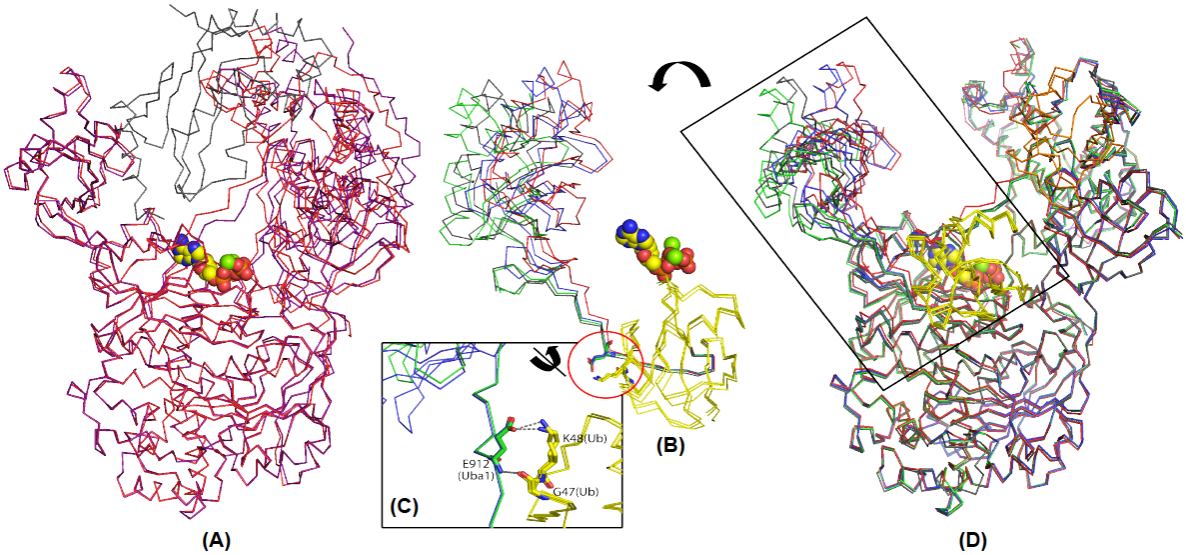


Figure 41: (A) Superposition of Mg-ATP-bound Uba1 structure in red and the Uba1-Ubc13 complex in blue depicting the proximal conformation of the UFD in both cases. (B) Narrowing down the motion of the UFD in the ubiquitin-bound states shown in either grey or green and (C) close-up view of the interaction site in the Uba1-Ubc13 complex (blue) and one of the ubiquitin-bound structures (green) with ubiquitin (in yellow). (D) Superposition of Uba1 in complex with Mg-ATP, Ubc13 and ubiquitin featuring the relatively open conformations of the UFD in the ubiquitin-bound complex.

Complexes of UFD with the respective E2 enzymes have been reported for the SUMO and NEDD8 systems. When we superimpose the recently published structure of human UBC9 in complex with the SUMO E1-UFD (Liu et al., 2017a)onto the full-length enzyme based on the common UFD domain present in both structures it becomes clear that instead of being present between the UFD and SCCH domain the E2 is projecting outwards (away from the SCCH) suggesting the requirement of a different orientation of

the UFD for the juxtaposition of the catalytic residues of SUMO E1 and UBC9. We could solve a crystal structure of Uba1 derived from *Chaetomium thermophilum* in complex with Ubc4 from the same organism (Figure 42B). Although the structure was only derived at a low resolution of 4.4 Å, it has three distinguishable features. First, this structure has the catalytic cysteine domain in its closed conformation. Second the N-terminal helices, which have been shown to be disordered in the closed conformation of E1s, are visible in this structure in an altered arrangement which aids SCCH domain movement. Third the UFD of *Ct*Uba1 is in the outward-facing conformation when in complex with *Ct*Ubc4. The later feature highlights the role of the UFD in the recruitment of the E2 enzyme, and this is the only published structure of a Ub E1 which has the E2 interacting only with UFD while being distal from the SCCH domain. When we superimposed the isolated Ubc13-UFD from the Uba1-Ubc13 complex with the chaetomium Uba1 complex (superimposition based on UFD) we see that the orientation of the Ubc13-UFD is similar to that observed in the *Ct*UFD-Ubc4 complex (Figure 42A-B). This observation suggests that in a three dimensional frame the rotation of UFD domain is not restricted to the axis coinciding with the viewing direction observed in the yeast Uba1 complexes (Figure 41), but it can also rotate around an axis perpendicular to the viewing direction as seen in the *Ct*Uba1-Ubc4 complex. The latter rotation could potentially represent an important role of the UFD in the recruitment of an E2 enzyme. Together with the rotation around the axis defined by the viewing direction the E2 can be presented to the catalytic cysteine domain by the UFD to provide an additional binding interface and to allow charging of the E2 with ubiquitin.

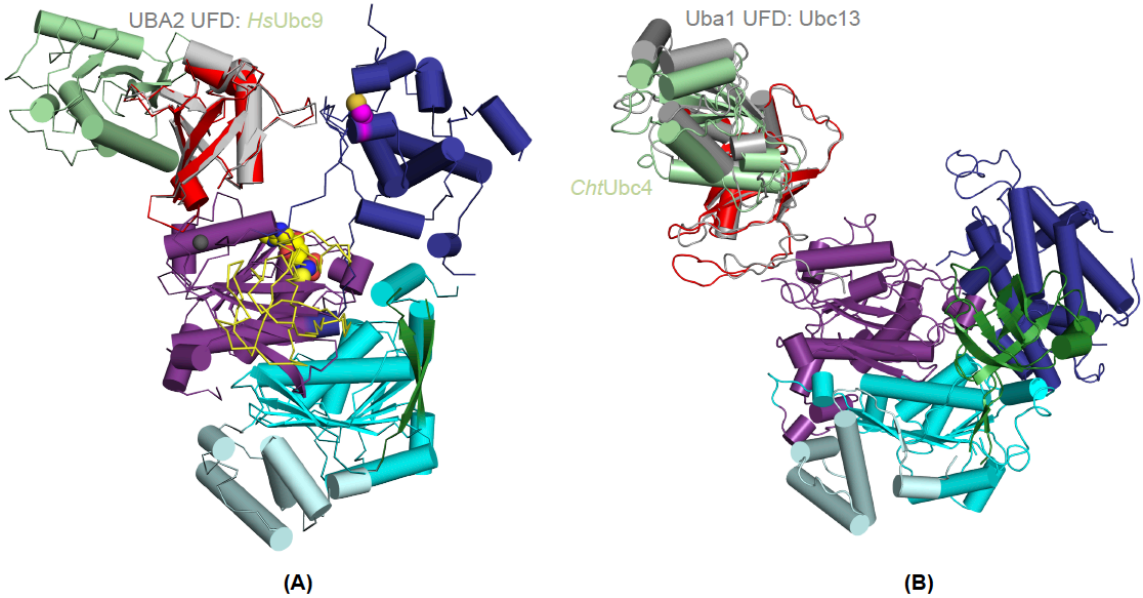


Figure 42: (A) Structural alignment of the SUMO E1 [PDB code: 3KYC] with a crystal structure of the UBA2 UFD (grey) in complex with Ubc9 (pale green) [PDB code: 5FQ2]. (B) Superimposition of the structure of *Ct*Uba1 bound to *Ct*Ubc4

(pale green) with an isolated complex of UFD-Ubc13 (grey) from the Uba1-Ubc13 structure.

Although we cannot rule out that crystal packing forces contribute to the various conformations of the UFD, the correlation between the binding of a specific substrate and the orientation of UFD is quite striking. Notably, the UFD motion of both the SUMO and NEDD8 E1 does not appear to be as flexible as seen in UBA1 due to the zinc coordinating site between the crossover loop and UFD linker that staples these two elements together. However, the structure of a trapped activation complex containing APPBP1-UBA3, two NEDD8s (one NEDD8 thioesterified to UBA3's catalytic C216 and the other NEDD8 noncovalently associated at the adenylation site), Mg-ATP and a catalytically inactive mutant of Ubc12 revealed a striking ~120° rotation of the E2 bound UFD (PDB code: 2NVU). This rotation can be the result of clash arising from the thioesterified NEDD8 (Huang et al., 2007). Therefore, the mechanism of UFD motion between NEDD8/SUMO E1 and UBA1 seems to be different where the former is regulated by thioesterified NEDD8/SUMO, the latter is under control of adenylation site ubiquitin.

2.4.16 A second E2-binding site in the E1

In the presence of Mg-ATP we obtained another crystal structure of Ubc13 bound to Uba1 at a resolution of 2.35 Å. Surprisingly in this structure there were two E2 molecules bound to a single Ub E1. The first Ubc13 was seen in a similar position as seen in the previously mentioned structure, i.e. it was held in place by both the UFD and the SCCH domain in the open conformation of the enzyme (Figure 35A). The second Ubc13 molecule was found residing at the adenylation site of Uba1, thus overlapping with the ubiquitin-binding site (Figure 43B). The appearance of the second Ubc13 on Uba1 probably reflects the E2 concentration in the crystallization experiment where Uba1 and Ubc13 were mixed in a 1:2 molar ratio in the presence of Mg-ATP. The other Ubc13-bound structure was determined using a 1:1 molar ratio, thus the second binding site was vacant. This suggests that the binding of the E2 between the SCCH and UFD is the preferred binding site on the Ub E1, exhibiting a higher affinity to this position. Moreover, the stability of this complex is further enhanced by the formation of the disulfide bridge between the catalytic cysteines of the E1 and E2.

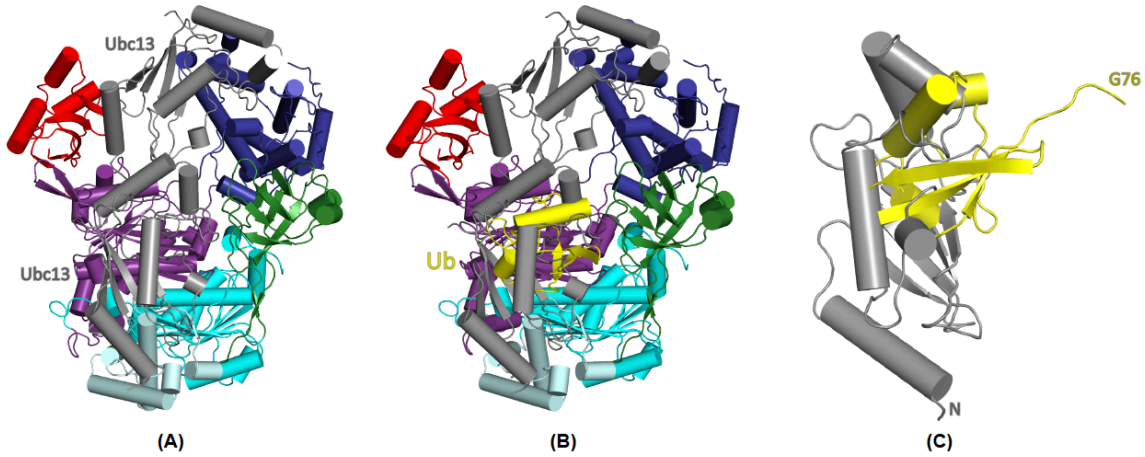


Figure 43: (A) Crystal structure of *Sc*Uba1 doubly loaded with *Sc*Ubc13. The two Ubc13 molecules are colored in grey. (B) Superimposition of this structure with the binding site of ubiquitin at the adenylation site indicates a competition between Ubc13 and ubiquitin for adenylation site binding. (C) Close-up view of the mutually exclusive binding of ubiquitin and Ubc13 to *Sc*Uba1.

Calculation of the interface areas between the covalently bound Ubc13 and Ubc13 located adenylation site with Uba1 using the PDBePISA server (Krissinel and Henrick, 2007) resulted in buried areas of 1800 Å^2 and 1200 Å^2, respectively. While only accounting for two thirds of the interface area observed for the canonical E1-E2 contact, this is still a rather extensive interaction surface for the second Ubc13, in excess of the 1000 Å^2 limit often taken as a threshold to distinguish molecular interaction surfaces from crystal contacts (Chen et al., 2013). In comparison, the interface area between ubiquitin bound at the adenylation site and Uba1 is somewhat more extended with 1600 Å^2.

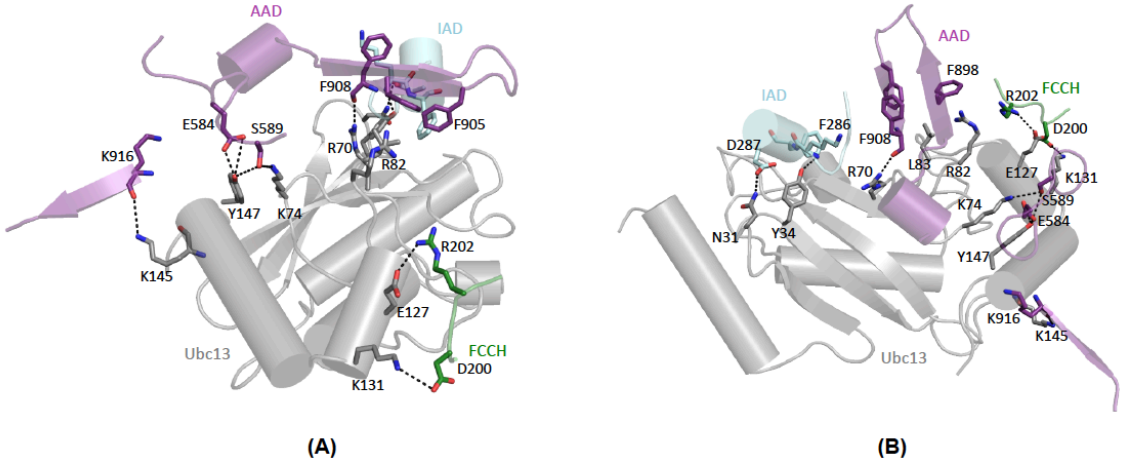


Figure 44: Interactions of Ubc13 bound at the adenylation site with Uba1 in two different orientations (A and B). Ubc13 is colored in grey and three domains of Uba1, IAD, FCCH and AAD, are colored in cyan, green and magenta, respectively.

Ubc13 bound at the adenylation site is involved in a threeway interaction with the FCCH domain, IAD and AAD of Uba1. The residues involved in the interaction are located in the β1-β2 loop, β4 strand, the loop containing the catalytic cysteine residue and the C-terminal helices α3 and α4. N31 and Y34 at the β1-β2 loop interact with D287 and F286 of the IAD. R70 and K74 hydrogen bond with F908 and S589 of the AAD, whereas L83 contacts F905 via van der Waals interactions. R82 appears to be forming a cation-π interaction with the side chain of F898. E127 and K131 from H3 and K145 and Y147 from the C-terminal helix are involved in a number of electrostatic interactions with the FCCH and AAD. Superimposition of the two Ubc13 molecules results in an RMS deviation of 0.40 Å with structural changes in the N-terminal region, the loop between β1-β2 and another loop preceding helix α3 of Ubc13 (Figure 42). Notably, the Ubc13 present in the canonical E2 binding site is not linked to Uba1 via a disulfide bridge, most likely as a result of radiation damage during X-ray data collection.

Strikingly, Wee et al. reported the existence of a second Ubc4 binding event to human Ub E1 while performing steady-state kinetics of E1 catalyzed Ubc4~Ub formation as a function of Ubc4 concentration (Wee et al., 2000). At higher Ubc4 concentrations, a decrease in the reaction velocity was detected which suggested substrate inhibition. The authors found that the K_m for productive Ubc4-binding and the K_i for inhibitory binding were in the same range, indicating similar binding affinities for the two sites on Uba1. In another report in the SUMO system, Wang et al. showed inhibition of SUMO adenylation at increased concentrations of Ubc9, which is in line with our structure, since binding of an E2 (Ubc13 in our structure and Ubc9 in the mechanistic study of Wang et al.) at the adenylation site will block ubiquitin or SUMO adenylation, respectively (Wang et al., 2010b). These earlier reports are in line with a second E2 binding event to the E1, however, this needs further confirmation.

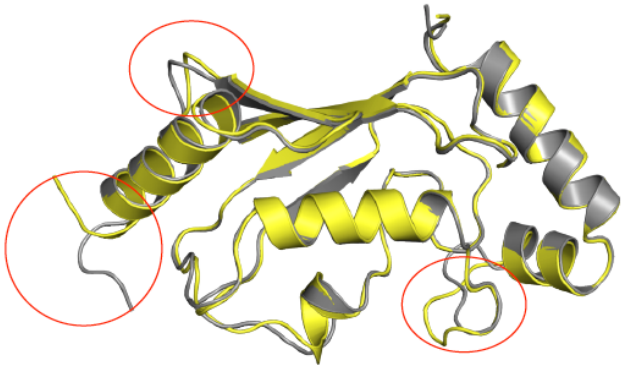


Figure 45: Structural alignment of Ubc13 (yellow) bound at the adenylation site with Ubc13 (grey) present as a disulfide linked substrate between the SCCH and UFD domain. The latter is the catalytically relevant position for E2 binding. Differences between the two structures are highlighted by red circles.

2.5 CONCLUSION

In this chapter, the molecular details of the catalytic cycle of Uba1 with the help of six unpublished structures and several published structures reported for *Sp*Uba1 and the SUMO E1 are presented. These structures provide valuable insights into the binding of all three substrates of Uba1 (ubiquitin, Mg-ATP and E2) and the Uba1-mediated formation of the products ubiquitin adenylate and thioesterified ubiquitin. Furthermore, how Uba1 utilizes movement of the UFD for the recruitment of an E2 enzyme and how the E2 positions itself in the central groove located between the UFD and the SCCH domain to bring its catalytic cysteine into close spatial proximity of the C600~ubiquitin (C632~ubiquitin in case of human Uba1) thioester to enable transthioesterification can be illuminated with the help of these structures. The E1 enzymes are very efficient enzymes as suggested by their catalytic turnover of ~2 per second for E1~Ubl formation. The multi-domain architecture of Uba1 supports its efficiency when catalyzing two separate reactions, in addition to maintaining its ability to transfer the activated ubiquitin to the cognate E2s with high fidelity. Flexible linkers connecting the domains play a crucial role in achieving an interdomain cross-talk. The UFD is linked to the AAD, the site for binding of both ubiquitin and ATP by an 18-residue loop forming a β-hairpin at the end of the AAD. The FCCH is linked to the IAD by two long antiparallel β-strands, whereas the SCCH domain is connected to the AAD by two linkers, the crossover loop, an extended 18-residue linker that traverses 40 Å from one side of the molecule to the other, and the shorter reentry loop. These linkers serve as sensors for the intramolecular communication and enable movement of the domains for the capture of a product formed on one domain of the enzyme by another domain. Another important feature of the catalytic activity of the E1 enzyme is its differential affinity for the products. The first catalytic half reaction results in the ubiquitin-adenylate and pyrophospahate. The former binds very tightly to the AAD as it is the substrate for the second catalytic half reaction while the later is released accompanied by remodeling of the adenylation active site. This serves as a measure to control product inhibition and keep the reaction moving in the forward direction. The same implies for the low affinity of AMP which is produced as a result of the thioesterification of ubiquitin. This ensures binding of the next ATP molecule in the nucleotide-binding pocket to move the reaction cycle forward. Thus, the E1 enzyme works as a well-devised nanomachine, which provides sufficient amounts of activated ubiquitin to the rest of the ubiquitylation machinery.

A significant number of Uba1crystal structures discussed in this chapter provide insights into the various catalytic states of the enzyme. This enables us to visualize the catalytic cycle of Uba1 in significant detail. When we put these pieces together we can imagine

the spatiotemporal distribution of Uba1 states under variable substrate concentrations. ATP does not appear to be a limiting substrate for Uba1 as it is the basic cellular energy source. However, the concentrations of ubiquitin or the different E2 enzymes can vary based on their transcriptional and translational rates, which can be altered by a number of stress conditions (Hanna et al., 2007). Furthermore, the availability of free ubiquitin depends on three factors: 1) The expression levels of ubiquitin, 2) the abundance of deubiquitinases that generate mature ubiquitin and 3) the functioning of the ubiquitin recycling element that maintains the free pool of ubiquitin in the cell. Taking into account the abundance and shortage of either ubiquitin or E2, the various routes of the Uba1 catalyzed reaction are presented in Figure 46.

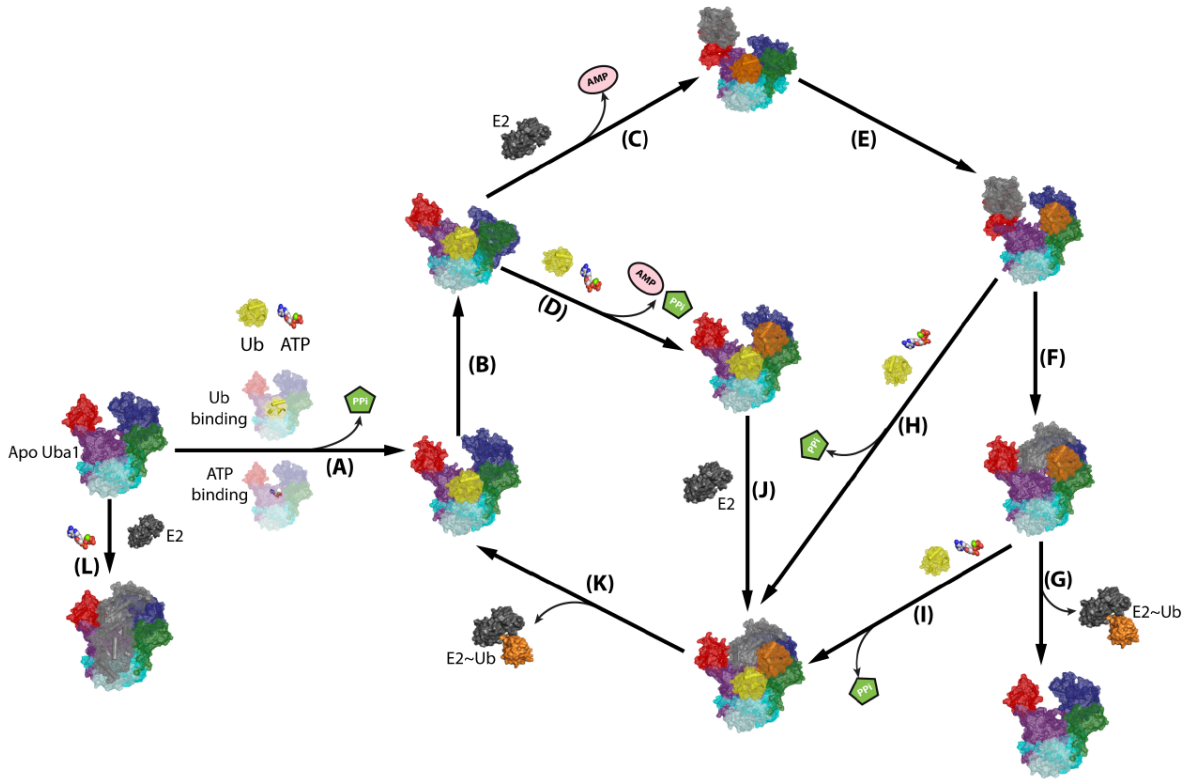


Figure 46: Catalytic cycle of the ubiquitin activating enzyme.

Although the binding of both ubiquitin and ATP to Uba1 is independent, when both substrates occupy their respective position at the adenylation site, the AMP•ubiquitin covalent adduct is formed as shown by reaction step (A). Pyrophosphate is released in this process. The low affinity of pyrophosphate facilitates its release together with a magnesium ion. The lack of an intact ATP in the pocket no longer required the β and γ-phosphate coordinating residues to be in their nucleotide-binding positions and favors the remodeling of the active site to facilitate the nucleophilic attack of the incoming catalytic cysteine along with the rotation of the entire SCCH domain in reaction (B). The closed conformation of apo *Sp*Uba1 suggests that the open and closed conformations

of Uba1 exist in equilibrium in the absence of its substrates, however, whether this motion is enough to reach the C-terminus of ubiquitin or the binding of ubiquitin or the ubiquitin adenylate enables extension of the catalytic cysteine to the adenylation active site, is still unclear. When E2 enzymes are present in significant amounts, the UFD can catch the N-terminal helix of the E2 while the thioesterification reaction is in progress as seen in the *Ct*Uba1-Ubc4 complex (Reaction C). Upon achieving the tetrahedral intermediate state, AMP is released and ubiquitin is captured via a thioester bond. The movement of Uba1 back into the open conformation as shown in reaction E can provide the E2 with the second binding interface for the transthioesterification (Reaction F) resulting in the charging of the E2 with ubiquitin and its release from Uba1 due to the clash of ubiquitin with Uba1 in the conformationally flexible E2~Ub configuration (G). However, in the absence of an E2 (reaction D) Uba1 can load itself with the second ubiquitin, which can undergo the adenylation reaction like in reaction A. Once formed the doubly loaded Uba1 can bind the E2 as shown in reaction (J) and with the release of the E2~Ub conjugate (reaction K) Uba1 returns back to the resulting state of reaction (A). E2 binding and the second ubiquitin adenylation can happen simultaneously at saturating concentrations of both ubiquitin and E2 (reaction H or I). In the event of ubiquitin stress and comparative abundance of E2, Uba1 can lock itself in an inhibited state by sequential binding of a first E2 via a disulfide linkage and the second E2 at the adenylation site. However, as Uba1 is a cytosolic protein and therefore is present in a reducing environment, a disulfide linkage of E1-E2 pair would require oxidative stress as reported previously. Remarkably, ATP-binding appears as a switch for the open conformation of Uba1 which may facilitate reactions (E) and (F). Altogether, we observe alternating conformational states of Uba1 linked to one type of reaction where adenylation takes place in the open conformation, thioesterification in the closed conformation while transthioesterification requires the open conformation again. The fluctuation between these two states drives the catalytic cycle of Uba1.

Chapter 3: Targeting the ubiquitin activating enzyme for cancer therapy.

3.1 ABSTRACT

The UPS is required to tightly control a number of cellular processes including cell cycle control and protein quality control. Hence it is not surprising that anomalies in the system are observed in case of abnormal cell growth or cancer (Bedford et al., 2011; Nalepa et al., 2006). In malignant cells, a higher UPS activity is encountered suggesting its role in supporting a higher metabolic rate. Therefore, targeting the UPS is seen as a therapeutically useful approach for the treatment of various forms of cancer. Indeed, three proteasome inhibitors have been approved for the treatment of multiple myeloma (Teicher and Tomaszewski, 2015). Inhibition of UBA1 the master regulator of the ubiquitylation pathway thus holds promise to provide a therapeutic avenue for cancer therapy. Moreover, in contrast to the proteasome, which only executes the degradative route of ubiquitylation, UBA1 controls every cellular ubiquitylation event, which will result in a more drastic cellular response upon inhibition. Previous studies have shown that inhibition of the E1 enzyme preferentially induces cell death in malignant cells over normal cells (Xu et al., 2010). Furthermore, tumor growth in a mouse model of leukemia could be delayed by injecting the UBA1 inhibitor PYZD-4409. These results encourage the development of potent UBA1 inhibitors as an anti-cancer strategy. In addition, these inhibitors can help overcome resistances which develop when patients are treated with proteasome inhibitors.

Although several inhibitors of UBA1 have been reported, most of them suffer from low inhibitory potency and unspecific activity. However, the adenosyl sulfamate class of inhibitors was found to display a high specificity against E1 enzymes and two inhibitors from this class are currently in Phase I clinical trials for the treatment of solid tumors. We obtained crystal structures of three adenosyl sulfamates, MLN7243, ABPA3 and MLN4924, covalently linked to Ub and in complex with yeast Uba1 (Misra et al., 2017). The yeast enzyme shares 65% sequence identity in its adenylation domain with human UBA1 and is thus a valid model system for the investigation of inhibitors targeting UBA1. We carried out *in vitro* biochemical assays to derive the potency of these inhibitors against UBA1 and validated yeast Uba1 as a good model for UBA1 to structurally characterize these compounds for designing next generation inhibitors. In addition, we carried out computational approaches to simulate inhibitor flexibility in the binding pocket as well as to estimate binding energies of the inhibitors, which nicely corroborate our experimental data. The *in silico* approach to profile the potency and selectivity of E1 enzyme inhibitors suggests an avenue to test adenosyl sulfamates derivatives ahead of their synthesis. The structures together with biochemical assays

and computational calculations rationalize inhibitor potency and selectivity and provide important clues for the future design of highly specific E1 inhibitors.

3.2 INTRODUCTION

The degradative machinery of the cell is comprised of two components: The UPS and the autophagy-lysosomal system (ALS) (Saeki, 2017; Zhan et al., 2016). Whereas UPS monitors the degradation of short-lived cellular proteins containing an oligo-ubiquitin tag, ALS acts as a disposal system for long-lived proteins, large chunks of aggregates as well as organelles like mitochondria. Both UPS and ALS were considered as independent pathways for a long time but recent studies suggest that ubiquitylation can target substrates for degradation via both pathways (Fan et al., 2016). The coordination between these two pathways is required for the maintenance of cellular metabolism. Under conditions of stress, the UPS cannot take the burden of unusual amounts of toxic misfolded proteins. To overcome this, aggreagation-prone proteins are transported along microtubules in a retrograde fashion to the microtubule-organizing center to be sequestered into a single large cellular garbage bin-like structure called the “aggresome” (Johnston et al., 1998; Rodriguez-Gonzales, 2009). Histone deacetylase 6 plays an important role during aggresome formation because it can bind to both polyubiquitinated proteins and the dynein proteins, thereby recruiting protein cargo to the dynein motor to transport misfolded proteins to aggresomes. In addition, it regulates microtubule dynamics by deacetylation of α-tubulin (Kawaguchi et al., 2003).

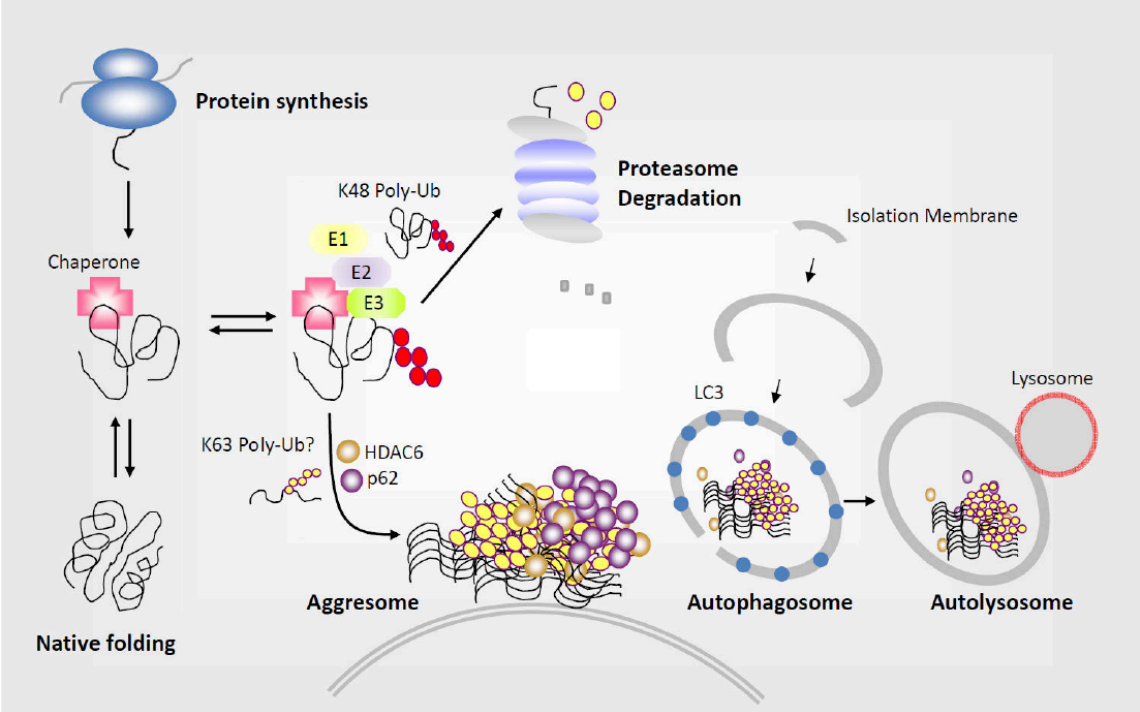


Figure 47: Two routes for the degradation of misfolded proteins: Proteasomal degradation and aggresome formation followed by degradation by the autophagy-lysosomal system.

The clearance of aggresomes is mediated by ALS and involves specific adaptors, including the p62/SQSTM-1/sequestosome (Liu et al., 2016). The autophagic adaptor p62 has a UBA domain that interacts with polyubiquitin chains of misfolded proteins and a PB1 domain that mediates self-aggregation to form condensed cargo-p62 complexes. Cargo-loaded p62 and its aggregated complexes are delivered to autophagic vacuoles through the specific interaction of p62 with Light Chain 3 II (LC3-II), an active form of LC3, on the surface of autophagic double membrane structures (Bjorkoy et al., 2006). Interestingly, both p62 and HDAC6 exhibit a preference for K63-linked polyubiquitin chains, suggesting that this form of ubiquitin modification may underlie the formation as well as autophagic degradation of protein aggregates (Tan et al., 2008). Formation of aggresomes and its disposal by ALS reduces the burden on UPS for clearing an overload of misfolded proteins.

Cancer cells produce proteins that induce both cell survival and proliferation, and/or inhibit mechanisms of cell death. This notion set the stage for preclinical testing of proteasome inhibitors as a means to shift this fine equilibrium towards cell death. Moreover, malignant cells are more dependent on the proteasome to remove misfolded or damaged proteins due to their genetic instability and rapid proliferation (Deshaies, 2014). Early studies showed that proteasome inhibitors induced apoptosis in leukaemic cell lines and were active in an *in vivo* model of Burkitt's lymphoma. Further *in vitro* investigations demonstrated that proteasome inhibition displayed a broad spectrum of anti-proliferative and pro-apoptotic activity against hematological tumors (Crawford et al., 2011). The proteasome inhibitor bortezomib (Velcade, PS-341), a break-through multiple myeloma treatment, moved rapidly through development from bench in 1994 to first approval in 2003 (Albanell and Adams, 2002). The United States Food and Drug Administration approved bortezomib for the treatment of multiple myeloma and mantle cell lymphoma. This success was followed by two other proteasome inhibitors, carfilzomib and MLN9708, which were approved for the treatment of multiple myeloma (Manasanch and Orlowski, 2017). Although bortezomib was a great step forward in multiple myeloma treatment, a sub-population of patients failed to respond to bortezomib and almost all patients relapsed after bortezomib was used as a single agent or in combination. Mutations and overexpression of the β-subunit of the proteasome was seen as a frequent cause of bortezomib resistance. Furthermore, hematologic tumor cells were found to be more sensitive to proteasome inhibitors compared to solid tumor cells (Ruschak et al., 2011).

To surpass the limitation of proteasome inhibitors in treating malignancies as well as to develop a diverse arsenal of cancer treatment, enzymes of the UPS upstream of the proteasome have been targeted (Bedford et al., 2011). Although genetic manipulations

like si/shRNA and more recently CRISPR/Cas9 systems are effective ways of testing the therapeutic potential of a target protein, small-molecule inhibitors are desirable as they allow temporal control and can be used later for medicinal use.

Malignant cells display a higher activity of UPS and accumulation of ubiquitylated proteins. The levels of UBA1 were shown to be similar in normal and malignant cells, indicating that this enzyme is more actively utilized in malignant cells (Xu et al., 2010). This observation suggests that malignant cells are more dependent on E1 activity and provides a rationale for targeting the Ub E1 as an anti-cancer strategy.

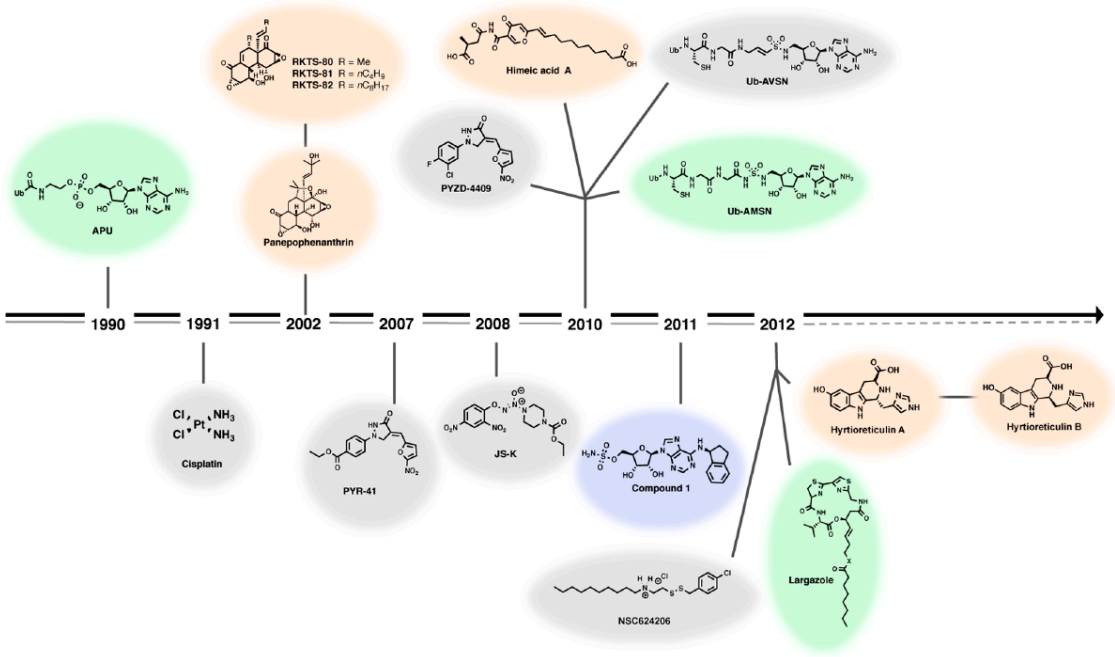


Figure 48: Timeline of the development of UBA1 inhibitors since 1990. UBA1 inhibitors act either by binding to the ATP binding pocket (green) or target the active site cysteine (grey). Mechanism-based inhibitors and inhibitors with unknown mechanisms are colored in blue and orange respectively.

An shRNA-mediated knockdown of UBA1 induced cell death in malignant cells. Previous reports demonstrated that UBA1 inhibition induces cell death through a mechanism linked to endoplasmic reticulum associated stress (ER stress). Upon ER stress, cells activate a series of complementary adaptive mechanisms to cope with protein-folding alterations, which together are known as the unfolded protein response (UPR) (Hetz, 2012). When cells undergo irreversible ER stress, this pathway eliminates damaged cells by apoptosis. This may serve as an explanation of the efficacy of proteasome and Ub E1 inhibitors against malignant cells as accumulation of intracellular proteins due to defect in the UPS can induce chronic ER stress in malignant cells. Furthermore, as Ub E1 inhibition can block both proteasome as well as aggresome mediated degradation, it will result in a more severe accumulation of toxic

aggregates of misfolded proteins. Therefore, this strategy might be advantageous against solid tumors where proteasome inhibitors are still ineffective. UBA1 inhibition can also turn beneficial to overcome proteasome resistance (An and Statsyuk, 2015; Xu et al., 2010).

Ever since 1990, multiple UBA1 inhibitors have been reported starting from a nonhydrolyzable mimetic of the the ubiquitin adenylate intermediate adenosyl-phospho-ubiquitinol (APU) (Figure 48) (Wilkinson et al., 1990). The chemotherapeutic agent, cisplatin also showed an inhibitory effect agsinst UBA1 which was abolished in the presence of sulfhydryl compounds like dithiothretol (DTT) and βME, thus supporting the hypothesis that inhibition is achieved by reaction with the active site cysteine (Isoe et al., 1992). Panepophenanthrin, a compound isolated from the *Panus radis* mushroom, was the first natural product to show UBA1 inhibition (Sekizawa et al., 2002). Many derivatives of this compound including RKTS-80-82 were found to inhibit Uba1~Ub thioester formation more potently in cell-free assays than the originally isolated molecule (Matsuzawa et al., 2006). However, the mechanism and selectivity of these compounds are largely unknown. Two inhibitors containing a pyrazolidine scaffold, PYR-41 and PYZD-4409, showed an inhibitory potency in the low μM range and were proposed to target the catalytic cysteine via a Michael addition (Xu et al., 2010; Yang et al., 2007). However, as these inhibitors were shown to cross-react with DUBs, their specificity is questionable (Chen et al., 2014a; Kapuria et al., 2011). A cellular screen of compounds performed to identify inhibitors of p27 proteolysis identified NSC624206, which inhibits the loading of adenylated ubiquitin onto the active site cysteine (Ungermannova et al., 2012a). JS-K, a nitric oxide prodrug, also inhibits Ub loading of the E1 cysteine, presumably by thiol nitrosylation (Kitagaki et al., 2009). Ub-AMSN and Ub-AVSN probes function similar to the SUMO-AMSN and SUMO-AVSN probes discussed in chapter 2 (section 2.4.7) (Lu et al., 2010) and thus inhibit Uba1 activity. Another natural product, largazole from the cyanobacteria genus *Symploca*, was implicated as UBA1 inhibitor by blocking ubiquitin adenylation (Ungermannova et al., 2012b). The mechanism of inhibition of several other inhibitors like Himeic acid and Hyrtioreticulin is unclear. Although the cornucopia of inhibitors provide a reasonable set of starting points, most of these inhibitors have a low inhibitory potency (in the μM range) and their specificity toward UBA1 is only poorly characterized.

Adenosyl sulfamates were described as cell-membrane permeable E1 inhibitors with the NEDD8 E1 selective inhibitor MLN4924 as first-in-class representative (Soucy et al., 2009). These compounds form N-acylsulfamates with Ub/Ubls (Nawrocki et al., 2012) in the presence of ATP and the E1 (Brownell et al., 2010; Chen et al., 2011) via a mechanism where the sulfamate NH_2-group attacks the E1~Ub/Ubl thioester to produce

a covalent Ub/Ubl-inhibitor adduct (Figure 49). The resulting adduct mimics the Ub/Ubl adenylate and binds tightly (with picomolar affinity) to the E1, thus causing its inhibition (Chen et al., 2011).

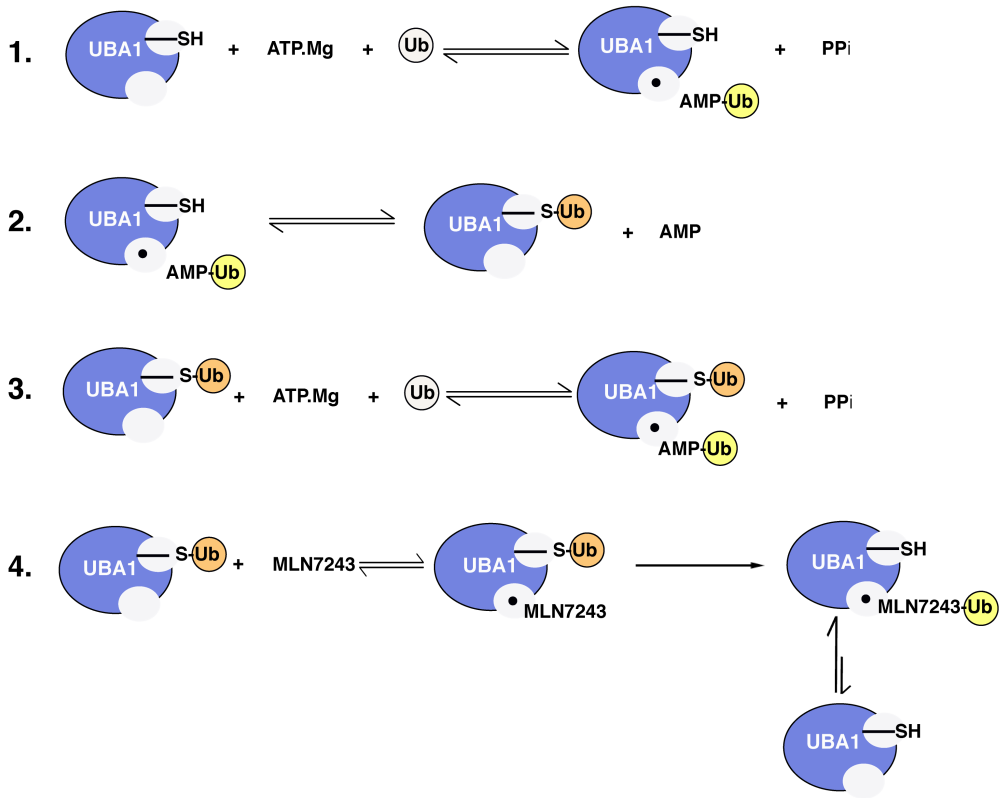


Figure 49: Substrate-assisted mechanism of adenosyl sulfamate inhibitors: The first three steps represent the normal function of UBA1 up to the formation of the doubly-loaded Uba1 complex, while the fourth step illustrates how MLN7243 (representative for all adenosyl sulfamates) inhibits the enzyme.

Although adenosyl sulfamates specifically inhibit E1 enzymes, their specificities differ vastly as reflected in low nanomolar to high micromolar IC_{50} values reported for these compounds towards various E1 enzymes (Figure 50). Recently, the adenosyl sulfamates MLN7243 (Traore et al., 2014), a UBA1-selective inhibitor, and ABPA3 (An and Statsyuk, 2015), a dual inhibitor of the Ub and NEDD8 E1s were reported. Despite the structure of MLN4924 bound to NEDD8 E1, knowledge of how chemically diverse adenosyl sulfamates occupy the ATP-binding pocket of UBA1 and display variable specificity towards different E1 enzymes had been missing. With this motivation, all three inhibitors were crystallized as ubiquitin•inhibitor covalent adducts in complex with yeast Uba1. In this chapter, the specificity determinants of adenosyl sulfamate inhibitors towards E1 enzymes by utilizing our structural data combined with biochemical approaches is addressed. The results are further substantiated by computational tools

to explain the experimental observations and together the study presents an avenue for the development of highly potent yet specific E1 inhibitors.

(A)

E1	UBA1	NEDD8 E1	SUMO E1	UBA6	ATG7
MLN4924 IC_{50} (µM)	1.5±0.71	$4.7±1.5•10^{-2}$	8.2±6.2	1.8	>10

(B)

E1	UBA1	NEDD8 E1
MLN4924 IC_{50} (µM)	22	$12•10^{-3}$

Figure 50. Inhibitory values reported for MLN4924 targeting different E1 enzymes. (A) IC_{50} values reported in Soucy et al., 2009, using a FRET based E1-E2 transthioesterification assay at an ATP concentration of 20 µM. The IC_{50} values for the E1 enzymes are in micromolar units, except for the NEDD8 E1 where the value is in the low nanomolar range. (B) Chen at al. reported IC_{50} values of MLN4924 using an ATP-PP_i exchange assay at an ATP concentration of 100 µM.

3.3 MATERIALS AND METHODS

3.3.1 Constructs and expression

A construct encoding the *Saccharomyces cerevisiae* Uba1 Δ9 variant (amino acids 10-1024) in pET28a was used to express the protein with an N-terminal hexahistidine-tag (Lee and Schindelin, 2008) as described in the previous chapter. For human UBA1 purification we used a *Homo sapiens* Δ40 UBA1 construct in the pET23b+ vector with a C-terminal His-tag. Tag-free yeast ubiquitin was cloned into the NdeI/HindIII restriction sites of the pET30a vector similar to human ubiquitin construct mentioned in the last chapter. Proteins were expressed in *Escherichia coli* BL21(DE3) RIL and BL21(DE3) strains for Uba1/UBA1 and ubiquitin, respectively, in TB media. For Uba1/UBA1 expression, cells were induced at an OD_{600} of 0.8-1 with 0.1 mM IPTG followed by overnight growth at 25° C and for yeast ubiquitin 0.4 mM IPTG was used for similar OD_{600} of expression culture followed by growth for 4 hours at 37° C. Yeast Uba1 was purified as described in section 2.3.3.

3.3.2 Human UBA1 purification

A three-step purification, Ni-NTA, ion exchange and size exclusion chromatography was carried out for UBA1 purification. After Ni-NTA chromatography, the fractions with the desired protein were pooled and diluted twofold with 20 mM Tris, pH 8.0 containing 5 mM βME and protease inhibitors. The resulting sample was passed through a pre-equilibrated Q-sepharose (GE, Healthcare) column. The column was washed with 20 mM Tris, pH 8.0, 100 mM NaCl, 5 mM βME buffer. A gradient from 100 to 400 mM NaCl using the same buffer was run to elute the protein. The fractions containing the protein were pooled and concentrated to 5 ml using Centricon plus-20 ultrafiltration device (50 kDa cutoff; Millipore) and injected onto a pre-equilibrated Superdex 26/ 60 200pg column (GE, Healthcare). The final buffer for this protein contained 25 mM Tris pH 7.5, 100 mM NaCl and 5 mM βME. The same purification strategy was used to purify the UBA1 A580T mutant derived from the Δ40 UBA1 construct by site-directed mutagenesis. All proteins were flash-frozen in liquid nitrogen and stored at -80° C.

3.3.3 Crystallization and structure determination

For crystallization, purified His_6-Uba1 and ubiquitin were incubated for 30 minutes at room temperature at a molar ratio of 1:2 in the presence of 2.5 mM Mg^{2+}-ATP and 1 mM of ABPA3, which was synthesized as described (An and Statsyuk, 2015) or purchased in the case of MLN7243 (Chemietek) and MLN4924 (Chemietek). The final Uba1 concentration was 12 mg ml^{-1}. Crystallization was performed via sitting drop

vapor diffusion by mixing 0.5 µl of protein and 0.5 µl of mother liquor containing 0.2 M Li_2SO_4, 0.1 M Bis-Tris pH 6.0 and 15% (w/v) PEG 3350 against mother liquor. Crystals belonged to the orthorhombic space group $P22_12_1$.

15% (v/v) glycerol was added to the mother liquor to cryo-preserve the crystals upon transfer to liquid nitrogen. Data collection was performed at 100 K either on beamline BL14.1 at BESSY, Berlin, Germany (Mueller et al., 2015) or beamline ID30A-3 at ESRF, Grenoble, France. Data were processed using XDS (Kabsch, 2010) and Aimless (Collaborative Computational Project, 1994). Data collection statistics are summarized in Table 9. The structures were solved by molecular replacement with PhaserMR (McCoy et al., 2007) using 4NNJ as the search model but omitting the thioesterified ubiquitin. Two molecules of the Uba1-Ub complex were detected in the asymmetric unit together with a single additional Ub covalently linked to C600 (Figure 57) for which the electron density was particularly prominent in the MLN4924-bound Uba1 structure. Manual rebuilding and ligand placement was accomplished with Coot and refinement was carried out with Refmac (Murshudov et al., 1997). Figures representing protein structures were generated with PyMOL (http://www.pymol.org).

3.3.4 Inhibition assays

5 µM *Sc*Uba1/*Hs*UBA1 protein was mixed with 5 µM LI-COR 800CW NHS ester dye labeled (https://www.licor.com/bio/products/reagents/irdye/800cw/nhs_ester) human ubiquitin and either 3.33% (final concentration) of DMSO or the inhibitor at the same concentration of DMSO in 25 mM Tris pH 7.5 and 200 mM NaCl. The reaction was started by adding 100 µM ATP and 200 µM $MgCl_2$ solution. After incubation at 37° C for one hour the reaction was quenched with 35 mM EDTA and the samples were analyzed on 12% SDS-PAGE gels under non-reducing conditions. The gels were scanned at 800 nm to visualize binding of thioesterified ubiquitin to UBA1. The gels were later stained with coomassie to document uniform loading. The infrared fluorescence intensities obtained by the 800 nm scan were quantified using the LI-COR Odyssey software. Each assay was repeated three times to obtain statistically significant IC_{50} values for each inhibitor for yeast Uba1 and human UBA1. Moreover, two or more batches of proteins were tested in the assay to control for batch to batch variations.

To test the mutant A580T in this assay, a similar set-up was used except that the incubation of the reaction was done at room temperature for one hour as the mutant showed bands on the gel in the higher molecular weight range after incubation at 37° C for one hour. This assay was performed for both wtUBA1 and A580T mutant in duplicates.

3.3.5 Molecular Dynamics Simulations and LIE calculations

The preparation of protein and ligand structures was carried out in MOE (Molecular Operating Environment) version 2015.1001. The three complexes of Uba1 containing MLN7243, MLN4924 and ABPA3 were superimposed by minimizing the root mean square deviation (rmsd) of their backbone atoms. For each complex, only chain A, corresponding to a single UBA1 moiety without ubiquitin, was used, along with the ligand and all water molecules within 4.5 Å of the chain; all other atoms were removed. If present, S-hydroxyl-L-cysteine in position 1006 was mutated to cysteine. Alternate locations were selected after visual inspection and the system was protonated using Protonate3D (Labute, 2009) with default settings at pH 7.4 and locally minimized with the Amber12:EHT force field implemented in MOE using an rms gradient of 0.1 kcal/(mol·Å). Furthermore, a complex of human UBA1 and MLN7243 was generated on the basis of the published homology model of human UBA1 without ligand (Correale et al., 2014). After preparation of the protein according to the presented workflow and modeling the side chains of M505 and R551 (corresponding to N471 and K519 in yeast Uba1) in conformations not blocking the binding pocket, MLN7243 was placed therein, in analogy to the binding mode observed in the corresponding yeast Uba1 complex structure.

Proteins and ligands were set up as described above. Atomic charges for the ligand were assigned with Antechamber/RESP (Wang et al., 2004) based on the electrostatic potential calculated with Gaussian 09 (Frisch et al., 2009) at the HF/6-31G* level. Systems were built using the tleap module of AMBER 14 (Case et al., 2014) with ff99SB force field parameters (Hornak et al., 2006). Missing force field parameters were estimated with the Amber module parmchk (Wang et al., 2006). Initially, 2,000 steps of minimization were performed on the complex in the generalized Born implicit solvent model (Tsui et al., 2001) using the Sander module. Sodium ions were added for neutralization and the complex was solvated in a TIP3P water box (Jorgensen et al., 1983) with a minimum protein-to-box border distance of 10 Å. The simulations were carried out with NAMD 2.10 (Phillips et al., 2005). System equilibration was performed by a 10,000 step minimization. After applying harmonic constraints (0.5 kcal/(mol·Å²)) to all non-solvent atoms, the system was heated from 100 to 300 K while gradually releasing the constraints over 500 ps in a box of constant volume. Thereafter, the entire system was allowed to move freely for 500 ps. For all simulations, periodic boundary conditions were applied to the systems. The electrostatic interactions were handled with the particle mesh Ewald methodology (PME) (Darden et al., 1993). A 12 Å cut-off was applied for the van der Waals interactions. Nosé-Hoover Langevin piston pressure control and Langevin dynamics were used for keeping the pressure and temperature

constant in the production run. 2 fs steps were used for integration, while the coordinates of the trajectories were saved every ps. MD simulations for the yeast Uba1 complexes were carried out for 50 ns each, whereas the human UBA1 structure was simulated for 25 ns. Simulations for the free ligand in a water box were run for 50 ns in each case. Interaction energies (van-der-Waals and electrostatic contribution) according to the LIE method (Åqvist et al., 2001) were calculated with cpptraj (Roe et al., 2013) using the entire trajectories.

3.3.6 Contributions to the work

Maximillian Kuhn carried out the MD simulations and LIE calculations under the supervision of Prof. Dr. Christoph Sotriffer at Institute of Pharmacy and Food Chemistry, Department of Chemistry and Pharmacy, University of Würzburg. Our collaborators, Heeseon An and Dr. Alexander Statsyuk (previously located at Northwestern University, Illinois, USA) provided us the inhibitor ABPA3. Mark Löbel, an earlier bachelor student in the lab performed replicates of the inhibition assays.

3.3 RESULTS AND DISCUSSION

3.4.1 Crystal structures of Uba1-ubiquitin~inhibitor covalent adducts

We determined crystals structures of MLN4924, MLN7243 and ABPA3 covalently linked to Ub and in complex with yeast Uba1 at resolutions of 2.8 Å, 2.7 Å and 2.3 Å, respectively, resulting in very well defined inhibitor-bound structures (Figure 51A-G). MLN4924, ABPA3 and MLN7243 occupy the ATP-binding pocket with the sulfamate-nitrogen covalently linked to G76 of Ub. Data collection and refinement statistics for the three structures are summarized in Table 9. Overall, the structures are very similar as reflected in pairwise rms deviations ranging from 0.2 to 0.3 Å for the ABPA3/MLN4924 and ABPA3/MLN7243 pair, respectively, however, subtle differences exist in the structures due to the differences in the chemical makeup of the compounds and minor structural changes in the protein induced by these chemical differences.

While MLN4924, MLN7243 and ABPA3 all form covalent adducts with Ub in the presence of Uba1/UBA1 and ATP via their sulfamate moiety, the different nucleobases and different stereochemistries at the C4' atom present in these compounds probably affect the rates of covalent adduct formation and binding of the covalent Ub•inhibitor adducts to Uba1/UBA1, thus determining the overall specificity (Figure 51A-C). While ABPA3 is an adenosine derivative substituted at the N6-position with a meta-ethynylbenzene moiety, MLN4924 and, in particular, MLN7243 differ significantly from adenosine by replacing the ribose with cyclopentane derivatives and the adenine with other ring systems consisting of fused 6 and 5 membered aromatic heterocycles, each carrying one bulky substituent. In addition, in MLN7243 the cyclopentane moiety and aromatic heterocycle are bridged by a unique amino-group.

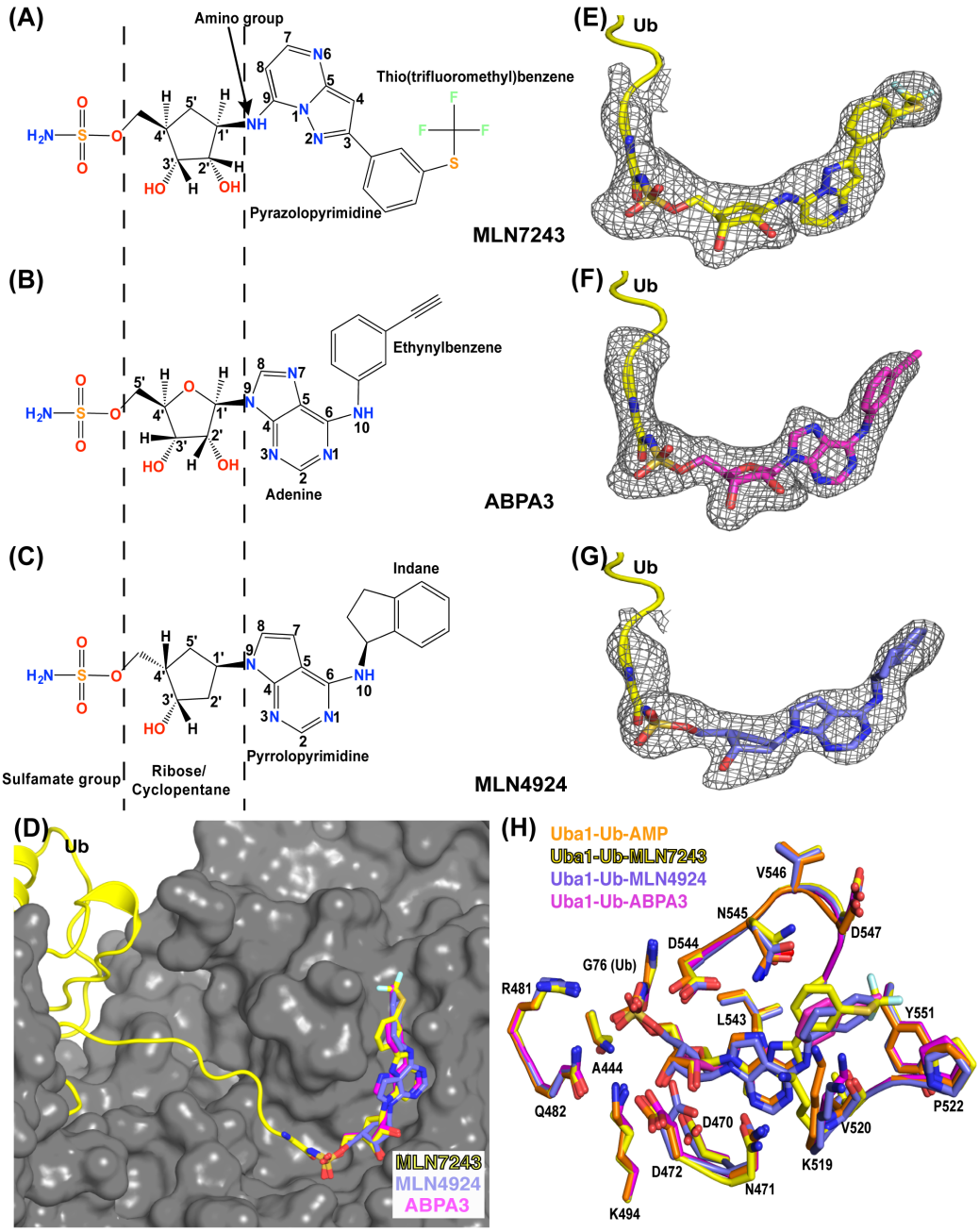


Figure 51. Chemical structures and omit maps of the inhibitors and their binding site. (A-C) Chemical structures of the inhibitors MLN7243 (A), ABPA3 (B) and MLN4924 (C), highlighting similarities and differences. (D) Superposition of the MLN4924(blue) / ABPA3(magenta) / MLN7243(yellow)-Ub (yellow ribbon) adducts in complex with Uba1 (gray surface). (E-G) Omit F_o-F_c electron density maps (green) contoured at rmsd of three after the inhibitor (MLN4924, MLN7243 and ABPA3) and the C-terminal di-glycine of ubiquitin were omitted. (H) Superpositions of MLN7243-Ub (yellow), ABPA3-Ub (magenta), MLN4924-Ub (blue) and the ubiquityl-adenylate (orange, PDB entry 4NNJ) in complex with Uba1.

	Uba1-Ub-ABPA3	Uba1-Ub-MLN4924	Uba1-Ub-MLN7243
PDB code	**5L6H**	**5L6I**	**5L6J**
Beamline	ESRF, ID30A-3	BESSY, BL14.1	BESSY, BL14.1
Wavelength (Å)	0.9686	0.9184	0.9184
Space group	P 2 2_1 2_1	P 2 2_1 2_1	P 2 2_1 2_1
Unit cell Parameters (Å)	72.0, 194.1, 230.2	72.5, 191.9, 230.2	72.5, 193.9, 230.2
R_{sym}[a]	0.107 (1.866)	0.133 (0.828)	0.130 (0.887)
R_{pim}[b]	0.061 (1.060)	0.093 (0.597)	0.103 (0.685)
$CC_{1/2}$	0.998 (0.549)	0.994 (0.689)	0.992 (0.595)
Redundancy	7.5 (7.8)	5.6 (5.2)	4.5 (4.6)
Unique reflections	143,279	82,916	91,638
Completeness	0.99 (1.0)	0.99 (0.92)	0.99 (0.97)
<I/σI>[c]	13.3 (1.3)	11.2 (1.9)	10.5 (1.7)
Resolution limits (Å)	20.00-2.30	20.00-2.76	20.00- 2.68
R^{d} /R_{free}[e]	0.173/0.216	0.176/0.219	0.173/0.223
Deviation from ideal values in			
Bond distances (Å)	0.023	0.013	0.018
Bond angles (°)	2.216	1.664	1.973
Planar groups (Å)	0.012	0.007	0.105
Chiral centers (Å^3)	0.128	0.089	0.079
Ramachandran statistics (%) (Preferred/Allowed/Outliers)	96.9/3/0.1	96.1/3.1/0.1	96.9/3/0.1
Overall average B factor (Å^2)	58.0	52.0	49.0

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i |I_i - \langle I \rangle| / \Sigma_{hkl}\Sigma_i I_i$ where I_i is the i^{th} measurement and $\langle I \rangle$ is the weighted mean of all measurements of I.

[b] $R_{pim} = \Sigma_{hkl}\, 1/(N-1)^{1/2} \Sigma_i |I_i - \langle I \rangle| / \Sigma_{hkl}\Sigma_i I_i$, where N is the redundancy of the data and $I(hkl)$ the average intensity.

[c] $\langle I/\sigma I \rangle$ indicates the average of the intensity divided by its standard deviation.

[d] $R = \Sigma_{hkl} ||F_o| - |F_c|| / \Sigma_{hkl}|F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes.

[e] R_{free} same as R for 5% of the data randomly omitted from the refinement. The number of reflections includes the R_{free} subset.

Table 9. Data collection and refinement statistics

3.4.2 Closer look at the inhibitor-Uba1 interfaces

In the following section the specific protein-inhibitor interactions are described in detail and are compared to each other. The hydroxyl-groups of the MLN7243 cyclopentane and ABPA3 ribose form two and one H-bonds with the side chains of D470 and K494, respectively. However, due to the lack of the 2'-OH-group, MLN4924 engages in only one H-bond with D470. MLN4924 features a 4'-endo compared to a 2'-endo pucker in the other structures and, due to the different stereochemistry at the C4'-atom, the O-C-C4' link between the sulfamate and cyclopentane is rotated by 90° (Figure 51G). The adenine of ABPA3 and pyrrolopyrimidine of MLN4924 form similar contacts as the adenine in the Ub-adenylate bound structure (Schäfer et al., 2014), where both the backbone carbonyl and amide-groups of V520 form two H-bonds with the purine (Figure 51H). ABPA3 forms two additional H-bonds between its purine N7 and C8 atoms and the N545 and D544 side chains of Uba1, respectively. Since MLN4924 has a pyrrolopyrimidine instead of adenine, it lacks the N7-nitrogen and hence does not form these H-bonds. In the MLN7243-bound Uba1 structure, the nucleobase is flipped, due to the unique amino-bridge, resulting in only one H-bond to V520. Both MLN4924 and ABPA3 terminate with a hydrophobic moiety, indane and ethynylbenzene, respectively, which are embedded in a hydrophobic pocket formed by V440, V520, L543 and A548 (Figure 52B-C).

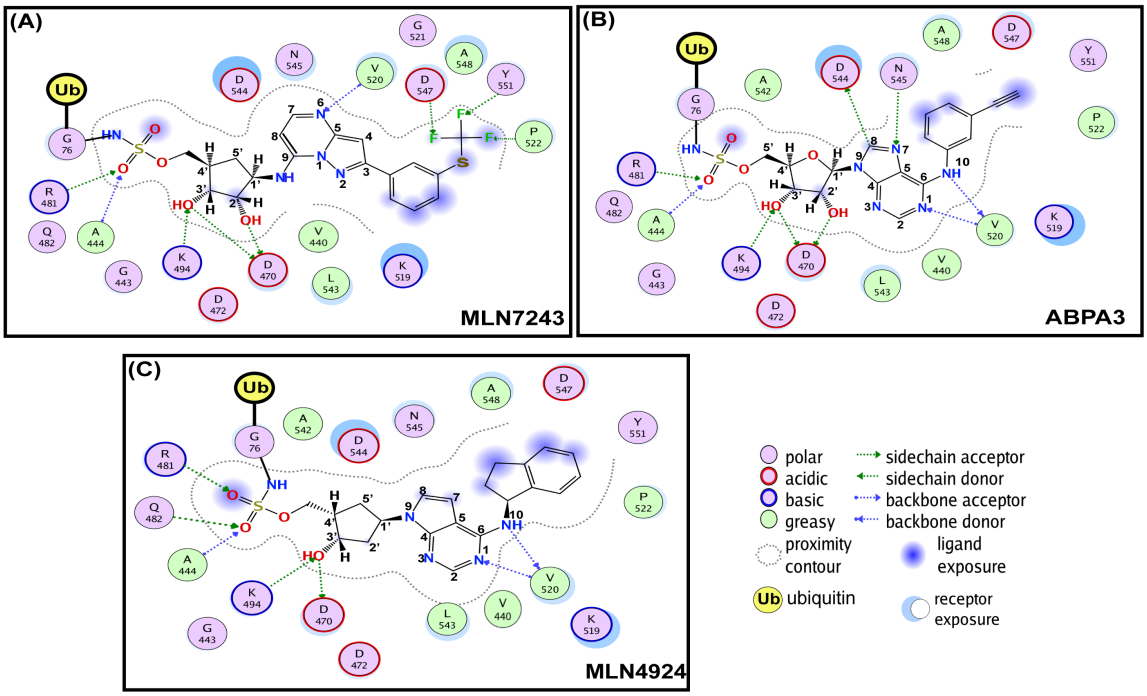


Figure 52: 2D-interaction map of the Uba1-inhibitor complexes. (A-C) Schematic 2D representations (generated with MOE) of Uba1-inhibitor interactions involving (A) MLN7243, (B) ABPA3 and (C) MLN4924. The legend (lower right) defines the different types of interactions.

In contrast to ABPA3 and MLN4924, MLN7243 contains a meta-thio(trifluoromethyl) benzene, which occupies the same hydrophobic patch, but engages in several H-bonds via its CF_3-group (Figure 53). The thio(trifluoromethyl)-group acts as a hook, rendering it structurally compatible with the UBA1-surface. The hydrophobic residues V440, V520, L543 and A548 together with Y551, P522 and D547 frame the pocket which is preserved in human UBA1 and can be pharmacologically targeted with a CF_3-group.

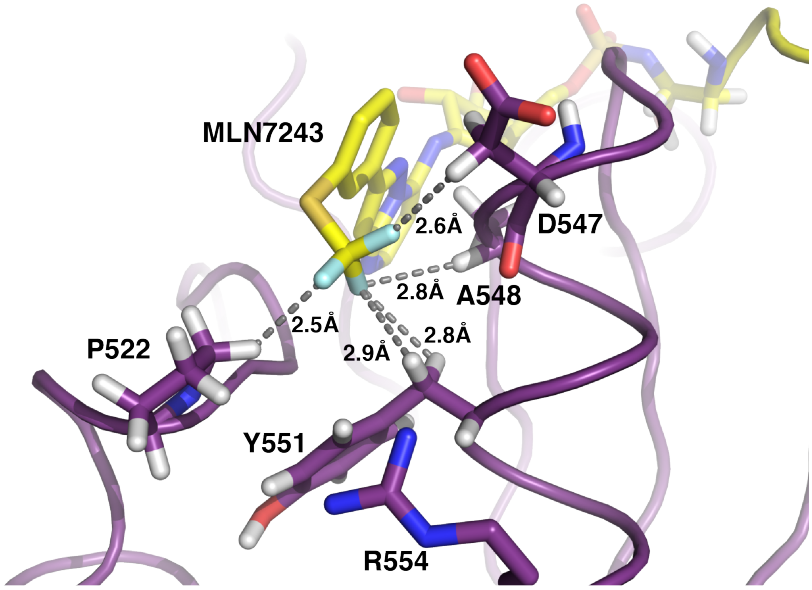


Figure 53. Interactions of the trifluoromethyl group of MLN7243 with Uba1. The trifluoromethyl moiety of MLN7243 (yellow) forms a network of hydrogen bonds with P522, D547, A548 and Y551 (purple). Hydrogen atoms are only shown for atoms forming putative interactions with the trifluoromethyl moiety.

3.4.2 Similarities between inhibitor-ubiquitin adduct and ubiquitin adenylate

The overall arrangement of the coordinating residues around AMP in the Ub-adenylate is not significantly altered in the inhibitor-bound structures (Figure 51H). The adenine present in AMP and ABPA3 forms H-bonds with V520, D544 and N545 of UBA1 and the hydroxyl groups of ribose are coordinated through the side chains of K494 and D470. These contacts are partially absent in the MLN7243 and MLN4924 structures due to the replacement of adenine with a pyrrolopyrimidine ring and the absence of a hydroxyl group in the cyclopentane ring of MLN4924, coupled to altered conformations of D472 and N545. In the MLN7243 and ABPA3 bound structures one of the sulfamate oxygen forms hydrogen bonds with R481 and A444, thus mimicking similar interactions with the α-phosphate of AMP, whereas, due to the aforementioned variation in the C4' stereochemistry and the spatial arrangement of sulfamate group in the MLN4924-bound structure, Q482 replaces R481 for this interaction and R481 coordinates another oxygen atom of the sulfamate group. These interactions were proposed to lower the pK_a of the sulfamate amino group, thus promoting its deprotonation and subsequent

nucleophilic attack on the thioester bond (Brownell et al., 2010). Ubiquitin adopts a virtually identical position in the inhibitor-bound and ubiquitin-adenylate structures as reflected by pairwise RMS deviations of less than or equal to 0.2 Å for the main chain atoms.

3.4.3 Comparative analysis with the NEDD8 E1 bound NEDD8•MLN4924 structure

A comparison of the crystal structure of the NEDD8•MLN4924-NEDD8 E1 complex (Brownell et al., 2010) with our structures reveals critical residues at the ATP-binding sites that may contribute to inhibitor specificity. In UBA3 the hydrophobic patch of UBA1 around the ATP-binding pocket is preserved with I75, I148, L166 and A171, however, it is extended by I170, W174 and V449, which surround the MLN4924 indane (Figure 54). Furthermore, M101, which is present in the NEDD8 E1, forms additional hydrophobic interactions with the MLN4924 pyrimidine, and homology modeling suggest that M505 of UBA1, corresponding to N471 of yeast Uba1, would occupy a position equivalent to that of M101 (Figure 54). The MLN4924-bound NEDD8 E1 structure features a unique H-bond between the Q149 side-chain oxygen and the N10-nitrogen of MLN4924. The equivalent position in UBA1 is occupied by a glycine, which cannot form this H-bond.

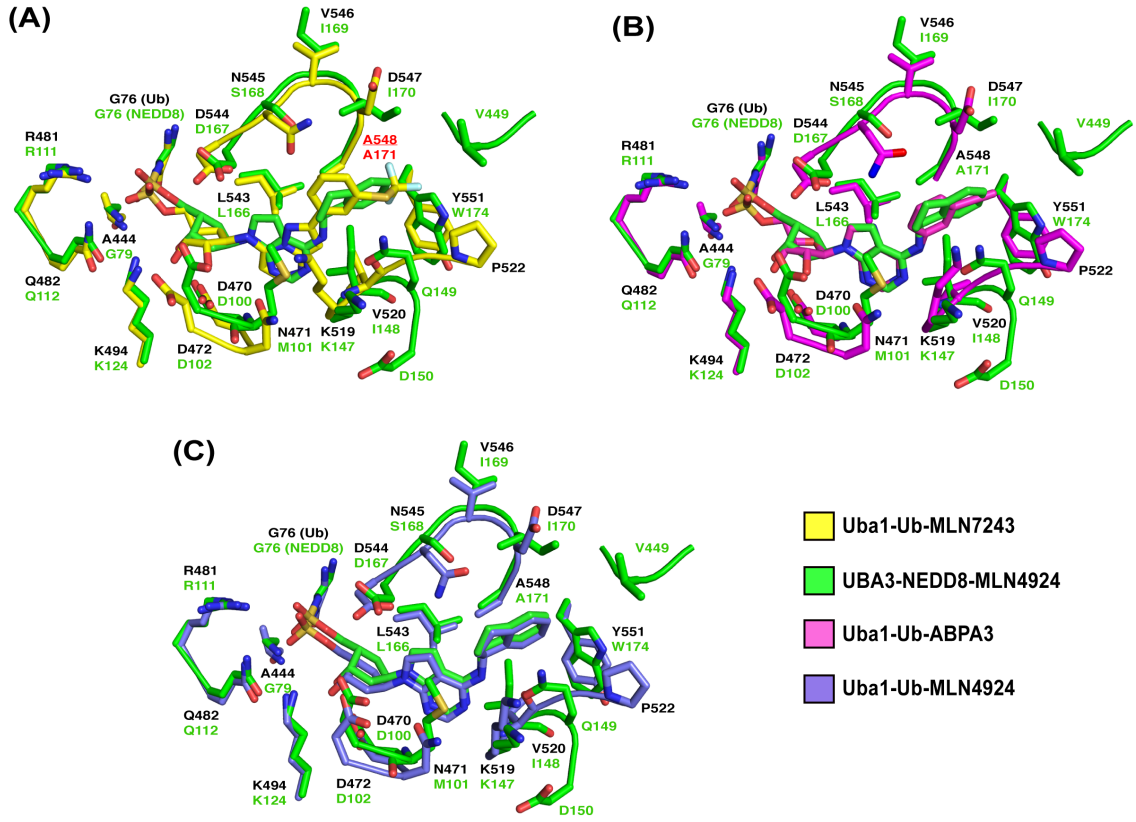
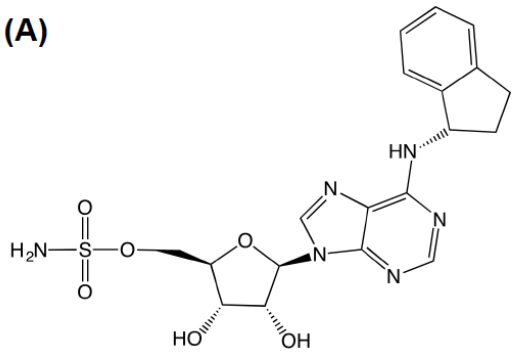


Figure 54: Comparative analysis with NEDD8 E1 bound MLN4924-NEDD8 adduct. (A-C) Superpositions of the MLN4924-NEDD8 E1 (green, PDB entry 3GZN) as well as the MLN7243 (yellow), ABPA3 (magenta) and MLN4924 (purple) Uba1

complexes.

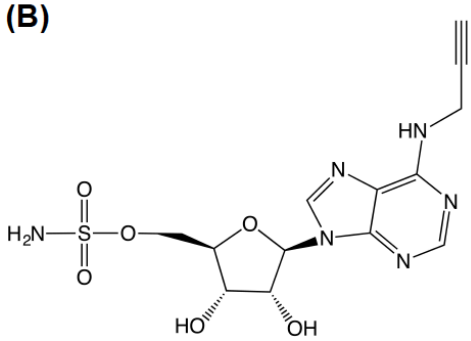
Overall, the different C4' stereochemistry of MLN4924, extended hydrophobic pocket of UBA3 and the unique interactions of MLN4924 with UBA3 together may contribute to the specificity of MLN4924 towards the NEDD8 E1. Since the N10-atoms of ABPA3 and MLN4924 bind similarly to UBA1, we expect that the ABPA3 N10 would also form a similar H-bond with Q149 of UBA3. Accordingly, the N6-ethynylbenzene moiety of ABPA3 would also be in contact with the same hydrophobic patch of UBA3, thus partly explaining the potency of MLN4924 and ABPA3 towards the NEDD8 E1. In contrast, MLN7243 benefits from the C-H···F bonds formed by the fluorines of the trifluoromethyl-group, and this group fits into the pocket lined by P522, D547, A548 and Y551 of UBA1 (Figure 53). When superimposed with the NEDD8 E1, MLN7243 shows only two possible C-H···F bonds with I170 and W174 of UBA3 plus an unfavorable contact (2.5 Å distance) with the backbone carbonyl of V449, which originates from the ubiquitin-fold domain (UFD) of UBA3 (Figure 54A). Although the potency of MLN7243 toward the NEDD8 E1 needs to be tested, our structural data suggest it to be more specific for UBA1.

(A)



Compound I

(B)



ABP1

(C)

E1	UBA1	NEDD8 E1	SUMO E1	UBA6	UBA7	ATG7
Compound 1 IC_{50} (μM)	5.2	$1 \cdot 10^{-2}$	6.4	0.92	4.0	>100

Figure 55. Specificity of Compound I and ABP1 towards different E1 enzymes. (A) Compound-I is identical to ABPA3 except for an indane at the N10-position, and its covalent adduct with Ub presumably adopts a similar orientation as ABPA3 when bound to UBA1. (B) ABP1 is another adenosyl sulfamate analogue with a smaller propargyl group at the N10 position. This avoids steric clashes with M405 of Atg7 and M97 of SUMO E1 and the compound therefore inhibits these enzymes *in vivo* and *in vitro*. (C) IC_{50} values reported for compound I against E1 enzymes derived from an ATP-PP_i exchange assay at an ATP concentration of 1 mM (Chen et al., 2011).

As ABPA3 closely resembles compound I (Chen et al., 2011) (Figure 55A), the ABPA3-bound Uba1 structure also suggests how the Ub-compound I covalent adduct may bind to UBA1. Most likely the indane moiety of compound I occupies the same space in Uba1 as the ethynylphenyl group of ABPA3 and forms similar hydrophobic contacts. Moreover, the presence of an additional hydroxyl group in the ribose rings and adenine as nucleobase in both compounds compared to MLN4924 allows for more favorable interactions with the ATP-binding pocket of both UBA1 and NEDD8 E1. The adenine ring of compound I most likely forms H-bonds with D544 and N545, while the pyrrolopyrimidine of MLN4924 cannot form H-bonds with D544 and N545. An additional H-bond is predicted to form between the 2'-OH-group of compound I and D470 of UBA1, similar to ABPA3. We also cannot exclude the possibility that the inverted C4'-stereochemistry of MLN4924 further contributes to its lower potency toward UBA1. These additional interactions may partly explain the better potency of these compounds against NEDD8 E1 in comparison to MLN4924 (Milhollen et al., 2012).

Looking at published data (see Figure 55C), it appears that compound I (and likewise ABPA3) is more specific for the NEDD8 E1 than MLN4924, which is in line with our structures as compound I will form additional contacts to the NEDD8 E1 ATP-binding pocket with its intact adenosine moiety. However, the specificity determinants for MLN4924 lie in other regions than the adenylation domain itself (*e,g*. in the catalytic cysteine domain in the transition state).

3.4.4 UBA1~UB thioester inhibition assays

To check the inhibitory effect of the compounds on the yeast and human ubiquitin activating enzymes we performed *in vitro* biochemical assays where we labeled the ubiquitin with the LI-COR IR 800CW NHS ester infrared dye and monitored the formation of UBA1~Ub thioester under non-reducing conditions (Figure 56A-C). Consistent with our structural analyses MLN7243 appeared to be the most potent compound against UBA1 with an IC_{50} value of 0.4±0.1 µM, whereas MLN4924 showed a rather poor IC_{50} value of 43±19 µM in these assays. ABPA3 turned out to be the second best inhibitor as characterized by an IC_{50} value of 0.5±0.1 µM. To evaluate our inhibitor bound *Sc*Uba1 structures we carried out the same assays with the yeast Uba1 protein. The resulting IC_{50} values for MLN7243, ABPA3 and MLN4924 were 0.2±0.1 µM, 0.24±0.01 µM and 45±3 µM, respecively which are quite comparable to the values achieved for the human UBA1 protein, thus underlining the significance of our structural data (Figure 56D-F).

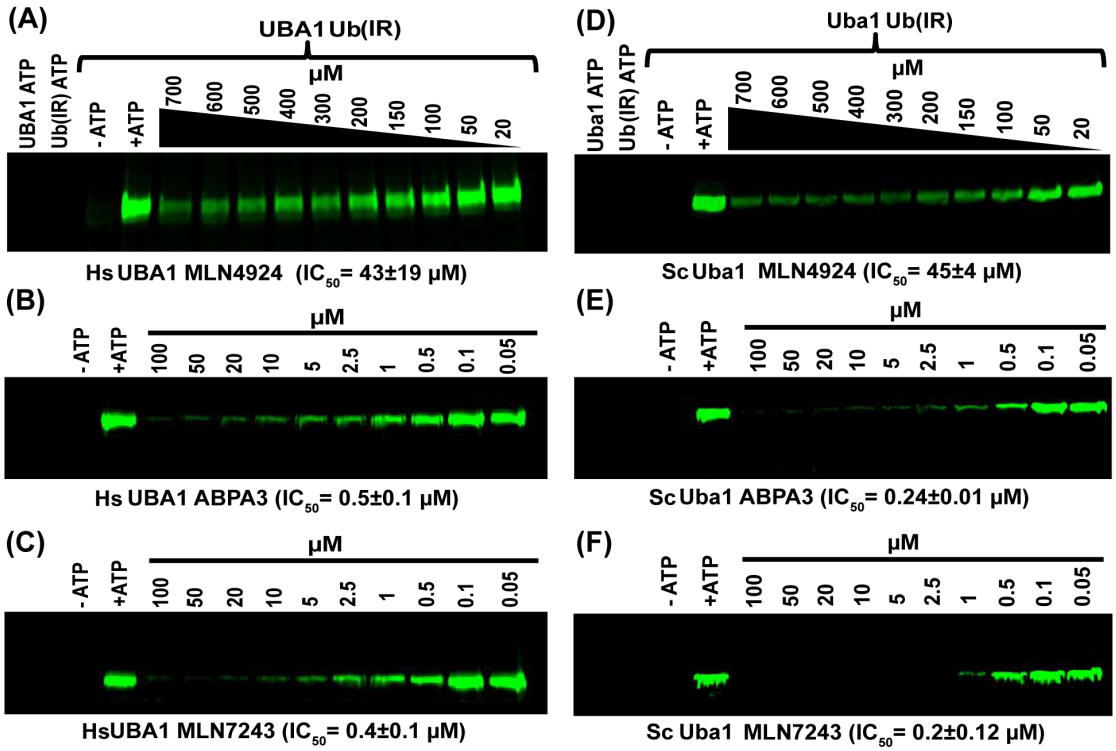


Figure 56: *In vitro* UBA1~Ub thioester inhibition assays. (A-C) 800 nm scan showing IR dye labeled ubiquitin bound to human UBA1 in fluorescent green color in the presence of a concentration gradient of MLN4924, ABPA3 and MLN7243, respectively. (D-F) Yeast Uba1 inhibition assay with (D) MLN4924 ,(E) ABPA3 and (F) MLN4924. IC_{50} values obtained for either human UBA1 or yeast Uba1 with MLN4924, ABPA3 and MLN7243 are depicted. The values represent the mean of triplicate measurements ± S.D.

Remarkably, we could observe a second ubiquitin molecule in the structures which was covalently linked to the active site cysteine. The C600~Ub thioester was best defined in the Ub•MLN4924 adduct structure (Figure 57). Electron density for the thioesterified ubiquitin was also present in the ABPA3-bound Uba1 structure, but it was too weak in the MLN7243-bound structure to allow modeling of the third ubiquitin. The presence of a thioesterified ubiquitin can be explained by a competition between ATP and the inhibitors as the crystallization conditions contained a 2.5-fold molar excess of ATP over the inhibitor concentrations. The position of the thioesterified ubiquitin is equivalent to that observed in the complex of Uba1 with acyl-adenylated ubiquitin (PDB entry 4NNJ) and, as in this structure, no thioesterified ubiquitin could have bound to the second Uba1 present in the asymmetric unit due to severe steric clashes with a symmetry related molecule in the crystal.

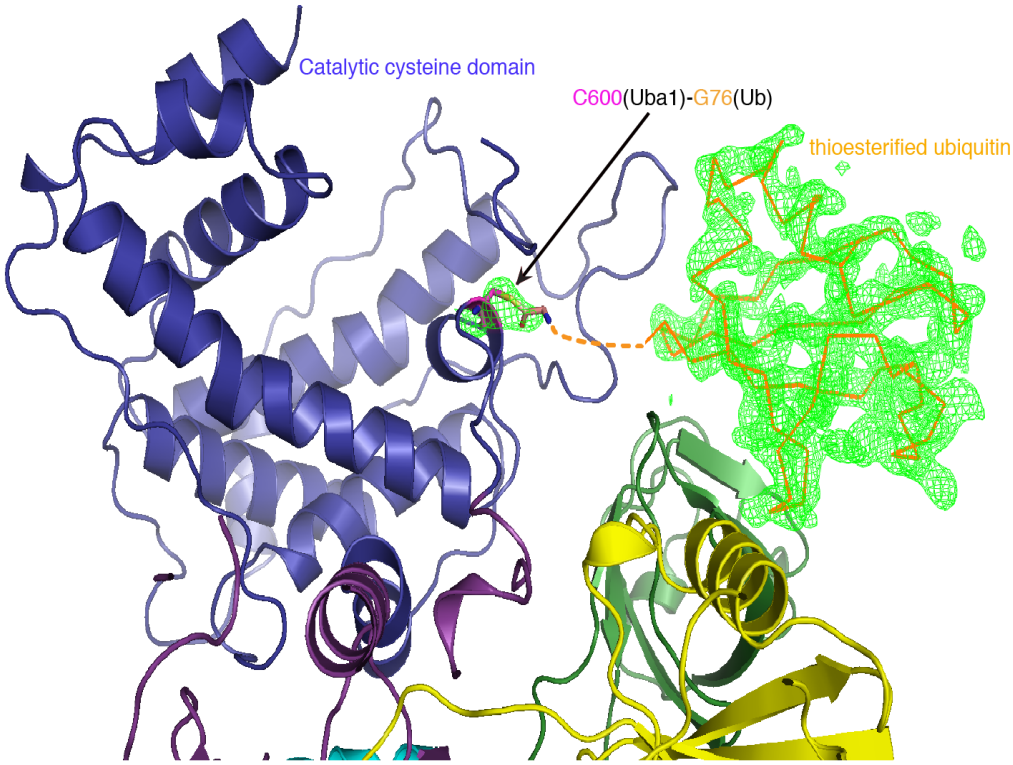
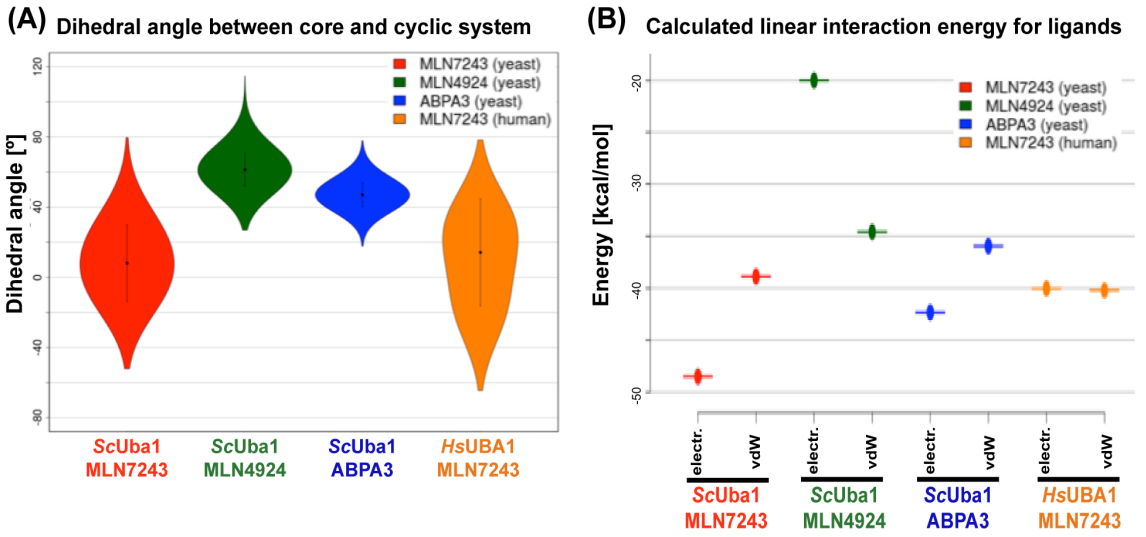


Figure 57. Presence of an additional ubiquitin thioesterified to the active site cysteine. Omit F_o-F_c electron density map (green) contoured at an rmsd of three after omission of this ubiquitin molecule and C600 in the MLN4924-bound Uba1 structure. The C-terminal G76 of ubiquitin is covalently linked to the catalytic cysteine (C600), as demonstrated by the omit map. Residues 73-75 of ubiquitin (approximate position indicated by the dashed line) only displayed very weak density and were omitted from the final model.

3.4.5 MD simulations and LIE calculations

Molecular dynamics (MD) simulations were carried out to assess the differences with respect to the binding mode of the three reported yeast Uba1 complex structures of MLN7243, MLN4924 and ABPA3 and, additionally, of human UBA1 in complex with MLN7243. In all cases, ubiquitin was not included in the simulations. While all key hydrogen bonds (formed via the ligands' hydroxyl groups and their core cyclic system, *i.e.* either pyrrolopyrimidine or adenine) were found to be stable throughout the simulations which had a duration of 50 ns, deviations for the preferentially adopted conformation of the cyclic substituents (corresponding to the 3-thiotrifluoromethylphenyl group in MLN7243, the indane moiety in MLN4924 and the 3-ethynylphenyl group in ABPA3, respectively) were observed. In yeast Uba1, the 3-thiotrifluoromethylphenyl group of MLN7243 preferably adopts a coplanar conformation with respect to the core with a median dihedral angle of 8°, whereas the cyclic substituents of the other ligands adopt more twisted angles (61° in MLN4924 and 47° in ABPA3, respectively; Figure

58A). In human UBA1, the preferred angle is similar to that found with the yeast protein, however, it exhibits a broader distribution (Figure 58A). The overall rather planar conformations of MLN7243 should allow for more suitable hydrophobic contacts compared to the other two ligands. Furthermore, the SCF_3 group can adopt the hook conformation, corresponding to a dihedral angle < -60°, measured along the rotatable sulphur-carbon-bond. This conformation occurs during 75% (for yeast Uba1) and 70% (for human UBA1) of the simulation time, emphasizing the shape complementarity within the protein surface.



(C)

Protein	Ligand	electrostatic			van der Waals		
		$<E_{int,free}^{el}>$	$<E_{int,bound}^{el}>$	ΔE_{int}^{el}	$<E_{int,free}^{vdW}>$	$<E_{int,bound}^{vdW}>$	ΔE_{in}^{vdW}
Yeast Uba1	MLN7243	-55.7	-104.2	-48.5	-22.9	-61.8	-39.0
	MLN4924	-53.4	-73.4	-20.0	-22.2	-56.8	-34.6
	ABPA3	-65.0	-107.3	-42.3	-21.3	-57.3	.36.0
Human UBA1	MLN7243	-55.7	-95.8	-40.0	-22.9	-63.2	-40.2

Figure 58: MD simulations of the protein-ligand complexes and linear interaction energy calculations. (A) Violin plots of the dihedral angle distribution measured between the cyclic substituent of the ligand and each ligand's central bicyclic system. Dots represent the median, lines the median absolute deviation. (B) Graphical representation of the linear interaction energies (unscaled electrostatic and van-der-Waals interaction energy contributions ΔE_{int}^{el} and ΔE_{int}^{vdW}, respectively) for ligand complexes with yeast Uba1 and human UBA1 (MLN7243). (C) Table for linear interaction energies for yeast and human UBA1-ligand complexes. The electrostatic ($<E_{int}^{el}>$) and van-der-Waals ($<E_{int}^{vdW}>$) interaction energies in kcal/mol are summarized for each ligand in the free and in the bound state. The gain in energy upon binding is given by the difference $\Delta E_{int}^{el/vdW} = <E_{int,bound}^{el/vdW}> - <E_{int,free}^{el/vdW}>$. The standard errors of the mean are ≤ 0.15 kcal/mol for all reported values.

Linear interaction energy (LIE) calculations (Åqvist et al., 2001) predicted MLN7243 to bind to yeast Uba1 slightly more potently than ABPA3, with regard to both the electrostatic (ΔE_{int}^{el}) and the van-der-Waals (ΔE_{int}^{vdW}) contributions (Figure 58B-C). These ligands exhibit a similar LIE profile in the complexed state. Nevertheless, since ABPA3 is capable of forming more suitable interactions with surrounding water molecules due to its higher hydrophilicity, its energy gain upon binding is attenuated. MLN4924 is clearly inferior to the other two ligands, mainly because of its missing hydroxyl group and the resulting inability to form electrostatic interactions on a comparable level. Since these results are in good qualitative agreement with the experimentally measured inhibitory values, MLN7243 was also investigated as the best ligand on the basis of the human UBA1 homology model published earlier (Correale et al., 2014). In this complex the electrostatic contribution was found to be slightly less favorable. Considering the highly conserved binding pocket, the reasons for this remain to be elucidated. Nevertheless, the calculations indicate that MLN7243 is likely to be a comparatively potent inhibitor for the human protein as well. However, it has to be taken into account that structural uncertainties in the homology model of the human UBA1 may affect the obtained results.

3.4.6 A548T mutation negatively affects MLN7243-binding to UBA1

A548 is highly conserved in the E1 family (Fig. 25) and mutation of the equivalent residue to either a Thr or an Asp in the human NEDD8 E1 (A171T/D) was found in MLN4924-resistant cell lines following prolonged MLN4924-treatment (Milhollen et al., 2012; Toth et al., 2012). The mutation of this, so called, gateway-keeper residue impaired NEDD8 E1 inhibition, due to faster off-rates of the covalent NEDD8•MLN4924-adduct from the mutant NEDD8 E1. Given the conservation of this alanine, its mutation to a larger residue could also be associated with drug resistance to the mechanistically similar UBA1-inhibitors ABPA3 and MLN7243. To investigate this, we purified human A580T mutant corresponding to the A171T mutant of NEDD8 E1 and investigated the inhibitory effect of MLN7243 against this mutant (Figure 59). Not surprisingly, the mutant also displayed a resistance phenotype in our *in vitro* assays. However, the difference in the IC_{50} values was observed to be ~6 fold higher in comparison to the wild type enzyme. This difference is much more drastic when compared to the two-fold increase reported for MLN4924 against the A171T mutant (Milhollen et al., 2012; Toth et al., 2012).

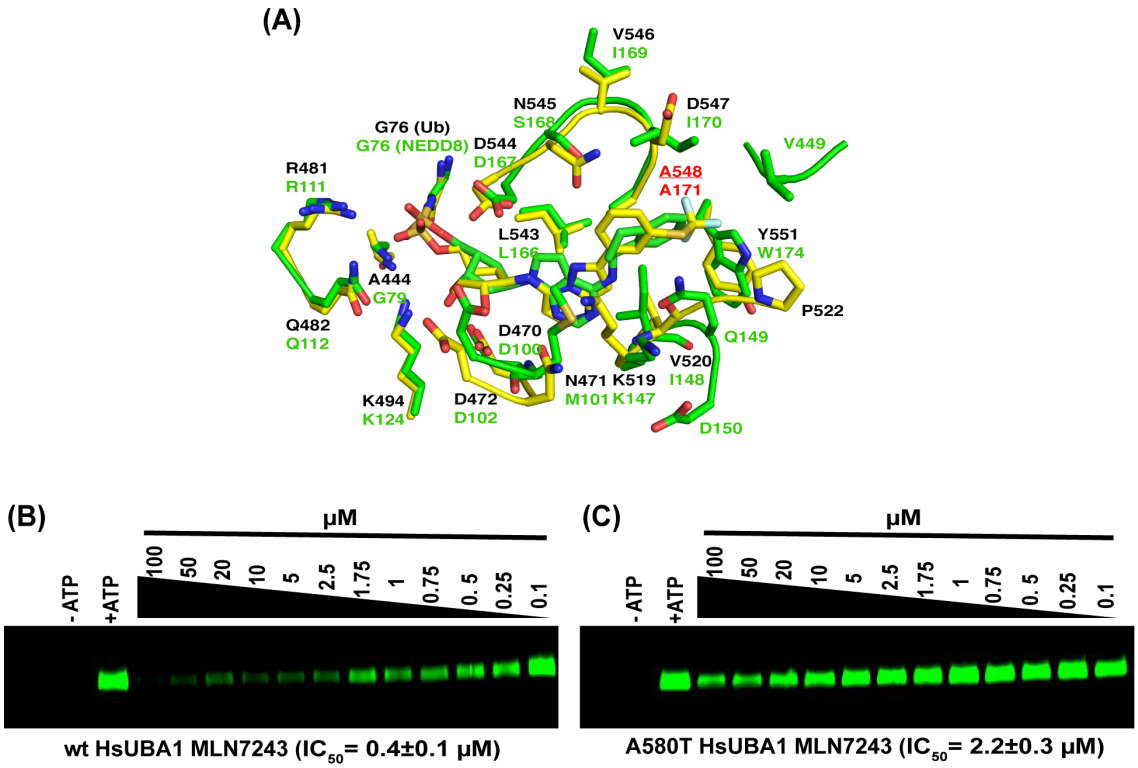


Figure 59: Inhibition assay with wtUBA1 and the A580T mutant, which corresponds to A171 of NEDD8E1 and A548 of *Sc*Uba1 (highlighted as red text in panel A). (B-C) UBA1~UB thioester inhibition assay representing the IC_{50} values for the wild type and A580T UBA1. The values represent the mean of duplicate measurements ± S.D.

Another inhibitor-resistant mutation reported for NEDD8 E1 is I310 which lies close to the C-terminus of NEDD8. As described, this mutant has a higher affinity for ATP while requiring more NEDD8 for activation. This residue is conserved in both UBA1 and the SUMO E1, therefore we imagine that this mutation will impart a similar resistance phenotype to both enzymes. Interestingly, when we analyzed this residue after superimposing the structures of the SUMO E1-AMP•SUMO adduct mimic (PDB code: 3KYC) and tetrahedral intermediate mimic (PDB code: 3KYD), it was obvious that this residue moves away from the C-terminus of SUMO1 in the latter structure. This indicates that this conserved isoleucine plays a role in Ub/ubl modifier binding to E1 enzymes rather than during catalysis. As this residue lies far from the ATP binding site in the known structures of E1 enzymes, it seems unlikely that this residue contributes significantly to inhibitor/ATP binding.

3.4.7 Unique features of the ATP-binding pocket of E1 enzymes can serve as specificity regulators

Although the adenylation site of the E1 enzymes is highly conserved, especially in the ATP binding pocket region, with key residues responsible for ATP coordination being the same, the surrounding residues in the pocket display variations in their amino acid compositions. When we look closely in this region, we find striking differences in the pocket lining the nucleobases. As the nucleobases serve as specificity determinants as seen for MLN7243, it should be possible to develop highly specific, adenosyl sulfamate-based E1-inhibitors using structure-based drug design approaches,.

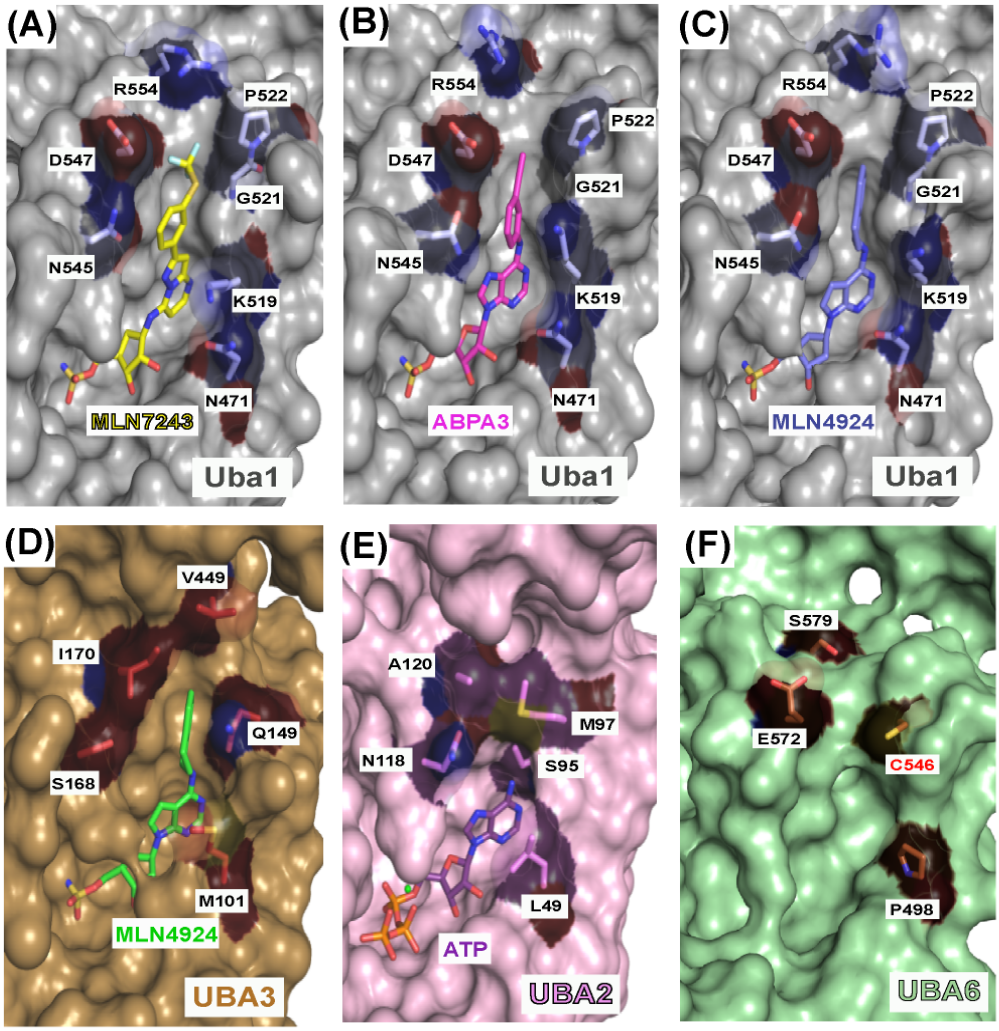


Figure 60: Surface representation of the nucleobase-binding pockets of different E1 enzymes. Binding pockets with key differences being highlighted are shown for (A) MLN7243 (yellow), (B) ABPA3 (magenta), (C) MLN4924 (purple) bound to Uba1 (gray surface), (D) MLN4924 in complex with the NEDD8 E1 (brown surface) [PDB code: 3GZN], (E) ATP-bound SUMO E1 (pink surface) [PDB code: 1Y8Q] and of an apo-UBA6 homology model (green surface).

Differences in the ATP-binding pockets of different E1 enzymes can thus be exploited by complementing the unique pharmacophores presented by each of these enzymes (Figure 60). For example, R554 in Uba1 can be targeted by extending MLN7243 with a hydrogen-bond acceptor, or even a negatively charged group, thus rendering the inhibitor more specific for UBA1. Remarkable, the homology model of UBA6 presents a surface exposed cysteine residue in this pocket. This cysteine can be targeted for covalent binding inhibitors as reported for kinases (Lanning et al., 2014). The E1 ATP-binding pockets also differ in their secondary structure elements, which influence the respective available E1 surface interacting with the inhibitors (Figure 61). Whereas in Uba1 the loop between β-strand 19 and α-helix 26 is interrupted by a 3_{10} helix and features three β-turns and one γ-turn, this loop is shorter in UBA3 (PDB code: 3GZN) and UBA2 (PDB code: 1Y8Q) and only contains one γ-turn and two β-turns as well three β-turns and one γ-turn, respectively. Atg7 (PDB code: 3VH4) has the longest insertion in this region, which folds into the longer α-helix 6 preceded by one γ-turn and two β-turns. UBA5 (PDB code: 3H8V) (Bacik et al., 2010) represents the second longest insertion with α-helix 4 being framed by three β-turns. In the SUMO and NEDD8 E1 enzymes there is an additional contact with the inhibitor provided by the UFD (represented by residues 474-478 and 448-451, respectively) to this pocket (Figure 61A). As illustrated in detail in chapter 2, the UFD has been shown to undergo domain movement from a closed conformation to an open conformation to accommodate the E2 enzyme in the site between the catalytic cysteine domain and the UFD itself. Assuming that this domain oscillates between the open and closed states, the closed state will induce clashes between this residue and the inhibitor-binding pocket, thereby lowering the affinity of this compound. When we superimpose the NEDD8 E1-bound NEDD8•MLN4924 adduct (PDB code: 3GZN) which represents a closed conformation of the UFD with the E2-bound NEDD8 E1 structure (PDB code: 2NVU), which depicts an open state of this domain, we find residues from this domain in close proximity to the inhibitor pocket in the closed state and Asp89 from the E2 enzyme posing a charge mismatch with this pocket in the open state. Since in the known structures of Uba1 the UFD has not yet been visualized in proximity of the ATP-binding pocket, it is tempting to speculate whether the residues from the UFD can promote/obstruct binding of the adenosyl sulfamate inhibitors in case of the SUMO E1 and NEDD8 E1. Additionally, in the superposition (Figure 61A) M97 of the SUMO E1 and M405 of Atg7 pose clashes to the N6-substituents in MLN4924 and ABPA3 as well as to the CF_3S-substituted phenyl group of MLN7243, hence providing possible explanations as to why these compounds are ineffective against these two enzymes (Brownell et al., 2010; An and Statsyuk, 2013; An and Statsyuk, 2015).

Interestingly, the adenosyl sulfamate ABP1 (An and Statsyuk, 2013), in which the propargyl-group is directly linked to the exocyclic adenine N6-nitrogen (Figure 55B), displayed a global decrease of Ub, NEDD8 and SUMO conjugation levels in cell-based assays, whereas ABPA3, instead of inhibiting, rather activated SUMO-conjugation in cells (An and Statsyuk, 2015). In the ATP-bound SUMO E1 structure (Lois and Lima, 2005) (PDB code: 1Y8Q) M97 of SAE2 is positioned just above the exocyclic amino-group of the adenine, and presumably forms unfavorable interactions with the N10-ethynylphenyl-group of ABPA3 (Figure 61A). Most likely, the smaller propargyl-group of ABP1 fits better into the SUMO E1 pocket, leading to SUMO E1 inhibition *in cellulo*. In summary, similar to ATP-competitive kinase inhibitors, variations in the E1 ATP-binding pockets and inhibitor structures impart their potency and selectivity (Knight and Shokat, 2005). As mentioned above, R554, which is conserved in human UBA1, can be utilized as an additional pharmacophore of Uba1 to extend MLN7243 in this direction and achieve higher specificity towards Uba1.

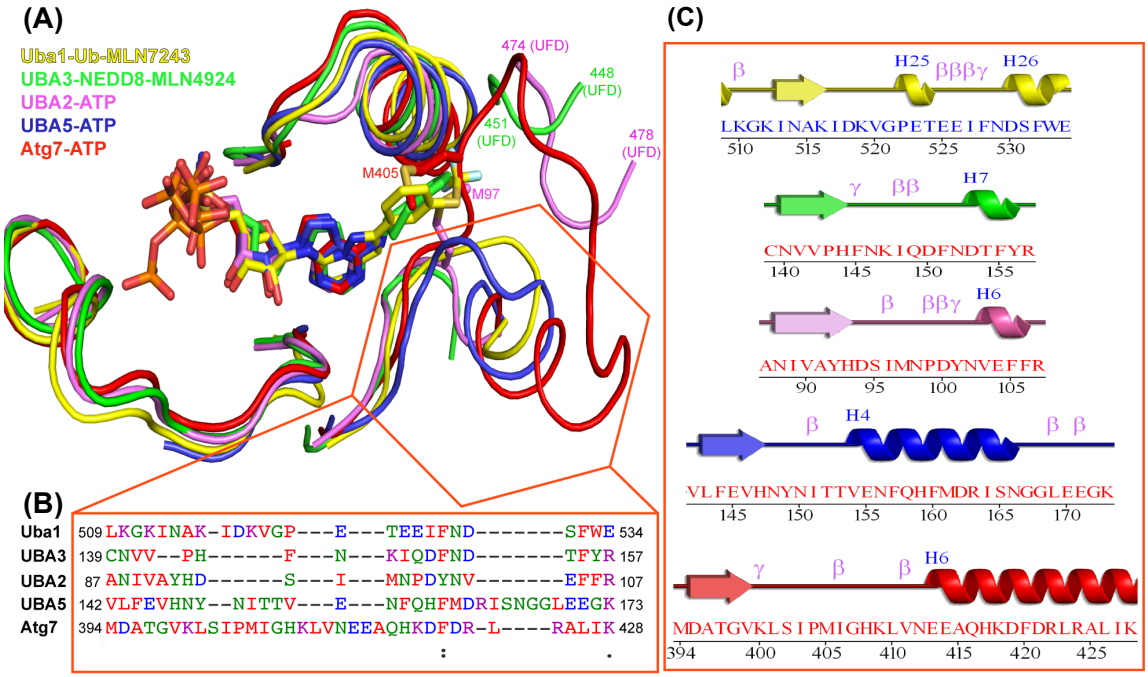


Figure 61: Differences at the adenylation active site among E1 enzymes.
(A) Superposition of the adenylation site of various E1 enzymes bound to either an inhibitor or ATP. (B) The most divergent region is highlighted by the red hexagon and the corresponding primary sequences are aligned. (C) The secondary structure composition for this region as generated by PDBsum (de Beer et al., 2014) indicates diverging secondary structure elements in the E1 family. These differences in the nucleotide-binding pockets can be exploited to achieve higher selectivity towards individual E1 enzymes.

3.5 CONCLUSION

In this chapter, insights into the observed specificities of inhibitors targeting the ubiquitin activating and related E1 enzymes via crystal structures of the specific NEDD8-E1 inhibitor MLN4924, the dual NEDD8/Ub E1 inhibitor ABPA3, and the selective Ub-E1 inhibitor MLN7243 bound to yeast Uba1 are presented and these structures are augmented with biochemical and computational studies. Although our crystal structures quite accurately depict the bound state of ubiquitin•inhibitor adduct to Uba1, due to the complex mechanism of inhibition of these compounds, it is hard to describe all specificity determinants of these compounds, especially when the structures of the transition states of the enzyme bound to inhibitors are unavailable where the catalytic cysteine domain is expected to be in close proximity to the adenylation domain. However, considering that the total rate (including the rate-limiting step) of Ub•inhibitor formation is much faster than the rate of Ub•inhibitor release, it is clear that a high binding affinity of the final Ub•inhibitor adduct to the E1 enzyme will undoubtedly contribute significantly to its specificity.

In the closed state presented in the SUMO E1 and, more recently, Ub E1, H21 corresponding to residues 475 to 481 move away from the adenylation site, therefore residues R481, N478 and Q482 which are responsible for phosphate backbone coordination are not positioned in the closed conformation to stabilize ATP (Lv et al., 2017b; Olsen et al., 2010). Furthermore, the N-terminal helices H1 and H2 are disordered, most likely moving away from the adenylation site due to the clash arising with the SCCH domain, taking R21 out of the picture. As shown in the Biomol green assay, the double mutant R21/ 481A impairs the adenylation activity of the enzyme and, in the closed state, both residues are distal from the adenylation site, giving rise to the notion that the closed state has an impaired binding to ATP. It is tempting to speculate that this is the reason why adenosyl sulfamate inhibitors occupy the ATP binding site in the closed state of the enzyme as they do not require either R21 or R481 to be present in this pocket. Upon superimposition of the MLN7243•Ub adduct bound Uba1 structure with the NSC624206 bound *Sp*Uba1 structure (PDB entry 5UM6) representing the closed state of the enzyme, we observe serious clashes for the nucleobase and the ribose moiety of the inhibitor with residues N538 (N545 for ScUba1), E540 (D547), Y544 (Y551) and the main chain carbonyl group of V513 (V520) colliding with the nucleobase and residues L536 (L543) and D463 (D470) interfering with the ribose (Figure 62). This observation suggests that although this structure provides valuable insights into the conformations adopted in the closed state of Uba1, it is not compatible with the substrate/inhibitor-bound form of the enzyme.

Both MLN7243 and MLN4924 are currently in phase 1 and phase 1b clinical trials, respectively, for the treatment of advanced solid tumors (Sarantopoulos et al., 2016), thus underlining their therapeutic potential. Our studies can be extended further to design more potent and selective E1 inhibitors. While UBA1 inhibition has been proposed to overcome the limitation of proteasomal inhibitors against solid tumors, both SUMO and NEDD8 E1s have already been shown to be potential targets for cancer therapy. In addition, specific inhibition of individual E1s can further improve our

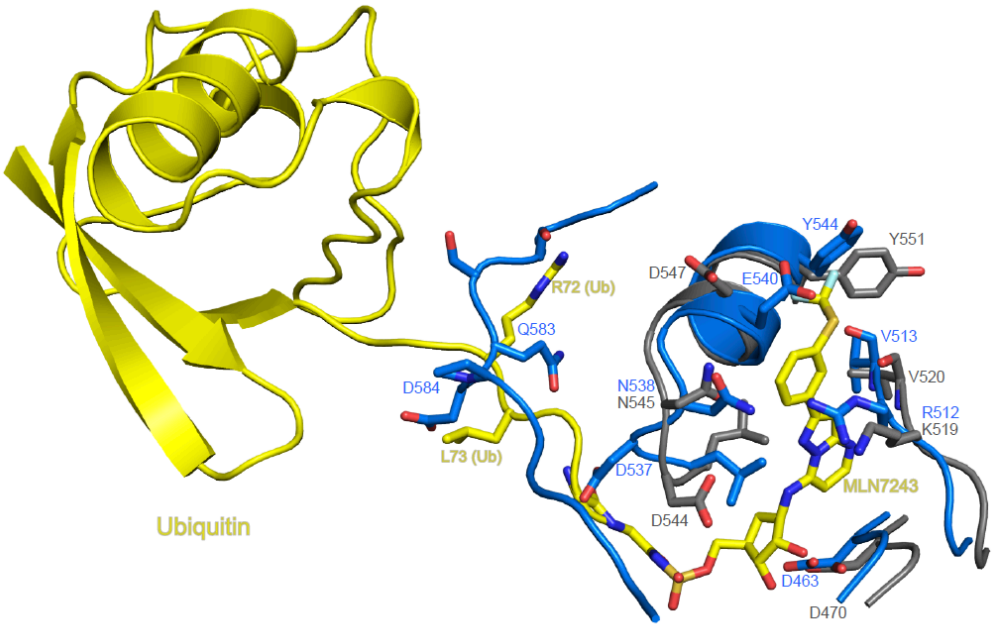


Figure 62: Representation of clashes between the MLN7243•ubiquitin adduct (yellow) in complex with *Sc*Uba1 (grey, PDB code: 5L6J) and *Sp*Uba1-NSC624206 (blue, PDB code: 5UM6) where the latter represents the closed state of the enzyme.

understanding of the diverse roles Ub/Ubl post-translational modifications play in eukaryotes.

In summary, our structures can be used to design derivative inhibitors featuring nucleobase and ribose isosteres that display improved pharmacokinetic properties (St Jean and Fotsch, 2012), potencies and selectivities. The MLN4924•NEDD8 adduct has been shown to not be able to inhibit UBA1, thus, by combining adenosyl sulfamtes with the C-terminal tails of Ub/Ubl modifiers, one can impart high specificity for individual E1 enzymes to these inhibitors. As the inhibitor-resistant mutants reported so far are either located in the ATP-binding pocket or the Ubl-binding site, developing inhibitors that target both these sites will offer a significant advantage to prevent the evolution of resistance mutants. The reported inhibitor resistant mutants should be taken into account to develop second-generation inhibitors which may turn out to be very effective

in combination therapies. While on the one hand very specific E1 inhibitors will greatly improve our understanding of the Ub/Ubl post-translational modifications, they might be disadvantageous from a therapeutic perspective, as treatment with a highly specific inhibitor targeting a single E1 enzyme will presumably lead to the more rapid emergence of resistant mutations. To overcome this limitation pan-specific E1 inhibitors like ABPA3 and compound I, targeting two or more E1 enzyme may be more advantageous since these compounds would be more difficult to overcome via the occurrence of mutations. The structures presented here together with those of MLN4924•NEDD8 bound to the NEDD8 E1, SUMO•AMP mimicking intermediates bound to SUMO E1 (Olsen et al., 2010), ATP-UBA5 (Bacik et al., 2010; Oweis et al., 2016) and yeast Atg7CTD-Atg8-MgATP (Noda et al., 2011) provide a platform to develop potent adenosyl sulfamate inhibitors targeting individual E1s which will enable the improvement of these inhibitors for basic research and drug discovery purposes.

4. CONCLUDING DISCUSSION

This thesis presents comprehensive insights into the catalytic cycle of UBA1 with the help of high resolution crystal structures and functional studies. In addition, it rationalizes how different adenosyl sulfamate inhibitors achieve diverse specificities for E1 enzymes. While the former addresses a fundamental problem in enzymology, the latter provides a perspective into translational research. UBA1 is an interesting example of evolutionary enhancement of a basic enzymatic transformation, namely the adenylation at the C-terminus of a peptide, by incorporating additional features such as the thioesterification capability and interaction sites for downstream components of the same pathway to evolve the initial reaction in a pathway, which is present in all eukaryotes. Moreover, this increases the efficiency of the enzyme-catalyzed reaction due to the integration of the steps following adenylation in a single but multifunctional enzyme. In addition to assigning different functionalities to individual domains connected via flexible linkers, the substrates and products of the individual steps can regulate the catalytic activity of the enzyme by acting as sensors for interdomain communication. Our structures illustrate that this mechanism is an important part of the catalytic cycle of the enzyme.

“Knowing the enemy is important before we plan to defeat it”, this saying comes to mind when we realize how adenosyl sulfamates utilize the enzymatic mechanism of Uba1 to inhibit its activity. Adenosyl sulfamates require the product of the second catalytic half reaction of UBA1, *i. e*. the Uba1~Ub thioester to be able to inhibit the activity of the enzyme. This way the compound utilizes the enzymatic activity to trap it in an inhibitory state. Several E1 enzymes have been reported as valid targets for anti-cancer therapeutics. MLN4924, the first adenosyl sulfamate to be reported is a potent inhibitor of the NEDD8 activating enzyme (Soucy et al., 2009). This compound has advanced to a phase 1b trial for the treatment of acute myelogenous leukemia underlining its therapeutic potential. The SUMO E1 has been shown to enable cells to tolerate Myc hyperactivation (Kessler et al., 2012). Myc is an oncogene, which is required for tumor maintenance and progression in several types of malignancies. Thus, inhibition of the SUMO E1 in Myc-driven tumors appears to be a promising strategy against cancer. FAT10, the ubiquitin-like modifier activated by UBA6 is overexpressed in several types of cancer including glioma, colorectal, liver or gastric cancer (Liu et al., 2014; Sun et al., 2014; Yuan et al., 2012). FAT10ylation of tumor the suppressor protein p53 is a signal for proteasomal degradation. Increasing evidence suggests that FAT10 itself promotes tumorigenesis by serving as a proto-oncogene and its upregulation during inflammatory processes is involved in tumorigenesis. Therefore, by targeting UBA6, the only E1 enzyme in the FATylation pathway, represents a novel avenue to overcome

carcinogenesis. Likewise, ATG7 and UBA5 have also been proposed to be critical for cancer cell survival and proliferation (da Silva et al., 2016; da Silva et al., 2013). Therefore, a rational drug design targeting individual E1 enzymes using available structures and the adenosine sulfamate scaffold represents a promising opportunity to identify new derivatives of this compound class that can shut down individual Ub/Ubl pathways. This will not only be desirable from a human health perspective but having such inhibitors will also help to better understand those ubiquitin-like systems, like UFMylation and FATylation, which are still underexplored due to their late discovery.

5. FUTURE PERSPECTIVE

5.1 Tetrahedral transition state for Uba1

The catalytic intermediate state of Uba1 where the catalytic cysteine attacks the ubiquitin C-terminus to form the ubiquitin thioester is an interesting state of the enzyme which has not yet been structurally visualized. Although a recently published structure of *Sp*Uba1 in its apo-form represents a closed state of the enzyme, the active site cysteine is still 23 Å away from the catalytic center in this structure. Moreover, upon superimposition of this structure with MLN7243-ubiquitin-bound Uba1 structure, we could visualize serious clashes for both ubiquitin and the ligand. This suggests that the enzyme in the substrate bound form will adopt altered arrangements in these regions. Therefore, we want to crystallize Uba1 in the presence of an electrophilic probe that can capture the enzyme in its closed conformation while trapping the catalytic cysteine in a tetrahedral intermediate state.

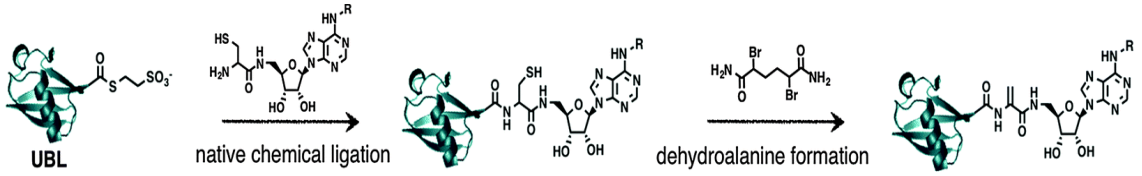


Figure 63: Synthesis of Ub probe 3 designed to capture Ub E1 in its tetrahedral intermediate state.

For this purpose, we have collaborated with Dr. Alexander Statsyuk at the University of Houston, Texas, who has reported a chemical probe capable of trapping the enzyme in its transition state (An and Statsyuk, 2016). This compound is named Ub probe 3 and its synthesis is shown in Figure 63. We will try to co-crystallize Uba1 in the presence of this probe in hope to determine a transition state structure of the enzyme, which will complete the structural characterization of the catalytic states of Ub E1 by providing this missing piece. Moreover, this structure will give us important insights into additional specificity determinants for adenosyl sulfamate inhibitors as this will visualize the closed conformation of the enzyme with which the inhibitors have to initially interact with.

5.2 Fragment based screening

Fragment based drug discovery is an emerging approach of finding new compounds that bind and modulate the activity of a protein of interest (Erlanson et al., 2016). This method presents several advantages over traditional HTS strategy. First, most fragment libraries consist of hundreds to few thousands of compounds in comparison to the multimillion compound libraries used in HTS. This renders the usage and maintenance

of the library tremendously simpler. Second, the compounds of these libraries are smaller, having less than 20 non-hydrogen atoms, and are therefore termed fragments. The use of fragments enables one to cover a much larger chemical space with a much smaller number of compounds. Fragments typically have three additional properties: They have fewer than three hydrogen bond acceptors and donors, fewer than three rotatable bonds and a clogP of three or below translating into a much higher solubility. Together these properties are called 'the rule of 3' for selecting fragments analogous to Lipinski's rule of 5 for drug-like compounds.

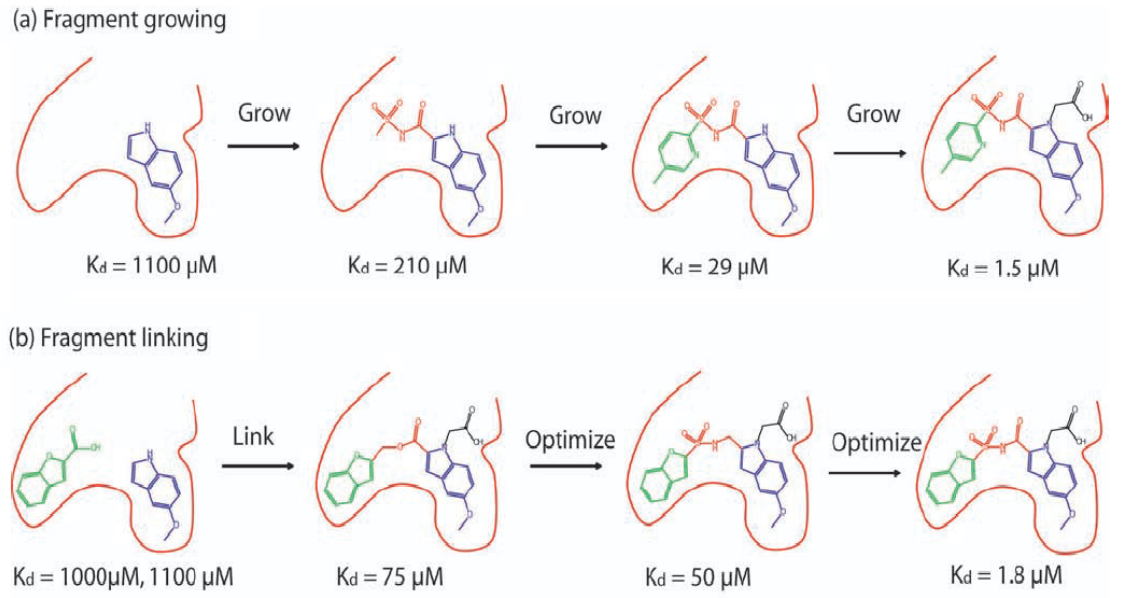


Figure 64: Fundamental concepts of fragment based drug discovery. The affinity of fragments towards the target can be improved by fragment growing or linking.

Due to the presence of fewer interactions, fragments usually display low affinities but their high solubility can be used to compensate for the former limitation. Fragments can be grown in the respective binding pocket to achieve higher affinities and specificities. If multiple fragments are bound near the same binding pocket, they can be linked to accomplish the same goal.

We are in the process of carrying out a fragment based screening campaign on Uba1 in collaboration with Dr. Manfred Weiss at Helmholtz Zentrum Berlin. Our collaborators have developed a fragment library consisting of one hundred compounds (Huschmann et al., 2016) which were tested for their ability to bind to Uba1 in thermal shift and X-ray crystallographic studies. The preliminary experiments identified L-canavanine as promising lead compoud, which binds to Uba1 and inhibits its activity.

5.2.1 Thermal Shift Assays

Thermal-shift assays represent a very fast and efficient way to determine whether specific compounds stabilize or destabilize a target protein. Thermofluor is used in order to get a first idea of any interaction between a fragment and a target protein since the binding of a compound to a protein typically results in a stabilization, however, binding of some compounds leads to destabilization. The fluorescence of a dye interacting with hydrophobic regions of the proteins is used as the read-out to determine the melting temperature of the protein of interest. For this method SYPRO orange (SYPRo; Sigma-Aldrich) is typically used, which displays excitation and emission maxima at 490 nm and 594 nm, respectively, when bound to the proteins. Most hydrophobic residues are buried in the interior of the protein and therefore SYPRo is not able to bind until the protein is in its unfolded state.

As a positive control ATP was used which showed a stabilizing effect coupled to an increase in the melting temperature of 3.5 °C when present at a concentration of 9.2 mM (see Figure 65A). The aforementioned hit, L-canavanine also showed stabilization effect on Uba1, however, a concentration of 184 mM was needed to achieve a thermal shift of roughly 2 °C (Figure 65B) coupled to a change from a one-state to a two-state transition.

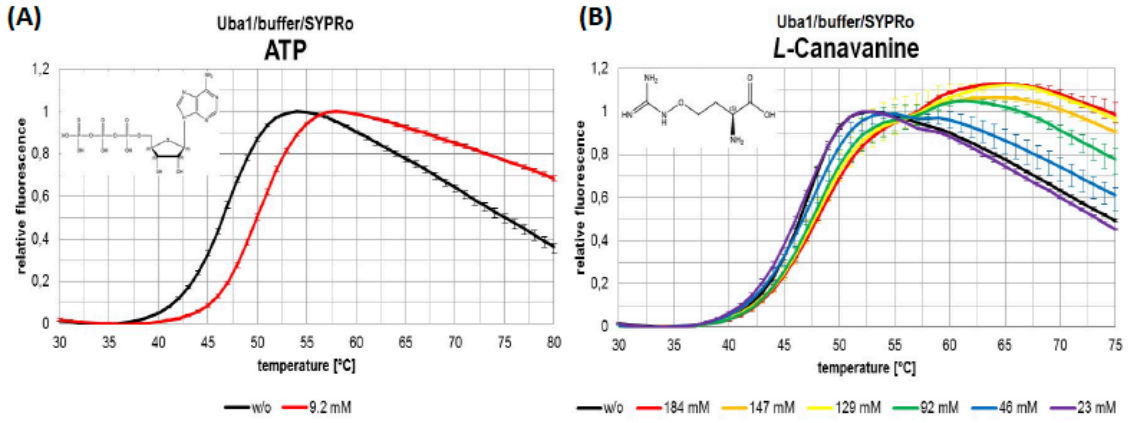


Figure 65: The observed thermal shifts of Uba1 when incubated with (A) ATP and (B) L-canavanine.

5.2.2 X-ray structures of fragment bound complexes

X-ray crystallography is the ideal approach to find the position of fragments that are bound to a protein. However, because of the small size of the fragments, it is usually not easy to find bound fragments especially when diffraction data only extend to low resolution. This problem is further exacerbated if the occupancy of the compound is low as a consequence of either a low solubility and/or a weak affinity of the fragment to the target protein. We could obtain two crystal structures of L-canavanine bound to Uba1

where in one case crystallization was carried out in the presence of ATP and in the other ubiquitin was included instead of ATP. These crystals diffracted to resolutions of 3.5 Å and 3.1 Å respectively.

In the absence of Ub, the non-proteinogenic amino acid *L*-canavanine binds into the ATP-binding pocket at exactly the position where the triphosphate moiety would be located (Figure 66A-B). No electron density for the base or the ribose of ATP could be observed in this dataset, although the compound was present at a concentration of 50 mM (Figure 67B), which was apparently displaced by L-canavanine due to the compound being present at a much higher concentration of 2.5 mM. L-canavanine was stabilized by several hydrogen bonds which were formed with Arg21, Arg481, Asn478 and Asp544 of Uba1 which usually interact with the triphosphate moiety of ATP (D544 via Mg^{2+} coordination).

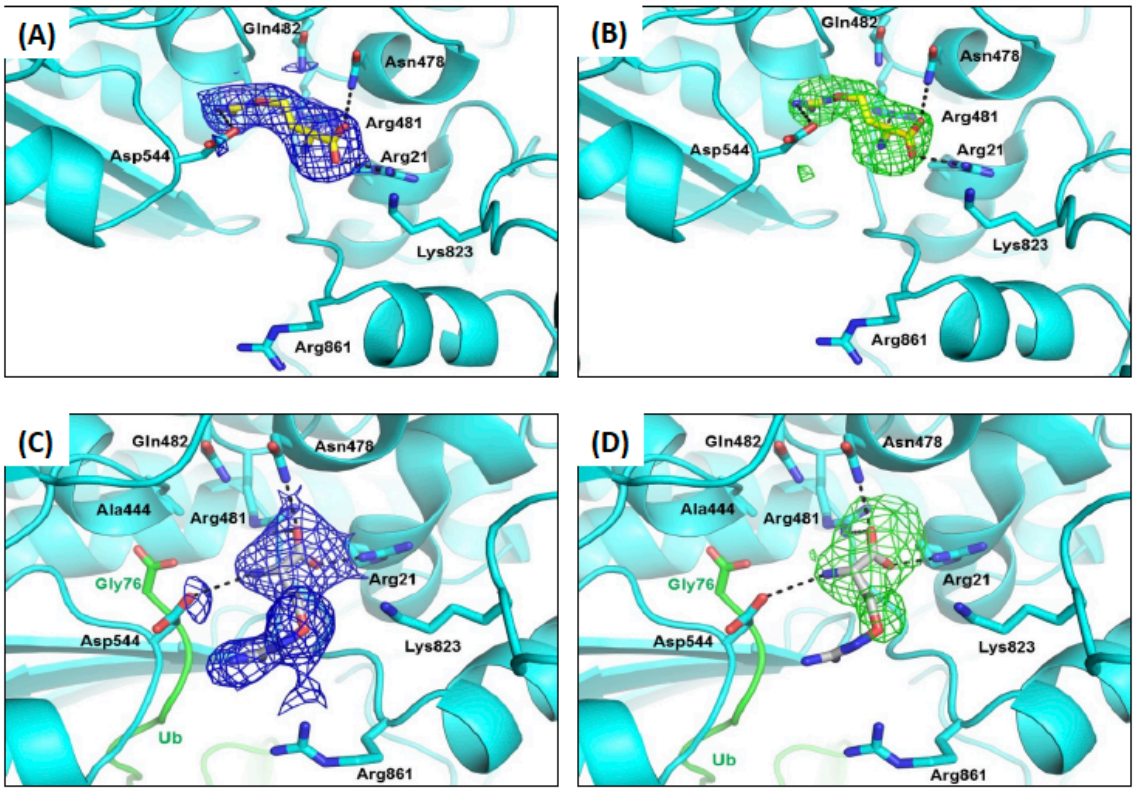


Figure 66: (A-B) $2F_o$-F_c and omit maps of *L*-canavanine in the Uba1-ATP crystals at a resolution of 3.5 Å. (A) illustrates the $2F_o$-F_c map at one times the rmsd and a carve radius of 0.7 Å. (B) shows the F_o-F_c omit map displayed at three times the rmsd and a carve radius of 0.7 Å. *L*-Canavanine is represented as sticks and its C-atoms are colored in yellow. Hydrogen bonds are indicated by dashed lines. (C-D) $2F_o$-F_c and omit maps of *L*-canavanine in the Uba1-Ub crystals at resolution of 3.1 Å. (C) illustrates the $2F_o$-F_c map at one times the rmsd and a carve radius of 1.3 Å. (B) shows the F_o-F_c omit map displayed at three times the rms deviation and a carve radius of 2 Å. *L*-Canavanine is represented as sticks and its C-atoms are colored in grey. Hydrogen bonds are indicated by dashed lines.

The second crystal form in which L-canavanince could be located represents the Uba1-Ub complex where ATP is absent from the crystallizations conditions. Again, *L*-canavanine occupied the ATP-binding pocket (Figure 66C-D). However, the guanidinium group of the fragment adopted a different orientation in this structure, presumably enforced by a clash with G76, the C-terminal residue of Ub (see superimposition of both *L*-canavanines in Figure 67B). While the interactions of the carboxyl group of *L*-canavanine with Uba1's Arg21 and Asn478 stayed intact, the hydrogen bond between the guanidinium group to Arg481 was lost. Instead, a hydrogen bond interaction with Asp544 was formed and Arg481 additionally stabilized the carboxyl group of *L*-canavanine. As indicated by its missing density in Figure 66D, the guanidinium group might be flexible, striving to form hydrogen bonds with either Asp544 or Arg861 with both residues being present at a distance of ~4 Å. The latter residue is forced to move closer towards the ATP-binding pocket, thus making the interaction with *L*-canavanine possible. The movement of Arg861 (colored in cyan and purple in Figure 67), was presumably caused by the binding of Ub to Uba1 changing the side-chain conformation of this residue as also observed in the ubiquitin bound Uba1 structure.

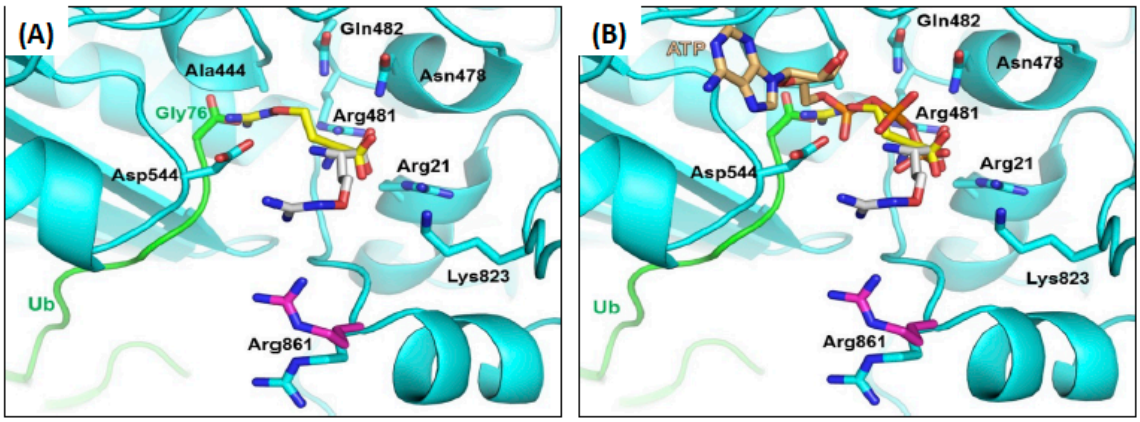


Figure 67: Superimposition of *L*-canavanine bound in the presence and absence of ATP. (A) Both *L*-canavanines (from the Uba1-ATP crystal in yellow and the Uba1-Ub crystal in grey) are shown as sticks. Ub is colored in green and is illustrated as a cartoon with the C-terminus represented as sticks. Arg861 is shown in cyan and purple in the Uba1-ATP and Uba1-Ub structures, respectively. In (B), ATP is illustrated as sticks and colored in bronze while it is absent in (A) to show how *L*-canavanine occupies the same binding site as the phosphate groups.

5.2.3 Activity assays

L-canavanine differs from *L*-arginine in the replacement of the δ-carbon atom with an oxygen atom (Figure 68, top). As L-canavanine was found to bind in the ATP binding pocket of Uba1, we expected it to show an inhibitory effect on the activity of the enzyme in an ATP-competitive manner. We tested both amino acids (*L*-canavanine and *L*-

arginine) in our UBA1~Ub thioester assay to check whether they can modulate the activity of the enzyme (Figure 68). Interestingly, while L-canavanine showed a dose-dependent inhibition starting at 10 mM in the presence of 0.5 mM ATP, L-arginine did not show any effect on Uba1 activity in this concentration range. Moreover, the inhibition by L-canavanine was lost when a higher concentration of ATP (2.5 mM) was used (data not shown).

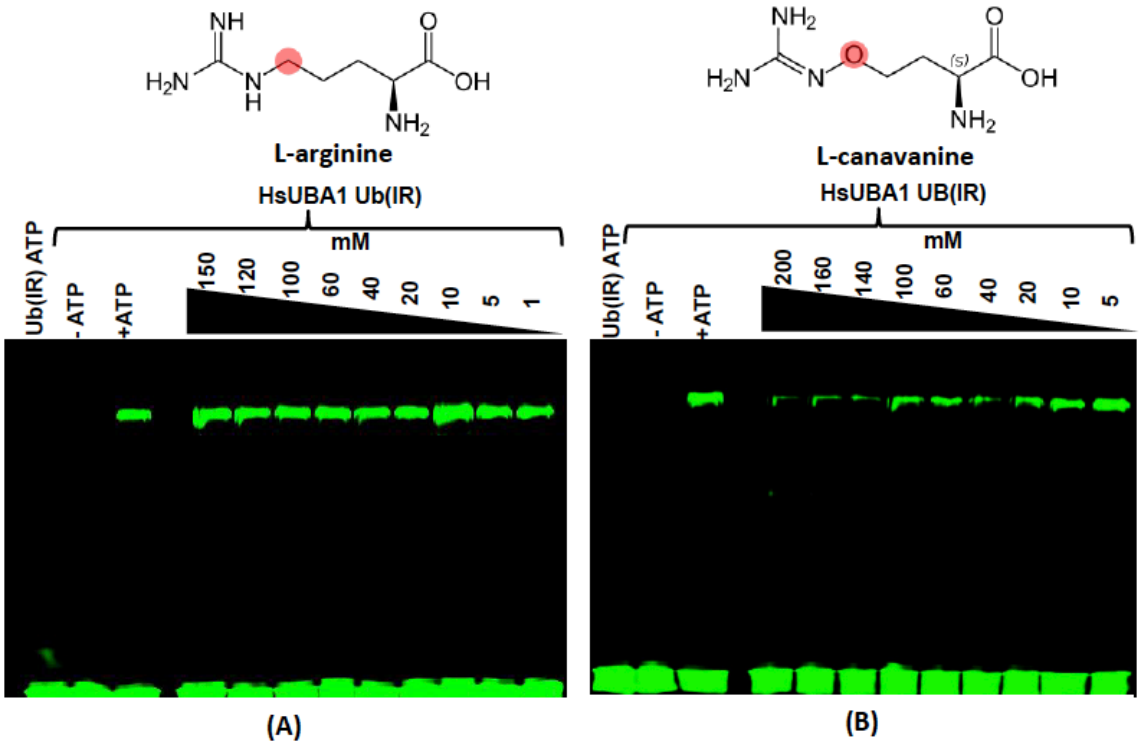


Figure 68: Dose-dependent activity assay with L-arginine and L-canavanine. Ub(IR labelled) and ATP concentrations were reduced to 5 µM and 0.5 mM, respectively. The fragments were incubated with *Hs*UBA1 for 10 min before Ub(IR) was added. The incubation was performed for 60 min at 37 °C. After that, 10 µl of SDS sample buffer including TCEP was added. 6 µl of each sample were loaded on a 12% SDS-PAGE and scanned at 800 nm to detect IR-labeled Ub. The strong bands near the bottom of the gel correspond to free IR-labeled Ub.

5.2.4 Fragment evolution

One limitation of fragment based screening is that it further is needed to transform the fragment into a drug or drug-like compound that may show significant affinity and specificity toward the protein of interest. This includes synthesis of derivatives of the initial fragment compounds with additional functional groups and then performing detailed structure-activity relationship (SAR) studies with these compounds. Both fragment growing and linking methods can be utilized where the latter requires a proximity of two initial fragments. As L-canavanine is bound in the ATP-binding pocket several possibilities for linking are obvious including adenosine, ubiquitin and adenosyl

sulfamates. We first want to combine canavanine with the C-terminus of ubiquitin as both have amino acid features. Moreover, the heptameric C-terminus peptides and its derivatives have previously been shown to inhibit UBA1 activity (Zhao et al., 2012). Peptides that combine L-canavanine with the C-terminal VLRLRGG sequence of ubiquitin or variations thereof will be used as a starting point (Figure 69B).

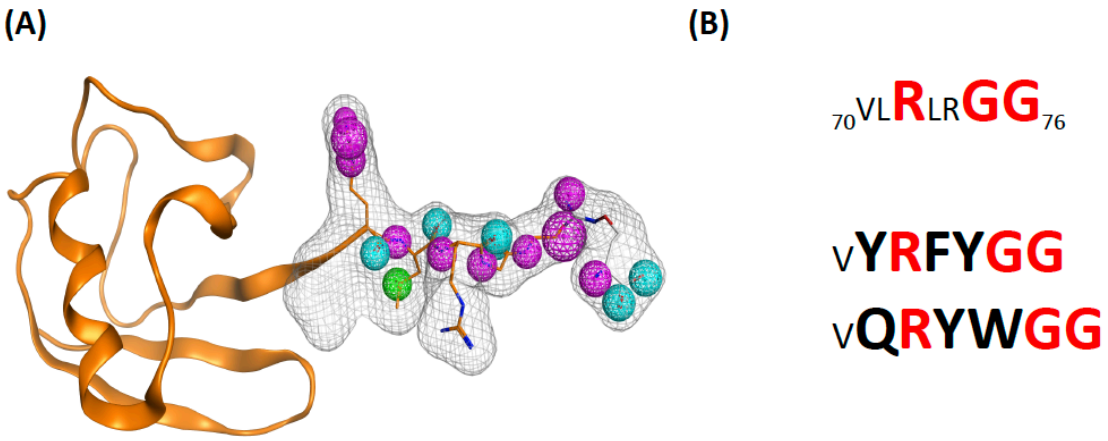


Figure 69: (A) Pharmacophore design combining L-canavanine (grey) and the C-terminal amino acids of ubiquitin (orange). The magenta spheres represent hydrogen bond donors while the cyan spheres show hydrogen bond donor groups. The green sphere reflects a hydrophobic region. (B) The natural C-terminal sequence of ubiquitin corresponding to residues 70-76 and its variants that were reported to show improved inhibition (Zhao et al., 2012).

Further, we plan to search peptidomimentic compounds that mimic the ubiquitin C-terminus coupled to L-canavanine. For this purpose, we have designed pharmacophore model for this region (Figure 69A), which can be searched against available compound databases like ZINC and emolecules (Irwin and Shoichet, 2005). The variants of the ubiquitin C-terminus that performed better in inhibition assays can also be included for pharmacophore design to find more hits in this search.

5.3 Targeting UBA1 of eukaryotic pathogens

As ubiquitin is highly conserved and essential throughout eukaryotes, targeting the Uba1 of eukaryotic pathogens may be an effective strategy to control several vector-born diseases like malaria, sleeping sickness and leishmaniasis. This approach can also help in controlling fungal infections like candidiasis. Figure 70 highlights the conserved features in the context of the yeast Uba1 structure. Not surprisingly, the adenylation active site including the ATP-binding pocket of and the region around the catalytic cysteine show high conservation, but other than that it appears to be quite variable, especially in the IAD and 4HB. Specifically, targeting the regions that are different form the human ortholog may lead to possible therapeutic advantage.

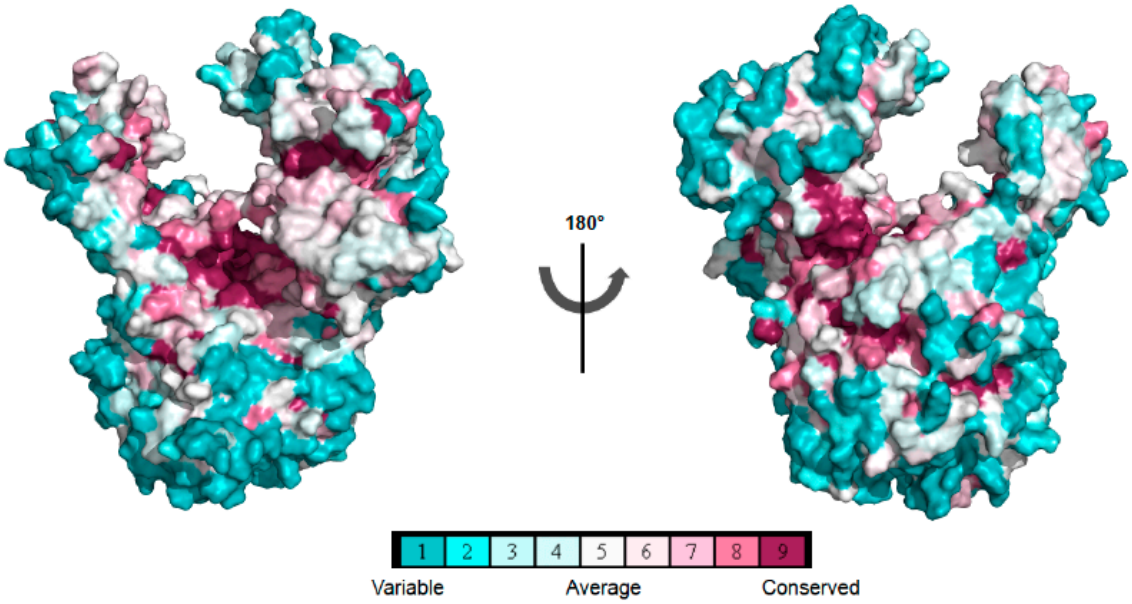


Figure 70: Depiction of surface conservation among six Uba1 sequences displayed onto the *Sc*Uba1 structure as calculated with the Consurf server. The most conserved regions are colored in red. The cluster of highly conserved region on the left side represents the adenylation active site and the same on the right side marks region around the catalytic cysteine at the SCCH domain.

The sequence conservation between Uba1 of four eukaryotic pathogens and their counterparts in *Saccharomyces cerevisiae* and humans indicates the existence of pathogen-specific regions of Uba1. This endeavor will undoubtedly benefit from structural studies of the ubiquitin activating enzymes derived from these parasitic organisms. In their absence the available homology models might still be suitable starting points for based drug design efforts, which could be combined with HTS or fragment based screening approaches utilizing the recombinantly produced enzymes.

	***Sc*Uba1**	***Hs*UBA1**
***Plasmodium falciparum* Uba1**	36% identity 53% similarity	35% identity 55% similarity
***Trypanosoma brucei* Uba1**	38% identity 54% similarity	38% identity 54% similarity
***Leishmania sp.* Uba1**	37% identity 52% similarity	36% identity 53% similarity
***Candida albicans* Uba1**	67% identity 83% similarity	53% identity 70% similarity

Table 10: Summary of identity and similarity scores for four eukaryotic pathogens with either *Saccharomyces cerevisiae* Uba1 or human UBA1.

Comparison of the ubiquitin system with the best-studied post-translational modification system, phosphorylation, yields strikingly similar numbers of enzymes involved in these processes (Cohen and Tcherpakov, 2010). Whereas phosphorylation is comprised of more than five hundred protein kinases and approximately one hundred phosphatases, ubiquitylation consists of around six hundred enzymes (E1, E2 and E3) and hundred deubiquitinases. Gleevec or imatinib, tyrosine kinase inhibitors used for the treatment of chronic myelogenous leukemia and acute lymphocytic leukemia, respectively as well as bortezomib, the proteasome inhibitor discussed earlier, were both approved by the FDA around the same time (Gleevec in 2001 and bortezomib in 2003). However, as per now more than twenty protein kinases inhibitors are in use for the treatment of various kinds of malignancies, whereas the inhibitors of the UPS used for medicinal purposes only sum up to 3, all of which are proteasome inhibitors. This huge gap between the two systems does not correlate with the role and significance of UPS in human health conditions. We envision that targeting the E1 enzymes with adenosyl sulfonamides can narrow this gap and pave the way for novel methods for the treatment of various cancer types. Interestingly, kinases, the proteasome and E1 enzymes, all utilize ATP as an energy source. The huge array of ATP analogs developed for kinase inhibition can therefore serve as starting points for inhibiting ATPase activity of the proteasome associated with the regulatory 19 S particle and block adenylation of Ub/Ubl proteins by E1 enzymes. In addition, a search for allosteric inhibitors targeting regions other than the ATP-binding pocket through high throughput screening (HTS) methods may lead to interesting compounds.

6. LIST OF PDB ENTRIES GENERATED IN THIS THESIS

PDB code	Description
5L6H	**Uba1 in complex with Ub-ABPA3 covalent adduct at 2.3 Å resolution**
5L6I	**Uba1 in complex with Ub-MLN4924 covalent adduct at 2.76 Å resolution**
5L6J	**Uba1 in complex with Ub-MLN7243 covalent adduct at 2.68 Å resolution**

The crystal structures discussed in Chapter 2 (Table 7), have not been submitted to the PDB yet.

7. LIST OF PUBLICATIONS

Peer reviewed publications

Mohit Misra, Maximilian Kuhn, Mark Löbel, Heeseon An, Alexander V. Statsyuk. Christoph Sotriffer and Hermann Schindelin. Dissecting the specificity of adenosyl sulfamate inhibitors targeting the ubiquitin-activating enzyme. Structure 25, 1-10, July 2017

Congress Contributions

Oral Presenations:

- Delivered a talk on **"Fragment based screening to target the ubiquitin activating enzyme for cancer therapy"** at the **German Crystallographic Society (Deutsche Gesellschaft für Kristallographie, DGK) 2017 in Karlsruhe, Germany (27th-30th March 2017).**

- Oral presentation at the **Eureka 2016**, **11th International PhD symposium at Rudolf Virchow Center, Würzburg, Germany (12th-13th October 2016).**

Poster Presentations:

- "Snapshots of the catalytic cycle of the ubiquitin activating enzyme" at the **"Ubiquitin and Ub-like modifiers" EMBO conference in Cavtat, Croatia (18th-22nd September 2015).**

- Presented poster at the **15th International Conference on the Crystallization of Biological Macromolecules (ICCBM 15) in Hamburg, Germany (17th-20th September 2014).**

8. LIST OF EXPRESSION CONSTRUCTS

Insert	Vector/Tag-s/Protease cleavage site	Source
Δ9 Uba1	pET 28a /N term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubiquitin	pET 30a/ tag free	*Saccharomyces cerevisiae*
Ubiquitin	pET 30a/ tag free	*Homo sapiens*
Δ40 UBA1	pET23b+/ C-term-His_6 tag	*Homo sapiens*
Δ48 UBA1	pET23b+/ C-term-His_6 tag	*Homo sapiens*
Full Length UBA1	pET21d/ N-term-His_{12} tag	*Homo sapiens*
Δ48 UBA1	pET21d/ N-term-His_{12} tag	*Homo sapiens*
Ubc1	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc2	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc3	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc4	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc5	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc6	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc7	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc8	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc9	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc10	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc11	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc12	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*
Ubc13	pETM11/ N-term-His_6 tag/ TEV cleavage site	*Saccharomyces cerevisiae*

MoeB	pET15b	*Escherichia coli*
MoaD	pTXB1/ C-terminal Intein tag/ self cleavable	*Escherichia coli*
Ubiquitin	pBADM11/ N-term-His_6 tag/ TEV cleavage site (SAA linker between TEV cleavage site and N-terminus of ubiquitin)	*Saccharomyces cerevisiae*

The mutants were generated using the above mentioned constructs and are listed in Table 5.

9. ACKNOWLEDGMENT

It is a pleasure to finally being able to pen down my gratitude for the people who have been invaluable support for me. There are so many people to name, whom I would like to thank for being a part of my journey and perhaps, I may miss a few names, but if you are reading this text, I hope you know that I am very grateful for your generous support.

The first name that comes to mind is of course my supervisor Prof. Hermann Schindelin. He has influenced me in so many ways that probably I can't even describe. I owe him BIG for his patience while explaining me the very fundamental concepts and for his belief and support to pursue my doctoral project. He listened to my stupidest ideas and encouraged me when I needed it. His exceptional grasp of the subject and enthusiasm for science is an inspiration for me.

I would like to thank Dr. Gabriele Blum-Oehler and the entire GSLS team for their continuous support from the very beginning. Especially Jenny, Bianca and Felizitas for being the nicest souls and very supportive. We call Gabi, "The Fairy" for she has the solution to all our problems and she is a very kind hearted person. I am really grateful for her backup in critical times.

I am grateful to Prof. Dr. Caroline Kisker for never letting me feel the lack of a doctor mother. Her constant support has been very uplifting for me. Like home, Hermann and Caroline run a family at work and it has been a wonderful experience being a part of it.

Dr. Sonja Lorenz has been a very friendly mentor and I feel lucky to have her in my doctoral committee. She kept me in touch with the recent findings of the ubiquitin field and has been very encouraging. The EMBO conference in 2015 was a memorable trip with her. I am also grateful to Prof. Dr. Alexander Buchberger for guiding me in my doctoral committee meetings and for making me understand the complexity of the ubiquitin system through his questions and comments in our ubiquitin journal club.

Monika did the very first experiments with me in the lab and I am very happy to still be able to work with her. She is a very dedicated technician and her organizational skills are out of the box. On top of that, she is a very kind and humble person and I am really thankful for her assistance and patience throughout.

Dr. Antje Schäfer introduced me with the lab and the project. Being part of the UBA1 sub-group of the lab, we shared fruitful scientific discussions and ideas and I am grateful for her scientific input, support and guidance.

I have been fortunate to have Dr. Vikram Babu Kasaragod as a friend and a confidant. He has backed me up during crucial phases of my PhD and has been vital for me. I found tremendous moral support from Dr. Florian Sauer in times when I desperately needed some. I greatly enjoyed his camaraderie in the lab and am grateful to him for timely suggestions and critique. Dr. Daniel Grabarczyk has been a friend and well

wisher. Scientific discussions with him over coffee every afternoon has been a cherishable experience. I present my special thanks to Dr. Bernhard Froehlich, who has always been quick in solving all kinds of technical problems I faced. He has been very kind and cheerful and has provided delicious coffee for good period of time. It has been great to have the “exuberant” Ingrid around who other than being cordial also helped me with scientific and personal matters. I am also thankful to Dr. Carolyn Delto for being very considerate and caring and for sharing her experiences whenever needed. Dr. Petra Hanzelmann has been a role model for her dedication for science and her constant presence in the lab especially in the weekends has been very comforting. I enjoyed great companionship with Theresa, GuBo (Gudrun & Bodo) and Sebastian both inside and outside the lab. I would like to thank Teresa Frank and Andrea Heinzmann who have been very obliging and cooperative in administrative matters.

Last but not the least I would like to show my gratitude towards my parents who provided unconditional love and support. Though suffering from the hardships of long distance parenthood, they never complained and have been very enthusiastic in nurturing my dreams. My caring elder sister, Dr. Manjari Misra has been an exemplary figure throughout my studies. I offer my gratitude to my dearest uncle late Kamla Kant Misra to whom I dedicate this thesis. He was a strong support and always poured fatherly love on me.

10. ABBREVIATION

°	Degree
°C	Degree celcius
Å	Ångström
3D	Three dimensional
A280	Absorbance at 280 nm
A_{600}	Optical density at 600 nm
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
Amp	Ampicillin
BESSY	Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung m. b. H.
Cam	Chloramphenicol
CCP4	Collaborative computational project number 4
CV	Column volume
Da	Dalton
DUB	Deubiquitinases
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetate
ERAD	Endoplasmic-reticulum associated protein degradation
ESRF	European synchrotron radiation facility
Fig.	Figure
FPLC	Fast protein liquid chromatography
GTP	Guanosine triphosphate
H_2O_2	hydrogen peroxide
HEPES	4-(2-hydroxyethyl)-1- piperazineethanesulfonic acid
IκB	Inhibitor of kappa B
IPTG	Isopropyl-β-thiogalactoside
IR	Infrared
Kan	Kanamycin
K_D	Dissociation constant
kDa	Kilo Dalton
LB	Lysogen broth

M	Molar
min	Minute
mL	Milliliter
mM	Millimolar
Moco	Molybdenum cofactor
MR	Molecular Replacement
MWCO	Molecular weight cut off
n. a.	Not available
n.d.	Not determined
NF-κB	Nuclear Factor-κB
nm	Nanometer
nM	Nanomolar
OD	Optical Density
PAGE	Polyacrylamide gel electophoresis
PDB	Protein Data Bank
PEG	Poly ethylene glycol
PH	Pleckstrin homology
pI	Isoelectric point
PISA	Protein Interfaces, Surfaces andAssemblies
PSM	Prestained Marker
PTM	Post translational modification
rmsd	Root mean square deviation
RT	Room temperature
Sc	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
SPR	Surface Plasmon resonance
SYPRo	SYPRO orange
TB	Terrific Broth
TCEP	Tris-(2-carboxyethyl)-phosphin
TEV	Tobacco Etch Virus
TLS	Translation, liberation, screw
TRIS	Trishydroxymethylaminomethane
tRNA	Transfer Ribonucleic acid
Ub	Ubiquitin
Ub•AMP	Ubiquitin adenylate
Ub(a)	adenylation site ubiquitin
Ub(t)	thioesterified ubiquitin

UBA1~Ub	thioesterified ubiquitin bound to UBA1
UPS	Ubiquitin Proteasome System
v/v	volume per volume
w/v	Weight per volume
WT	Wild-type
β-ME	β -mercaptoethanol
ε	Extinction coefficient
µL	Microliter
µM	Micromolar

11. AFFIDAVIT

I hereby confirm that my thesis entitled “Knowing then defeating: The ubiquitin activating enzyme, a promising target for cancer therapy” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied 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.

Würzburg,
Place, Date **Signature**

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation “Erst verstehen, dann beseigen: Das Ubiquitin-aktivierende Enzym, ein vielversprechender Kandidat in der Krebstherapie” eigenständig, d.h. insbesondere selbständig 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.

Würzburg,
Ort, Datum **Signature**

12. CURRICULUM VITAE

MOHIT MISRA

✉ mohitmisra21@gmail.com +491793613721
Permanant address: A-308, Shiva Gardens Apartment,
Vrindavan Yojna, Phase II, Raibareilly Road,
Lucknow, 226025, U.P., India
Current address: Höchbergerstr. 36, Würzburg 97082
Bayern, Germany

Academic Qualifications

Doctor of Philosophy **Expected 2017**
Biomedicine/ Structural Biology
University of Würzburg

Master of Science **2011**
Biotechnology
Jamia Millia Islamia, New Delhi
Marks: 75.3% (Distinction)

Postgraduate Diploma 'A' level **2010**
Bioinformatics
NIELIT, Kolkata

Bachelor of Science **2008**
Zoology, Chemistry, Botany
University of Lucknow
Marks: 73.6% (First division)

Research Experience

Structural and biochemical insights into the catalytic cycle of ubiquitin activating enzyme and targeting the enzyme for cancer therapy. **February 2013- present**
Doctoral thesis project.

Rudolf Virchow Center for Experimental Biomedicine
University of Würzburg
Supervisor: Prof. Dr. Hermann Schindelin

Methods: X-ray crystallography, protein biochemistry, protein-protein/ligand interactions, enzymology.

Array platform technology for multiplexed diagnosis. **Aug 2011- Dec 2012**
Indo-Finnish collaborative project.

Junior Research Fellow
International Center for Genetic Engineering and Biotechnology, New Delhi, India
Supervisor: Dr. Navin Khanna, Group leader: Recombinant Gene Products group

Methods: Protein expression and purification in *Pichia pastoris* and *E. coli*. ELISA, Protein refolding, Hybridoma

Cloning, heterologous expression and tag free purification of Rice MAP Kinases. **Jan 2011-June 2011**

Master Dissertation project.
National Institute of Plant Genome Research,
New Delhi, India
Supervisor: Dr. Alok K. Sinha, Staff Scientist VI

Methods: PCR, Cloning, Expression and purification of proteins from *E. coli*, Yeast-two hybrid.

***In Silico* protein-protein interaction prediction between MAPKs and MKKs in Rice.** **May 2010-Aug 2010**

Summer training project.

National Institute of Plant Genome Research,
New Delhi, India
Supervisor: Dr. Alok K. Sinha, Staff Scientist VI

Methods: Homology modeling, Protein-protein docking, Softwares like MODELLER, Accelrys Discovery Studio (Biovia), Bioinformatics tools.

Awards and Honors:

- Awarded **Career Development Fellowship** by Graduate School of Life Sciences, University of Würzburg.
- **First prize in Scientific writing contest in Eureka 2015**, 10th International PhD symposium organized at University of Würzburg, Germany.
- Awarded **PhD fellowship** by **Graduate School of Life Sciences**, University of Würzburg from February 2013 till December 2016.
- **25th rank all over India in National Eligibility Test (NET) LS conducted by CSIR** (Counsil of Scientific and Industrial Research), a premier National R & D organisation, **Department of Scientific and Industrial Research, Government of India.**
- **Qualified Graduate Aptitude Test in Engineering (GATE)** conducted by **Indian Institute of Technology (IIT), Madras (90 percentile).**
- Won **First prize in 6th National Children Science Congress 1998** for the project presentation on the topic "**Plant Diversity Conservation and Care"** and was certified as **'State Child Scientist'.**

Würzburg,

Place, Date **Signature**

13. REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, 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.

Aichem, A., and Groettrup, M. (2016). The ubiquitin-like modifier FAT10 in cancer development. Int J Biochem Cell Biol *79*, 451-461.

Albanell, J., and Adams, J. (2002). Bortezomib, a proteasome inhibitor, in cancer therapy: from concept to clinic. Drug Future *27*, 1079-1092.

An, H., and Statsyuk, A.V. (2013). Development of activity-based probes for ubiquitin and ubiquitin-like protein signaling pathways. J Am Chem Soc 135, 16948-16962.

An, H., and Statsyuk, A.V. (2015). An inhibitor of ubiquitin conjugation and aggresome formation. Chemical Science 6, 5235-5245.

An, H., and Statsyuk, A.V. (2016). Facile synthesis of covalent probes to capture enzymatic intermediates during E1 enzyme catalysis. Chem Commun *52*, 2477-2480.

Ashida, H., Kim, M., and Sasakawa, C. (2014). Exploitation of the host ubiquitin system by human bacterial pathogens. Nat Rev Microbiol *12*, 399-413.

Åqvist, J., and Marelius, J. (2001). The Linear Interaction Energy Method for Predicting Ligand Binding Free Energy. Comb Chem High T Scr *4*, 613-626.

Bacik, J.P., Walker, J.R., Ali, M., Schimmer, A.D., and Dhe-Paganon, S. (2010). Crystal structure of the human ubiquitin-activating enzyme 5 (UBA5) bound to ATP: mechanistic insights into a minimalistic E1 enzyme. J Biol Chem *285*, 20273-20280.

Bedford, L., Lowe, J., Dick, L.R., Mayer, R.J., and Brownell, J.E. (2011). Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. Nature Reviews Drug Discovery *10*, 29-46.

Berndsen, C.E., and Wolberger, C. (2014). New insights into ubiquitin E3 ligase mechanism. Nat Struct Mol Biol *21*, 301-307.

Bialas, J., Groettrup, M., and Aichem, A. (2015). Conjugation of the Ubiquitin Activating Enzyme UBE1 with the Ubiquitin-Like Modifier FAT10 Targets It for Proteasomal Degradation. Plos One *10*.

Bjorkoy, G., Lamark, T., and Johansen, T. (2006). p62/SQSTM1 - A missing link between protein aggregates and the autophagy machinery. Autophagy *2*, 138-139.

Bohnsack, R.N., and Haas, A.L. (2003). Conservation in the mechanism of nedd8 activation by the human AppBp1-Uba3 heterodimer. J Biol Chem *278*, 26823-26830.

Bohren, K.M., Nadkarni, V., Song, J.H., Gabbay, K.H., and Owerbach, D. (2004). A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. J Biol Chem *279*, 27233-27238.

Bossis, G., and Melchior, F. (2006). Regulation of SUMOylation by reversible oxidation of SUMO conjugating enzymes. Mol Cell *21*, 349-357.

Broemer, M., Tenev, T., Rigbolt, K.T.G., Hempel, S., Blagoev, B., Silke, J., Ditzel, M., and Meier, P. (2010). Systematic In Vivo RNAi Analysis Identifies IAPs as NEDD8-E3 Ligases. Mol Cell *40*, 810-822.

Brownell, J.E., Sintchak, M.D., Gavin, J.M., Liao, H., Bruzzese, F.J., Bump, N.J., Soucy, T.A., Milhollen, M.A., Yang, X., Burkhardt, A.L., et al. (2010). Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. Mol Cell 37, 102-111.

Buetow, L., and Huang, D.T. (2016). Structural insights into the catalysis and regulation of E3 ubiquitin ligases. Nat Rev Mol Cell Bio *17*, 626-642.

Burch, T.J., and Haas, A.L. (1994). Site-Directed Mutagenesis of Ubiquitin - Differential Roles for Arginine in the Interaction with Ubiquitin-Activating Enzyme. Biochemistry *33*, 7300-7308.

Burroughs, A.M., Iyer, L.M., and Aravind, L. (2009). Natural history of the E1-like superfamily: Implication for adenylation, sulfur transfer, and ubiquitin conjugation. Proteins-Structure Function and Bioinformatics *75*, 895-910.
Cajee, U.F., Hull, R., and Ntwasa, M. (2012). Modification by Ubiquitin-Like Proteins: Significance in Apoptosis and Autophagy Pathways. Int J Mol Sci *13*, 11804-11831.
Cappadocia, L., and Lima, C.D. (2017). Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. Chem Rev.
Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A. , Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker,R.C., Wang, J., Wolf, R.M., Wu, X., and P.A. Kollman (2014), AMBER 14, University of California, San Francisco.
Chan, T.F., Carvalho, J., Riles, L., and Zheng, X.F.S. (2000). A chemical genomics approach toward understanding the global functions of the target of rapamycin protein (TOR). Proc Natl Acad Sci U S A *97*, 13227-13232.
Chavarria, N.E., Hwang, S.M., Cao, S.Y., Fu, X., Holman, M., Elbanna, D., Rodriguez, S., Arrington, D., Englert, M., Uthandi, S.*, et al.* (2014). Archaeal Tuc1/Ncs6 Homolog Required for Wobble Uridine tRNA Thiolation Is Associated with Ubiquitin-Proteasome, Translation, and RNA Processing System Homologs. Plos One *9*.
Chen, C., Meng, Y., Wang, L., Wang, H.X., Tian, C., Pang, G.D., Li, H.H., and Du, J. (2014a). Ubiquitin-activating enzyme E1 inhibitor PYR41 attenuates angiotensin II-induced activation of dendritic cells via the IBa/NF-B and MKP1/ERK/STAT1 pathways. Immunology *142*, 307-319.
Chen, J., Sawyer, N., and Regan, L. (2013). Protein-protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area. Protein Sci *22*, 510-515.
Chen, J.J., Tsu, C.A., Gavin, J.M., Milhollen, M.A., Bruzzese, F.J., Mallender, W.D., Sintchak, M.D., Bump, N.J., Yang, X., Ma, J., et al. (2011). Mechanistic studies of substrate-assisted inhibition of ubiquitin-activating enzyme by adenosine sulfamate analogues. J Biol Chem 286, 40867-40877.
Chen, J.X., Yang, L., Chen, H.X., Yuan, T., Liu, M.G., and Chen, P. (2014b). Recombinant adenovirus encoding FAT10 small interfering RNA inhibits HCC growth in vitro and in vivo. Exp Mol Pathol *96*, 207-211.
Chiu, Y.H., Sun, Q., and Chen, Z.J.J. (2007). E1-L2 activates both ubiquitin and FAT10. Mol Cell *27*, 1014-1023.
Choe, K.N., and Moldovan, G.L. (2017). Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. Mol Cell *65*, 380-392.
Christiano, R., Nagaraj, N., Frohlich, F., and Walther, T.C. (2014). Global Proteome Turnover Analyses of the Yeasts S-cerevisiae and S-pombe. Cell Reports *9*, 1959-1965.
Ciechanover, A., Elias, S., Heller, H., and Hershko, A. (1982). Covalent Affinity Purification of Ubiquitin-Activating Enzyme. J Biol Chem *257*, 2537-2542.
Ciechanover, A., Heller, H., Katzetzion, R., and Hershko, A. (1981). Activation of the Heat-Stable Polypeptide of the Atp-Dependent Proteolytic System. P Natl Acad Sci-Biol *78*, 761-765.
Cohen, P., and Tcherpakov, M. (2010). Will the Ubiquitin System Furnish as Many Drug Targets as Protein Kinases? Cell *143*, 686-693.
Cohen-Kaplan, V., Livneh, I., Avni, N., Cohen-Rosenzweig, C., and Ciechanover, A. (2016). The ubiquitin-proteasome system and autophagy: Coordinated and independent activities. Int J Biochem Cell Biol *79*, 403-418.
Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50, 760-763.

Cook, J.C., and Chock, P.B. (1991). Immunocytochemical Localization of Ubiquitin-Activating Enzyme in the Cell-Nucleus. Biochem Biophys Res Commun *174*, 564-571.
Cook, J.C., and Chock, P.B. (1992). Isoforms of Mammalian Ubiquitin-Activating Enzyme. J Biol Chem *267*, 24315-24321.
Correale, S., de Paola, I., Morgillo, C. M., Federico, A., Zaccaro, L., Pallante, P., Galeone, A., Fusco, A., Pedone, E., Luque, F. J. and Catalanotti, B. (2014). Structural Model of the hUbA1-UbcH10 Quaternary Complex: In Silico and Experimental Analysis of the Protein-Protein Interactions between E1, E2 and Ubiquitin, PLoS ONE *9*, e112082.
Crawford, L.J., Walker, B., and Irvine, A.E. (2011). Proteasome inhibitors in cancer therapy. J Cell Commun Signal *5*, 101-110.
da Silva, S.R., Paiva, S.L., Bancerz, M., Geletu, M., Lewis, A.M., Chen, J.J., Cai, Y.F., Lukkarila, J.L., Li, H.L., and Gunning, P.T. (2016). A selective inhibitor of the UFM1-activating enzyme, UBA5. Bioorg Med Chem Lett *26*, 4542-4547.
Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N • log(N) method for Ewald sums in large systems. J Chem Phys *98*, 10089-10092.
da Silva, S.R., Paiva, S.L., Lukkarila, J.L., and Gunning, P.T. (2013). Exploring a New Frontier in Cancer Treatment: Targeting the Ubiquitin and Ubiquitin-like Activating Enzymes. J Med Chem *56*, 2165-2177.
de Beer, T.A., Berka, K., Thornton, J.M., and Laskowski, R.A. (2014). PDBsum additions. Nucleic Acids Res 42, D292-296.
Dassa, B., Yanai, I., and Pietrokovski, S. (2004). New type of polyubiquitin-like genes with intein-like autoprocessing domains. Trends Genet *20*, 538-542.
Dauden, M.I., Kosinski, J., Kolaj-Robin, O., Desfosses, A., Ori, A., Faux, C., Hoffmann, N.A., Onuma, O.F., Breunig, K.D., Beck, M.*, et al.* (2017). Architecture of the yeast Elongator complex. Embo Reports *18*, 264-279.
Deshaies, R.J. (2014). Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. BMC Biol *12*.
Deshaies, R.J., and Joazeiro, C.A.P. (2009). RING Domain E3 Ubiquitin Ligases. Annu Rev Biochem *78*, 399-434.
Dikic, I., Wakatsuki, S., and Walters, K.J. (2009). Ubiquitin-binding domains - from structures to functions. Nat Rev Mol Cell Bio *10*, 659-671.
Dlamini, N., Josifova, D.J., Paine, S.M., Wraige, E., Pitt, M., Murphy, A.J., King, A., Buk, S., Smith, F., Abbs, S.*, et al.* (2013). Clinical and neuropathological features of X-linked spinal muscular atrophy (SMAX2) associated with a novel mutation in the UBA1 gene. Neuromuscul Disord *23*, 391-398.
Dohmen, R.J. (2004). SUMO protein modification. Bba-Mol Cell Res *1695*, 113-131.
Doris, K.S., Rumsby, E.L., and Morgan, B.A. (2012). Oxidative Stress Responses Involve Oxidation of a Conserved Ubiquitin Pathway Enzyme. Mol Cell Biol *32*, 4472-4481.
Duda, D.M., Walden, H., Sfondouris, J., and Schulman, B.A. (2005). Structural analysis of Escherichia coli ThiF. J Mol Biol *349*, 774-786.
Durfee, L.A., Kelley, M.L., and Huibregtse, J.M. (2008). The basis for selective E1-E2 interactions in the ISG15 conjugation system. J Biol Chem *283*, 23895-23902.
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.
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.
Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W., and Jhoti, H. (2016). Twenty years on: the impact of fragments on drug discovery. Nature Reviews Drug Discovery *15*, 605-619.
Fan, T., Huang, Z.X., Chen, L., Wang, W., Zhang, B.Y., Xu, Y., Pan, S.Z., Mao, Z.F., Hu, H., and Geng, Q. (2016). Associations between autophagy, the ubiquitin-

proteasome system and endoplasmic reticulum stress in hypoxia-deoxygenation or ischemia-reperfusion. Eur J Pharmacol *791*, 157-167.
Fan, W., Cai, W., Parimoo, S., Schwarz, D.C., Lennon, G.G., and Weissman, S.M. (1996). Identification of seven new human MHC class I region genes around the HLA-F locus. Immunogenetics *44*, 97-103.
Farre, J.C., and Subramani, S. (2016). Mechanistic insights into selective autophagy pathways: lessons from yeast. Nat Rev Mol Cell Bio *17*, 537-552.
Finley, D., Bartel, B., and Varshavsky, A. (1989). The Tails of Ubiquitin Precursors Are Ribosomal-Proteins Whose Fusion to Ubiquitin Facilitates Ribosome Biogenesis. Nature *338*, 394-401.
Finley, D., Ciechanover, A., and Varshavsky, A. (1984). Thermolability of Ubiquitin-Activating Enzyme from the Mammalian-Cell Cycle Mutant Ts85. Cell *37*, 43-55.
Friedmann, J.S., Koop, B.F., Raymond, V., and Walter, M.A. (2001). Isolation of a ubiquitin-like (UBL5) gene from a screen identifying highly expressed and conserved iris genes. Genomics *71*, 252-255.
Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009). Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT.
Furukawa, K., Mizushima, N., Noda, T., and Ohsumi, Y. (2000). A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes. J Biol Chem *275*, 7462-7465.
Ghaboosi, N., and Deshaies, R.J. (2007). A conditional yeast E1 mutant blocks the ubiquitin-proteasome pathway and reveals a role for ubiquitin conjugates in targeting Rad23 to the proteasome. Mol Biol Cell *18*, 1953-1963.
Ghaemmaghami, S., Huh, W., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.
Giannakopoulos, N.V., Luo, J.K., Papov, V., Zou, W.G., Lenschow, D.J., Jacobs, B.S., Borden, E.C., Li, J., Virgin, H.W., and Zhang, D.E. (2005). Proteomic identification of proteins conjugated to ISG15 in mouse and human cells. Biochem Biophys Res Commun *336*, 496-506.
Glickman, M.H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev *82*, 373-428.
Goehring, A.S., Rivers, D.M., and Sprague, G.F. (2003a). Attachment of the ubiquitin-related protein Urm1p to the antioxidant protein Ahp1p. Eukaryot Cell *2*, 930-936.
Goehring, A.S., Rivers, D.M., and Sprague, G.F. (2003b). Urmylation: A ubiquitin-like pathway that functions during invasive growth and budding in yeast. Mol Biol Cell *14*, 4329-4341.
Goldknopf, I.L., and Busch, H. (1977). Isopeptide Linkage between Nonhistone and Histone-2a Polypeptides of Chromosomal Conjugate-Protein-A24. Proc Natl Acad Sci U S A *74*, 864-868.
Grenfell, S.J., Trauschazar, J.S., Handleygearhart, P.M., Ciechanover, A., and Schwartz, A.L. (1994). Nuclear-Localization of the Ubiquitin-Activating Enzyme, E1, Is Cell-Cycle-Dependent. Biochem J *300*, 701-708.

Gribble, F.M., Loussouarn, G., Tucker, S.J., Zhao, C., Nichols, C.G., and Ashcroft, F.M. (2000). A novel method for measurement of submembrane ATP concentration. J Biol Chem *275*, 30046-30049.
Groen, E.J., and Gillingwater, T.H. (2015). UBA1: At the Crossroads of Ubiquitin Homeostasis and Neurodegeneration. Trends in molecular medicine *21*, 622-632.
Haas, A.L., Bright, P.M., and Jackson, V.E. (1988). Functional Diversity among Putative E2 Isozymes in the Mechanism of Ubiquitin-Histone Ligation. J Biol Chem *263*, 13268-13275.
Haas, A.L., and Rose, I.A. (1982). The Mechanism of Ubiquitin Activating Enzyme - a Kinetic and Equilibrium-Analysis. J Biol Chem *257*, 329-337.
Haas, A.L., Rose, I.A., and Hershko, A. (1981). Purification of the Ubiquitin Activating Enzyme Required for Atp-Dependent Protein-Degradation. Fed Proc *40*, 1691-1691.
Haas, A.L., Warms, J.V., Hershko, A., and Rose, I.A. (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. J Biol Chem *257*, 2543-2548.
Haas, A.L., Warms, J.V.B., and Rose, I.A. (1983). Ubiquitin Adenylate - Structure and Role in Ubiquitin Activation. Biochemistry *22*, 4388-4394.
Haglund, K., and Dikic, I. (2005). Ubiquitylation and cell signaling. EMBO J *24*, 3353-3359.
Hanna, J., Meides, A., Zhang, D.P., and Finley, D. (2007). A ubiquitin stress response induces altered proteasome composition. Cell *129*, 747-759.
Hershko, A., and Ciechanover, A. (1982). Mechanisms of Intracellular Protein Breakdown. Annu Rev Biochem *51*, 335-364.
Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., and Rose, I.A. (1980). Proposed Role of Atp in Protein Breakdown - Conjugation of Proteins with Multiple Chains of the Polypeptide of Atp-Dependent Proteolysis. P Natl Acad Sci-Biol *77*, 1783-1786.
Hetz, C. (2012). The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat Rev Mol Cell Bio *13*, 89-102.
Hong, S.B., Kim, B.W., Lee, K.E., Kim, S.W., Jeon, H., Kim, J., and Song, H.K. (2011). Insights into noncanonical E1 enzyme activation from the structure of autophagic E1 Atg7 with Atg8. Nat Struct Mol Biol *18*, 1323-U1332.
Hoppe, T. (2005). Multiubiquitylation by E4 enzymes: 'one size' doesn't fit all. Trends Biochem Sci *30*, 183-187.
Hori, T., Osaka, F., Chiba, T., Miyamoto, C., Okabayashi, K., Shimbara, N., Kato, S., and Tanaka, K. (1999). Covalent modification of all members of human cullin family proteins by NEDD8. Oncogene *18*, 6829-6834.
Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins *65*, 712-725.
Huang, B., Lu, J., and Bystrom, A.S. (2008a). A genome-wide screen identifies genes required for formation of the wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine in Saccharomyces cerevisiae. Rna-a Publication of the Rna Society *14*, 2183-2194.
Huang, D.T., Hunt, H.W., Zhuang, M., Ohi, M.D., Holton, J.M., and Schulman, B.A. (2007). Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. Nature *445*, 394-398.
Huang, D.T., Miller, D.W., Mathew, R., Cassell, R., Holton, J.M., Roussel, M.F., and Schulman, B.A. (2004). A unique E1-E2 interaction required for optimal conjugation of the ubiquitin-like protein NEDD8. Nat Struct Mol Biol *11*, 927-935.
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. Mol Cell *17*, 341-350.
Huang, D.T., Zhuang, M., Ayrault, O., and Schulman, B.A. (2008b). Identification of conjugation specificity determinants unmasks vestigial preference for ubiquitin within the NEDD8 E2. Nat Struct Mol Biol *15*, 280-287.

Huschmann, F.U., Linnik, J., Sparta, K., Uhlein, M., Wang, X.J., Metz, A., Schiebel, J., Heine, A., Klebe, G., Weiss, M.S.*, et al.* (2016). Structures of endothiapepsin-fragment complexes from crystallographic fragment screening using a novel, diverse and affordable 96-compound fragment library. Acta Crystallogr F *72*, 346-355.
Ichimura, Y., Kirisako, T., Takao, T., Satomi, Y., Shimonishi, Y., Ishihara, N., Mizushima, N., Tanida, I., Kominami, E., Ohsumi, M.*, et al.* (2000). A ubiquitin-like system mediates protein lipidation. Nature *408*, 488-492.
Ikeda, F., and Dikic, I. (2008). Atypical ubiquitin chains: new molecular signals - 'Protein modifications: Beyond the usual suspects' review series. Embo Reports *9*, 536-542.
Irwin, J.J., and Shoichet, B.K. (2005). ZINC - A free database of commercially available compounds for virtual screening. J Chem Inf Model *45*, 177-182.
Isoe, T., Naito, M., Shirai, A., Hirai, R., and Tsuruo, T. (1992). Inhibition of Different Steps of the Ubiquitin System by Cisplatin and Aclarubicin. Biochim Biophys Acta *1117*, 131-135.
Jedrzejowska, M., Jakubowska-Pietkiewicz, E., and Kostera-Pruszczyk, A. (2015). X-linked spinal muscular atrophy (SMAX2) caused by de novo c.1731C>T substitution in the UBA1 gene. Neuromuscul Disord *25*, 661-666.
Jin, J., Li, X., Gygi, S.P., and Harper, J.W. (2007a). Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature *447*, 1135-1138.
Jin, J.P., Li, X., Gygi, S.P., and Harper, J.W. (2007b). Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature *447*, 1135-U1117.
Johnson, E.S. (2004a). Protein modification by SUMO. Annu Rev Biochem *73*, 355-382.
Johnson, E.S. (2004b). Protein modification by SUMO. Annu Rev Biochem *73*, 355-382.
Johnson, E.S., Schwienhorst, I., Dohmen, R.J., and Blobel, G. (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. EMBO J *16*, 5509-5519.
Johnston, J.A., Ward, C.L., and Kopito, R.R. (1998). Aggresomes: A cellular response to misfolded proteins. J Cell Biol *143*, 1883-1898.
Jorgensen, W.L., Chandrasekhar J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J Chem Phys *79*, 926-935.
Judes, A., Bruch, A., Klassen, R., Helm, M., and Schaffrath, R. (2016). Sulfur transfer and activation by ubiquitin-like modifier system Uba4*Urm1 link protein urmylation and tRNA thiolation in yeast. Microb Cell *3*, 554-564.
Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.
Kaiser, S.E., Mao, K., Taherbhoy, A.M., Yu, S.S., Olszewski, J.L., Duda, D.M., Kurinov, I., Deng, A., Fenn, T.D., Klionsky, D.J.*, et al.* (2012). Noncanonical E2 recruitment by the autophagy E1 revealed by Atg7-Atg3 and Atg7-Atg10 structures. Nat Struct Mol Biol *19*, 1242-+.
Kaiser, S.E., Riley, B.E., Shaler, T.A., Trevino, R.S., Becker, C.H., Schulman, H., and Kopito, R.R. (2011). Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. Nat Methods *8*, 691-U129.
Kapuria, V., Peterson, L.F., Showalter, H.D.H., Kirchhoff, P.D., Talpaz, M., and Donato, N.J. (2011). Protein cross-linking as a novel mechanism of action of a ubiquitin-activating enzyme inhibitor with anti-tumor activity. Biochem Pharmacol *82*, 341-349.
Karaduman, R., Chanarat, S., Pfander, B., and Jentsch, S. (2017). Error-Prone Splicing Controlled by the Ubiquitin Relative Hub1. Mol Cell *67*, 423-+.
Kaur, J., and Debnath, J. (2015). Autophagy at the crossroads of catabolism and anabolism. Nat Rev Mol Cell Bio *16*, 461-472.
Kawaguchi, Y., Kovacs, J.J., McLaurin, A., Vance, J.M., Ito, A., and Yao, T.P. (2003). The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. Cell *115*, 727-738.

Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol *22*, 159-180.
Kessler, J.D., Kahle, K.T., Sun, T.T., Meerbrey, K.L., Schlabach, M.R., Schmitt, E.M., Skinner, S.O., Xu, Q.K., Li, M.Z., Hartman, Z.C.*, et al.* (2012). A SUMOylation-Dependent Transcriptional Subprogram Is Required for Myc-Driven Tumorigenesis. Science *335*, 348-353.
Kitagaki, J., Yang, Y., Saavedra, J.E., Colburn, N.H., Keefer, L.K., and Perantoni, A.O. (2009). Nitric oxide prodrug JS-K inhibits ubiquitin E1 and kills tumor cells retaining wild-type p53. Oncogene *28*, 619-624.
Klassen, R., Grunewald, P., Thuring, K.L., Eichler, C., Helm, M., and Schaffrath, R. (2015). Loss of anticodon wobble uridine modifications affects tRNA(Lys) function and protein levels in Saccharomyces cerevisiae. PLoS One *10*, e0119261.
Komander, D., Clague, M.J., and Urbe, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat Rev Mol Cell Bio *10*, 550-563.
Komatsu, M., Chiba, T., Tatsumi, K., Iemura, S., Tanida, I., Okazaki, N., Ueno, T., Kominami, E., Natsume, T., and Tanaka, K. (2004). A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. EMBO J *23*, 1977-1986.
Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.
Kulathu, Y., Garcia, F.J., Mevissen, T.E.T., Busch, M., Arnaudo, N., Carroll, K.S., Barford, D., and Komander, D. (2013). Regulation of A20 and other OTU deubiquitinases by reversible oxidation. Nature Communications *4*.
Kulkarni, M., and Smith, H.E. (2008). E1 Ubiquitin-Activating Enzyme UBA-1 Plays Multiple Roles throughout C. elegans Development. PLoS Genet *4*.
Kumar, S., Yoshida, Y., and Noda, M. (1993). Cloning of a Cdna Which Encodes a Novel Ubiquitin-Like Protein. Biochem Biophys Res Commun *195*, 393-399.
Labute, P. (2009). Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. Proteins, *75*, 187-205.
Lake, M.W., Wuebbens, M.M., Rajagopalan, K.V., and Schindelin, H. (2001). Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. Nature *414*, 325-329.
Lamb, C.A., Yoshimori, T., and Tooze, S.A. (2013). The autophagosome: origins unknown, biogenesis complex. Nat Rev Mol Cell Bio *14*, 759-774.
Lanning, B.R., Whitby, L.R., Dix, M.M., Douhan, J., Gilbert, A.M., Hett, E.C., Johnson, T.O., Joslyn, C., Kath, J.C., Niessen, S.*, et al.* (2014). A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. Nat Chem Biol *10*, 760-767.
Lee, I., and Schindelin, H. (2008). Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. Cell *134*, 268-278.
Lee, J.G., Baek, K., Soetandyo, N., and Ye, Y.H. (2013). Reversible inactivation of deubiquitinases by reactive oxygen species in vitro and in cells. Nature Communications *4*.
Lee, T.V., Ding, T., Chen, Z., Rajendran, V., Scherr, H., Lackey, M., Bolduc, C., and Bergmann, A. (2008). The E1 ubiquitin-activating enzyme Uba1 in Drosophila controls apoptosis autonomously and tissue growth nonautonomously. Development *135*, 43-52.
Lehmann, C., Begley, T.P., and Ealick, S.E. (2006). Structure of the Escherichia coli ThiS-ThIF complex, a key component of the sulfur transfer system in thiamin biosynthesis. Biochemistry *45*, 11-19.
Leidecker, O., Matic, I., Mahata, B., Pion, E., and Xirodimas, D.P. (2012). The ubiquitin E1 enzyme Ube1 mediates NEDD8 activation under diverse stress conditions. Cell Cycle *11*, 1142-1150.
Leidel, S., Pedrioli, P.G.A., Bucher, T., Brost, R., Costanzo, M., Schmidt, A., Aebersold, R., Boone, C., Hofmann, K., and Peter, M. (2009). Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA. Nature *458*, 228-U229.

Leimkuhler, S., Buhning, M., and Beilschmidt, L. (2017). Shared Sulfur Mobilization Routes for tRNA Thiolation and Molybdenum Cofactor Biosynthesis in Prokaryotes and Eukaryotes. Biomolecules *7*.
Leimkuhler, S., Wuebbens, M.M., and Rajagopalan, K.V. (2001). Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. J Biol Chem *276*, 34695-34701.
Lenschow, D.J., Lai, C., Frias-Staheli, N., Giannakopoulos, N.V., Lutz, A., Wolff, T., Osiak, A., Levine, B., Schmidt, R.E., Garcia-Sastre, A.*, et al.* (2007). IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses. Proc Natl Acad Sci U S A *104*, 1371-1376.
Li, W., Tu, D.Q., Brunger, A.T., and Ye, Y.H. (2007). A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. Nature *446*, 333-337.
Liakopoulos, D., Doenges, G., Matuschewski, K., and Jentsch, S. (1998). A novel protein modification pathway related to the ubiquitin system. EMBO J *17*, 2208-2214.
Liu, B., Lois, L.M., and Reverter, D. (2017a). Structural analysis and evolution of specificity of the SUMO UFD E1-E2 interactions. Sci Rep *7*.
Liu, L., Dong, Z., Liang, J., Cao, C., Sun, J., Ding, Y., and Wu, D. (2014). As an independent prognostic factor, FAT10 promotes hepatitis B virus-related hepatocellular carcinoma progression via Akt/GSK3 beta pathway. Oncogene *33*, 909-920.
Liu, W.J., Ye, L., Huang, W.F., Guo, L.J., Xu, Z.G., Wu, H.L., Yang, C., and Liu, H.F. (2016). p62 links the autophagy pathway and the ubiqutin-proteasome system upon ubiquitinated protein degradation. Cell Mol Biol Lett *21*.
Liu, X.P., Zhao, B., Sun, L.M., Bhuripanyo, K., Wang, Y.Y., Bi, Y.T., Davuluri, R.V., Duong, D.M., Nanavati, D., Yin, J.*, et al.* (2017b). Orthogonal ubiquitin transfer identifies ubiquitination substrates under differential control by the two ubiquitin activating enzymes. Nature Communications *8*.
Lois, L.M., and Lima, C.D. (2005). Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. EMBO J *24*, 439-451.
Lorenz, S., Cantor, A.J., Rape, M., and Kuriyan, J. (2013). Macromolecular juggling by ubiquitylation enzymes. BMC Biol *11*.
Lu, X.Q., Olsen, S.K., Capili, A.D., Cisar, J.S., Lima, C.D., and Tan, D.S. (2010). Designed Semisynthetic Protein Inhibitors of Ub/Ubl E1 Activating Enzymes. J Am Chem Soc *132*, 1748-+.
Luders, J., Pyrowolakis, G., and Jentsch, S. (2003). The ubiquitin-like protein HUB1 forms SDS-resistant complexes with cellular proteins in the absence of ATP. EMBO Rep *4*, 1169-1174.
Lukasiak, S., Schiller, C., Oehlschlaeger, P., Schmidtke, G., Krause, P., Legler, D.F., Autschbach, F., Schirmacher, P., Breuhahn, K., and Groettrup, M. (2008). Proinflammatory cytokines cause FAT10 upregulation in cancers of liver and colon. Oncogene *27*, 6068-6074.
Lv, Z.Y., Rickman, K.A., Yuan, L.M., Williams, K., Selvam, S.P., Woosley, A.N., Howe, P.H., Ogretmen, B., Smogorzewska, A., and Olsen, S.K. (2017a). S-pombe Uba1-Ubc15 Structure Reveals a Novel Regulatory Mechanism of Ubiquitin E2 Activity. Mol Cell *65*, 699-+.
Lv, Z.Y., Yuan, L.M., Atkison, J.H., Aldana-Masangkay, G., Chen, Y., and Olsen, S.K. (2017b). Domain alternation and active site remodeling are conserved structural features of ubiquitin E1. J Biol Chem *292*, 12089-12099.
Lydeard, J.R., Schulman, B.A., and Harper, J.W. (2013). Building and remodelling Cullin-RING E3 ubiquitin ligases - 'Ubiquitylation: mechanism and functions' review series. Embo Reports *14*, 1050-1061.
Manasanch, E.E., and Orlowski, R.Z. (2017). Proteasome inhibitors in cancer therapy. Nature Reviews Clinical Oncology *14*, 417-433.

Marelja, Z., Stocklein, W., Nimtz, M., and Leimkuhler, S. (2008). A novel role for human Nfs1 in the cytoplasm: Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis. J Biol Chem *283*, 25178-25185.
Martensen, P.M., and Justesen, J. (2004). Small ISGs coming forward. J Interferon Cytokine Res *24*, 1-19.
Matsuzawa, M., Kakeya, H., Yamaguchi, J., Shoji, M., Onose, R., Osada, H., and Hayashi, Y. (2006). Enantio- and diastereoselective total synthesis of (+)-panepophenanthrin, a ubiquitin-activating enzyme inhibitor, and biological properties of its new derivatives. Chem-Asian J *1*, 845-851.
Matunis, M.J., Coutavas, E., and Blobel, G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. J Cell Biol *135*, 1457-1470.
Maupin-Furlow, J.A. (2014). Prokaryotic ubiquitin-like protein modification. Annu Rev Microbiol *68*, 155-175.
McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr *63*, 32-41.
Mcgrath, J.P., Jentsch, S., and Varshavsky, A. (1991). Uba1 - an Essential Yeast Gene Encoding Ubiquitin-Activating Enzyme. EMBO J *10*, 227-236.
McNally, T., Huang, Q., Janis, R.S., Liu, Z., Olejniczak, E.T., and Reilly, R.M. (2003). Structural analysis of UBL5, a novel ubiquitin-like modifier. Protein Sci *12*, 1562-1566.
Merlet, J., Burger, J., Gomes, J.E., and Pintard, L. (2009). Regulation of cullin-RING E3 ubiquitin-ligases by neddylation and dimerization. Cell Mol Life Sci *66*, 1924-1938.
Meulmeester, E., and Melchior, F. (2008). Cell biology: SUMO. Nature *452*, 709-711.
Mevissen, T.E.T., Hospenthal, M.K., Geurink, P.P., Elliott, P.R., Akutsu, M., Arnaudo, N., Ekkebus, R., Kulathu, Y., Wauer, T., El Oualid, F*., et al.* (2013). OTU Deubiquitinases Reveal Mechanisms of Linkage Specificity and Enable Ubiquitin Chain Restriction Analysis. Cell *154*, 169-184.
Mevissen, T.E.T., and Komander, D. (2017). Mechanisms of Deubiquitinase Specificity and Regulation. Annu Rev Biochem *86*, 159-192.
Milhollen, M.A., Thomas, M.P., Narayanan, U., Traore, T., Riceberg, J., Amidon, B.S., Bence, N.F., Bolen, J.B., Brownell, J., Dick, L.R., et al. (2012). Treatment- Emergent Mutations in NAE beta Confer Resistance to the NEDD8-Activating Enzyme Inhibitor MLN4924. Cancer Cell 21, 388-401.
Milletti, F.; Storchi, L.; Sforna, G.; Cruciani, G. New and original p*K*a prediction method using grid molecular interaction fields. *J. Chem. Inf. Model.* **2007**, *47*, 2172-2181.
Lu, X., Olsen, S.K., Capili, A.D., Cisar, J.S., Lima, C.D., and Tan, D.S. (2010). Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes. J Am Chem Soc *132*, 1748-1749.
Mishra, S.K., Ammon, T., Popowicz, G.M., Krajewski, M., Nagel, R.J., Ares, M., Jr., Holak, T.A., and Jentsch, S. (2011). Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature *474*, 173-178.
Misra, M., Kuhn, M., Lobel, M., An, H., Statsyuk, A.V., Sotriffer, C., and Schindelin, H. (2017). Dissecting the Specificity of Adenosyl Sulfamate Inhibitors Targeting the Ubiquitin-Activating Enzyme. Structure *25*, 1120-+.
Mitchell, M.J., Wilcox, S.A., Watson, J.M., Lerner, J.L., Woods, D.R., Scheffler, J., Hearn, J.P., Bishop, C.E., and Graves, J.A.M. (1998). The origin and loss of the ubiquitin activating enzyme gene on the mammalian Y chromosome. Hum Mol Genet *7*, 429-434.
Mizushima, N., Noda, T., and Ohsumi, Y. (1999). Apg16p is required for the function of the Apg12p-Apg5p conjugate in the yeast autophagy pathway. EMBO J *18*, 3888-3896.
Mizushima, N., Noda, T., Yoshimori, T., Tanaka, Y., Ishii, T., George, M.D., Klionsky, D.J., Ohsumi, M., and Ohsumi, Y. (1998a). A protein conjugation system essential for autophagy. Nature *395*, 395-398.

Mizushima, N., Sugita, H., Yoshimori, T., and Ohsumi, Y. (1998b). A new protein conjugation system in human - The counterpart of the yeast Apg12p conjugation system essential for autophagy. J Biol Chem *273*, 33889-33892.
Molecular Operating Environment (MOE), Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016.
Mueller, E.G. (2006). Trafficking in persulfides: delivering sulfur in biosynthetic pathways. Nat Chem Biol *2*, 185-194.
Mueller, U., Forster, R., Hellmig, M., Huschmann, F.U., Kastner, A., Malecki, P., Puhringer, S., Rower, M., Sparta, K., Steffien, M., et al. (2015). The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. Eur Phys J Plus 130.
Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.
Nagata, T., Nakamura, M., Kawauchi, H., and Tanigawa, Y. (1998). Conjugation of ubiquitin-like polypeptide to intracellular acceptor proteins. Bba-Mol Cell Res *1401*, 319-328.
Nakai, Y., Nakai, M., and Hayashi, H. (2008). Thio-modification of yeast cytosolic tRNA requires a ubiquitin-related system that resembles bacterial sulfur transfer systems. J Biol Chem *283*, 27469-27476.
Nakamura, M., and Tanigawa, Y. (2000). Protein tyrosine phosphorylation induced by ubiquitin-like polypeptide in murine T helper clone type 2. Biochem Biophys Res Commun *274*, 565-570.
Nakamura, M., and Tanigawa, Y. (2003). Characterization of ubiquitin-like polypeptide acceptor protein, a novel pro-apoptotic member of the Bcl2 family. Eur J Biochem *270*, 4052-4058.
Nakamura, M., Xavier, R.M., Tsunematsu, T., and Tanigawa, Y. (1995). Molecular-Cloning and Characterization of a Cdna-Encoding Monoclonal Nonspecific Suppressor Factor. Proc Natl Acad Sci U S A *92*, 3463-3467.
Nalepa, G., Rolfe, M., and Harper, J.W. (2006). Drug discovery in the ubiquitin-proteasome system. Nature Reviews Drug Discovery *5*, 596-613.
Nawrocki, S.T., Griffin, P., Kelly, K.R., and Carew, J.S. (2012). MLN4924: a novel first-in-class inhibitor of NEDD8-activating enzyme for cancer therapy. Expert Opin Investig Drugs 21, 1563-1573.
Noda, N.N., Satoo, K., Fujioka, Y., Kumeta, H., Ogura, K., Nakatogawa, H., Ohsumi, Y., and Inagaki, F. (2011). Structural basis of Atg8 activation by a homodimeric E1, Atg7. Mol Cell *44*, 462-475.
Noma, A., Sakaguchi, Y., and Suzuki, T. (2009). Mechanistic characterization of the sulfur-relay system for eukaryotic 2-thiouridine biogenesis at tRNA wobble positions. Nucleic Acids Res *37*, 1335-1352.
Okumura, A., Lu, G.S., Pitha-Rowe, I., and Pitha, P.M. (2006). Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15. Proc Natl Acad Sci U S A *103*, 1440-1445.
Okumura, A., Pitha, P.M., and Harty, R.N. (2008). ISG15 inhibit Ebola VP40VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. Proc Natl Acad Sci U S A *105*, 3974-3979.
Olsen, S.K., Capili, A.D., Lu, X., Tan, D.S., and Lima, C.D. (2010). Active site remodelling accompanies thioester bond formation in the SUMO E1. Nature *463*, 906-912.
Olsen, S.K., and Lima, C.D. (2013). Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. Mol Cell *49*, 884-896.
Oweis, W., Padala, P., Hassouna, F., Cohen-Kfir, E., Gibbs, D.R., Todd, E.A., Berndsen, C.E., and Wiener, R. (2016). Trans-Binding Mechanism of Ubiquitin-like Protein Activation Revealed by a UBA5-UFM1 Complex. Cell Rep *16*, 3113-3120.

Owerbach, D., McKay, E.M., Yeh, E.T., Gabbay, K.H., and Bohren, K.M. (2005). A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. Biochem Biophys Res Commun *337*, 517-520.
Padala, P., Oweis, W., Mashahreh, B., Soudah, N., Cohen-Kfir, E., Todd, E.A., Berndsen, C.E., and Wiener, R. (2017). Novel insights into the interaction of UBA5 with UFM1 via a UFM1-interacting sequence. Sci Rep *7*.
Pajares, M., Jimenez-Moreno, N., Dias, I.H.K., Debelec, B., Vucetic, M., Fladmark, K.E., Basaga, H., Ribaric, S., Milisav, I., and Cuadrado, A. (2015). Redox control of protein degradation. Redox Biology *6*, 409-420.
Park, S., Isaacson, R., Kim, H.T., Silver, P.A., and Wagner, G. (2005). Ufd1 exhibits the AAA-ATPase fold with two distinct ubiquitin interaction sites. Structure *13*, 995-1005.
Patterson, C. (2002). A new gun in town: the U box is a ubiquitin ligase domain. Sci STKE *2002*, pe4.
Pelzer, C., Kassner, I., Matentzoglu, K., Singh, R.K., Wollscheid, H.P., Scheffner, M., Schmidtke, G., and Groettrup, M. (2007). UBE1L2, a novel E1 enzyme specific for ubiquitin. J Biol Chem *282*, 23010-23014.
Petroski, M.D., and Deshaies, R.J. (2005). Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. Cell *123*, 1107-1120.
Phillips, J., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R., Kalé, L., and Schulten, K. (2005). Scalable Molecular Dynamics with NAMD. J Comput Chem *26*, 1781-1802.
Pickart, C., Wu, P.Y., Eddins, M., and Wolberger, C. (2003). A conserved catalytic residue in the ubiquitin conjugating (E2) enzyme family. FASEB J *17*, A972-A972.
Pugh, D.J., Ab, E., Faro, A., Lutya, P.T., Hoffmann, E., and Rees, D.J. (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. BMC Struct Biol *6*, 1.
Rajagopalan, K.V., and Johnson, J.L. (1992). The Pterin Molybdenum Cofactors. J Biol Chem *267*, 10199-10202.
Ramelot, T.A., Cort, J.R., Yee, A.A., Semesi, A., Edwards, A.M., Arrowsmith, C.H., and Kennedy, M.A. (2003). Solution structure of the yeast ubiquitin-like modifier protein Hub1. J Struct Funct Genomics *4*, 25-30.
Rani, N., Aichem, A., Schmidtke, G., Kreft, S.G., and Groettrup, M. (2012). FAT10 and NUB1L bind to the VWA domain of Rpn10 and Rpn1 to enable proteasome-mediated proteolysis. Nature Communications *3*.
Redman, K.L., and Rechsteiner, M. (1989). Identification of the Long Ubiquitin Extension as Ribosomal Protein-S27a. Nature *338*, 438-440.
Rehman, S.A.A., Kristariyanto, Y.A., Choi, S.Y., Nkosi, P.J., Weidlich, S., Labib, K., Hofmann, K., and Kulathu, Y. (2016). MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. Mol Cell *63*, 146-155.
Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res *42*, W320-324.
Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A.L. (1994). Inhibitors of the Proteasome Block the Degradation of Most Cell-Proteins and the Generation of Peptides Presented on Mhc Class-I Molecules. Cell *78*, 761-771.
Rodrigo-Brenni, M.C., and Morgan, D.O. (2007). Sequential E2s drive polyubiquitin chain assembly on APC targets. Cell *130*, 127-139.
Rodriguez-Gonzales (2009). Role of the aggresome pathway in cancer: targeting historic deacetylase 6-dependent protein degredation. (vol 8, pg 2557, 2008). Cancer Res *69*, 4092-4092.
Roe, D. R., and Cheatham III, T.E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput *9*, 3084-3095.

Rotin, D., and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. Nat Rev Mol Cell Bio *10*, 398-409.
Rudolph, M.J., Wuebbens, M.M., Rajagopalan, K.V., and Schindelin, H. (2001). Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. Nat Struct Biol *8*, 42-46.
Ruschak, A.M., Slassi, M., Kay, L.E., and Schimmer, A.D. (2011). Novel Proteasome Inhibitors to Overcome Bortezomib Resistance. J Natl Cancer Inst *103*, 1007-1017.
Saeki, Y. (2017). JB Special Review-Recent Topics in Ubiquitin-Proteasome System and Autophagy Ubiquitin recognition by the proteasome. J Biochem *161*, 113-124.
Sarantopoulos, J., Shapiro, G.I., Cohen, R.B., Clark, J.W., Kauh, J.S., Weiss, G.J., Cleary, J.M., Mahalingam, D., Pickard, M.D., Faessel, H.M., et al. (2016). Phase I Study of the Investigational NEDD8-Activating Enzyme Inhibitor Pevonedistat (TAK-924/MLN4924) in Patients with Advanced Solid Tumors. Clin Cancer Res 22, 847- 857.
Schafer, A., Kuhn, M., and Schindelin, H. (2014). Structure of the ubiquitin-activating enzyme loaded with two ubiquitin molecules. Acta Crystallogr D Biol Crystallogr *70*, 1311-1320.
Scheel, H., Tomiuk, S., and Hofmann, K. (2005). Bioinformatical evidence for a prokaryotic ubiquitin-like protein modification system. FEBS J *272*, 155-156.
Schlieker, C.D., Van der Veen, A.G., Damon, J.R., Spooner, E., and Ploegh, H.L. (2008). A functional proteomics approach links the ubiquitin-related modifier Urm1 to a tRNA modification pathway. Proc Natl Acad Sci U S A *105*, 18255-18260.
Schmidtke, G., Aichem, A., and Groettrup, M. (2014). FAT10ylation as a signal for proteasomal degradation. Biochim Biophys Acta *1843*, 97-102.
Schmidtke, G., Kalveram, B., Weber, E., Bochtler, P., Lukasiak, S., Hipp, M.S., and Groettrup, M. (2006). The UBA domains of NUB1L are required for binding but not for accelerated degradation of the ubiquitin-like modifier FAT10. J Biol Chem *281*, 20045-20054.
Schmitz, J., Chowdhury, M.M., Haenzelmann, P., Nimtz, M., Lee, E.Y., Schindelin, H., and Leimkuhler, S. (2008a). The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins. Biochemistry *47*, 6479-6489.
Schmitz, J., Chowdhury, M.M., Hanzelmann, P., Nimtz, M., Lee, E.Y., Schindelin, H., and Leimkuhler, S. (2008b). The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins. Biochemistry *47*, 6479-6489.
Schulman, B.A., and Harper, J.W. (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat Rev Mol Cell Biol *10*, 319-331.
Sekizawa, R., Ikeno, S., Nakamura, H., Naganawa, H., Matsui, S., Iinuma, H., and Takeuchi, T. (2002). Panepophenanthrin, from a mushroom strain, a novel inhibitor of the ubiquitin-activating enzyme. J Nat Prod *65*, 1491-1493.
Shah, J.J., Jakubowiak, A.J., O'Connor, O.A., Orlowski, R.Z., Harvey, R.D., Smith, M.R., Lebovic, D., Diefenbach, C., Kelly, K., Hua, Z*., et al.* (2016). Phase I Study of the Novel Investigational NEDD8-Activating Enzyme Inhibitor Pevonedistat (MLN4924) in Patients with Relapsed/Refractory Multiple Myeloma or Lymphoma. Clin Cancer Res *22*, 34-43.
Shang, F., Deng, G., Wu, C.N., Gong, X., Nowell, T.R., Smith, D., Laursen, R.A., and Taylor, A. (2000). The origin of the two isoforms of ubiquitin-activating enzyme (E1). FASEB J *14*, A1528-A1528.
Silva, G.M., Finley, D., and Vogel, C. (2015). K63 polyubiquitination is a new modulator of the oxidative stress response. Nat Struct Mol Biol *22*, 116-123.
Skaug, B., and Chen, Z.J. (2010). Emerging role of ISG15 in antiviral immunity. Cell *143*, 187-190.
Smit, J.J., and Sixma, T.K. (2014). RBR E3-ligases at work. Embo Reports *15*, 142-154.

Soucy, T.A., Smith, P.G., Milhollen, M.A., Berger, A.J., Gavin, J.M., Adhikari, S., Brownell, J.E., Burke, K.E., Cardin, D.P., Critchley, S.*, et al.* (2009). An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature *458*, 732-U767.
St Jean, D.J., Jr., and Fotsch, C. (2012). Mitigating heterocycle metabolism in drug discovery. J Med Chem 55, 6002-6020.
Stankovic-Valentin, N., Drzewicka, K., Konig, C., Schiebel, E., and Melchior, F. (2016). Redox regulation of SUMO enzymes is required for ATM activity and survival in oxidative stress. EMBO J *35*, 1312-1329.
Stephen, A.G., Trauschazar, J.S., Ciechanover, A., and Schwartz, A.L. (1995). Phosphorylation of the Ubiquitin-Activating Enzyme E1 Is Cell-Cycle Dependent. Mol Biol Cell *6*, 755-755.
Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Res *26*, 423-440.
Strong, M., Sawaya, M.R., Wang, S.S., Phillips, M., Cascio, D., and Eisenberg, D. (2006). Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proc Natl Acad Sci U S A *103*, 8060-8065.
Streich, F.C., Jr., and Haas, A.L. (2010). Activation of ubiquitin and ubiquitin-like proteins. Subcell Biochem 54, 1-16.
Su, H.L., and Li, S.S. (2002). Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. Gene *296*, 65-73.
Sugaya, K., Ishihara, Y., and Inoue, S. (2015). Nuclear localization of ubiquitin-activating enzyme Uba1 is characterized in its mammalian temperature-sensitive mutant. Genes Cells *20*, 659-666.
Sun, G.H., Liu, Y.D., Yu, G., Li, N., Sun, X., and Yang, J. (2014). Increased FAT10 expression is related to poor prognosis in pancreatic ductal adenocarcinoma. Tumor Biol *35*, 5167-5171.
Swatek, K.N., and Komander, D. (2016). Ubiquitin modifications. Cell Res *26*, 399-422.
Taherbhoy, A.M., Tait, S.W., Kaiser, S.E., Williams, A.H., Deng, A., Nourse, A., Hammel, M., Kurinov, I., Rock, C.O., Green, D.R.*, et al.* (2011). Atg8 transfer from Atg7 to Atg3: a distinctive E1-E2 architecture and mechanism in the autophagy pathway. Mol Cell *44*, 451-461.
Tan, J.M.M., Wong, E.S.P., Dawson, V.L., Dawson, T.M., and Lim, K.L. (2008). Lysine 63-linked polyubiquitin potentially partners with p62 to promote the clearance of protein inclusions by autophagy. Autophagy *4*, 251-253.
Tanida, I., Mizushima, N., Kiyooka, M., Ohsumi, M., Ueno, T., Ohsumi, Y., and Kominami, E. (1999). Apg7p/Cvt2p: A novel protein-activating enzyme essential for autophagy. Mol Biol Cell *10*, 1367-1379.
Tatsumi, K., Sou, Y.S., Tada, N., Nakamura, E., Iemura, S., Natsume, T., Kang, S.H., Chung, C.H., Kasahara, M., Kominami, E.*, et al.* (2010). A Novel Type of E3 Ligase for the Ufm1 Conjugation System. J Biol Chem *285*, 5417-5427.
Tatsumi, K., Yamamoto-Mukai, H., Shimizu, R., Waguri, S., Sou, Y.S., Sakamoto, A., Taya, C., Shitara, H., Hara, T., Chung, C.H.*, et al.* (2011). The Ufm1-activating enzyme Uba5 is indispensable for erythroid differentiation in mice. Nature Communications *2*.
Taylor, S.V., Kelleher, N.L., Kinsland, C., Chiu, H.J., Costello, C.K., Backstrom, A.D., McLafferty, F.W., and Begley, T.P. (1998). Thiamin biosynthesis in Escherichia coli - Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation. J Biol Chem *273*, 16555-16560.
Teicher, B.A., and Tomaszewski, J.E. (2015). Proteasome inhibitors Commentary. Biochem Pharmacol *96*, 1-9.
Tokgoz, Z., Bohnsack, R.N., and Haas, A.L. (2006). Pleiotropic effects of ATP center dot Mg2+ binding in the catalytic cycle of ubiquitin-activating enzyme. J Biol Chem *281*, 14729-14737.
Toth, J.I., Yang, L., Dahl, R., and Petroski, M.D. (2012). A gatekeeper residue for NEDD8-activating enzyme inhibition by MLN4924. Cell Rep 1, 309-316.

Traore, T., Huck, J.H., Shi, J.S., Sappal, D.S., Duffey, J.D., Yang, Y.Y., Kadakia, E.K., Chakravarty, A.C., Stringer, B.S., Ishii, Y.I., *et al.* (2014). Pre-clinical in vivo characterization of MLN7243, an investigational ubiquitin activating enzyme inhibitor, in solid tumor models. Eur J Cancer *50*, 85-85.
Tsui, V., and Case, D.A. (2001). Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. Biopolymers *56*, 275-291
Tsuji, H., Matsudo, Y., Tsuji, S., Hanaoka, F., Hyodo, M., and Hori, T.A. (1990). Isolation of Temperature-Sensitive Cho-K1 Cell Mutants Exhibiting Chromosomal Instability and Reduced DNA-Synthesis at Nonpermissive Temperature. Somat Cell Mol Genet *16*, 461-476.
Ungermannova, D., Parker, S.J., Nasveschuk, C.G., Chapnick, D.A., Phillips, A.J., Kuchta, R.D., and Liu, X.D. (2012a). Identification and Mechanistic Studies of a Novel Ubiquitin E1 Inhibitor. J Biomol Screen *17*, 421-434.
Ungermannova, D., Parker, S.J., Nasveschuk, C.G., Wang, W., Quade, B., Zhan, G., Kuchta, R.D., Phillips, A.J., and Liu, X.D. (2012b). Largazole and Its Derivatives Selectively Inhibit Ubiquitin Activating Enzyme (E1). Plos One *7*.
van der Veen, A.G., and Ploegh, H.L. (2012). Ubiquitin-Like Proteins. Annual Review of Biochemistry, Vol 81 *81*, 323-357.
van Wijk, S.J.L., and Timmers, H.T.M. (2010). The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. FASEB J *24*, 981-993.
Varshavsky, A. (1997). The ubiquitin system. Trends Biochem Sci *22*, 383-387.
Vertegaal, A.C. (2010). SUMO chains: polymeric signals. Biochem Soc Trans *38*, 46-49.
Vijaykumar, S., Bugg, C.E., and Cook, W.J. (1987). Structure of Ubiquitin Refined at 1.8 a Resolution. J Mol Biol *194*, 531-544.
Walden, H., Podgorski, M.S., Huang, D.T., Miller, D.W., Howard, R.J., Minor, D.L., Holton, J.M., and Schulman, B.A. (2003a). The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. Mol Cell *12*, 1427-1437.
Walden, H., Podgorski, M.S., Huang, D.T., Miller, D.W., Howard, R.J., Minor, D.L., Jr., Holton, J.M., and Schulman, B.A. (2003b). The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. Mol Cell *12*, 1427-1437.
Walden, H., Podgorski, M.S., and Schulman, B.A. (2003c). Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8. Nature *422*, 330-334.
Walters, K.J., Lech, P.J., Goh, A.M., Wang, Q.H., and Howley, P.M. (2003). DNA-repair protein hHR23a alters its protein structure upon binding proteasomal subunit S5a. Proc Natl Acad Sci U S A *100*, 12694-12699.
Wang, C.Y., Xi, J., Begley, T.P., and Nicholson, L.K. (2001). Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. Nat Struct Biol *8*, 47-51.
Wang, J., and Chen, Y. (2010). Role of the Zn(2+) motif of E1 in SUMO adenylation. J Biol Chem *285*, 23732-23738.
Wang, J., Taherbhoy, A.M., Hunt, H.W., Seyedin, S.N., Miller, D.W., Miller, D.J., Huang, D.T., and Schulman, B.A. (2010a). Crystal structure of UBA2(ufd)-Ubc9: insights into E1-E2 interactions in Sumo pathways. PLoS One *5*, e15805.
Wang, J.H., Cai, S., and Chen, Y.A. (2010b). Mechanism of E1-E2 Interaction for the Inhibition of Ubl Adenylation. J Biol Chem *285*, 33457-33462.
Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model *25*, 247–260.
Wang J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. (2004). Development and testing of a general AMBER force field. J. Comput. Chem. *25*, 1157-1174.

Wauer, T., Swatek, K.N., Wagstaff, J.L., Gladkova, C., Pruneda, J.N., Michel, M.A., Gersch, M., Johnson, C.M., Freund, S.M.V., and Komander, D. (2015). Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. EMBO J *34*, 307-325.
Wee, K.E., Lai, Z.H., Auger, K.R., Ma, J.H., Horiuchi, K.Y., Dowling, R.L., Dougherty, C.S., Corman, J.I., Wynn, R., and Copeland, R.A. (2000). Steady-state kinetic analysis of human ubiquitin-activating enzyme (E1) using a fluorescently labeled ubiquitin substrate. J Protein Chem *19*, 489-498.
Welchman, R.L., Gordon, C., and Mayer, R.J. (2005). Ubiquitin and ubiquitin-like proteins as multifunctional signals. Nat Rev Mol Cell Bio *6*, 599-609.
Wenzel, D.M., Stoll, K.E., and Klevit, R.E. (2011). E2s: structurally economical and functionally replete (vol 433, pg 31, 2010). Biochem J *433*, 535-535.
Whitby, F.G., Xia, G., Pickart, C.M., and Hill, C.P. (1998). Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. J Biol Chem *273*, 34983-34991.
Wilkinson, K.D. (2005). The discovery of ubiquitin-dependent proteolysis. Proc Natl Acad Sci U S A *102*, 15280-15282.
Wilkinson, K.D., Smith, S.E., Oconnor, L., Sternberg, E., Taggart, J.J., Berges, D.A., and Butt, T. (1990). A Specific Inhibitor of the Ubiquitin Activating Enzyme - Synthesis and Characterization of Adenosyl-Phospho-Ubiquitinol, a Nonhydrolyzable Ubiquitin Adenylate Analog. Biochemistry *29*, 7373-7380.
Wilson, V.G., and Heaton, P.R. (2008). Ubiquitin proteolytic system: focus on SUMO. Expert Rev Proteomics *5*, 121-135.
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.
Wishart, T.M., Mutsaers, C.A., Riessland, M., Reimer, M.M., Hunter, G., Hannam, M.L., Eaton, S.L., Fuller, H.R., Roche, S.L., Somers, E.*, et al.* (2014). Dysregulation of ubiquitin homeostasis and beta-catenin signaling promote spinal muscular atrophy. J Clin Invest *124*, 1821-1834.
Wolberger, C. (2014). Mechanisms for regulating deubiquitinating enzymes. Protein Sci *23*, 344-353.
Xie, S.T. (2014). Characterization, crystallization and preliminary X-ray crystallographic analysis of the Uba5 fragment necessary for high-efficiency activation of Ufm1. Acta Crystallogr F *70*, 765-768.
Xirodimas, D.P., Saville, M.K., Bourdon, J.C., Hay, R.T., and Lane, D.P. (2004). Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. Cell *118*, 83-97.
Xu, G.W., Ali, M., Wood, T.E., Wong, D., Maclean, N., Wang, X., Gronda, M., Skrtic, M., Li, X., Hurren, R.*, et al.* (2010). The ubiquitin-activating enzyme E1 as a therapeutic target for the treatment of leukemia and multiple myeloma. Blood *115*, 2251-2259.
Xu, J.J., Zhang, J.H., Wang, L., Zhou, J., Huang, H.D., Wu, J.H., Zhong, Y., and Shi, Y.Y. (2006). Solution structure of Urm1 and its implications for the origin of protein modifiers. Proc Natl Acad Sci U S A *103*, 11625-11630.
Yang, X.F., Brownell, J.E., Xu, Q., Zhu, F.Y., Ma, J.Y., Loke, H.K., Rollins, N., Soucy, T.A., Minissale, J.J., Thomas, M.P.*, et al.* (2013). Absolute Quantification of E1, Ubiquitin-Like Proteins and Nedd8-MLN4924 Adduct by Mass Spectrometry. Cell Biochem Biophys *67*, 139-147.
Yang, Y.L., Kitagaki, J., Dai, R.M., Tsai, Y.C., Lorick, K.L., Ludwig, R.L., Pierre, S.A., Jensen, J.P., Davydov, I.V., Oberoi, P.*, et al.* (2007). Inhibitors of ubiquitin-activating enzyme (E1), a new class of potential cancer therapeutics. Cancer Res *67*, 9472-9481.
Yuan, J., Tu, Y.Y., Mao, X.G., He, S.M., Wang, L., Fu, G.Q., Zong, J.H., and Zhang, Y.S. (2012). Increased Expression of FAT10 is Correlated with Progression and Prognosis of Human Glioma. Pathology & Oncology Research *18*, 833-839.
Yuan, W.M., and Krug, R.M. (2001). Influenza B virus NS1 protein inhibits conjugation of the interferon (IFN)-induced ubiquitin-like ISG15 protein. EMBO J *20*, 362-371.

Zacksenhaus, E., and Sheinin, R. (1990). Molecular-Cloning, Primary Structure and Expression of the Human-X Linked A1s9 Gene Cdna Which Complements the Ts A1s9 Mouse L-Cell Defect in DNA-Replication. EMBO J *9*, 2923-2929.
Zhan, J., He, J., Zhou, Y., Wu, M., Liu, Y., Shang, F., and Zhang, X. (2016). Crosstalk Between the Autophagy-Lysosome Pathway and the Ubiquitin-Proteasome Pathway in Retinal Pigment Epithelial Cells. Curr Mol Med *16*, 487-495.
Zhao, B., Choi, C.H.J., Bhuripanyo, K., Villhauer, E.B., Zhang, K.Y., Schindelin, H., and Yin, J. (2012). Inhibiting the Protein Ubiquitination Cascade by Ubiquitin-Mimicking Short Peptides. Org Lett *14*, 5760-5763.
Zheng, N. (2003). A closer look of the HECTic ubiquitin ligases. Structure *11*, 5-6.