Stony Brook University

Insights into ubiquitin activation and transfer to E2 from the structure of the Uba1-ubiquitin complex

A Dissertation Presented

by

Imsang Lee

to

The Graduate School

in Partial fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

August 2007

Stony Brook University

The Graduate School

Imsang Lee

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree,
hereby recommend acceptance of this dissertation.

Hermann Schindelin, Ph.D.
Advisor
Professor, Rudolf-Virchow-Center, University of Würzberg, Germany
Formerly Associate Professor, Department of Biochemistry and Cell Biology,
Stony Brook University

Robert S. Haltiwanger, Ph.D.
Co-advisor
Professor, Department of Biochemistry and Cell Biology,
Stony Brook University

Nicolas Nassar, Ph.D.
Chairperson of Defense
Assistant Professor, Department of Physiology and Biophysics,
Stony Brook University

Nisson Schechter, Ph.D.
Outside Examiner
Professor, Department of Psychiatry and Behavioral Sciences,
Stony Brook Univeristy

This dissertation is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Insights into ubiquitin activation and transfer to E2
from the structure of the Uba1-ubiquitin complex

By

Imsang Lee

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2007

Covalent attachment of a small, highly conserved protein called ubiquitin is a predominant mechanism for regulating protein function in eukaryotes and its defective regulation is manifest in diseases that range from developmental abnormalities and autoimmunity to neurodegenerative diseases and cancer. Generally, ubiquitin and ubiquitin-like proteins (Ubls) are conjugated via their C termini to their targets by parallel, but specific, cascades involving three classes of enzymes known as E1, E2, and E3. E1 activating enzymes play key roles in the transfer cascades: each E1 activates its cognate Ubl by first catalyzing a Ubl C-terminal adenylation, followed by formation of an E1~Ubl thioester intermediate, and ultimately generating a thioester-linked E2~Ubl product.

The 2.7 Å resolution crystal structure of the complex between yeast Uba1, a 114 kDa monomeric E1 and ubiquitin shows modular nature of E1 enzymes with activities specified by individual domains. These domains pack together creating a large groove in the middle and Uba1 selectively recruits ubiquitin into the groove through a bipartite recognition mechanism, involving the acidic cleft that recognizes the positively charged ubiquitin C-terminal sequence through electrostatic interactions and specific contacts with side chains of the ubiquitin C-terminus, and the hydrophobic surface on the adenylation domain that interacts with the canonical hydrophobic patch of ubiquitin defined by residues Leu8, Ile44, and Val70. Marked conformational changes in the C-terminal ubiquitin-fold domain (UFD), including movement of the linker connecting the domain to the rest of the enzyme, suggest a conformation-dependent mechanism for the activation and transthioesterification functions of Uba1.

Although the overall domain arrangement, adenylation active site location, and the position of the catalytic cysteine are similar in all three E1 enzymes (ubiquitin-, NEDD8-, and SUMO-E1), the detailed architecture and positioning of the individual domains are distinctive in each E1. As such, it appears that activation mechanisms, including Ubl interactions, conformational changes, and E2 recruitment, may be specific to each conjugation pathway, suggesting that each E1 family enzyme has developed a unique solution for accomplishing their ultimate goal: activate and transfer the correct Ubl to its cognate E2.

Table of Contents

List of Abbreviations

Å: Angstrom, 10^{-10} m

Amp: Ampicillin

AMP: Adenosine monophosphate

ATP: Adenosine triphosphate

Aos1: Activation of sentrin/SUMO protein 1

Apg12: Autophagy-related protein 12

APPBP1: Amyloid beta precursor protein-binding protein 1

AXR1: Auxin resistance protein 1

Cα: Alpha carbon

CUE: Coupling of ubiquitin conjugation to endoplasmic reticulum degradation

dNTP: deoxynucleotide triphophate

DTT: Dithiothreitol

DUB: Deubiquinating enzyme

E1: Ubiquitin-like protein activating enzyme

E2: Ubiquitin-like protein conjugating enzyme

E3: Ubiquitin-like protein ligase

ECR1: E1 C-terminal related protein 1

EDTA: Ethylenediamine-tetraacetic acid

ER: Endoplasmic reticulum

GTP: Guanosine triphosphate

HECT: homologous to E6AP C terminus

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

His-tag: Histidine affinity tag

IPTG: Isopropyl β-D-thiogalactopyranoside

ISG15: Interferon stimulated gene, 15 kDa

Kan: Kanamycin

kDa: Kilodalton

Km: Michaelis constant

LB: Luria Bertani

MME: Monomethyl ether

Moco: Molybdenum cofactor

MWCO: Molecular weight cut-off

NCS: Non-crystallographic symmetry

NEDD8: Neural precursor cell expressed developmentally down-regulated protein 8

Ni-NTA: Nickel-nitrilotriacetic acid

NMR: Nuclear magnetic resonance

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

PEG: Polyethylene glycol

PPi: Inorganic pyrophosphate

RING: Really interesting new gene

Rmsd: Root mean square deviations

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SUMO: Small ubiquitin-related modifier

TRIS: Tris (hydroxymethyl) aminomethane

Uba2: SUMO-1 activating enzyme subunit 2

UBA3: Ubiquitin-activating enzyme E1C isoform 3

UBD: Ubiquitin-binding domains

Ubl: Ubiquitin-like proteins

UEV: Ubiquitin E2 variant

UIM: Ubiquitin-interacting motif

Ula1: Ubiquitin-like activation protein 1

ULP: Ubl-specific protease

WT: wild-type

List of Figures

List of Tables

Acknowledgements

I am grateful to all the members of the Schindelin and Kisker group. I would like to acknowledge the staff of NSLS, especially at beamlines X25, X26C, and X29. I thank my committee members, Dr. Nicolas Nassar, Dr. Robert Haltiwanger, and Dr. Nisson Schechter for their suggestions and guidance. I would like to express sincere gratitude for the help that was always provided by my mentor, Dr. Michael Lake and my extraordinary advisors, Dr. Hermann Schindelin and Dr. Caroline Kisker from whom I have learned a great deal. I would like to especially thank my friend Kyoung Eun Lee, who has been an inspiration during my graduate studies. Finally, I dedicate this thesis to my brother Joonsang, my sister Yewon, and my parents Hyunsook and Seungjo. This work would not have been possible without their unconditional love and support.

MAIN INTRODUCTION

A. Post-translational Modification by Ubiquitin-Like Proteins

Covalent posttranslational modifications can greatly expand the diversity and functional extent of an organism's proteome. Examples of these modifications include low molecular weight compounds such as phosphate, methyl, acetyl, or glycosyl groups. In addition, entire proteins can also be attached covalently to protein substrates. The classic example of a protein that covalently modifies other proteins is ubiquitin [1], a 76-residue polypeptide that is highly conserved among eukaryotes, but is absent from eubacteria and archaea.

Ubiquitin is typically attached to protein substrates via an isopeptide linkage between its C-terminus and the ε-amino group of a lysine residue in the target protein [2]. However, ubiquitin has also been found isopeptide-bonded to protein N-termini [3] and thioester-bonded to cysteine residues in certain targets [4]. Proteins can be modified by either a single ubiquitin (monoubiquitination) or by polyubiquitin chains, in which multiple ubiquitin molecules are linked to each other such that the C-terminus of one ubiquitin is covalently linked the amino group of a lysine side chain in the previous ubiquitin. Polyubiquitin chains linked via each of ubiquitin's seven lysines have been found *in vivo* and a single target has been found to be simultaneously modified by a mixture of linkages [5, 6].

Among the many functions of ubiquitin, the best understood is the targeting of protein substrates for degradation by the 26S proteasome. For this purpose, ubiquitin is attached to the substrate in the form of polyubiquitin chain that is then recognized by

specific receptors within the proteasome, or by adaptor proteins that can subsequently bind to the proteasome [7, 8]. However, it has recently become clear that ubiquitin is involved in a variety of other vital processes at different loci within the cell, ranging from the nucleus to the plasma membrane. These include cell cycle progression, organelle biogenesis, apoptosis, regulated cell proliferation, cellular differentiation, quality control in the endoplasmic reticulum (ER), protein transport, inflammation, antigen processing, DNA repair, and stress responses [9-12]. Thus, the cell uses ubiquitin to modify many other proteins to modulate their functions.

Since the discovery of ubiquitin about 30 years ago, an entire family of small proteins related to ubiquitin (called ubiquitin-like proteins, or Ubls) has been identified and new members are still being added (Figure 1.1 and Table 1.1). Although not necessarily displaying high sequence similarity, the Ubls all possess essentially the same three-dimensional structure, the ubiquitin or β-grasp fold [13], which consists of two parts: (1) A globular domain comprised of a five-stranded mixed β-sheet and an α-helix, and (2) a flexible C-terminal tail terminating in a glycine-glycine dipeptide (Figure 1.2). The function of these modifiers derives from the ability of their C-terminal glycine to participate in a highly ordered sequence of covalent interactions that usually culminates in their ligation to the target. Ubiquitin's closest relative, NEDD8 (Rub1 in *Saccharomyces cerevisiae*), activates cullin-RING (Really Interesting New Gene) ubiquitin ligases involved in the cell cycle, signaling, and embryogenesis [14]. Autophagy is regulated by Atg12 and Atg8, the latter of which modifies a lipid to modulate membrane dynamics [15]. SUMO family members (SUMO-1, -2, -3 in higher eukaryotes; Smt3 in *S. cerevisiae*) regulate transcription, DNA repair, nuclear transport,

chromosomal function, and signal transduction [16]. Like ubiquitin, SUMO-2, SUMO-3, and Smt3 have been found in polymeric chains, but their chain function is largely unknown, and it is not clear whether other Ubls function via chains [16, 17]. Other, less characterized Ubls include ISG15, FAT10, Ufm1 and Urm1 [17, 18].

Structures of ubiquitin and SUMO bound to their cognate Ubl-binding domains reveal how these proteins provide their targets with new molecular interaction surfaces [19-22], and structures of SUMO-modified targets demonstrate how one Ubl can mediate either positive or negative effects on intermolecular interactions, depending on the context. SUMO modification of RanGAP1 promotes its interaction with the nuclear pore component Nup358/RanBP2 [23], whereas interactions between SUMO and its conjugated thymine DNA glycosylase induces a conformational change antagonistic to DNA binding [24].

Ubls are generally ligated to their targets by related but distinct enzymatic cascades involving the sequential action of an E1 (Ubl-activating enzyme), an E2 (Ubl-conjugating enzyme), and an E3 (Ubl-protein ligase) [18] (Figure 1.3). These enzymes function in a regulated, hierarchical cascade to select a specific Ubl, identify its target (including the exact site of modification), and catalyze Ubl conjugation. Generally, each Ubl has its own set of enzymes, including a unique E1 and one or multiple related E2s and E3s. Structures reveal that many of these enzymes are themselves modular, with conserved domains carrying out related functions in different pathways. The modular nature of these cascades has provided many opportunities for the regulation of pathway-specific features throughout evolution.

Finally, there is a mechanism of considerable physiological and pathological importance, in which Ubls can be removed by deubiquitinating enzymes (DUBs). DUBs have turned ubiquitin and many of the Ubls into dynamic modifiers whose attachments are tightly regulated both spatially and temporally [25].

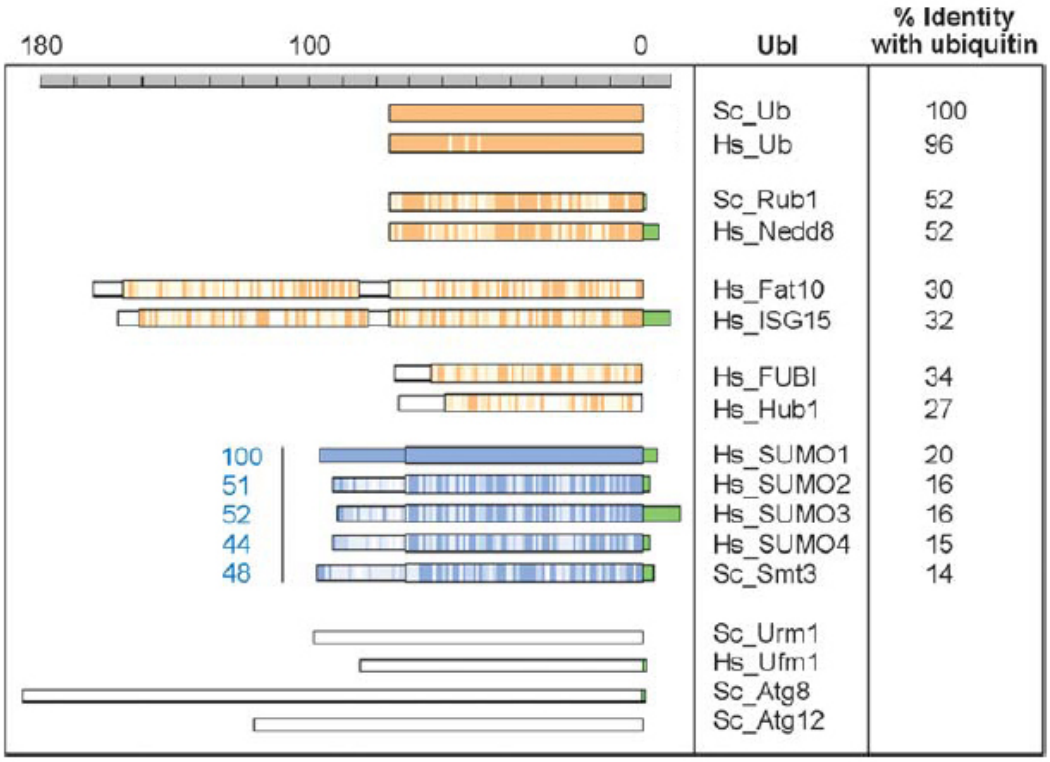


Figure 1.1 The eukaryotic Ubl protein family. The known Ubls from the yeast *S. cerevisiae* (Sc) and *Homo sapiens* (Hs) are depicted under the scale bar on the left (amino acid number, with mature C termini positioned at zero). The unprocessed C-terminal extensions of the precursors are depicted in green. Colored bars (more closely related to ubiquitin, *orange*; SUMO-related, *blue*) represent levels of amino acid conservation: dark bars, identical amino acids between yeast and human; light bars, conservative substitutions; white bars, non-conservative substitutions. The percent identity of each Ubl to yeast ubiquitin (Sc_Ub) is shown at the right. Percent identity between individual human SUMO isoforms and yeast Smt3 is reported separately with blue numbers on the left. Significant amino acid conservation between ubiquitin and the Urm1, Ufm1, Atg8 and Atg12 cannot be detected on the primary structure level. This figure has been reproduced from Kerscher et al [26].

Table 1.1 Ubiquitin-like proteins and their substrates

Modifier protein	Substrates
Ubiquitin	
Lys 48 chain	Cytoplasmic, nuclear and ER proteasome substrates
Lys 63 chain	Ribosomal protein L28, TRAF6, PCNA and endocytic cargo
Mono-ubiquitin	Histones, endocytic and endosomal factors and cargo
NEDD8/Rub1	Cullin subunits of SCF Ub ligases (E3s)
SUMO	Many diverse substrates, including PML, Ran-GAP1, p53, PCNA, IκBα
ISG15	Several, including Serpin2a
Apg12	Autophagy protein Apg5

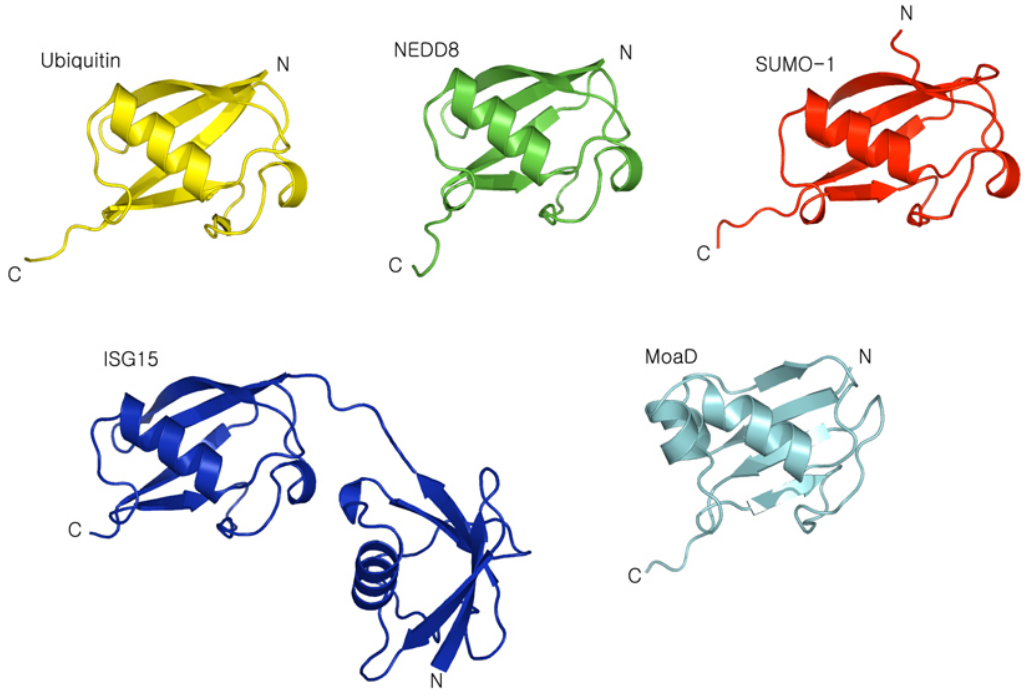


Figure 1.2 Ubl (Ubiquitin-like protein) structures. Structures of ubiquitin [27], NEDD8 [28], SUMO-1 [29], ISG15, which has two ubiquitin-like domains [30], and MoaD [31] are shown. All of them are structurally homologous to ubiquitin and are shown in the same orientation with the exception of the N-terminal Ubl domain of ISG15. The N- and C-termini of each protein are labeled.

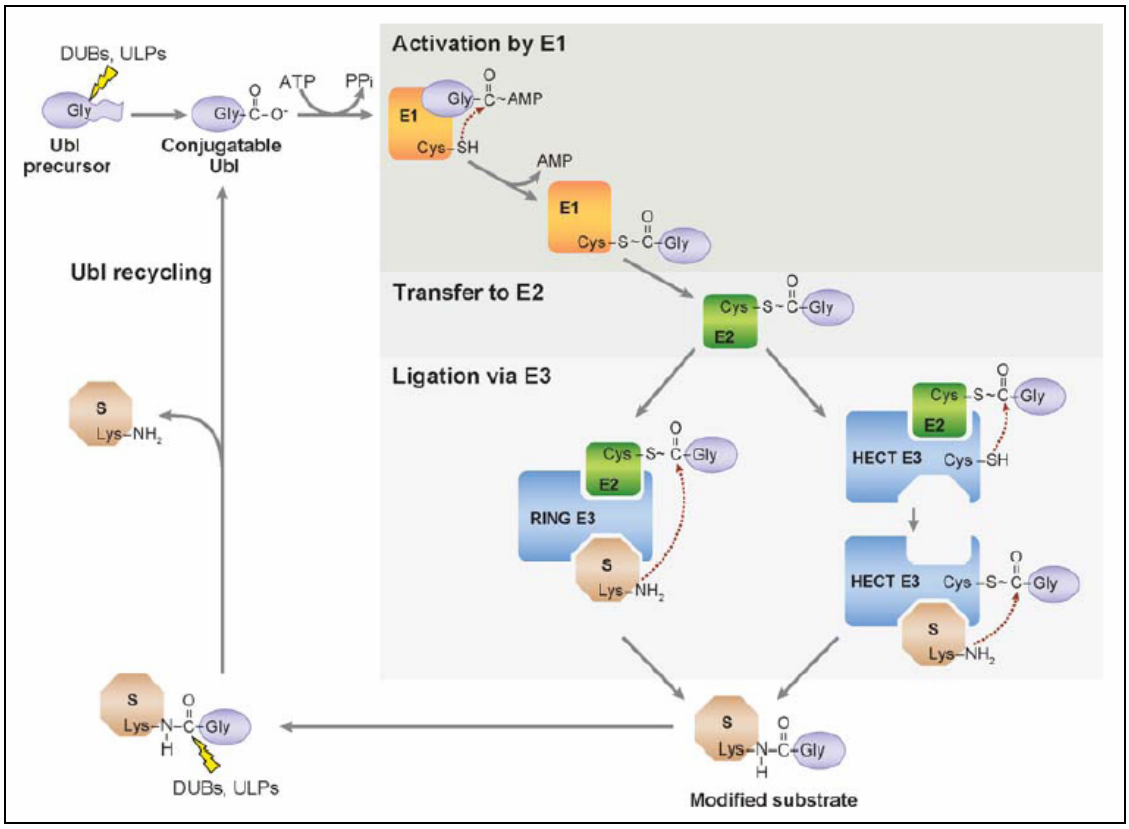


Figure 1.3 Generalized Ubl-conjugation pathway. Precursor Ubls are processed by either DUBs (deubiquitinating enzymes) or ULPs (Ubl-specific protease) to expose, if necessary, a C-terminal glycine in the mature Ubl. The processed Ubl can be activated with ATP by the E1, or Ubl-activating enzyme. The E1 adenylates the Ubl C-terminal carboxylate, forming a high-energy Ubl-AMP intermediate. This intermediate is attacked (indicated by the *red arrow*) and covalently bound by the catalytic cysteine of the E1, creating a thioester linkage and releasing AMP. The Ubl is transferred to the catalytic cysteine of an E2, or Ubl-conjugating enzyme, via a transthioesterification reaction. The Ubl can then be ligated to a substrate with the aid of an E3, or Ubl-protein ligase. The adaptor-like RING E3 catalyze the modification by binding simultaneously to the Ubl~E2 thioester complex and the substrate to be modified. This positions the amino group of a substrate lysine near the E2~Ubl thioester, thus catalyzing transfer of the Ubl to substrate. HECT E3s catalyze substrate ligation in two steps. First, the Ubl is transferred to a catalytic cysteine of the HECT E3 via transthioesterification. Then, the E3~Ubl thioester complex transfers the Ubl to the substrate. The DUBs and ULPs can remove Ubl from substrates. This figure was reproduced from Kerscher et al [26].

B. Ubl Conjugation Machineries: E1-E2-E3 Cascades

The ligation of ubiquitin and Ubls to their many substrates is performed by pairs of ubiquitin-conjugating (E2) and ligating (E3) enzymes. Many different E2 and especially E3 enzymes exist in the cell because each pair can recognize and modify only a subset of the vast number of different substrates. In contrast, the entrance of Ubls into these ligation pathways is performed by ubiquitin-activating (E1) enzymes, and in general, there is only one E1 enzyme for each different Ubl. For ubiquitin itself, the E1 (as well as a newly discovered second E1 in humans [32]) is a single chain protein of about 115 kDa whose sequence displays a weakly conserved two-fold repeat which is related to the MoeB/ThiF proteins involved in Molybdenum cofactor (Moco) and thiamine biosynthesis, respectively. For many of the other Ubls, the E1 is a heterodimer where each subunit corresponds to one half of a single-chain E1.

E1 enzymes activate their respective Ubls in a three-step process. The C-terminal carboxylate group of the Ubl, which usually ends with a pair of glycine residues, is first activated by adenylation; Ubl is then covalently joined to a conserved cysteine side chain of E1 yielding an E1~Ubl thioester. An assembly-line like process ensues, repeating the first step of adenylation of a second Ubl molecule to produce a fully-loaded E1 bearing two Ubls, one in the form of a non-covalently bound adenylate and the other as a covalently linked thioester. At this point, other enzymes in the pathway become involved, and the Ubl, which is covalently attached to E1 is transferred to form a thioester complex with an E2 enzyme, followed by the eventual transfer of ubiquitin to a target protein.

E2 Enzymes and Their Interactions

E2s accept a Ubl from the E1 and typically interact with an E3 to promote Ubl transfer to the target. The SUMO, NEDD8, and ISG15 pathways each have their own dedicated E2 and a handful of E3s, whereas the ubiquitin pathway is expansive. The human genome encodes tens of E2s and hundreds of E3s for ubiquitin, allowing for a multitude of distinct ubiquitination events.

All E2s share a ~150-residue catalytic core domain that contains the active site cysteine residue required for thioester formation with ubiquitin, and this domain is structurally conserved across many species. Although several E2s contain extensions at either or both ends of their core domains, several E2s for ubiquitin, the E2 for SUMO (Ubc9), and the E2 for ISG15 (UbcH8) consist exclusively of a catalytic core domain. Thus, the catalytic core domain is minimally sufficient for all E2 activities. In addition, several E2 catalytic core domains are regulated by posttranslational modification with SUMO [33]. Detailed structural studies have revealed the basis for many E2 activities and how they function together in a cascade.

Structures of E1-E2 and E2-E3 complexes revealed that E1- and E3-binding sites on E2s partially overlap [34-37]. Accordingly, E2 binding to E1 or E3 was found to be mutually exclusive in competitive binding experiments with three human E2-E3 pairs [38]. The cascade can proceed from one step to the next because E1 and E3 interact with different forms of E2. E1s bind free E2s, and release E2~Ubl thioesters, whereas E3s preferentially associate with E2~Ubl complexes, as evidenced by the direct association of the SUMO E3, RanBP2/Nup358, with SUMO [23].

Despite the labile nature of the E2~Ubl thioester complexes (the ~ indicates that a covalent complex has been formed), structural information has been obtained from changes in NMR signals upon forming complexes between ubiquitin and two different ubiquitin E2s [39, 40]. The complexes showed significant perturbations in resonances corresponding to the ubiquitin C-terminal tail and regions around the E2 catalytic cysteine, including the long central α-helix and nearby loops. An NMR-based docking model of the Ubc1~ubiquitin thioester intermediate suggests that the ubiquitin C-terminal tail extends through a shallow tunnel in the E2 structure, terminating at the catalytic cysteine [39]. Since the ubiquitin interface is complementary to the known E1- and E3-binding sites, these interactions may be preserved during the E1-catalyzed formation of E2~Ubl complexes and also during the E3-catalyzed transfer of the Ubl from the E2.

Ubc1

As stated earlier, all E2s share a ~150-residue catalytic core domain that is structurally conserved throughout many species. For example, 11 E2 enzymes have been identified in *S. cerevisiae*, and at least 25 are known in mammals [26]. E2 enzymes are divided into three classes in which class I enzymes are the simplest and are comprised exclusively of the core catalytic domain that contains the active site cysteine residue required for thioester formation with ubiquitin. Several E2 proteins are more complex than the class I members and have either N- or C-terminal extensions. Class II E2 proteins have a C-terminal extension or a “tail”, whereas class III E2 proteins have an additional N-terminal sequence [41]. One of the key functions of the class II E2 conjugating enzymes is the creation of the polyubiquitin chain required for protein

labeling and subsequent degradation. For example the mammalian E2 protein E2-25K is able to synthesize free Lys48-linked polyubiquitin chains in the absence of an E3 enzyme [42]. The E2 proteins Ubc1 and Ubc3 (Cdc34) from *S. cerevisiae* are able to assemble polyubiquitin chains in conjunction with an auto-ubiquitination activity [43].

Ubc1 is unique among yeast E2s in that it contains not only a catalytic core domain but also a ubiquitin-associated (UBA) domain at its C-terminus [44] (Figure 1.4). Previous studies have shown that a deletion of the UBA domain on Ubc1 alters the pattern of its autoubiquitination, but its function has not been assessed in an E3-dependent reaction [43]. UBA domains bind ubiquitin [20], and thus a reasonable hypothesis is that the UBA domain of Ubc1 aids in its ability to form polyubiquitin chains.

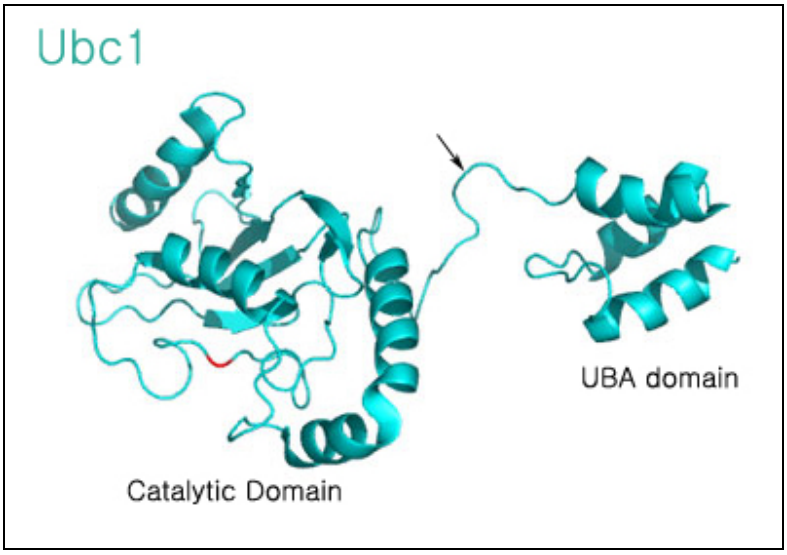


Figure 1.4 Structure of yeast Ubc1 [44]. As can be seen from the ribbon diagram, Ubc1 contains two structurally distinct domains connected by a flexible tether. The black arrow points to the tether. Ubc1's catalytic cysteine (Cys88) is colored in red.

E3 Enzymes and Their Interactions

A substantial fraction of the cellular proteins is regulated by Ubl modifications. Proteomic studies have identified thousands of such proteins in *S. cerevisiae* alone [6]. Target substrates are typically selected for modification by a vast repertoire of E3 Ubl-protein ligases, with ~600 human proteins having motifs associated with E3s [45]. E3s interact with select E2 partners, thereby identifying targets as substrates for particular Ubl modifications. E3s are classified on the basis of their E2~Ubl-binding domains and there are two main types of E3 for ubiquitin, the RING class and the HECT class [46].

The RING E3s contain a subunit or domain with a RING motif, which coordinates a pair of zinc ions. RING E3s function at least in part as adaptors: They bind the E2~Ubl and substrate protein simultaneously and position the substrate lysine nucleophile in close proximity to the reactive E2~Ubl thioester bond, thus facilitating the transfer of the Ubl. A recent study suggests that the RING E3s also triggers subtle conformational changes in the bound E2, stimulating Ubl release from the E2 cysteine and transfer to substrate [47]. Nevertheless, RING E3s do not appear to be directly involved in the chemical reactions associated with the transfer of ubiquitin from the E2 to the target substrate.

Catalysis of ubiquitination by the HECT E3s follows a mechanism distinct from that of RING E3s. In HECT E3s, the ubiquitin is first transferred from the E2s to an active site cysteine in the conserved HECT domain of the E3 in a transthioesterification reaction. The thioester-linked ubiquitin is then transferred to the substrate resulting in isopeptide bond formation. HECT E3s have a bilobed architecture, with a large distance between the E2-binding site in the N-terminal lobe and the active site cysteine in the C-

terminal lobe. For the E3s to be catalytically active, both lobes must be brought together, which may require movements of up to 50 Å [48]. Notably, when mutations that restrict movement in the hinge between the two lobes in the HECT E3s are introduced, catalytic activity decreased markedly [49].

E1 Enzymes

Each Ubl has its own dedicated E1 [50-54], which is essential for all subsequent conjugation steps [52, 55, 56]. E1s play at least three critical functions in initiating Ubl conjugation cascades. First, the E1 selects the correct Ubl for the pathway, by associating non-covalently with the Ubl. Second, the E1 activates the Ubl C-terminus, which allows the ensuing reactions to proceed. Third, the E1 coordinates the Ubl with the correct pathway, by transferring the Ubl to its cognate E2 or set of E2s in the case of ubiquitin. The enzymology of the E1 for ubiquitin has been studied extensively. Ubiquitin's E1 carries out these three functions through four enzymatic steps (Figure 1.5) [57-60].

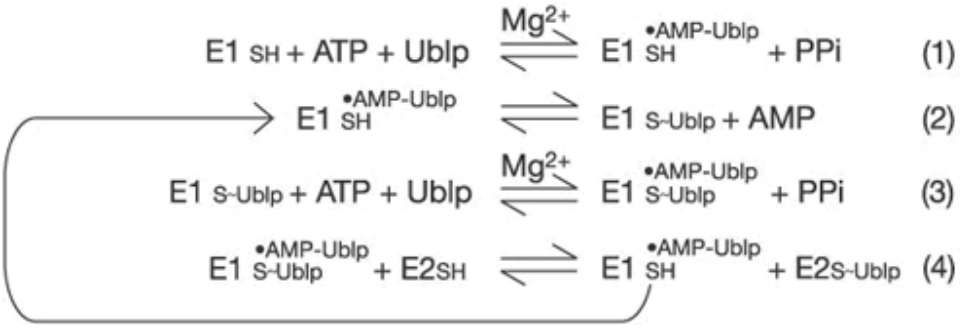


Figure 1.5 Generalized reaction scheme of E1 enzymes. (Reaction 1) A Ubl's dedicated E1 first binds ATP, Mg^{2+}, and the Ubl, and then catalyzes adenylation of the Ubl C-terminus. (Reaction 2) The E1 catalytic cysteine attacks the Ubl-adenylate, forming a thioester-linked E1-Ubl complex. (Reaction 3) The E1 repeats the adenylation reaction on a second Ubl molecule, resulting in the E1 carrying two molecules, the first bound covalently and the second non-covalently. (Reaction 4) The fully-loaded E1 binds E2 and

promotes transfer of its thioester-bound Ubl to the E2 catalytic cysteine. This figure was reproduced from Walden et al [61].

To distinguish between the different types of E1-Ubl interactions, covalent complexes will be designated with a tilde (~), and non-covalent complexes will be specified with a hyphen (-). The first reaction catalyzed by E1 is the adenylation of Ubl's C-terminus [57, 58, 62]. In addition to adding a good leaving group (AMP) on the Ubl's C-terminus, this step has also been suggested to be important for selection of the correct Ubl. The adenylated ubiquitin, a mixed acyl-phosphate anhydride between the C-terminal glycine of ubiquitin and AMP derived from the α/β cleavage of ATP, forms a tight non-covalent complex with the E1 [59, 62, 63]. If the downstream steps of the conjugation cascade are blocked, either with a nonhydrolyzable analog of the ubiquitin~adenylate [64], or by chemically blocking E1's catalytic cysteine with iodoacetamide [60], the adenylated ubiquitin remains stably associated with the E1. Studies of E1s for SUMO and NEDD8 demonstrate that the first step of their reaction cycle also results in the formation of an Ubl~adenylate intermediate [53, 65].

The adenylation reaction has been most extensively studied for ubiquitin, where the reaction proceeds via a strictly ordered mechanism, with ATP binding preceding ubiquitin binding, prior to the formation of the ubiquitin~adenylate [59, 60, 63, 66]. Recent studies of the E1 for NEDD8 [65], and of mutant versions of ubiquitin [67] demonstrate that it is also possible for the E1s to function with random addition of ATP and the Ubl. Thus, ubiquitin E1's requirement for ordered substrate addition is not a structural requirement for the catalytic competence of E1s, but reflects differences in the E1's affinities for ATP and ubiquitin as leading and trailing substrates [65].

Following the adenylation of ubiquitin's C-terminus, the E1's catalytic cysteine attacks the adenylate and forms a thioester with ubiquitin's C-terminus and the resulting E1~ubiquitin thioester serves as the proximal donor of activated ubiquitin in the formation of the E2~Ubl thioesters which are required in all ubiquitin conjugation reactions. Formation of the E1~ubiquitin thioester is rapid, with a turnover rate of ~2 per second [59]. Next, the E1 adenylates a second molecule of ubiquitin, so the E1 is loaded with two molecules of ubiquitin – the first bound covalently as a thioester, and the second one bound non-covalently as an adenylate [59]. A previous kinetic study demonstrated that the E1 for NEDD8 also proceeds through an analogous, doubly-loaded intermediate [65], thus suggesting that the mechanism of all E1s is similar. Formation of the second Ubl~adenylate prior to transfer of the Ubl to E2 may serve several functions. First, it primes E1 to carry out another round of catalysis immediately upon transfer of a molecule of Ubl to E2 [59]. Second, studies of the E1 for ubiquitin show that binding of the second molecule of ubiquitin, in the form of a ubiquitin~adenylate promotes transfer of the thioester-linked ubiquitin to E2 [68].

The final step catalyzed by E1 is the transfer of ubiquitin to E2, the next component of the reaction cascade. This involves the non-covalent association of the E1 with the E2, and a transthioesterification reaction in which the Ubl is transferred from the catalytic cysteine of the E1 to that of the E2 enzyme.

It has been suggested that the E1 plays a major role in bringing the Ubl together with the correct E2 [46]. Support for this hypothesis is derived from findings that E1s associate non-covalently with Ubl~adenylate [59], and that E1s also associate non-covalently with the E2 during the reaction [69]. Consistent with this theory, while

NEDD8 normally cannot be transferred to an E2 for ubiquitin, a NEDD8 mutant that is activated by the E1 for ubiquitin can be transferred to a ubiquitin E2 [28]. In addition, mutations of residues on either E1 or E2 that are involved in the non-covalent E1-E2 interaction diminish E2~Ubl thioester bond formation [35, 70].

Evolutionary origins of Ubl adenylation from bacterial biosynthetic enzymes

Ubls and their activating enzymes appear to have evolved from ancient and conserved biosynthetic pathways, which exist in bacteria as well as eukaryotes. MoaD and ThiS are structural homologues of ubiquitin that play a critical role in molybdenum cofactor (Moco) and thiamine biosynthesis, respectively [31, 71-74]. Like ubiquitin and other Ubls, MoaD and ThiS are adenylated at their C-terminus prior to downstream steps in the pathways [75, 76]. Unlike ubiquitin and Ubls such as NEDD8, SUMO, ISG15 and Apg8, however, the ultimate functions of MoaD and ThiS are not to be conjugated to other proteins. Instead, MoaD and ThiS are temporarily modified themselves, carrying sulfur atoms at their C-terminus as thiocarboxylates, and are involved in sulfur transfer in the Moco and thiamine biosynthetic pathways [76, 77]. Thus, a common feature of Ubl and the bacterial biosynthetic pathways is the involvement of the Ubl C-terminus in several chemical reactions. Interestingly, MoeB and ThiF, the enzymes that catalyze adenylation of the C-terminus of MoaD and ThiS proteins, respectively, share significant sequence homology to domains of the E1s for ubiquitin and Ubls [52, 53]. This sequence homology includes a Gly-X-Gly-X-X-Gly nucleotide binding motif, suggesting that the C-terminal adenylation reaction is similar for ubiquitin, Ubl, MoaD and ThiS pathways.

Three crystal structures of MoeB-MoaD complexes revealed the structural basis for Ubl adenylation (Figure 1.6).

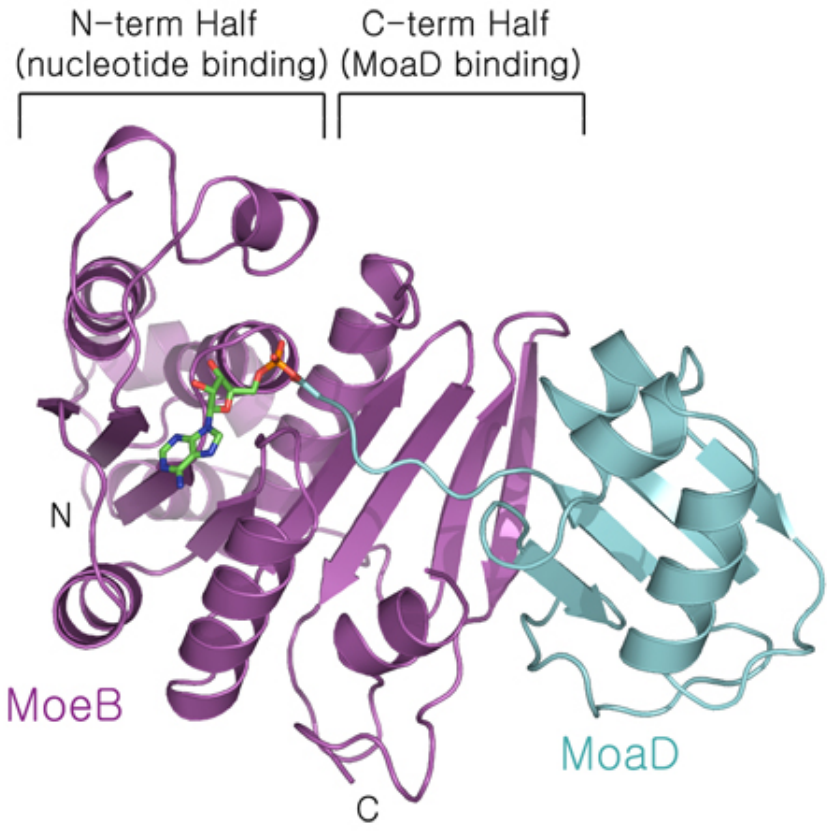


Figure 1.6 Structure of the MoeB-MoaD~adenylate complex [31]. MoeB is shown in magenta, MoaD in cyan and AMP in green. The N-terminal subdomain of MoeB adopts a variation of the Rossmann fold and is involved in ATP binding and catalyzes the adenylation reaction. The C-terminal subdomain contains an antiparallel, four-stranded β-sheet involved in binding to the globular domain of MoaD.

Modular nature of the Ubl-activating enzyme

Analyses of primary sequences of Ubl-activating enzyme, in conjunction with previously determined crystal structures described below, suggest that Ubl-activating enzymes are modular proteins that have evolved multiple distinct domains to carry out their specific functions. The common feature of all Ubl-activating enzymes is a region of high sequence homology to MoeB and ThiF, which suggests that the common function associated with E1s is the C-terminal adenylation of Ubls. Additional sequences at the N- and C-termini, and in the middle of the MoeB/ThiF homology domain are likely to correspond to new domains that have evolved for specific functions of a particular Ubl-activating enzyme.

The four Ubls whose E1s are most closely related are ubiquitin, NEDD8, SUMO and ISG15. The E1s for these modifiers are 110-120 kDa, and either exist as single polypeptides or heterodimeric complexes (Figure 1.7). In the heterodimeric E1s, one subunit corresponds to roughly the N-terminal half of the single-chain E1s, and the other corresponds to roughly the C-terminal half. The E1s for ubiquitin and ISG15 are single polypeptides, whereas the E1s for NEDD8 and SUMO family members are ~110 kDa heterodimers, with one protein (depending on species they are named APPBP1, AXR1 or Ula1 for NEDD8 and Sae1 or Aos1 for SUMO) homologous to the N-terminal half of the E1s for ubiquitin and ISG15, and the other protein (UBA3 or ECR1 for NEDD8 and Sae2 or Uba2 for SUMO) homologous to the C-terminal half of the E1 for ubiquitin.

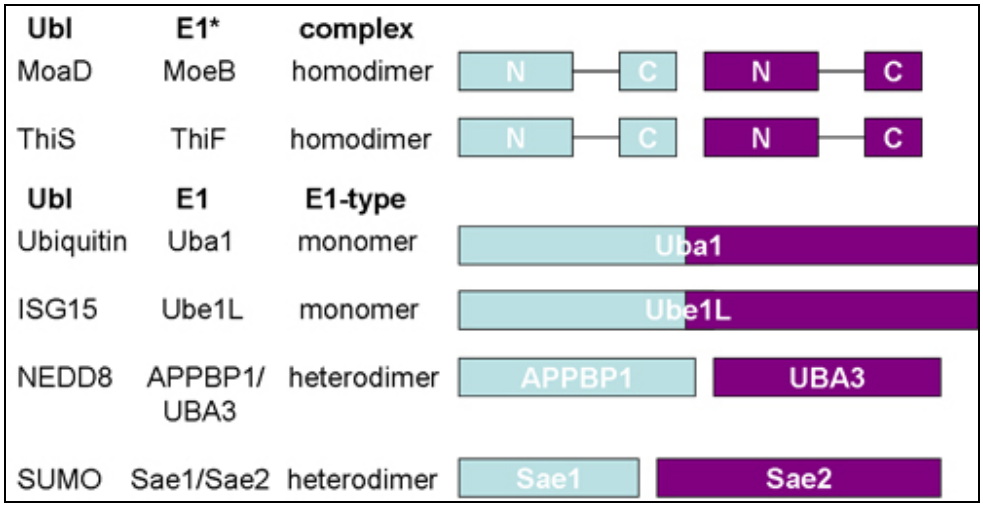


Figure 1.7 Comparison of the activating enzymes for Ubiquitin and Ubls. The activating enzymes for *E. coli* MoaD and ThiS, MoeB and ThiF are ~27 kDa, and the crystal structures of MoeB and ThiF revealed a homodimer [31, 72, 73]. The sequences corresponding to the N- and C-terminal subdomains in the MoeB structure are referred to as 'N' and 'C', respectively. The E1s for ubiquitin, ISG15, NEDD8 and SUMO are ~110-120 kDa proteins or heterodimeric complexes that contain a twofold repeat that corresponds to the sequence and structure of MoeB. (*) denotes that MoeB and ThiF are bacterial ancestors of the E1 enzymes.

The N- and C-terminal halves of the E1s for ubiquitin, NEDD8, SUMO and ISG15 are partially homologous to each other, and the region of sequence homology between the two halves is also the region of sequence homology to MoeB and ThiF (Figure 1.8). The crystal structures of MoeB, APPBP1-UBA3 (NEDD8-E1) and Sae1-Sae2 (SUMO-E1) complex revealed that this homologous region adopts a Rossmann fold, a structure frequently found in nucleotide binding proteins [29, 31, 61, 78]. Indeed, the Gly-X-Gly-X-X-Gly (X denotes any amino acid) ATP binding motif [79] resides in the homology region in the C-terminal half of E1s. In addition to playing a role in nucleotide binding, this region is also involved in dimerization as shown in the homodimer interface in the symmetric MoeB-MoaD crystal structure [31], and in the protein-protein interface in the APPBP1-UBA3 [61] and Sae1-Sae2 complexes [29]. Despite the presence of two

MoeB repeats in E1 enzymes, the E1s for ubiquitin, NEDD8 and SUMO contain only one functional active site in their C-terminal halves compared to two for the MoeB homodimer. Hence, for Uba1 the N-terminal MoeB homology region, which lacks the ATP binding motif will be referred to as the "inactive" adenylation domain (IAD) and the C-terminal MoeB homology region as the "active" adenylation domain (AAD) hereafter. (Figure 1.9)

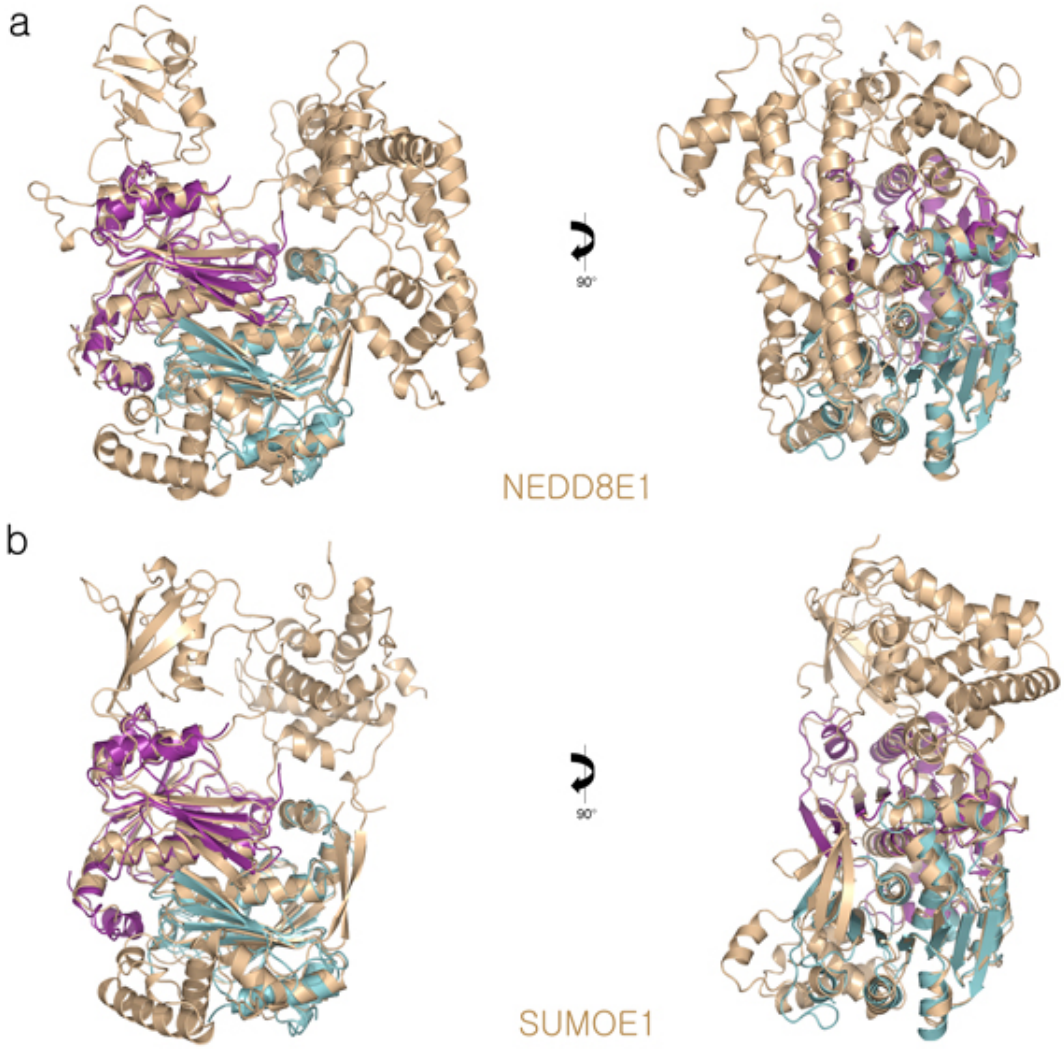


Figure 1.8 The structures of APPBP1-UBA3 (NEDD8-E1) [80] and Sae1-Sae2 (SUMO-E1) [29] complex. Ribbon representation of (a) NEDD8-E1 and (b) SUMO-E1 with the MoeB homodimer (cyan and purple) superimposed on the two adenylation domains.

At the location of a highly mobile, 8-residue linker in MoeB, both the NEDD8-E1 and SUMO-E1 contain large insertions (up to 220 residues) in each MoeB-related region. The structures of the two E1s suggest that these insertions have evolved to carry out E1-specific functions such as transfer of Ubl to the catalytic cysteine of cognate E2s. Indeed, the insertion in the C-terminal half of the E1 sequences contains the catalytic cysteine. Here, the two parts of the catalytic cysteine domain will be referred to as the first catalytic cysteine half-domain (FCCH) and the second catalytic cysteine half-domain (SCCH), even though the two halves differ in molecular weight and show variations in the sequence and the secondary structure composition in both the ubiquitin-E1 and its homologues (Figure 1.9). 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 [75].

Finally, another E1-specific domain is found at the C-terminus of all E1 enzymes. This domain will be termed the ubiquitin-fold domain (UFD) due to its structural similarity to ubiquitin and other Ubls.

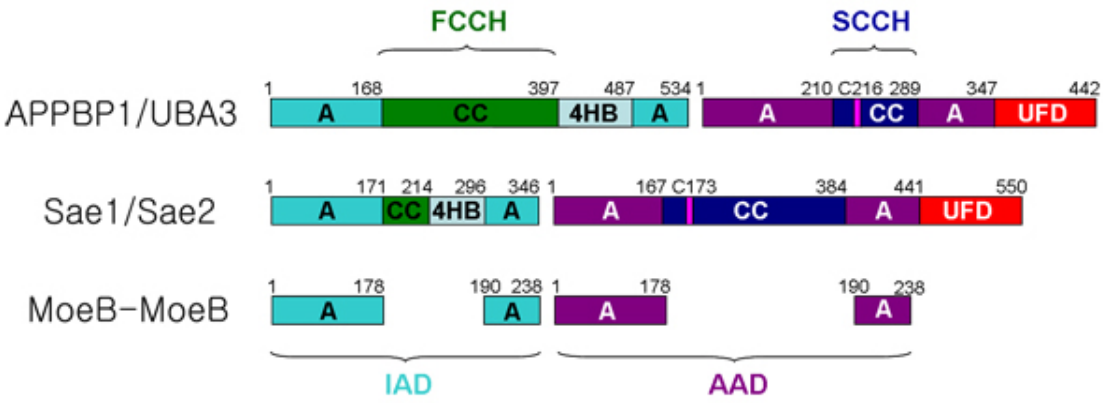


Figure 1.9 Schematic domain representation observed in NEDD8-E1, SUMO-E1, and the crystallographic MoeB dimer. Numbers and color coding indicate the domain boundaries. The adenylation half-domains are shown in cyan and purple and the catalytic cysteine half-domains are shown in forest and blue. E1 catalytic cysteines are shown in pink. A, adenylation domain; CC, catalytic cysteine domain; UFD, ubiquitin-fold domain; IAD, inactive adenylation domain; AAD, active adenylation domain; FCCH, first catalytic cysteine half-domain; SCCH, second catalytic cysteine half-domain; 4HB, four helix bundle domain.

Uba1

The ubiquitin-activating enzyme is highly conserved in yeast [52], plants [50], and humans [81]. The E1 enzyme plays an essential role in yeast including for example during sporulation and cell proliferation, since deletion of the *uba1* gene, which encodes for the yeast E1 enzyme is lethal [52]. Moreover, a hypomorphic allele of UBA1 was identified, which impairs ubiquitin conjugation to substrate proteins [82]. Several mammalian cell lines with mutations that render the UBA1 gene temperature-sensitive showed a severe defect in ubiquitin conjugation and decrease in protein turnover when shifted to a non-permissive temperature [51, 55, 56]. Although it has so far been assumed that a single activating enzyme for ubiquitin exists, very recent studies suggest the presence of a second ubiquitin E1 enzyme in humans [32, 83].

Uba1 consists of three building blocks: first, the adenylation half-domains (IAD and AAD) composed of two MoeB/ThiF-homology motifs, one of which binds ATP and ubiquitin; second, the catalytic cysteine half-domains (FCCH and SCCH) inserted into each of the adenylation half-domains, which contains the E1 active site cysteine (Cys 600 in yeast Uba1); and third, the C-terminal ubiquitin-fold domain (UFD), which recruits specific E2s.

CHAPTER 2

Structure of the non-covalent Uba1-ubiquitin complex

I. Introduction

The E1 activity represents an essential step during Ubl conjugation. Each Ubl has a dedicated E1, or activating enzyme, that initiates its conjugation cascade. First, E1 associates with the Ubl and catalyzes the adenylation of the Ubl C-terminus in an ATP-dependent process. Second, E1 forms a thioester between its conserved catalytic cysteine and the Ubl. Finally, the E1~Ubl thioester complex then recruits an E2 to facilitate transfer of the thioester-linked Ubl to a conserved E2 cysteine (transthioesterification). The energy stored in the E2~Ubl thioester is utilized to conjugate Ubl to target lysine ε-amino groups, either directly or through complexes mediated by E3s. Despite the central role of the ubiquitin-activating enzyme (Uba1 in yeast) in this cascade, a crystal structure of a ubiquitin-activating enzyme had not been available prior to the work presented here.

To gain structural and functional insights into the selection of ubiquitin by Uba1, Uba1~ubiquitin thioester formation, and Uba1-mediated transfer of ubiquitin to E2, I crystallized the Uba1-ubiquitin complex and determined its structure at 2.7 Å resolution using molecular replacement. In this chapter, I describe the structure in order of the three E1 activities: adenylation, thioester bond formation and Ubl transfer to E2. The structural data along with biochemical analyses of Uba1 mutants provide mechanistic insights into Uba1 function.

II. Materials and Methods

A. Cloning and protein expression of Uba1

The *S. cerevisiae uba1* gene was cloned into the NheI and EcoRI sites of the pET28a vector (Novagen) using Uba1 sense, 5'-GCAGCCATATGGCTAGCGCCGCCG GAGAAATCG-3' and Uba1 antisense, 5'-GTTAGCAGCCGGATCGAATTCTCATAG ATGAATGG-3' oligonucleotides as primers. The expression construct introduced a hexahistidine tag at the N-terminus of the protein. The pET28a vector of the T7 pET system (Novagen) containing the *S. cerevisiae uba1* gene was transformed into Rosetta (DE3) cells by heat shock. After overnight growth, a single colony was picked from a plated culture onto an LB-agar plate containing 50 μg/ml kanamycin and 34 μg/ml chloramphenicol and used to inoculate a 15 ml-LB culture in the presence of the same concentrations of kanamycin and chloroamphenicol. The culture was incubated at 37 °C overnight, and the entire culture was used to inoculate 3 liter of LB medium containing both antibiotics. Expression of the protein was induced by the addition of 0.3 mM IPTG to the growing culture when the cells reached an OD_{600} of 0.6-0.8 after incubation at 37 °C. After addition of IPTG, the *E. coli* culture was incubated at 16°C for 18 hours.

	Restriction Sites	Vector	Antibiotic resistance
WT-His-Uba1(10-1024)*	NheI, EcoRI	pET28a	Kan
C600A-His-Uba1 FL	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(1-361)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(10-361)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(405-573)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(1-573)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(10-573)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(597-860)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(1-924)	NheI, EcoRI	pET28a	Kan
WT-His-Uba1(10-924)	NheI, EcoRI	pET28a	Kan
WT-Uba1 FL	NcoI, XhoI	pET16b	Amp
WT-Uba1-His FL	NheI, EcoRI	pET21b	Amp
C600A-Uba1-His FL	NheI, EcoRI	pET21b	Amp
WT-Uba1-Intein FL	NheI, SapI	pTXB1	Amp
C600A-Uba1-Intein FL	NheI, SapI	pTXB1	Amp
WT-Intein-Uba1 FL	SapI, EcoRI	pTYB11	Amp
C600A-Intein-Uba1 FL	SapI, EcoRI	pTYB11	Amp

Table 2.1 Summary of Uba1 constructs used for expression and crystallization. Numbers inside the parentheses stand for the terminal residues of Uba1. Truncated domain constructs in pET28a (Rows 3 – 10) were also attempted in other vectors. The His-tagged proteins were purified following similar protocols as described in Materials and Methods section C, and the intein-tagged proteins following the protocol outlined in section D. WT: wild-type, C600A: Cys600Ala mutant of Uba1, FL: full-length, (*) denotes the construct which yielded crystals.

B. Cloning and expression of Ubc1 and ubiquitin

The *S. cerevisiae* Ubc1 and ubiquitin proteins were cloned into the NdeI and SapI sites of the pTXB1 vector (New England Biolabs). The resulting constructs contained an intein tag at the C-terminus of either protein. The pTXB1 vector of the T7 IMPACT^{TM} system (New England Biolabs) containing either the *S. cerevisiae ubc1* or *ubiquitin* genes was transformed as described above into BL21 (DE3) cells. A single colony was picked from a culture plated onto an LB-agar plate containing 100 μg/ml ampicillin and used to

inoculate 15 ml of LB medium plus ampicillin. The culture was incubated at 37 °C overnight, and with the complete culture, 3 liter of LB medium were inoculated. The expression of either protein was induced by the addition of 0.3 mM IPTG to the growing culture when the cells had reached an OD_{600} of 0.6-0.8 after incubation at 37 °C. After addition of IPTG, cells were incubated at 16°C for 18 hours.

C. Protein purification for Uba1

Optimized purification protocol

All procedures were performed at 4 °C. The cells were harvested by centrifugation (16 min at 12,000 x g) and then resuspended in lysis buffer (50 mM Tris, pH 7.6, 500 mM NaCl, 5 mM β-mercaptoethanol, 25 mM imidazole, complete protease inhibitor (EDTA free, Roche), 5% glycerol). The cell walls were ruptured by passing them twice through a French Pressure Cell at a pressure of 14,000 psi, and the lysate was further centrifuged for 36 minutes at 75,000 x g to remove cell debris. The supernatant was loaded onto a column containing 7.5 ml of Ni-NTA beads (Qiagen). The column was thoroughly washed by the addition of 100 ml wash buffer (50 mM Tris, pH 7.6, 500 mM NaCl, 25 mM imidazole) prior to elution with 250 mM imidazole in the same buffer. The eluted sample was then dialyzed against a buffer containing 25 mM Tris, pH 7.6, 150 mM NaCl and 3 mM β-mercaptoethanol to remove the imidazole. The His-tag was cleaved off using thrombin (Sigma) at 4 °C overnight, and uncleaved protein was separated by a second subtractive Ni^{2+} affinity column. Solid $(NH_4)_2SO_4$ was added slowly under gentle stirring up to a final concentration of 1.0 M to the pooled protein after the Ni^{2+} affinity column. The resulting solution was centrifuged (7 min at 3,200 x g)

and the supernatant was loaded onto a HiLoad 16/10 phenyl sepharose column (Amersham) equilibrated in 50 mM Tris, pH 7.6, 1.0 M $(NH_4)_2SO_4$ and 10 mM dithiothreitol (DTT) prior to elution with a descending linear gradient to 200 mM $(NH_4)_2SO_4$ in the same buffer. The pooled fractions from the hydrophobic interaction column were concentrated and then purified on a HiLoad Superdex 26/60 S200 column (Amersham) equilibrated with 25 mM Tris, pH 7.6, 200 mM NaCl and 5 mM DTT. The fractions containing the desired protein were concentrated using a Centricon Plus-20 (30000 MWCO) concentrator (Millipore) to 5 mg/ml as determined by UV/Vis spectroscopy using a calculated extinction coefficient of 70,250 $M^{-1}cm^{-1}$ at 280 nm. 20 µl protein aliquots were prepared and flash-frozen by pipetting them directly into liquid nitrogen and stored at -80 °C.

Initial purification protocol

All procedures are the same as in the protocol described above except that the hydrophobic interaction (Phenyl Sepharose) column was not utilized. Instead, the pooled protein from the Ni^{2+} affinity column was loaded onto a Mono Q 10/10 column (Amersham) after passing it over a PD-10 desalting column (Amersham) and chromatographed using a NaCl gradient (Buffer A: 50 mM Tris, pH 7.6, 100 mM NaCl, 5 mM DTT; Buffer B: Buffer A containing 1 M NaCl instead of 100 mM NaCl). The protein sample was concentrated and loaded onto a HiLoad Superdex 26/60 S200 column (Amersham).

D. Protein purification for Ubc1 and ubiquitin

The cells were harvested by centrifugation (16 min at 12,000 x g) and then resuspended in lysis buffer (20 mM Tris, pH 8.0, 350 mM NaCl, complete protease inhibitor (EDTA free, Roche)) and lysed by passing them twice through a French pressure cell at a pressure of 14,000 psi. After centrifugation (40,000 x g, 35 min), the supernatant was loaded onto a chitin affinity column equilibrated in the buffer described above. The column was then washed with a buffer containing 20 mM Tris, pH 8.0, 1 M NaCl prior to on-column cleavage overnight at 4 °C in 20 mM Tris, pH 8.0, 350 mM NaCl, 100 mM DTT. The target protein was eluted using additional cleavage buffer without DTT. The pooled fractions were concentrated using a Centricon Plus-20 (5000 MWCO) concentrator (Millipore) and applied to a HiLoad Superdex 26/60 S200 column (Amersham) equilibrated with 25 mM Tris, pH 7.6, 200 mM NaCl and 1 mM DTT. The purified protein was concentrated to 10 mg/ml as determined by UV/Vis spectroscopy using calculated molar extinction coefficients of 19,940 $M^{-1}cm^{-1}$ for Ubc1 and 1,490 $M^{-1}cm^{-1}$ for ubiquitin at 280 nm. 20 µl protein aliquots were prepared for each protein and flash-frozen in liquid nitrogen and stored at -80 °C.

E. Confirmation of the target protein (Uba1)

To verify whether the final product of purification in fact is *S. cerevisiae* Uba1, the purified protein was subjected to mass spectrometry using the Voyager DE-STR (Applied Biosystems) mass spectrometer of the Proteomics Center in the Stony Brook University Medical Center. A mass spectrum of the peptide mixture resulting from the

digestion of sample by trypsin provided a fingerprint, which was searched against primary sequence databases.

F. Crystallization and data collection

Initial crystallization trials were performed with full-length Uba1 (residues 1-1024). Small quasi-crystals were obtained by equilibrating a mixture containing 1 µl of protein (5 mg/ml) and 1 µl of reservoir solution consisting of 0.05 M cadmium sulfate, 1.0 M sodium acetate and 0.1 M HEPES, pH 7.5 against the reservoir solution in hanging drop vapor diffusion experiments. However, all standard optimization trials, including micro- and macroseeding, failed to provide diffraction-quality crystals. Generation of Uba1 and Uba1-ubiquitin complexes suitable for crystallization involved testing more than 50 different variant proteins of full-length and deletion mutants of Uba1 (Table 2.1). During this process, I obtained crystals of the complex lacking the N-terminal residues 1-9 (Δ9Uba1) in a couple of crystallization conditions (1: 0.4 M potassium nitrate, 20% PEG 8000 and 0.1 M glycyl-glycine, pH 8.5, 2: 0.5 M L-proline and 20% PEG 5000 MME) using a Honeybee 961 crystallization robot (Genomic Solutions). All attempts with the other constructs were unsuccessful. Secondary structure analysis of the extreme N-terminus of Uba1 predicted that the first 9 residues lack any identifiable secondary structure and are very likely unstructured. The examination of crystal contacts after the structure had been solved revealed that the region is located next to a symmetry-related molecule, potentially interfering with crystallization.

Crystals of the Uba1-ubiquitin complex were grown by incubating Δ9Uba1 and ubiquitin together at concentrations of 5 mg/ml and 10 mg/ml in a molar ratio of 1.0:1.3

at 4 °C for 1 hour, followed by hanging drop vapor diffusion at room temperature against a reservoir containing 0.5 M L-proline and 17-20 % PEG 5000 MME. Crystals suitable for data collection appeared after 5-7 days and grew to a final size of ~700 x 100 x 80 μm^3 (Figure 2.1). Crystals were transferred into mother liquor supplemented with 20% PEG 400 and 5% sucrose and were flash frozen in liquid nitrogen. Diffraction data were collected at a temperature of 100 K to a resolution of 2.7 Å at beamline X29 at the National Synchrotron Light Source at Brookhaven National Laboratory at a wavelength of 1.1 Å using ADSC Quantum-315 detector. Diffraction data were indexed, integrated and scaled using HKL2000.

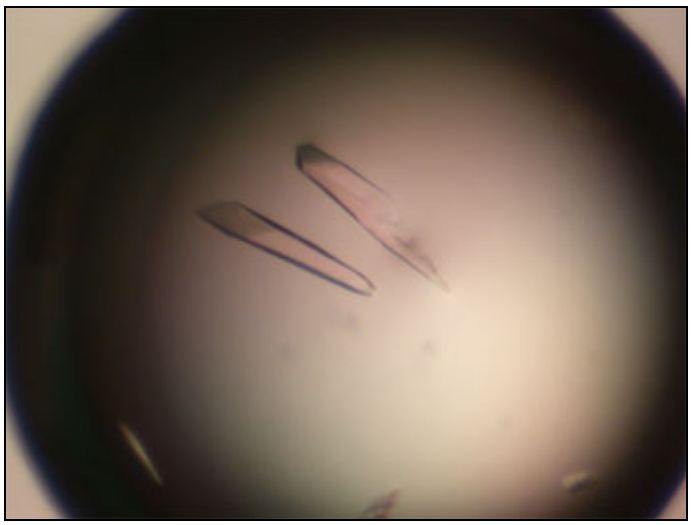


Figure 2.1 Crystals of the Uba1-ubiquitin complex

G. Generation of the transthioesterification models for Uba1

For the pre-transthioesterification model, the $Ubc1^{core}$ domain (residues 1-150, PDB entry 1TTE [44]) were modeled onto the Uba1-ubiquitin complex structure in O [84], first by least-squares superposition of the UBA3 UFD -$Ubc12^{core}$ complex (PDB

entry 1Y8X [35]) with the UFD of Uba1, followed by least-squares superposition of the Ubc1^{core} with the Ubc12^{core}. The structure of the UBA3 UFD and Ubc12^{core} was subsequently removed.

For the post-transthioesterification model, the Ubc1~ubiquitin intermediate was modeled onto the Uba1-ubiquitin complex structure in O, first by superposition of the UBA3 UFD -Ubc12^{core} complex with the UFD of Uba1, followed by superposition of the Ubc1~ubiquitin thioester complex (PDB entry 1FXT [39]) with the Ubc12^{core}. The structures of the UBA3 UFD and Ubc12^{core} were subsequently removed.

H. Generation of Uba1 variants

The QuikChange kit from Stratagene was used to generate the Asp544Ala, Tyr586Ala, Asp591Tyr, Phe898Ala, Phe905Ala, Asp782Ala, Ala913Pro, Ser914Pro, Glu1004Ala, Asp1014Lys/Glu1016Lys as well as domain deletion mutant ΔFCCH. Primers used to generate these mutants are shown in Table 2.2. For each desired mutation, 125 ng of both primers (forward and reverse complement) were added to 50 ng of double-stranded template DNA. For all mutagenesis reactions, the wild-type *uba1* gene inserted into a pET28a vector was used as a template. Additionally, 5 µl of 10x reaction buffer and 1 µl of dNTP mix were added and the total volume was adjusted to 50 µl with ddH_2O. Finally, 1 µl of Pfu Turbo DNA polymerase (2.5 U/µl) was added. This mixture was subjected to 16 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. Next, 1 µl of the DpnI enzyme (10 U/µl) was added to the reaction mixture and incubated for 2 hours at 37°C in order to digest the non-mutated

parental DNA template. Finally, 20 µl of the DpnI digested reaction mixture was transformed into 100 µl of chemically competent DH5α cells and plated onto LB agar plate containing 50 µg/ml kanamycin. DNA sequences from isolated plasmids of the resulting colonies were verified by automated sequencing.

Mutant	Primer sequences
D544A	5'-CCAACGCTCTAGCCAATGTCGACGC-3' 5'-GCGTCGACATTGGCTAGAGCGTTGG-3'
Y586A	5'-CCAAGATTGACTGAATCAGCGTCTTCTTCTAGAGACCC-3' 5'-GGGTCTCTAGAAGAAGACGCTGATTCAGTCAATCTTGG-3'
S589A	5'-CTGAATCATACTCTTCTGCGAGAGACCCACCAGAAAAG-3' 5'-CTTTTCTGGTGGGTCTCTCGCAGAAGAGTATGATTCAG-3'
D591Y	5'-GAATCATACTCTTCTTCTAGATATCCACCAGAAAAGTCTATCCC-3' 5'-GGGATAGACTTTTCTGGTGGATATCTAGAAGAAGAGTATGATTC-3'
T601A	5'-CTATCCCATTGTGTGCGCTACGTTCTTTCCC-3' 5'-GGGAAAGAACGTAGCGCACACAATGGGATAG-3'
H611A	5'-CCCAAACAAGATTGATGCGACCATTGCCTGGGCC-3' 5'-GGCCCAGGCAATGGTCGCATCAATCTTGTTTGGG-3'
D782A	5'-GAAAATTCAAGTTAATGCCGATGATCCGGATCC-3' 5'-GGATCCGGATCATCGGCATTAACTTGAATTTTC-3'
D782N	5'-GAAAATTCAAGTTAATAATGATGATCCGGATCCAAATGCC-3' 5'-GGCATTTGGATCCGGATCATCATTATTAACTTGAATTTTC-3'
F898A	5'-GCAATATAAGAATGGCGCGGTTAATTTAGCTTTGCC-3' 5'-GGCAAAGCTAAATTAACCGCGCCATTCTTATATTGC-3'
F905A	5'-GTTAATTTAGCTTTGCCAGCGTTCGGTTTTTCGGAACC-3' 5'-GGTTCCGAAAAACCGAACGCTGGCAAAGCTAAATTAAC-3'
A913P	5'-CGGAACCAATTCCGTCACCAAAGGG-3' 5'-CCCTTTGGTGACGGAATTGGTTCCG-3'
S914P	5'-GTTTTTCGGAACCAATTGCTCCGCCAAAGGGAGAATATAACAAC-3' 5'-GTTGTTATATTCTCCCTTTGGCGGAGCAATTGGTTCCGAAAAAC-3'
A913P/S914P	5'- GGTTTTTCGGAACCAATTCCGCCGCCAAAGGGAGAATATAAC-3' 5'- GTTATATTCTCCCTTTGGCGGCGGAATTGGTTCCGAAAAACC-3'
E1004K	5'-CTACAATGATTCTCAAAATTTGCGCAGATG-3' 5'-CATCTGCGCAAATTTTGAGAATCATTGTAG-3'
D1014K/E1016K	5'-CAGATGACAAGGAAGGAGAGAAAGTTAAAGTTCCTTTCATTACC-3' 5'-GGTAATGAAAGGAACTTTAACTTTCTCTCCTTCCTTGTCATCTG-3'
ΔFCCH	5'-GTTGGACCCAACGGGTGAAGAAGTCAAAGTACCCCGTAAAATC-3' 5'-GATTTTACGGGGTACTTTGACTTCTTCACCCGTTGGGTCCAAC-3'

Table 2.2 Oligonucleotide primers utilized to generate the Uba1 variants

I. Uba1 activity assays

Multiple-turnover Uba1-Ubc1 transthioesterification assay

Uba1 wild-type and mutants (100 nM), ubiqutin (4 μM), Ubc1 (2 μM) were incubated in 25 mM Tris, pH 7.5, 50 mM NaCl, 2.5 mM ATP, 5 mM $MgCl_2$ at room temperature for the indicated times. The reactions were terminated by adding 6x SDS-gel loading buffer without DTT, and the samples were resolved via SDS-PAGE. The reaction mixtures were detected by Western blot using a mouse monoclonal antibody against ubiquitin (Santa Cruz Biotechnology) and the IRDye 800CW labeled secondary antibody (LI-COR Biosciences), in combination with an Odyssey infrared imaging system (LI-COR Biosciences).

Single-turnover Uba1-Ubc1 transthioesterification assay

Uba1 (1 μM) was incubated with ubiquitin (4.1 μM) in 25 mM Tris, pH 7.5, 50 mM NaCl, 2.5 mM ATP, 5 mM $MgCl_2$ for 30 min at room temperature. The charging reaction was treated with 35 mM EDTA for 15 min at room temperature (added as 1/7 the volume of the charging reaction). This reaction was diluted into an equal volume of chase mixes containing a final concentration of 25 mM Tris, 50 mM NaCl, pH 7.5 in the presence of Ubc1 (1.5 μM final concentration). Samples were taken at different time points after the start of the chase reaction and stopped by mixing with non-reducing 6x SDS-gel loading buffer. The samples were fractionated by SDS-PAGE and analyzed by Western blot following the same protocol as above.

Uba1~ubiquitin thioester formation assay

Uba1 (1 μM) was incubated with ubiquitin (2.7 μM) in 25 mM Tris, pH 7.5, 50 mM NaCl, 2.5 mM ATP, 5 mM $MgCl_2$ at room temperature for the indicated times. The reaction was stopped by adding non-reducing 6x SDS-PAGE sample buffer, and analyzed by SDS-PAGE and Western blot as described above.

III. Results

A. Preparation of Uba1 for Crystallization

The initial three step purification of the WT-His-Uba1 (pET28a) construct (see Materials and Methods section C) could not prevent a significant contamination with *E. coli* host proteins and degradation products of the overexpressed Uba1 protein (Figure 2.2) and consequently did not produce an Uba1 sample with sufficiently high purity for crystallization. However, replacing the ion exchange with hydrophobic interaction chromatography yielded protein samples with much higher purity suitable for crystallization. The mass spectrometric (MS) analysis of the final product from the optimized purification protocol confirmed that it is indeed *S. cerevisiae* Uba1 (Figure 2.3).

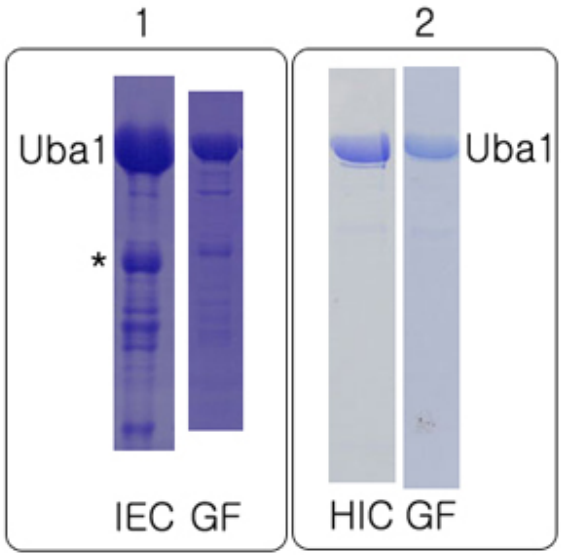


Figure 2.2 SDS-PAGE analyses of the Uba1 sample from the initial protocol (1) and from the optimized protocol (2). Full-length bands of Uba1 are labeled. A major contaminating *E. coli* host protein is labeled with (*). IEC: Ion exchange chromatography, HIC: Hydrophobic interaction chromatography, GF: Gel filtration

Match to: **gi|4715; Score: 220**
ubiquitin-activating enzyme [Saccharomyces cerevisiae]

Nominal mass (M_r): **114221**; Calculated pI value: **5.00**
NCBI BLAST search of gi|4715 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Saccharomyces cerevisiae

Variable modifications: Carbamidomethyl (C),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **50**
Number of mass values matched: **27**
Sequence Coverage: **41%**

Matched peptides shown in **Bold Red**

```
   1 MSSNNSGLSA AGEIDESLYS RQLYVLGKEA MLKMQTSNVL ILGLKGLGVE
  51 IAKNVVLAGV KSMTVFDPEP VQLADLSTQF FLTEKDIGQK RGDVTRAKLA
 101 ELNAYVPVNV LDSLDDVTQL SQFQVVVATD TVSLEDKVKI NEFCHSSGIR
 151 FISSETRGLF GNTFVDLGDE FTVLDPTGEE PRTGMVSDIE PDGTVTMLDD
 201 NRHGLEDGNF VRFSEVEGLD KLNDGTLFKV EVLGPFAFRI GSVKEYGEYK
 251 KGGIFTEVKV PRKISFKSLK QQLSNPEFVF SDFAKFDRAA QLHLGFQALH
 301 QFAVRHNGEL PRTMNDEDAN ELIKLVTDLS VQQPEVLGEG VDVNEDLIKE
 351 LSYQARGDIP GVVAFFGGLV AQEVLKACSG KFTPLKQFMY FDSLESLPDP
 401 KNFPRNEKTT QPVNSRYDNQ IAVFGLDFQK KIANSKVFLV GSGAIGCEML
 451 KNWALLGLGS GSDGYIVVTD NDSIEKSNLN RQFLFRPKDV GKNKSEVAAE
 501 AVCAMNPDLK GKINAKIDKV GPETEKIFND SFWESLDFVT NALDNVDART
 551 YVDRRCVFYR KPLLESGTLG TKGNTQVIIP RLTESYSSSR DPPEKSIPLC
 601 TLRSFPNKID HTIAWAKSLF QGYFTDSAEN VNMYLTQPNF VEQTLKQSGD
 651 VKGVLESISD SLSSKPHNFE DCIKWARLEF EKKFNHDIKQ LLFNFPKDAK
 701 TSNGEPFWSG AKRAPTPLEF DIYNNDHFHF VVAGANLRAY NYGIKSDDSN
 751 SKPNVDEYKS VIDHMIIPEF TPNANLKIQV NDDDPDPNAN AANGSDEIDQ
 801 LVSSLPDPST LAGFKLEPVD FEKDDDTNHH IEFITACSNC RAQNYFIETA
 851 DRQKTKFIAG RIIPAIATTT SLVTGLVNLE LYKLIDNKTD IEQYKNGFVN
 901 LALPFFGFSE PIASPKGEYN NKKYDKIWDR FDIKGDIKLS DLIEHFEKDE
 951 GLEITMLSYG VSLLYASFFP PKKLKERLNL PITQLVKLVT KKDIPAHVST
1001 MILEICADDK EGEDVEVPFI TIHL
```

Figure 2.3 Mascot search results. Mascot is a search engine which uses mass spectrometry data to identify proteins from primary sequence databases (www.matrixscience.com). Amino acids colored in red where those identified by mass spectrometry of tryptic peptides.

B. Structure Determination

The Uba1-ubiquitin crystals belong to space group $P2_12_12_1$ with unit cell dimensions of a = 115.36 Å, b = 118.56 Å, c = 207.57 Å, and contained two molecules of the Uba1-ubiquitin complex per asymmetric unit (Table 2.3). The packing density of the crystals is fairly low with a calculated Matthew's coefficient (Vm) of 3.0 $Å^3$/Dalton

corresponding to a solvent content of 58 %. Initial attempts to solve the structure of Uba1 by molecular replacement utilizing truncated forms of NEDD8-E1 and full-length monomeric and dimeric forms of *E. coli* MoeB failed. However, a search model consisting of a truncated form of the SUMO-E1 (see below) allowed the location of the equivalent part in Uba1. In detail, the Uba1 structure was solved by sequential molecular replacement using two separate models. First, the adenylation domain was positioned correctly by MOLREP using the two corresponding domains from the Sae1-Sae2-Mg-ATP complex where the inactive adenylation domain consisting of Sae1's residues 10-178 and 233-345 and its active counterpart of Sae2's residues 7-210 (PDB entry 1Y8Q [29]) using data between 50 and 4.0 Å resolution. Subsequently, the catalytic cysteine domain was located using the homologous domain (residues 629-889) of the mouse ubiquitin-activating enzyme (PDB entry 1Z7L [85]) with PHASER [86] utilizing data between 15.0 and 3.5 Å resolution. Together these search models correspond to approximately 60% of the entire sequence. At this stage, twofold non-crystallographic symmetry (NCS) averaging with masks generated around individual domains was performed at 2.7 Å resolution and residues 12-782 and 797-904 of Uba1 were built in these maps, which represent about 90% of the Uba1 protein, but did neither include the C-terminal UFD nor the bound ubiquitin. For the generation of the masks, the missing domains were included on the basis of the corresponding SUMO-E1 parts.

The maps were subsequently improved through iterative cycles of refinement against the 2.7 Å amplitudes and SIGMAA-weighted phase combination using CCP4. Ubiquitin (PDB entry 1UBQ [27]) was located manually by fitting it into a region of significant difference density located near the adenylation domain. At the later stages of

the refinement, inspection of the SIGMAA weighted $2F_O$-F_C electron density maps revealed that the two copies of Uba1 have conformational differences within the C-terminal residues 906-1024 (UFD). At this point tight NCS restraints between the two molecules were relaxed to allow independent refinement of the two molecules. All model building was carried out using O [84] followed by refinement using both CNS [87] and REFMAC5 [88]. The last few cycles of refinement were carried out with TLS restraints as implemented in REFMAC5. The TLS bodies were defined according to the domain architecture of Uba1 (TLS group1: residues 12-177 and 263-426 (Uba1), TLS group2: residues 178-262 (Uba1), TLS group3: residues 427-596 and 862-916 (Uba1), TLS group4: residues 597-860 (Uba1), TLS group5: residues 920-1024 (Uba1), TLS group6: residues 1-76 (ubiquitin)). The refined model contains residues 12-785 and 794-1024 of Uba1 and all residues (1-76) of ubiquitin for complex A, and 11-646, 650-787 and 797-1024 of Uba1 as well as 1-76 of ubiquitin for complex B. Residues 786-793 of Uba1 in complex A and 647-649 and 788-796 in complex B are not visible in the electron density maps and are presumably disordered.

Table 2.3 Data and Refinement Statistics

Data Collection Statistics	
Space group $P2_12_12_1$ Unit cell dimensions (Å): a = 115.36, b = 118.56, c = 207.57	
Molecules/asymmetric unit	2
X-ray source	NSLS, beamline X29
Wavelength (Å)	1.1000
Resolution limits (Å)	50 - 2.7
No. observations	503,923
No. unique observations	77,706
Completeness (%) (last shell)	98.5 (86.9)
R_{merge}[a] (%) (last shell)	11.5 (41.2)
I/σI	15.9 (1.7)
Mean redundancy	6.5

Refinement Statistics	
Resolution (Å)	20 – 2.7
No. protein/solvent atoms	17,034/184
R (R_{free}[a]) (%)	19.4 (24.6)
Rms bond length deviations (Å)	0.013
Rms bond angle deviations (°)	1.392
Rms chiral volume deviations ($Å^3$)	0.099
Rms planar group deviations (Å)	0.004
Average B-factors (Protein/solvent atoms, $Å^2$)	38.4/33.8
Ramachandran statistics[a] (%)	95.0/4.1/0.9

[a]$R_{merge} = \sum_{hkl}\sum_{k}|I(k) - [I]|/\sum_{hkl}\sum_{k}I(k)$, where $I(k)$ is the value of the k^{th} measurement of the intensity of a reflection, $[I]$ is the mean value of the intensity of that reflection, and the summation is of all the measurements. Brackets denote the highest resolution shell (2.8-2.7 Å). $R = \sum_{hkl}|F_{obs} - F_{cal}|/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure factors, respectively, for all data (no σ cutoff). R_{free} = R calculated with 5% of the reflection data chosen randomly and omitted from the start of refinement. Ramachandran statistics have been determined with MolProbity [89] and refer to the percentage of residues in the core/allowed/disallowed regions of the Ramachandran diagram.

C. Quality of the model

The working and free R factors are 0.194 and 0.246, respectively for all data between 20 and 2.7 Å resolution. A representative portion of a SIGMAA weighted $2F_O$-F_C electron density map is shown in Figure 2.4. The model possesses an excellent overall stereochemistry with 95.0% of all residues in favored regions and 4.1% in the additionally allowed regions of the Ramachandran diagram based on an analysis with MolProbity [89].

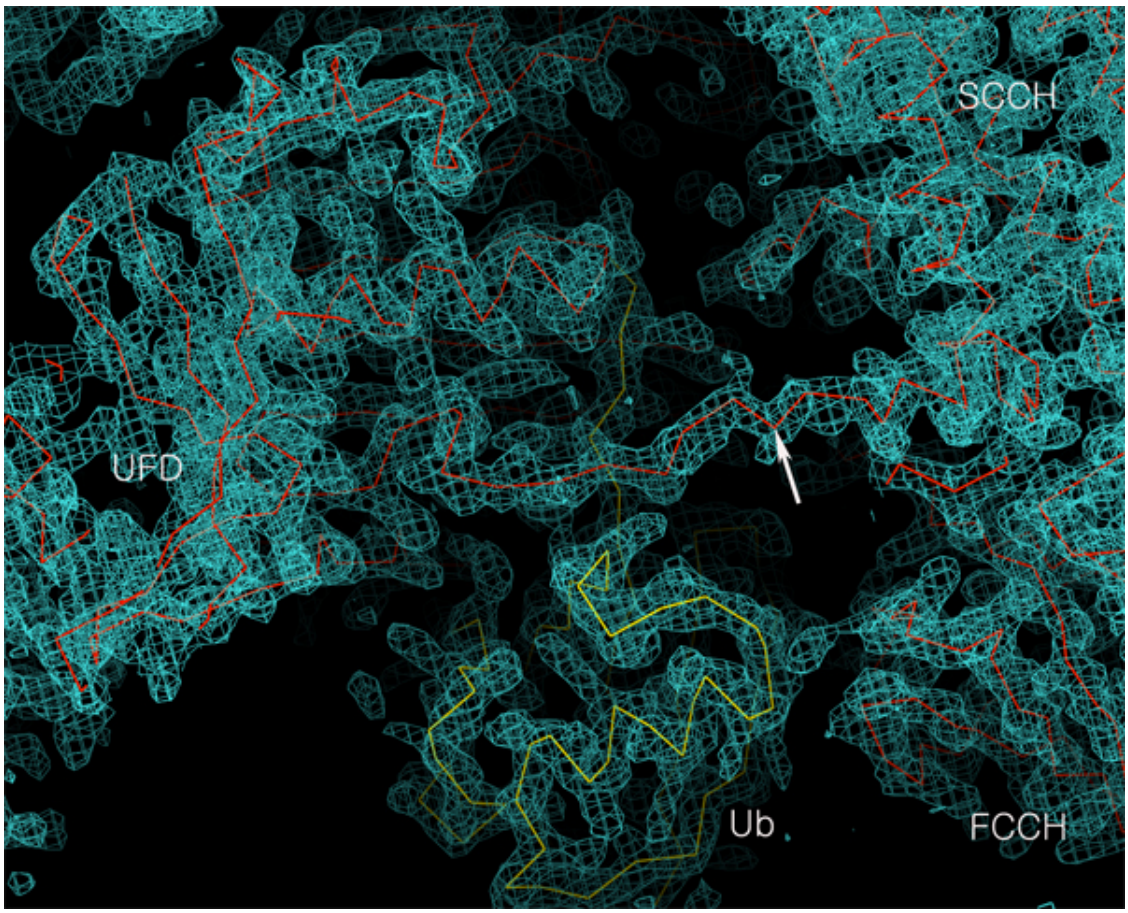


Figure 2.4 SIGMAA weighted $2F_O$-F_C electron density map (blue mesh) contoured at 1σ showing a region of the Uba1-ubiquitin complex. Uba1 (UFD on the left and catalytic cysteine domain (FCCH and SCCH) on the right) and ubiquitin backbones are colored in red and yellow, respectively. The white arrow points to the crossover loop.

D. Overall Structure of the Uba1-ubiquitin Complex

Although the crystals only gave useful diffraction to 2.7 Å, virtually all residues are clearly defined and there is no significant disorder in the protein itself. There are two molecules of the 120 kDa Uba1-ubiquitin complex in each asymmetric unit and they exhibit an almost identical set of structural features except in the C-terminal UFD with a root-mean-square deviation (rmsd) of 1.08 Å over 933 (out of a total of 1024) aligned Cα atoms. Residues excluded from the superposition are located in the UFD. These residues can be separately superimposed resulting in a rmsd of 0.60 Å over 96 aligned Cα atoms, while the ubiquitin molecules (residues 1-76) can be superimposed with a rmsd of 0.35 Å.

Uba1 consists of a complex arrangement of six structural domains referred to as IAD, AAD, FCCH, SCCH, 4HB and UFD as defined in the introduction, with overall dimensions of 85 Å x 90 Å x 60 Å (Figure 2.5). In the complex structure, four Uba1 domains, AAD, FCCH, SCCH and UFD, pack together to generate a large central canyon (~40 Å wide), which recruits a ubiquitin molecule snugly in a manner resembling a baseball in a mitt. The large size of the canyon in the Uba1 suggests it may function to accommodate an E2 as well.

Three Uba1 domains, UFD, FCCH and SCCH, are linked with the respective adjacent domains in the Uba1 sequence by flexible linkers. UFD is linked to AAD by an 18-residue loop forming a β-hairpin at the end of the AAD (UFD linker). FCCH is linked to IAD by two long antiparallel β-strands, whereas SCCH is linked to AAD by an extended 18-residue linker that traverses 40 Å from one side of the molecule to the other (crossover loop). This crossover loop divides the canyon into two clefts that are

continuous both below and above the loop. As in the previous E1 structures [29, 61], when viewed facing the E1 catalytic cysteine located centrally above AAD, ubiquitin's globular domain binds in the right cleft (cleft 2) with its C-terminal tail extending under the crossover loop to approach the adenylation active site in the left cleft (cleft 1) (Figure 2.5). The above described structural features suggest that Uba1 may undergo large-scale conformational changes during the course of its functional cycle, an idea that is supported by other structural details that will be described later.

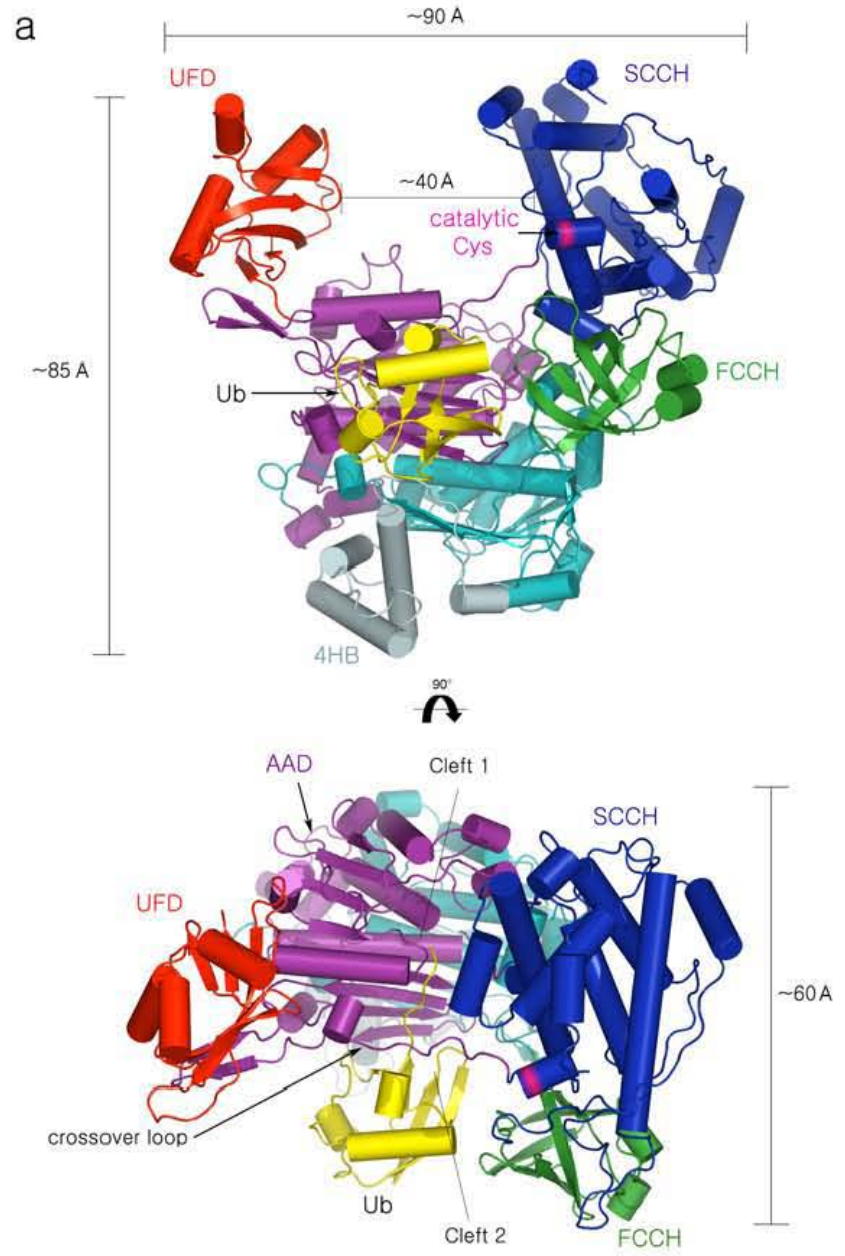

a
~90 A
UFD
SCCH
~40 A
catalytic
Cys
~85 A
Ub
FCCH
4HB
90°
AAD
Cleft 1
SCCH
UFD
~60 A
crossover loop
Ub
Cleft 2
FCCH
b
FCCH
SCCH
Uba1
1 169 268 356 404 594 C600 860 928 1024
A CC 4HB A A CC A UFD
APPBP1/UBA3
1 168 397 487 534 1 210 C216 289 347 442
A CC 4HB A A CC A UFD
Sae1/Sae2
1 171 214 296 346 1 167 C173 384 441 550
A CC 4HB A A CC A UFD
MoeB-MoeB
1 178 190 238 1 178 190 238
A A A A
IAD
AAD

Figure 2.5 Overall structure of the Uba1-ubiquitin complex. (a) Cartoon diagram of the front and top views of the Uba1-ubiquitin (yellow) complex with α-helices as cylinders and β-strands as curved arrows. The six domains of Uba1 are colored as follows: Adenylation (A) domains in cyan (IAD) and purple (AAD), the two subdomains of the catalytic cysteine (CC) domain in forest (FCCH) and blue (SCCH), the ubiquitin-fold domain (UFD) in red and the four helix bundle (4HB) domain in pale cyan. The catalytic cysteine is shown in pink. The overall size of the complex is about ~85 Å by ~90 Å by ~60 Å and there is a ~40 Å gap between the UFD and SCCH domains. (b) Domain architectures of Uba1, APPBP1/UBA3 (NEDD8-E1), Sae1/Sae2 (SUMO-E1), and the MoeB dimer colored according to the Uba1 structure.

The AAD ("Active" Adenylation Domain)

The AAD consists of residues 405-596 and 860-927 and can be superimposed with the MoeB monomer with a rmsd of 1.23 Å for 195 aligned Cα atoms (out of 249 residues in MoeB). This domain consists of eight β-strands that form a continuous β-sheet surrounded by eight α-helices. In the N-terminal half of the domain, four β-strands are all parallel and shows a variation of the Rossmann fold. 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 2.6). The first of these 3_{10} helices (η7) contains five residues with the sequence Ser-Asn-Leu-Asn-Arg (residues 477 to 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 Gly-X-Gly-X-X-Gly, 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 as will be discussed later.

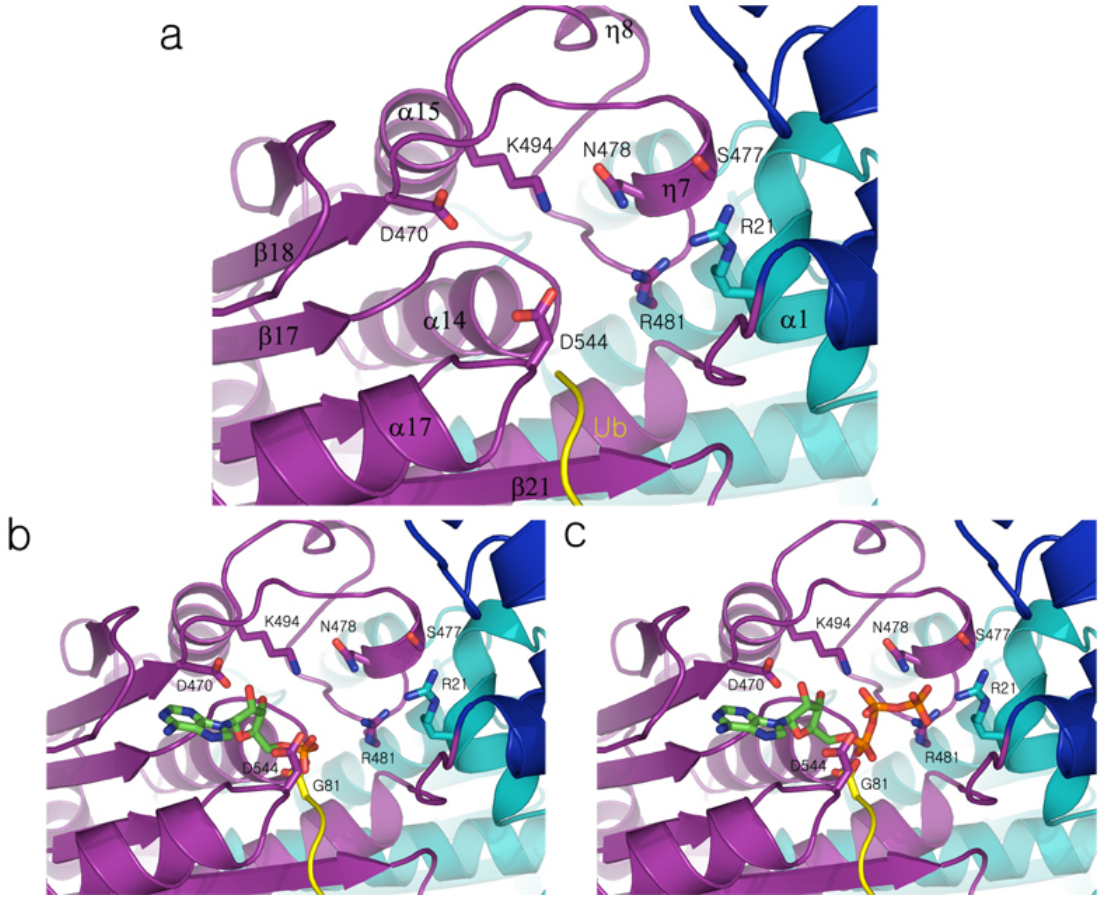


Figure 2.6 The adenylation active site of Uba1. (a) The apo-structure as observed in the crystal structure, (b) the adenylation model and (c) the ATP complex model. Ubiquitin's C-terminal tail is shown in yellow and conserved residues are shown in stick representation. The models were generated by superposition of MoeB-MoaD~adenylate (PDB entry 1JWB [31]) and MoeB-MoaD-ATP complex (PDB entry 1JWA [31]) on the AAD of Uba1.

The area of strongest structural conservation between the AADs of Uba1, UBA3, Sae2 and MoeB is the ATP-binding region (Figure 2.6). All side chains that interact directly with ATP are identical among these proteins. These residues include Asp470, Ser477, Asn478, Arg481, Lys494, and Asp544 from this domain. They are in nearly identical positions compared to structures of MoeB-MoaD complexes [31]. The high level of conservation in the ATP-binding sites of E1s and MoeB likely reflects stringent requirements for the catalytic mechanism of the adenylation reaction.

The crossover loop starts near the end of this domain (residues 582-596) and passes over the adenylation active site, leading to Cys600 and the site of thioester bond formation. In SUMO-E1, NEDD8-E1, and MoeB, two Cys-X-X-Cys motifs are found in this region, which are responsible for coordinating a zinc atom through their thiolates. This zinc-binding site is quite distant from the adenylation active site, thus suggesting a structural rather than a catalytic role for the metal. However, in Uba1 and all other ubiquitin-E1s, the first Cys-X-X-Cys motif has been replaced by Ser-X-X-Ser, while the second motif does not exist at all, thus strongly suggesting that ubiquitin-E1s do not bind a corresponding zinc ion at this site (Figure 2.28). Implications that arise from this difference will be discussed later.

The IAD ("Inactive" Adenylation Domain)

The IAD consists of residues 1-169 and 357-404 and can be superimposed with the MoeB monomer with a rmsd of 1.60 Å for 202 aligned Cα atoms (out of a total of 249). It exhibits an identical set of structural features as the AAD (and MoeB), except that it lacks the Gly-X-Gly-X-X-Gly 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 (Figure 2.7). However, the IAD contributes the conserved Arg21 to the adenylation active site located in the AAD, just as the second monomer in the MoeB homodimer contributes the equivalent residue (Arg14) to the ATP binding site (Figure 2.6). Arg21 from the IAD and Arg481 from the AAD extend into the ATP binding site and presumably form a salt-bridge with the γ-phosphate group and more importantly stabilize the developing negative charge on the β-phosphate

during hydrolysis of the α-β phosphodiester bond which accompanies formation of the ubiquitin~adenylate. The 4HB of Uba1 is equivalent to a short four-helix bundle domain in APPBP1 and Sae1 and reveals a slightly different topology compared to them (Figure 2.8).

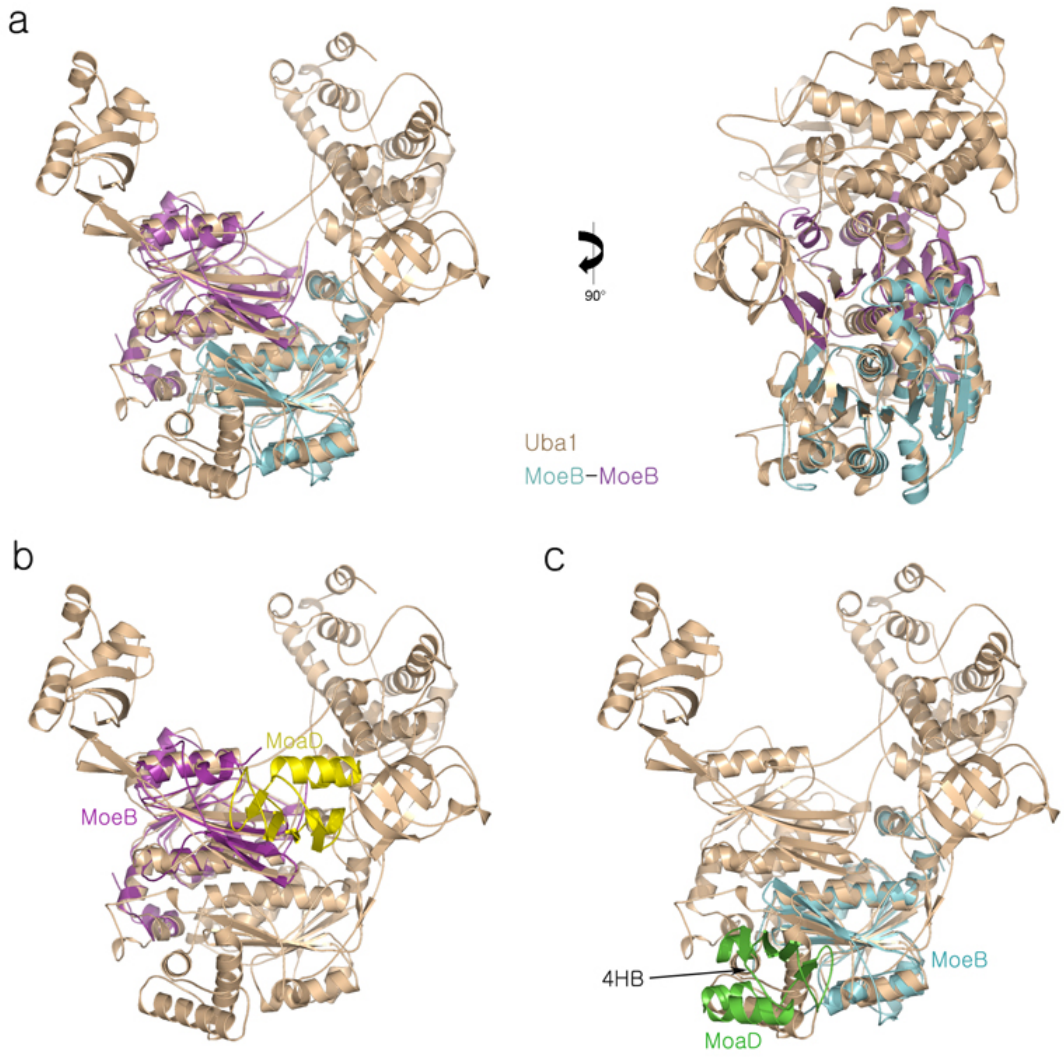


Figure 2.7 Structural evidence for a single adenylation active site in Uba1. (a) Ribbon representation of Uba1 (wheat) and the MoeB homodimer (cyan and purple) superimposed via the two adenylation domains. (b) One MoeB monomer (purple) and its associated MoaD (yellow). MoaD is located in a comparable position as ubiquitin in the Uba1-ubiquitin complex. (c) The second MoeB monomer (cyan) and its associated MoaD (green). The "inactive" adenylation domain lacks the nucleotide binding motif, and the conserved four helix bundle precludes ubiquitin-binding.

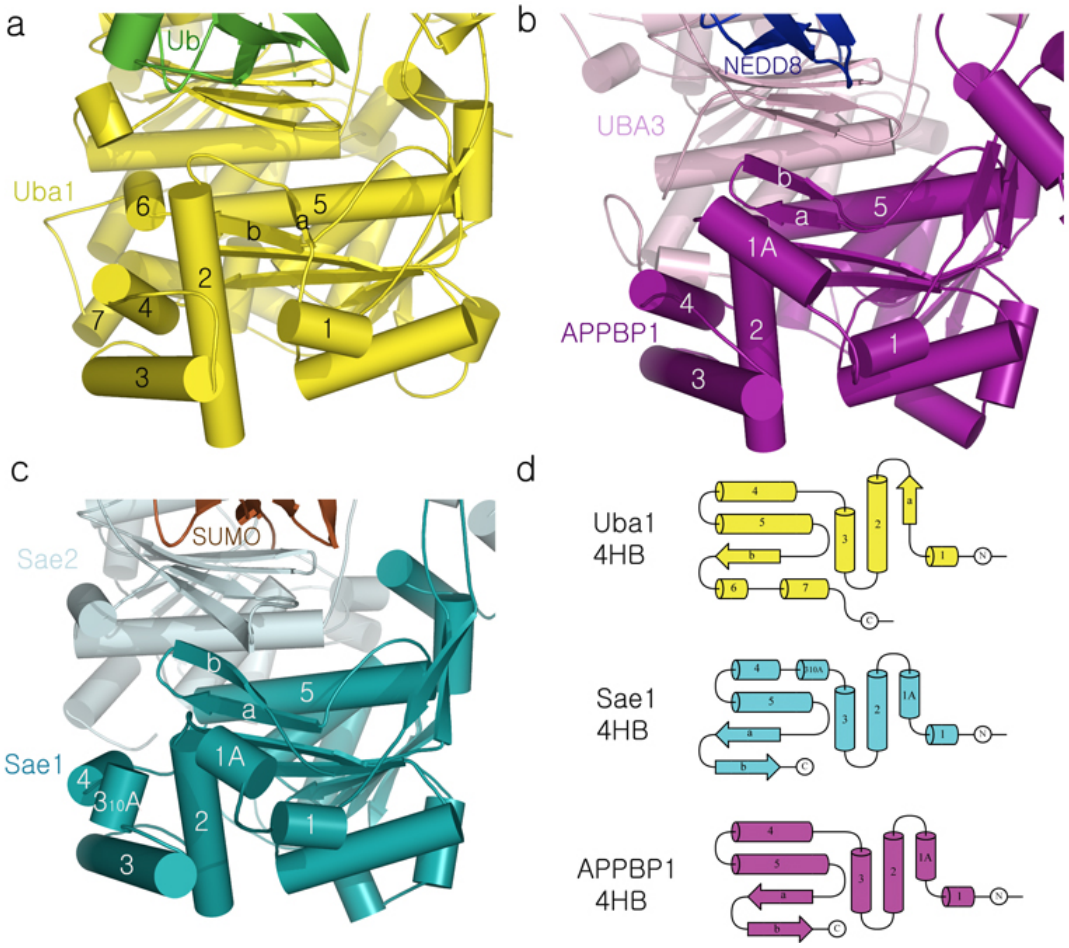


Figure 2.8 Comparison of the 4HB domains. Close-up view of the 4HBs of (a) Uba1, (b) APPBP1 [80] (PDB entry 1R4N), and (c) Sae1 [29] (PDB entry 1Y8R). All ribbon representations correspond to the same orientation. (d) Secondary structure diagram for each 4HB domain is shown in the same color code. The 4HB of Uba1 displays the same overall architecture as the other 4HBs, although some of the secondary structure elements are not conserved (conserved secondary α-helices have been numbered in the same way).

The FCCH (First Catalytic Cysteine Half-domain)

This domain consists of residues from 175 to 265 of Uba1 and this portion of E1 differs significantly between the E1s for different Ubls. For SUMO-E1, the FCCH consists of only two antiparallel β-strands and a disordered region (residues 180-202 of Sae1). It is larger in the ubiquitin-E1 (~90 residues), but even much larger in the

NEDD8-E1 (~230 residues), where this half-domain is almost entirely helical. By contrast, the FCCH of Uba1 is about the same size of the globular body of ubiquitin and is essentially an all β-structure with three pairs of anti-parallel β-sheets forming a barrel (Figure 2.9). The closest structural homologues, identified using DALI (http://www.ebi.ac.uk/dali), are a domain from the large subunit of initiation factor eIF2 from *Pyrococcus abyssi* (Z = 6.0; PDB entry 1KK1) and a domain from elongation factor EF-Tu from *E. coli* (Z = 6.0; PDB entry 1EFC). The FCCH does not contain the catalytic cysteine residue, and its role in various Ubl-E1 enzymes has not yet been determined. However, the fact that the domain is connected to the rest of the enzyme by two long flexible linkers suggests it may undergo conformational changes during translocation of the adenylated C-terminus of ubiquitin to the E1 active site cysteine for E1~ubiquitin thioester formation (Figure 2.10).

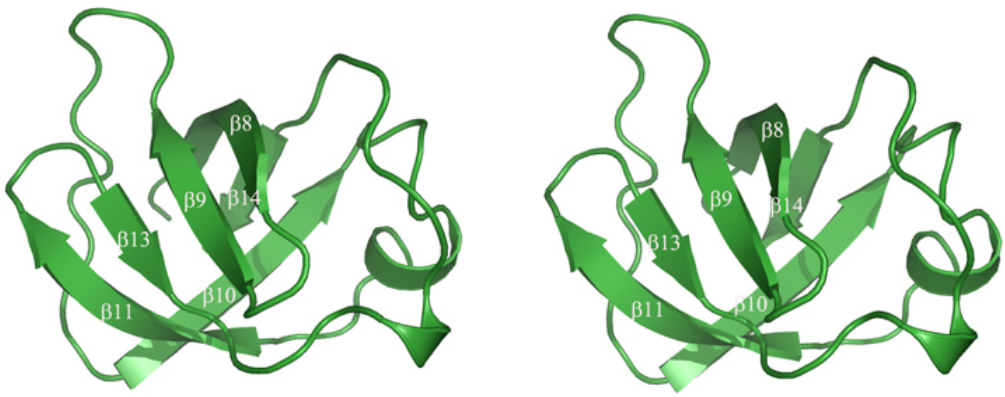


Figure 2.9 Stereo view of the FCCH of Uba1. This domain exhibits a 6-stranded β-barrel structure.

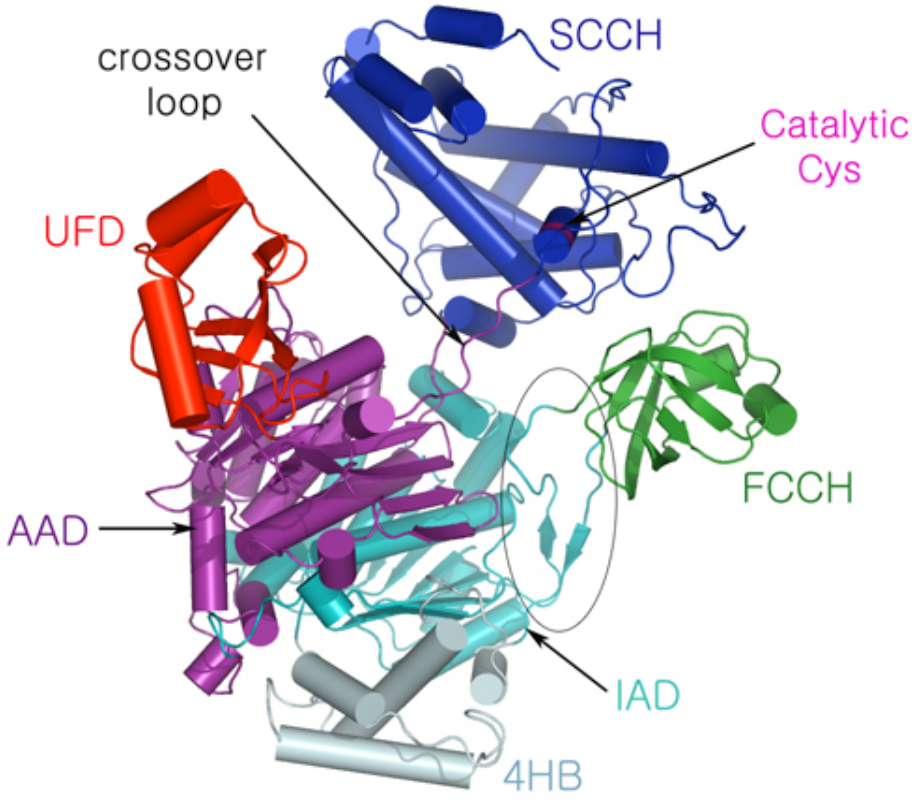


Figure 2.10 The FCCH of Uba1 is connected by two extended linkers to the IAD, which are highlighted by the black oval. Domains are colored according to Figure 2.5.

The SCCH (Second Catalytic Cysteine Half-domain)

The SCCH is built around a short core motif (~80 residues), which is present in all Ubl-E1 proteins and includes the catalytic cysteine residue. In NEDD8-E1, this core region represents the entire SCCH. In SUMO-E1, the SCCH is expanded by an insertion of ~120 residues. An even larger, unrelated insertion is present in all ubiquitin-E1 (Figure 2.11). Superpositions of the SCCH of the ubiquitin-E1 with the equivalent half-domains of the NEDD8-E1 and SUMO-E1 shows considerable homology among them as reflected in the DALI scores of 11.3 and 8.5, which translate into a rmsd of 1.8 Å and 1.3 Å for 72 and 74 aligned Cα atoms, respectively. The function of the insertion in the SCCH is

presently unclear, however, in the ubiquitin-E1, it may be involved in interactions with the incoming E2s, thus facilitating E1-E2 transthioesterification (Figure 2.12). This will be discussed in more detail later.

The SCCH of Uba1 is predominantly helical and the shape of the domain can be described as a distorted "U" with a large, central cleft in the middle. The cleft is bridged by a long and poorly structured region that lacks electron density for 8 residues (786-793). The topology of the SCCH is rather complex. Neither of the two "arms" of the "U" is built up from an uninterrupted stretch of amino acids. The rather complicated fold places the N- and C-terminal ends of the half-domain in close proximity. The active site cysteine is located near the N terminus of the domain, just upstream of a very short α-helix (α18). The SCCH appears perched over the rest of the molecule, burying a relatively small interface with the FCCH (500 $Å^2$ of total buried surface area) and with the adenylation domains (800 $Å^2$ of total buried surface area).

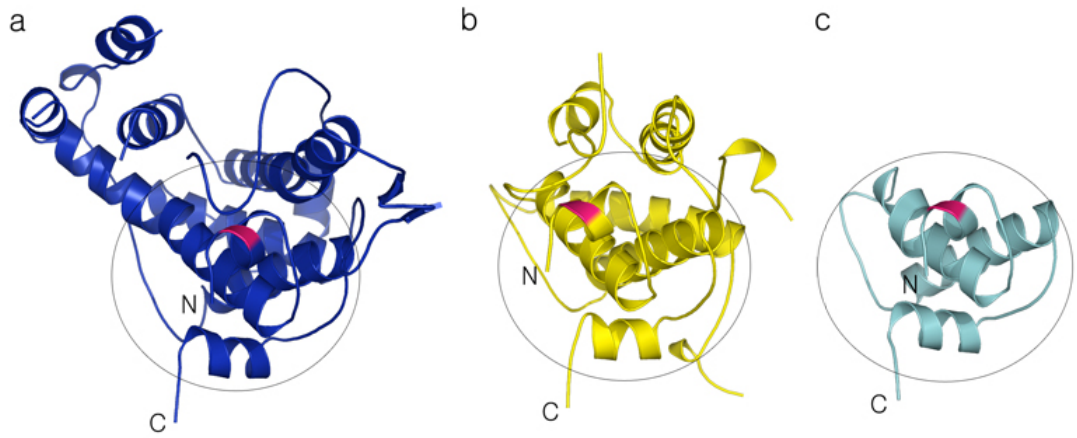


Figure 2.11 Comparison of the structures of the SCCHs of (a) ubiquitin-E1, (b) SUMO-E1, and (c) NEDD8-E1. The black ovals indicate that the NEDD8-E1 SCCH represents the common core of the fold. The coordinates for SUMO-E1 and NEDD8-E1 were taken from PDB entries 1Y8R [29] and 1R4N [80], respectively. The location of the catalytic cysteine is highlighted in pink, and the N- and C-termini are labeled.

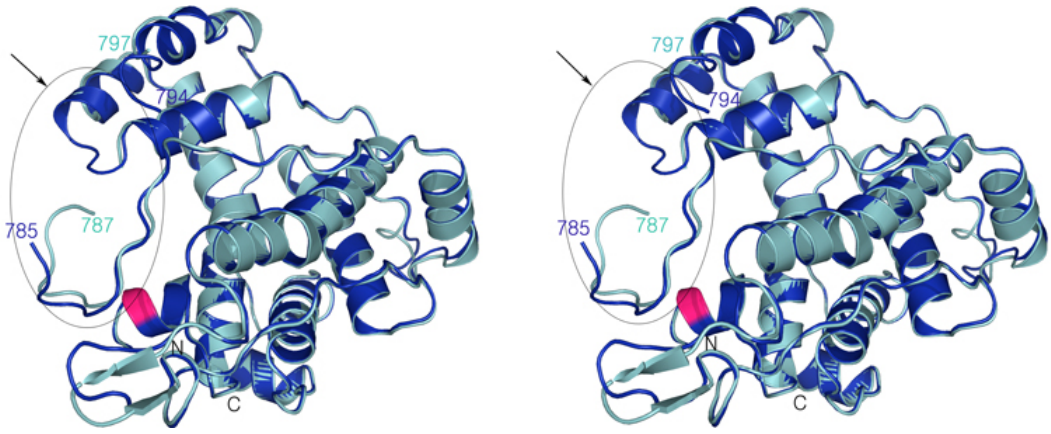


Figure 2.12 Stereo view of a superposition of the two SCCH domains present in the asymmetric unit. The black oval indicates the presumed area where the SCCH might contact an incoming E2 during transthioesterification. The catalytic cyteines are again highlighted in pink, and the N- and C-termini are labeled.

The UFD (Ubiquitin-fold Domain)

The UFD consists of residues 928-1024 of Uba1. The domain superimposes with the structure of ubiquitin (rmsd of 3.1 Å for 61 Cα atoms out of a total of 76 Cα atoms in ubiquitin) and adopts an α/β structure. However, the Uba1 UFD structure reveals that it is topologically different from ubiquitin. The two major differences between the overall structures of the Uba1 UFD and ubiquitin are: (1) While the C-terminal antiparallel β-strands (β30 and β32) correspond to the C-terminal strands (β3 and β5) in ubiquitin, they run in opposite directions. (2) Two α helices (α32, α33) and extended loops are inserted between β29 and β30 of the Uba1 UFD (Figure 2.13). As mentioned previously, UFD is connected to the rest of the enzyme by an extened 18-residue linker (UFD linker), which spans a distance of ~19 Å (Figure 2.14).

The UFDs from ubiquitin-, SUMO-, and NEDD8-E1 enzymes are structurally similar (Figure 2.22a). The Uba1 UFD can be superimposed with the Sae2 UFD (rmsd of 2.2 Å for 74 Cα atoms) and with the UBA3 UFD (rmsd of 2.0 Å for 65 Cα atoms). Other

structural homologues, identified using DALI (http://www.ebi.ac.uk/dali), are the Elongin B component of the SOCS E3s (Z = 3.9; PDB entry 1VCB), and a domain from the tubulin-binding cofactor B (Z = 3.7; PDB entry 1T0Y).

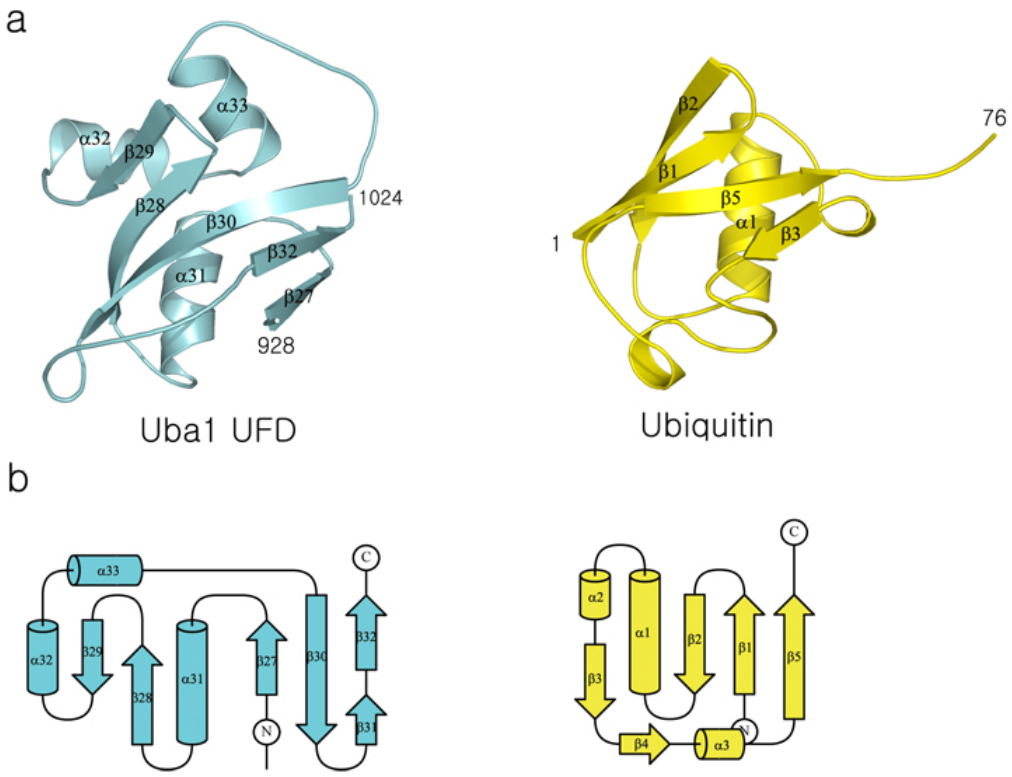


Figure 2.13 Comparison between Uba1 UFD and ubiquitin. (a) Both share the common β-grasp fold. N- and C-terminal residues are numbered. (b) A topology diagram for each is shown.

Previous studies indicate that the C-terminal UFD of NEDD8- and SUMO-E1 are involved in the recruitment of their cognate E2s and the ensuing transthioesterification reaction in which the Ubl is transferred from the E1 active site cysteine to the catalytic cysteine of the E2 [29, 35] The recent discovery of a second ubiquitin-E1 (Uba6) in vertebrates and sea urchin showed that the UFD provides the E1 with the capability of

interacting and charging different sets of E2, thus contributing to selectivity in E2 recruitment [32].

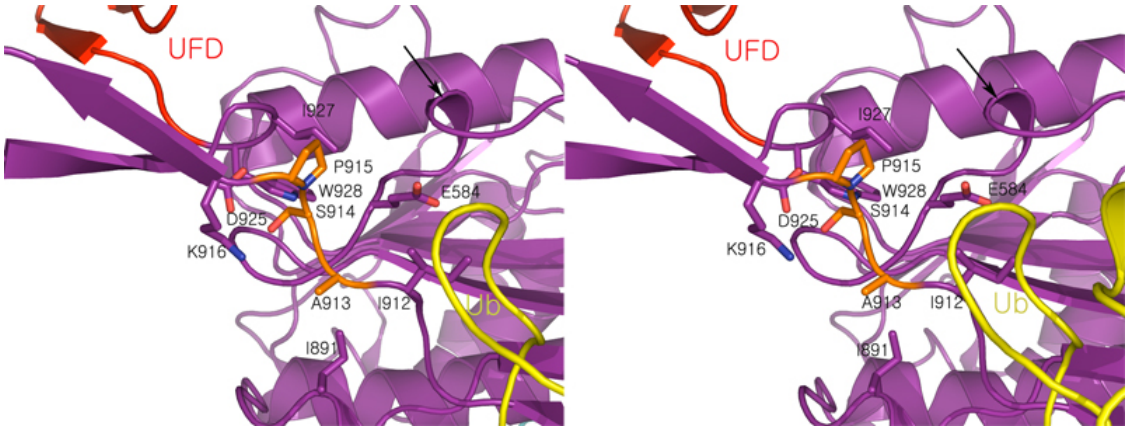


Figure 2.14 Stereo view showing the area of the UFD linker. The suspected hinge region is colored in orange and adjacent residues are shown in stick representation. The black arrow points to the crossover loop.

E. Model of how Uba1 adenylates ubiqutin

The two adenylation domains of Uba1 contain the repeat sequence found in the N- and C-terminal halves of all E1, which also represents the region of sequence homology found in MoeB. As discussed above, these structures consist of a mixed eight-stranded β-sheet surrounded by eight α-helices, and also include regions where sequence homology is weak. Both domains contain the two subdomains found in MoeB, with the N-terminal half of each adopting a variation of Rossmann fold βαβαβ topology. The two adenylation domains pack together in the same way as the two MoeB molecules within the MoeB homodimer (Figure 2.7).

The MoeB homodimer contains a perfect twofold symmetry, so each molecule of MoeB binds to one molecule of MoaD, forming a heterotetrameric complex. Thus, two molecules of MoaD can be adenylated simultaneously by the MoeB homodimeric

complex. In contrast, Uba1 and other E1s appear to have evolved directionality in the reaction by restricting the activity to a single adenylation active site in cleft 1, composed primarily of the C-terminal half of the E1. It contains the glycine-rich nucleotide binding motif corresponding to the ATP binding site in MoeB. By contrast, the IAD lacks the Gly-X-Gly-X-X-Gly ATP binding residues in its MoeB-homology region. Also, the second subdomain of the AAD is accessible from cleft 2 thus allowing interactions with ubiquitin, much like the corresponding C-terminal subdomain of MoeB is exposed to interact with MoaD. By contrast, the second subdomain of Uba1's IAD is buried by the four-helix bundle, whose sequence is conserved in the N-terminal portions of other E1s. The presence of the 4HB prevents the interaction with a second molecule of ubiquitin.

Although there is limited sequence homology between MoeB and E1s in cleft 2, the Ubl-binding region, there is significant structural homology in the region where ubiquitin binds in cleft 2. The structure of the adenylation domain of Uba1 was aligned with the MoeB-MoaD~adenylate complex, to allow modeling of the structures of ubiquitin and AMP in the complex. These models (Figure 2.6) suggest that both clefts in the E1 canyon are involved in the formation of the Ubl~adenylate, with residues involved in catalyzing the adenylation reaction in cleft 1, and extensive hydrophobic contacts with Ubl present in cleft 2. Residues important for the activation reaction as shown by the MoeB-MoaD~adenylate structure are conserved in the structure of ubiquitin-E1. These include a hydrophobic patch involving Uba1 Val440, Gly441, Val520, Leu543 and Ala548 that according to the model contact the adenine base, Asp470 and Asp472, which ligate the ribose, and Asp544 which coordinates the Mg^{2+} ion. The 3_{10}-helix (η7) contains Arg481, that would contact the β- and γ-phosphate groups of ATP and just as one

monomer of MoeB contributes a key arginine side chain to coordinate the ATP bound to the opposite monomer, Uba1's Arg21 is in the same position relative to the adenylation active site, thus alleviating the developing negative charge on the β-phosphate during the formation of the ubiquitin~adenylate.

F. Ubiquitin recognition by Uba1: Implications for Ubl-E1 interactions

The Uba1-ubiquitin interactions result in the burial of ~3200 $Å^2$ of exposed surface area corresponding to 33% of ubiquitin's total surface area. Similar values are observed in the case of the NEDD8-E1, which buries ~3350 $Å^2$ of total surface area with NEDD8 (34% of NEDD8) [80]. In contrast, the SUMO-E1-SUMO interface buries only ~1650 $Å^2$ (20% of SUMO) [29]. Most of the interactions involving ubiquitin binding are localized to the four-stranded β sheet (β21-β24) preceding the UFD domain and the crossover loop. The details of the interface between Uba1 and ubiquitin are described by dividing it into three areas: (1) The hydrophobic interface between the "canonical" hydrophobic patch of ubiquitin and the conserved AAD of Uba1, (2) the interactions between ubiquitin's C-terminal tail and the AAD as well as the crossover loop, and (3) the polar interface between ubiquitin and the FCCH (Figure 2.15 and 2.16).

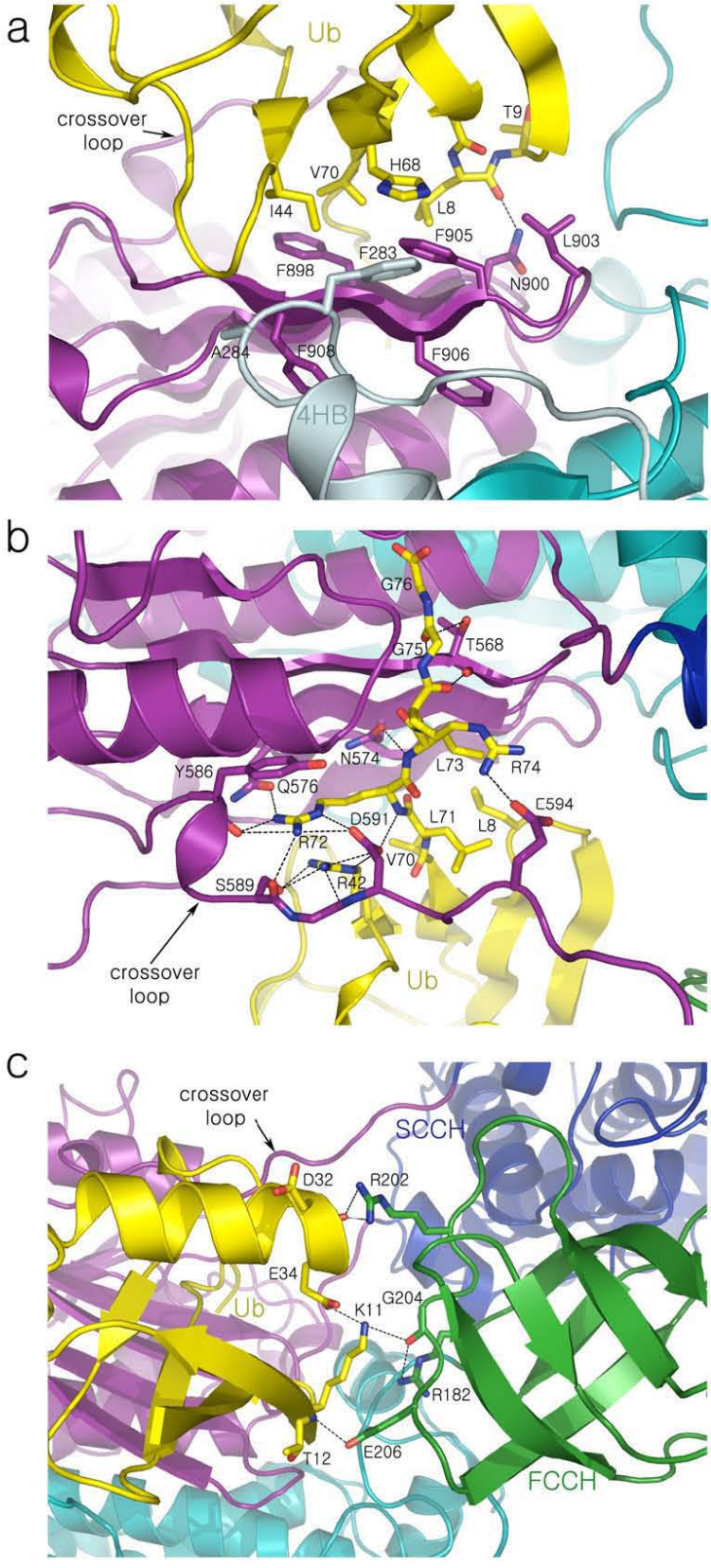

a
Ub
crossover loop
T9
V70
H68
I44
L8
F905
L903
F283
F898
N900
A284
F908
F906
4HB
b
G76
G75
T568
N574
L73
R74
Y586
Q576
E594
D591
L71
L8
R72
V70
S589
R42
crossover loop
Ub
c
crossover loop
SCCH
D32
R202
E34
Ub
K11
G204
R182
T12
E206
FCCH

Figure 2.15 Detailed views of the Uba1-ubiquitin interface. Domains are colored according to Figure 2.5 including the corresponding carbon atoms with nitrogen atoms in blue and oxygen atoms in red. Dashed lines indicate hydrogen bonds. (a) Interface between the hydrophobic surface on ubiquitin and the AAD. (b) Interactions between ubiquitin's C-terminal tail with the AAD and crossover loop. (c) Interface between ubiquitin and the FCCH.

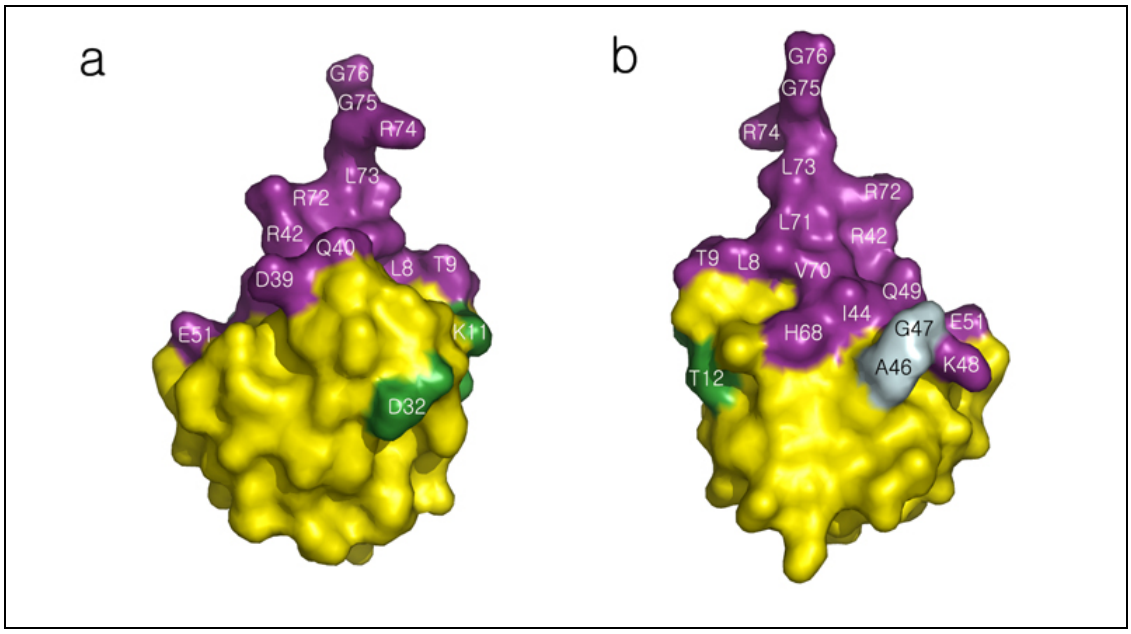


Figure 2.16 Surface regions of ubiquitin which interact with its E1. (a, b) Two views of ubiquitin are shown which differ in a 180° rotation around the vertical axis. Ubiquitin residues contacting the AAD, FCCH, and 4HB are shown in purple, forest and pale cyan, respectively.

The structure indicates extensive hydrophobic contacts between a hydrophobic patch on the surface of the AAD (β23 and β24), involving Phe898, Leu903, and Phe905 and a hydrophobic patch that is absolutely conserved between ubiquitin and NEDD8 including Leu8, Ile44, and Val70 (Figure 2.15a). These residues in ubiquitin are essential for viability in yeast [90]. Mutation of Leu8 to alanine, or Leu8 in combination with Val70 to alanine, reduces ubiquitin conjugate formation by more than 50% compared to wild-type [91]. In addition to hydrophobic contacts, this interaction is stabilized by a

hydrogen bond between Uba1's Asn900 and the carbonyl oxygen of ubiquitin's Leu8. Phe283 and Ala284, which are located in the turn between the first and the second α-helix (α6 and α7) of the 4HB, interact with Ala46 and Gly47 present in the turn between the third and fourth β-strand (β3 and β4) of ubiquitin and support it from underneath.

Besides their essential role in initiating Ubl conjugation cascades, E1s select the correct Ubl for the pathway. Each Ubl has a dedicated E1, which exhibits remarkable specificity. For example, despite the fact that ubiquitin and NEDD8 are nearly 60% identical and have strikingly similar structures, they are distinguished by their respective E1s. Ubiquitin-E1 only activates ubiquitin, and NEDD8-E1 only activates NEDD8. This specificity is crucial because the E1 also transfers the Ubl to its cognate E2, thereby linking the Ubl with its correct downstream pathway. Therefore, an important question is how E1s specifically recognize only their particular Ubl. The structure of the Uba1-ubiquitin complex suggests a rationale for part of the preference of E1s for their particular Ubls. Residue 72 is the only known determinant of selectivity in Ubls for their E1s. Residue 72 is an arginine in ubiquitin and an alanine in NEDD8 (Figure 2.33). In SUMO family members, the corresponding residue is either a glutamate or a glutamine. Evidence that residue 72 is the key specificity determinant comes from the following observations: (1) Mutation of ubiquitin Arg72 to leucine reduces formation of the ubiquitin~adenylate by ~1,000-fold [67]. (2) The Ala72 to arginine mutation in NEDD8 allows it to be activated at a similar rate as ubiquitin by the E1 for ubiquitin [28]. (3) The Arg72 to leucine mutation in ubiquitin allows it to be activated by the E1 for NEDD8 [65].

In the model, ubiquitin's C-terminal tail extends away from the hydrophobic patch, and sits in a shallow groove in the AAD, underneath the crossover loop. At the beginning of the tail, ubiquitin's Leu71 interacts with Uba1's Pro592 in the crossover loop. Next, Uba1's Gln576 from β22 and Tyr586, Ser589, and Asp591 from the crossover loop create a pocket that holds the Arg72 side chain of ubiquitin by forming several hydrogen bonds and a salt bridge (Figure 2.15b). As expected, the Arg72Ala mutant of ubiquitin decreases the Ubc1~ubiquitin thioester formation significantly, in agreement with the structure (Figure 2.17). By contrast, in the NEDD8-E1 complex, UBA3's hydrophobic Leu206 and Tyr207 interact with Ala72 of NEDD8 and in SUMO-E1, Sae2's Arg119 and Tyr159 contact Glu93 of SUMO-1. Notably, in the NEDD8-E1 structure, UBA3 cannot tolerate an arginine at NEDD8's position 72 because of repulsion from UBA3's Arg190, thus preventing the misactivation of ubiquitin by NEDD8-E1 [80]. An additional salt bridge is found between Uba1's Glu594 and Arg74 of ubiquitin, which has also been shown to play a critical role in ubiquitin activation; mutation of Arg74 to leucine reduces the rate of ATP:PPi exchange [67].

At the end of the tail, ubiquitin's C-terminus is inserted into the deep ATP binding pocket in Uba1. Proper positioning of the C-terminal Gly-Gly dipeptide appears to be accomplished by hydrogen bonded interactions. The carbonyl oxygen of Gly75 of ubiquitin forms hydrogen bonds with Oγ1 and the amide nitrogen of Thr568, which is conserved in all E1 family enzymes. The C-terminus of ubiquitin extends over β21 of Uba1, which acts as a structural scaffold. Sequence alignments using E1 enzymes show a preference for small amino acids (Gly or Ala) at the center of β21, which appear to allow the insertion of the Gly-Gly motif of the Ubls into the adenylation active site.

Finally, one face of ubiquitin's globular domain interacts with the FCCH (Figure 2.15c). In this interface, ubiquitin contacts the subdomain comprised of Uba1's residues 175-265 that forms a wall of the broad, deep groove in the Uba1 structure. This portion of the interface is unique to eukaryotic E1s and is not found in distal bacterial ancestors such as MoeB. The nature of this interface is predominantly polar, with four residues (Lys11, Thr12, Gln31, and Asp32) from ubiquitin forming hydrogen bonds with three residues (Arg202, Gly204, and Glu206) from Uba1 also burying ~560 $Å^2$ of surface area. In case of NEDD8-E1, this polar interface is more extensive, involving the whole acidic face of the helix α1 and subsequent loop [80]. In SUMO-E1, this interface appears to be absent.

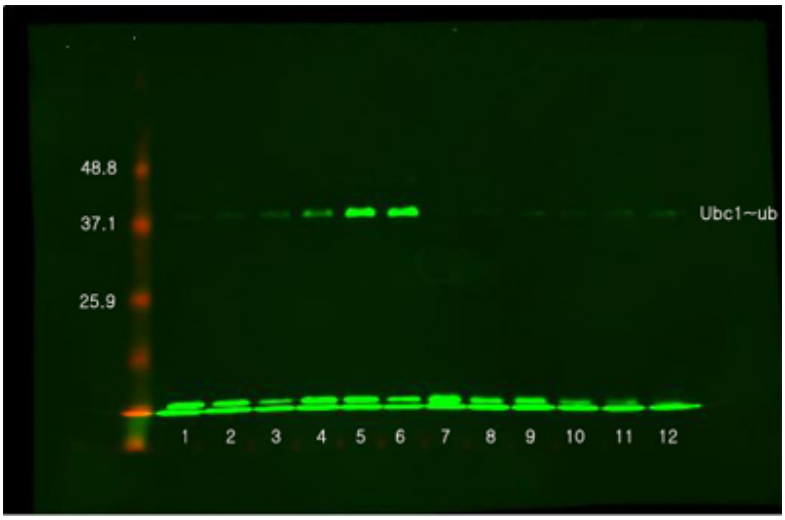


Figure 2.17 Arginine 72 of ubiquitin is a critical specificity determinant for Uba1 binding. Effects of the R72A substitution in ubiquitin on forming the covalent Ubc1~Ub complex was monitored by following the time-course for forming the reducible complex of Ubc1 with ubiquitin. Proteins were resolved by SDS-PAGE, and Ubc1~Ub was visualized by Western blot. Lanes 1 - 6 represent Ub wild-type after 1 – 6 minutes and lanes 7 - 12, the Ub R72A variant using the same time intervals.

G. Structural insights into formation of the E1~Ubl thioester intermediate

The SCCH contains the E1 catalytic cysteine that forms a thioester with the C-terminus of ubiquitin and promotes transfer of ubiquitin to its E2. Formation of the thioester complex between Uba1 and ubiquitin likely proceeds through a nucleophilic attack on the ubiquitin~adenylate by the E1 active site cysteine. This reaction likely involves deprotonation of E1's active site cysteine by a general base catalyst. It is unclear whether the catalytic cysteine Cys600 requires assistance from accessory catalytic residues. The basic residue that is closest (less than 5 Å) to the active site cysteine in yeast Uba1 is Arg603, and this residue is conserved as a lysine among ubiquitin-E1s from different species. However, Arg603 is unlikely to act as a general base, both chemically and structurally. Asp782 is conserved as negatively charged residue in all ubiquitin-E1s (Figure 2.35), and its potential role as a general base was analyzed. Mutation of Asp782 to alanine or asparagine has little effect on the Uba1~ubiquitin thioester formation, rendering such a model improbable (data not shown).

To identify conserved residues, all available SCCH structures were superimposed (Figure 2.18). Only two basic residues, His611 and Arg841, are present in all 3 structures, and both of them seem to be too far away for a direct involvement in cysteine deprotonation. Moreover, His611 is separated from the active site cysteine by Thr601, a strictly conserved residue, which has been shown to be important for the function of the NEDD8-E1 [61, 92]. In the present model, the cysteine sulfur points away from the threonine, but this rotamer position could easily change during catalysis. The apparent lack of a convincing general base residue near the active site cysteine does not preclude a

possibility of realignment of either of these residues through a conformational change during the reaction.

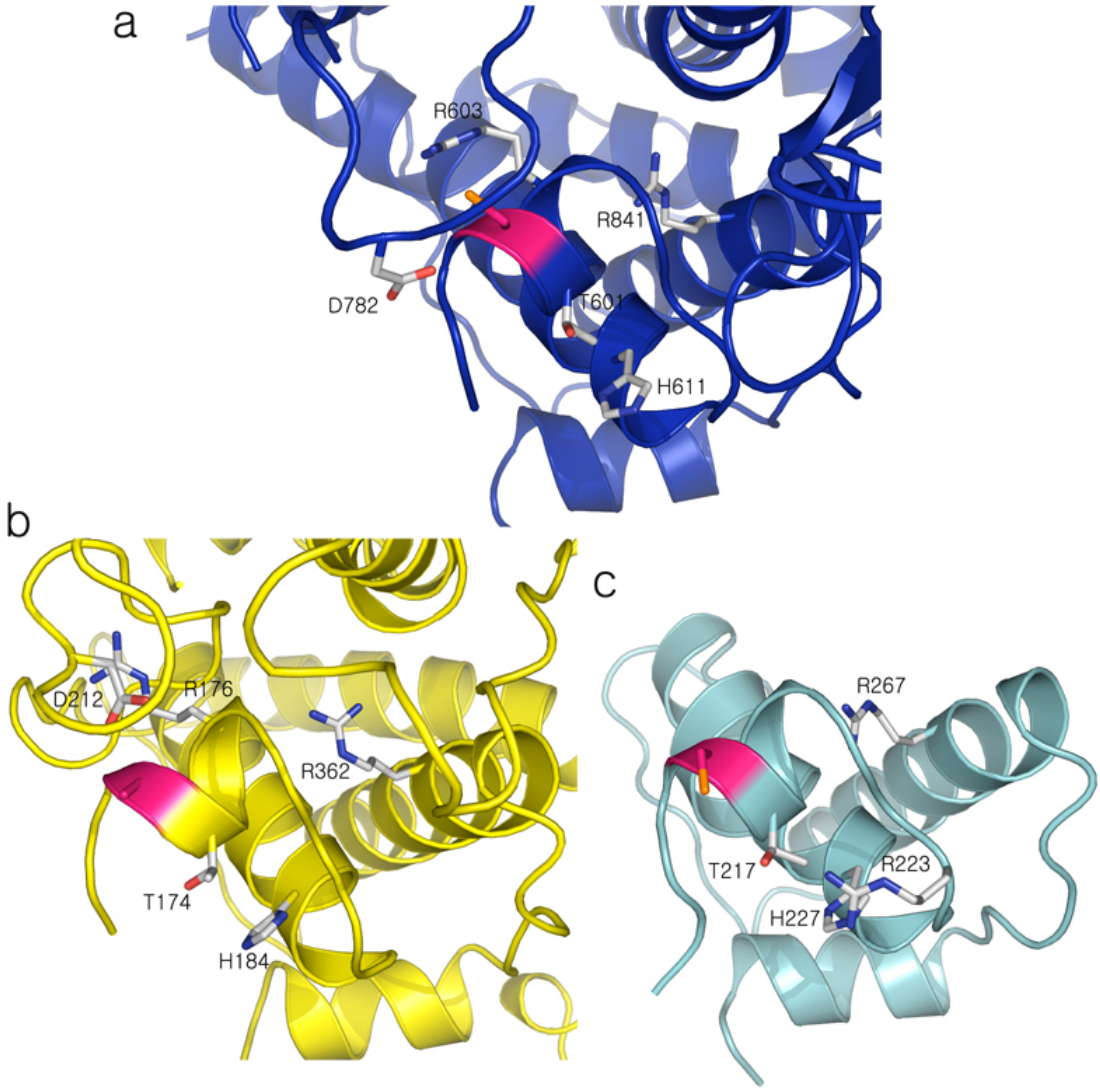


Figure 2.18 Comparison of the E1 catalytic cysteine active sites from (a) Uba1, (b) SUMO-E1, and (c) NEDD8-E1. Amino acid side chains discussed in the text are shown and labeled. The coordinates for SUMO-E1 and NEDD8-E1 were taken from PDB entries 1Y8R [29] and 2NVU [92], respectively. The catalytic cysteines (which has been replaced with Ala in the SUMO-E1) are colored in pink.

Another requirement for E1~Ubl thioester formation is that the active site cysteine needs to be in close proximity of the ubiquitin C-terminus. In the crystal

structure, the thiol of Cys600 is ~35 Å away from the adenylation active site, thus strongly suggesting that a conformational change in the complex would be required for the juxtaposition of the ubiquitin C-terminus and the active site cysteine thiol (Figure 2.19). The gap can be reduced to only ~15 Å, taking into account that the C-terminal five residues of ubiquitin are flexible [27, 93], and could, in principle, extend directly toward the active site cysteine, while remaining non-covalently associated with Uba1 via hydrophobic interactions. The remainder of the gap could be closed by a change in the relative orientations of the adenylation, the SCCH, and the FCCH domains. Indeed, the AAD and the SCCH are linked by two extended loops (residues 582-598 and 861-867) and the IAD and the FCCH are connected by two long antiparallel β-strands (residues 169-175 and 260-266), which could serve as hinges that allow the domains to rotate with respect to one another. Additionally, a local conformational change might occur around the catalytic cysteine.

On the basis of the structure, we hypothesize ubiquitin~adenylate slides out from underneath the crossover loop coupled with conformational changes bringing the catalytic cysteine closer, as a prerequisite to allow formation of the thioester linkage. At this point the catalytic cysteine could pull ubiquitin off its binding site in the AAD, leading to rebinding in the additional binding site near the SCCH. We are currently testing this model via a Uba1 variant that can crosslink the two loops connecting the AAD and the SCCH domains (Figure 2.19).

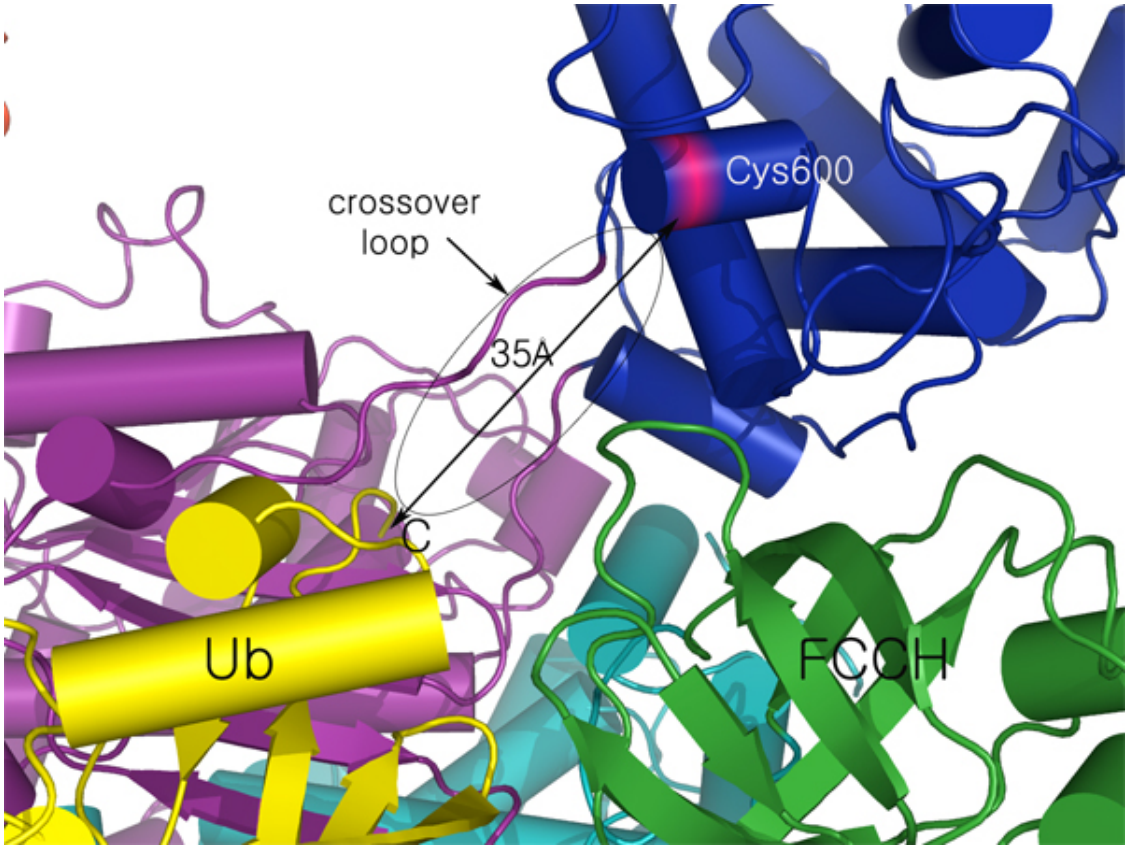


Figure 2.19 Separation between the ubiquitin C-terminus and the Uba1 catalytic cysteine. The double-sided black arrow marks the distance between the ubiquitin C-terminus and Cys600 of Uba1. The black oval indicates the region where a hypothetical "sliding" of the ubiquitin's C-terminal tail could occur.

H. Structural and mechanistic insights into Ubl transfer from E1 to E2

The final function of E1 is to coordinate the Ubl with its correct E2. This involves E1 interacting non-covalently with one of the cognate E2 enzymes, and a subsequent transthioesterification reaction in which the Ubl is transferred from E1's catalytic cysteine to that of the E2. For ubiquitin, two E1 enzymes sit at the top of the cascade, transferring ubiquitin to many different E2s, one at a time. The fully-loaded E1 has a strong affinity for E2s, with dissociation constants (K_d) in the subnanomolar to nanomolar range depending on the E2 [69]. However, the ubiquitin-E1 appears to have a

low affinity for E2s in the absence of ubiquitin, as E1 readily separates from the E2s during purification [66]. The low affinity of the free E1 for E2s might facilitate the rapid cycling of E1, so that it can charge ubiquitin's many different E2s [66]. The situation may differ for the Ubls SUMO, NEDD8 and other family members, because there is only one known E2 for each of these cascades. Although the K_m is not a measure of complex stability, it should be pointed out that the K_m for NEDD8's E2, Ubc12, during the reaction is similar to those for ubiquitin-E2s during their reactions [65]. However, both Ubc12 and the E2 for SUMO family members, Ubc9, have been shown to associate with their E1s in the free state [70].

Previously, the region of an E2 that is involved in the non-covalent interaction with an E1 has been mapped by mutational analysis. A mutational analysis of Ubc9 from *S. cerevisiae*, the E2 for the SUMO family member Smt3, mapped the E1 binding site to the N-terminal helix (H1) and the loop between the first and second β-strand (β1β2 loop) [70]. The location of the E1 binding site is probably conserved among E2s for ubiquitin and other Ubls, because two studies of E2s for ubiquitin revealed that mutations in H1 diminished E2~Ub thioester bond formation [94, 95]. In addition, in another ubiquitin E2, Ubc13, a protein-protein interaction that blocks access to this surface impairs E2~Ub thioester bond formation [96-98].

Recently, the crystal structure of a complex between the C-terminal domain (UFD) from NEDD8's heterodimeric E1 (APPBP1-UBA3) and the catalytic core domain of Ubc12 was reported [35]. It revealed that the E2 binds to the concave surface formed by the UFD's twisted β-sheet, in a manner which resembles ubiquitin's interactions with ubiquitin-binding domains (UBDs) (Figure 2.20).

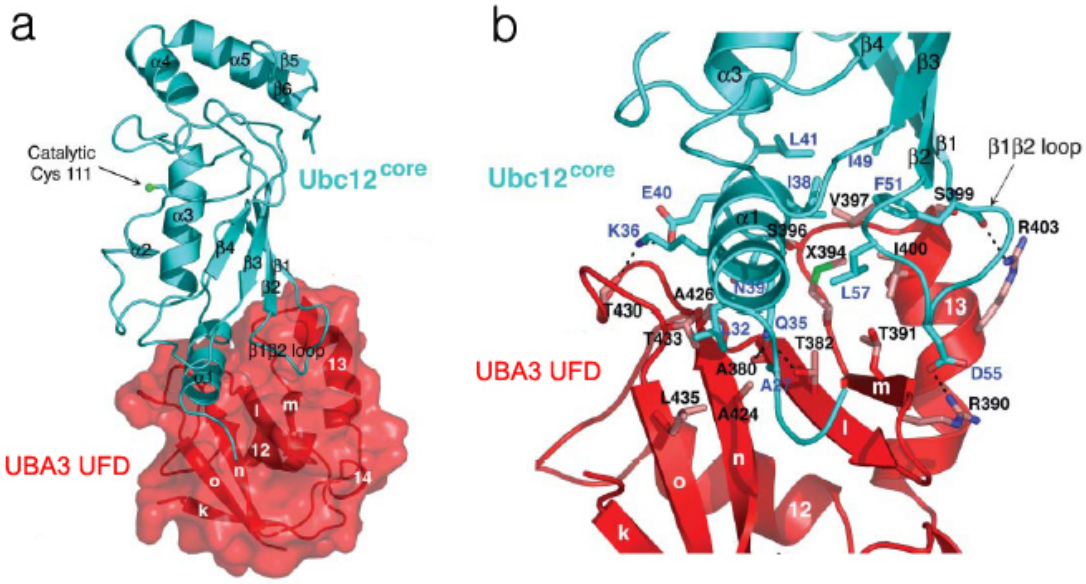


Figure 2.20 Details of the UBA3 UFD-Ubc12^{core} interface [35]. (a) Overall structure of the complex, with the UFD of UBA3 in red cartoon representation overlaid with a transparent surface and the Ubc12^{core} shown in a cartoon representation in cyan. Secondary structure elements are labeled, and the position of Ubc12's catalytic cysteine is represented by a green sphere. (b) Close-up view of interactions between the UBA3 UFD and the Ubc12^{core}. The side chains of the UBA3 UFD are shown in pink and labeled in black and those of the Ubc12^{core} in cyan and labeled in blue. Hydrogen bonds are represented by dashed lines. This figure was reproduced from Huang et al [35].

E2 recognition: Specificity elements in the E1 UFD-E2 interface

In the Uba3 UFD-Ubc12 structure [35], the N-terminal helix (H1) of Ubc12 makes several contacts with the surface of the β-sheet in the UFD. This surface is composed primarily of residues with small and uncharged side chains (Ala380, Thr382, Thr384, Thr391, Ala424, Ala426, and Leu435), which facilitates interactions with Gln31, Leu32, Gln35, and Asn39 in helix H1. One aspect that should be mentioned here is that Ubc12 has a unique 26-residue N-terminal extension upstream of its ~150-residue conserved E2 core domain, and that this extension fits into the groove conserved in UBA3, but not in other E1s. About 50% of the binding energy between the NEDD8-E1

and Ubc12 comes from this interaction (B. Schulman, personal communication). Thus, the NEDD8-E1 has an additional mechanism to select its E2 partner [99]. However, E2s for ubiquitin-E1 (13 in yeast) and for SUMO-E1 (1 in yeast, Ubc9) do not have extensions at their N-termini and most of the interactions must occur between UFD and H1 of E2 and the β1β2 loop.

Although it is not possible to directly discern the structural basis for specificity from this model, it is possible to identify differences between UFD and E2 sequences within the postulated interface which may participate in specificity (Figure 2.34 and 2.35). Notably, the E2-binding grooves of Uba1's UFD and to some extent, Sae2's UFD have distinct charge distributions (Figure 2.21). Both feature a generally acidic surface, however Uba1's UFD is more acidic with three conserved acidic residues (Glu1004, Asp1014 and Glu1016) clustered in the last two C-terminal β-strands (β30 and β31). When the Uba1 UFD is superimposed onto the Uba3 UFD-Ubc12 structure, these three residues appear to make contacts with Lys5 and Lys9 (*S. cerevisiae* Ubc1 numbering) from helix H1 of the E2. The alignment of yeast E2 sequences reveals that these two lysines are conserved in most E2 sequences, except in Ubc3 and Ubc9. Notably in Ubc9, which is the only SUMO-conjugating enzyme, the second lysine is replaced by a glutamate. Mutations of the conserved acidic residues from the E2-binding surface of UFD significantly decrease formation of the Ubc1~Ub thioester product (Figure 2.22). The β1β2 loop of ubiquitin-E2s likely makes interactions with residues from β29 of Uba1's UFD, however, the prominent structural difference in this region between ubiquitin-E2s makes it difficult to assess these potential interactions.

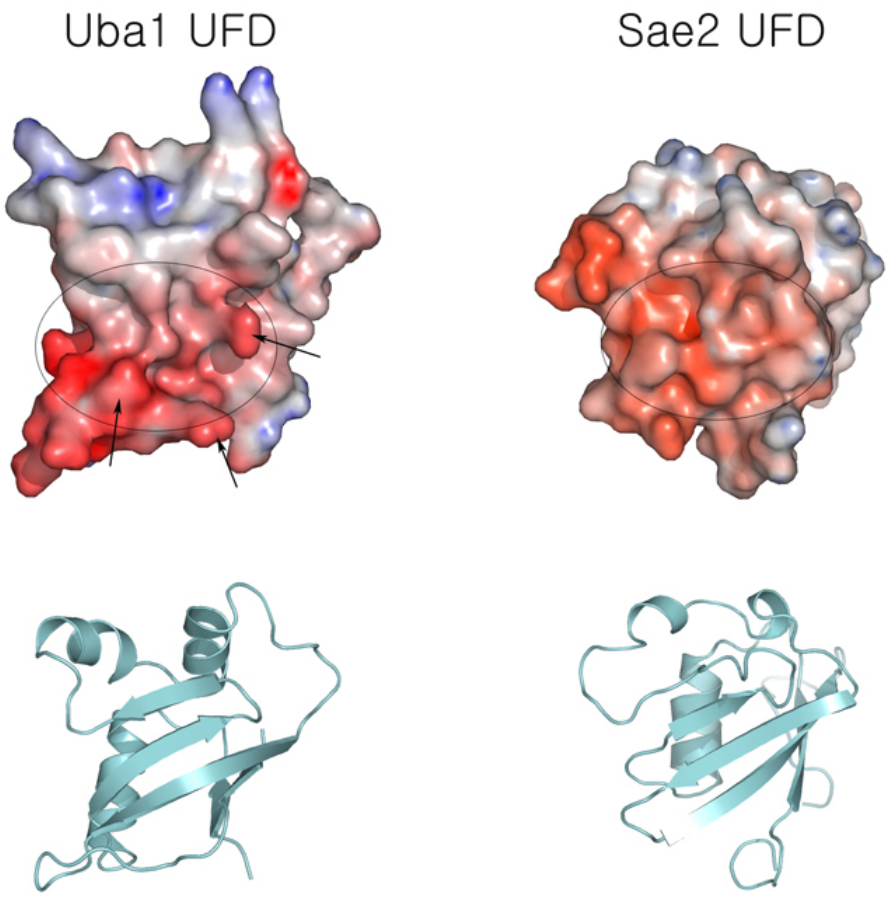


Figure 2.21 Electrostatic surface representation of UFDs from Uba1 and Sae2 viewed into the postulated E2-binding groove (black oval). The three conserved acidic residues that supposedly interact with basic residues from the N-terminal helix (H1) of the E2s are indicated by the black arrows for the Uba1 UFD. The corresponding ribbon images below are shown for orientation purposes.

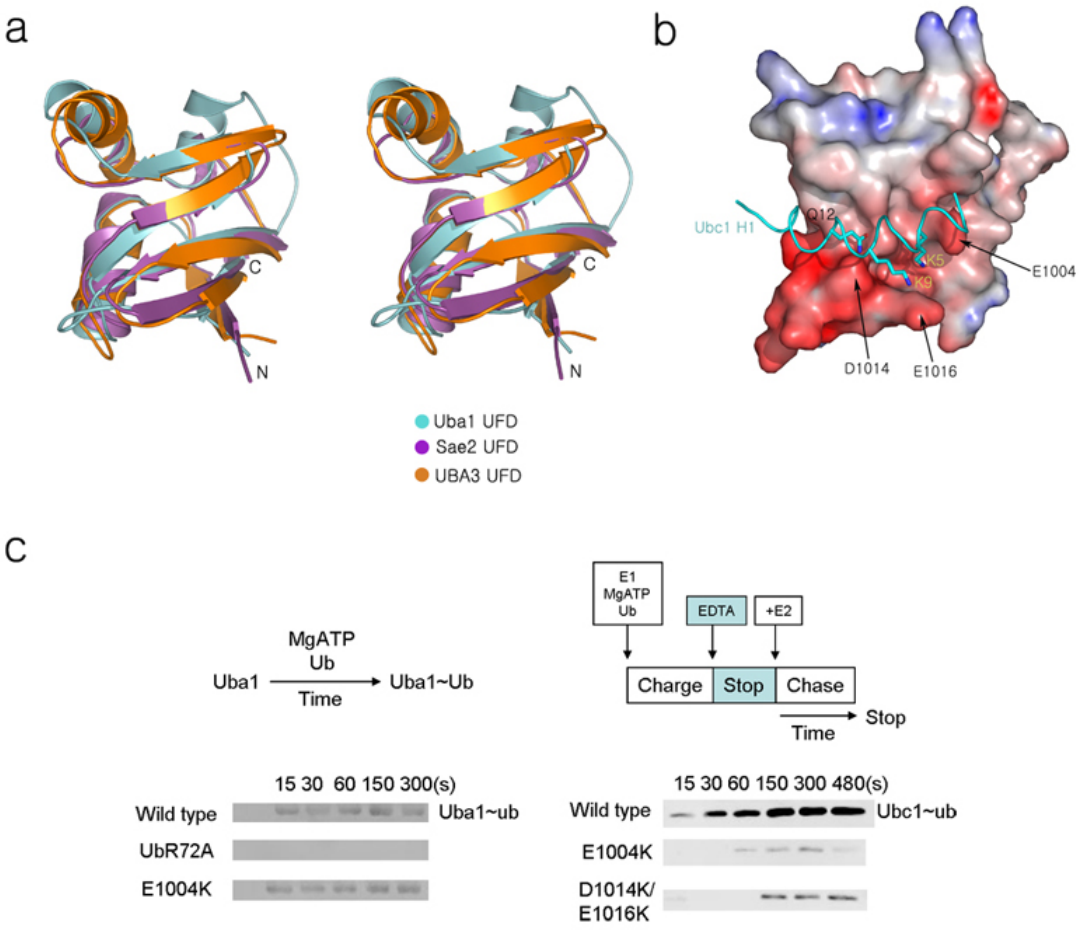


Figure 2.22 Structure of the Uba1 UFD and its role in the transthioesterification reaction. (a) Stereo view of the superposition of the UFDs from the ubiquitin, SUMO and NEDD8 E1 enzymes. (b) Electrostatic footprint and surface complementarity of charged residues of helix 1 (H1) of Ubc1 and Uba1's UFD. Helix H1 of Ubc1 was modeled as described in Materials and Methods, section G, and the conformations of the Lys5, Lys9, and Gln12 side chains were adjusted to avoid steric clashes. The black arrows point to the side chains from UFD that presumably contact the residues mentioned above. (c) Mutational analysis of the thioester formation (left) and transthioesterification (right) reactions. The E1004K variant does not prevent Uba1~Ub thioester formation, but strongly reduces ubiquitin transfer to the E2. The D1014K/E1016K double mutant also blocks transthioesterification.

Although differences in residues predicted to be at the interface between the E2 enzyme and the UFD for the complexes analyzed here are evident, it appears as if specificity in this system is unlikely to reflect one or a small number of changes in amino acids at the interface between the E2 and the UFD. In other words, E2s do not appear to have a single specificity determinant for E1, such as the Arg72 of ubiquitin. First, an in-depth analysis of E2 sequences (especially the N-terminal helix which binds UFDs) failed to identify classes of residues that might segregate E2s into distinct classes related to the E1s that they function with [100]. Second, extensive alanine scanning mutagenesis of the Uba3 UFD-Ubc12 interaction indicates that many residues will contribute to the interaction. For example, mutation of 8 of 9 interface residues in Ubc12 reduced or abolished its charging by NEDD8 [35, 92]. Likewise, mutations of 4 of 5 residues on the interaction surface of UBA3's UFD reduced charging of Ubc12 [35]. Thus, the structural analysis of at least one ubiquitin E1-E2 complex will be required to further address this question.

Transthioesterification model for the ubiquitin-E1

Upon docking the structures of a complex between the UFD of the E1 and E2s onto full-length structures of apo or single Ubl-loaded E1s for NEDD8 and SUMO, it becomes apparent that the E2 would bind to the opposite side of and face away from E1's catalytic cysteine (Figure 2.23). Thus, significant conformational changes would be required to enable the E1 and E2 catalytic cysteine to face and approach each other in both of these E1 enzymes. Indeed, the trapped $(NEDD8)_2$-E1 complex structure shows that the UFD changes its conformation dramatically to allow transthioesterification

(Figure 2.24). During this process, the Ubc12 binding site on the AAD which interacts with the UFD and hence is buried, is unmasked [92].

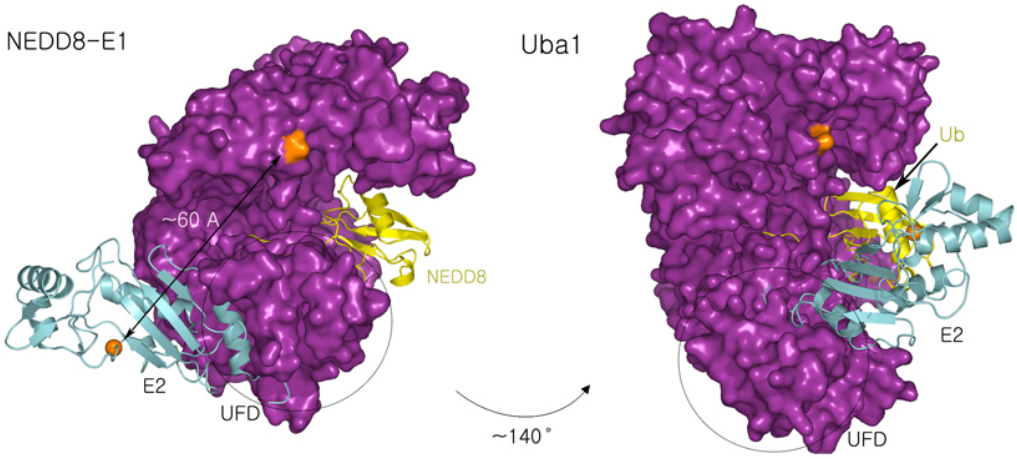


Figure 2.23 Surface representation of the NEDD8-E1 and Uba1 with the docked Ubc12, the NEDD8-E2. Ubc12 (cyan), NEDD8 (yellow), and ubiquitin (yellow) are shown in ribbon representation. Notably, the docked E2 is in a favorable orientation for transthioesterification in case of Uba1. The E1 and E2 catalytic cysteines are colored in orange. For NEDD8-E1, the distance of ~60 Å between the two cysteines is indicated. The corresponding distance in the Uba1-E2 complex is significantly shorter with ~30 Å. The black ovals highlight different orientations of the UFD in the two E1s.

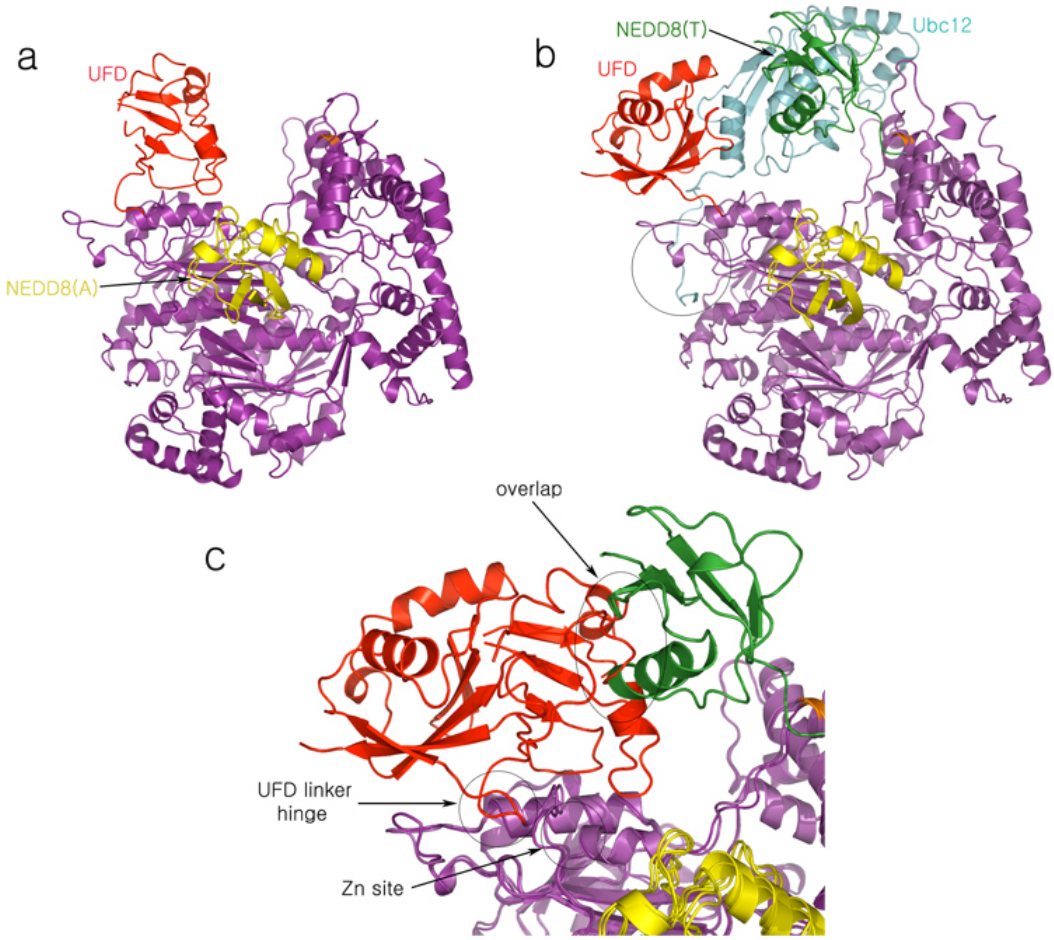


Figure 2.24 Structures of the NEDD8-E1 in complex with (a) NEDD8 (yellow) and (b) Ubc12 (cyan) and thioester-linked NEDD8 (forest) with another NEDD8 in the adenylation active site (yellow) bound in the same orientation as in (a) [80, 92]. UFD is colored in red, and the rest of the E1 in purple. Catalytic cysteines are highlighted in orange. (c) Superposition of the two complexes. The hinge region located near the Zn-binding site is identified by an arrow. The black ovals in (b) and (c) indicate the unique E1-E2 binding site in NEDD8-E1 and a structural overlap between the NEDD8(T) and the conformation of the UFD shown in (a).

In contrast, a detailed structural analysis of Uba1 suggests that the ubiquitin-E1 will likely go through a distinct conformational change, largely due to a difference in the architecture of the protein. First, the "canyon" of Uba1 is much wider (~40 Å) compared to other E1s (Figure 2.25), and the UFD in the single ubiquitin-loaded E1 structure

presented here is in a comparable position as the UFD of the doubly NEDD8-loaded E1 (Figure 2.26). Also, when one predicts where the thioester-linked ubiquitin would be bound according to the orientation of the thioester-linked NEDD8 in the doubly-loaded complex, it appears as if would not clash with the UFD as seen in the case of NEDD8-E1 due to the presence of the much larger gap in the E1 for ubiquitin. Second, the E2-binding groove on the Uba1 UFD faces toward the SCCH, and as a consequence, the E2 catalytic cysteine also faces into the direction of E1's large central groove, which contains both the ubiquitin-binding site and the E1's active site cysteine (Figure 2.23). Third, in contrast to other E1s, a relatively small change in the UFD linker can bring the E2 active site cysteine close to the E1 cysteine. Already in the structure presented here, the gap between the E1 and E2 catalytic cysteines is reduced from ~38 Å to ~27 Å based on the two different orientations of the UFD linker observed in the asymmetric unit (Figure 2.27). Finally, there are notable differences in the size and directionality of the β hairpin near the end of UFD linker. In contrast to the SUMO-E1 and NEDD8-E1, Uba1 has a longer antiparallel β hairpin (residues 915-927) which is facing "outward" (Figure 2.28). Uba1 and all ubiquitin E1s are bereft of the zinc-binding motif, which coordinates a zinc ion in the vicinity of this hairpin in other E1 enzymes. One implication of the absence of the zinc ion is that it causes the UFD linker in Uba1 to be more stretched out which contributes to the wide gap between the SCCH and the UFD. Due to these structural differences it is likely that the ubiquitin-E1 does not require such a dramatic conformational change in the UFD linker hinge as observed for the NEDD8-E1 and probably also the SUMO-E1 and hence, a smaller conformational change in the orientation of UFD is required during the transthioesterification reaction cycle.

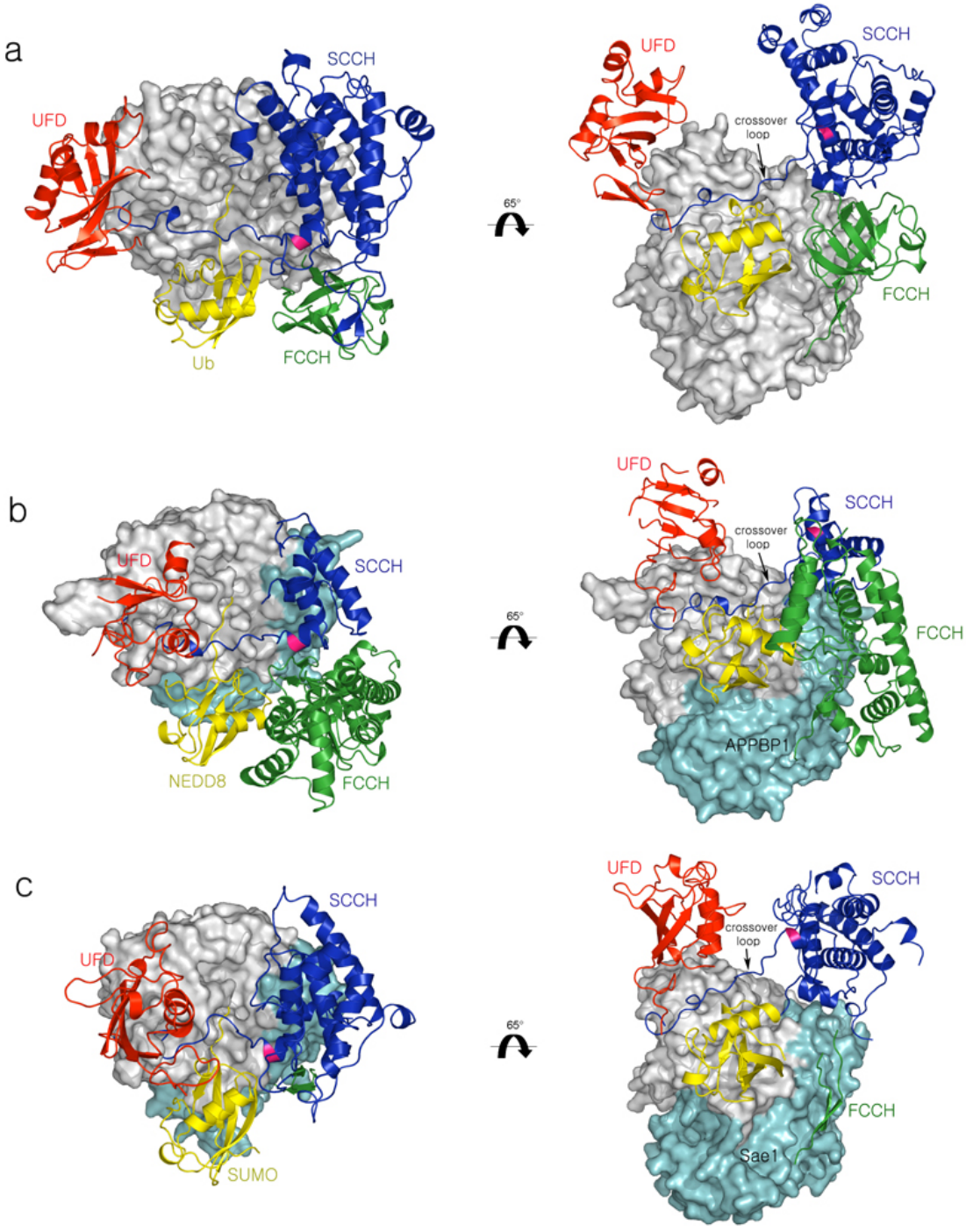


Figure 2.25 Comparison between the ubiquitin, NEDD8, and SUMO E1 enzymes. (a) Ribbon and surface representations for the Uba1-ubiquitin complex. The catalytic cysteine is shown in pink. The "canyon" between the SCCH and the UFD is considerably wider compared to the other complexes. (b) Ribbon and surface representations of the APPBP1-UBA3-NEDD8 complex (PDB entry 1R4M). APPBP1 is colored in cyan. (c) The corresponding representation of the Sae1-Sae2-SUMO complex (PDB entry 1Y8R).

The FCCH is mostly disordered in the structure. Sae1 is colored in cyan. Notably, the β-sheet surfaces of UFDs are all facing into different directions.

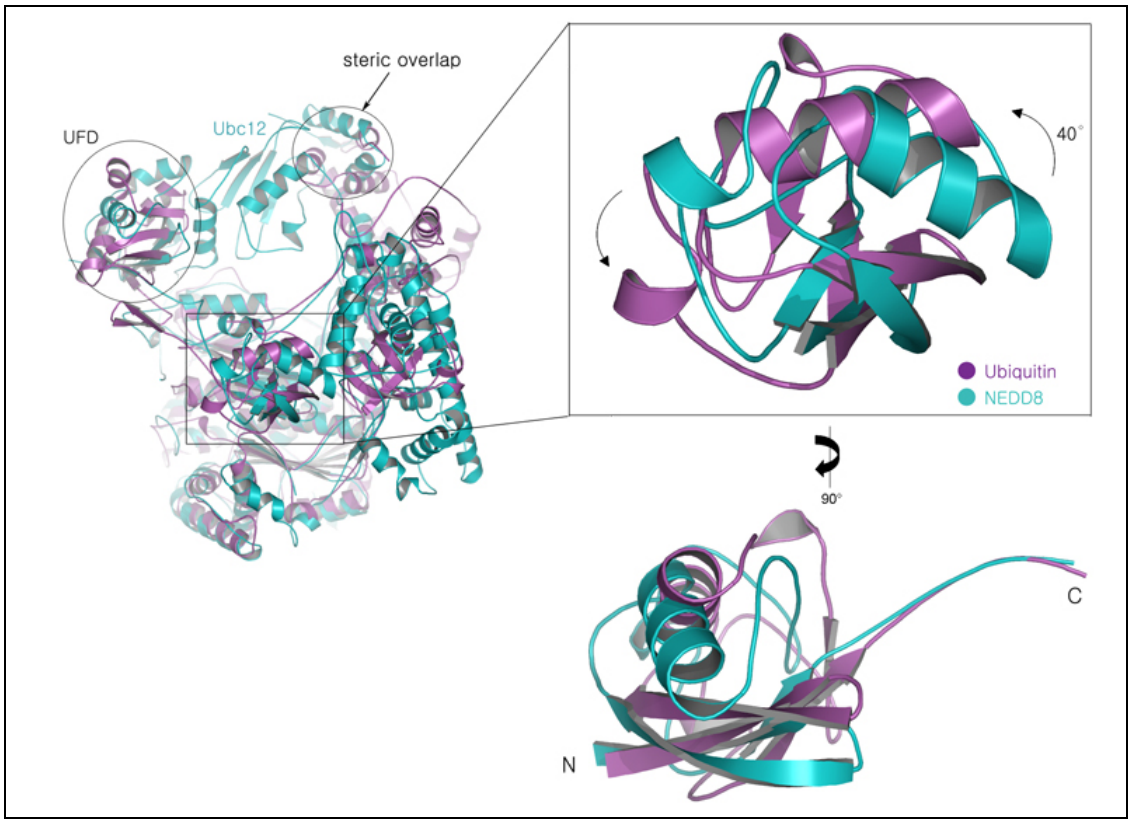


Figure 2.26 Ribbon representation of the trapped NEDD8 activation complex (teal) (PDB entry 2NVU, [92]) and the Uba1-ubiquitin complex (magenta) superimposed via the two adenylation domains. The black ovals indicate that the UFD of a single ubiquitin loaded-Uba1 is in a comparable position with the UFD of the NEDD8 activation complex, and that the bound Ubc12 of the NEDD8 activation complex clashes with the SCCH of Uba1. Ubiquitin bound in the adenylation active site is rotated by ~40° compared to NEDD8.

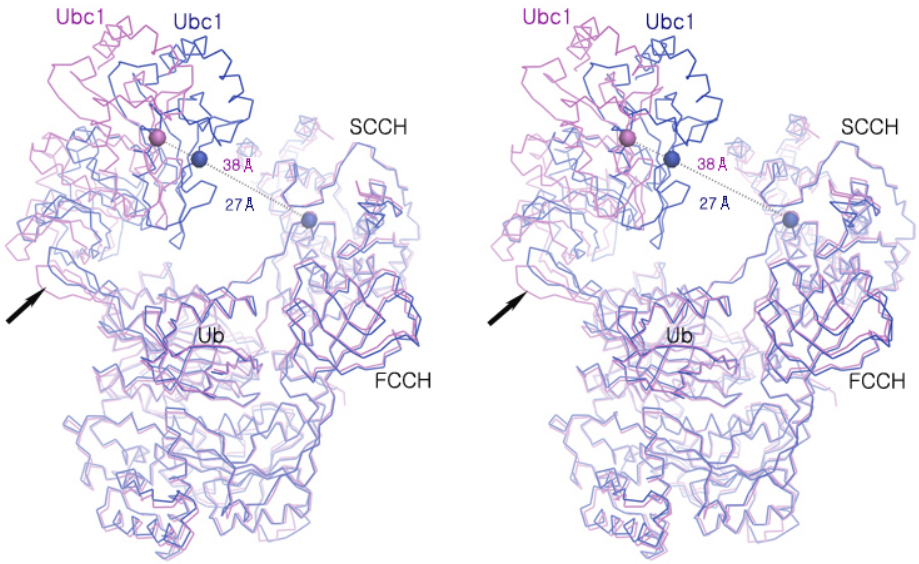


Figure 2.27 Stereo view showing the conformational changes in UFD and the UFD linker observed in the crystal structures. The two copies of the Uba1-ubiquitin complex (magenta and blue, respectively) present in the asymmetric unit were superimposed. The complex with Ubc1 was modeled as described in Materials and Methods, section G. The catalytic cysteines of Uba1 and Ubc1 are represented as spheres and dashed lines which are labeled with the distances between them. The black arrows point to the UFD linker to highlight the conformational differences.

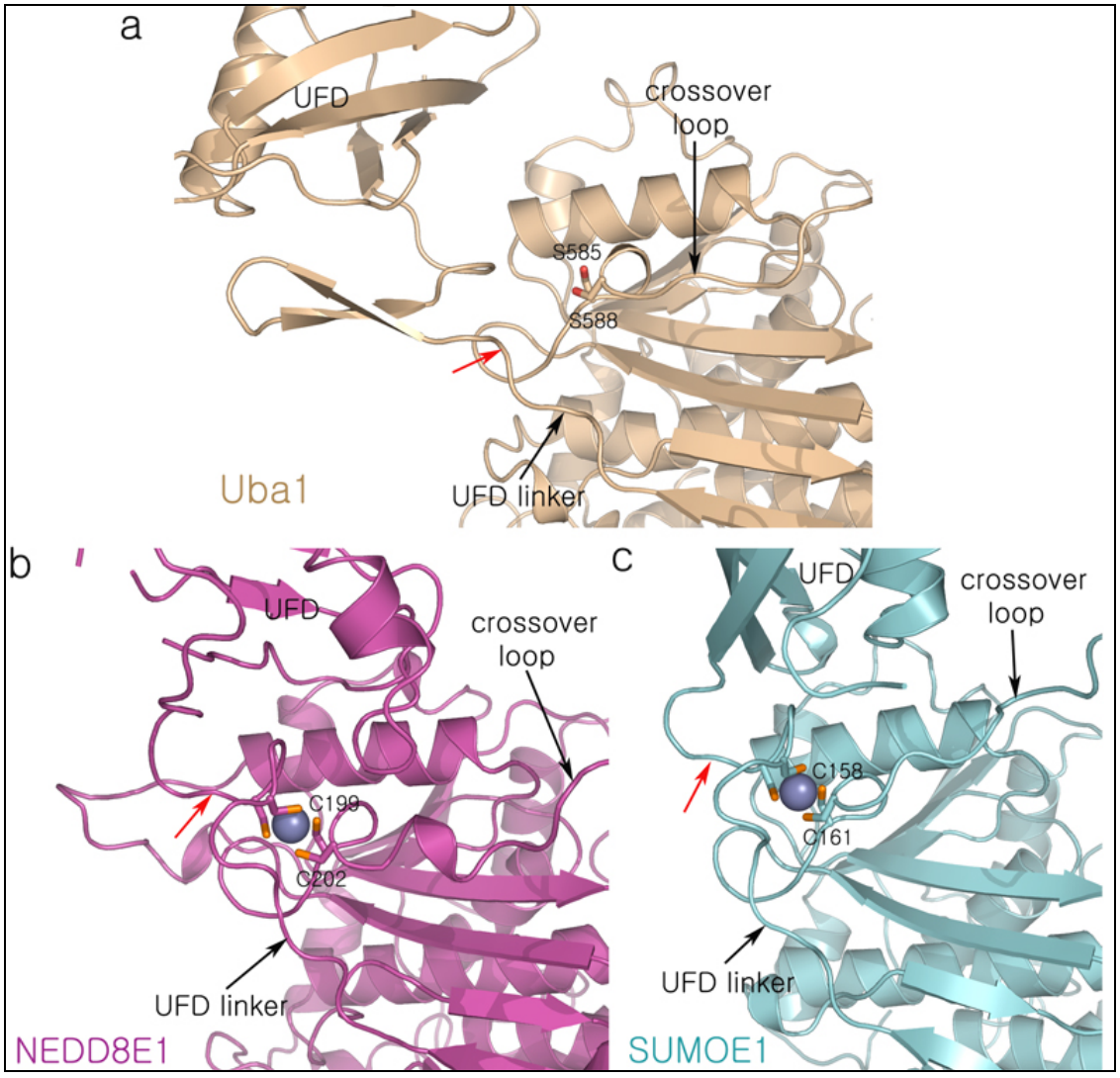


Figure 2.28 Structural comparison of the UFD linkers in E1s and their contribution to the hinge mechanism. (a) In Uba1, the first Cys-X-X-Cys motif on the crossover loop is replaced by Ser-X-X-Ser. (b,c) In NEDD8- and SUMO-E1, a zinc ion is coordinated by the two Cys-X-X-Cys motifs in the same manner. Zinc ions in (b) and (c) are shown as blue spheres. Red arrows indicate the hinge location on the UFD linker in each E1.

Based on the single ubiquitin-loaded Uba1 crystal structure, a structural model of a transthioesterification complex for the ubiquitin-E1 has been constructed (Figure 2.29). The model shows that a significantly smaller rotation of the UFD by ~40° in contrast to the ~120° rotation for the UBA3 UFD around residues in the hinge region (residues 912-

914) would bring the Uba1's and the Ubc1's catalytic cysteines into close spatial proximity as reflected by a drop in the inter-cysteine distance from ~38 Å to ~8 Å. (Figure 2.29). Two major structural features of Uba1 and Ubc1 and other E2s also contribute to this different behavior. For Uba1, because the UFD is connected to the adenylation domain by an extended loop of 18 amino acids (UFD linker), a small change at the hinge translates into a large shift in the orientation of the whole UFD. For Ubc1, because its core domain has an oblong tower-like structure which is anticipated for the core domains of all E2s, a small rotation at the base translates into a large translation for the catalytic cysteine.

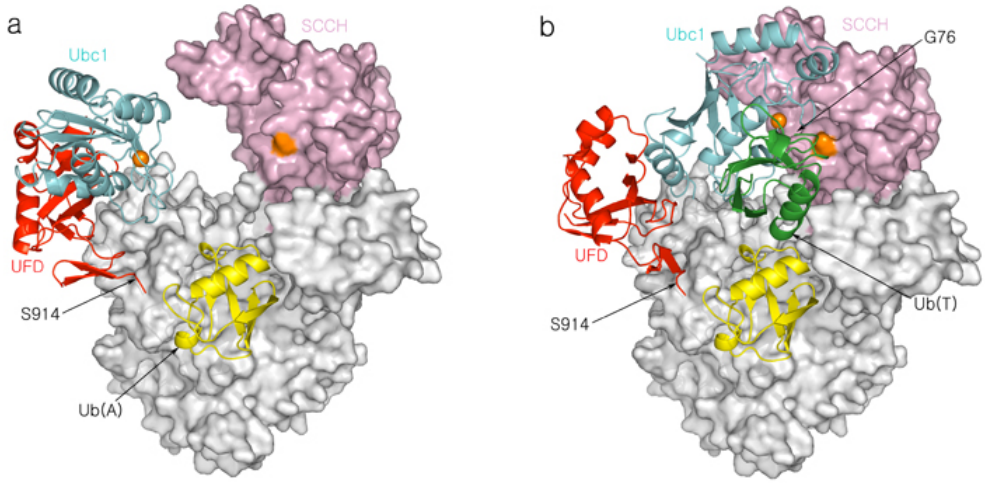


Figure 2.29 Model of the Uba1 transthioesterification complex. Surface representation: IAD, AAD, FCCH and SCCH; ribbon representation: UFD, Ub(A), Ub(T) and Ubc1. White: IAD, AAD and FCCH; Pink: SCCH; Red: UFD; Yellow: Ub(A); Forest: Ub(T); Cyan: Ubc1; Orange: Uba1's and Ubc1's catalytic cysteines (Cys600 and Cys88, respectively). (a) Pre-transthioesterification step. Ubc1 was modeled as described in Materials and Methods, section G. (b) A rotation around the "hinge" would allow the two catalytic cysteines to come into close proximity to facilitate the transthioesterification reaction. Ub(T): Ub thioester-linked to Cys88 of Ubc1. This model is based on the Ubc1-Ub thioester intermediate (PDB entry 1FXT [39]). Residues 779-785 from SCCH were removed to allow for a clearer presentation of Cys600 of Uba1 and the C-terminal tail of Ub(T).

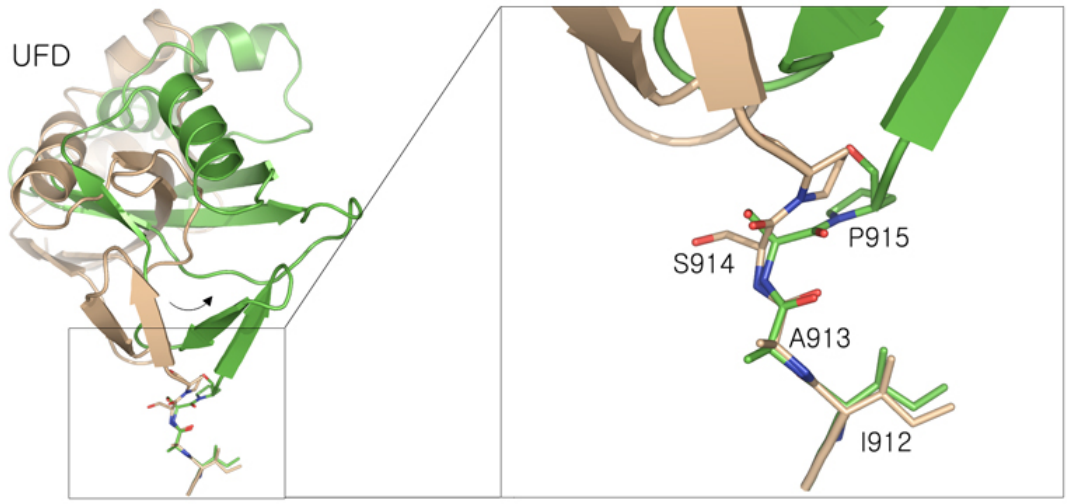


Figure 2.30 The UFD linker hinge. The model predicts that the UFD needs to rotate approximately 40° to assume a position that is competent for transthioesterification.

Three aspects of the Uba1 structure are consistent with the possibility that the UFD could rotate relative to the remainder of the E1. First, a comparison between the two Uba1 molecules present in the asymmetric unit revealed that the UFD linker connecting the UFD to the adenylation domain is particularly mobile, thus suggesting that domain rotations play a role during catalysis. Second, the three central residues in the linker, Ile912, Ala913, and Ser914, make no specific side chain contacts, neither to the remainder of the E1 structure nor to ubiquitin. Third, the UFD buries a relatively small interface with the adenylation domain (400 $Å^2$ of total surface area), suggesting that the UFD may act as an independent unit, rather than as part of a tighter complex.

In addition, the transthioesterification model reveals a potential interface between the SCCH and Ubc1. It appears that the central loop region (residues 775-795) of the SCCH would be in close contact with the incoming E2 and that the transthioesterification reaction may require it to move away from E2 to avoid a steric clash (Figure 2.31).

Indeed, the middle part of the loop region on the SCCH (residue 786-793) is disordered in the crystal structure, indicating that this loop is characterized by high conformational flexibility. These observations raise the possibility that the SCCH of the ubiquitin-E1 may have an intrinsic affinity for its E2s.

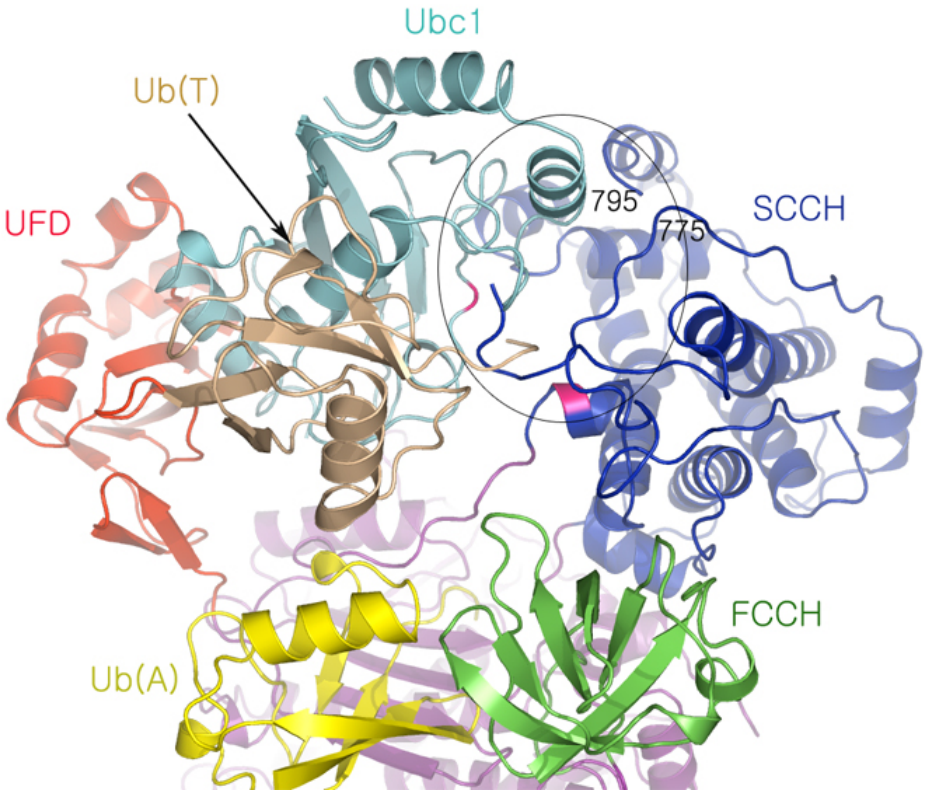


Figure 2.31 The putative interface between the E2, Ubc1 and the SCCH of the ubiquitin-E1 during E1-E2 transthioesterification. Ubc1 and Ub(T) were modeled as in Figure 2.29. The black oval highlights the interface. The central loop (residues 775-795) of SCCH is in close contact with Ubc1 and the C-terminal tail of thioester-linked ubiquitin, Ub(T). Uba1's and Ubc1's catalytic cysteines are colored in pink.

UFD linker hinge flexibility

To test the importance of the relative flexibility of this region on the ubiquitin transfer from E1 to E2, I constructed mutant forms of Uba1, each of which contains the intact UFD necessary for the proper E2 binding. Single turnover pulse-chase Uba1-Ubc1

transthioesterification assays were used in order to specifically examine effects of mutations in Uba1 on transfer to Ubc1, without sensing effects of mutations on any other functions of Uba1, such as the loading of ubiquitin. The activity was measured by Western blot analysis using anti-ubiquitin antibodies that detect all proteins included in the assay mixture that are bound to ubiquitin.

Based on the fact that a mutation to proline restricts the rotation about the polypeptide backbone [101], I singly mutated the central residues of the UFD hinge (Ala913 and Ser914) to proline, and also constructed a corresponding double proline mutation. Ala913Pro has little effect on Ubc1~Ub thioester formation as assessed by the single turnover pulse-chase experiment. In contrast, Ser914Pro and the double mutant Ala913Pro/Ser914Pro decreased the transthioesterification activity by more than 50% compared to the wild-type (Figure 2.32). Modeling of these mutants and subsequent inspection of the rotational freedom around the phi (φ) and psi (ψ) backbone angles of these positions revealed a moderate loss in allowable conformations for the UFD, as measured by the ability of the E2 catalytic cysteine to approach to within 10 Å of the thiol group of Cys600. Notably, the Ser914Pro single and the Ala913Pro/Ser914Pro double mutant exhibit a loss of rotational freedom; however they can still compensate by increasing the amount of rotation around the backbone psi (ψ) bond, which could be the reason that both still retain some activity.

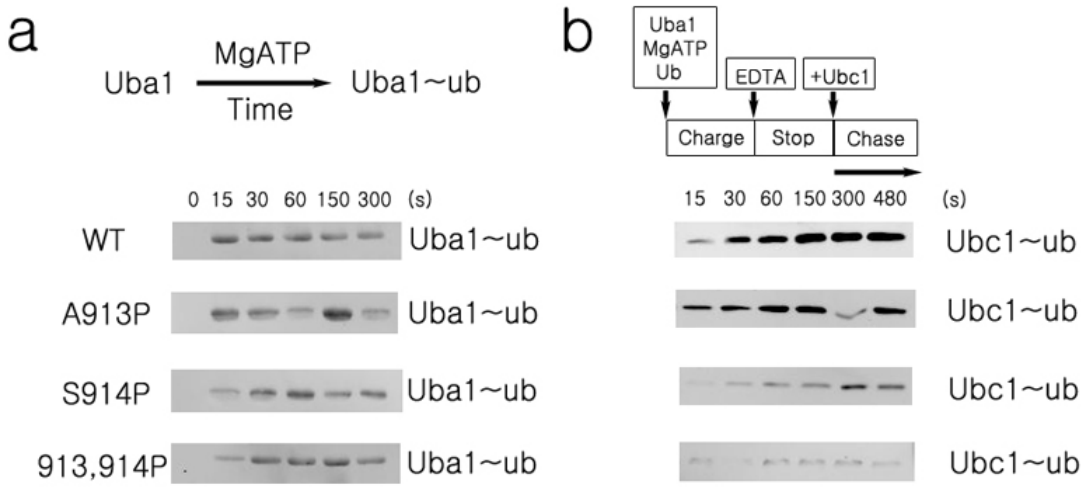


Figure 2.32 Mutational analysis of the UFD linker. (a) Effects of mutations on forming the Uba1~ub complex, monitored by following the time-course for forming the reducible complex of Uba1 with ubiquitin. Proteins were resolved by SDS-PAGE, and Uba1~ub was visualized by Western blot. (b) Effects of mutations on ubiquitin transfer from Uba1's catalytic cysteine Cys600 to the catalytic cysteine from Ubc1. Single turnover pulse-chase Uba1-Ubc1 transthioesterification assays were used in order to specifically examine effects of mutations in Uba1 on transfer to Ubc1. Wild-type and proline mutant versions of Uba1 used in each reaction are denoted on the left.

Supplementary figures

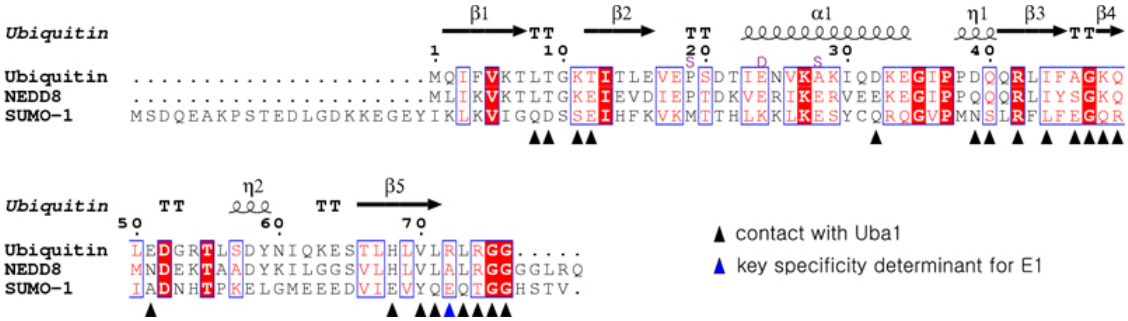


Figure 2.33 Sequence alignment of human ubiquitin, NEDD8, and SUMO-1. Sequence numbers are based on human ubiquitin. The three residues in which yeast and human ubiquitin differ are shown in purple on the top of the sequence at positions 19, 24 and 29. The secondary structure diagram for ubiquitin is shown on the top. Identical residues are colored in red.

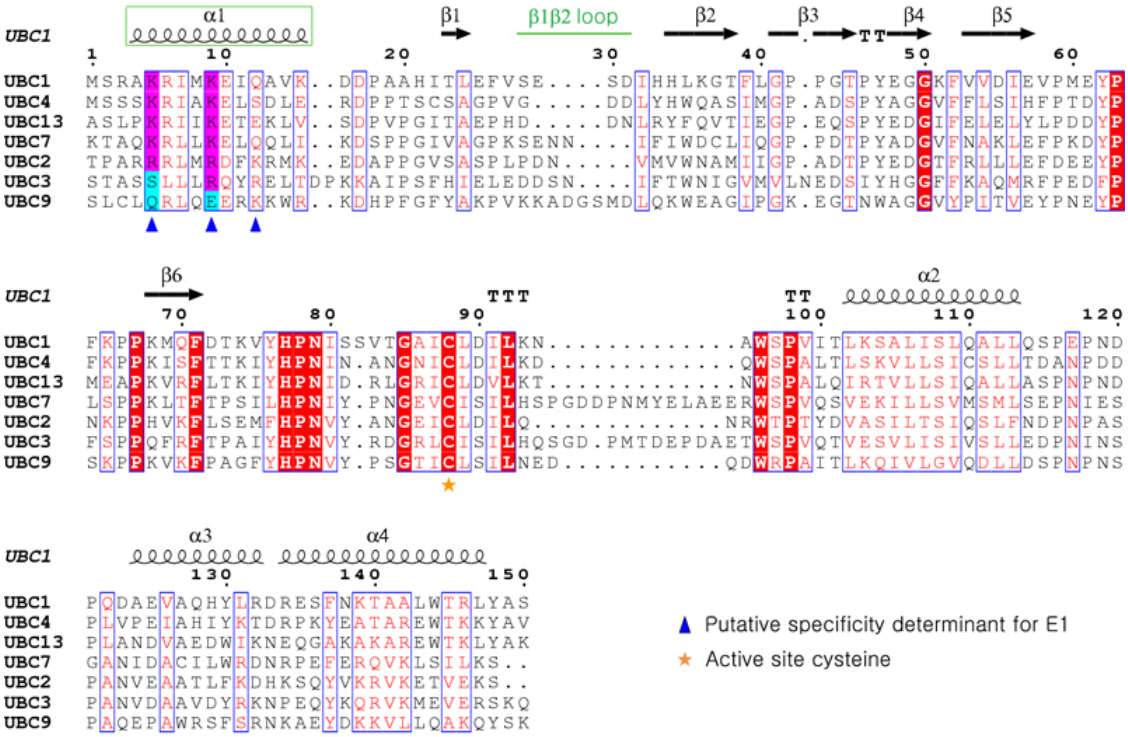
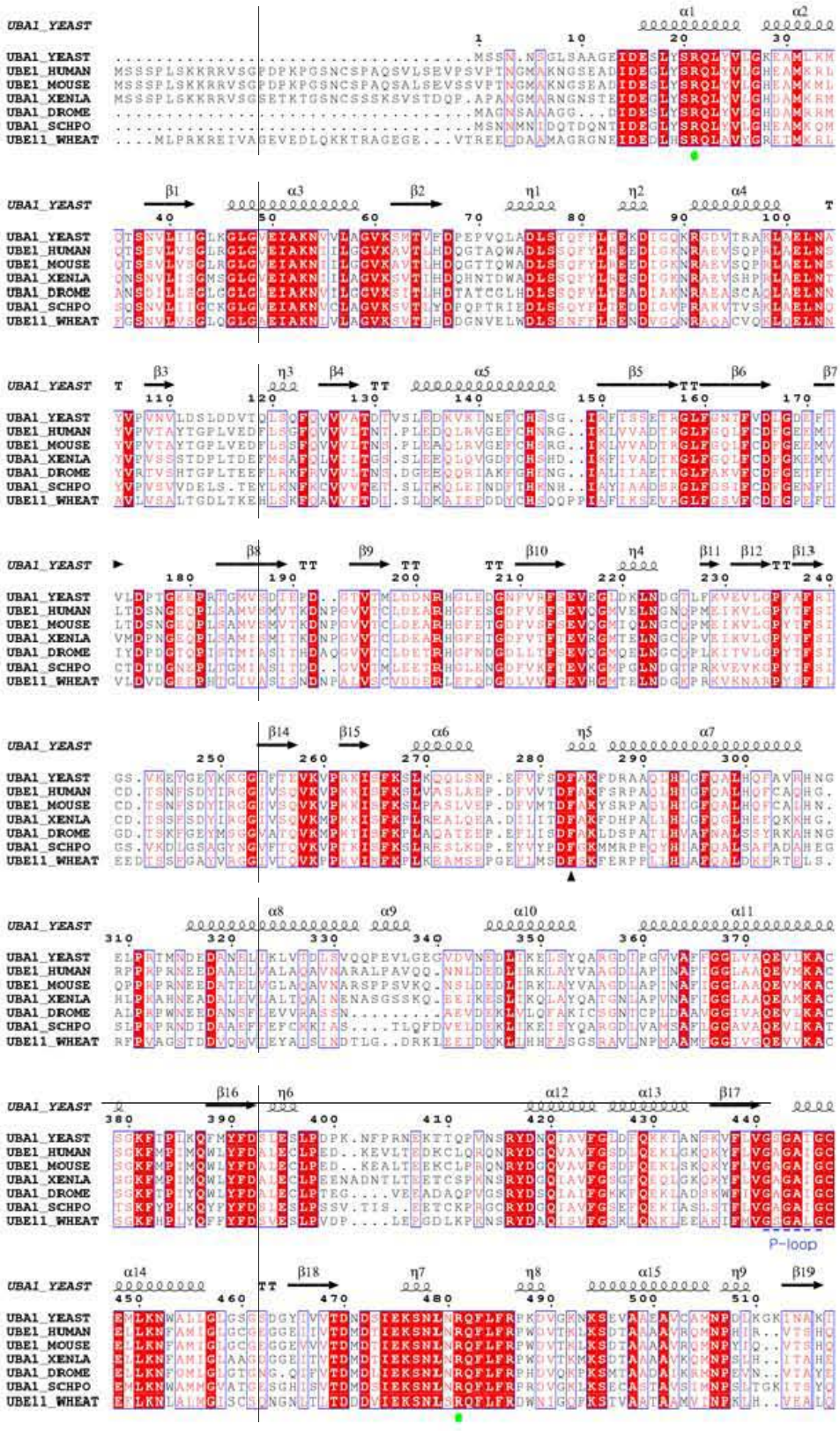


Figure 2.34 Multiple sequence alignment of several yeast ubiquitin-specific E2 enzymes and the E2 for SUMO (Ubc9). The numbering is based on *S. cerevisiae* Ubc1. The secondary structure diagram for *S. cerevisiae* Ubc1 is shown on the top. Identical residues are colored in red. Magenta and cyan highlight residues that are conserved among the ubiquitin E2s, but not in Ubc9, the SUMO-conjugating E2. The E1-interacting α-helix H1 and β1β2 loop are highlighted. The alignment only includes the E2 catalytic core domain (residues 1-150).


UBA1_YEAST
UBA1_YEAST
UBE1_HUMAN
UBE1_MOUSE
UBA1_XENLA
UBA1_DROME
UBA1_SCHPO
UBE11_WHEAT
P-loop

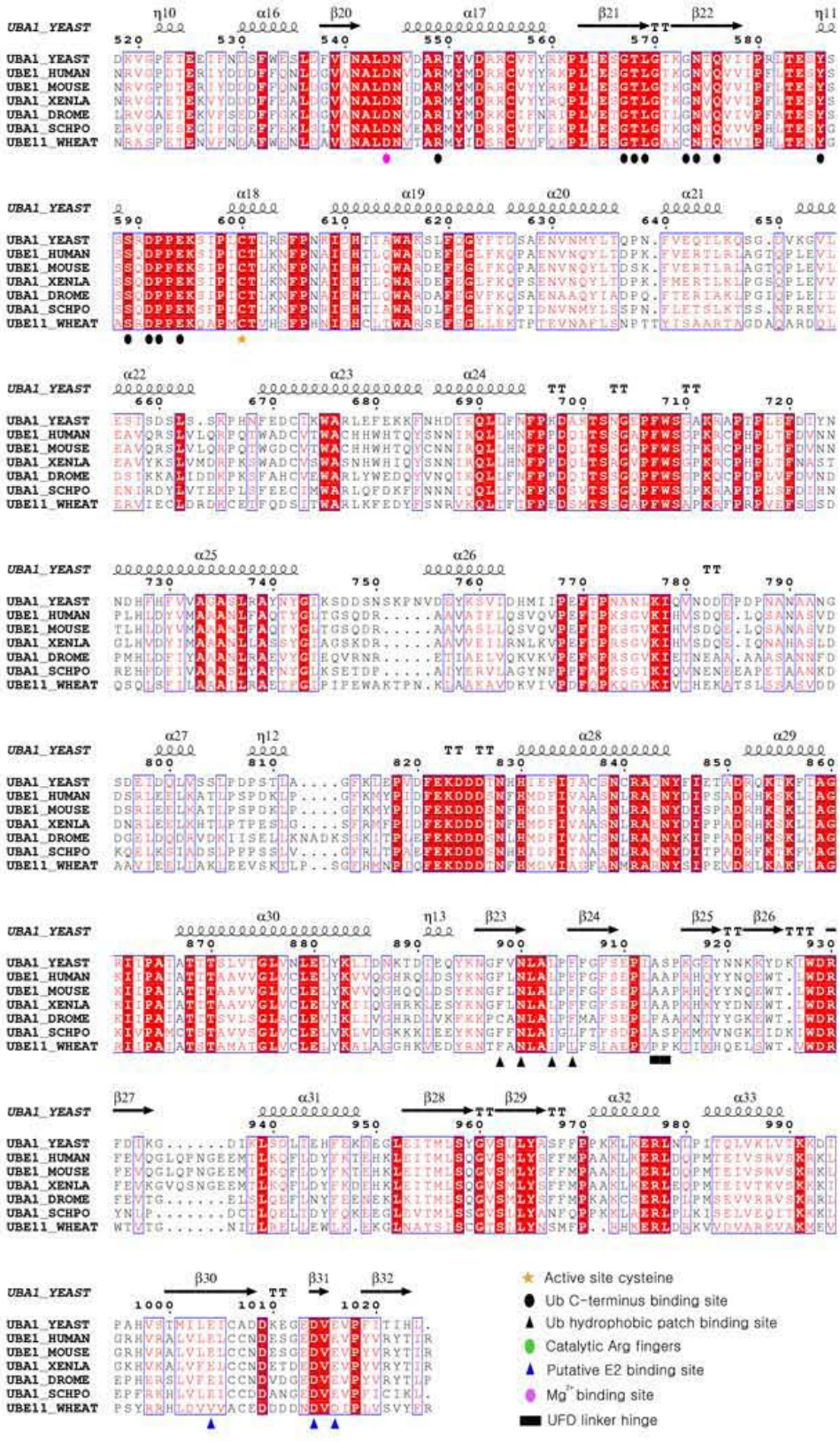

η10 α16 β20 α17 β21 β22 η11
UBA1_YEAST
520 530 540 550 560 570 580
UBA1_YEAST
UBE1_HUMAN
UBE1_MOUSE
UBA1_XENLA
UBA1_DROME
UBA1_SCHPO
UBE11_WHEAT
α18 α19 α20 α21
590 600 610 620 630 640 650
α22 α23 α24
660 670 680 690 700 710 720
α25 α26
730 740 750 760 770 780 790
α27 η12 α28 α29
800 810 820 830 840 850 860
α30 η13 β23 β24 β25 β26
870 880 890 900 910 920 930
β27 α31 β28 β29 α32 α33
940 950 960 970 980 990
β30 β31 β32
1000 1010 1020
Active site cysteine
Ub C-terminus binding site
Ub hydrophobic patch binding site
Catalytic Arg fingers
Putative E2 binding site
Mg2+ binding site
UFD linker hinge

Figure 2.35 Multiple sequence alignment of the E1 for ubiquitin from the following organisms: YEAST, *Saccharomyces cerevisiae*; HUMAN, *Homo sapiens*; MOUSE, *Mus musculus*; XENLA, *Xenopus laevis*; DROME, *Droshophila melanogaster*; SCHPO, *Schizosaccharomyces pombe*; WHEAT, *Triticum aestivum*. The numbers refer to *S. cerevisiae* Uba1, while the secondary structure diagram for *S. cerevisiae* Uba1 is shown on the top. Identical residues are colored in red. Alignments were generated using ClustalW (http://www.ebi.ac.uk/clustalw) and were visualized with ESPript (http://esript.ibcp.fr).

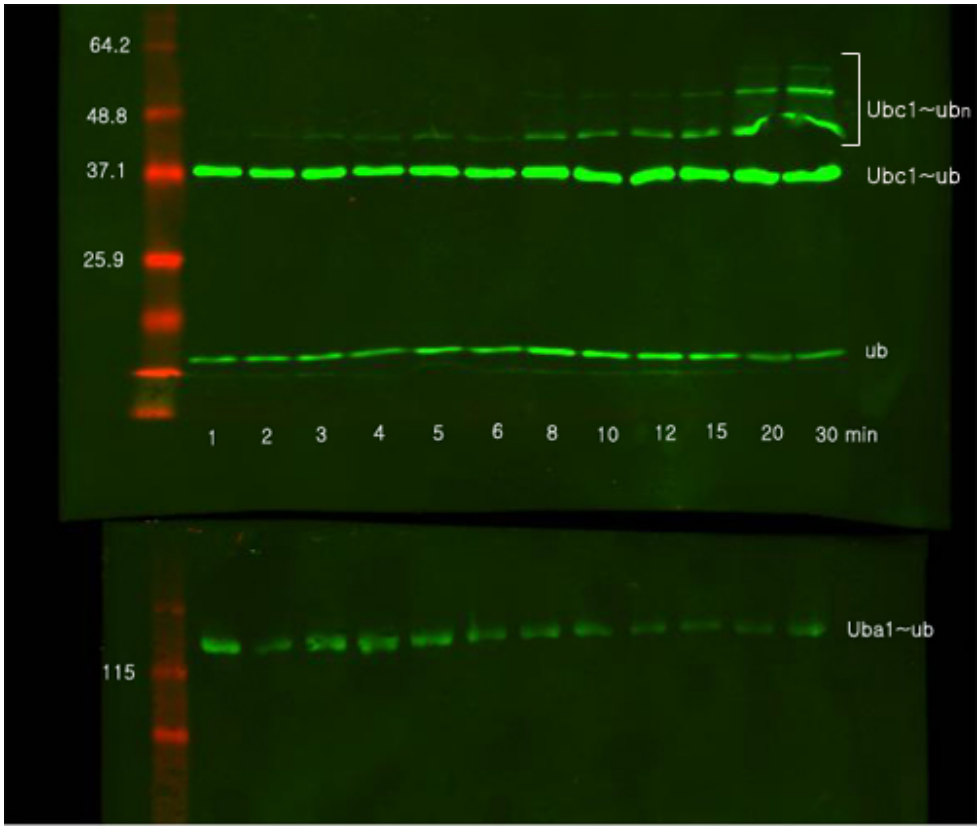


Figure 2.36 Ubc1 promotes the formation of polyubiquitin chains. (Top) Purified Ubc1 (2 µM) was incubated with Uba1 (1 µM), ubiquitin (4.1 µM), ATP (2.5 mM) and $MgCl_2$ (5 mM) for the indicated time at room temperature. Reactions were stopped by addition of SDS sample buffer without DTT and analyzed by SDS-PAGE and Western blot. (Bottom) The same reactions as shown on top in the absence of Ubc1.

IV. Discussion

Selectivity of ubiquitin and Ubl pathways

The Uba1-ubiquitin structure reveals that Uba1 interacts with a multipurpose binding site on ubiquitin. Ubiquitin's Leu8/Ile44/Val70 hydrophobic patch involved in interaction with Uba1 is absolutely conserved in NEDD8. In ubiquitin, this hydrophobic patch was originally identified as the site of proteasome binding [91]. More recently, ubiquitin's hydrophobic patch has been shown to be involved in interactions with the ubiquitin-associated (UBA) domain, ubiquitin-interacting motif (UIM), and CUE (coupling of ubiquitin conjugation to endoplasmic reticulum degradation) domains found in ubiquitin recognition machineries involved in endocytosis, ER protein sorting, vacuolar protein sorting, and other functions [102-107]. The finding that an E1 and these ubiquitin recognition domains all interact with a common surface suggests that Ubl conjugation and effector machineries have coevolved to distinguish their cognate Ubl.

The Uba1-ubiquitin structure reveals two mechanisms for establishing specificity. Comparison of the Uba1-ubiquitin with the MoeB-MoaD, APPBP1-UBA3-NEDD8, and Sae1-Sae2-SUMO structures [29, 31, 80] suggest that global differences in Ubl sequences and structures likely account for much of the observed specificity. The Uba1-ubiquitin structure also reveals how very closely related Ubls, such as ubiquitin and NEDD8, can be distinguished. Uba1 has several residues in the crossover loop that make favorable interactions with residues in the C-terminal tail of ubiquitin, including ubiquitin's Arg72.

E2 recognition: Multiple E2 binding sites on E1

The transthioesterification model suggests the presence of an additional E2 binding site on the SCCH of Uba1 besides the UFD. Recent studies on NEDD8 modification indicate that the UFD rotates ~120° during the E1~NEDD8 thioester formation step, which allows the binding of E2 near the SCCH of the E1 [35, 92]. However, in this structure, the E1 and E2 catalytic cysteines are still ~20 Å apart. Hence, the transthioesterification between E1 and E2 enzyme appear to require an additional translocation of the E2 toward the E1 catalytic cysteine.

The SCCHs of the ubiquitin-, NEDD8-, and SUMO-E1 vary widely in their lengths and sequences. However, both ubiquitin- and SUMO-E1 have a large surface loop near the catalytic cysteine in a similar position. A new study revealed that Ubc9 directly binds to this loop region in the SCCH of Sae2, thus indicating that the intrinsic affinity between the E2 and the SCCH of E1 drives the further translocation of E2 towards the catalytic cysteine of E1 [108] (Figure 2.37). Therefore, it is very likely that ubiquitin-specific E2s also have an intrinsic affinity for the SCCH of ubiquitin-E1.

The E1-catalyzed step is critical in defining Ubl specificity and fidelity for downstream cascades by maintaining its cognate E2 charged with the correct Ubl. It is an appealing idea that an additional E2 specificity is achieved within the E1-E2 transthioesterification complex by the interface between E2 and the SCCH and multiple interaction interfaces between E1 and E2 collaborate during the E1-E2 transthioesterification cycle. Structural analysis of a trapped ubiquitin-E1 complex

comprising E1, E2 and thioester-linked ubiquitin will be required to further address this question.

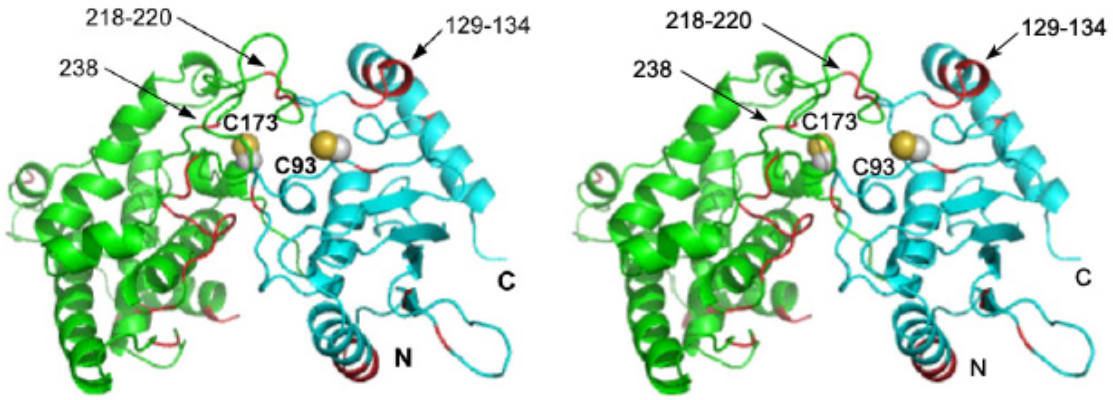


Figure 2.37 Stereo view of the NMR-derived complex between Ubc9 and the SCCH of SUMO-E1. The SCCH and Ubc9 are shown in green and cyan, respectively. Residues that show significant chemical shift changes upon complex formation are indicated in red. The two catalytic cysteines are indicated with their side chains in space-filling representation. This figure was reproduced from [108].

UFDs: Common protein-protein interaction scaffolds in Ubl pathways

The structure of the UFD of Uba1 reveals a ubiquitin-like fold and the UFD's interactions with E2 share common features with Ubls' and ubiquitin-fold proteins' interactions with their partners. For example, ubiquitin binds UIM, CUE, UBA, Npl4 Zn finger (NZF), Gga and Tom1 (GAT), ubiquitin E2 variant (UEV), and other ubiquitin-binding domains (UBDs) with its Leu8/Ile44/Val70 hydrophobic patch exposed on its β-sheet [20] (Figure 2.38). The large number of different interactions that occur between ubiquitin and UBDs or ubiquitin-pathway enzymes probably explains why the Ile44 face of ubiquitin has been highly conserved throughout evolution. The Uba1 UFD and ubiquitin-E2s are thought to make similar interactions although they do not appear to be hydrophobic.

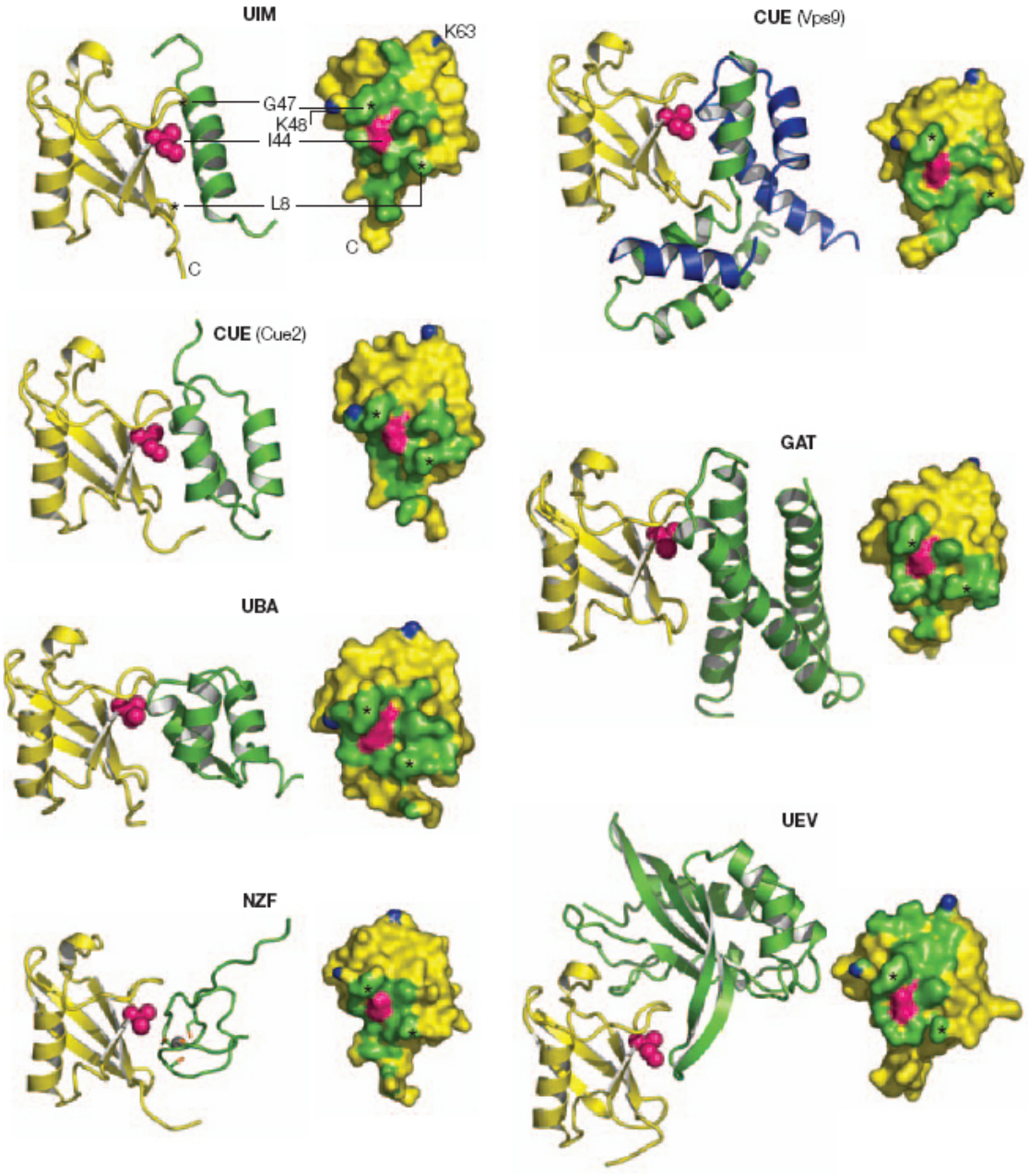


Figure 2.38 Ubiquitin-UBD complex structures. For each UBD, the left panel (ribbon diagrams) shows ubiquitin (yellow) always in the same orientation, and the UBDs are colored green (green and blue for the CUE dimer). Ile44 of ubiquitin is shown in magenta, in space-filling representation. The right panel shows ubiquitin with the UBD contact surfaces colored in green. The Ile44 side chain is again shown in magenta and is always part of the contact surface. The amino groups of Lys48 and Lys63, which are two of the key sites for polyubiquitin chain conjugation, are colored blue. Leu8 and Gly47, which move to accommodate the different UBD partners, are highlighted with asterisks. This figure was reproduced from Hicke et al [20].

A few examples of interactions between an Ubl and its binding partner raise the intriguing possibility that some of the function of the E1 UFD may be related to structural mimicry of ubiquitin and Ubls. Ubiquitin interacts with one of its E2, UbcH5, in two ways. Like other Ubls, ubiquitin forms a covalent thioester complex with UbcH5 as an intermediate in the conjugation cascade. However, ubiquitin also forms a non-covalent complex with UbcH5. The NMR studies of the complex revealed that UbcH5 recognizes the surface centered on Leu8/Ile44/Val70 of ubiquitin [40]. Interestingly, ubiquitin binds to a surface on UbcH5 formed predominantly by residues in β-strands 1-3, a surface distinct from Uba1 UFD's binding surface, which is formed by the N-terminal helix of Ubc1 (Figure 2.39). Although the non-covalent UbcH5-ubiquitin interaction does not directly mimic the E1-E2 complex, it is still notable that both E1 UFD and ubiquitin use a common surface for recognizing E2 enzymes.

In addition to the UFD of E1s, many other proteins in ubiquitin and Ubl pathways adopt structures resembling ubiquitin. As examples, the Elongin B component of the SOCS E3s resembles ubiquitin [109], as does a domain from tubulin-binding cofactor B [110] and the proteasome-associated protein Rad23 [111] (Figure 2.40). As the molecular mechanisms of UFD proteins are further resolved, it will be interesting to see whether these structures are coincidental, or whether they reflect functional mimicry.

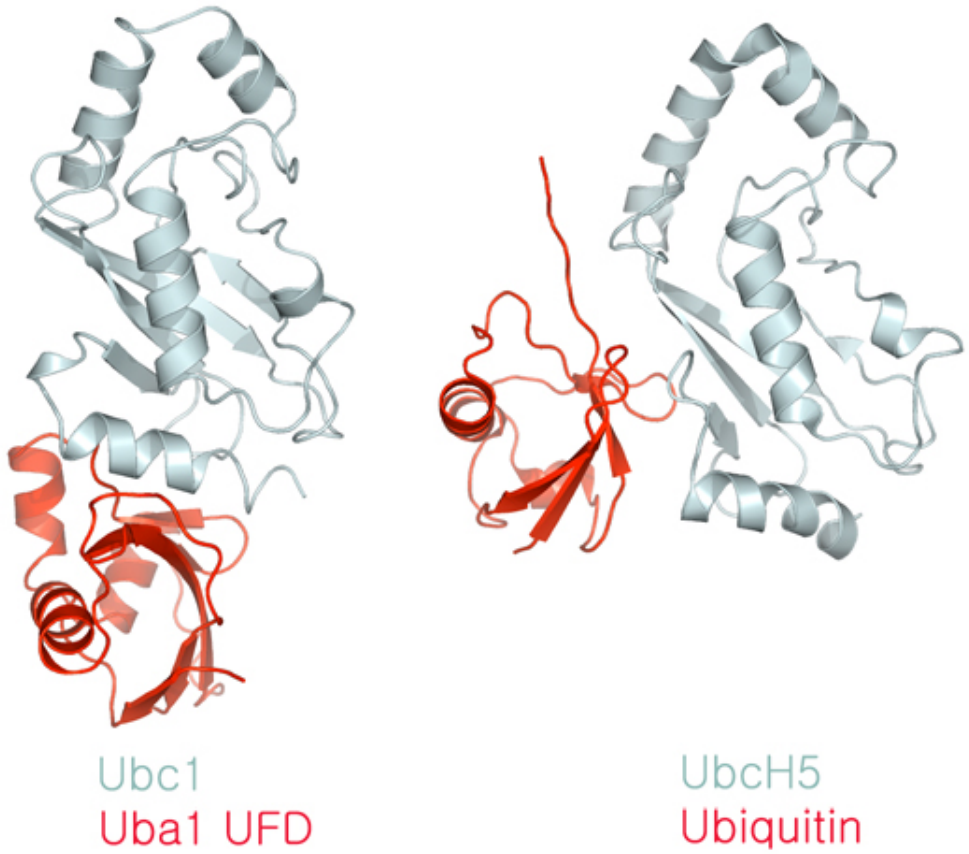


Figure 2.39 Comparison between Ubc1-Uba1 UFD and UbcH5-ubiquitin complexes with the E2s in the same orientation. Uba1 UFD and ubiquitin are shown in red and Ubc1 and UbcH5 are shown in pale cyan. The coordinates for the UbcH5-ubiquitin complex were taken from PDB entry 2FUH [40].

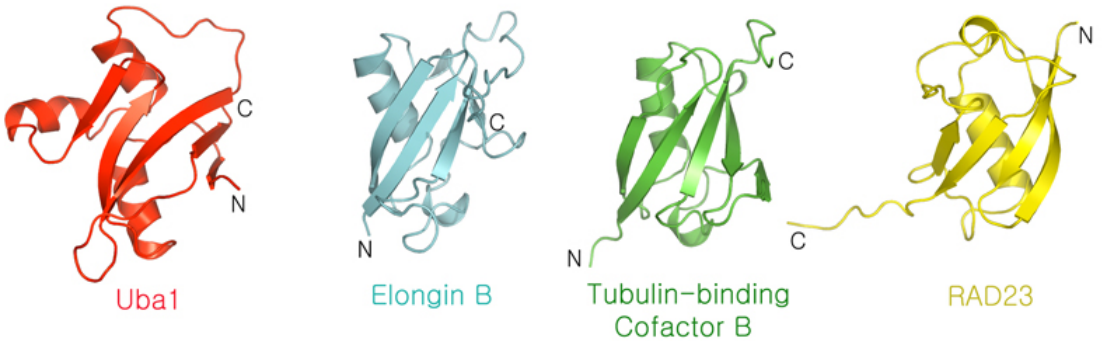


Figure 2.40 The UFD of Uba1 and its structural homologues. The structures that are displayed in this figure are: Elongin B component of the VHL-ElonginB-ElonginC complex (PDB entry 1VCP [109]); Ubl domain from tubulin-binding cofactor B (PDB entry 1T0Y [110]); Ubl domain from RAD23 (PDB entry 1P1A [112]). The N- and C-termini are labeled in each structure.

Release of the E2~Ubl thioester from E1

After the transthioesterification reaction, the Ubc1~ubiquitin thioester product is released. Elimination of the covalent tether between the thioester-linked ubiquitin and Uba1 would allow rotation of the UFD back to the conformation observed in the single ubiquitin-loaded Uba1 structure. In the NEDD8-E1 structure, superposition of the UFD-Ubc12~NEDD8 thioester model onto the APPBP1-UBA3-NEDD8 structure showed that a significant portion of NEDD8 in the Ubc12~NEDD8 thioester model clashes with the APPBP1-UBA3 surface, thus raising the possibility that the UFD rotation alone could facilitate the product release. However, such a clash is not likely in Uba1 due to the larger gap between the SCCH and UFD in the single ubiquitin-loaded Uba1 compared to NEDD8- and SUMO-E1 counterparts. One possibility is that transthioesterification could cause structural changes in the SCCH, especially the central loop region (residues 775-795), which would transmit conformational changes to the putative binding interface between the SCCH and E2, thus contributing to the dissociation of the E2.

CONCLUDING DISCUSSION

E1 enzymes are thought to have evolved from the bacterial enzymes MoeB and ThiF, which catalyze a similar adenylation reaction on respective Ubl proteins. E1s that catalyze protein conjugation have a more a complex architecture with additional domains that catalyze E1- and E2-thioester transfer following adenylation. The structure of Uba1 reveals its modular nature, which is similar to that of the E1s for NEDD8 and SUMO with activities specified by individual domains: (1) adenylation domains, (2) catalytic cysteine domains, and (3) a C-terminal domain with ubiquitin-like fold that binds E2. The Uba1 consists of a distinct arrangement of these domains connected by long, flexible linkers, creating a large groove in the middle. This groove is divided into two clefts by a crossover loop leading from the adenylation domain to the catalytic cysteine, accommodating E1's protein substrates, Ubl and E2. This architecture appears to be "primed" for large scale conformational changes. Indeed, a comparison of the two copies of Uba1 in the asymmetric unit indicates a ~10° rotation of UFD. In addition, structural comparisons between different E1s and different E1/Ubl complexes revealed that the domains are flexibly tethered to each other, with relative rotations of ~10 to 20° about the hinge loops between domains [29, 61, 80, 99].

The adenylation active site is separated from the E1 catalytic cysteine by ~35 Å in the structures of Uba1, NEDD8-, and SUMO-E1, indicating a requirement for conformational changes in the Ubl C-terminal tail and/or between the E1 adenylation domain and catalytic cysteine domains. Moreover, the E1 catalytic cysteine is exposed in

these structures, leaving the identity of residues involved in catalyzing the formation of an E1~Ubl thioester complex in question. In addition, a recent study on NEDD8 modification revealed that the thioester-bound NEDD8 is flexibly tethered to the E1, raising the possibility that a similar situation is also present in the ubiquitin- and SUMO-E1s [92]. This flexibility may be important for E1 domain rotations that accompany the formation of E1~Ubl complex and also for the subsequent transthioesterification reaction.

In addition to features shared among the E1s from different pathways, structural analysis of Uba1 has also revealed a recognition mechanism of the E1 for ubiquitin. Interactions with ubiquitin are extensive involving three different faces of ubiquitin and Uba1 features key residues contacting the Arg72 of ubiquitin, which has been described as being crucial for ubiquitin discrimination. The key questions on the mechanism of Uba1 function include the nature and timing of the Uba1 conformational changes during the steps of thioester formation with the C-terminus of ubiquitin and the details of protein-protein interactions between E1 and E2 during the transthioesterification reaction. Uba1 has a much wider groove between its SCCH and UFD compared to other E1s, and it is likely to result in a distinct UFD movement during the transthioesterification cycle. Also, the UFD and SCCH of Uba1 probably have an intrinsic affinity for its E2s, thus contributing to selectivity in E2 recruitment. However, these speculations will have to be addressed in more detail in future studies. The Uba1 structure, combined with biochemical studies detailing E1 and E2 interactions allowed us to generate a structural model of E1/E2/ubiquitin assembly that provides a framework for the design and interpretation of future experiments to address these questions.

REFERENCES

1. Goldstein G, Scheid M, Hammerling U, Schlesinger DH, Niall HD, Boyse EA: **Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells.** *Proc Natl Acad Sci U S A* 1975, **72:**11-15.

2. Goldknopf IL, Busch H: **Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24.** *Proc Natl Acad Sci U S A* 1977, **74:**864-868.

3. Kuo ML, den Besten W, Bertwistle D, Roussel MF, Sherr CJ: **N-terminal polyubiquitination and degradation of the Arf tumor suppressor.** *Genes Dev* 2004, **18:**1862-1874.

4. Cadwell K, Coscoy L: **Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase.** *Science* 2005, **309:**127-130.

5. Kirkpatrick DS, Hathaway NA, Hanna J, Elsasser S, Rush J, Finley D, King RW, Gygi SP: **Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology.** *Nat Cell Biol* 2006, **8:**700-710.

6. Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP: **A proteomics approach to understanding protein ubiquitination.** *Nat Biotechnol* 2003, **21:**921-926.

7. Elsasser S, Gali RR, Schwickart M, Larsen CN, Leggett DS, Muller B, Feng MT, Tubing F, Dittmar GA, Finley D: **Proteasome subunit Rpn1 binds ubiquitin-like protein domains.** *Nat Cell Biol* 2002, **4:**725-730.

8. Young P, Deveraux Q, Beal RE, Pickart CM, Rechsteiner M: **Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a.** *J Biol Chem* 1998, **273:**5461-5467.

9. Sun L, Chen ZJ: **The novel functions of ubiquitination in signaling.** *Curr Opin Cell Biol* 2004, **16:**119-126.

10. Liu YC: **Ubiquitin ligases and the immune response.** *Annu Rev Immunol* 2004, **22:**81-127.

11. Hicke L, Dunn R: **Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins.** *Annu Rev Cell Dev Biol* 2003, **19:**141-172.

12. Hoege C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S: **RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO.** *Nature* 2002, **419:**135-141.

13. Burroughs AM, Balaji S, Iyer LM, Aravind L: **Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold.** *Biol Direct* 2007, **2:**18.

14. Pan ZQ, Kentsis A, Dias DC, Yamoah K, Wu K: **Nedd8 on cullin: building an expressway to protein destruction.** *Oncogene* 2004, **23:**1985-1997.

15. Ohsumi Y, Mizushima N: **Two ubiquitin-like conjugation systems essential for autophagy.** *Semin Cell Dev Biol* 2004, **15:**231-236.

16. Johnson ES: **Protein modification by SUMO.** *Annu Rev Biochem* 2004, **73:**355-382.

17. Schwartz DC, Hochstrasser M: **A superfamily of protein tags: ubiquitin, SUMO and related modifiers.** *Trends Biochem Sci* 2003, **28:**321-328.

18. Pickart CM, Eddins MJ: **Ubiquitin: structures, functions, mechanisms.** *Biochim Biophys Acta* 2004, **1695:**55-72.

19. Harper JW, Schulman BA: **Structural complexity in ubiquitin recognition.** *Cell* 2006, **124:**1133-1136.

20. Hicke L, Schubert HL, Hill CP: **Ubiquitin-binding domains.** *Nat Rev Mol Cell Biol* 2005, **6:**610-621.

21. Hecker CM, Rabiller M, Haglund K, Bayer P, Dikic I: **Specification of SUMO1- and SUMO2-interacting motifs.** *J Biol Chem* 2006, **281:**16117-16127.

22. Song J, Zhang Z, Hu W, Chen Y: **Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation.** *J Biol Chem* 2005, **280:**40122-40129.

23. Reverter D, Lima CD: **Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex.** *Nature* 2005, **435:**687-692.

24. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M: **Crystal structure of thymine DNA glycosylase conjugated to SUMO-1.** *Nature* 2005, **435:**979-982.

25. Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, Bernards R: **A genomic and functional inventory of deubiquitinating enzymes.** *Cell* 2005, **123:**773-786.

26. Kerscher O, Felberbaum R, Hochstrasser M: **Modification of proteins by ubiquitin and ubiquitin-like proteins.** *Annu Rev Cell Dev Biol* 2006, **22:**159-180.

27. Vijay-Kumar S, Bugg CE, Cook WJ: **Structure of ubiquitin refined at 1.8 A resolution.** *J Mol Biol* 1987, **194:**531-544.

28. Whitby FG, Xia G, Pickart CM, Hill CP: **Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes.** *J Biol Chem* 1998, **273:**34983-34991.

29. Lois LM, Lima CD: **Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1.** *Embo J* 2005, **24:**439-451.

30. Narasimhan J, Wang M, Fu Z, Klein JM, Haas AL, Kim JJ: **Crystal structure of the interferon-induced ubiquitin-like protein ISG15.** *J Biol Chem* 2005, **280:**27356-27365.

31. Lake MW, Wuebbens MM, Rajagopalan KV, Schindelin H: **Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex.** *Nature* 2001, **414:**325-329.

32. Jin J, Li X, Gygi SP, Harper JW: **Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging.** *Nature* 2007, **447:**1135-1138.

33. Pichler A, Knipscheer P, Oberhofer E, van Dijk WJ, Korner R, Olsen JV, Jentsch S, Melchior F, Sixma TK: **SUMO modification of the ubiquitin-conjugating enzyme E2-25K.** *Nat Struct Mol Biol* 2005, **12:**264-269.

34. Zheng N, Wang P, Jeffrey PD, Pavletich NP: **Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases.** *Cell* 2000, **102:**533-539.

35. Huang DT, Paydar A, Zhuang M, Waddell MB, Holton JM, Schulman BA: **Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1.** *Mol Cell* 2005, **17:**341-350.

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

37. Zhang M, Windheim M, Roe SM, Peggie M, Cohen P, Prodromou C, Pearl LH: **Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex.** *Mol Cell* 2005, **20:**525-538.

38. Eletr ZM, Huang DT, Duda DM, Schulman BA, Kuhlman B: **E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer.** *Nat Struct Mol Biol* 2005, **12:**933-934.

39. Hamilton KS, Ellison MJ, Barber KR, Williams RS, Huzil JT, McKenna S, Ptak C, Glover M, Shaw GS: **Structure of a conjugating enzyme-ubiquitin thiolester intermediate reveals a novel role for the ubiquitin tail.** *Structure* 2001, **9:**897-904.

40. Brzovic PS, Lissounov A, Christensen DE, Hoyt DW, Klevit RE: **A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination.** *Mol Cell* 2006, **21:**873-880.

41. Jentsch S, Seufert W, Sommer T, Reins HA: **Ubiquitin-conjugating enzymes: novel regulators of eukaryotic cells.** *Trends Biochem Sci* 1990, **15:**195-198.

42. Chen Z, Pickart CM: **A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin.** *J Biol Chem* 1990, **265:**21835-21842.

43. Hodgins R, Gwozd C, Arnason T, Cummings M, Ellison MJ: **The tail of a ubiquitin-conjugating enzyme redirects multi-ubiquitin chain synthesis from the lysine 48-linked configuration to a novel nonlysine-linked form.** *J Biol Chem* 1996, **271:**28766-28771.

44. Merkley N, Shaw GS: **Solution structure of the flexible class II ubiquitin-conjugating enzyme Ubc1 provides insights for polyubiquitin chain assembly.** *J Biol Chem* 2004, **279:**47139-47147.

45. Dye BT, Schulman BA: **Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins.** *Annu Rev Biophys Biomol Struct* 2007, **36:**131-150.

46. Pickart CM: **Mechanisms underlying ubiquitination.** *Annu Rev Biochem* 2001, **70:**503-533.

47. Ozkan E, Yu H, Deisenhofer J: **Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases.** *Proc Natl Acad Sci U S A* 2005, **102:**18890-18895.

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

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

50. Hatfield PM, Callis J, Vierstra RD: **Cloning of ubiquitin activating enzyme from wheat and expression of a functional protein in Escherichia coli.** *J Biol Chem* 1990, **265:**15813-15817.

51. Zacksenhaus E, Sheinin R: **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* 1990, **9:**2923-2929.

52. McGrath JP, Jentsch S, Varshavsky A: **UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme.** *Embo J* 1991, **10:**227-236.

53. Johnson ES, Schwienhorst I, Dohmen RJ, Blobel G: **The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer.** *Embo J* 1997, **16:**5509-5519.

54. Lammer D, Mathias N, Laplaza JM, Jiang W, Liu Y, Callis J, Goebl M, Estelle M: **Modification of yeast Cdc53p by the ubiquitin-related protein rub1p affects function of the SCFCdc4 complex.** *Genes Dev* 1998, **12:**914-926.

55. Ciechanover A, Finley D, Varshavsky A: **Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85.** *Cell* 1984, **37:**57-66.

56. Finley D, Ciechanover A, Varshavsky A: **Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85.** *Cell* 1984, **37:**43-55.

57. Ciechanover A, Elias S, Heller H, Hershko A: **"Covalent affinity" purification of ubiquitin-activating enzyme.** *J Biol Chem* 1982, **257:**2537-2542.

58. Hershko A, Ciechanover A, Rose IA: **Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown.** *J Biol Chem* 1981, **256:**1525-1528.

59. Haas AL, Rose IA: **The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis.** *J Biol Chem* 1982, **257:**10329-10337.

60. Haas AL, Warms JV, Hershko A, Rose IA: **Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation.** *J Biol Chem* 1982, **257:**2543-2548.

61. Walden H, Podgorski MS, Schulman BA: **Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8.** *Nature* 2003, **422:**330-334.

62. Ciechanover A, Heller H, Katz-Etzion R, Hershko A: **Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system.** *Proc Natl Acad Sci U S A* 1981, **78:**761-765.

63. Haas AL, Warms JV, Rose IA: **Ubiquitin adenylate: structure and role in ubiquitin activation.** *Biochemistry* 1983, **22:**4388-4394.

64. Wilkinson KD, Smith SE, O'Connor L, Sternberg E, Taggart JJ, Berges DA, Butt T: **A specific inhibitor of the ubiquitin activating enzyme: synthesis and characterization of adenosyl-phospho-ubiquitinol, a nonhydrolyzable ubiquitin adenylate analogue.** *Biochemistry* 1990, **29:**7373-7380.

65. Bohnsack RN, Haas AL: **Conservation in the mechanism of Nedd8 activation by the human AppBp1-Uba3 heterodimer.** *J Biol Chem* 2003, **278:**26823-26830.

66. Hershko A, Heller H, Elias S, Ciechanover A: **Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown.** *J Biol Chem* 1983, **258:**8206-8214.

67. Burch TJ, Haas AL: **Site-directed mutagenesis of ubiquitin. Differential roles for arginine in the interaction with ubiquitin-activating enzyme.** *Biochemistry* 1994, **33:**7300-7308.

68. Pickart CM, Kasperek EM, Beal R, Kim A: **Substrate properties of site-specific mutant ubiquitin protein (G76A) reveal unexpected mechanistic features of ubiquitin-activating enzyme (E1).** *J Biol Chem* 1994, **269:**7115-7123.

69. Haas AL, Bright PM, Jackson VE: **Functional diversity among putative E2 isozymes in the mechanism of ubiquitin-histone ligation.** *J Biol Chem* 1988, **263:**13268-13275.

70. Bencsath KP, Podgorski MS, Pagala VR, Slaughter CA, Schulman BA: **Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation.** *J Biol Chem* 2002, **277:**47938-47945.

71. Rudolph MJ, Wuebbens MM, Rajagopalan KV, Schindelin H: **Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation.** *Nat Struct Biol* 2001, **8:**42-46.

72. Duda DM, Walden H, Sfondouris J, Schulman BA: **Structural analysis of Escherichia coli ThiF.** *J Mol Biol* 2005, **349:**774-786.

73. Lehmann C, Begley TP, Ealick SE: **Structure of the Escherichia coli ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis.** *Biochemistry* 2006, **45:**11-19.

74. Wang C, Xi J, Begley TP, Nicholson LK: **Solution structure of ThiS and implications for the evolutionary roots of ubiquitin.** *Nat Struct Biol* 2001, **8:**47-51.

75. Leimkuhler S, Wuebbens MM, Rajagopalan KV: **Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor.** *J Biol Chem* 2001, **276:**34695-34701.

76. Taylor SV, Kelleher NL, Kinsland C, Chiu HJ, Costello CA, Backstrom AD, McLafferty FW, Begley TP: **Thiamin biosynthesis in Escherichia coli. Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation.** *J Biol Chem* 1998, **273:**16555-16560.

77. Pitterle DM, Johnson JL, Rajagopalan KV: **In vitro synthesis of molybdopterin from precursor Z using purified converting factor. Role of protein-bound sulfur in formation of the dithiolene.** *J Biol Chem* 1993, **268:**13506-13509.

78. Rossmann MG, Moras D, Olsen KW: **Chemical and biological evolution of nucleotide-binding protein.** *Nature* 1974, **250:**194-199.

79. Walker JE, Saraste M, Runswick MJ, Gay NJ: **Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold.** *Embo J* 1982, **1:**945-951.

80. Walden H, Podgorski MS, Huang DT, Miller DW, Howard RJ, Minor DL, Jr., Holton JM, Schulman BA: **The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1.** *Mol Cell* 2003, **12:**1427-1437.

81. Handley PM, Mueckler M, Siegel NR, Ciechanover A, Schwartz AL: **Molecular cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme E1.** *Proc Natl Acad Sci U S A* 1991, **88:**258-262.

82. Swanson R, Hochstrasser M: **A viable ubiquitin-activating enzyme mutant for evaluating ubiquitin system function in Saccharomyces cerevisiae.** *FEBS Lett* 2000, **477:**193-198.

83. Pelzer C, Kassner I, Matentzoglu K, Singh RK, Wollscheid HP, Scheffner M, Schmidtke G, Groettrup M: **UBE1L2, a novel E1 enzyme, specific for Ubiquitin*.** *J Biol Chem* 2007.

84. Jones TA, Zou JY, Cowan SW, Kjeldgaard M: **Improved methods for building protein models in electron density maps and the location of errors in these models.** *Acta Crystallogr A* 1991, **47 (Pt 2):**110-119.

85. Szczepanowski RH, Filipek R, Bochtler M: **Crystal structure of a fragment of mouse ubiquitin-activating enzyme.** *J Biol Chem* 2005, **280:**22006-22011.

86. Storoni LC, McCoy AJ, Read RJ: **Likelihood-enhanced fast rotation functions.** *Acta Crystallogr D Biol Crystallogr* 2004, **60:**432-438.

87. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al: **Crystallography & NMR system: A new software suite for macromolecular structure determination.** *Acta Crystallogr D Biol Crystallogr* 1998, **54:**905-921.

88. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by the maximum-likelihood method.** *Acta Crystallogr D Biol Crystallogr* 1997, **53:**240-255.

89. Davis IW, Murray LW, Richardson JS, Richardson DC: **MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes.** *Nucleic Acids Res* 2004, **32:**W615-619.

90. Sloper-Mould KE, Jemc JC, Pickart CM, Hicke L: **Distinct functional surface regions on ubiquitin.** *J Biol Chem* 2001, **276:**30483-30489.

91. Beal R, Deveraux Q, Xia G, Rechsteiner M, Pickart C: **Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting.** *Proc Natl Acad Sci U S A* 1996, **93:**861-866.

92. Huang DT, Hunt HW, Zhuang M, Ohi MD, Holton JM, Schulman BA: **Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity.** *Nature* 2007, **445:**394-398.

93. Vijay-Kumar S, Bugg CE, Wilkinson KD, Cook WJ: **Three-dimensional structure of ubiquitin at 2.8 A resolution.** *Proc Natl Acad Sci U S A* 1985, **82:**3582-3585.

94. Sullivan ML, Vierstra RD: **Cloning of a 16-kDa ubiquitin carrier protein from wheat and Arabidopsis thaliana. Identification of functional domains by in vitro mutagenesis.** *J Biol Chem* 1991, **266:**23878-23885.

95. Pitluk ZW, McDonough M, Sangan P, Gonda DK: **Novel CDC34 (UBC3) ubiquitin-conjugating enzyme mutants obtained by charge-to-alanine scanning mutagenesis.** *Mol Cell Biol* 1995, **15:**1210-1219.

96. McKenna S, Spyracopoulos L, Moraes T, Pastushok L, Ptak C, Xiao W, Ellison MJ: **Noncovalent interaction between ubiquitin and the human DNA repair protein Mms2 is required for Ubc13-mediated polyubiquitination.** *J Biol Chem* 2001, **276:**40120-40126.

97. Moraes TF, Edwards RA, McKenna S, Pastushok L, Xiao W, Glover JN, Ellison MJ: **Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2-hUbc13.** *Nat Struct Biol* 2001, **8:**669-673.

98. VanDemark AP, Hofmann RM, Tsui C, Pickart CM, Wolberger C: **Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer.** *Cell* 2001, **105:**711-720.

99. Huang DT, Miller DW, Mathew R, Cassell R, Holton JM, Roussel MF, Schulman BA: **A unique E1-E2 interaction required for optimal conjugation of the ubiquitin-like protein NEDD8.** *Nat Struct Mol Biol* 2004, **11:**927-935.

100. Winn PJ, Religa TL, Battey JN, Banerjee A, Wade RC: **Determinants of functionality in the ubiquitin conjugating enzyme family.** *Structure* 2004, **12:**1563-1574.

101. Ramachandran GN, Sasisekharan V: **Conformation of polypeptides and proteins.** *Adv Protein Chem* 1968, **23:**283-438.

102. Chen L, Shinde U, Ortolan TG, Madura K: **Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly.** *EMBO Rep* 2001, **2:**933-938.

103. Fisher RD, Wang B, Alam SL, Higginson DS, Robinson H, Sundquist WI, Hill CP: **Structure and ubiquitin binding of the ubiquitin-interacting motif.** *J Biol Chem* 2003, **278:**28976-28984.

104. Kang RS, Daniels CM, Francis SA, Shih SC, Salerno WJ, Hicke L, Radhakrishnan I: **Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding.** *Cell* 2003, **113:**621-630.

105. Polo S, Sigismund S, Faretta M, Guidi M, Capua MR, Bossi G, Chen H, De Camilli P, Di Fiore PP: **A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins.** *Nature* 2002, **416:**451-455.

106. Prag G, Misra S, Jones EA, Ghirlando R, Davies BA, Horazdovsky BF, Hurley JH: **Mechanism of ubiquitin recognition by the CUE domain of Vps9p.** *Cell* 2003, **113:**609-620.

107. Shih SC, Katzmann DJ, Schnell JD, Sutanto M, Emr SD, Hicke L: **Epsins and Vps27p/Hrs contain ubiquitin-binding domains that function in receptor endocytosis.** *Nat Cell Biol* 2002, **4:**389-393.

108. Wang J, Hu W, Cai S, Lee B, Song J, Chen Y: **The Intrinsic Affinity between E2 and the Cys Domain of E1 in Ubiquitin-like Modifications.** *Mol Cell* 2007, **27:**228-237.

109. Stebbins CE, Kaelin WG, Jr., Pavletich NP: **Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function.** *Science* 1999, **284:**455-461.

110. Lytle BL, Peterson FC, Qiu SH, Luo M, Zhao Q, Markley JL, Volkman BF: **Solution structure of a ubiquitin-like domain from tubulin-binding cofactor B.** *J Biol Chem* 2004, **279:**46787-46793.

111. Watkins JF, Sung P, Prakash L, Prakash S: **The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function.** *Mol Cell Biol* 1993, **13:**7757-7765.

112. Ryu KS, Lee KJ, Bae SH, Kim BK, Kim KA, Choi BS: **Binding surface mapping of intra- and interdomain interactions among hHR23B, ubiquitin, and polyubiquitin binding site 2 of S5a.** *J Biol Chem* 2003, **278:**36621-36627.