University of Alberta

A Structural and Functional Characterization of Ubc13, Mms2, and Uev1a: Insights into the Mechanism of Alternative Ubiquitin Chain Synthesis

By

Trevor F. Moraes 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry

Edmonton, Alberta

Fall, 2004

Your file Votre référence
ISBN: 0-612-95989-9
Our file Notre référence
ISBN: 0-612-95989-9

My Nana, Mai.

ACKNOWLEDGEMENTS

I would like to thank the following organizations and people:

(i) The Canadian Institutes for Health Research (CIHR), the Alberta Heritage Foundation for Medical Research (AHFMR), and the University of Alberta for financial support over the last five years.

(ii) Dr. Mike Ellison and Dr. Mark Glover for giving me a research environment that was interesting and enjoyable, as well as for supervision and direction throughout my graduate career.

(iii) Members of the Ellison and Glover labs, particularly Ross and Scott for their crystallographic expertise and Sean "Hippo" McKenna, Bobby "the Badger" Varelas, and Chris Ptak for their valuable input with respect to the Ub field in experimental procedures and in the editing parts of the three manuscripts.

(iv) Dr. Wei Xiao and Landon Pastushok from the University of Saskatchewan for providing me with the Ubc13/Uev system upon which most of my research is based.

(v) Dr. Bill Plaxton for providing the a solid foundation in protein biochemistry.

(vi) My parents and siblings Frank, Ira, Theo and Sabrina for their unwavering support over 29 years.

(vii) Most importantly I need to thank Kori and Jai who make my time outside the lab the most valuable part of my life.

TABLE OF CONTENTS:

List of Tables:

List of Figures:

LIST OF SYMBOLS, NOMENCLATURE, OR ABBREVIATIONS:

3D	three-dimensional
Ala	alanine
AMP	adenosine monophosphate
Arg	arginine
Asn	asparagine
Asp	aspartate, aspartic acid
ATP	adenosine triphosphate
BCA	bicinchoninic acid
C-terminal	carboxy-terminal
cDNA	complimentary DNA
Cys	cysteine
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DUB	deubiquitinating enzyme
E1	ubiquitin activating enzyme
E2	ubiquitin conjugating enzyme
E3	ubiquitin protein ligase
E6AP	E6-associated protein
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
Gln	glutamine
Glu	glutamate, glutamic acid
Gly	glycine
GST	glutathione *S*-transferase
HECT	homolgous to carboxy-terminal domain of E6-AP
His	histidine
HSQC	heteronuclear single quantum coherence
IKK	IκB kinase
Ile	isoleucine

IPTG	isopropyl-β-D-thiogalactopyranoside
ITC	isothermal titration calorimetry
K_D	dissociation constant
LB	Luria-Bertani broth
Leu	leucine
Lys	lysine
Met	methionine
N-terminal	amino-terminal
NF-κB	nuclear factor kappa B
NMR	nuclear magnetic resonance
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PCNA	proliferating cell nuclear antigen
Phe	phenylalanine
PMSF	phenylmethylsulfonylfluoride
Poly-Ub	poly-ubiquitin
PP_i	pyrophosphate
Pro	proline
RING	really interesting new gene
RMSD	root mean square deviation
S. cerevisiae	*Saccharomyces cerevisiae*
Ser	serine
SCF	Skp1-Cullin/Cdc53-F-box protein
SDS	sodium dodecyl sulfate
ssDNA	single-stranded DNA
SUMO	small ubiquitin-related modifier
TCA	trichloroacetic acid
Thr	threonine
Trp	tryptophan

Tyr	tyrosine
*Ub	^{35}[S]-ubiquitin
Ub	ubiquitin
Ub_2	diubiquitin
Ub_3	triubiquitin
Ub_n	ubiquitin chain
Ubc	ubiquitin conjugating enzyme
Uev	ubiquitin conjugating enzyme variant
Val	valine
Wt	wildtype

Chapter 1:

Introduction

1.1 Biological Implications of Protein Ubiquitination

Post-translational modifications of intracellular proteins are used in almost every cellular pathway to initiate and propagate signals. Examples include protein acetylation[1], glycosylation[2], lipidation[3], methylation[4], and phosphorylation[5]. Protein phosphorylation is the best characterized of these modifications and has been found to have a diverse array of functions including: dictating the stability of proteins, allosteric activation, subcellular localization and moderating protein:protein interactions[5]. Modifications of proteins through acetylation, glycosylation, lipidation and methylation appears to have limited biological roles, however as more information is acquired many additional functions have been attributed to them. An excellent example of this is the post-translational modification of surface exposed Lys residues by the protein ubiquitin (Ub). Ubiquitination is best known for its function to target substrates for rapid degradation by the 26S proteasome, thereby down-regulating the activity of central regulatory proteins[6;7]. The degradation of poly-ubiquitinated proteins is a critical component within many different signaling pathways including cell cycle control[8;9], transcription[10], apoptosis and the immune response[11]. Not surprisingly, defects in Ub-dependent proteolysis have been implicated as a causative factor in cancers and are being researched in anti-cancer therapeutics [12;13]. Furthermore, numerous inherited diseases including Alzheimer's disease (AD), Parkinson's disease (PD), amylotrophic lateral sclerosis (ALS) and

Huntington's disease (HD) have also been linked to defects in the ubiquitin proteasome system[14].

Recently, a subset of the ubiquitination field has been explored that involves non-proteasomal processes that can be signaled by mono-ubiquitination or by non-canonical poly-Ub chains. Mono-ubiquitination is involved in a variety of processes from membrane transport to transcriptional regulation[15]. The formation of non-canonical poly-Ub chains is the focus of this dissertation.

1.2 The Ubiquitin System- The E1, E2, E3 cascade.

The mechanism utilized to covalently conjugate Ub to proteins requires three enzymatic steps in which Ub is passed sequentially as an activated thiolester from a Ub activating enzyme (E1) to an Ub conjugating enzyme (E2), and finally, with the help of an Ub ligase (E3), to a target[7] (**Figure 1.1**). In the first step, the E1 adenylates the carboxy-terminus of Ub (Gly76) activating it for the formation of a thiolester bond to the active-site Cys of E1[16]. In the second step, Ub is transferred from the E1 to the active-site Cys of an E2 in a transthiolesterfication reaction. Lastly the E2, often in coordination with an E3, facilitates the formation of an isopeptide bond between the carboxy-terminus of Ub and a lysine ε-amino group on a target protein. In a step-wise manner, the conjugated-Ub can then be the site of further ubiquitination on its surface lysine residues creating a poly-Ub chain on the original target[17] (**Figure 1.1**)[16].

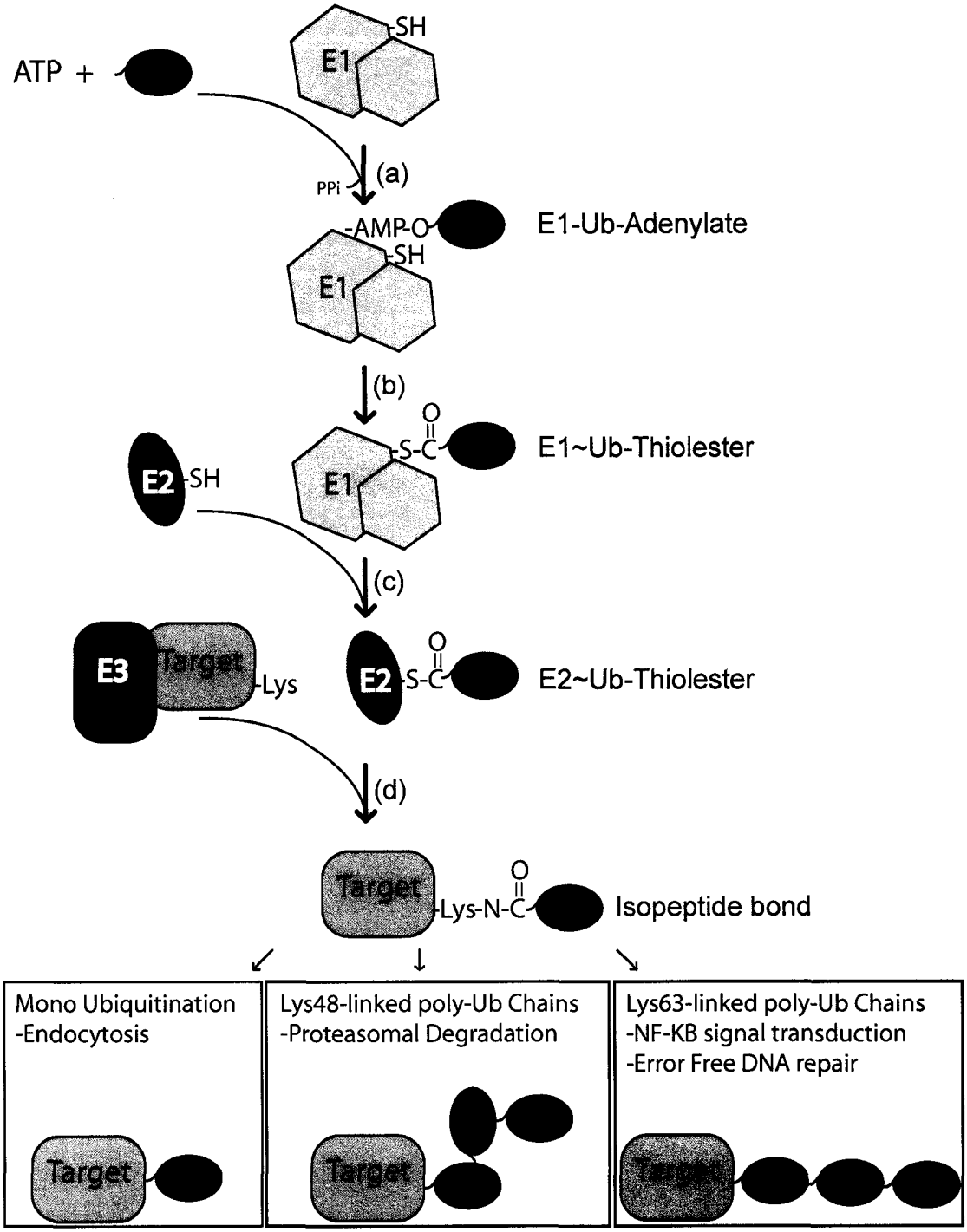


Figure 1.1 The Protein Ubiquitination Cascade. (a) The E1 (green) is responsible for the ATP-dependent activation of Ub (red), resulting in the formation of an E1-Ub~adenylate. (b) Ub is then transferred to the active site Cys of the E1, creating an E1~Ub thiolester. (c) Ub is subsequently transferred to the active site Cys of an E2 (blue) in a transthiolesterification reaction. (d) With the aid of an E3 (purple), Ub forms an isopeptide bond with a surface-exposed Lys on the target substrate. Three different poly-Ub chain linkages and their biological outcomes are shown.

Within the ubiquitin system there exists only one type of E1 (Uba1) and the selective targeting of specific substrates is facilitated by the existence of multiple E2s (13 known in *S. cerevisiae*, >19 in humans)[16;17] and an even larger number of E3 enzymes[16;17]. Each E2 can interact with a number of E3s and together, the type of Ub modification (either mono-Ub or topology of poly-Ub chains) can be specified[18]. There also exist de-ubiquitinating (DUB) enzymes that can disassemble poly-Ub chains undoing the action of the E1/E2/E3 cascade[19]. Together, the E1, E2, and E3 proteins are responsible for the ubiquitination of a vast number of proteins.

1.3 Ubiquitin

Ub is a highly conserved 76 amino acid protein found in all eukaryotes[7]. Other than the flexible carboxy-terminal residues including the catalytic Gly76, the core domain of ubiquitin is extremely well ordered and tightly packed through a network of hydrogen bonds[16;17]. Within the core domain, Lys48 is the site of canonical proteasomal poly-Ub chain formation[20]. There are six other Lys residues in the core domain, four of which have also been implicated in poly-Ub chain assembly (Lys6, Lys11, Lys29, and Lys63)[21-25]. The primary sequence of Ub, including the aforementioned residues, is very highly conserved throughout eukaryotes (only 3 residues differ between *S.cerevisiae* and Human at positions 19, 24, 28)[26] presumably conserving specific and distinct surfaces that are required for the interaction with E1, E2, E3 and other Ub related machinery. Indeed, this is the case as surfaces on Ub have been identified for mediating endocytosis[27], interaction with the proteasome[28],

interaction with E1 enzymes[29], association with E2 enzymes[30;31] and interacting with Uevs[32](**Figure 1.2A**)[33].

Ub has a highly ordered ββαββαβ fold that has been classified as a superfold based on its structural homology with numerous other proteins[34]. Although some family members are completely unrelated to Ub or Ub-related processes, certain members namely "Ub-like" proteins, appear to function in an analogous manner to the Ub-conjugation system. There are 11 Ub-like small molecules which can be covalently attached to amino residues post-translationally[17], of which SUMO and Nedd8 are best characterized.

1.4 Ubiquitin-like Proteins, SUMO and Nedd8

The small Ub-related modifier SUMO (sentrin/Smt3p) has 18% sequence homology to Ub. The overall structure is similar to Ub with the exception that SUMO contains of a long and flexible amino-terminus (**Figure 1.2C**)[35;36]. Furthermore, the canonical lysine side chain for poly-Ub chain assembly, Lys48 is Gln69 in SUMO. Given the mechanism of Ub chain assembly this accounts for the observation that only a single SUMO molecule, as opposed to chains, are found conjugated onto target substrates[37;38]. SUMO modification (sumoylation) is accomplished by an E1 homologue, AOS1-UBA[39], and a SUMO specific E2, Ubc9[40]. Sumoylated proteins appear to be primarily localized to the nucleus or nuclear envelope. Some examples include: PML, a RING finger-containing E3 enzyme, involved in nuclear localization[41], and RanGAP1, the Rab GTPase-activating protein, in which sumoylation directs it to the nuclear pore

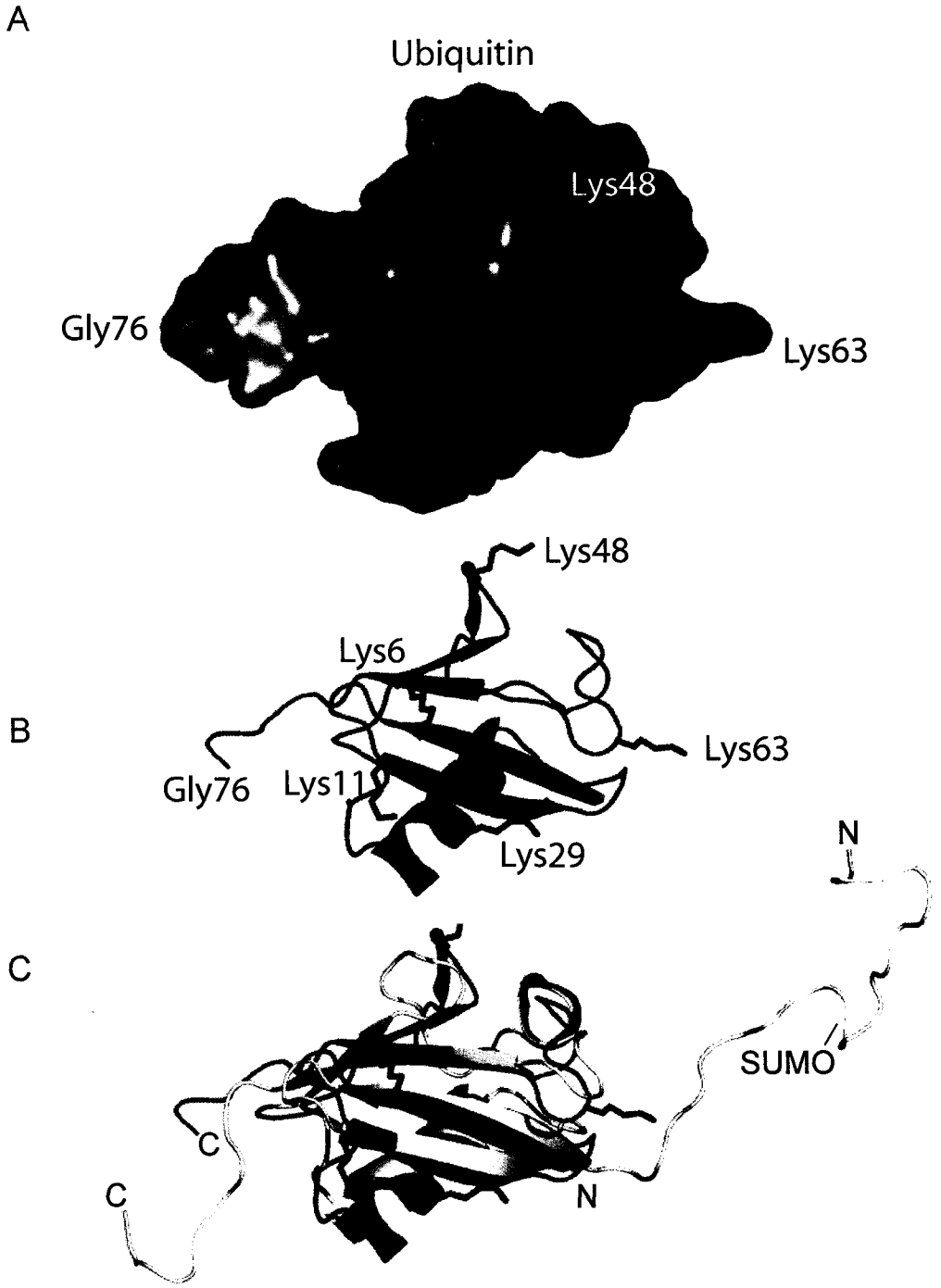


Figure 1.2 The Structure of Ubiquitin. **A.** The surface representation of Ub (shaded in red) with residues that are involved in binding E2s (grey) and UEVs (purple). **B.** Ub secondary structure displayed as a cartoon representation. Principal residues involved in the formation of poly-Ub chains have been labelled, including the carboxy- terminal Gly76 and surface exposed Lys6, Lys11, Lys29, Lys48 and Lys63 residues. **C**. An overlay of Ub with the Ub-like protein SUMO.

complex[42]. Furthermore given that sumoylation of target proteins at Lys residues involved in ubiquitination, it has been hypothesized that sumoylation antagonizes ubiquitination, thereby affecting the activity of crucial regulatory proteins. A particular example involves the proliferating cell nuclear antigen, PCNA, whose sumoylation blocks ubiquitination and hence DNA repair activity[38].

Another well characterized Ub-like protein, Nedd8 (Rub1), shares 60% sequence homology to Ub. The crystal structure of Nedd8 closely resembles that of Ub [43;44], however, a key amino acid residue difference (Ub-R72; Nedd8-A72) near the carboxy-terminus of Nedd8 fails to react with the E1 enzyme specific for Ub activation (UBA1) allowing for the possibility of involvement in alternative pathways[43;44]. Nedd8 protein modification employs an E1 dimer, APPBP1-UBA3[39;45], and a Nedd8 specific E2, Ubc12[45;46]. The similarities between the Ub-like proteins provides a wealth of information regarding the Ub system as structures of analogous participants in the Nedd8 and SUMO pathways have been determined, complimenting regions of the Ub system that have yet to be explored.

1.5 Ub in Chains.

1.5.1 Traditional Poly-Ubiquitin chains.

Cellular proteins that are conjugated with Lys48-linked poly-Ub chains are typically recognized and rapidly degraded by the 26S proteasome[17]. The structures of Lys48-linked poly-Ub chains of various lengths have been studied using crystallography[47-49] and nuclear magnetic resonance (NMR)[50]. Various

conformations of the chains were observed suggesting that Lys48-linked poly-Ub chains appear to be dynamic, with the Ub molecules behaving as rigid units connected by flexible tails. Furthermore, these dynamic chains appear to alternate between two distinct conformations, "open" and "closed". The "closed" conformation may serve to sequester key hydrophobic residues (Leu8, Ile44, and Val70) required for interaction and recognition by the proteasome[50]. These results have led to the hypothesis that proteins that recognize poly-Ub chains, bind these chains in the "open" conformation, thus stabilizing the interaction.

1.5.2 Alternative Poly-Ubiquitin Chains.

Although Lys48-linked poly-Ub chains comprise the predominant Ub chain conformation, recent studies have shown that alternative Lys residues within Ub (at positions 6, 11, 29, and 63) can also serve as sites for Ub chain formation[21]. Many of these alternative chains do not serve to target substrates for degradation, but rather function in activation/inactivation or localization. Recently the BRCA1-BARD1 complex has been shown to catalyze the formation of Lys6-linked poly-Ub chains onto BRCA1 *in vivo*[21]. The BRCA1-BARD1-mediated Lys6-linked poly-Ub chains are recognized and deubiquitinated by 26S proteasome but do not signal for target degradation. UbcH5A, a 120 kDa protein complex that behaves as an E1, E2, and E3 can synthesize unanchored Lys29-linked poly-Ub chains[51;52]. In addition, enzymes that catalyze the formation of Lys11 poly-Ub chains have also been observed, but their roles remains to be discovered.

Lys63-linked poly-Ub chains represent the best characterized non-canonical chains and have been shown to play a critical role in a variety of processes. Lys63 chains are necessary in the post-replicative DNA repair pathway[22] and in the function in NF-κB signal transduction[53], and ribosome function[54]. Ubiquitination and endocytosis of several yeast plasma membrane proteins requires Lys63-linked poly-Ub chains, ultimately leading to their down-regulation via lysosomal degradation[55]. Lys63-linked chains have also been implicated in the stress response[24] and mitochondrial DNA inheritance[56].

The existence of poly-Ub chains linked through surface exposed Lys residues other than Lys48 suggests that mechanisms exist which allow for the differential recognition of specific poly-Ub chain types. Unfortunately, no proteins have been shown to specifically recognize these alternative poly-Ub chain linkages other than for the purpose of deubiquitinating them[21;57].

1.6 Catalytic Mechanism of Ubiquitination.

1.6.1 E1 -Ubiquitin Activating Enzyme.

The ubiquitin-activating enzyme, or E1, catalyzes the first step in the ubiquitination cascade in which it must first form a thiolester intermediate with Ub in order to efficiently transfer Ub to an E2 enzyme[17]. The initial chemistry of the E1 reaction involves the hydrolysis of ATP in order to form an Ub-adenylate at the carboxy-terminus of Ub Gly76[58;59]. The adenylate subsequently donates the Ub to the conserved active-site Cys of E1 forming a thiolester intermediate[58-60]. Finally, the

E1~Ub thiolester undergoes a second round of adenylate formation, yielding a complex which contains two molecules of activated Ub[60]. The activation of Ub is a very efficient process ensuring that the production of activated~Ub is not the rate-limiting step in the ubiquitination process as there are higher concentrations of total E2s versus E1 in cells. Studies have show that this process is extremely efficient; the ATP binding event to transthiolesterification (Ub transfer from E1 to the E2) occurs at rates that are 10- to 100- fold faster than the entire catalytic rate of substrate ubiquitination[60].

Recent insight into the molecular framework of E1 enzymes has come from the crystal structures of MoeB-MoaD[61] and APPBP1-UBA3[62;63]. The MoeB-MoaD protein complex is involved in the *Escherichia coli* (*E. coli*) molybdenum cofactor biosynthesis pathway. These proteins function in an evolutionarily conserved manner to Ub activation in eukaryotes by E1 (MoeB) and Ub (MoaD) proteins[64]. The crystal structures of the MoeB-MoaD complex in its apo, ATP-bound, and MoaD-adenylate forms provide important insights, particularly with respect to the details of the adenylation reaction: ATP binds to a nucleotide binding pocket and a conserved Asp residue co-ordinates a Mg^{2+} ion required for ATP hydrolysis[61]. In addition, the conserved nature of the hydrophobic protein-protein interactions between MoeB and MoaD suggests a model for the formation of the E1 adenylate ternary complex.

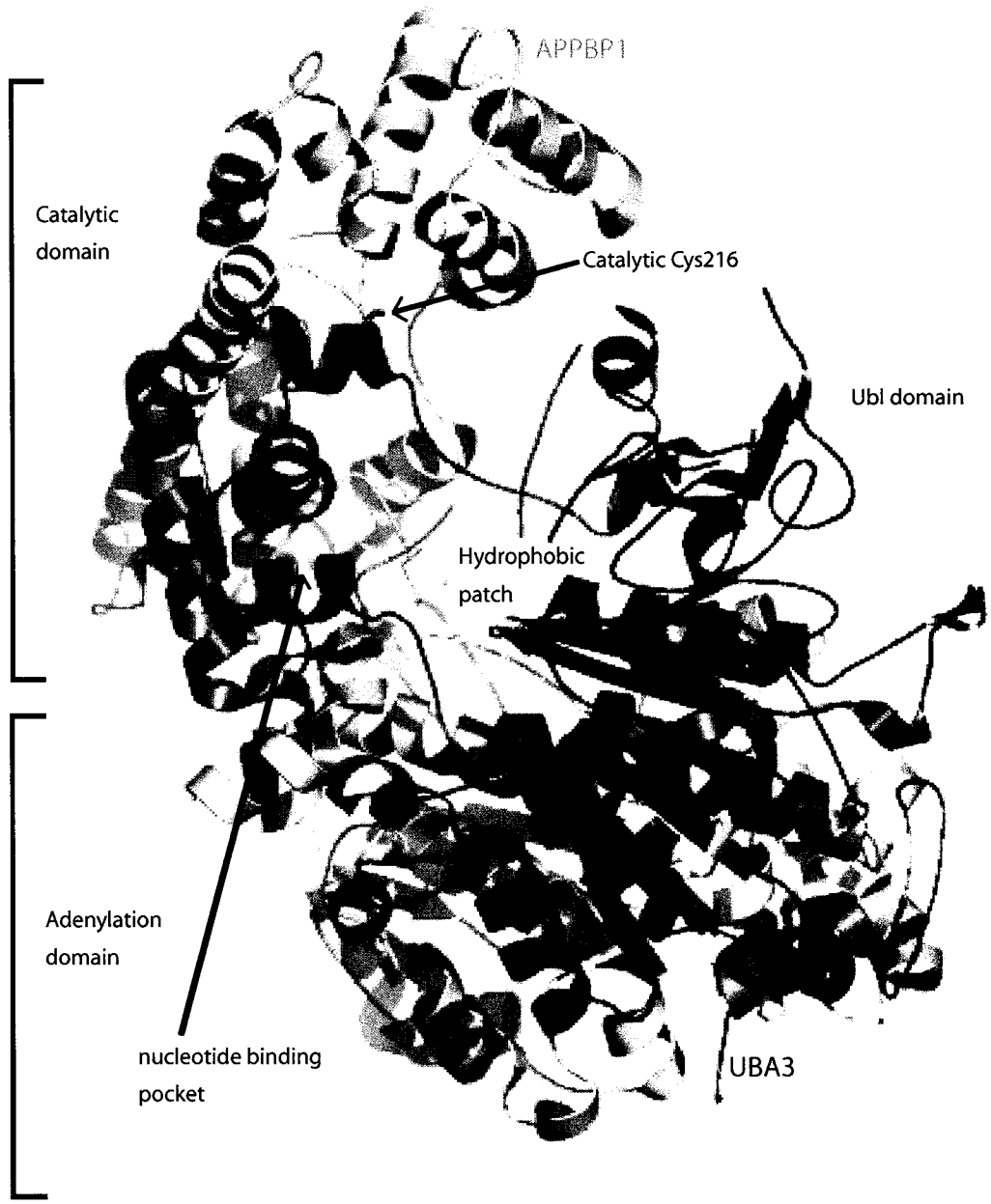


Figure 1.3 The structure of an E1 complex. The structure of the APPBP1-UBA3 complex illustrated in ribbon diagrams, APPBP is yellow, whereas UBA3 is purple. Domains and the catalytic Cys216 are highlighted. The adenylation, catalytic Cys, and Ub-like domains together contribute to a broad, deep groove in the structure that is divided into two clefts which coordinate nucleotide and protein binding.

The structure of APPBP1-UBA3, a heterodimeric E1 enzyme for the Ub-like protein Nedd8, revealed that E1 activity is specified by unique domains (**Figure 1.3**). The APPBP1 (amino-terminal half of the ubiquitin E1) acts as an adenylation domain which resembles the MoeB enzyme from bacteria (*103*), whereas UBA3 (the carboxy-terminal half of ubiquitin E1) contains the catalytic domain with the active-site Cys that is responsible for thiolester formation, and a domain which is responsible for E2 recognition that has structural similarity to the structure of Ub[62]. These domains contribute to the formation of single groove with two adjacent clefts that facilitate the coordination of a nucleotide and Nedd8/Ub binding to the E1 in such a manner that each reaction drives the subsequent reaction in an "assembly-line" fashion. Together these studies complete a structural depiction of the three functional steps of the E1: Ub-adenylate formation, E1~Ub thiolester formation, and E2 binding[18;65].

1.6.2 E2 -Ub Conjugating Enzyme

All E2 enzymes conjugate Ub through a covalent thiolester bond via transfer from E1. The common features of E2s, include a conserved active-site Cys and core domain of 150 residues (Figure 1.4 and 1.5). The molecular structures of E2s determined throught X-ray crystallography[66-72] and NMR[73] have confirmed that the fold of the core E2 domain is highly conserved amongst family members, consisting of a four-stranded anti-parallel β-sheet flanked by four α-helices and a short 3-10 helix. The active-site Cys sits on a loop that connects β4 to the 3–10 helix and is

shadowed by residues of the α2-α3 loop, together creating a shallow catalytic groove (**Figure 1.4 and Figure 2.1**). The most highly conserved residues amongst the E2 family members cluster to regions surrounding the active site (**Figure 1.5**). The most divergent residues are found on the surface of the E2 opposite the active-site, for which specific biological functions have not been clearly defined. An obvious distinguishing feature of E2s are the presence or absence of amino- and carboxy-terminal extensions, which have been used to categorizes E2s into four classes: type I (Ubc4, Ubc5, and Ubc13 in *S. cerevisiae*) lack extensions, type II (Ubc1 and Ubc8) have carboxy terminal extensions, type III (Ubc7, Ubc9, Ubc11, and Ubc12) have amino terminal extension, and type IV (Rad6, Cdc34, and Ubc6) have both carboxy- and amino-terminal additions to the core domain[74]. Furthermore, the presence and amino acid composition of two loop region insertions centered at residues 35 and 110 are also thought to play an important role in mediating the diverse actions of the E2s[16;75;76].

1.6.3 E2 Interactions with Ub, Substrates and E3s

The interactions of E2s with other protein components of the Ub system have also been explored by NMR and X-ray crystallography. The covalent E2~Ub thiolester[30;31;77] or E2~Ub oxyester (which involves the mutation of the active-site Cys to Ser to create a more stable thiolester analog)[78] intermediates have been analyzed by solution NMR experiments, illustrating that the carboxy-terminal tail of Ub and regions surrounding the active-site Cys of the E2 represent the principal areas

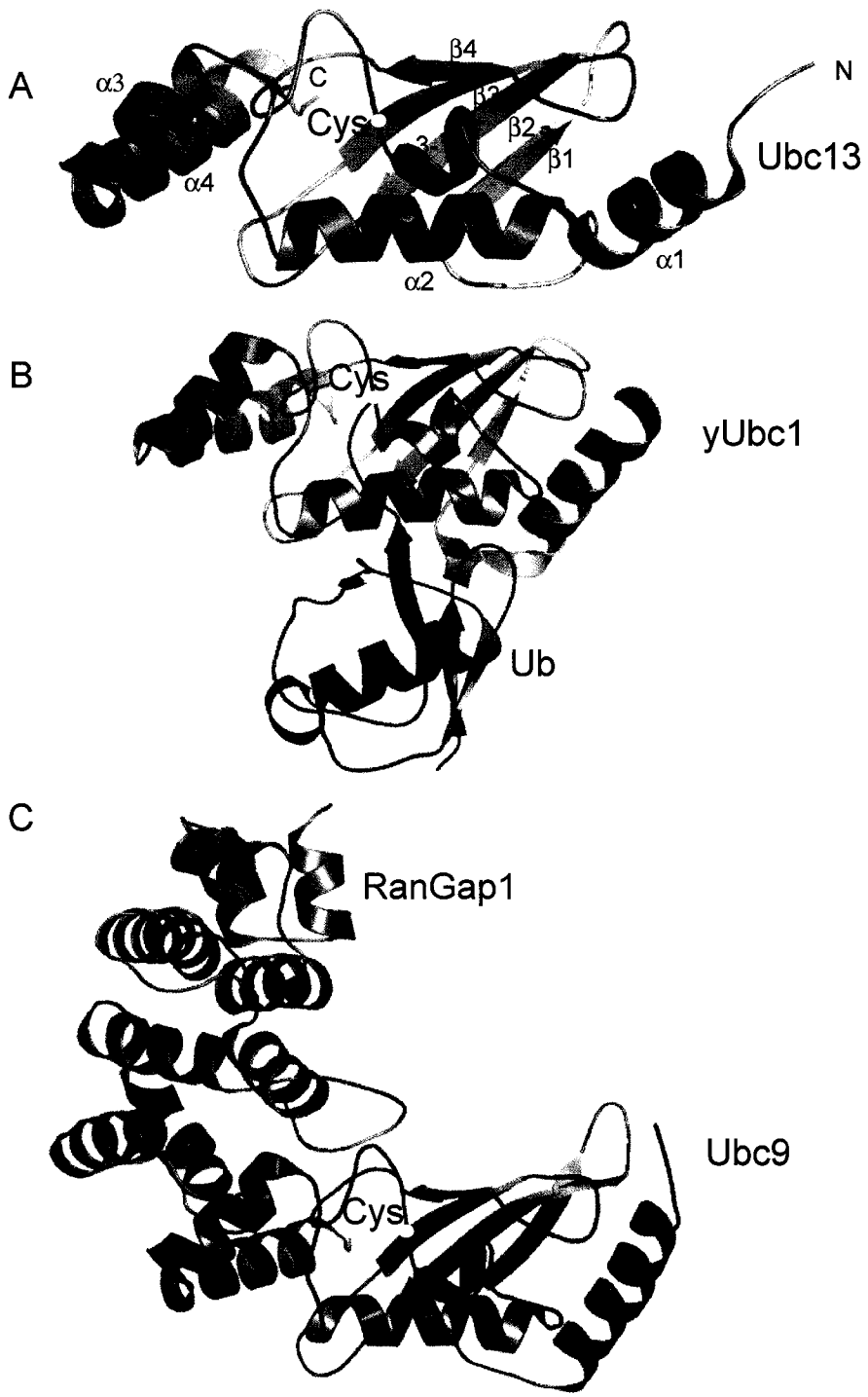


Figure 1.4 Structural properties of E2 enzymes. Three different examples of the conserved E2 fold are shown, including **(A)** Ubc13, **(B)** yeast Ubc1 protein which is modelled with the thiolester linked Ub, and **(C)** yeast Ubc9 bound to substrate RanGap1. Each E2 consists of a four-stranded antiparallel β-sheet flanked by four α-helices, labeled accordingly in the case of Ubc13. The active site Cys is also indicated in each case.

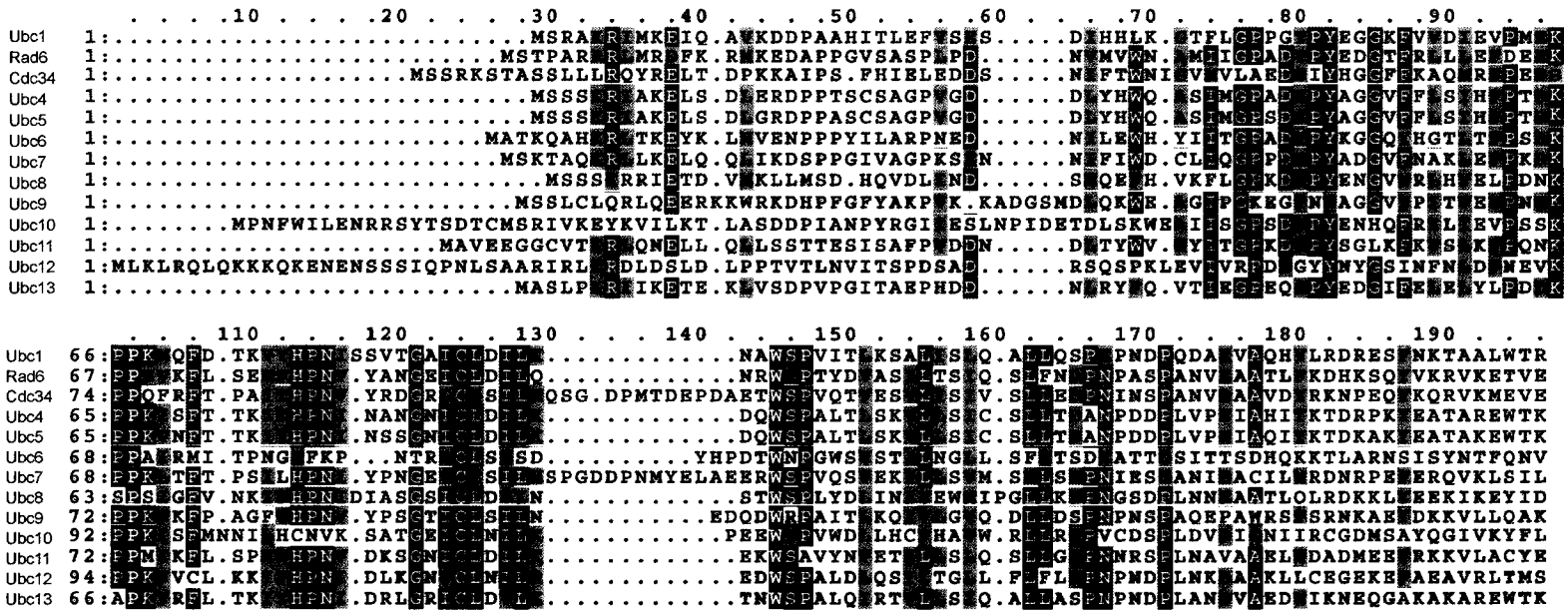


Figure 1.5 Sequence alignments of the core domains from S. cerevisiae E2 enzymes. The core E2 domain without C-terminal extensions from Ubc1, Rad6, CDC34, Ubc4, Ubc5, Ubc6, Ubc7, Ubc8, Ubc9, Ubc10, Ubc11, Ubc12, and Ubc13, were used for alignment using Indonesia multiple sequence alignment software. Residues with significant sequence similarity are boxed in black and moderately conserved residues are boxed in grey. Residue numbering for each E2 is shown on the left . The active-site Cys is highlighted witha red box. A structural alignment of selected E2s and Uevs can be found in figure 2.2.

of interaction. Further NMR chemical shifts changes in the E2 and Ub upon E2~Ub thiolester formation provide surfaces on Ub and the E2 that are likely the sites of interactions. This data was used to create a model of the E2~Ub thiolester[77;30] (**Figure 1.4B**)[31]. E2s are also involved in many non-covalent protein-protein interactions with targets and E3s; examples of each have been determined through X-ray crystallography. The direct interaction between E2s and target proteins is best demonstrated by the crystal structure of Ubc9 in complex with a large fragment of its substrate RanGAP1[37] (**Figure 1.4C**). Ubc9 makes specific contacts with a sequence recognition motif (tetrapeptide: *-K-x-D/E where *=hydrophobic and K is the sumoylated lysine residue) for sumoylation that positions the ε-amino group of the lysine residue 3.5 Å from the active-site cysteine. The substrate lysine approaches the active site from the side opposite to the covalently attached Ub, as described above.

The crystallographically determined crystal structures of an E2 (UbcH7) bound to an E3 ligase (c-cbl and E6-AP)[79;80] demonstrate that the specificity of this interaction is dictated by only a few residues in the E2 while utilizing two loops congregating at one end of the E2 structure (near the amino- terminus)[81] (**Figure 1.7**)[18]. Other non-covalent interactions have suggested the self-association and heterodimerization of E2s [76;82-86], these results become particularly interesting as non-covalent interaction between Ub or Ub-like modifiers and their cognate E2s have also been observed[78;87♣]. The crystallographic data have allowed for the development of

♣ Personal communication S.A. McKenna.

a hypothesis that E2-E2 associations, potentially associated with a specific E3, may somehow provide a scaffold upon which poly-Ub chains are built.

The ultimate function of an E2, after accepting ubiquitin from E1, is to transfer Ub onto a target or on a HECT-E3 (described below). E2 enzymes can often autoubiquitinate on specific Lys residues [30;33;85;88;89]. Autoubiquitination may represent a non-specific transfer of Ub to a nearby primary amine, but it is also possible that this modification represents a mechanism whereby the E2 activity may be regulated[90]. The E2 mechanism utilized to form a isopeptide bond either onto itself or a target protein is the subject of Chapter 4.

1.6.4 UEVs -Ub Conjugating Enzyme Variants.

Ubiquitin-conjugating enzyme variants (UEVs) are atypical E2s that share significant sequence identity with E2s, but lack the characteristic active-site Cys residue required for thiolester formation[91-94]. In spite of the inability to form a thiolester linkage with Ub, some UEVs (Mms2, Uev1a, and Tsg101/Vps23p) possess the ability to bind Ub non-covalently, and function as either co-factors in poly-Ub chain formation or as Ub sensors[82;91;92;95;96]. Recent evidence has demonstrated that Mms2 forms a heterodimer with an E2 (Ubc13), and together they function in the error-free DNA postreplicative repair mechanism in both humans and yeast[82;93;95;97;98]. Another UEV, Uev1a (Croc1)[99], is involved in NF-κB activation, and also forms a heterodimer with Ubc13[53;100;101]. These E2-UEV heterodimers possess the unique catalytic ability to synthesize non-canonical Lys63-linked poly-Ub chains[82;102;103].

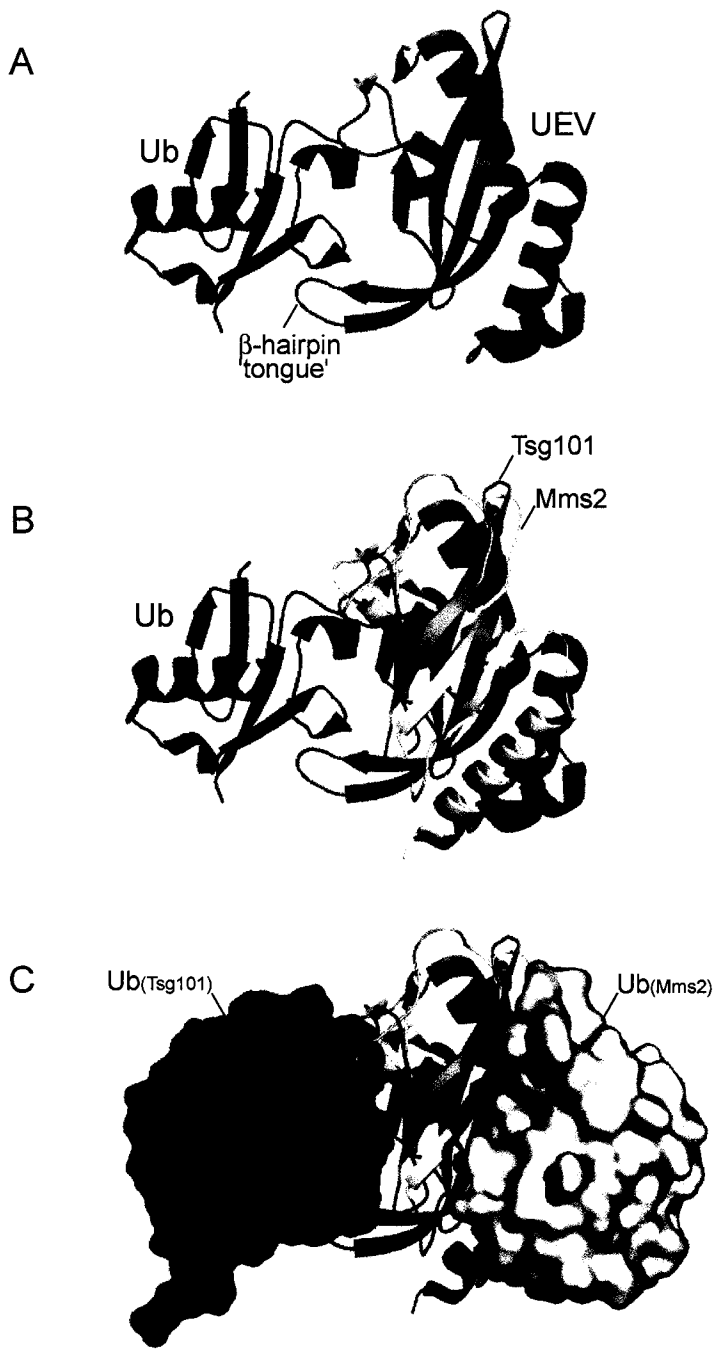


Figure 1.6 Comparison of UEV structures and Ub binding sites. (**A**) Crystal structure of Tsg101 (green) bound to Ub (red) where the β-hairpin "tongue" is labelled. (**B**) The crystal structure of hMms2 (yellow) is superimposed onto the structure of Tsg101 (aligned using the program Pymol). (**C**) Ub binding sites determined by crystallography and NMR are modelled onto Tsg101 and Mms2.

A third UEV, Tsg101 (Vps23p in *S.cerevisiae*), shares sequence similarity with Mms2 and Uev1a in its amino-terminal region, whereas the remaining two-thirds of the protein are involved in interactions with other proteins[91](**Figure1.6**)[92;104;105]. Tsg101, plays a central role in the cellular vacuolar protein sorting pathway, and is responsible for the sorting of membrane-associated proteins through a series of endosomal compartments prior to delivery to the lysosome[106]. Tsg101 does not interact with an E2 to form poly-Ub chains, but instead recognizes and binds mono-ubiquitinated substrates for eventual incorporation into the lysosome[96;107]. The solution and crystal structures of the Tsg101 UEV domain highly resemble those of the Mms2 (Chapter1) and Uev1a (Chapter2) crystal structures, with the exception that Tsg101 possesses an extended β-hairpin that links β-strands 1 and 2. All three UEVs illsutrate conserved Ub binding activity, remarkably they bind through very different interfaces [104] (**Figure 1.6**)[32;105]. Mms2 and Uev1a use a face that permits contact with Ubc13 in order to place the Lys63 of Ub close to the active-site of Ubc13[31;33;66], whereas Tsg101 uses the extended β-hairpin region allowing it to independently bind PTAP peptide ligands[32]. Although UEVs share similar structure, sequence and an affinity for Ub, they have adapted unique binding surfaces. A phylogenetic analysis of UEVs and E2 illustrates distinguishes the two but demonstrates a solid clustering of all UEV sequences within the E2 family tree[99].

1.6.5 Substrate Specific E3 Ligases.

The E3 enzyme complexes function to recognize both substrate and the E2 required for poly-Ub chain formation[17;18;108]. In order to provide substrate specificity there are a number of E3s most of which can be divided into two families, HECT and RING finger E3s[108]. RING finger E3s act as a scaffold for both the E2 thiolester intermediate and the substrate such that transfer of Ub from the E2 to the substrate can procede[109]. On the other hand, HECT domain E3s contain a catalytic Cys residue which accepts Ub from an E2 prior to acting as the proximal Ub donor to substrate[110]. Thus, E3 proteins serve whether directly or indirectly to catalyze the transfer of activated Ub onto the target substrate or the elongating poly-Ub chain.

RING finger E3s are further subcategorized into multi-subunit RING E3s and single polypeptide E3s. The crystal structure of the multi-subunit SCF ubiquitin ligase has recently been determined[111;112]. The SCF complex orients target proteins through a complicated scaffold arrangement where the RING finger protein Rbx1 has been implicated in binding the cognate E2 and a variable F-box protein presents the specific target to the E2 active-site[111;112]. Additionally, the crystal structure of the single subunit RING E3, c-cbl, bound to an E2 (UbcH7) gives an example of the surfaces of the RING and the E2 that are involved in this association[79](**Figure 1.8A**). Through a RING domain interaction, monomeric RING E3s are able to bind to their cognate E2 and present substrates for ubiquitination via a substrate-binding domain. The RING motif is generally defined by a conserved pattern of eight cysteine and

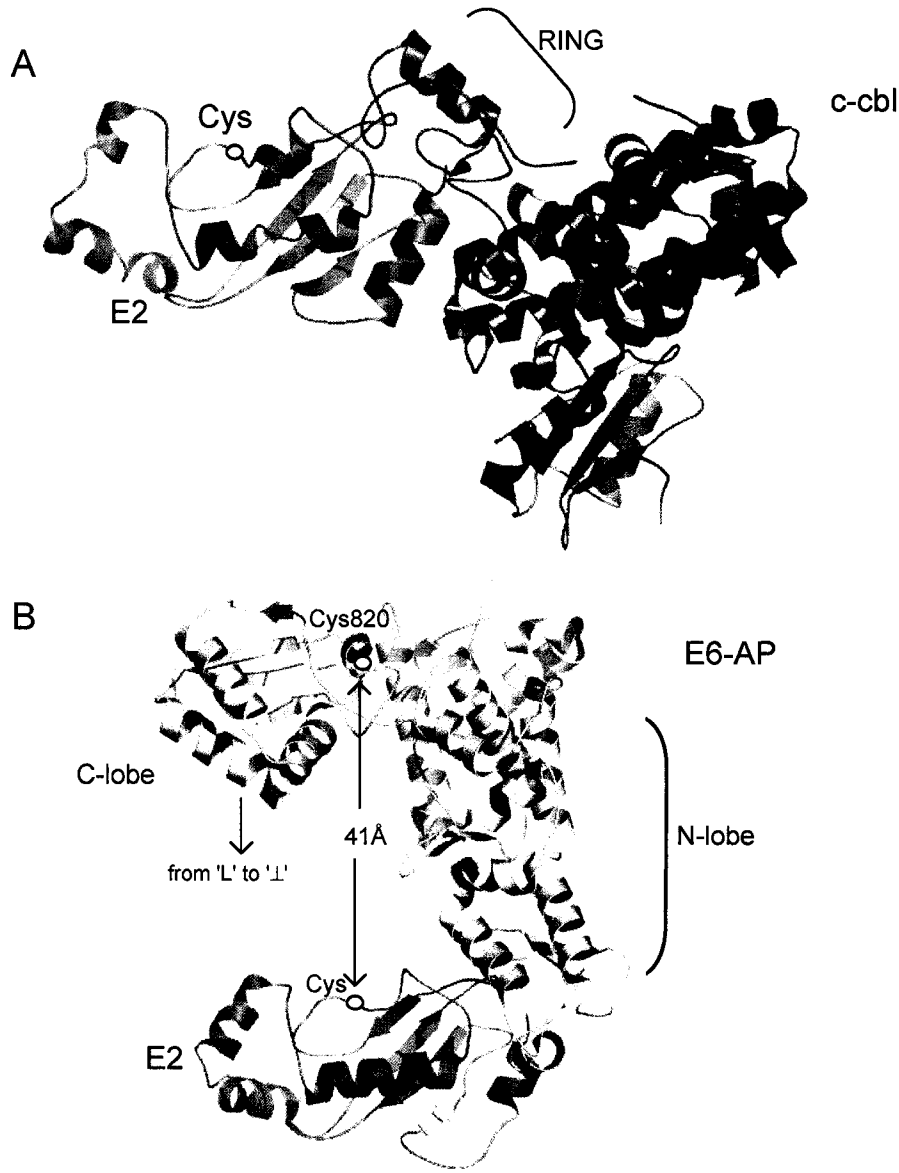


Figure 1.7 HECT and RING E3 interactions with an E2. A. The interaction between UbcH7 (E2, blue) and c-Cbl (RING E3, purple) is shown, with the c-Cbl RING domain indicated. **B.** The interactions between UbcH7 (E2, blue) and the E6AP HECT domain (gold) where the N-lobe, and C-lobe are indicated. The active-sites on each protein are indicated with a yellow sphere.

histidine residues: Cys-X_2-Cys-X_{9-39}-Cys-X_{1-3}-His-X_{2-3}-Cys/His-X_2-Cys-X_{4-48}-Cys-X_2-Cys, where the X can be any amino acid. The cysteines and histidines are arranged in pairs in the primary sequence and commonly bind two Zn^{2+} atoms in an interleaved fashion, forming two separate Zn^{2+} binding sites[113;114]. Proteolysis studies have shown that the ~50-residue RING motif of BRCA1 and BARD1 are part of a larger ~100-residue RING domain[115]. Furthermore, the residues flanking the RING motif of BRCA1 form anti-parallel α-helices that are involved in the dimerization of RING domains. The RING domain of c-cbl and BRCA1 provide a shallow groove into which the amino-terminal helix and first two loops of an E2 can bind[79;♦].

The HECT (homologous to the E6-AP carboxy-terminus)[116] E3s contain a 350 residue carboxy-terminal domain that includes a strictly conserved active-site Cys[117]. This Cys is the site of a third high-energy thiolester intermediate within the Ub cascade. All HECT E3s use a similar catalytic mechanism and typically require Ubc5/6/7 as their accompanying E2[80;116;118]. Specific substrate specificity is determined by the HECT E3 amino terminal binding domain[17]. The structure of a HECT E3 bound to an E2 ubiquitination complex (E6-AP•UbcH7) was the first E2-E3 structure and provided important insights into the mechanism and specificity of interaction mediated by HECT E3s[80](**Figure 1.7B**). The catalytic HECT domain of E6AP is composed of two lobes (amino- and carboxy-terminal, N-Lobe and C-Lobe, respectively) with a broad catalytic cleft at the junction of the two lobes creating an 'L' shape. The E2 binds the amino-terminal lobe such that the corresponding E2-E3

♦ Personal communication from R. Klevit

complex is U-shaped with the highly-conserved active-site Cys820 of E6AP sitting 41 Å from the catalytic Cys of the E2[80]. Another recent structure of a HECT E3, WWP1/AIP5, suggests an alternate conformation of the HECT domain, where instead of adopting an 'L' shape, the carboxy terminal lobe packs against the middle of the amino-terminal lobe creating an inverted T (⊥)[119]. These structures demonstrate the large-scale structural rearrangement necessary for efficient transthiolesterification of Ub from E2 to E3.

1.7 Ubc13, Mms2 and Uev1a Biological Function.

1.7.1 Ubc13•Uev1a in NF-κB Signal Transduction.

NF-κB (nuclear factor-kappaB) comprises dimeric transcription factors of the Rel family of DNA-binding proteins that recognize a common sequence motif. NF-κB is found in essentially all cell types. In response to infections, inflammation, and other stresses, NF-κB regulates a wide variety of genes including cytokines, chemokines, adhesion molecules, acute phase proteins, inducible effector proteins, and regulators of apoptosis and cell proliferation, which together play a crucial role in the recognition of pathogens by innate or adaptive immune responses[120-122]. NF-κB is normally sequestered in the cytoplasm of non-stimulated cells and consequently requires translocation to the nucleus to carry out it's function. The subcellular location of NF-κB is controlled by a family of inhibitory proteins IκBs, that bind NF-κB and mask its nuclear localization signal, thereby preventing nuclear uptake. Exposure of cells to a variety of extracellular stimuli including tumor necrosis factor-

α (TNF-α) and interleukin-1 (IL-1), leads to an intricate signal cascade resulting in the rapid phosphorylation, ubiquitination, and ultimately proteolytic degradation of IκB, that frees NF-κB to translocate to the nucleus where it regulates gene transcription (**Figure 1.8**).

It is within the NF-κB intricate signal cascade, the Uev1a•Ubc13 heterodimer plays an early and non-destructive role. IKK activation is required for the downstream phosphorylation of IκB. This is accomplished by a protein complex that includes the protein kinase TAK1[100;123-125]. In turn, TAK1 activation has been demonstrated to be dependent upon a RING-finger E3 protein ligase, TRAF6[53;100], which oligomerizes on the cytoplasmic side of the cell membrane upon stimuli of extracellular receptors[126]. Most importantly, TAK1 activation by Traf6 occurs only after non-canonical Lys63-linked poly-Ub chains have been conjugated onto Traf6 [53;100;127]. The mechanism by which these Lys63-linked chains are responsible for activation of TAK1 has yet to be determined. The E2-UEV heterodimer Ubc13•Uev1a is responsible for the alternative chain assembly onto an oligomerized TRAF6. Furthermore, these chains do not lead to degradation by the 26S proteasome. The RING-finger domain of TRAF6 is required for the ubiquitination event[53] and is responsible for interaction with Ubc13[100;128]. Recently, Lys63 ubiquitination of TRAF2 by the Ubc13•Uev1a complex has been shown to be redundant to TRAF6 poly-ubiquitination in this pathway[129]. In B and T cells, the pathway is somewhat more complicated, the regulatory subunit of IKK (NEMO) is also ubiquitinated with Lys63-linked poly-Ub chains by Ubc13•Uev1a[101]. Lys63 linked poly-Ub chains in

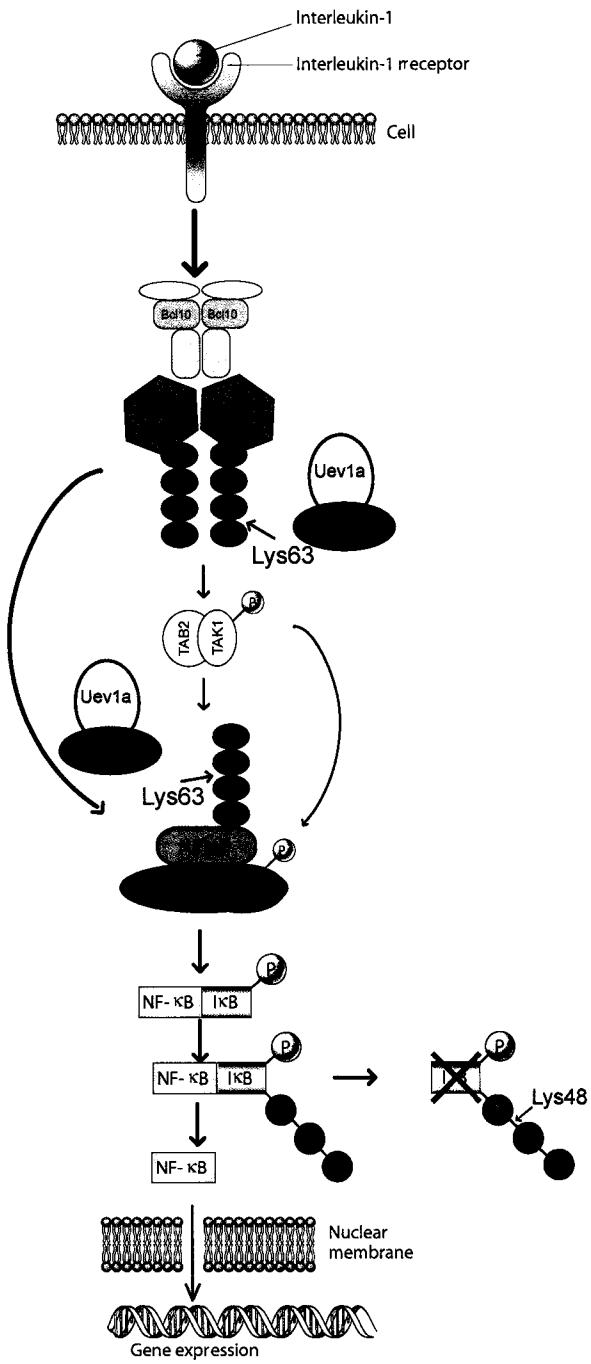


Figure 1.8 The Role of Uev1a in NF-κB Signal Transduction. In response to IL-1 binding to its receptor, a set of adaptor proteins bind to the receptor oligomerize and recruit TRAF6 (E3) molecules to also oligomerize. Traf6 oligomerization allows for it to become modified by Lys63 poly-Ub chains, which is dependent on Uev1a/Ubc13 an UEV/E2 complex. Lys63 poly-Ubiquitinated Traf6 activates the TAK1 kinase leading to a signal cascade and the eventual release of the NF-κB transcription factor into the nucleus.

the NF-κB pathway do not signal for proteasomal destruction; rather they function to activate the catalytic domain of a kinase (TAK1). Therefore the assembly of Lys63-linked poly-Ub chains by Uev1a•Ubc13 represents the propagation of a signal cascade needed to activate NF-κB pathways through non-proteasomal mechanisms.

1.7.2 Ubc13•Mms2 in the Rad6-Dependent Error-Free DNA Repair Pathway.

The Mms2 gene was first identified in knockout screens that resulted in sensitivity to DNA damaging agents such as methyl methanesulphonate MMS[130]. More recently, a null *mms2* phenotype in *S. cerevisiae* demonstrated that Mms2 functions in the error-free component of the Rad6 DNA repair pathway[93]. The eukaryotic genome has a number of repair mechanisms to cope with environmental or spontaneous DNA damage including: direct repair, base excision repair, nucleotide excision repair, double-strand break repair, and cross-link repair[131-133]. In *S. cerevisiae*, these DNA repair processes can be subdivided into three classes based on genetic epistatic analyses of DNA repair mutants: the *Rad3*, *Rad6*, and *Rad52* groups[133;134]. The *Rad52* group of mutants demonstrates responsibility for double-stranded break repair through mitotic or meiotic recombination[135] and nucleotide excision repair is mediated by the *Rad3* pathway[136]. Chromosomes are monitored and DNA lesions are usually repaired via these nucleotide or base-excision repair mechanisms, however in case of moderate to extreme DNA damage these pathways become overwhelmed, leading to lesions that persist during DNA replication. Stalled replication forks are responsible for the generation of single-strand DNA (ssDNA),

and these lesions can be lethal[137]. At these lesions, the post-replicative repair proteins within the *Rad6* group perform their function [134;138].

Rad6-dependent post-replicative repair of the ssDNA breaks functions either through an error-prone or error-free mechanism. The error-prone mechanism employs the activity of Polζ (Rev3), a mutagenic polymerase that fills in damaged regions with low fidelity[139], whereas the error-free mechanism functions through the activity of Polδ (Pol30) which uses sequence information of the undamaged strand at the replication fork to repair the lesion[134]. The error-free and prone pathways are strongly dependent on Rad6 and its function as a Ub-conjugating enzyme (Ubc2)[140;141]. Recent work has demonstrated that a target of this E2 is the proliferating cell nuclear cell antigen (PCNA)[38]: a trimeric, ring-shaped protein complex that functions as a sliding clamp[142] and processivity factor for DNA polymerases, including Polδ [134]. PCNA may be modified in three different ways by the protein ubiquitination machinery, and each modification may alter its function in DNA repair by recruiting different repair components to the DNA (**Figure 1.9**).

Rad6 is recruited to regions of DNA damage by an interaction with Rad18, a RING-finger E3 ssDNA-binding protein[98;143]. Together, there proteins are responsible for the mono-ubiquitination of PCNA at Lys164[38]. This mono-ubiquitination event is hypothesized to target stalled replication forks to initiate error-prone DNA repair[38;144]. The second ubiquitination event occurs after the initial Ub is ligated to PCNA at Lys164. At this point additional Ub molecules can be sequentially tethered onto the initial Ub via non-canonical Lys63-linkages, through

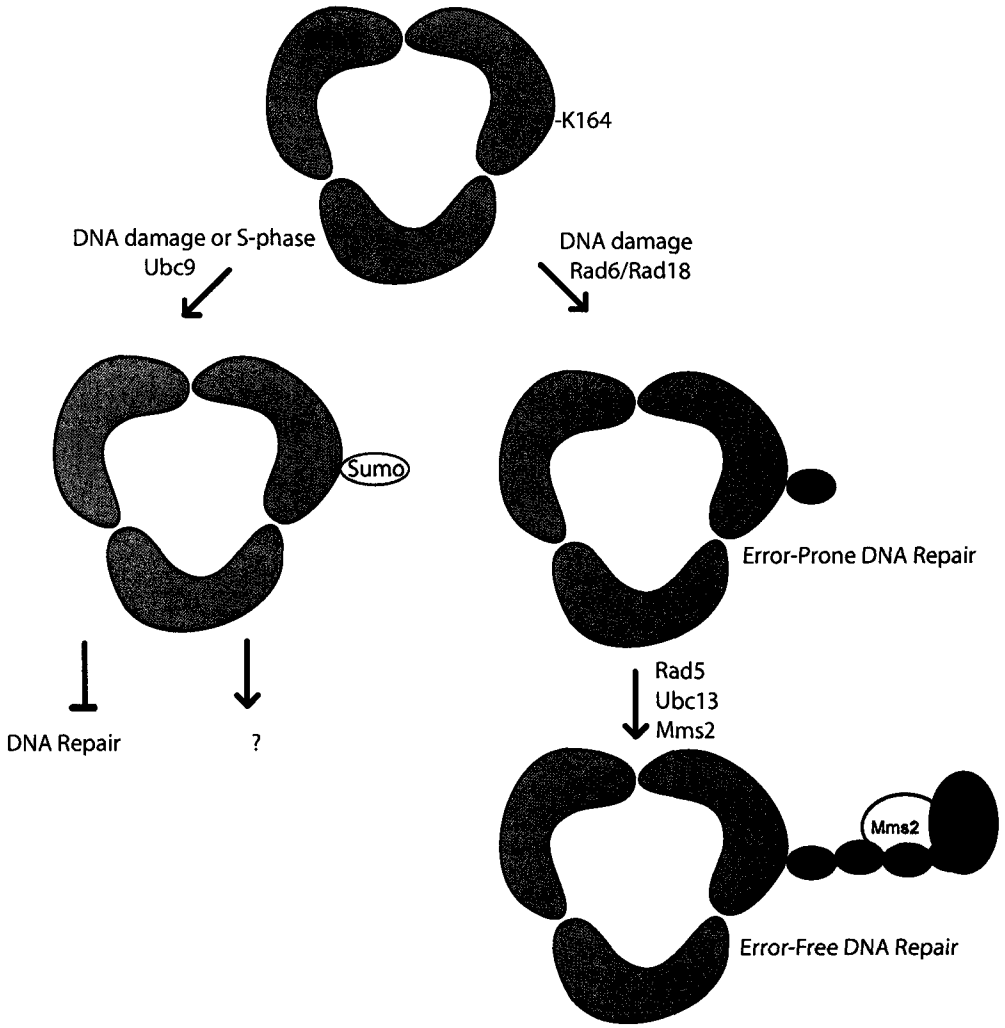


Figure 1.9 The Role of Mms2 in DNA Repair Pathways. PCNA forms a trimeric ring which binds DNA and acts as a processivity factor for DNA polymerases. Upon DNA damage, PCNA is monoubiquitinated on Lys164 in a reaction which is dependent on the Rad6/Rad18 (E2/E3) proteins. Monoubiquitinated PCNA has been suggested to target stalled replication forks to initiate error-prone repair. The Ub on PCNA may be further modified by Lys63-linked poly-Ub chains in an Mms2/Ubc13/Rad5 (Uev/E2/E3) dependent manner. This modification causes the induction of error-free DNA repair pathways. In a competing reaction, Lys164 may be modified by SUMO in both normal and damaged cells during S phase, and requires Ubc9 (E2). The specific role of each type of ubiqutination is unknown at present.

the action of the heterodimer Ubc13•Mms2 and Rad5, a RING-finger E3 DNA-binding protein which recruits the heterodimer to ssDNA[98;145]. Mms2, Ubc13 and Rad5 are crucial for the initiation of error-free post-replicative DNA repair, which accounts for previous observations that the proteolytic activity of proteasomes is not required for the DNA repair function[82]. The interplay between the two Ub conjugation activities (Rad6 and Ubc13/Mms2) is further mediated by physical interactions which have been observed between the two RING-finger containing ssDNA binding proteins Rad18 and Rad5[98;145]. Although the role of this alternative Ub chain remains to be determined (*e.g.* polymerase recruitment), the discovery of PCNA as a target represents the first protein substrate recognized and ubiquitinated by components of the Rad6 pathway.

Finally, PCNA is also modified in a third manner by SUMO on precisely the same residue (Lys164). This post-translational modification occurs via another E2 Ubc9 and is thought to signal normal DNA replication via inhibition of ubiquitination of PCNA[38]. Thus the role of Mms2 and Ubc13 in the error-free branch of the DNA repair process may be critical, it is subject to substantial regulation that is still the subject of much discovery.

1.8 A Comparison of the Mechanisms of Ubiquitination and Acetylation.

Lysine side chains are not only the site of ubiquitination and Ub-like modifications but also serve as sites for acetylation and methylation which raises the interesting possibility that there maybe interplay between these modifications in

certain biological processes [146-150]. The post-translational modification of Lys residues through ubiquitination or acetylation both result in the synthesis of an isopeptide bond that is generated from the reaction between the lysine ε-amino group and a high-energy thiolester[18](**Figure 1.10**). Therefore, knowledge of the catalytic mechanism of acetylation by the histone acetyl transferases (HATs) should provide insights into mechanism of isopeptide bond formation within ubiquitination[18]. HATs use acetyl-CoA as the acetyl donor that is activated and transferred to lysine residues via the formation of a high-energy thiolester linkage to coenzyme A in an analogous fashion to the activation of the carboxy-terminus of Ub by a thiolester bond. Furthermore, there is a subclass of HATs that transfer the acetyl group to a catalytic Cys residue before transfer to the substrate's Lys that is strikingly similar to the transfer of Ub from E1 to E2 (or E2 to HECT E3)[151]. Isopeptide bond formation in both the ubiquitination and acetylation reactions proceeds through nucleophilic attack by the ε-amino group onto the electrophilic carbon center of the thiolester bond covalently linking Ub or acetyl-coA[18]. Based on the structure of HAT GCN5 in complex with a histone H3 peptide and coenzyme A (equivalently the E2, Ub and substrate), a mechanism has been proposed in which a Glu activates a water to act as general base abstracting a proton from the ε-amino group of the Lys. The nucleophilic Lys can then attack the electrophilic carbon creating a tetrahedral intermediate and an oxyanion that is stabilized by a main chain amide group[152]. Although E2s and HATs are both single domain proteins with a catalytic cleft housing their respective active sites that engage both their peptide substrates and

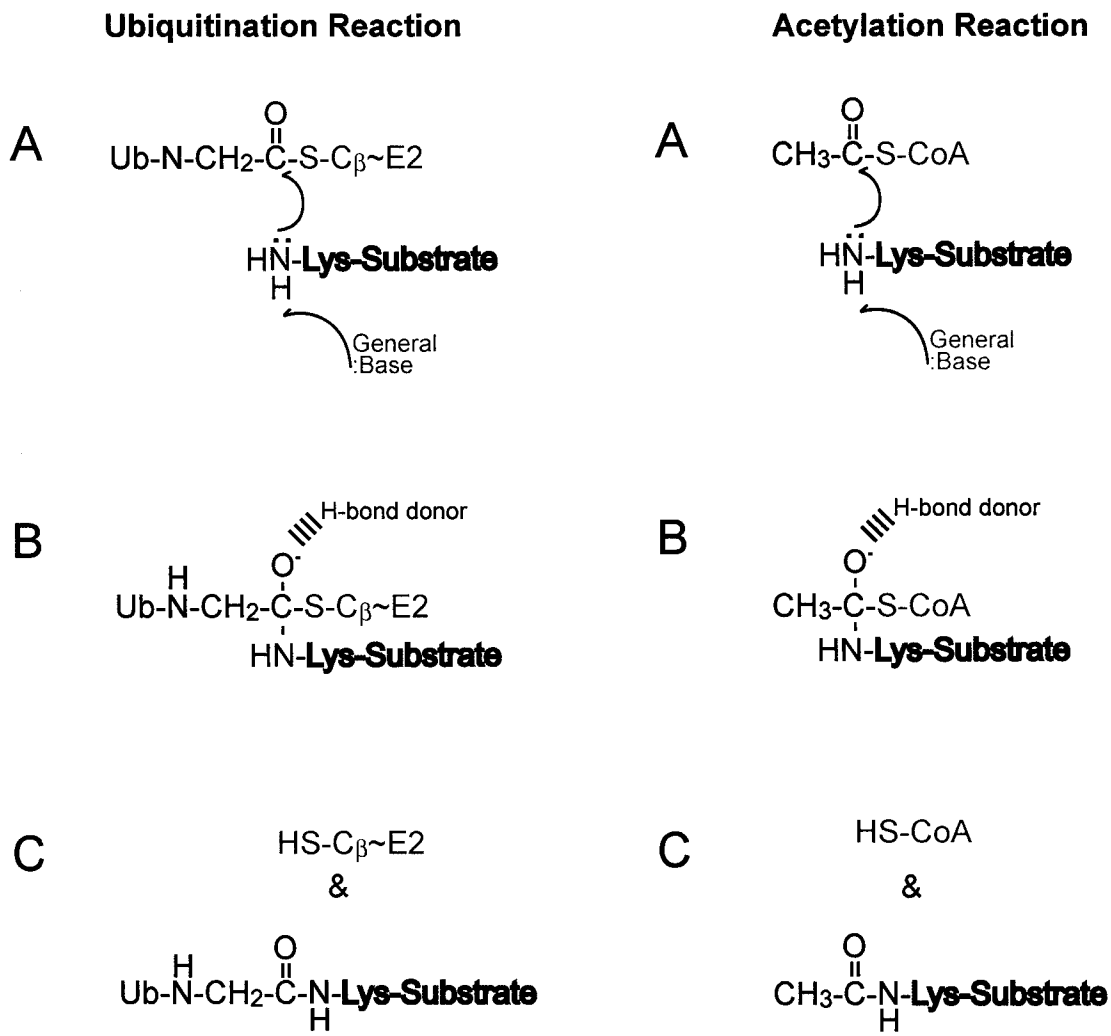


Figure 1.10 A Schematic of Lysine Ubiquitination and Acetylation: Isopeptide Bond Synthesis. Comparison of the ubiquitination reaction catalyzed by an E2/E3 and the acetylation reaction performed by histone acetyl transferaces. Similarities include: **A.** The role of a general base to create a nucleophilic ε-amino group on the substrate Lys. **B.** The formation and stabilization of a tetrahedral indemediate via hydrogen bonding to the oxyanion. **C.** The synthesis of an isopeptide bond linking either Ub or an acetyl group to a substrate Lysine.

donor Ub or acetyl groups respectively, their catalytic sites and secondary structural topology are quite different[18]. Therefore the activation of the substrate lysine ε-amino group and the stabilization of the tetrahedral transition state intermediate is probably achieved through a variation of this mechanism. Even within the HATs there exist variations on the mechanism of acetylation[153] and we would expect to see differences in the mechanism of isopeptide bond formation between the four classes of E2s.

1.9 Overview:

The focus of this dissertation is the determination of the mechanism of alternative K63-linked poly-Ub chains. formation specifically with respect to Lys63 linked Ubs. Lys-63 linked poly-Ub chains are synthesized through the enzymatic activity of the mms2-ubc13 heterodimer. Structural information regarding Ubc13 and Mms2 was obtained by X-ray crystallography (Chapter 2). This structural determination was followed up by structural and functional comparisons of Mms2 and Uev1a, two proteins of the Ub pathway that distinguish between the functions of DNA repair and NF -κB pathways (Chapter 3). Finally, X-ray crystallography was used in combination with biochemical and functional analysis to probe catalytic residues required for Lys63 linked isopeptide bond formation within the Mms2•Ubc13 heterodimer (Chapter 4). These results will be discussed in Chapter 5, with respect to the mechanism of Lys63 poly-Ub chain formation and within a general framework for apoly-Ub chain catalysis.

1.10 Reference List

1. Kouzarides, T. (2000) *EMBO J* **19,** 1176-9
2. Haltiwanger, R. S. and Lowe, J. B. (2004) *Annu Rev Biochem* **73,** 491-537
3. Casey, P. J. (1995) *Science* **268,** 221-5
4. Szyf, M., Pakneshan, P., and Rabbani, S. A. (2004) *Cancer Lett* **211,** 133-43
5. Johnson, L. N. and Lewis, R. J. (2001) *Chem Rev* **101,** 2209-42
6. Pickart, C. M. (1997) *FASEB J* **11,** 1055-66
7. Hershko, A. and Ciechanover, A. (1998) *Annu Rev Biochem* **67,** 425-79
8. Reed, S. I. (2003) *Nat Rev Mol Cell Biol* **4,** 855-64
9. King, R. W., Deshaies, R. J., Peters, J. M., and Kirschner, M. W. (1996) *Science* **274,** 1652-9
10. Muratani, M. and Tansey, W. P. (2003) *Nat Rev Mol Cell Biol* **4,** 192-201
11. Jesenberger, V. and Jentsch, S. (2002) *Nat Rev Mol Cell Biol* **3,** 112-21
12. Adams, J. (2004) *Cancer Cell* **5,** 417-21
13. Robinson, P. A. (2004) *Curr Protein Pept Sci* **5,** 163-76
14. Petrucelli, L. and Dawson, T. M. (2004) *Ann Med* **36,** 315-20
15. Hicke, L. (2001) *Nat Rev Mol Cell Biol* **2,** 195-201
16. Pickart, C. M. (2001) *Annu Rev Biochem* **70,** 503-33
17. Glickman, M. H. and Ciechanover, A. (2002) *Physiol Rev* **82,** 373-428
18. Passmore, L. A. and Barford, D. (2004) *Biochem J* **379,** 513-25
19. Wing, S. S. (2003) *Int J Biochem Cell Biol* **35,** 590-605
20. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) *Science* **243,** 1576-83
21. Nishikawa, H., Ooka, S., Sato, K., Arima, K., Okamoto, J., Klevit, R. E., Fukuda, M., and Ohta, T. (2004) *J Biol Chem* **279,** 3916-24

22. Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995) *Mol Cell Biol* **15,** 1265-73

23. Johnson, E. S., Ma, P. C., Ota, I. M., and Varshavsky, A. (1995) *J Biol Chem* **270,** 17442-56

24. Arnason, T. and Ellison, M. J. (1994) *Mol Cell Biol* **14,** 7876-83

25. Baboshina, O. V. and Haas, A. L. (1996) *J Biol Chem* **271,** 2823-31

26. Ozkaynak, E., Finley, D., Solomon, M. J., and Varshavsky, A. (1987) *EMBO J* **6,** 1429-39

27. Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M., and Hicke, L. (2001) *J Biol Chem* **276,** 30483-9

28. Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M., and Pickart, C. (1996) *Proc Natl Acad Sci U S A* **93,** 861-6

29. Burch, T. J. and Haas, A. L. (1994) *Biochemistry* **33,** 7300-8

30. Hamilton, K. S., Ellison, M. J., Barber, K. R., Williams, R. S., Huzil, J. T., McKenna, S., Ptak, C., Glover, M., and Shaw, G. S. (2001) *Structure (Camb)* **9,** 897-904

31. McKenna, S., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., Spyracopoulos, L., and Ellison, M. J. (2003) *J Biol Chem* **278,** 13151-8

32. Sundquist, W. I., Schubert, H. L., Kelly, B. N., Hill, G. C., Holton, J. M., and Hill, C. P. (2004) *Mol Cell* **13,** 783-9

33. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6

34. Mayer, R. J., Landon, M., and Layfield, R. (1998) *Fold Des* **3,** R97-9

35. Mossessova, E. and Lima, C. D. (2000) *Mol Cell* **5,** 865-76

36. Sheng, W. and Liao, X. (2002) *Protein Sci* **11,** 1482-91

37. Bernier-Villamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002) *Cell* **108,** 345-56

38. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002) *Nature* **419,** 135-41

39. Liakopoulos, D., Doenges, G., Matuschewski, K., and Jentsch, S. (1998) *EMBO J* **17,** 2208-14

40. Johnson, E. S. and Blobel, G. (1997) *J Biol Chem* **272,** 26799-802

41. Kamitani, T., Nguyen, H. P., Kito, K., Fukuda-Kamitani, T., and Yeh, E. T. (1998) *J Biol Chem* **273,** 3117-20

42. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997) *Cell* **88,** 97-107

43. Whitby, F. G., Xia, G., Pickart, C. M., and Hill, C. P. (1998) *J Biol Chem* **273,** 34983-91

44. Rao-Naik, C., delaCruz, W., Laplaza, J. M., Tan, S., Callis, J., and Fisher, A. J. (1998) *J Biol Chem* **273,** 34976-82

45. Gong, L. and Yeh, E. T. (1999) *J Biol Chem* **274,** 12036-42

46. Yeh, E. T., Gong, L., and Kamitani, T. (2000) *Gene* **248,** 1-14

47. Cook, W. J., Jeffrey, L. C., Carson, M., Chen, Z., and Pickart, C. M. (1992) *J Biol Chem* **267,** 16467-71

48. Cook, W. J., Jeffrey, L. C., Kasperek, E., and Pickart, C. M. (1994) *J Mol Biol* **236,** 601-9

49. Phillips, C. L., Thrower, J., Pickart, C. M., and Hill, C. P. (2001) *Acta Crystallogr D Biol Crystallogr* **57,** 341-4

50. Varadan, R., Walker, O., Pickart, C., and Fushman, D. (2002) *J Mol Biol* **324,** 637-47

51. Mastrandrea, L. D., You, J., Niles, E. G., and Pickart, C. M. (1999) *J Biol Chem* **274,** 27299-306

52. You, J. and Pickart, C. M. (2001) *J Biol Chem* **276,** 19871-8

53. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z. J. (2000) *Cell* **103,** 351-61

54. Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M., and Finley, D. (2000) *Cell* **102,** 67-76

55. Galan, J. M. and Haguenauer-Tsapis, R. (1997) *EMBO J* **16,** 5847-54

56. Fisk, H. A. and Yaffe, M. P. (1999) *J Cell Biol* **145,** 1199-208

57. Brummelkamp. T.R., Nijman, S. M. B., Dirac, A. M. G., and Bernards, R. (2003) *Nature* **424,** 797-801

58. Ciechanover, A., Heller, H., Katz-Etzion, R., and Hershko, A. (1981) *Proc Natl Acad Sci U S A* **78,** 761-5

59. Haas, A. L., Warms, J. V., Hershko, A., and Rose, I. A. (1982) *J Biol Chem* **257,** 2543-8

60. Haas, A. L., Warms, J. V., and Rose, I. A. (1983) *Biochemistry* **22,** 4388-94

61. Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V., and Schindelin, H. (2001) *Nature* **414,** 325-9

62. Walden, H., Podgorski, M. S., and Schulman, B. A. (2003) *Nature* **422,** 330-4

63. Walden, H., Podgorski, M. S., Huang, D. T., Miller, D. W., Howard, R. J., Minor, D. L. Jr, Holton, J. M., and Schulman, B. A. (2003) *Mol Cell* **12,** 1427-37

64. Rajagopalan, K. V. (1997) *Biochem Soc Trans* **25,** 757-61

65. VanDemark, A. P. and Hill, C. P. (2003) *Nat Struct Biol* **10,** 244-6

66. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001) *Cell* **105,** 711-20

67. Jiang, F. and Basavappa, R. (1999) *Biochemistry* **38,** 6471-8

68. Lin, Y., Hwang, W. C., and Basavappa, R. (2002) *J Biol Chem* **277,** 21913-21

69. Worthylake, D. K., Prakash, S., Prakash, L., and Hill, C. P. (1998) *J Biol Chem* **273,** 6271-6

70. Cook, W. J., Jeffrey, L. C., Xu, Y., and Chau, V. (1993) *Biochemistry* **32,** 13809-17

71. Cook, W. J., Jeffrey, L. C., Sullivan, M. L., and Vierstra, R. D. (1992) *J Biol Chem* **267,** 15116-21

72. Cook, W. J., Martin, P. D., Edwards, B. F., Yamazaki, R. K., and Chau, V. (1997) *Biochemistry* **36,** 1621-7

73. Miura, T., Klaus, W., Ross, A., Guntert, P., and Senn, H. (2002) *J Biomol NMR*

22, 89-92

74. Ptak, C., Gwozd, C., Huzil, J. T., Gwozd, T. J., Garen, G., and Ellison, M. J. (2001) *Mol Cell Biol* **21,** 6537-48

75. Liu, Y., Mathias, N., Steussy, C. N., and Goebl, M. G. (1995) *Mol Cell Biol* **15,** 5635-44

76. Varelas, X., Ptak, C., and Ellison, M. J. (2003) *Mol Cell Biol* **23,** 5388-400

77. Hamilton, K. S., Ellison, M. J., and Shaw, G. S. (2000) *J Biomol NMR* **18,** 319-27

78. Miura, T., Klaus, W., Gsell, B., Miyamoto, C., and Senn, H. (1999) *J Mol Biol* **290,** 213-28

79. Zheng, N., Wang, P., Jeffrey, P. D., and Pavletich, N. P. (2000) *Cell* **102,** 533-9

80. Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse, J. M., and Pavletich, N. P. (1999) *Science* **286,** 1321-6

81. VanDemark, A. P. and Hill, C. P. (2002) *Curr Opin Struct Biol* **12,** 822-30

82. Hofmann, R. M. and Pickart, C. M. (1999) *Cell* **96,** 645-53

83. Pickart, C. M. and Rose, I. A. (1985) *J Biol Chem* **260,** 7903-10

84. Gwozd, C. S., Arnason, T. G., Cook, W. J., Chau, V., and Ellison, M. J. (1995) *Biochemistry* **34,** 6296-302

85. Ptak, C., Prendergast, J. A., Hodgins, R., Kay, C. M., Chau, V., and Ellison, M. J. (1994) *J Biol Chem* **269,** 26539-45

86. Chen, P., Johnson, P., Sommer, T., Jentsch, S., and Hochstrasser, M. (1993) *Cell* **74,** 357-69

87. Liu, Q., Jin, C., Liao, X., Shen, Z., Chen, D. J., and Chen, Y. (1999) *J Biol Chem* **274,** 16979-87

88. Banerjee, A., Gregori, L., Xu, Y., and Chau, V. (1993) *J Biol Chem* **268,** 5668-75

89. Arnold, J. E. and Gevers, W. (1990) *Biochem J* **267,** 751-7

90. Yamanaka, A., Hatakeyama, S., Kominami, K., Kitagawa, M., Matsumoto, M., and Nakayama, K. (2000) *Mol Biol Cell* **11,** 2821-31

91. VerPlank, L., Bouamr, F., LaGrassa, T. J., Agresta, B., Kikonyogo, A., Leis, J., and Carter, C. A. (2001) *Proc Natl Acad Sci U S A* **98,** 7724-9

92. Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M., Rich, R. L., Myszka, D. G., and Sundquist, W. I. (2001) *Cell* **107,** 55-65

93. Broomfield, S., Chow, B. L., and Xiao, W. (1998) *Proc Natl Acad Sci U S A* **95,** 5678-83

94. Sancho, E., Vila, M. R., Sanchez-Pulido, L., Lozano, J. J., Paciucci, R., Nadal, M., Fox, M., Harvey, C., Bercovich, B., Loukili, N., Ciechanover, A., Lin, S. L., Sanz, F., Estivill, X., Valencia, A., and Thomson, T. M. (1998) *Mol Cell Biol* **18,** 576-89

95. Brusky, J., Zhu, Y., and Xiao, W. (2000) *Curr Genet* **37,** 168-74

96. Katzmann, D. J., Babst, M., and Emr, S. D. (2001) *Cell* **106,** 145-55

97. Xiao, W., Lin, S. L., Broomfield, S., Chow, B. L., and Wei, Y. F. (1998) *Nucleic Acids Res* **26,** 3908-14

98. Ulrich, H. D. and Jentsch, S. (2000) *EMBO J* **19,** 3388-97

99. Villalobo, E., Morin, L., Moch, C., Lescasse, R., Hanna, M., Xiao, W., and Baroin-Tourancheau, A. (2002) *Mol Biol Evol* **19,** 39-48

100. Wang, C., Deng, L., Hong, M., Akkaraju, G. R., Inoue, J., and Chen, Z. J. (2001) *Nature* **412,** 346-51

101. Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., Xiao, W., and Dixit, V. M. (2004) *Nature* **427,** 167-171

102. Hofmann, R. M. and Pickart, C. M. (2001) *J Biol Chem* **276,** 27936-43

103. Ashley, C., Pastushok, L., McKenna, S., Ellison, M. J., and Xiao, W. (2002) *Gene* **285,** 183-91

104. Pornillos, O., Alam, S. L., Davis, D. R., and Sundquist, W. I. (2002) *Nat Struct Biol* **9,** 812-7

105. Pornillos, O., Alam, S. L., Rich, R. L., Myszka, D. G., Davis, D. R., and Sundquist, W. I. (2002) *EMBO J* **21,** 2397-406

106. Hicke, L. (2001) *Cell* **106,** 527-30

107. Hewitt, E. W., Duncan, L., Mufti, D., Baker, J., Stevenson, P. G., and Lehner, P. J. (2002) *EMBO J* **21,** 2418-29

108. Deshaies, R. J. (1999) *Annu Rev Cell Dev Biol* **15,** 435-67

109. Joazeiro, C. A. and Weissman, A. M. (2000) *Cell* **102,** 549-52

110. Zheng, N. (2003) *Structure (Camb)* **11,** 5-6

111. Schulman, B. A., Carrano, A. C., Jeffrey, P. D., Bowen, Z., Kinnucan, E. R., Finnin, M. S., Elledge, S. J., Harper, J. W., Pagano, M., and Pavletich, N. P. (2000) *Nature* **408,** 381-6

112. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W., and Pavletich, N. P. (2002) *Nature* **416,** 703-9

113. Borden, K. L. (2000) *J Mol Biol* **295,** 1103-12

114. Freemont, P. S. (2000) *Curr Biol* **10,** R84-7

115. Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C., and Klevit, R. E. (2001) *Nat Struct Biol* **8,** 833-7

116. Huibregtse, J. M., Scheffner, M., Beaudenon, S., and Howley, P. M. (1995) *Proc Natl Acad Sci U S A* **92,** 5249

117. Scheffner, M., Nuber, U., and Huibregtse, J. M. (1995) *Nature* **373,** 81-3

118. Nuber, U., Schwarz, S., Kaiser, P., Schneider, R., and Scheffner, M. (1996) *J Biol Chem* **271,** 2795-800

119. Verdecia, M. A., Joazeiro, C. A., Wells, N. J., Ferrer, J. L., Bowman, M. E., Hunter, T., and Noel, J. P. (2003) *Mol Cell* **11,** 249-59

120. Karin, M. and Ben-Neriah, Y. (2000) *Annu Rev Immunol* **18,** 621-63

121. Ghosh, S. and Karin, M. (2002) *Cell* **109 Suppl,** S81-96

122. Ghosh, S., May, M. J., and Kopp, E. B. (1998) *Annu Rev Immunol* **16,** 225-60

123. Ishitani, T., Ninomiya-Tsuji, J., Nagai, S., Nishita, M., Meneghini, M., Barker, N., Waterman, M., Bowerman, B., Clevers, H., Shibuya, H., and Matsumoto, K. (1999) *Nature* **399,** 798-802

124. Takaesu, G., Kishida, S., Hiyama, A., Yamaguchi, K., Shibuya, H., Irie, K.,

Ninomiya-Tsuji, J., and Matsumoto, K. (2000) *Mol Cell* **5,** 649-58

125. Takaesu, G., Surabhi, R. M., Park, K. J., Ninomiya-Tsuji, J., Matsumoto, K., and Gaynor, R. B. (2003) *J Mol Biol* **326,** 105-15

126. Cao, Z., Xiong, J., Takeuchi, M., Kurama, T., and Goeddel, D. V. (1996) *Nature* **383,** 443-6

127. Chen, Z. J., Parent, L., and Maniatis, T. (1996) *Cell* **84,** 853-62

128. Wooff, J., Pastushok, L., Hanna, M., Fu, Y., and Xiao, W. (2004) *FEBS Lett* **566,** 229-33

129. Shi, C. S. and Kehrl, J. H. (2003) *J Biol Chem* **278,** 15429-34

130. Prakash, S., Sung, P., and Prakash, L. (1993) *Annu Rev Genet* **27,** 33-70

131. Wood, R. D., Mitchell, M., Sgouros, J., and Lindahl, T. (2001) *Science* **291,** 1284-9

132. Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaccmaz, K., and Linn, S. (2004) *Annu Rev Biochem* **73,** 39-85

133. Hoeijmakers, J. H. (2001) *Nature* **411,** 366-74

134. Broomfield, S., Hryciw, T., and Xiao, W. (2001) *Mutat Res* **486,** 167-84

135. Dudas, A. and Chovanec, M. (2004) *Mutat Res* **566,** 131-67

136. Prakash, S. and Prakash, L. (2000) *Mutat Res* **451,** 13-24

137. Barbour, L. and Xiao, W. (2003) *Mutat Res* **532,** 137-55

138. Ulrich, H. D. (2002) *Eukaryot Cell* **1,** 1-10

139. Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996) *Science* **272,** 1646-9

140. Jentsch, S., McGrath, J. P., and Varshavsky, A. (1987) *Nature* **329,** 131-4

141. Silver, E. T., Gwozd, T. J., Ptak, C., Goebl, M., and Ellison, M. J. (1992) *EMBO J* **11,** 3091-8

142. Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994) *Cell* **79,** 1233-43

143. Bailly, V., Prakash, S., and Prakash, L. (1997) *Mol Cell Biol* **17,** 4536-43

144. Matunis, M. J. (2002) *Mol Cell* **10,** 441-2

145. Ulrich, H. D. (2003) *J Biol Chem* **278,** 7051-8

146. Giandomenico, V., Simonsson, M., Gronroos, E., and Ericsson, J. (2003) *Mol Cell Biol* **23,** 2587-99

147. Gronroos, E., Hellman, U., Heldin, C. H., and Ericsson, J. (2002) *Mol Cell* **10,** 483-93

148. Ito, A., Kawaguchi, Y., Lai, C. H., Kovacs, J. J., Higashimoto, Y., Appella, E., and Yao, T. P. (2002) *EMBO J* **21,** 6236-45

149. Li, M., Luo, J., Brooks, C. L., and Gu, W. (2002) *J Biol Chem* **277,** 50607-11

150. Freiman, R. N. and Tjian, R. (2003) *Cell* **112,** 11-7

151. Yan, Y., Harper, S., Speicher, D. W., and Marmorstein, R. (2002) *Nat Struct Biol* **9,** 862-9

152. Rojas, J. R., Trievel, R. C., Zhou, J., Mo, Y., Li, X., Berger, S. L., Allis, C. D., and Marmorstein, R. (1999) *Nature* **401,** 93-8

153. Draker, K. A. and Wright, G. D. (2004) *Biochemistry* **43,** 446-54

Chapter 2:

Crystal Structure of the Human Ubiquitin Conjugating Enzyme Complex, hMms2•hUbc13.

2.1 Summary:

The ubiquitin conjugating enzyme complex Mms2•Ubc13 plays a key role in post-replicative DNA repair in yeast, and the NF-κB signal transduction pathway in humans. This complex assembles novel poly-ubiquitin chains onto as yet uncharacterized protein targets. This chapter focuses on the determination of the crystal structure of a complex between hMms2 and hUbc13 at 1.85 Å resolution and the structure of free hMms2 at 1.9 Å resolution. The structures reveal that the hMms2 monomer undergoes a localized conformational change upon its interaction with hUbc13. The nature of the interface provides a physical basis for the preference of Mms2 for Ubc13 over a variety of other structurally similar ubiquitin conjugating enzymes. The structure of the hMms2•hUbc13 complex provides the conceptual foundation for understanding the mechanism of poly-ubiquitin chain assembly and for its interactions with the RING-finger proteins Rad5 and TRAF6/2. Subsequent studies involving the structure of the yeast Ubc13•Mms2 complex and models predicting the sites of Ub association through NMR spectrocopy and mutagenesis are also discussed.

2.2 Introduction:

In its traditionally viewed role, covalent attachment of Ub serves as a tag that commits short-lived or otherwise damaged proteins to degradation by the proteasome[1-3]. Recently, a novel, highly conserved, non-proteolytic ubiquitination pathway has been uncovered that is unconventional in both mechanism and function. The central player in this pathway is a heterodimer composed of an enzymatically active ubiquitin conjugating enzyme Ubc13 (E2) and an enzymatically inactive ubiquitin conjugating enzyme variant Mms2 (UEV).

Heterodimeric complexes between Ubc13 and a UEV are implicated in at least two important cellular processes. In *Saccharomyces cerevisiae,* the heterodimer plays an essential role in the process of post-replication DNA repair (PRR)[4-6], while in humans, it regulates the NF-κB signal transduction pathway[7-10]. PRR is a DNA damage tolerance mechanism that facilitates the completion of DNA synthesis and cell division in the presence of damaged DNA[11]. PRR is composed of an error-prone pathway and an error-free pathway[12]. While the error-prone pathway involves several specialized DNA polymerases for translesion DNA synthesis, error-free PRR involves a variety of other of proteins, including the Mms2•Ubc13 heterodimer and the RING finger protein Rad5. It was also demonstrated that yeast Rad5 directly interacts with Ubc13 and another RING finger protein Rad18[5].

The NF-κB signal transduction pathway is involved in a number of cellular reponses including apoptosis, inflammation and immunity[7-9]. Upon stimulation of cells by various agonists, IκB, the inhibitor of NF-κB, is rapidly phosphorylated by

the IκB kinase (IKK) complex and then degraded by the 26S proteosome in a ubiquitin dependent manner. In an interesting reversal of its generally accepted role in degradation, ubiquitin actually activates IKK as a post-translational modification that requires the RING finger protein TRAF6 and the Ubc13-Mms2 heterodimer. In the presence of Traf6, the heterodimer has been shown to catalyze the formation of Lys63 linked ubiquitin chains[7] onto what has now been characterized as Traf6[8] and another target the regulatory subunit of IKK, Nemo[9].

2.3 Experimental Procedures.

2.3.1 Expression and purification of recombinant Mms2 and Ubc13.

Both human *UBC13* and *MMS2* open reading frames were PCR amplified and cloned as *Bam*HI-*Sal*I fragments into the corresponding sites of a GST fusion vector pGEX6 (Pharmacia). In each case, the last codon of N-terminal GST sequence was separated from the first codon of Mms2 and Ubc13 by intervening DNA that encoded the linker peptide, Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro-Leu-Gly-Ser. The linker peptide contained the cleavage site for the PreScission Protease (Pharmacia) that cleaves between Gln and Gly. Cleavage results in the separation of GST from Mms2 or Ubc13, which contain Gly-Pro-Leu-Gly-Ser appended to the N-terminus of the first codon. The DNA sequences were derived from plasmid-borne cloned versions for Mms2[13] and Ubc13[14] respectively. The DNA sequences of the human *UBC13* and *MMS2* coding regions were verified by sequencing each recombinant plasmid.

Proteins were expressed in the *E. coli* strain *BL21(DE$_3$)-RIL* (Stratagene) that

contained extra copies of the argU, ileY, and leuW tRNA genes in addition to the pGEX6-derived plasmids as described above. Two liter cultures were grown at 37°C to OD_{590}= 0.4 in LB media containing ampicillin (50 μg/ml) followed by induction with isopropyl ß-D-thiogalactopyranoside (IPTG) (0.4 mM) for 10 hours at 30°C. Cells were harvested by centrifugation, and stored at –80°C. All subsequent steps were performed at 4°C.

Cell pellets were resuspended in 50 mL of disruption buffer (20 mM Tris/Cl pH 7.9, 10 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, 1 mM DTT, 0.3 M ammonium sulfate, 1 mM PMSF), lysed by two passages through a French Press, followed by centrifugation (40,000 rpm for 45 minutes). The supernatant was dialyzed against 4 L of PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH 7.3) overnight at 4°C and clarified through a 0.45 μm low protein binding filter (Millipore). The filtered lysate was applied slowly to a 5 mL Glutathione Sepharose 4B RediPack column (Pharmacia) equilibrated with 50 mL of 1XPBS buffer. The column was washed three times with 30 mL of PBS buffer, and the retained protein was eluted with three washes of 5 mL of Glutathione Elution buffer (10 mM reduced glutathione, 50 mM Tris/Cl pH 8.0). Glutathione was removed from protein samples by dialysis against 4 L of PreScission Cleavage Buffer (50 mM Tris/Cl pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT) for 4 hours followed by the addition of PreScission Protease (Pharmacia) (40 units). Fusion proteins were completely cleaved after a 10 hour incubation at 4°C. Following cleavage, the sample was reapplied to the Glutathione Sepharose 4B column and eluted as described above. In

this case, cleaved Ubc13 and Mms2 proteins appeared in the 1XPBS flow-through, while the GST tag and the PreScission Protease (which is also fused to GST) remained bound to the column. The flow-through was concentrated to 2 mL using an Ultrafree Centrifugal Filter Device (Millipore- 10 kDa molecular mass cutoff) and applied to a Hi-Load 16/60 Superdex 75 column (Pharmacia) equilibrated with 200 mL of Superdex 75 buffer (50 mM HEPES pH 7.5, 75 mM NaCl, 1 mM EDTA, 1 mM DTT). Proteins were eluted at a flow rate of 1 mL/min, and collected in 1 mL fractions. Both Ubc13 and Mms2 eluted between 69-80 mL and were judged to be pure by SDS-PAGE. Samples were subsequently pooled and concentrated.

2.3.2 Selenium Incorporated Protein Expression.

Selenomethionine-substituted hMms2 (Se-hMms2) was expressed by incorporation of selenomethionine into bacterial growth media under growth conditions that repress methionine biosynthesis[15]. Proteins were expressed in the *E. coli* strain *BL21(DE$_3$)-RP* (Stratagene). Colonies of human Mms2 previously streaked on LB/CHL/AMP were used to innoculate 100 mL of LB/CHL/AMP grown overnight at 37°C. Cells were harvested by centrifugation for 30min @ 3500 rpms and resuspended in 2 L minimal media (5xM9 salts: 12.8g/L Na_2HPO_4, 3.0g/L KH_2PO_4, 1.0 g/L NH_4Cl, 0.5 g/L NaCl, plus 2 mM $MgSO_4$ 0.2 mM $CaCl_2$, 1% glucose and the appropriate antibiotic) at 37 °C to OD_{600}=0.4. Cells were then transferred to 25°C where selected amino acids (100mg Seleno-$_L$-Methionine, 200mg Lysine-HCl, 200mg Threonine, 100mg Leucine, 200mg Phenylalanine, 100mg

Isoleucine, and 100mg Valine) were added. After 30 min cells were induced with IPTG (0.4 mM) and grown for an additional 10 hours at 25 °C. Purification followed the above protocol.

2.3.3 Protein crystallization.

All crystals were grown at 20-24 °C using the hanging drop vapor diffusion technique. Crystals of hMms2-hUbc13 and Se-hMms2-hUbc13 were grown by mixing 1 μL of complex at 6 mg•mL^{-1} in 50 mM Hepes pH 7.5, 75 mM NaCl, 1mM EDTA, and 0.5 mM DTT with 1 μL of well solution (20% PEG 8000, and 100 mM citrate pH 6.5), which was then equilibrated against the well solution. hMms2-hUbc13 crystallised in the space group $P2_12_12_1$ (a=44.1, b=74.1, c=91.5 Å, $\alpha=\beta=\gamma=90^{o}$). Crystals of monomeric hMms2 were grown in the same way as the complex except the well solution was 16% PEG 6000, 25 mM NaCl and 100 mM citrate pH 6.5. The space group of the hMms2 crystals was $C222_1$ (a=42.2, b=54.5, c=108.2 Å, $\alpha=\beta=\gamma=90^{o}$).

2.3.4 Structure determination.

Diffraction data for native hMms2 to 1.9 Å resolution were collected from a single crystal at 105 K with an RaxisIV image plate detector and a Rigaku RU-H3R rotating anode generator. Diffraction data for the selenium derivative of hMms2-Ubc13 (2.0 Å) and the native hMms2-hUbc13 heterodimer (1.85 Å) were collected from single crystals at NSLS, beamline X12C using a Brandeis 1K CCD detector.

Data were scaled and integrated with DENZO and SCALEPACK[16]. Multi-wavelength anomalous dispersion (MAD)[17] phases to 2.0 Å were calculated for the Se-hMms2-hUbc13 derivative with SOLVE[18] and improved by solvent flattening using DM[19]. Maps calculated with these phases gave excellent electron density for hMms2 and slightly poorer but interpretable density for hUbc13. The program ARP/wARP (version 5.1)[19;20] was used to initially build and refine 235 of the 290 amino acid residues in the hMms2-hUbc13 complex and the remainder were built manually using O[21]. The structure of monomeric hMms2 was solved by molecular replacement using MOLREP[22] and hMms2 derived from the heterodimer as the search model. Rigid body refinement of the initial solution gave a starting R-factor of 39.0% and a correlation coefficient of 63.9% using reflections from 8-4 Å.

Structures of the hMms2-hUbc13 complex and monomeric hMms2 were refined through several cycles. Manual rebuilding was performed with reference to Sigma A-weighted[23] 2|Fo| - |Fc| electron density maps calculated in CNS[24] and experimental, MAD-phased electron density maps for the hMms2-Ubc13 complex. Both structures were refined with CNS using torsion angle dynamics and maximum likelihood targets[25]. The quality of the model was assessed by PROCHECK[26]. The fraction of residues in the most favored and additionally allowed regions of the Ramachandran plot, respectively, are 87.7 % and 11.5 % for the heteromeric hMms2-hUbc13 complex and 87.8 % and 11.3 % for the hMms2 monomer. No residues in either structure were in the disallowed region.

2.4 Results and Discussion

2.4.1 Overview of the hMms2•hUbc13 structure

Purified full-length hMms2•hUbc13 complex, and monomeric hMms2 were crystallized and their structures determined by MAD phasing and molecular replacement. hUbc13 adopts a canonical E2 fold, consisting of a four-stranded anti-parallel β-sheet flanked by four α-helices[27] (**Figure 2.1**). Although hMms2 shares significant sequence similarity to the ubiquitin conjugating enzymes[6], it is strikingly atypical in that it lacks the active-site Cys required for Ub conjugation and the two most carboxy-terminal helices, and it contains an extended amino-terminal peptide tail.

2.4.2 hMms2-hUbc13 interface

hMms2 and hUbc13 interact through a surprisingly narrow interface approximately 30 Å long and 10 Å wide that buries a total of 1500 $Å^2$ of solvent accessible surface area. Equilibrium sedimentation analysis of the complex indicates that it has a dissociation constant (K_D) of approximately 2 μM (near the limit of this procedure. A more accurate value has been determined by isothermal titration calorimetry). The interface involves the amino-terminal extension and α-helix (α1) and the L1 loop of hMms2 which packs against the exposed face of the hUbc13 β-sheet (β strands 2-4 and loops L1 and L4).

The interior of the interface is composed of hydrophobic side chains that pack against other hydrophobic residues and the mainchain of the dimerization partner (**Figure 2.2** and **2.4**). The critical hydrophobic interface residues, Phe57, Tyr34 and

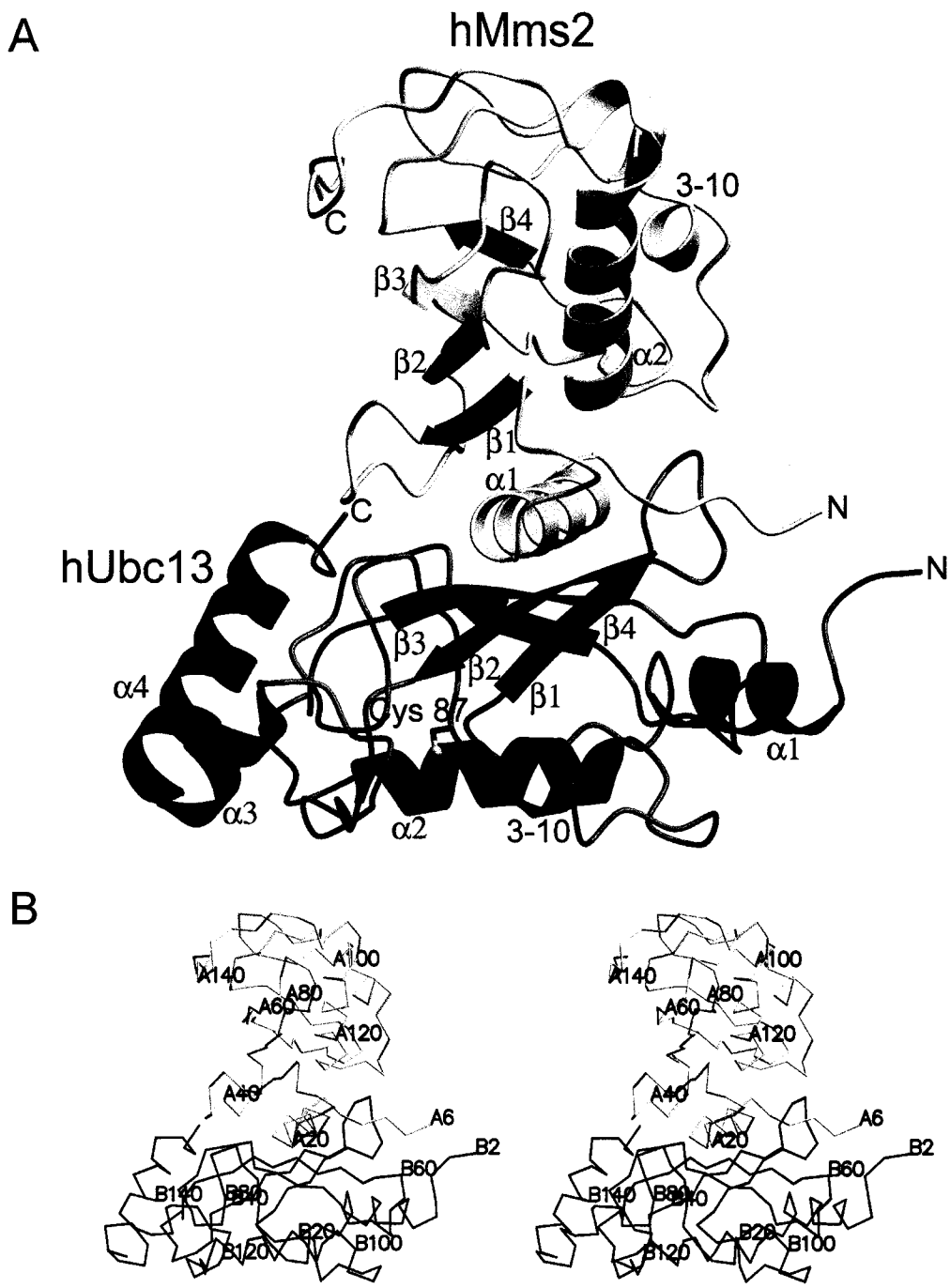


Figure 2.1 The Overall Structure of the hMms2•hUbc13 Complex.
A. Ribbon diagram of the hMms2•Ubc13 complex (hMms2 in yellow, hUbc13 in blue). All secondary structural elements and the active site cysteine of hUbc13 (Cys87) are labeled. **B.** A stereoview (divergent) of the C_α backbone of the hMms2•hUbc13 complex where the A chain is hMms2 and the B chain is hUbc13.

Table 2.1

Summary of Crystallographic Data for the Complex of Mms2•Ubc13

Data collection

	X12C-NSLS (Brookhaven National Lab)				RaxisIV 100
	Se-hMms2-hUbc13 complex			Native-complex	Native-hMms2
	λ_1 (Peak)	λ_2 (Inflection)	λ_3 (Remote)		
Wavelength (Å)	0.981958	0.982127	1.02213	0.90006	
Resolution (Å)	21-2.0	21-2.0	21-2.0	40-1.85	20-1.9
(last shell)	(2.03-2.00)	(2.07 –2.00)	(2.07 –2.00)	(1.88-1.85)	(1.93-1.90)
Completeness (%)	98.0	94.4	95.0	98.3	97.9
Observations	113204	107661	106576	110669	53554
Unique Reflections	39267	42419	37856	26066	9995
mosaicity (°)	0.4	0.4	0.4	0.2	0.5
$\langle I \rangle / \sigma(I)$(last shell)	22.0 (6.5)	18.9 (4.4)	21.6 (6.7)	37.1 (3.2)	36.9 (9.7)
R_{sym}[a](last shell)	0.041 (0.137)	0.037 (0.164)	0.034 (0.129)	0.024 (0.27)	0.05 (0.144)
Se atoms found (total)	7 (8)			--	--
Figure of Merit	0.60 (SOLVE 'Z-score' = 45.8)[b]			--	--

Refinement Statistics

	complex hMms2-hUbc13	hMms2
Resolution Range (Å)	40 – 1.85	20 –1.9
R-factor	21.5	21.3
R_{free}[c]	24.4	24.6
R.M.S. Deviations		
Bond length (Å)	0.0065	0.0049
Bond Angles (°)	1.55	1.27
# of protein residues	288 residues	137 residues
# of Water molecules	190	99

[a] $R_{sym} = \Sigma_{i,h} |I(i,h) - \langle I(h) \rangle| / \Sigma_{i,h} |I(i,h)|$ where $I(i,h)$ and $\langle I(h) \rangle$ are the ith and mean intensity of reflection h.

[b] Z-score, quality criteria according to program SOLVE[18].

[c] R_{free} is calculated using 5% of the reflections[28].

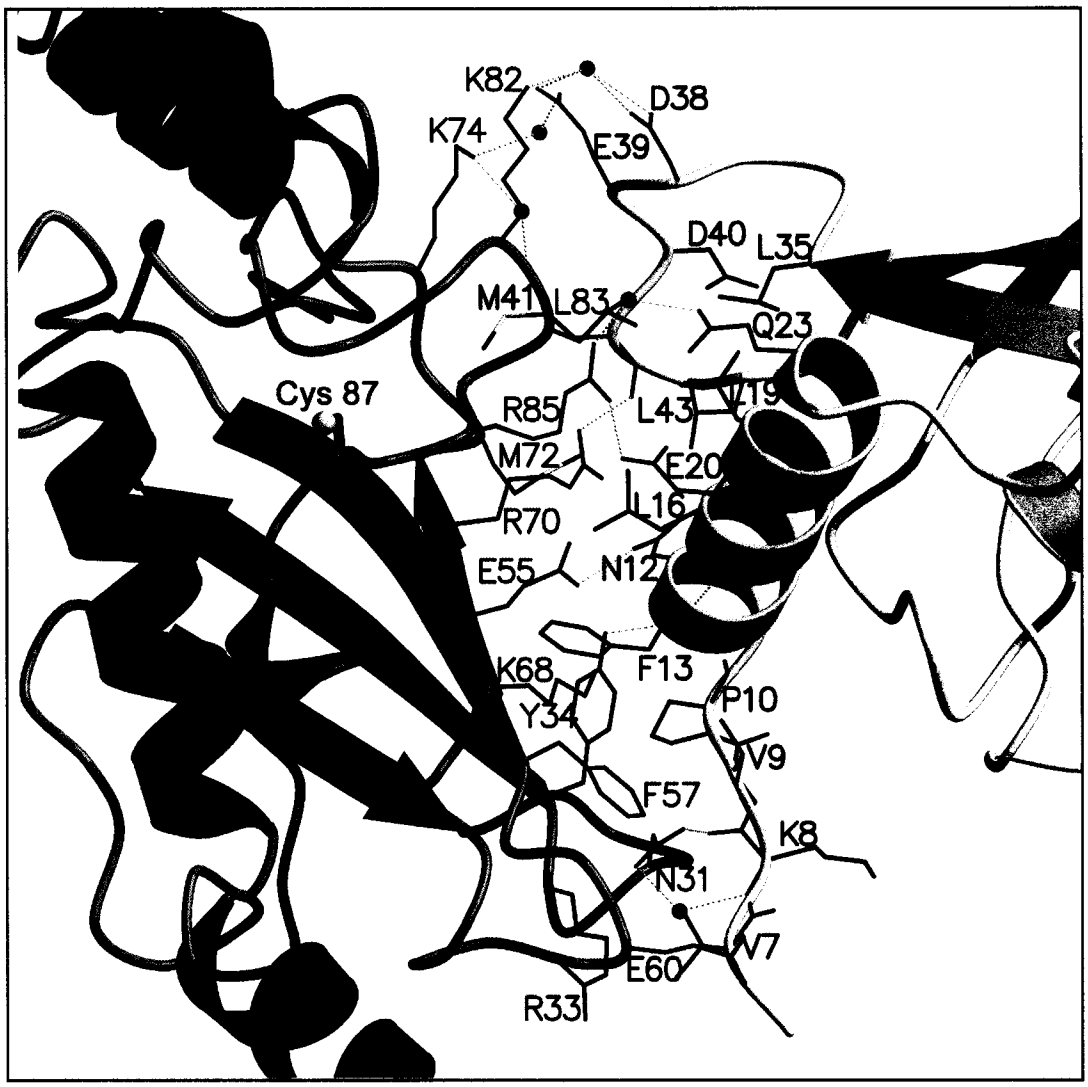


Figure 2.2 The Interface Created by the hMms2•hUbc13 Complex. The interface of hMms2(yellow) and hUbc13(blue) is oreinted ~90° clockwise with respect to Figure 2.1. Residues in the interface are labelled (hMms2 in yellow and hUbc13 in blue) and the active-site Cys is illustrated as a ball and stick and shaded yellow.

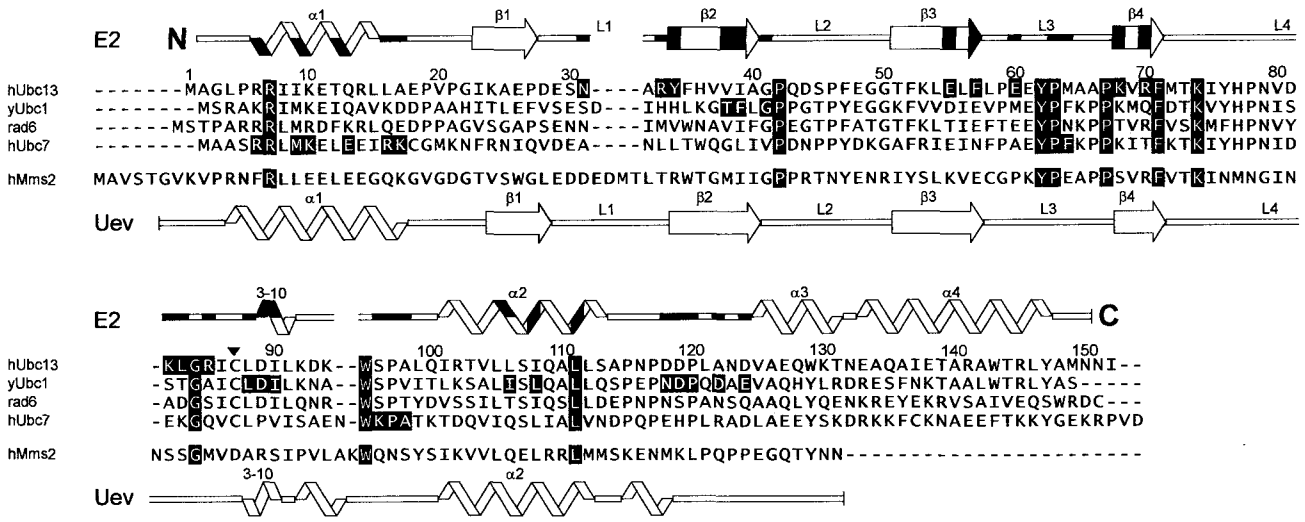


Figure 2.3 E2/UEV Structure-Based Sequence Alignment. Sequence alignments of the interface residues in hUbc13 and hMms2 with homologues / orthologues in humans and yeast. Residues in the hMms2•hUbc13 interface are colored in yellow and blue respectively. Residues in the proposed hUbc13•E3(RING) interface colored in green. Residues in the E2 that interact with ubiquitin are colored in red. Absolutely and partially conserved residues between the E2 homologues are shaded in black and grey respectively. Secondary structural elements are shown for the E2s and UEV above and below the sequence alignment respectively.

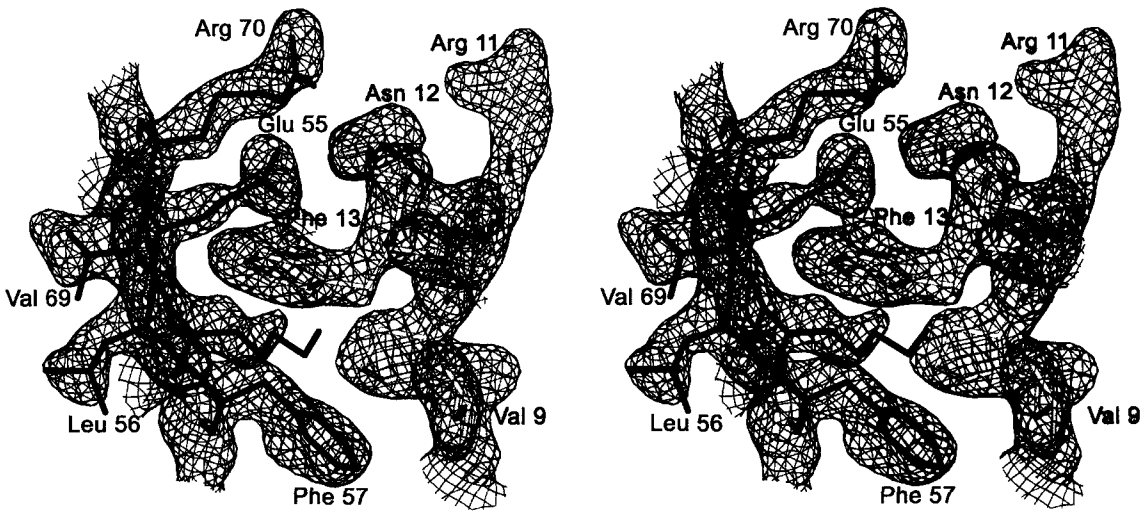


Figure 2.4 Electron Density Maps of hUbc13 in complex with Mms2. Stereo view (convergent/cross-eyed) of the experimental, MAD-phased complex to 2.0 Å resolution with the electron density map contoured at 1.2 σ. The electron density is superimposed on the refined model. Subsequent studies have demonstrated that Phe13 from Mms2 and Phe57 from Ubc13 contribute a significant amount of energy to the binding interaction. The density of the interdigitated hydrophobic residues helps explains this

Leu83, in hUbc13 are not conserved in other E2 proteins, explaining in part why hMms2 only binds to hUbc13 and not other E2s (**Figure 2.3**). In contrast, the solvent-exposed border of the interface is largely composed of hydrophilic residues that make a number of direct and water-mediated electrostatic inter-subunit contacts.

Comparison of the monomeric and hUbc13-bound forms of hMms2 reveals a conformational rearrangement in hMms2 upon complex formation. In monomeric hMms2, Val9 forms the amino-terminus of α1. Upon binding to hUbc13, Val9 unravels from α1, and packs against Phe57 of hUbc13. This repositions the region of hMms2 amino-terminal to α1 to contact L1 and L3 in hUbc13 (**Figure 2.5**). Furthermore, Phe13 undergoes a *t*→ *-g* rotation about χ_1 that allows it to pack against Phe57 in hUbc13.

2.4.3 Interactions between hMms2•hUbc13, ubiquitin and E3s

Ubiquitin dependent degradation via the 26S proteosome is facilitated by a poly-ubiquitin chain attached to the targeted substrate in which ubiquitin molecules are joined by isopeptide bonds that link Lys48 of one ubiquitin with the carboxy-terminus of the next ubiquitin in the chain[29]. In some cases, this chain building function has been found to reside exclusively within the E2[30]. In contrast, hMms2•hUbc13 catalyses the formation of poly-ubiquitin chains that are linked by isopeptide bonds between Lys63 and the carboxy-terminus of the next monomer in the chain[31]. The heterodimer is both necessary and sufficient to link two ubiquitin molecules at Lys63. Recent work has shown that this reaction involves the transfer of

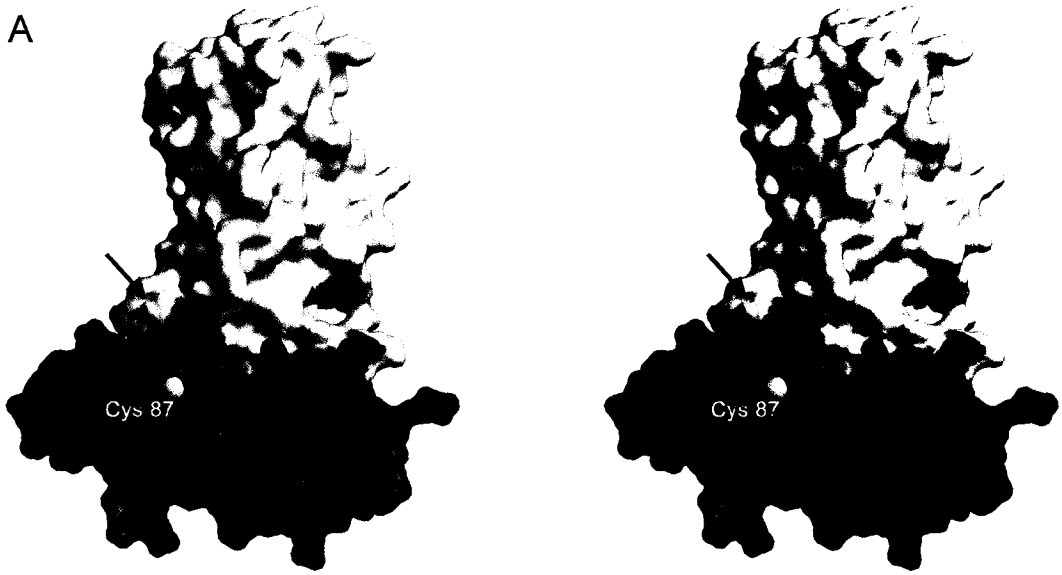
Figure 2.5 Comparison of free and bound forms of hMms2. A stereo view of the overlay of the free hMms2 with the structure of hMms2 bound in the heterodimer complex. The secondary structure is illustrated with a ribbon diagram of the heterodimer (hMms2 in yellow, hUbc13 in blue and the free hMms2 in red). The surface of hUbc13 is shown in grey.

ubiquitin from the active-site of Ubc13 to Lys63 of a ubiquitin molecule that is positioned on the surface of Mms2[32]. Additionally, the RING-finger E3 Traf6 is required to extend the length of the chain beyond two units[7]. Thus, chain assembly likely involves the placement of two ubiquitin molecules and one Traf6 molecule on the surface of the heterodimer.

The probable locations of the thiolester-linked ubiquitin and Traf6 on the surface of Ubc13 can be inferred from other studies. Independent HSQC-NMR experiments have revealed a common binding interface between two different E2s and their thiol-ester-linked ubiquitins[33-35]. This interface involves residues in L4 immediately carboxy-terminal to the active site cysteine, as well as residues in the loop between helix α2 and α3 (**Figure 2.3**). Similarly, the approximate interface of Traf6 with Ubc13 can be inferred from the crystallographic structure of the ring-finger E3 c-Cbl in complex with the human E2 UbcH7[36]. At this interface, L1 and L2 of UbcH7 make contact with residues in and around the ring motif of c-Cbl. Notably, a number of these contact residues are conserved in hUbc13 (**Figure 2.3**) and the Mms2-Ubc13 E3s, Traf6 and Rad5 (**Figure 2.6B**).

A molecular surface representation of the hMms2-hUbc13 heterodimer reveals two clefts at either side of the heterodimer interface (**Figure 2.6A**). Both clefts narrow to channels as they approach the active-site of hUbc13, one from the right and one from the left. We note that the ubiquitin molecule that donates its Lys63 can be accommodated at either cleft without steric interference from the thiolester ubiquitin or Traf6. Occupation of the right cleft however, would present Traf6 the

A



B

			*							*			*	*						*	*		*								
c-Cbl	C	K	I	C	…	C	T	S	C	L	T	S	W	Q	…	C	P	F	C	R											
Rad18	C	H	I	C	…	C	S	L	C	I	R	T	H	L	…	C	P	L	C	F											
Traf6	C	P	I	C	…	C	K	A	C	I	I	K	S	I	…	C	P	Q	C	R											
Rad5	C	P	N	C	…	C	E	K	C	L	F	E	Y	I	…	C	P	N	C	N											

Figure 2.6 Molecular surface of hMms2-hUbc13 heterodimer. **A** The molecular surface of hMms2 and hUbc13 are colored in yellow and blue, respectively, in approxiamately the same orientation as in Figure 2.1. The potential ubiquitin binding site and E3 RING finger binding residues on Ubc13 are colored in red and green as in Figure 2.2. The active site cysteine is colored yellow. Arrows indicated clefts leading to the active site where potential targets or non-covalent Ubs may be positioned for ubiquitination. **B** A sequence alignment of the consensus sequence of RING-finger domain of E3's involved in the Mms2/Ubc13 pathways (TRAF6, RAD5 and RAD18) and a known E3 RING finger protein c-Cbl. The * symbol represents residues in the c-Cbl ring finger that are in the interface between c-Cbl and UbcH7[34].

opportunity of participating directly in chain catalysis. The position of this non-covalent ubiquitin will present a key piece of the puzzle in determining both the mechanism of Lys63-linked ubiquitin chain assembly and target recognition.

2.4.4 A Comparison Between the Human and Yeast Ubc13•Mms2 structures.

The yeast Ubc13•Mms2 crystal structure[37] was determined independently by Vandemark et al. (Cell 2001) in a parallel time frame to the human Ubc13•Mms2 structure described above. There are few structural differences between these two heterodimers, overall they are virtually superimposable with an RMSD of 0.9 Å. The few interesting differences include several conservative amino acid changes, for example a Phe57 and Met72 in human Ubc13 are a Tyr and Leu in yUbc13 respectively (**Figure 2.7**). These two residues are in the interface between Ubc13 and Mms2 and may account for the slight discrepancy in binding affinity reported for the two homologs (human 50nM[38], yeast 300nM[39]). Differing residues in the Mms2 homologs that contribute to the interaction between Mms2 and Ubc13 include Met41 (hMms2) that is a Leu in yMms2. Overall the conservation in structure of the heterodimer is supported by biochemical data that suggests that these proteins function identically in the synthesis of Lys63 linked Ub chains.

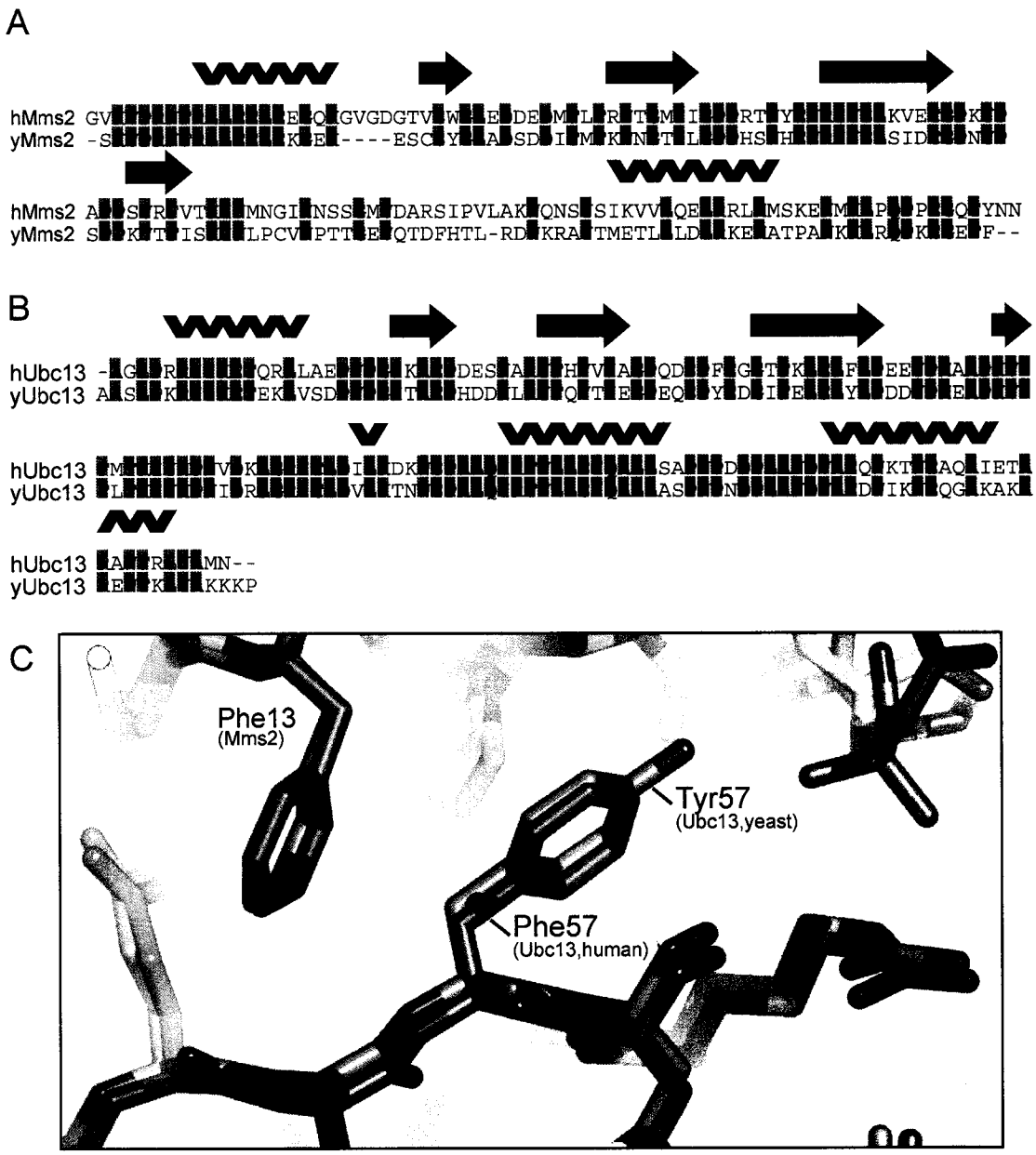


Figure 2.7 Comparison of *S. cerevisiae* and Human Mms2 and Ubc13. Residues shaded in grey are identical and the secondary structural elements are shown above the alignement. **A.** A comparison of sequence between yeast and human Mms2 illustrates the 50% identity between the sequences leadin to and RMSD of 0.7Å, regions that bind Ub and Ubc13 are extremely well conserved. **B.** A comparison of sequence between yeast and human Ubc13 illustrates the 70% identity between the sequences, regions that bind Ub and Mms2 are extremely well conserved. The RMSD between the two structures is 0.8 Å. **C.** An overlay of the of the yeast (light blue) and human (green) Ubc13•Mms2 complexes illustrating the slight variation in the interaface residues.

2.4.5 Models of the Ubiquitin-bound Human and Yeast Ubiquitin Conjugation Complex for Ubc13•Mms2.

Two dimensional ^{1}H-^{15}N NMR was used to map residues involved in: a) the interaction between human Mms2 and Ubc13 and b) the interactions between each of the subunits (Mms2 and Ubc13 alone) or heterodimer with non-covalently bound acceptor Ub or a thiolester linked donor Ub (work done by S.Mckenna 1999-2004). These NMR derived constrants were used in coordination with an unbiased docking program (BiGGER)[40] to assemble the four components of the catalytic complex[38;41]. The derived model is very similar to the Ubs docked onto yUbc13•Mms2 heterodimer which was based on mutagenesis data including: a) the Ubc13 (A110R) substitution located on the surface of α3, near the center of the predicted interaction between Ubc13 and the donor Ub, that resulted in a 4-fold reduction in the rate of isopeptide bond formation, b) the Ubc13 (D81A) substitution is situated near the predicted position of the acceptor Ub Lys63 which resulted in a diminished affinity between the heterodimer and Ub *in vitro*, conversely c) the Mms2 (E12R) substitution located near the on opposite cleft has no effect in weaken the interaction between Ub and the heterodimer(**Figure 2.8**)[37].

Both models envision the accepting Ub molecule sitting on a concave face of Mms2 with its carboxy-terminus situated far from the vestigal active site of Mms2. In combination with Ubc13, the concave surface on the left of the active-site narrows to form a channel that approaches the active-site of Ubc13. The side-chain Lys63 for the acceptor-Ub lies within the channel, placing the ε-amino group within 3 Å of the

catalytic Cys of Ubc13 whereas the Lys48 residue is buried in the 2800 $Å^2$ protein:protein interface created between Mms2•Ubc13 and the acceptor-Ub. Furthermore, the arrangement of the Ubs on the heterodimer poses no steric problem for the interaction of Ubc13 with a functionally specific E3. The model suggests that the UEV non-covalently orients Ub in specific manner to promote the formation of Lys63 linked Ub chains by the E2 Ubc13 and that the selection of targets for Lys63 linked chains are most likely specified by the individual E3 that can be associated with this complex.

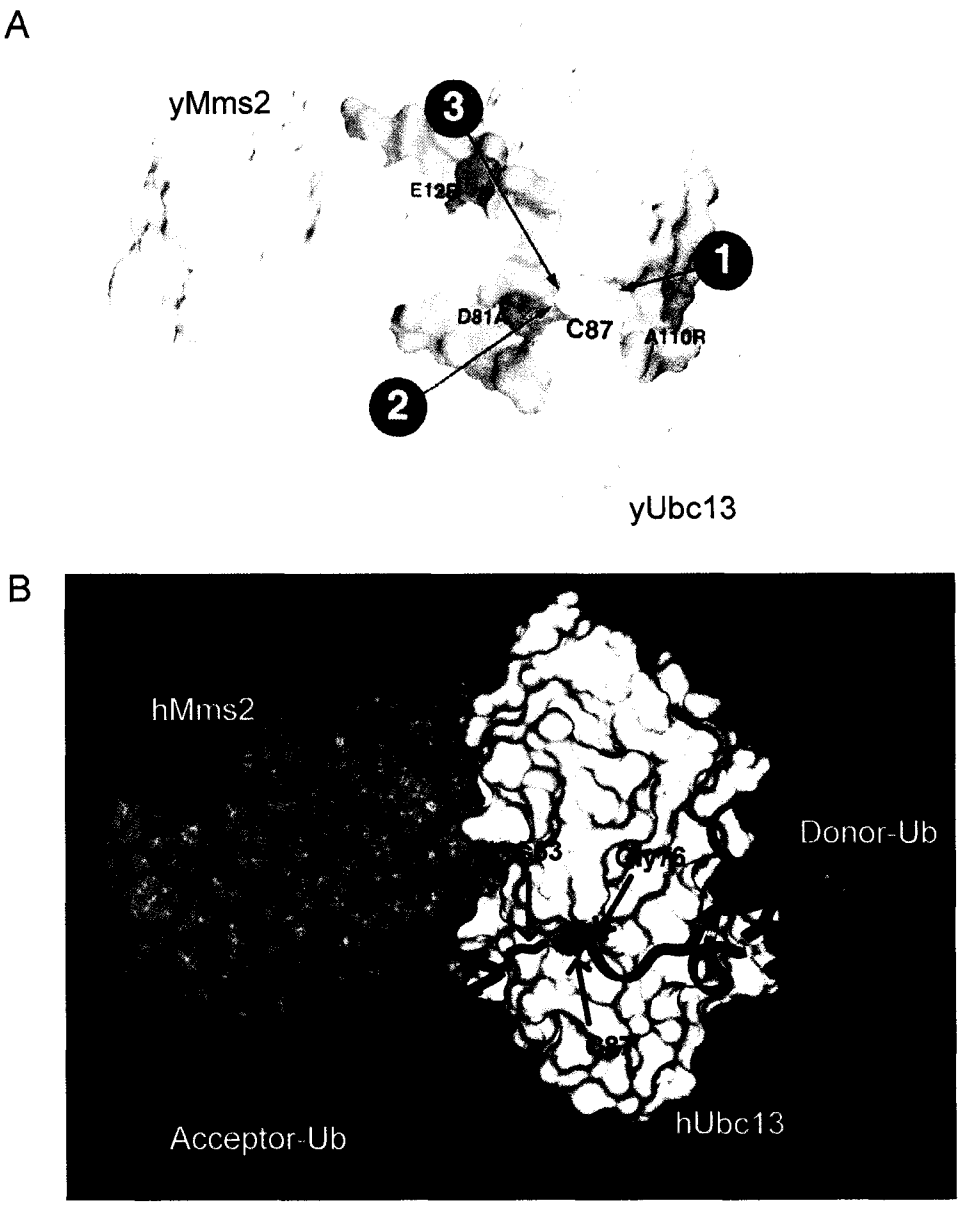


Figure 2.8 Model of the Tetrameric Ub-Conjugating Enzyme Complex. **A.** Surface representation of the *S.cerevisiae* Ubc13•Mms2 complex highlighting the three cavities that lead to the active-site. The three point mutations in each cavity used to develop the yeast tetrameric model are shown. **B**. The NMR-derived model of the human (Mms2•Ub)•Ubc13~Ub complex. The interaction between either the acceptor-Ub or donor-Ub (red, ribbon) and the surface of the Ubc13(yellow) Mms2(blue) heterodimer are illustrated. Lys63 of the acceptor-Ub, Gly76 of the donor-Ub, the active-site Cys (C87) and the hypothesized E3 binding residues (in white) are highlighted. This figure has been adapted from McKenna *et al.*(2003)[40] and Vandemark *et al.* (2001)[36].

2.5 Reference List

1. Hochstrasser, M. (1996) *Annu. Rev. Genet* **30,** 405-439
2. Hershko, A. and Ciechanover, A. (1998) *Annu Rev Biochem* **67,** 425-79
3. Pickart, C. M. (1997) *FASEB J* **11,** 1055-66
4. Hofmann, R. M. and Pickart, C. M. (1999) *Cell* **96,** 645-53
5. Ulrich, H. D. and Jentsch, S. (2000) *EMBO J* **19,** 3388-97
6. Broomfield, S., Chow, B. L., and Xiao, W. (1998) *Proc Natl Acad Sci U S A* **95,** 5678-83
7. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z. J. (2000) *Cell* **103,** 351-61
8. Wang, C., Deng, L., Hong, M., Akkaraju, G. R., Inoue, J., and Chen, Z. J. (2001) *Nature* **412,** 346-51
9. Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., Xiao, W., and Dixit, V. M. (2004) *Nature* **427,** 167-171
10. Shi, C. S. and Kehrl, J. H. (2003) *J Biol Chem* **278,** 15429-34
11. Xiao, W., Chow, B. L., Fontanie, T., Ma, L., Bacchetti, S., Hryciw, T., and Broomfield, S. (1999) *Mutat Res* **435,** 1-11
12. Xiao, W., Chow, B. L., Broomfield, S., and Hanna, M. (2000) *Genetics* **155,** 1633-41
13. Xiao, W., Lin, S. L., Broomfield, S., Chow, B. L., and Wei, Y. F. (1998) *Nucleic Acids Res* **26,** 3908-14
14. Yamaguchi, T., Kim, N. S., Sekine, S., Seino, H., Osaka, F., Yamao, F., and Kato, S. (1996) *J Biochem (Tokyo)* **120,** 494-97
15. Doublié, S. (1997) *Methods Enzymol.* **276,** 523-530
16. Otwinowski, Z. and Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* 276, 301-326. 97.
17. Hendrickson, W. A. and Ogata, C. M. Phase determination from multiwavelength anomalous diffraction experiments. Meth. Enzymol 276, 494-522. 97.

18. Terwilliger, T. C. Solve: an automated crystallographic structure solution program for MIR and MAD. Acta Crystallogr. A43, 1-5. 87.

19. Collaborative Computational Project, Number 4. 1994. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* D50, 760-763. 94.

20. Lamzin, V. S. and Wilson, KS. Automated refinement of protein models. Acta Cryst. D49, 129-149. 93.

21. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgarrd, M. O. *Acta Cryst.* A47, 110-119. 91.

22. Vagin, A. and Teplyakov, A. (2000) *Acta Crystallogr D Biol Crystallogr* **56 Pt 12,** 1622-4

23. Read, R. J. *Acta Crystallogr. A* 42, 140-149. 86.

24. Brunger, A. T. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Cryst D 54, 905-921. 98.

25. Murshudov, G. N., Vagin, A., and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst* D53, 240-255. 97.

26. Laskowski, R. A., MAcArthur, M. W., Moss, D. S., and Thornton, J. M. PROCHECK. *J. App. Cryst.* 26, 283. 93.

27. Cook, W. J., Jeffrey, L. C., Carson, M., Chen, Z., and Pickart, C. M. (1992) *J Biol Chem* **267,** 16467-71

28. Brünger, A. T. The Free R value: a Novel Statistical Quality for Assessing the Accuracy of Crystal Structures. *Nature* 355, 472-474. 92.

29. Pickart, C. M. (2000) *Trends Biochem Sci* **25,** 544-8

30. Scheffner, M., Smith, S., and Jentsch, S. (1998) The Ubiquitin-Conjugation System. Ubiquitin and the Biology of the Cell. Plenum Press, New York

31. Hofmann, R. M. and Pickart, C. M. (2001) *J Biol Chem* **276,** 27936-43

32. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6

33. Liu, Q., Jin, C., Liao, X., Shen, Z., Chen, D. J., and Chen, Y. (1999) *J Biol Chem* **274,** 16979-87

34. Miura, T., Klaus, W., Gsell, B., Miyamoto, C., and Senn, H. (1999) *J Mol Biol* **290,** 213-28

35. Hamilton, K. S., Ellison, M. J., Barber, K. R., Williams, R. S., Huzil, J. T., McKenna, S., Ptak, C., Glover, M., and Shaw, G. S. (2001) *Structure (Camb)* **9,** 897-904

36. Zheng, N., Wang, P., Jeffrey, P. D., and Pavletich, N. P. (2000) *Cell* **102,** 533-9

37. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001) *Cell* **105,** 711-20

38. McKenna, S., Hu, J., Moraes, T., Xiao, W., Ellison, M. J., and Spyracopoulos, L. (2003) *Biochemistry* **42,** 7922-30

39. Ulrich, H. D. (2003) *J Biol Chem* **278,** 7051-8

40. Palma, P. N., Krippahl, L., Wampler, J. E., and and Moura, J. J. (2000) *Proteins* **39,** 372-384

41. McKenna, S., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., Spyracopoulos, L., and Ellison, M. J. (2003) *J Biol Chem* **278,** 13151-8

Chapter 3

A Structural and Functional Comparison of Uev1a and Mms2

3.1 Summary

The ubiquitin (Ub) conjugating enzyme variants (UEV) Uev1a and Mms2 interact with Ubc13 to form heterodimeric complexes with different biological functions. Uev1a•Ubc13 participates in signaling pathways leading to NF-κB activation while Mms2•Ubc13 functions in DNA repair. In light of the 90% identity that exists between Mms2 and Uev1a, the structural features that contribute to these functional differences must be relatively few. In the present work, we compared the monomeric and heterodimeric structures of Mms2 and a derivative of Uev1a (reported here) that lacks its unique amino-terminal extension. Based both on thermodynamic measurements and enzyme assays, we found that the amino-terminus of Uev1a weakened the interaction between Uev1a and either Ub or Ubc13 but stimulated poly-Ub chain assembly. This observation suggested to us that the amino-terminal extension of Uev1a stimulated chain assembly by facilitating dissociation of the Uev1a-Ubc13-Ub_n complex that is an obvious prerequisite for catalytic turnover. We tested this idea by introducing amino acid substitutions at the Mms2•Ubc13 interface that were designed to favor the dissociation of the heterodimer. Two of these replacements proved to have a positive influence on poly-Ub chain assembly. Therefore, this subtle property of the chain assembly mechanism may contribute at least in part to the biological differences attributed to Uev1a and Mms2.

3.2 Introduction

Protein ubiquitination defines a class of post-translational modifications in which ubiquitin (Ub) becomes covalently linked to a target protein. In many instances the covalently linked Ub becomes the site of further ubiquitination leading to the synthesis of a poly-ubiquitin (poly-Ub) chain. These poly-Ub chains are composed of isopeptide bonds in which the carboxy terminus of one Ub is linked to a lysine side chain of another Ub. The specific lysine residue employed in these chain linkages determines the fate of the target protein to which the chain is attached[1-3]. Lys48-linked poly-Ub chains direct the protein target to the proteasome for degradation[4]. On the other hand, Lys63-linked poly-Ub chains stimulate the signaling function of specific proteins involved in NF-κB activation[5-7] and function in DNA repair pathways[8;9]. The ability of different chain configurations to direct proteins to different fates doubtless stems from their unique topologies[10;11].

The initial steps involved in the assembly of either chain type are identical. The carboxy-terminus of Ub is first activated through the action of an ATP dependent ubiquitin activating enzyme (E1) resulting in an E1~Ub thiolester. Ub is then transferred from the E1 to an Ub conjugating enzyme (E2) through a transthiolation reaction that generates an E2~Ub thiolester. At this point, the pathway diverges mechanistically to yield the two different chain types. The details of Lys48 chain assembly remain an important question. By comparison, there is a better understanding of the mechanism that leads to Lys63-linked chains. The formation of an isopeptide bond between the carboxy-terminus of one Ub and Lys63 is catalyzed

by a heterodimer that is composed of the E2 Ubc13 and an Ub conjugating enzyme variant (UEV)[8]. The UEV monomer is structurally homologous to Ubc13 but lacks the active-site cysteine that is required for E2~Ub thiolester formation[12;13]. UEVs form a strong interaction with Ubc13 in both the Ub charged and uncharged form[13]. A key role of the UEV is to orient Lys63 of the non-covalently bound Ub close to the carboxy terminus of the Ub thiolester linked to Ubc13.[14;15].

Two functionally distinct UEVs (Mms2 and Uev1a) have been identified that poly-ubiquitinate different targets. Mms2 poly-ubiquitinates the DNA repair protein PCNA[9] while Uev1a poly-ubiquitinates the signal transduction proteins Traf6/2 or NEMO, that are required for NF-κB activation[5-7]. Both UEVs are largely identical, differing at only twelve amino acid positions in their core domains and at their amino-termini, where Uev1a possesses a 30 amino acid extension compared to the 5 amino acid extension found in Mms2.

To determine how these differences might affect their structure and function we compared the crystal structures of Mms2 and Uev1a both alone and in association with Ubc13 and also compared their relative abilities to support poly-Ub chain assembly *in vitro*. Together, our data indicates that relative to Mms2, the amino-terminal extension of Uev1a functions to weaken its interactions with both Ub[16] and Ubc13 and that the weakening of these interactions functions to stimulate poly-Ub chain assembly.

3.3 Experimental Procedures:

3.3.1 Protein Expression. Uev1a, Ubc13, Mms2 open reading frames (ORF) were cloned into a GST fusion vector pGEX6 as previously described [13;16]. ORFs for the Uev1a derivatives (Uev1aΔ30, and Uev1a K11R, Uev1aK15R, Uev1aK26R, Uev1aK30R) were synthesized by PCR based site-directed mutagenesis using the Uev1a-pGex6 vector as a template. ORFs for the Ubc13 derivatives were synthesized by PCR based site directed mutagenesis using the Ubc13-pGex6 vector as a template. GST-Uev1a and GST-Mms2 were expressed and purified as previously described[13] except that they were not treated with precission protease. E1 (Uba1) and ^{35}S-Ub were expressed and purified as previously described[17].

3.3.2 Protein purification and crystallization. Over-expression and purification of human Uev1a, Mms2 and Ubc13 was performed as described [13;16]. Crystallization conditions for Uev1a (Uev1aΔ30) were obtained from a Hauptman-Woodward Institute screen. From this initial screen, crystals of Uev1aΔ30 were grown at 20-24 °C using the hanging drop vapor diffusion technique. Crystals of Uev1aΔ30 were obtained by mixing 6 mg/mL in 50 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EDTA and 1 mM dithiothreitol (DTT) with 1 μL of well solution (20% (w/v) PEG 6000, 100 mM sodium thiosulfate pentahydrate, 100 mM citrate -pH 6.5) and then equilibrated against the well solution. Crystals of Uev1aΔ30•Ubc13 were obtained by mixing 1 μL of 10 mg/mL of protein in 50 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EDTA and 1 mM dithiothreitol (DTT) with 1 μL of well solution (18% (w/v) PEG 6000, 100

mM citrate -pH 6.5) and then equilibrated against the well solution.

3.3.3 Structure determination. Diffraction data for Uev1aΔ30 (1.7 Å) were collected from a single crystal at 105 K at Beamline 8.3.1 at the Advanced Light Source in Berkeley, California. Diffraction data for Uev1aΔ30•Ubc13 (2.5 Å) were collected from a single crystal at 105 K with a RaxisIV image plate detector and a Rigaku RU-H3R rotating anode generator. Data were scaled and integrated with DENZO and SCALEPACK[18]. The structure of Uev1aΔ30 was solved by molecular replacement using MOLREP[19] with Mms2 (PDB accession code: 1J74) as the search model. Rigid body refinement of the initial solution gave a starting R-factor of 36.7% and a correlation coefficient of 0.68 using reflections from 8-4 Å. Structures of the Uev1aΔ30•Ubc13 complex and monomeric Uev1aΔ30 were refined through several cycles. Manual rebuilding was performed with reference to Sigma A-weighted[20] 2|Fo| – |Fc| electron density maps calculated in CCP4[21]. Refinement was performed with REFMAC5[22] using torsion angle dynamics and maximum likelihood targets. The quality of the model was assessed by PROCHECK[23].

3.3.4 Ubiquitin Conjugation Reactions. Reactions (0.5 mL final volume) containing E1 (20 nM), ^{35}S-Ub (2.5 μM), and E2 (250 nM) in an ATP cocktail (Reaction Buffer: 10 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 5 mM ATP, 0.6 units/mL inorganic phosphatase) were incubated at 30 °C for 1, 1.5 or 16 hr. The specific protein concentrations, reaction components, and reaction times used are given in the

figure legends of Figure 3 and 4. Reactions were quenched using DTT and terminated by the addition of trichloroacetic acid (final, 10%) and processed for SDS-PAGE (12%). SDS-PAGE gels were analyzed by autoradiography.

3.3.5 Isothermal Titration Calorimetry. A VP-ITC MicroCalorimeter (Microcal, Northampton, MA) was used to analyze binding of Uev1a and Uev1aΔ30 with Ubc13 or Ubc13 derivatives with Mms2. Proteins were dialyzed against 50 mM Tris, pH 7.5, 150 mM NaCl, and 1 mM ETDA and degassed prior to analysis. The E2 (7 μM) was injected into the sample cell, and the Uev (~70 μM solution) was placed into the syringe, or vice versa. The dialysis buffer was placed in the reference cell. Either the Uev or E2 was titrated against dialysis buffer to obtain the heat of dilution. The following parameters were used in the titration: 30 °C, 10 μl injections, and 4 min between injections with stirring at 305 rpm. The data was fit to a titration curve using the program Microcal Origin (v 5.0) to extract thermodynamic parameters.

3.4 Results

3.4.1 The Structures of Uev1aΔ30 and Uev1aΔ30•Ubc13 Heterodimer. To gain insight into the biological function of Uev1a and Ubc13, we solved the structure of Uev1aΔ30 and Uev1aΔ30•Ubc13 and compared these with the previously determined structures of Mms2 and Mms2•Ubc13. The structure of Uev1aΔ30 was solved by molecular replacement using a poly-alanine model of Mms2 (PDB accession code: 1J74). The density for the core 140 residues of Uev1aΔ30 is well ordered (**Figure 3.1**) and adopts a very similar structure to its homologue Mms2 with an RMSD of 0.5Å. This structural homology was not surprising given that outside of the distinct amino-termini, Uev1a and Mms2 differ at only 12 amino acid positions (**Figure 3.2**).

As with Mms2, Uev1aΔ30 possesses a non-canonical Ubc fold that consists of an amino-terminal α-helix followed by four anti-parallel β-strands, a 3-10 helix and a supporting α-helix, but lacks the active site cysteine and the two carboxy-terminal helices found in the core domain of Ubcs (**Figure 3.1**) [12;24]. Based on a structural alignment, the residues involved in binding Ub and Ubc13 are identical in both Uev1aΔ30 and Mms2 (**Figure 3.2**). Interestingly, the positions of the 12 amino acids that differ between Uev1aΔ30 and Mms2 cluster on a surface of each protein that does not overlap with the Uev1a and Mms2 surfaces used in Ub and Ubc13 interactions (**Figure 3.1B**).

The structural similarities observed for both Uev1aΔ30 and Mms2 indicated that they would exhibit similar modes of interaction with Ubc13. This assumption

Table 3.1 Crystallographic Statistics for Uev1aΔ30 and Uev1aΔ30•Ubc13

Data collection

	Uev1aΔ30	Uev1aΔ30•Ubc13
Resolution (Å)	55-1.60	22-2.5
Completeness (%)	77.8	94
Observations	125536	34521
Unique reflections	17395	9870
Mosaicity (°)	0.3	1.1
<I>/ σ(I)	5.5	13.8
R_{sym} [a]	0.077	0.086
Cell (space group)	43.1 54.1 109.8 α=β=γ=90° (c2221)	44.9 75.1 91.2 α=β=γ=90° (p212121)

Refinement

Refinement		
Resolution (Å)	55-1.6	22-2.5
R_{fac}	19.8	24.9
R_{free}	21.3	29.6
R.M.S Deviations		
Bond length (Å)	0.012	0.01
Bond Angle (°)	1.24	1.4
# residues	140	286
# water molecules	80	0

[a] $R_{sym} = \Sigma_{i,h} |I(i,h) - \langle I(h)\rangle| / \Sigma_{i,h} |I(i,h)|$ where $I(i,h)$ and $\langle I(h)\rangle$ are the ith and mean intensity of reflection h.

[b] R_{free} is calculated using 5% of the reflections[30].

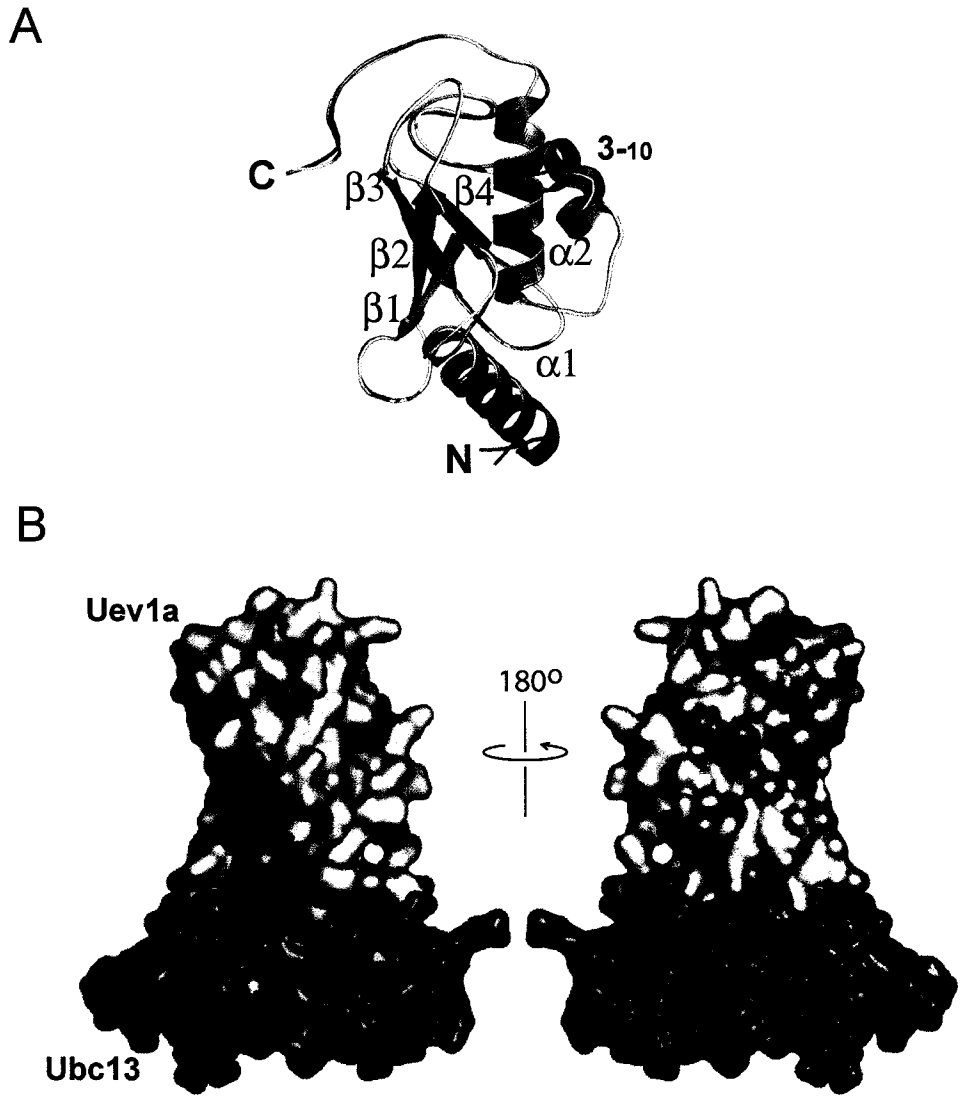


Figure 3.1 A Structural Comparison of Mms2 and Uev1a. A. Overlay of the Uev1a and Mms2 crystal structures. Shown are ribbon representations of Mms2 in green and Uev1aΔ30 in olive that were aligned using O^{29}. The RMSD between these structures was 0.4Å. **B.** Structure of the Uev1aΔ30•Ubc13 complex. A surface representation of Uev1aΔ30 bound to Ubc13 is shown in two views rotated 180° relative to one another. Ubc13 is shown in teal and its active-site cysteine is highlighted in yellow. Uev1a is shown in olive and its amino-terminus is highlighted in blue. Residues that vary between Mms2 and Uev1a, not including their respective amino-termini, are highlighted in green. Residues involved in the UEV•Ub non-covalent interaction and in the Ubc13~Ub thiolester are highlighted in red on both the Uev1aΔ30 and Ubc13 structures.

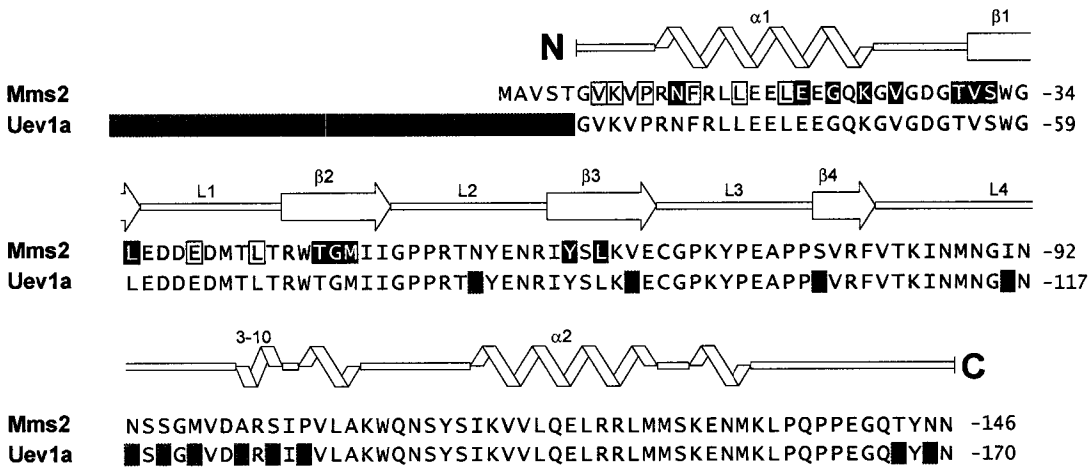


Figure 3.2 Sequence Analysis of Mms2 and Uev1a. Shown is the structure based sequence alignment of Mms2 and Uev1a. Mms2 residues highlighted in red and yellow have been shown to interact with Ub and Ubc13 respectively[14]. Mms2 residues within the interface of Mms2 and Ubc13 are boxed[12]. Uev1a residues highlighted in green identify the amino acid differences between Uev1a and Mms2.

was tested directly by solving the Uev1aΔ30•Ubc13 crystal structure and comparing it to that of Mms2•Ubc13. When superimposed these structures were virtually identical with an RMSD of 0.39Å. Comparing the structures of Uev1aΔ30 to Uev1aΔ30•Ubc13 shows that upon binding of Ubc13 the amino-terminal helix of Uev1aΔ30 unravels by one turn in a similar fashion to that observed for Mms2 when it binds Ubc13 [12]. This observation suggests that the respective amino-termini of Uev1a and Mms2 become repositioned upon binding to Ubc13. Comparing the core structures reveals no obvious reason for their different biological roles, therefore we must conclude that the amino terminus must play a role to distinguish the functions of these two UEVs.

3.4.2 The amino-terminus of Uev1a weakens the interaction with Ubc13. Overall, the comparison of the Uev1aΔ30 and Mms2 structures either alone or in complex with Ubc13 showed them to be virtually identical. In spite of their high degree of identity these proteins have been found to play distinct biological roles that should be accounted for within their structures. The most obvious structural distinction between these two UEVs lies in their respective amino-termini. Structural comparisons described above indicate that the conformational change at the amino-termini of Uev1a would function to reposition its amino-terminal extension upon binding Ubc13. Previous studies have also shown that the Uev1a•Ub (93±15 nM) interaction is two fold weaker than that of Mms2•Ub (211±14 nM) [16]. It is unlikely that the difference in interaction can be accounted for by the 12 amino acids that differ in the core domain because they cluster to a surface not involved in Ub binding,

therefore it is likely due to the amino-terminal differences. These observations suggest that the distinct amino-termini of these proteins might affect the interaction with Ubc13. We tested this possibility by measuring the dissociation constants (K_D) of the Uev1a•Ubc13 and Uev1aΔ30•Ubc13 using isothermal titration calorimetry (ITC) (**Figure 3.3**).

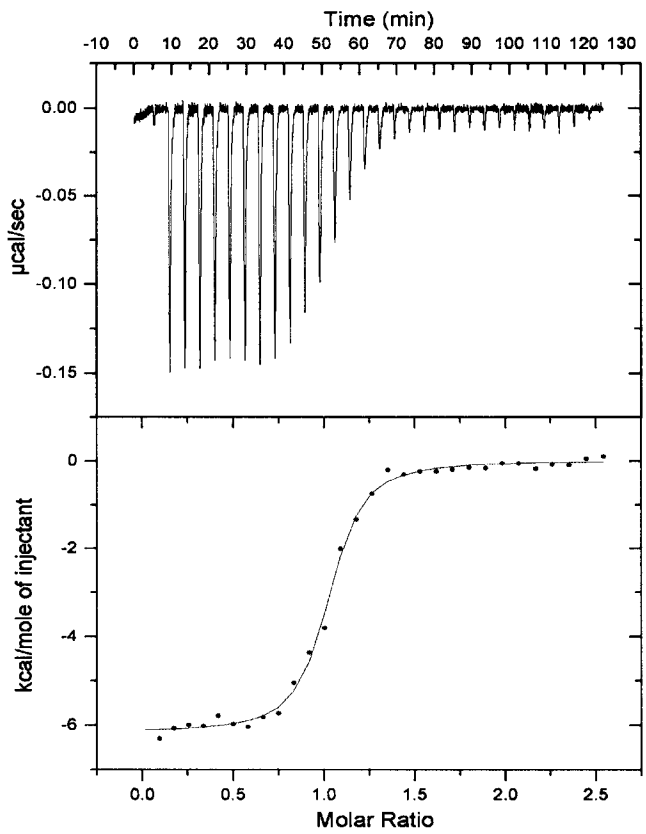
This analysis showed that the interaction between Uev1a•Ubc13 (63 nM) was slightly weaker than when the amino terminus of Uev1a was deleted Uev1aΔ30•Ubc13 (48 nM), a value that was virtually identical to that previously determined for of Mms2•Ubc13 (49 nM). Taken together, these observations specify that the amino terminal extension of Uev1a destabilizes interactions with both Ub[16] and Ubc13 relative to that of Mms2.

3.4.3 The amino-terminal extension of Uev1a directs poly-ubiquitination.

We next examined whether the differences between Mms2 and Uev1a in binding Ub and Ubc13 affected their catalytic properties *in vitro*. We used the same *in vitro* assay used to assess Lys63 linked Ub_2 synthesis by the Mms2•Ubc13 complex [13]. The reactions consisted of purified yeast E1, ATP, and Mg^{2+} as well as purified radiolabeled ^{35}S-Ub. To this mixture was added combinations of various Ubc13, Mms2 and Uev1a derivatives. After incubation the synthesis of poly-Ub chains was monitored using SDS-polyacrylamide electrophoresis followed by autoradiography.

When E1, Ub and Ubc13 are incubated together the principal reaction product produced was the monoubiquitinated Ubc13-Ub conjugate. This product results from the direct transfer of Ub from the Ubc13~Ub thiolester to Lys92 of Ubc13. When

A



B

	K_D
Uev1a	**62 ± 6nM**
Mms2*	**49 ± 7nM**
Uev1aΔ30	**47 ± 9nM**

Figure 3.3 ITC Analysis of Uev1a•Ubc13. Stability of the Uev1a•Ubc13 interaction is dependent upon the amino-terminal extension of Uev1a. **A**: Shown in the top panel is a representative thermogram for the titration of Ubc13 (sample cell) with Uev1a (titrant) using Isothermal Titration Calorimetry (ITC). The graph shows the heat evolved per injection. The ITC titration curve generated from this ITC experiment is shown in the bottom panel where the heat evolved per injection is shown as a function of the molar ratio of Uev1a to Ubc13. **B**: UEV•Ubc13 dissociation constants (K_D). Shown are the K_D values determined by ITC for the interactions between Ubc13 and various UEV derivatives including Mms2, Uev1a and Uev1aΔ30. The Mms2•Ubc13 K_D value (*) has been previously determined[16] and is included for the sake of comparison.

Lys92 of Ubc13 is mutated to Arg92 (Ubc13K92R) this monoconjugate band disappears[13]. To simplify the interpretation of the reaction products, subsequent reactions employed the Ubc13K92R derivative (**Figure 3.4, 3.5**).

Mms2 in combination with Ubc13K92R results in the synthesis of Lys63 linked Ub_2 (**Figure 3.4**) as the primary reaction product as previously described[13]. Alternatively when Uev1a is combined with Ubc13 a series of products that correspond to free chains and poly-Ub chains conjugated to Uev1a are synthesized (**Figure 3.4**). Therefore, from the results in figure 3, it is apparent that Uev1a is more effective at facilitating chain assembly than Mms2. Significantly, when Uev1aΔ30 is combined with Ubc13K92R the pattern of chain assembly is virtually indistinguishable from that of Mms2. Therefore the magnitude and type of chain assembly depends upon the UEV amino-terminal extension. It is likely that the Uev1a-Ub conjugates arise from the linkage of Ub to Lys residues within the amino-terminal extension (K11, K15, K26 and K30). This conclusion is based on the observation that single substitutions significantly reduce the levels of Uev1a-Ub conjugate (specifically K15 and K30). The ability of the amino-terminal extension to stimulate chain assembly does not appear to be an inherent property of its sequence or structure. For example, when GST is appended to the amino terminus of Mms2 (GST-Mms2) it allows for the synthesis of free poly-Ub chains (**Figure 3.5**). Similarly, when GST is appended to the amino-terminus of Uev1a (GST-Uev1a) the formation of free poly-Ub chains are favored over the Uev1a-Ub conjugate in a manner that is indistinguishable from Mms2. Taken together these observations

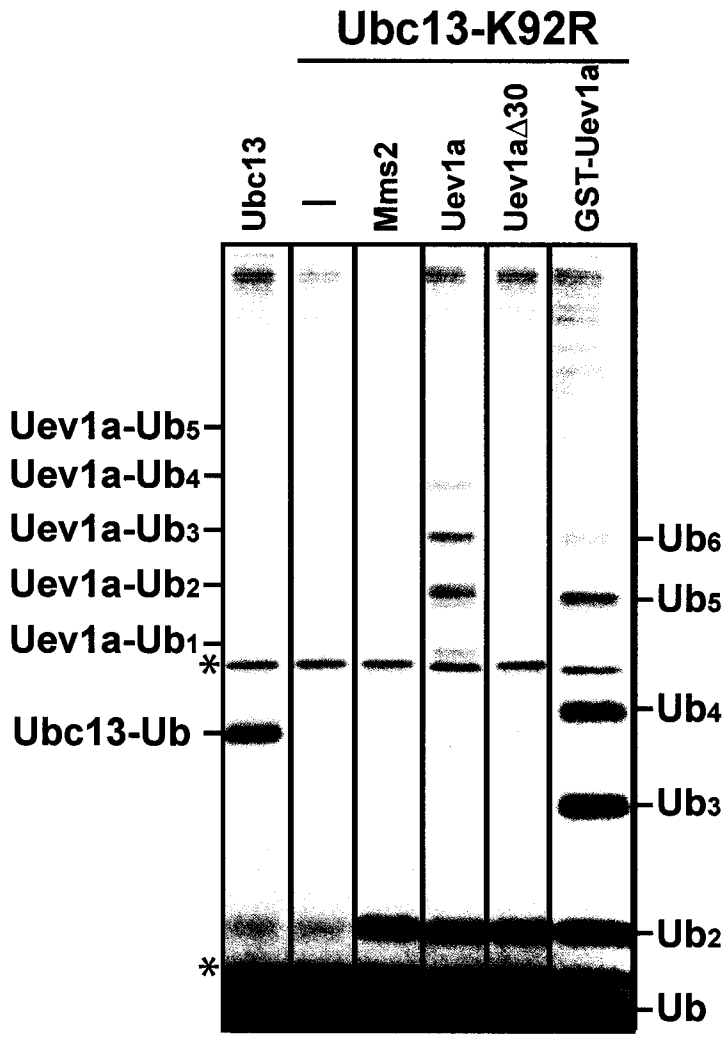


Figure 3.4 *In vitro* Ubiquitination Assay with Uev1a. The *in vitro* ubiquitination reactions included purified E1 (25nM) and ^{35}S-Ub (2.5mM) along with combinations of Ubc13, Uev1a and Mms2 derivatives (250nM) that are specified below. Reactions were incubated at 30°C for 90 min. and subsequently loaded onto and ran over an SDS polyacrylamide gel. Reaction products containing ^{35}S-Ub were subsequently visualized by autoradiography and included free poly-Ub chains (Ub_n) and poly-Ub chains conjugated to Uev1 (Uev1a-Ub_n). The * indicates contaminants present in the ^{35}S-Ub preparation. The amino terminal extension of Uev1a promotes poly-Ub chain assembly. *In vitro* reactions were carried out as described above and included Ubc13 alone and Ubc13K92R alone or in combination with either Mms2, Uev1a, Uev1aΔ30 or GST-Uev1a. GST was appended to the amino-terminus of full length Uev1a to generate the GST-Uev1a derivative.

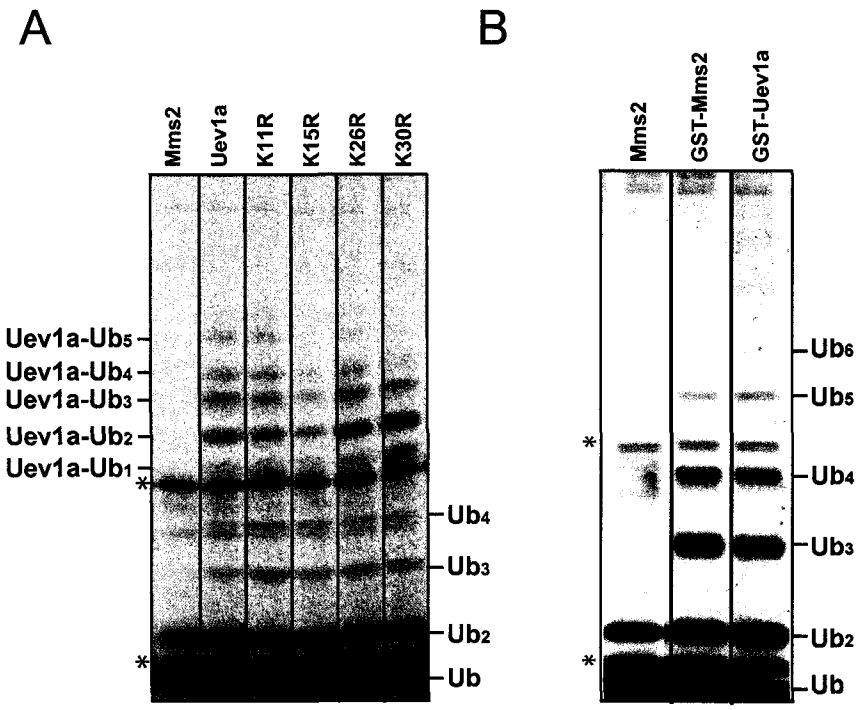
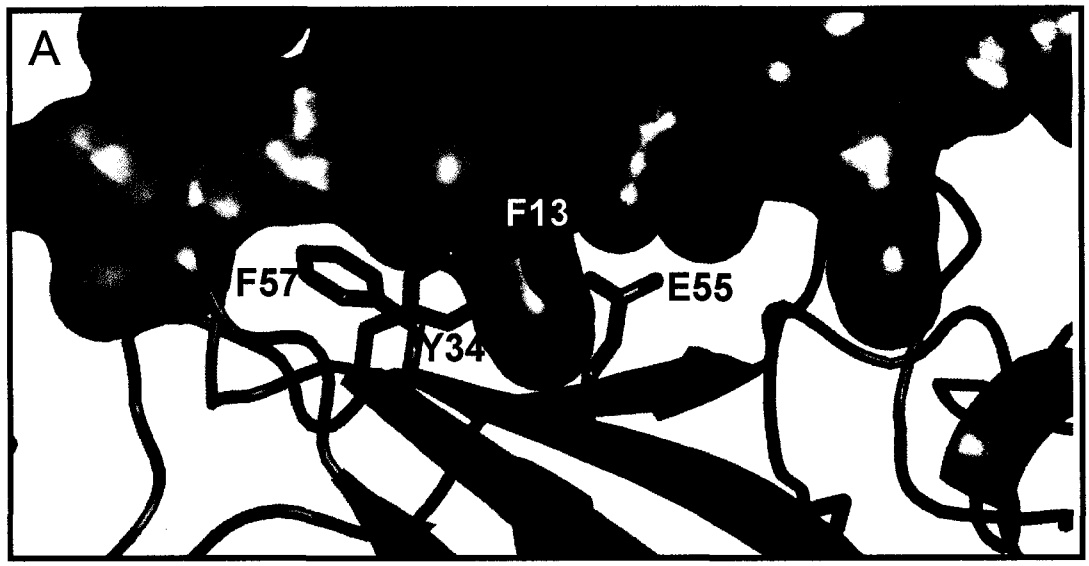


Figure 3.5 The Amino-Terminus of UEVs Functions to Stimulate Poly-Ub Chain Assembly. The *in vitro* ubiquitination reactions included purified E1 (25nM) and ^{35}S-Ub (2.5mM) along with combinations of Ubc13, Uev1a and Mms2 derivatives (250nM) that are specified below. Reactions were incubated at 30°C for 60 min. and subsequently loaded onto and ran over an SDS polyacrylamide gel. Reaction products containing ^{35}S-Ub were subsequently visualized by autoradiography and included free poly-Ub chains (Ub_n) and poly-Ub chains conjugated to Uev1a (Uev1a-Ub_n). The * indicates contaminants present in the ^{35}S-Ub preparation. **A**: Identification of K residues in the amino terminal extension of Uev1a that define sites of poly-Ub chain conjugation. *In vitro* ubiquitination reactions were carried out using Ubc13K92R in combination with either Mms2, Uev1a or one of a number of K to R derivatives of Uev1a that included K11R, K15R, K26R and K30R. These K residues are found within the amino terminal extension of Uev1a. **B:** A GST-Mms2 fusion protein supports poly-Ub chain assembly. In vitro ubiquitination reactions were carried out using Ubc13K92R in combination with either Mms2, GST-Mms2 or GST-Uev1a. As in the case of the GST-Uev1a, GST was appended to the amino terminus of full length Mms2 to generate the GST- Mms2 derivative.

demonstrate that amino-terminal extensions on UEVs, independent of sequence or structure, stimulate poly Ub chain assembly.

3.4.4 Destabilization of the Mms2•Ubc13 interface stimulates poly-Ub chain assembly. Based on the fact that the amino-terminal extension weakens the interaction between Uev1a and Ubc13, we reasoned that amino acid substitutions that destabilize the Ubc13•Mms2 interface may stimulate chain assembly. We introduced single amino acid substitutions into Ubc13 that, based on the crystal structure of the Mms2•Ubc13 complex, would be expected to destabilize this interaction (**Figure 3.6**). Three Ubc13 derivatives carrying the single amino acid substitutions Y34A, E55A or F57A were created. The measured K_Ds indicate that these substitutions affect the interaction to various degrees (**Figure 3.6B**). For example, F57A abolishes the interaction between Mms2 and Ubc13, whereas Y34A and E55A decreased the interaction 8-fold and 21-fold respectively.

We tested the ability of each of these derivatives to assemble poly-Ub chains in the presence of Mms2. As expected, complete disruption of the interface eliminates poly-Ub chain assembly supporting the notion that formation of the Mms2•Ubc13 complex is a prerequisite for efficient Ub_2 synthesis. When the other two Ubc13 derivatives were employed in the reactions an increase in the assembly of Ub_n chains greater than two Ubs in length was observed (**Figure 3.6C**). Furthermore, while the E55A substitution functioned to destabilize the Mms2•Ubc13 to a greater degree than the Y34A substitution, it also stimulated poly-Ub chain assembly better



B

	K_D (nM)	Ub_2	Ub_n
Ubc13	49 ± 6	18.4	3.7
Y34A	410 ± 35	21.4	6.1
E55A	1050 ± 100	20.8	7.3
F57A	ND	1.25	1.1

C

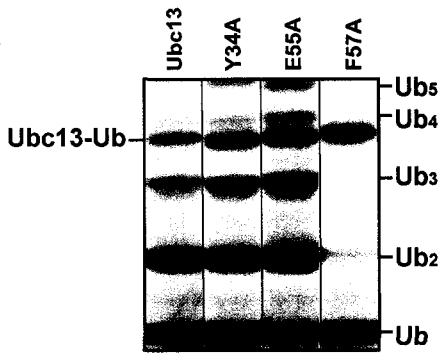


Figure 3.6 Limited Destabilization of the Mms2•Ubc13 Interaction Stimulates Poly-Ub Chain Assembly. A: The Mms2•Ubc13 interface. Ubc13 is shown in teal as a ribbon diagram while Mms2 is shown in yellow as a space-filled representation. Key residues in the interface are also indicated and include F13 in Mms2 as well as Y34, E55, and F57 in Ubc13. **B:** Amino acid substitutions at Ubc13 interface residues affect its interaction with Mms2. K_Ds as determined by ITC for the interaction between Mms2 and either Ubc13 or one the Ubc13 derivatives that include Y34A, E55A, and F57A are given. ND (not determined) refers to an interaction that was too weak to be detectable by ITC. The percentage of total Ub incorporated into Ub_2 or Ub_n (n=3, 4 and 5) poly-Ub chains as determined from the *in vitro* ubiquitination assays displayed in panel C is also given. **C**: Perturbation of the Mms2•Ubc13 interface stimulates poly-Ub chain assembly. Reactions included purified E1 (25nM), ^{35}S-Ub (2.5mM), Mms2 (250nM) along with either Ubc13 (250nM) or one of the Ubc13 derivatives (250nM) as indicated. Samples were quenched using DTT and subsequently loaded onto and ran over an SDS polyacrylamide gel. Reaction products containing ^{35}S-Ub were subsequently visualized by autoradiography.

than the Y34A substitution. These observations suggest that destabilization of the Mms2•Ubc13 interface contributes to increased poly-Ub chain assembly.

3.5 Discussion

The different biological pathways in which Uev1a and Mms2 participate must be accounted for by the limited structural differences that exist between them. The structural comparison of each of these Uevs alone or complexed with Ubc13 illustrate that they differ only in their amino-termini and at 12 amino acid positions that cluster to a surface within the core. Notably this cluster does not overlap with the UEV surfaces defining the Ub or Ubc13 binding sites (**Figure 3.1**), consistent with the notion that they play no direct role in poly-Ub chain catalysis. The amino acid differences between these UEVs raises the possibility that the surface cluster serves as a binding site for interacting partners specific to each UEV. This is supported by a highly non-conservative surface substitution that occurs at analogous positions 79 and 104 in Mms2 and Uev1a in which a serine in Mms2 is replaced by a structurally well-ordered phenylalanine.

The most likely structural determinant that differentiates these UEVs in terms of Ub chain catalysis is their amino-termini (5 residues in Mms2 and 30 residues in Uev1a - **Figure 3.2**). In both the case of Mms2 and Uev1aΔ30, the amino-terminus undergoes a rearrangement that places them close to the amino-terminal region of Ubc13. From the observations presented here we have concluded that the amino-terminus of Uev1a affects poly-Ub chain assembly by affecting the strength of

interaction between the constituents of the catalytic complex ((UEV•Ub)•Ubc13~Ub). The evidence for this conclusion is based on the following: 1) Full length Uev1a directs the assembly of both free and conjugated Ub chains to a greater extend than either Uev1aΔ30 or Mms2. 2) Appending GST to the amino-terminus of either Uev1a or Mms2 stimulates the synthesis of free poly-Ub chains in both cases. Therefore the extension facilitates poly-Ub assembly in a manner that is independent of its sequence. 3) The amino-terminal extension of Uev1a weakens the interaction between itself an either Ub[16] or Ubc13. 4) Amino acid substitutions directed at weakening the interaction between Mms2 and Ubc13 stimulate poly-Ub chain assembly.

These observations can be understood in terms of classical enzyme catalysis. Rates of enzyme catalysis is based not only on the affinity of an enzyme for its substrate but also on the efficacy of releasing its product, so called turnover. In the case of poly-Ub chain assembly an important factor in chain elongation would be the ability to clear the active site of two covalently linked Ub molecules (the product).

In the UEV based mechanism of Ub chain assembly there are two Ub binding sites[14]. One binding site on Ubc13 accommodates a donor Ub linked to the active site while the acceptor Ub occupies another binding site partitioned between Ubc13 and the UEV that orients the ubiquitin's Lys63 adjacent to the active site cysteine of Ubc13. By all accounts the Ub thiolester binding site on Ubc13 is weak, whereas the binding affinity of the acceptor Ub is measurably stronger[14]. Upon catalysis, two Ub molecules linked through a Lys63 residue (Ub_2) reside in the active site of the

UEV•Ubc13 complex. In order to promote poly-Ub chain assembly this product must be expelled. As Ub_2 is bound by both the UEV and Ubc13 in complex, then a less tight complex would promote dissociation thereby freeing the active site for subsequent thiolester formation, freeing the product (poly-Ub), and alleviating the non-covalent binding site within the E2•UEV complex to acquire the product as the acceptor (target). As such, the lower affinity Uev1a exhibits for both Ub and Ubc13 permits the complex to more efficiently release its product therefore enabling it to be more active at poly-Ub chain assembly by comparison to Mms2. Support for this premise also derives from the observation that unlike their human counterparts yeast Mms2 and Ubc13 show a weaker affinity for each other[25], but support greater levels of poly-Ub chain assembly[26].

The observations made with the yeast Mms2•Ubc13 complex indicated the possibility that the weaker interaction between these proteins functions to stimulate poly-Ub chain assembly. We tested this possibility by introducing amino acid substitutions into Ubc13 that destabilized its interaction with Mms2. Those substitutions that did not completely disrupt the interaction showed an increase in the assembly of Ub_n chains (**Figure 3.5**). This indicated that the affinity between a Uev and Ubc13 defines the level of poly-Ub chain assembly.

The presence of conjugation sites within the Uev1a extension does not appear to be the principal determinant in the stimulation of poly-Ub chain assembly given that a GST-Uev1a fusion protein also shows a similar affect. In this case, GST appears to occlude the K residues from accepting Ub such that the GST-Uev1a fusion

redirects poly-Ub chain assembly from conjugated chains to Ub_n chains. Furthermore, the GST-Mms2 fusion protein shows an equivalent capacity to direct Ub_n chain assembly as observed for GST-Uev1a. It remains to be determined whether or not conjugated poly-Ub chains may be assembled onto Uev1a *in vivo*. If so, these conjugated chains may aid in determining the functional specificity of Uev1a in an analogous fashion to that observed for the ubiquitinated TRAF proteins[27].

Curiously, upon complex formation with Ubc13 the amino terminal region of both UEVs is likely removed from the site of catalysis and is repositioned close to the region on Ubc13 that has been implicated in the binding of E3-RING finger domain containing proteins such as Rad5 in the case of the yeast Mms2•Ubc13 complex [25] and Traf6 or Traf2 in the case of the Uev1a•Ubc13 complex [28]. This raises the possibility that the amino-termini of each Uev may function to distinguish the interactions between an E3 and a given UEV•Ubc13 complex. Alternatively, we have demonstrated that the effect of the amino-terminal extension on the interaction between a UEV and Ubc13 affects the catalytic properties of the heterodimer. Of further consideration is a splice variant of Uev1a, Uev1B that results in an even larger amino-terminal extension. Although the function of this homolog is not known, it has lost the ability to interact with Ubc13. Taken together we can conclude that the amino-terminus of UEVs define their biological function.

3.6 Reference List

1. Pickart, C. M. (2000) *Trends Biochem Sci* **25,** 544-8
2. Hicke, L. (2001) *Nat Rev Mol Cell Biol* **2,** 195-201
3. Passmore, L. A. and Barford, D. (2004) *Biochem J* **379,** 513-25
4. Voges, D., Zwickl, P., and Baumeister, W. (1999) *Annu Rev Biochem* **68,** 1015-68
5. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z. J. (2000) *Cell* **103,** 351-61
6. Shi, C. S. and Kehrl, J. H. (2003) *J Biol Chem* **278,** 15429-34
7. Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., Xiao, W., and Dixit, V. M. (2004) *Nature* **427,** 167-171
8. Hofmann, R. M. and Pickart, C. M. (1999) *Cell* **96,** 645-53
9. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002) *Nature* **419,** 135-41
10. Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C., and Fushman, D. (2004) *J Biol Chem* **279,** 7055-63
11. Varadan, R., Walker, O., Pickart, C., and Fushman, D. (2002) *J Mol Biol* **324,** 637-47
12. Moraes, T. F., Edwards, R. A., McKenna, S., Pastushok, L., Xiao, W., Glover, J. N., and Ellison, M. J. (2001) *Nat Struct Biol* **8,** 669-73
13. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6
14. McKenna, S., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., Spyracopoulos, L., and Ellison, M. J. (2003) *J Biol Chem* **278,** 13151-8
15. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001) *Cell* **105,** 711-20
16. McKenna, S., Hu, J., Moraes, T., Xiao, W., Ellison, M. J., and Spyracopoulos, L. (2003) *Biochemistry* **42,** 7922-30
17. Varelas, X., Ptak, C., and Ellison, M. J. (2003) *Mol Cell Biol* **23,** 5388-400

18. Otwinowski, Z. and Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* 276, 301-326. 97.

19. Vagin, A. and Teplyakov, A. (2000) *Acta Crystallogr D Biol Crystallogr* **56 Pt 12,** 1622-4

20. Read, R. J. *Acta Crystallogr. A* 42, 140-149. 86.

21. Collaborative Computational Project, Number 4. 1994. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* D50, 760-763. 94.

22. Murshudov, G. N., Vagin, A., and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst* D53, 240-255. 97.

23. Laskowski, R. A., MAcArthur, M. W., Moss, D. S., and Thornton, J. M. PROCHECK. *J. App. Cryst.* 26 , 283. 93.

24. Pickart, C. M. (2001) *Annu Rev Biochem* **70,** 503-33

25. Ulrich, H. D. (2003) *J Biol Chem* **278,** 7051-8

26. Hamilton, K. S., Ellison, M. J., Barber, K. R., Williams, R. S., Huzil, J. T., McKenna, S., Ptak, C., Glover, M., and Shaw, G. S. (2001) *Structure (Camb)* **9,** 897-904

27. Wang, C., Deng, L., Hong, M., Akkaraju, G. R., Inoue, J., and Chen, Z. J. (2001) *Nature* **412,** 346-51

28. Wooff, J., Pastushok, L., Hanna, M., Fu, Y., and Xiao, W. (2004) *FEBS Lett* **566,** 229-33

29. Jones, T. A. , Zhou, J. Y., Cowan, S. W., and Kjeldgarrd, M. O. *Acta Cryst.* A47, 110-119. 91.

30. Brünger, A. T. The Free R value: a Novel Statistical Quality for Assessing the Accuracy of Crystal Structures. *Nature* 355, 472-474. 92.

Chapter 4:

A Structural and Functional Analysis of Active Site Residues in hUbc13.

4.1 Summary:

Ubiquitin conjugating enzymes or E2s mediate the covalent attachment of Ub to lysine residues on targets through formation of an isopeptide bond. E2s can catalyze the formation of chains of isopeptide linkages (polyUb chains) suggesting that the properties required to carry out isopeptide bond formation are housed within the E2 catalytic domain. The E2 hUbc13 was used to probe the mechanism of isopeptide bond formation within ubiquitination. hUbc13 builds Lys63 linked ubiquitin chains in the presence of its partner Mms2. Using reconstituted components of the Ub system, a set of ubiquitination assays have identified an Asp and Asn residue that are critical for isopeptide bond formation within this E2. Crystal structures of the inactive hUbc13 derivatives confirm that the Asn and Asp do not affect the active site structurally but rather must have catalytic roles in isopeptide bond formation, specifically acting as a hydrogen bond donor and as a general base, respectively. These residues are highly conserved in E2s supporting a general catalytic mechanism for E2s within ubiquitination.

4.2 Introduction:

Over the last two decades, the ubiquitin (Ub) system has emerged as a central and diversely functional post-translational regulator of eukaryotic cell physiology[1-4]. Underlying this richness of function is the system's capacity to selectively ubiquitinate a range proteins targets with a variety of poly-Ub chains of various structural configurations that owe their existence to the isopeptide bond[5]. An isopeptide bond links the carboxy-terminal carbon of Ub to the e-amino nitrogen on a lysine residue of other proteins. The reaction is catalyzed by the Ub conjugating enzymes (E2s). The E2s accept Ub in an activated form from Ub activating enzyme (E1) by a transthiolation reaction that transfers Ub from the E1 active site Cys to the E2 active site Cys, which, subsequently transfers Ub from the E2 to the target lysine[6]. Curiously, some E2s are capable of linking Ub molecules together as chains that are tethered to the targeted protein. Furthermore, chain configurations can vary depending on which Ub lysine is used as the linkage site. For example, Ub molecules that are joined in tandem through Lys48 selectively target proteins for degradation[7], while Ub molecules that are tandemly joined through Lys63 function in a non-degradative manner within NF-κB[8 9]signal transduction and error-free DNA repair[10].

Although the structures of Ub[11;12], and a variety of E2s[13;14] have been determined to high resolution, there is little structural understanding of the catalytic mechanism of Ub isopeptide bond formation. On the other hand, acetylation of lysine is a well-characterized analogous enzymatic reaction that is likely to hold much in common with ubiquitination[15]. Like the E2s, the acetyl transferases link acetate to the

ε-amino group of lysine through the formation of an isopeptide bond. While the acetyl transferases can vary from example to example in mechanistic detail, all are thought to involve a tetrahedral intermediate that is created when the nucleophilic amine nitrogen of lysine attacks the electrophilic α-carbon of an activated acetyl group. The formation of the intermediate is typically facilitated by two other catalytic participants: 1) a general base that abstracts a proton from the ε-amino group making the nitrogen more nucleophilic and 2) a hydrogen bond that stabilizes the oxyanion of the intermediate.

The Lys63-linked polyUb chain is a well understood example of isopeptide bond formation. The reaction is catalyzed by a heterodimer consisting of the E2 Ubc13 and the Ub conjugating enzyme variant (UEV) Mms2[16]. Curiously, the UEVs lack the active site cysteine found in the E2s[17]. The heterodimer is hypothesized to correctly positions two Ub molecules for covalent linkage[18;19]. One Ub molecule is covalently linked to the Ubc13 active site as a thiolester and the other Ub molecule is non-covalently positioned on Mms2 such that its Lys63 is juxtaposed to the activated carboxy-terminus of the thiolester Ub. The structure of the heterodimer has been determined crystallographically[19;20] and the placement of the two Ub molecules on the heterodimer surface has been determined to lower resolution by NMR spectroscopy[21] (**Figure 4.1**).

In the present study, we provide both structural and functional evidence that protein ubiquitination conforms to the canonical mechanism of isopeptide bond catalysis.

4.3 Experimental Procedures:

4.3.1 Ubc13 point mutations. Codon substitutions were introduced into the human Ubc13 open reading frame by PCR based site-directed mutagenesis using the Ubc13-pGex6 vector as a template (Ubc13-N79A, Ubc13-D89A, Ubc13-N116A and Ubc13-D119A). All mutations were verified by sequencing each open reading frame in both directions.

4.3.2 *In vivo* Assays of Ubc13 activity. A wildtype haploid *S. cerevisiae* strain HK580-10D (*MAT* α. *ade-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*) was received from Dr. H. Klein (New York University) and used as the recipient to delete the entire *UBC13* open reading frame by a one-step gene replacement method[22] using a *ubc13Δ::HIS3* cassette generated through PCR amplification as previously described[23]. The resulting strain WX905 was transformed with the two-hybrid vector pGAD424 carrying *hUBC13*.

The gradient plate assay was performed as a semi-quantitative measurement of relative methylmethane sulfonate (MMS) sensitivity[24]. 30 ml of molten YPD agar containing 0.025% MMS was poured into tilted square Petri dishes to create agar slants. After solidification for one hour, the petri dishes were returned flat and 30 ml of the same molten agar without MMS was poured to form the top layer. A 0.1 ml sample was taken from an overnight culture, mixed with 0.4 ml sterile water and 0.5 ml of molten YPD agar, and then immediately imprinted onto freshly made gradient plates via a microscope slide. Gradient plates were incubated at 30°C for 48 hours.

The *hMMS2*-coding region was PCR-amplified and cloned into pGBT9 (purchased from Clontech, Palo Alto, CA) as an amino-terminal fusion to $Gal4_{BD}$. The *hUBC13*-coding region was cloned into pGAD424 as an amino-terminal fusion to $Gal4_{AD}$. Yeast strain PJ69-4A (MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met::GAL7-lacZ- P. James, University of Wisconsin) containing both pGBT9*[hMMS2]* and pGAD424[20]-based constructs were plated onto SD-Trp-Leu-His plates supplemented with various concentrations of 3-aminotriazole (3-AT) to measure activation of the GAL1-HIS3 reporter gene as previously described[24].

4.3.3 ***In vitro*** **Assays of Ubc13 activity.** Ub chain assembly reactions were performed by incubating 250 nM E2 (hUbc13, hMms2, and /or hUbc13 derviatives) with 2.5 μM ^{35}S-Ub and 25 nM Uba1-6xHis at 30°C for 1 hr in ATP cocktail: 10 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 5 mM ATP, 0.6 units/mL inorganic phosphatase[25]. The reactions were terminated by the addition of 10 mM DTT, immediately followed by precipitation with 10% trichloroacetic acid. Samples were then resuspended and boiled for 5 min. in SDS loading buffer, followed by separation by SDS-PAGE and analysis by autoradiography as described previously [26].

Ubc13~Ub thiolester assays were performed by incubating 2 μM Ub, 1 μM E2 and 100 nM E1 at 30°C for 5 min to 60 min in 1.0 ml of ATP cocktail. Reactions were terminated by chelating Mg^{2+} with 50 mM EDTA (final) to inactivate E1[25;27]. Reactions (0.5 ml) were loaded onto a High Load Superdex 75 16/60 gel filtration

column (Pharmacia) equilibrated with 50 mM HEPES pH 7.5, 75 mM NaCl, 1 mM EDTA. The ^{35}S-Ub content of each eluted fraction (0.5 mL) was determined by scintillation counting and was found to distribute into three well-resolved peaks corresponding to ^{35}S-Ub and the hUbc13~^{35}S-Ub and Uba1~^{35}S-Ub thiolesters. Thiolester yields for all reactions were calculated from the specific radioactivity of the incorporated ^{35}S-Ub and were expressed as percentages of the total Ub in the reaction. Identical reactions were performed in the presence of the reducing agent DTT to normalize each reaction and estimate background ^{35}S-Ub CPM levels.

4.3.4 Protein purification and crystallization. Over-expression and purification of human Ubc13, Mms2 and yeast Uba1-6xHis have been described previously [18]. Mms2•Ubc13 heterodimers were purified by gel filtration chromatography as previously described[20]. All crystals were grown at 20-24 °C using the hanging drop vapor diffusion technique. Crystals of hMms2•hUbc13 and its derivatives were grown by mixing 1 μL of heterodimer at 8 mg•mL^{-1} in 50 mM Hepes pH 7.5, 150 mM NaCl, 1mM EDTA, and 1 mM DTT with 1 μL of well solution (20% PEG 6000, and 100 mM citrate pH 6.5), which was then equilibrated against the well solution.

4.3.5 Structure determination. Diffraction data for native hMms2•hUbc13-D89A were collected from a single crystal at 105 K using a RaxisIV image plate detector and with Rigaku RU-H3R rotating anode generator. Diffraction data for the derivatives of hMms2•hUbc13-N79A and hMms2•hUbc13-D119A were collected

from single crystals at Advanced Light Source (ALS, Berkeley, California), beamline 8.3.1 using a 2 x 2 CCD array (ADSC) detector. Data were scaled and integrated with Mosflm and SCALA[28]. Structures of the hMms2-hUbc13 derivatives were refined through several cycles. Manual rebuilding was performed with reference to Sigma A-weighted[29] 2|Fo| - |Fc| electron density maps calculated in Refmac5. All structures were refined with Refmac5[30]. The quality of the each of the models was assessed by PROCHECK[31].

4.4 Results

4.4.1 Candidate residues within the active site of hUbc13 are identified and substituted.

Four charged residues in hUbc13 were investigated due to their close proximity to the active site and their potential roles as general bases or hydrogen bond donors (**Figure 4.1**). N79, D89, N116, and D119 in hUbc13 were individually mutated to alanines creating hUbc13-N79A (N79A), hUbc13-D89A (D89A), hUbc13-N116A (N116A), and hUbc13-D119A (D119A) constructs.

4.4.2 Active site Asp residues and Asn79 in Ubc13 are required for function *In vivo.*

hUbc13 has previously been shown to be able to complement the function of yeast Ubc13 in error-free DNA repair, for which Lys63 ubiquitination activity is absolutely required [24]. Each of the four active site derivatives and hUbc13 were

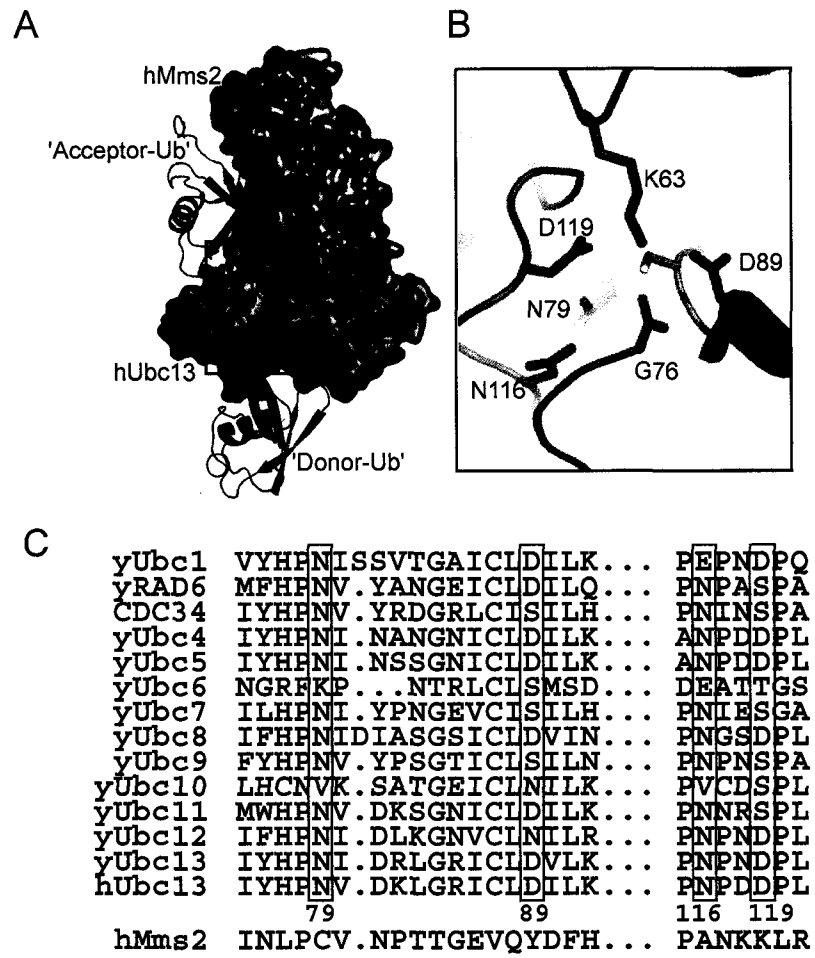


Figure 4.1 E2 Active-Site Structure. A. Surface representation of the heterodimer (hUbc13- teal; hMms2- olive). Ub molecules are depicted as red ribbons. The active-site C87 of hUbc13 is shown in yellow. K63 of the acceptor-Ub and the carboxy-terminal carboxyl group of the donor-Ub are depicted as sticks. **B.** Magnification of the active-site region showing the potentially important catalytic residues N79, D89, N116, D119. Secondary structural elements of hUbc13 are depicted as a teal ribbon. Also shown: K63 of the acceptor-Ub, the G76 carboxyl group of the donor-Ub, active-site C87. **C.** Sequence conservation of the active site.

expressed in a yeast *ubc13* knockout and tested for their ability to compliment the DNA repair activity. The active site derivatives N79A, D89A, and D119A were all sensitive to MMS whereas hUbc13 and N116A demonstrated the ability to compliment the null Ubc13 strain **(Figure 4.2)** suggesting that N116A does not play a critical role in the function of Ubc13. However each of the three other residues are necessary for biological function.

In order to build Lys63 linked ubiquitin chains, Ubc13 must be able to bind Mms2 as mutations in the interface of the hUbc13•hMms2 that disrupt their interaction also completely abolish Lys63 linked Ub chain building properties [19]. A yeast-2-hybrid assay was used to ensure that the failure of the active site mutants to complement yUbc13 was not due to an inability to bind Mms2. The yeast 2-hybrid assay demonstrated that all of the active site derivatives bind to hMms2 *in vivo*.

4.4.3 The formation of an isopeptide bond is dependent on the presence of the charged residues surrounding the active site cysteine.

The active site derivatives that failed to compliment the yUbc13 knockout were probed using *in vitro* ubiquitination assays to first test for the ability to catalyze the synthesis of an isopeptide bond. This assay has been previously employed to probe the activities of various yeast E2s [25;32] as well as hUbc13[18]. With respect to hUbc13 it was demonstrated that in the presence of purified ubiquitination reactions components (E1, Mg^{2+}, ATP, ^{35}S-Ub, and hMms2), hUbc13 supports the synthesis of a) hUbc13-Ub conjugate (hUbc13-Ub) and b) di-ubiquitin (Ub_2).

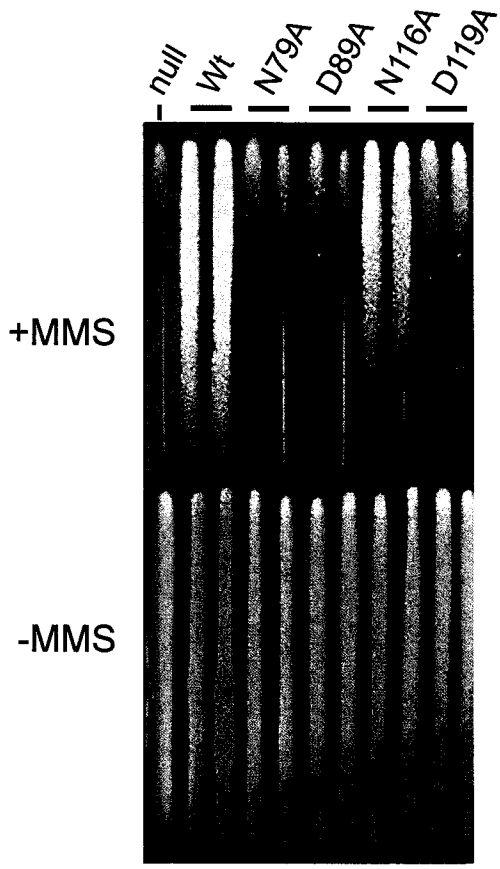


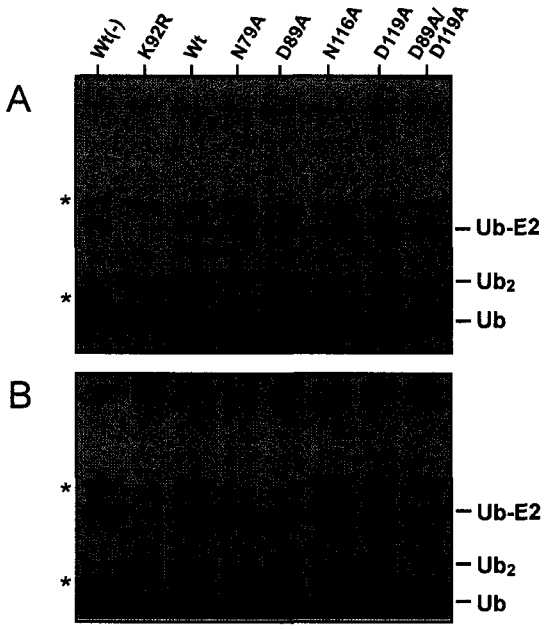
Figure 4.2 Sensitivity of Active-Site Mutants to MMS. Yeast expressing hUbc13 (Wt) or its active site mutants were streaked in duplicate onto gradient plates with and without MMS (see Methods). A plasmid lacking the Ubc13 open reading frame was used as a negative control(null).

We tested the ability of each derivative to support the synthesis of these products and compared them to hUbc13 **(Figure 4.3)**. The hUbc13-K92R (K92R) is a lysine to arginine derivative that renders hUbc13 unable to conjugate Ub at K92, but allows for thiolester formation, thereby not affecting Ub_2 formation[18] and is used as a negative control to compare hUbc13-Ub conjugate formation. The formation of Ub_2 or any poly-Ub chain assembly in hUbc13 is dependent on the presence of hMms2[18].

Relative to hUbc13, all derivatives supported the transfer of thiolester linked Ub to K92 of hUbc13 **(Figure 4.3A)**. Using the same experiment, we can also visualize the formation of an isopeptide bond by monitoring the production of Ub_2. N79A, D89A and D119A were all deficient in the ability to form Ub_2 when compared to Ubc13 **(Figure 4.3A)**. Overnight reactions illustrated that the D89A, D119A, and N79A were still unable to produce Ub_2 suggesting that the reaction was severely compromised (**Figure 4.3B**).

4.4.4 Thiol-ester formation is independent of the charged residues surrounding the Active Site.

The ubiquitination assays revealed that all three charged residues close in proximity to the active site affect Ub_2 synthesis. An *in vitro* thiolester assay monitored the formation of E2~Ub to determine whether defects in Ub_2 synthesis resulted from a defect in the transfer of Ub from E1 to E2 or rather in isopeptide bond catalysis. Both in the case of D89A and D119A the synthesis of E2~Ub thiolester was comparable to hUbc13 whereas N79A was less efficient but still functioned to



C

	%Ub_2 60min	%Ub_2 16hr
hUbc13	1.7	9.0
N79A	0.1	0.2
D89A	0.1	0.3
N116A	0.3	0.7
D119A	0.1	0.2
D89A/D119A	0.0	0.0

Figure 4.3 Ubiquitination by hUbc13. SDS-PAGE gels of reactions incubated for one hour (A) or 16 hr (B). The reaction products containing ^{35}S-Ub were visualized by autoradiography and include di-ubiquitin (Ub_2) and Ub conjugated to hUbc13 (E2-Ub). Contaminants present in the ^{35}S-Ub preparation are indicated (*). The quantity of ^{35}S-Ub_2 for each reaction was determined as a percentage of the total ^{35}S-Ub in each lane (C). A reaction with hUbc13(Wt) but lacking hMms2 was used as a control Wt(-).

accept a thiolester linked Ub (**Figure 4.4B**). N79A was unable to produce any Ub_2, despite the fact that wildtype levels of thiolester can be achieved after 1 hr. Therefore the inability to compliment yeast Ubc13 or build an isopeptide bond does not stem from an inability to accept Ub from the E1. This suggests that either these residues play some catalytic role in the synthesis of the isopeptide bond or they prevent its formation by hindering the ability of the acceptor Ub from being positioned near the active site.

4.4.5 The architecture of the hUbc13 active site is maintained in the D89A and N79A derivatives.

The hUbc13 active site derivatives were each crystallized in the presence of hMms2 as heterodimers. The active site derivatives were structurally aligned with the canonical hUbc13·hMms2 crystal structure and Ub bound tetrameric NMR model (**Table1**). All Ubc13 derivatives were purified in complex with hMms2 for crystallization confirming tight association with Mms2 *in vitro*.

N79A crystallized in spacegroup C2 and was solved by molecular replacement using the hUbc13·hMms2 as a search model (PDB accession code: 1J7D). The interface between hMms2 and N79A is intact and the overall structure is nearly identical (RMSD = 0.7 Å). The N79A mutation causes no major shifts in the location of any of the other charged residues with respect to the active site cysteine (**Figure 4.5A**).

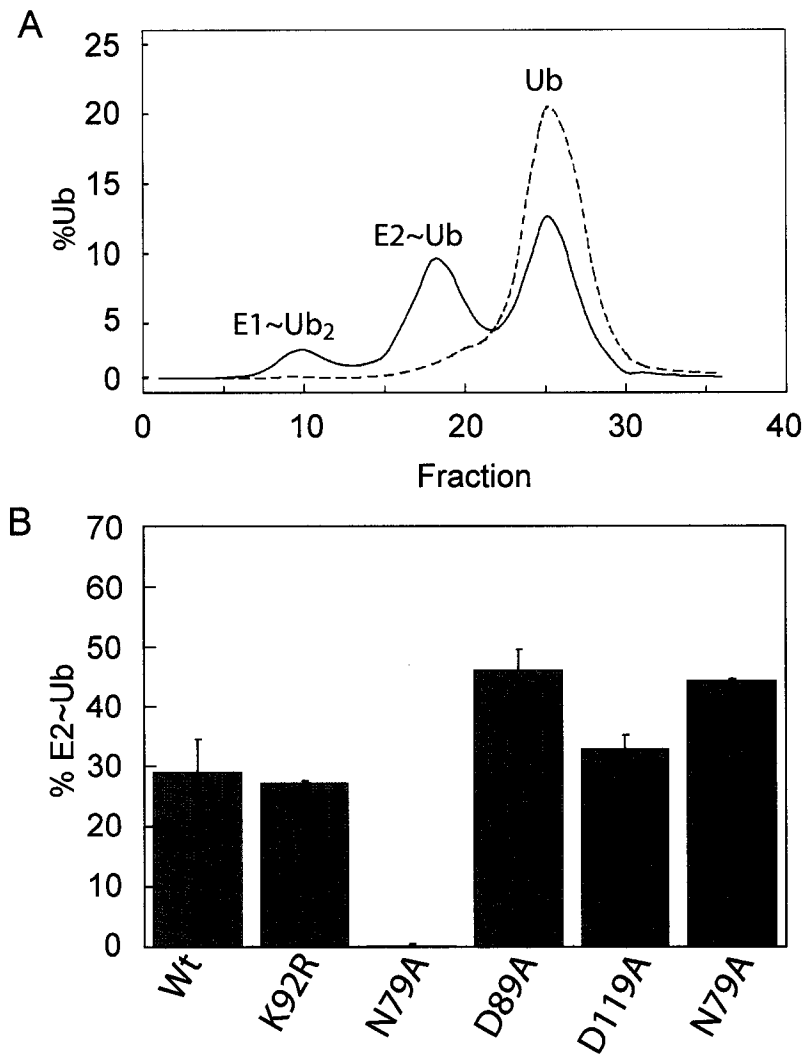


Figure 4.4 E2~Ub Thiolester Formation. A) A typical elution profile for a reaction containing wildtype hUbc13. Peaks containing ^{35}S-Ub correspond to free Ub, E2~Ub thiolester and E1~Ub thiolester. B) The incorporation of ^{35}S-Ub into thiolester was determined for each of the hUbc13 derivatives and is expressed as a percentage of the total Ub in the reaction.

This structural evidence supports previous reports from CD and NMR[33] that this Asn residue is not required for stabilizing the E2 structure but rather puts forward its role in stabilizing the oxyanion that develops during the formation of the tetrahedral intermediate.

The hUbc13 D89A mutant does not show perturbation of the overall architecture compared to the canonical active site of hUbc13 (RMS = 0.25 Å). Within this structure, D119 rotates slightly about χ_1 but is still positioned close to the modeled acceptor-Ub Lys63 residue when compared to hUbc13 (**Figure 4.5B**). Given the structural similarity between Ubc13 and the D89A derivative these observations point toward a catalytic role for this residue in isopeptide bond synthesis.

The hUbc13 D119A mutant shows a pronounced conformational change compared to wildtype. The loop proximal to the active site that contains residues N116 and D119 shifts position significantly. This loop is near to the 'acceptor-Ub' and the conformational shift involves disruption of hydrophobic contacts between L121 and the acceptor Ub in the model for Ub_2•Mms2•Ubc13. The γ–carbon of N116 translates 7 Å (moving from 7.6 to 15.2 Å away from the active site Cys87) (**Figure 5C**). While the position of N79 is not affected by the D119A mutation, D89 rotates 180° about χ_1, repositioning the negatively charged side chain ~9 Å from the ε-amino nitrogen of the acceptor-Ub Lys63 residue.

Table 4.1: Crystallographic Statistics for Ubc13 Active Site Derivatives.

	N79A	D89A	D119A
Data collection	ALS 8.3.1	RaxisIV	ALS 8.3.1
Wavelength (Å)	0.998	1.54	0.998
Completeness (%)	70	98.7	97.01
Unique reflections	10013	17091	31357
Mosaicity (°)	0.67	0.66	0.56
<I>/ σ(I)	15.6(5.6)	16.4(2.1)	29.4(3.4)
R_{sym}[a]	0.037(0.08)	0.055(0.46)	0.045(0.26)
Cell (space group)	C121	p212121	p212121
	86.4 42.9 93.7	44.7 73.8 91.5	44.2 74.1 92.2
	90 107 90	90 90 90	90 90 90

Refinement			
Resolution (Å)	13.42-2.3	27-2.1	46-1.7
R_{fac}	26.5	23.0	21.3
R_{free}	32.4	26.2	24.7
R.M.S Deviations			
Bond length (Å)	.032	0.006	0.01
Bond Angle (°)	2.544	0.99	1.24
# residues	288	288	288
# water molecules	5	190	150
Cα-R.M.S.Deviations hUbc13-hMms2 (Å)	0.780	0.25	0.32

[a] $R_{sym} = \Sigma_{i,h} |I(i,h) - \langle I(h) \rangle| / \Sigma_{i,h} |I(i,h)|$ where $I(i,h)$ and $\langle I(h) \rangle$ are the ith and mean intensity of reflection h.

[b] R_{free} is calculated using 5% of the reflections.

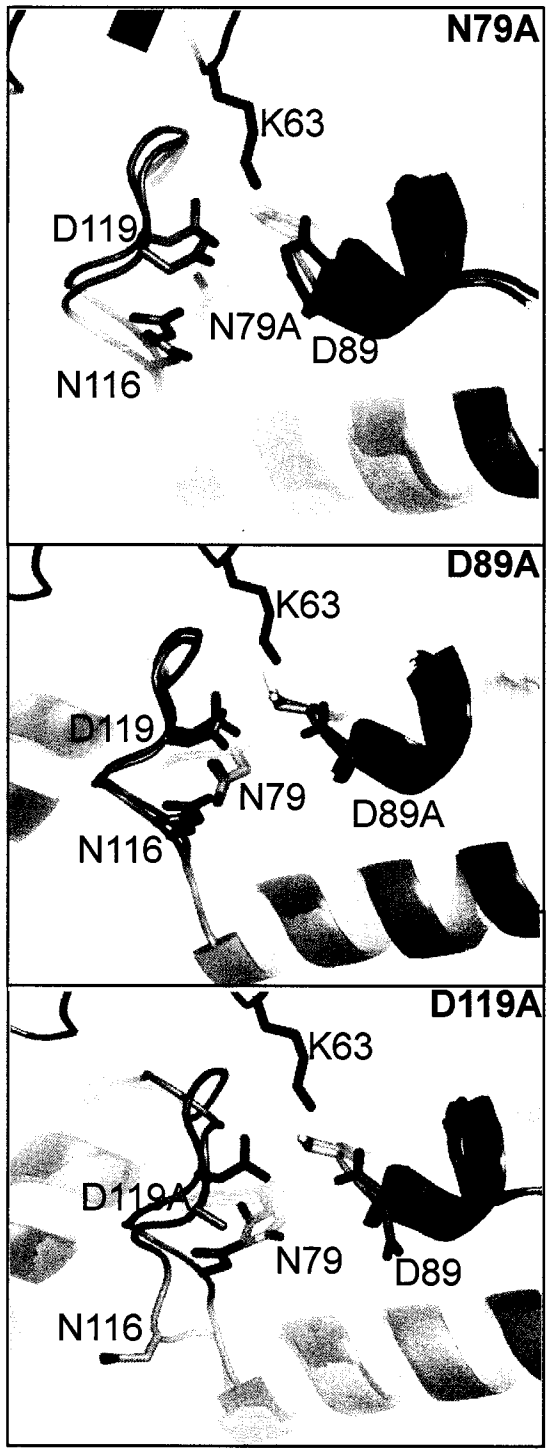


Figure 4.5 The Effect of Amino Acid Replacement on the Active-Site Structure. The ribbon structure the wild-type heterodimer (teal) is superimposed on each of the active-site replacements (grey) :N79A, D89A, and D119A. The C87 active-site sulfur is shown in yellow. The position of Ub Lys63 (K63) was determined relative to the wild-type heterodimer as previously described[19;20].

The hUbc13 D119A derivative alters the conformation of the active site in the region involved in binding the acceptor-Ub and is the most likely cause of its inability to synthesize Lys63 linked Ub chains which further results in its functional defect in vivo. On the other hand, the N79A and D89A derivatives display relatively no change in the architecture of the active site that might explain their deficiencies in the synthesis of an Ub-Ub isopeptide bond. This suggests that each of these two residues play a critical role in the catalysis of an isopeptide bond. From our observations, we propose that D89 in hUbc13 acts as the general base to ensure that the amino group of Lys63 is nucleophilic (**Figure 4.6**). Furthermore, our data supports the previous theory that N79 in hUbc13 (and homologous residues in other E2s) functions to stabilize the oxyanion that develops during ubiquitination. These two residues are catalytically necessary for isopeptide bond formation.

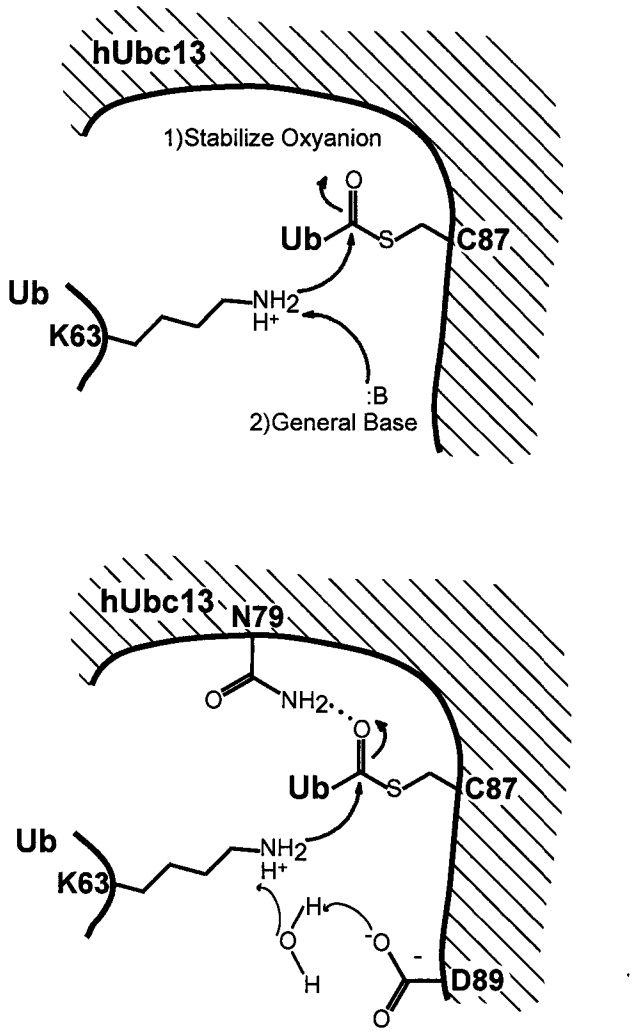


Figure 4.6 Catalysis of an Isopeptide bond. Shown is N79 as the oxyanion stabilizer and D89 as the general base in isopeptide bond formation.

4.5 Discussion:

The mechanism of isopeptide bond catalysis for the ubiquitination process can be inferred from the mechanism of lysine acetylation, as both mechanisms can employ thiolester intermediates and both involve the modification of surface Lys residues[15]. Based on the structure of *Tetrahymena* histone acetyl transferase (HAT) GCN5 in complex with a histone H3 peptide and coenzyme A[34], the acetylation reaction procedes via a nucleophilic attack by the lysine ε-amino group onto the electrophilic carbon center of the thiolester bond of acetyl-CoA. A general base is required to abstract a proton from the attacking ε-amino group and in this case a Glu activates a water molecule that acts as the general base. Furthermore the transition state tetrahedral reaction intermediate (oxyanion) is stabilized by a main chain amide group.

Several E2s possess the ability to synthesize an isopeptide poly-Ub linkage, and in an analogous fashion to the HAT proteins, these E2s must possess the necessary catalytic components to carry out this reaction. The Ubc13•Mms2 complex is one E2•UEV complex that is able to synthesize Lys63 linked poly-Ub in the absence of an E3[10;16;18]. Other advantages in using this system to probe the catalytic mechanism of isopeptide bond formation include the existence of a large body of structural data[18-21]. X-ray crystallographic and NMR data have allowed for the development of a model of the E2•UEV complex bound to a substrate acceptor-Ub and a donor- thiolester linked Ub.

Using this model as a guide, we have identified, N79 and D89, residues within hUbc13 that inhibit isopeptide bond formation. To further establish that these residues solely affect isopeptide bond catalysis they were screened for possible effects on catalyzing E2~Ub thiolester synthesis, Mms2 binding, and overall structural integrity of the E2. Furthermore, a role in catalysis of the isopeptide bind can be attributed to each of these residues in the mechanism of Lys63 linked poly-ubiquitination.

The side chain of Asn79 is hypothesized to stabilize the oxyanion of the transition state tetrahedral intermediate (**Figure 4.6**). This hypothesis has already been proposed for several E2s including Ubc9, E2K, Ubc5, and yUbc13[33]. The Ubc13 N79A derivative does show decreased thiolester production, however, at a timepoint equivalent to the amount of time used to assay isopeptide bond formation, the levels of N79A~Ub thiolester is comparable to hUbc13 (wildtype). Furthermore, the crystal structure of N79A illustrates that the role of the Asn is not necessary for E2 folding or coordinating the architecture of the active site supporting the notion that its function is catalytic.

We have identified D89 in Ubc13 as necessary in the mechanism of isopeptide bond formation, that functions as a general base. The non-conservative Ubc13 D89A derivative does not affect thiolester formation, or the structure of hUbc13·hMms2, yet fails to synthesize Ub_2. In our model, the function of D89 is to orient the ε-amino group of Lys63 on the acceptor Ub, and perhaps through a water molecule act as a general base to removed a proton from the Lys making it a nucleophile, preparing it for attack on the carbonyl carbon of the thiolester linked donor Ub. This Asp residue

is conserved in most E2s except for a subset of E2s that have extended loops within the core domains (cdc34, yUbc7), and those that conjugate Ub-like proteins (Ubc12, Nedd8) and (Ubc9, Sumo). Within the Ub-like subset of E2s, a comprehensive analysis of residues within Ubc9 has ruled out any specific residue as the general base[35]. There are multiple residues in the active site of Ubc9 that may function as a general base, such that their roles are redundant, and may function as the primary general base depending on the target that is brought into the active site (ie Ubc9-D127 for p53 and IkBα or Ubc9Y134 for RanGap1)[35]. A serine residue replaces the equivalent D89 in Ubc9 within a shallow active site accommodating its target (ie. RanGAP1 and Sumo). This serine residue is present in the other subset of E2s that contain an extended loop within the core domain (Cdc34, Ubc7). The role of this serine is still under investigation, however, recent work has shown it has a function in E2 ubiquitination activity in the context of other residues within the active site of Cdc34[25] and may be a site of phosphorylation.

The assembly of Lys63 linked Ub chains is a specialized function of the Ub system for which Ubc13 and its scaffolding partner Mms2 precisely orient the substrate acceptor-Ub toward the active-site Cys for isopeptide bond transfer. However, the superposition of structures of Ubc13 with other E2 crystal structures (yUbc1, UbcH7 and, yUbc10), demonstrates that the position of the Asp within the active site is conserved in known E2 structures, though the conformation of the side chain is not. The flexibility of this residue is confirmed by a notable chemical shift in

the Asp in the presence of thiolester linked Ub, suggesting by NMR (on Ubc1 and Ubc13[36]) that there are substrate induced conformational changes[21].

Different conformations of E2s have been observed in crystallographically determined structures. For the crystal structure of yUbc1[36], there exist two molecules in the asymmetric unit related by non-crystallographic symmetry. Differences between the two include a conformational change in the loop proximal to the active site; two residues that shift within their local regions significantly between the two monomers are the equivalent D119 residue and the D89 residue in hUbc13. This lends support to the idea that the equivalent D89 side chain in a subset of E2s is flexible and can be oriented close to the active site Cys in order to act as a general base.

It is possible that a Ubc13 D89A in comparison to wildtype repositions the thiolester linked Ub, leading to defects in isopeptide bond synthesis. The NMR analysis of the hUbc13~Ub thiolester implies the covalent linkage is flexible, even though the thiolester, as shown in the model of thiolester Ub in yUbc1~Ub and hUbc13~Ub, is planar with respect to the attacking Lys63 group[21]. Two pieces of evidence argue against the possibility of reorientation as an explanation for a deficiency in isopeptide bond formation. Over a long time period, the thiolester linked Ub would periodically be in the correct position for isopeptide bond formation as it is a flexible linkage. Therefore the *in vivo* complementation assays that show absolutely no growth and the *in vitro* isopeptide bond assays demonstrating virtually no Ub_2 production suggest that time has no effect on the ability to build the Lys63

linked poly-Ub chains. Thus we conclude that Ubc13 D89 functions catalytically in the synthesis of the isopeptide bond.

Although the derivative D119A also abolishes Ub_2 production, this is likely due to repositioning of the acceptor Ub. The loop in hUbc13 that houses D119 shifts significantly when an alanine is substituted. Furthermore, D119 remains static in the structure of the D89A derivative with respect to wildtype, suggesting that the residue is in its canonical orientation and still cannot function as the general base. From this observation we conclude that D119 does not function directly in catalysis.

From the data presented here as well as in previous studies, we can corroborate that N79 in the HPN motif of E2s may stabilize the developing negative charge on the carbonyl oxygen. Further to this we propose that D89 residue in hUbc13 functions as a general base, perhaps in ordering a water molecule to deprotonate the incoming Lys63 ε-amino group, creating a nucleophile. The experiments were conducted using a specialized E2 that builds Lys63 Ub chains. However, the conservation of both Asn79 and Asp89 residues in almost all E2s points to a general catalytic mechanism for E2 enzymes.

4.6 Reference List

1. Hershko, A. and Ciechanover, A. (1998) *Annu Rev Biochem* **67,** 425-79
2. Pickart, C. M. (2001) *Annu Rev Biochem* **70,** 503-33
3. Glickman, M. H. and Ciechanover, A. (2002) *Physiol Rev* **82,** 373-428
4. Hochstrasser, M. (1996) *Annu. Rev. Genet* **30,** 405-439
5. Pickart, C. M. (2001) *Mol Cell* **8,** 499-504
6. VanDemark, A. P. and Hill, C. P. (2002) *Curr Opin Struct Biol* **12,** 822-30
7. Bachmair, A. and Varshavsky, A. (1989) *Cell* **56,** 1019-32
8. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z. J. (2000) *Cell* **103,** 351-61
9. Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., Xiao, W., and Dixit, V. M. (2004) *Nature* **427,** 167-171
10. Hofmann, R. M. and Pickart, C. M. (1999) *Cell* **96,** 645-53
11. Phillips, C. L., Thrower, J., Pickart, C. M., and Hill, C. P. (2001) *Acta Crystallogr D Biol Crystallogr* **57,** 341-4
12. Cook, W. J., Jeffrey, L. C., Kasperek, E., and Pickart, C. M. (1994) *J Mol Biol* **236,** 601-9
13. Cook, W. J., Jeffrey, L. C., Sullivan, M. L., and Vierstra, R. D. (1992) *J Biol Chem* **267,** 15116-21
14. Cook, W. J., Martin, P. D., Edwards, B. F., Yamazaki, R. K., and Chau, V. (1997) *Biochemistry* **36,** 1621-7
15. Passmore, L. A. and Barford, D. (2004) *Biochem J* **379,** 513-25
16. Hofmann, R. M. and Pickart, C. M. (2001) *J Biol Chem* **276,** 27936-43
17. Broomfield, S., Chow, B. L., and Xiao, W. (1998) *Proc Natl Acad Sci U S A* **95,** 5678-83
18. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6

19. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001) *Cell* **105,** 711-20

20. Moraes, T. F., Edwards, R. A., McKenna, S., Pastushok, L., Xiao, W., Glover, J. N., and Ellison, M. J. (2001) *Nat Struct Biol* **8,** 669-73

21. McKenna, S., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., Spyracopoulos, L., and Ellison, M. J. (2003) *J Biol Chem* **278,** 13151-8

22. Rothstein, R. (1991) *Methods Enzymol* **194,** 281-301

23. Brusky, J., Zhu, Y., and Xiao, W. (2000) *Curr Genet* **37,** 168-74

24. Ashley, C., Pastushok, L., McKenna, S., Ellison, M. J., and Xiao, W. (2002) *Gene* **285,** 183-91

25. Varelas, X., Ptak, C., and Ellison, M. J. (2003) *Mol Cell Biol* **23,** 5388-400

26. Hodgins, R., Gwozd, C., Arnason, T., Cummings, M., and Ellison, M. J. (1996) *J Biol Chem* **271,** 28766-71

27. Haas, A. L., Murphy, K. E., and Bright, P. M. (1985) *J Biol Chem* **260,** 4694-703

28. Collaborative Computational Project, Number 4. 1994. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* D50, 760-763. 94.

29. Read, R. J. *Acta Crystallogr. A* 42, 140-149. 86.

30. Murshudov, G. N., Vagin, A., and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst* D53, 240-255. 97.

31. Laskowski, R. A., MAcArthur, M. W., Moss, D. S., and Thornton, J. M. PROCHECK. *J. App. Cryst.* 26 , 283. 93.

32. Ptak, C., Gwozd, C., Huzil, J. T., Gwozd, T. J., Garen, G., and Ellison, M. J. (2001) *Mol Cell Biol* **21,** 6537-48

33. Wu, P. Y., Hanlon, M., Eddins, M., Tsui, C., Rogers, R. S., Jensen, J. P., Matunis, M. J., Weissman, A. M., Wolberger, C. P., and Pickart, C. M. (2003) *EMBO J* **22,** 5241-50

34. Rojas, J. R., Trievel, R. C., Zhou, J., Mo, Y., Li, X., Berger, S. L., Allis, C. D., and Marmorstein, R. (1999) *Nature* **401,** 93-8

35. Bernier-Villamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002) *Cell* **108,** 345-56

36. Hamilton, K. S., Ellison, M. J., Barber, K. R., Williams, R. S., Huzil, J. T., McKenna, S., Ptak, C., Glover, M., and Shaw, G. S. (2001) *Structure (Camb)* **9,** 897-904

Chapter 5

5.1 General Discussion

The focus of this dissertation involves the structure and function of the E2-UEV heterodimeric complex with respect to the synthesis of Lys63 linked poly-Ub chains. The E2 Ubc13 has been identified in two distinct pathways that employ Lys63 linked poly-Ub chains; NF-κB signal transduction and post-replicative DNA repair. Ubc13 demonstrates the ability to interact with UEVs (Mms2 or Uev1a) that are necessary components of these pathways particularly with regard to the synthesis of Lys63 linked poly-Ub. To determine the mechanism of alternative Ub chain linkage synthesis by this particular E2, the following structural and biochemical studies were undertaken. First the structure of the Ubc13-UEV heterodimer complex was elucidated by X-ray crystallography. Second, using the structure as a template, NMR was utilized to model the associated Ubs onto this heterodimer. Third, biochemical assays and X-ray crystallographic studies were employed to identify catalytic residues necessary for poly-Ub chain synthesis.

Chapter 2 describes the crystallographically determined structure of the complex between Mms2 and Ubc13, and the structure of free hMms2[1]. These structures reveal that the Mms2 monomer has a similar fold to E2s but lacks a catalytic Cys and both carboxy- terminal helices hinting that any interaction with Ub could be novel in comparison to a canonical E2 interaction with Ub. The nature of the interface between Ubc13 and Mms2 provides a physical basis for the preference

of Mms2 for Ubc13 as a partner over a variety of other structurally similar E2s. The structure of the Mms2•Ubc13 complex provides the conceptual foundation for understanding the catalytic mechanism of poly-ubiquitin chain assembly. Subsequently this structure has been used as a basis for key NMR chemical shift mapping experiments that provided the binding sites for a non-covalently bound Ub associated with Mms2 and a thiolester Ub linked to Ubc13[2;3] and allowed the development of the (Mms2•Ub)•Ubc13~Ub tetrameric model that exists prior to poly-Ub chain synthesis.

Chapter 3 describes a structural and functional comparison between Mms2 and Uev1a to determine any distinguishing properties that may distinguish the two with respect to their biological roles[4]. Mms2 and Uev1a were found to adopt a similar core domain, however, an amino-terminal extension for Uev1a provides additional poly-Ub chain building sites compared to Mms2[5]. This function is attributed to the different binding affinity Uev1a has for Ubc13 and Ub[6]. From a structural perspective, the catalytic mechanism of isopeptide bond synthesis likely remains unchanged between the two heterodimer complexes.

Chapter 4 discusses the chemical mechanism of isopeptide bond formation linking two Ubs within the context of the Ubc13•UEV structure. Two residues surrounding the active site Cys of Ubc13 were identified as the general base and oxyanion stabilizer within the mechanism of Ub_2 synthesis[7]. The proposed mechanism is similar to the mechanism of isopeptide bond formation utilized in acetylation reactions.

The substance of this thesis has focused on the catalytic function of the Ubc13•UEV complex. The initial structural determination of the Mms2•Ubc13 heterodimer combined with the NMR based mapping of non-covalently bound and thiolester linked Ubs provides a basis for studying the catalytic mechanism of poly-Ub chain assembly. Currently, there are no structural data regarding the role an E3 would undertake in this particular mechanism. Many E3s function to stimulate poly-ubiqutination by arranging a target proximal to the E2. In this case, the UEV participates to orient Lys63 close to the active site of the E2 thereby stimulating Lys63 linked Ub chains formation. The E3 may be responsible for localizing this complex to specific pathways (and or targets) in which these alternative chains are required, for example; RAD5 in DNA repair, and Traf6 in NF-κB signal transduction. Both proteins contain RING domains and have been shown to interact with Ubc13. In this bottom-up approach to piecing together the components of the complex involving Lys63 linked Ub chain assembly, the next logical step would be to embark on a structural determination involving the heterodimer and E3. Several attempts have been made to isolate full length Traf6 for this purpose but with limited success with respect to solubility although expression levels have been good. Therefore, a more feasible option would be to isolate the RING component of RAD5 or Traf6 and approach the structural problem through crystallography or NMR chemical shift mapping as has recently been done for the BRCA1-BARD1 RING complex[8] and the CNOT-Ubc5 (RING/E2) complex.

The interaction between an E3 and the UEV-E2 complex may help define the individual roles of Uev1a and Mms2 in the cell[4]. On the basis of the E3/E2 structures determined for UbcH7 and Ubc13, the amino-terminal extension for Uev1a is located near the site of E3/E2 interface. In addition, the amino-terminal extension on Uev1a has been shown to affect Ub chain assembly with respect to Mms2 and Uev1aΔ30. To date, no structural data exist regarding this region of Uev1a except that to say that the region is flexible (preliminary HSQC-NMR)♠, and likely hinders crystallization. Ongoing NMR experiments using chemical shift data and NOE and perhaps a solution structure may help to further define this region, and delineate interactions with Ubc13 and RING E3s.

NMR is also currently being used to further substantiate the role of Asp89 as a general base in Ubc13's catalysis of Lys63 linked chains. The pKa of the Lys63 of the acceptor-Ub is being determined in wildtype Ubc13•Mms2 and the D89A•Mms2 derivative. This experiment should further corroborate the role of Asp89 in orienting and preparing the ε-amino group of Lys63 for attack on the thiolester linked Ub. Difficulties with this experimental procedure include overlap of the Lys63 13Cε, 1Hε chemical shifts in the NMR spectra collected for wildtype and mutant Ubc13s. Furthermore, the position of these peaks and their properties do not accurately represent them at the immediate moment before isopeptide bond synthesis as the thiolester linked Ub would undoubtedly affect the position and properties of the Lys63 residue and Asp89 (as has been already demonstrated by NMR)[3;6]. Thus, in

♠ Personal communication D.D.Hau

order to definitively examine the role of Asp89, the above NMR experiments would have to be done on a more stable or non-hydrolysable Mms2•Ubc13~Ub thiolester, for example a serine ester. To date attempts to create a Ubc13(C87S)~Ub ester bond have proven to be elusive. In order to further substantiate roles within the chemical catalytic mechanism of isopeptide bond synthesis we may need to await a trapped reaction intermediate structure of an E2 bound to substrate and thiolester linked Ub or strategically monitor the structural changes within the E2 active site during catalysis (a job for a highly motivated NMR spectroscopist, or a Laue crystallography).

5.2 Conclusion

This dissertation describes a significant contribution to the scientific community particularly with regards to the mechanism of Lys63 linked poly-Ub chain synthesis. Furthermore, the core E2•UEV structures presented here, to which other interacting components of the Ub system have been and continue to be added, provides the basis for the studying the E2 catalytic mechanism and the mechanism of substrate ubiquitination.

5.3 Reference List

1. Moraes, T. F., Edwards, R. A., McKenna, S., Pastushok, L., Xiao, W., Glover, J. N., and Ellison, M. J. (2001) *Nat Struct Biol* **8,** 669-73

2. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6

3. McKenna, S., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., Spyracopoulos, L., and Ellison, M. J. (2003) *J Biol Chem* **278,** 13151-8

4. Andersen, P. L., Zhou, H., Pastushok, L., Moraes, T. F., McKenna, S., Ziola, B., Ellison, M. J., Dixit, V. M., and Xiao, W. Distinct Regulation of Ubc13 Functions by the Two Uev Proteins Mms2 and Uev1A. Cell . submitted.

5. Moraes, T. F. , Hu, J., McKenna, S., Xiao, W., Glover, J. N., Ptak, C., and Ellison, M. J. The Amino Terminus of Uev1 Stimulates Poly-Ubiquitin Chain Assembly by Weakening Interactions between Components in the Catalytic Complex. J Biol Chem . 2004.

6. McKenna, S., Hu, J., Moraes, T., Xiao, W., Ellison, M. J., and Spyracopoulos, L. (2003) *Biochemistry* **42,** 7922-30

7. Moraes, T. F., Varelas, X., Pastushok, L., Mckenna, S. A., Ptak, C., Xiao, W., Glover, J. N. M., and Ellison, M. J. (2004) *J. Biol. Chem* **Submitted,**

8. Brzovic, P. S., Keeffe, J. R., Nishikawa, H., Miyamoto, K., Fox, D. 3rd, Fukuda, M., Ohta, T., and Klevit, R. (2003) *Proc Natl Acad Sci U S A* **100,** 5646-51

Appendix A

Towards Quantitative Prediction of Protein:Protein Interactions: A Systematic Empirical Approach that Correlates Binding Energy to the Structural Interface.

Summary:

This appendix contains the rationale, experimental procedures and details of 10 crystal structures of derivatives of the Ubc13•Mms2 complex that are being used to correlated thermodynamic parameters to structural detail within a protein:protein heterodimeric interface.

Introduction and Rationale:

The long term success for the design of drugs based on the rational principles of chemistry and structural biochemistry hinges on the ability to determine the fundamental rules of biomolecular interaction and to shape these empirical rules into a reliably predictive, computational approach. This project has focused on the systematic acquisition of thermodynamic data that will be used to refine and/or test existing algorithms that are currently being used to predict protein-protein and protein-ligand interactions. The force-field approach underlies all of these 'docking' strategies, however, it formally fails to account for the effects of entropy on the interacting system. Until such time as these entropic effects can be understood quantitatively and within a broader theoretical framework, the prospect for accurately

predicting biomolecular interactions and designing interacting compounds, will remain to a certain extent serendipitous.

This appendix describes a preliminary dataset that correlates the thermodynamic parameters of a protein:protein heterodimeric interface with that of structural information employing Isothermal Titration Calorimetry (ITC) and X-ray Crystallography. To study protein:protein interactions we have chosen to utilize a heterodimeric complex consisting of Mms2 and Ubc13 as a model system[1]. These two proteins are found in a DNA repair pathway whose function is dependent on their tight binding interaction. Currently, 74 ITC datasets have been pocessed and 10 heterodimeric complexes structures have been determined.

Long term Goal: Project Cyber Cell.

In contributing to the development of the virtual cell, we must soon take in account the independent rates of interaction between known protein structures. With this in mind we will need to combine known binding constants and concentrations with predicted binding constants determined through this methodology and others. This dataset not only provides a sounding board for testing the predicted parameters but also is a small step towards a dataset that will help rationally refine the current docking prediction software.

Experimental Procedures:

Protein Design, Expression, Purification:

Nine residues in each of Mms2 and Ubc13 were chosen to be mutated to alanine based on their potential contribution to the interaction between Mms2 and

Ubc13 using the crystal structure of the Mms2•Ubc13 heterodimer as a reference[1]. All Mms2 and Ubc13 constructs were made using a two-step PCR based site-directed mutagenesis strategy employing the Mms2-pGex and Ubc13-pGex vectors as templates. Mms2 and Ubc13-GST fusions were purified and subjected to precission protease cleavage to provide a final homogenous protein solution of either Ubc13 or Mms2 as previously described[2].

Isothermal Titration Calorimetry:

A VP-ITC MicroCalorimeter (Microcal, Northampton, MA) was used to analyze binding of Mms2 with Ubc13 (or derivatives). Proteins were dialyzed against 50 mM Tris, pH 7.5, 150 mM NaCl, and 1 mM ETDA and degassed prior to analysis. The dialysis buffer was placed in the reference cell. Either the UEV or E2 was titrated against dialysis buffer to obtain the heat of dilution. The following parameters were used in the titration: 30 °C, 10 μl injections, and 4 min between injections with stirring at 305 rpm. The data was fit to a titration curve using the program Microcal Origin (v 5.0) to extract thermodynamic parameters; only K_Ds will be listed in this appendix[α].

Protein Crystallization:

Heterodimers were assembled by mixing equal molar concentrations of Mms2 and Ubc13 and concentrated to 10mg/mL in 50 mM Hepes, pH 7.5, 150 mM NaCl, 1

[α] All protein purification and ITC data was perfoemed by J. Hu. For more details and ITC data please refer to J.Hu.

mM EDTA and 1 mM dithiothreitol (DTT). All crystals were grown in precisely the same conditions, 100mM Citrate pH 6.5, 20% PEG 6000 and 50 mM $MgCl_2$ at 22°C using the hanging drop method. Crystals were soaked into cryo-protectant using a 5% step wise gradient of glycerol to a maximum of 30% with 3 min. incubations between steps. Crystals were flash frozen in liquid nitrogen and screened before data collection at the Advanced Light Source in Berkeley, California.

Structure Determination:

Diffraction data for the derivatives of hMms2•hUbc13 were collected from single crystals at Advance Light Source (ALS, Berkeley, California), beamline 8.3.1 using a 2 x 2 CCD array (ADSC) detector. Each data set was integrated, scaled, and processed by the automated software package ELVES[3]. Merged.mtz amplitudes were subjected to truncate, CAD and standardized with respect to a unique R_{free} set of reflections using the hMms2•hUbc13 as the standard. All data was then rigid-body fit to 1JD4 (hMms2•hUbc13) model using CCP4 Refmac5 to give a initial R_{fac} and R_{free}. Manual rebuilding was performed with reference to Sigma A-weighted[4] 2|Fo| - |Fc| electron density maps calculated in Refmac5. All structures were refined with Refmac5[5].

Results and Discussion:

Residues in the interface of Ubc13 and Mms2 depicted in **Figure A.1** were mutated to alanines to test their contribution to the interface. There were 10 Mms2

constructs and 10 Ubc13 constructs (including wildtype) of which 74 combinations were subjected to ITC analysis. Of these combinations 8 Mms2-derivatives with wildtype-Ubc13 and 8 Ubc13-derivatives with wildtype-Mms2 were subjected to crystal screens; 11/16 crystallized and 10/16 data sets were collected and refined (**Table A.1**). The additional crystallized heterodimer (Ubc13•Mms2-V07A) did not survive the cryo-treatment.

None of the structures demonstrated gross conformational rearrangements but rather minor structural movements to compensate for either a loss of hydrophobic contact or a hydrogen bond partner for example the Mms2•Ubc13-R85A structure. Ubc13-R85 hydrogen bonds to Mms2-E20 in the wildtype heterodimer crystal structure. The R85A mutation causes the E20 residue in Mms2 to rotate 60° around $\chi 1$ moving it out of the interface (**Figure A.2**) and results in a 16 fold decrease in affinity (K_D), a change in enthalpy of 2400 cal/mol and an increase in entropy of 2.3 cal/mol/°C. This preliminary dataset of structures is a first step in systematically correlating a structural interface to thermodynamic binding parameters. Further analysis of these datasets are currently underway at the Uppsala University in Sweden by Lennart Nilsson.

Reference List

1. Moraes, T. F., Edwards, R. A., McKenna, S., Pastushok, L., Xiao, W., Glover, J. N., and Ellison, M. J. (2001) *Nat Struct Biol* **8,** 669-73
2. McKenna, S., Spyracopoulos, L., Moraes, T., Pastushok, L., Ptak, C., Xiao, W., and Ellison, M. J. (2001) *J Biol Chem* **276,** 40120-6
3. Holton, J. and Alber, T. (2004) *PNAS* **101,** 1537-1542

4. Read, R. J. *Acta Crystallogr. A* 42, 140-149. 86.

5. Murshudov, G. N., Vagin, A., and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst* D53, 240-255. 97.

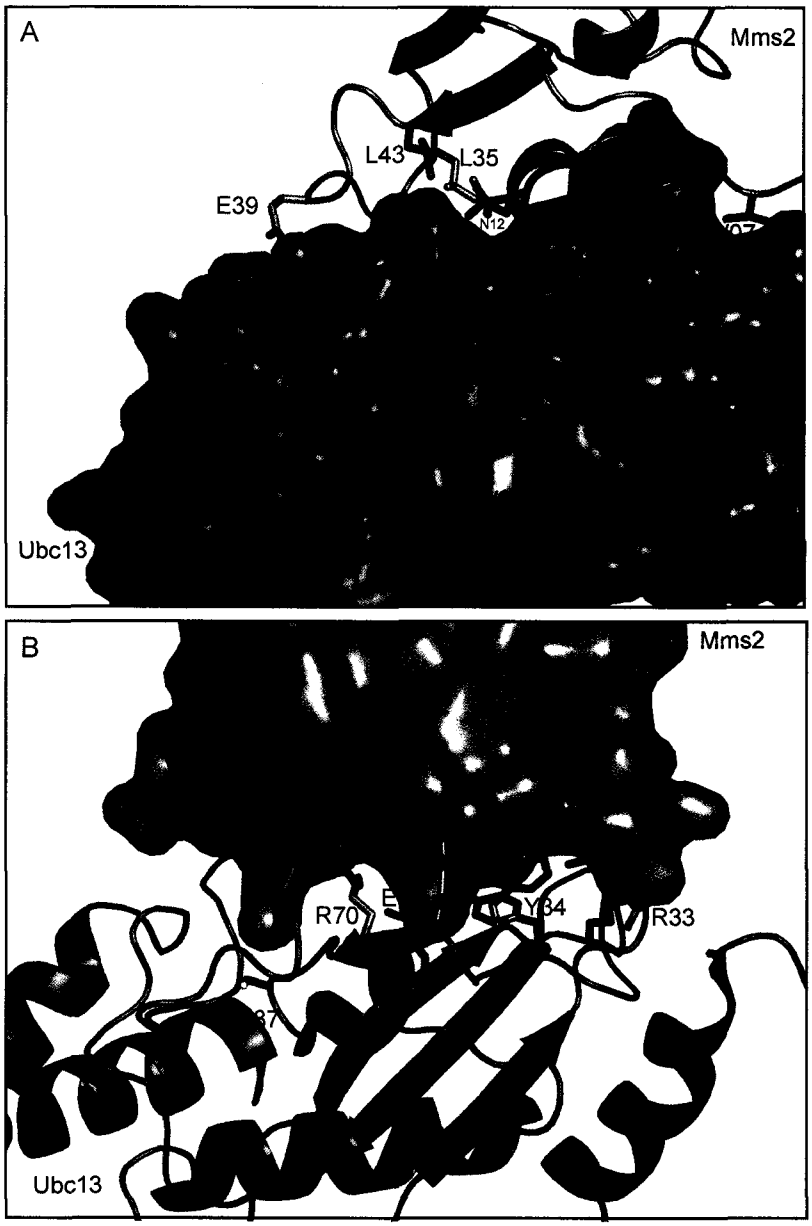


Figure A.1 Residues in the Interface of the Mms2·Ubc13 heterodimer. Mms2 is highlighted in gold and Ubc13 in teal with the active-site Cys87 residue colored yellow. **A.** The surface of Ubc13 and a ribbon representation of Mms2 depicting the nine residues in Mms2 that were individually mutated to alanines (V07, N12, F13, L16, L19, E20, L35, E39, and L43). **B.** The surface of Mms2 is shown with a ribbon representation of Ubc13 illustrating the nine residues in Ubc13 that were individually mutated to alanines (N31, R33, Y34, E55, F57, K68, R70, L83 and R85).

Table A.1 A Table of Crystallographic Statistics for the Crystallized Mms2•Ubc13 Interface Derivatives.

	Ubc13					Mms2				
Mutant Residue	**Asn31**	**Tyr34**	**Glu55**	**Lys68**	**Arg85**	**Asn12**	**Leu16**	**Glu20**	**Leu35**	**Leu43**
Completeness %	99.1	92.7	99.78	95.67	97.3	92.8	98.3	95.89	82	47.5
# of reflections	29139	31048	33264	32049	50288	23665	25665	24806	11274	14672
Resolution (Å)	46-1.76	40-1.68	46-1.68	40-1.68	12.7-1.3	39-1.84	45.6-1.84	46-1.84	47-2.3	46-1.8
R_{fac}/ R_{free}	**23.5/26.7**	**22.1/25.4**	**21.7/23.4**	**21.4/24.5**	**21.8/25.8**	**21.8/25.0**	**21.4/25.0**	**21.8/25.5**	**19.1/26**	**20.6/26.8**
RMS bond length(Å)/angle(°)	0.012/1.4	0.011/1.4	0.01/1.3	0.01/1.3	0.012/1.4	0.015/1.5	0.013/1.4	0.012/1.4	0.02/1.8	0.02/1.8
# residues/waters	289/50	289/190	289/190	289/190	289/0	289/190	289/190	289/190	289/190	289/190
RMSD with Ubc13•Mms2 (Å)	0.30	0.29	0.40	0.25	0.33	0.33	0.27	0.33	0.48	0.32
K_D (Wt=50nM)	**840nM**	**360nM**	**1070nM**	**212nM**	**830nM**	**34nM**	**1360nM**	**91nM**	**84nM**	**414nM**

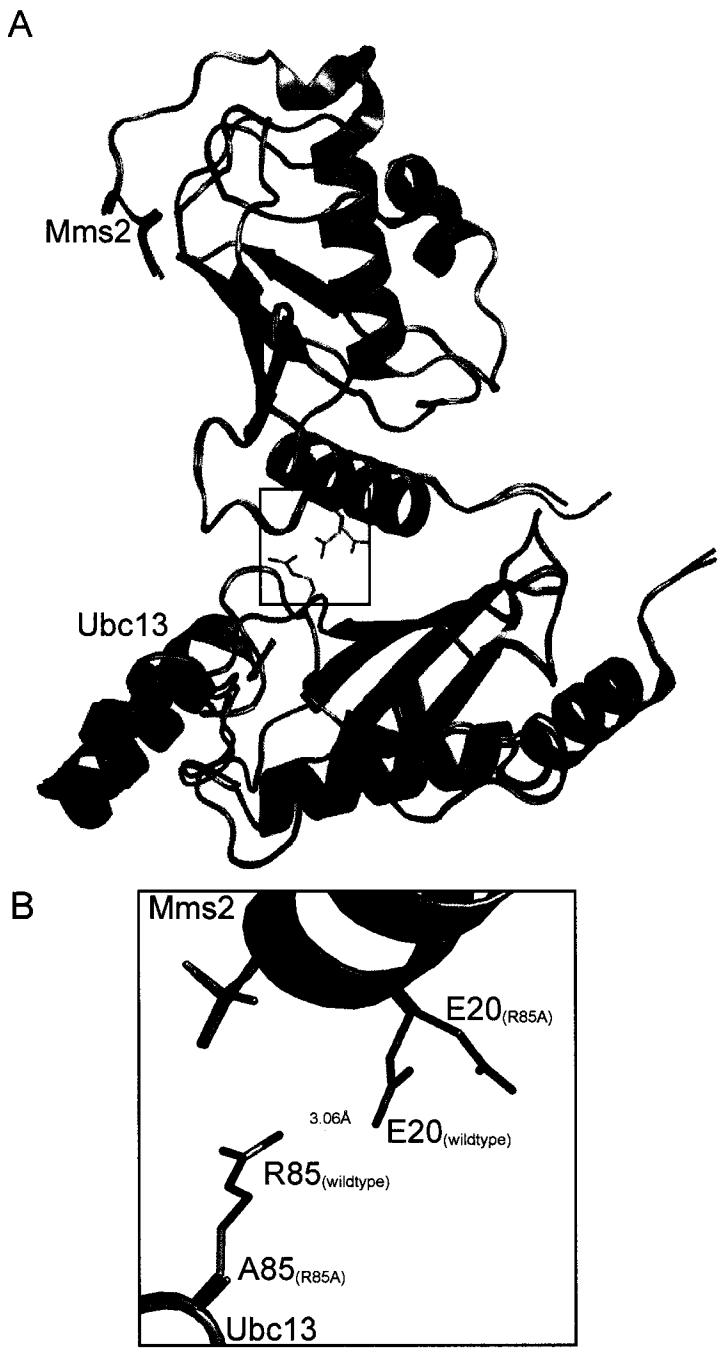


Figure A.2 Correlating Structural Rearrangements with Thermodynamics. **A.**The structure of the wildtype Mms2•Ubc13 heterodimer (gold and teal) is structurally aligned with the Mms2•Ubc13-R85A derivative (grey) with RMSD of 0.3Å demonstrating no major structural differences. **B.** A close-up of the R85 region of the crystal structure illustrates only one minor structural rearrangement. The E20 residue of Mms2 rotates ~60° around χ_1 as it is no longer connected by a salt-bridge to Ubc13's R85. Removing this interaction causes a 16 fold decrease in K_D (50 to 830 nM), a change in ethalpy upon association of 2400 cal/mol and an increase in entropy of 2.3 cal/mol/°C.