STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE U-BOX DOMAIN OF THE E4B UBIQUITIN LIGASE

By

Kyle Andrew Nordquist

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Biochemistry

August, 2011

Nashville, Tennessee

Approved:

Walter J. Chazin

Tina M. Iverson

Daniel C. Liebler

BethAnn McLaughlin

ZuWen Sun

For my beautiful daughter,

Hava Corinne,

To all our science experiments in the future…

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

None of the work completed in this dissertation would have been possible if not for an extraordinary collaborative effort and support spanning both work and play. I of course acknowledge the wonderful mentorship of my advisor, Dr. Walter Chazin. While times arose that we did not see eye-to-eye, I believe I have come to understand and appreciate the values that he has instilled in me and all his other students. I have no doubt that the many things I have learned under his guidance will benefit me in all my future endeavors.

The other members of the Chazin lab were instrumental in my success as a graduate student. First, Yoana Dimitrova's selflessness in helping out her fellow graduate students meant wonders. No matter how busy she was (which was always!), she would always take time out of her schedule to help me with a protocol or experiment. Sarah Soss has also been key during the latter stages of this work, and the many discussions detailing how ubiquitin is the king of all proteins have been insightful. Despite her multitude of undertakings, Susan Meyn displayed unbridled patience with me during the initial stages of my study, and it was certainly heartbreaking to all of us when she left our group. Young-Tae Lee and Brian Weiner made the initial NMR experiments and structure determination process possible. For all the other members of the Chazin lab, I thank you all for all the good times that we had as a group.

I also thank the members of my thesis committee for all the insightful discussions during my meetings. It was certainly difficult to get everyone in the same room sometimes, but your guidance has proved valuable.

It may have taken me a while to figure out that science is easier done through collaboration, but this recognition has fostered some amazing teamwork from outside of Vanderbilt. Rachel Klevit and her group have been instrumental throughout my thesis work, and I am anxious to see the conclusion of our current work. Kim Munro was nice enough to collect my AUC experiments, and Whitney Ridenour, while a Vanderbilt student, helped collect my Mass Spectrometry data.

Throughout my stay in Nashville, I have been more than blessed with the friends that I have made outside of Vanderbilt. My "East Nasty" connection has been the epicenter of many good times, and I honestly don't know how I will cope without our eccentric group. Tyler, Rick, and Stephen have been unbelievable friends, and their understanding, support, and generosity, especially over the past year and a half, have meant more than I can describe. I can only wish that I am half the friend you three were to me.

Of course, thanking my family goes without saying. Mom and Dad, you never told me what I couldn't do, and that simple attitude has been appreciated. I thank my luck every day of being raised by such supportive parents. Erik, I am grateful to call you my brother and best friend, and I thank you endlessly for instilling your early gift of music appreciation in me.

And finally, to my daughter Hava: much like your grandparents taught me, you can do anything you desire. Watching you grow these past 6 years has been more exciting than any science experiment ever could be. Coming home to a smiling child is a fantastic recipe to cure any scientific blues. I can't wait for everything to come, I love you.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATP	Adenosine trihosphate
AUC	Analytical Ultracentrifugation
Cdc	Cell Division Cycle protein
CFTR	Cystic Fibrosis Trasnmembrane Regulator
CHIP	Carboxy-terminal of Hsc70-Interacting Protein
DNA	Deoxyribonucleic Acid
ER	Endoplasmic Reticulum
ERAD	Endoplasmic Reticulum Associated Degradation
HECT	Homologous to E6-Carboxy Terminus
Hsp	Heat Shock Protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
NMR	Nuclear Magnetic Resonance
MALS	Multiangle Light Scattering
PTM	Post-translational modification
Prp19	Precursor RNA processing 19 protein
RING	Really Interesting New Gene
SAXS	Small Angle X-ray Scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ub	Ubiquitin
UbcH5	Ubiquitin conjugating human 5 protein
U-box	Ufd2-box

CHAPTER I

INTRODUCTION

Overview

Cellular signals propagated by the small protein modifier Ubiquitin (Ub) are indispensible for maintaining cellular homeostasis. Ub signaling plays major roles in protein localization, DNA repair, transcription activation, chaperone response, and cell stress most frequently resulting in targeting substrates for degradation by the 26S proteasome (*1-4*). Flaws in ubiquitination can be devastating to cellular regulation and lead to physiological defects like the neurological disorder Parkinson's Disease, as well as a variety of cancers such as breast, ovarian, and specific classes of lymphoma (*5*).

The role of ubiquitin in targeted protein degradation is now well established. Ubiquitin mediated proteolysis plays a significant role in a range of important cellular functions (*6*). Endoplasmic Reticulum Associated Degradation (ERAD), which is essential for degradation of misfolded or unfolded protein species, is arguably the most well studied example. Ub-meidated proteolysis is also essential to progression through the cell cycle, in particular at the conclusion of mitosis where poly-ubiquitination and subsequent degradation of cyclins signals transition from mitosis to G1 (reviewed in (*7*)). The levels of the well-known tumor suppressor p53 are kept low by ubiquitin-dependent proteolysis its degradation must be inhibited to initiate p53 signaling reviewed in (*8*).

Recent investigations of the Ub system also reveal critical roles in non-proteolytic signaling pathways. This includes activation of histones, where monoubiquitination

signals for the chromatin unwinding that allows transcription associated factors access to bind DNA and intiation of transcription (*9*). Another important role for non-proteolytic Ub signaling is internalization and endocytosis of various membrane bound proteins (*1*). During DNA double strand break repair, ubiquitination associated proteins such as BRCA1, RNF8, and Ubc13 are recruited to sites of chromatin damage where it has been suggested that Ub^{K63} chains help escalate the DNA damage response (*10*).

Unlike small molecule post-translational modifications, ubiquitination requires multiple enzymatic reactions. Precise coordination between three classes of enzymes (E1, E2, and E3) governs the fate of the target substrate. More specifically, the exact type of ubiquitination that occurs is dependent upon which particular E2 and E3 enzymes are involved. It is currently believed that the specific combination of E2-E3 pairs determines whether a substrate is mono- or poly-ubiquitinated, and if the latter, what topology of Ub chain is generated (*11-13*). While previous studies have characterized the steps of the ubiquitination cascade, current research suggests that there is still a wide gap in knowledge in understanding the molecular mechanisms that govern ubiquitination.

The research described in this dissertation analyzes the mammalian E3 ubiquitin ligase, E4B, and assesses its influence on the E2~Ub conjugate. E4B is a unique monomeric U-box E3 ligase that provides a powerful tool for the study of the mechanism of ubiquitination. The results described here provide insight into the structure, E2 binding characteristics, and the activation of the E2~Ub conjugate by an E3 ligase. This work provides a foundation for future studies directed to further understand the basis for ubiquitin-substrate specificity.

Ubiquitination and cellular signaling

Ubiquitin is a 76 amino acid β-grasp fold protein that was originally described as a lymphocyte differentiation factor (*14*) (Figure 1.1). Ubiquitin was so named because of its universal presence in most cellular species in addition to its extremely high degree of conservation. Groundbreaking work by Ciecanhover, Hershko, and Rose demonstrated Ub's function as a unique ATP-dependent proteolysis factor

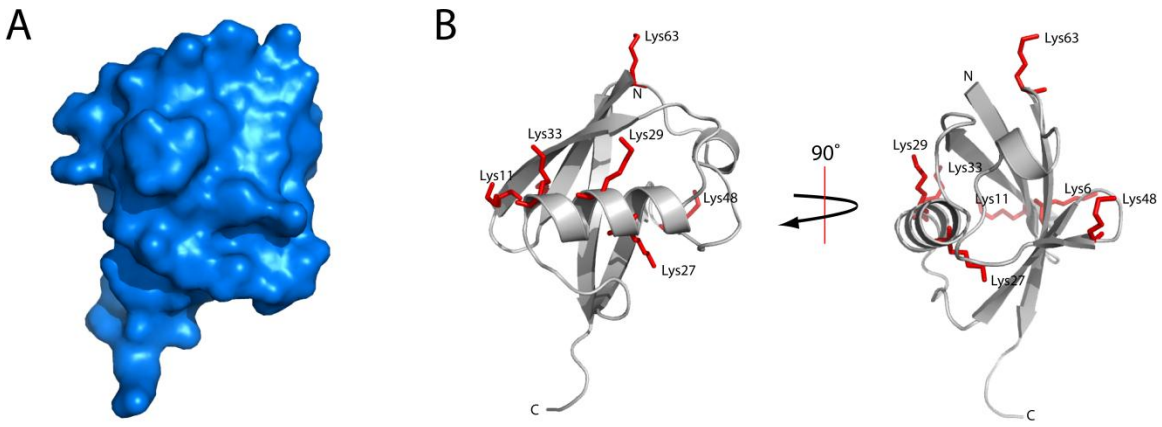

Figure 1.1. Structure of Ubiquitin. A) Surface diagram B) Cartoon representation showcasing lysine side chains. Produced from PDB ID 1UBQ (*15*).

crucial for protein degradation in the absence of lysosomes (*16*, *17*). Ubiquitin was shown to uniquely modify other proteins through covalent linkage and generate high molecular weight species, an observation that was necessary for the eventual degradation of the substrate (*18*). Further work established the role of the 26S proteasome in this system. (*19*).

Ubiquitin contains several unique structural features that enable it to interact with a vast number of proteins in the cell, as well as with itself. For example, the majority of protein-protein interactions mediated by ubiquitin are done so through a critical hydrophobic patch composed of residues Leu8, Ile44, and Val70. The C-terminal tail of ubiquitin is exposed, and allows for the covalent attachment of Ub to a variety of substrates via linkage with the ε-amino side chain of a lysine residue. Ubiquitin itself contains seven lysine residues (6, 11, 27, 29, 33, 48, and 63) which form the basis for poly-ubiquitin chain synthesis. Together, ubiquitin can mediate a diverse array of protein-protein interactions as well as propagate a variety of cellular signals through polyubiquitin chain formation.

Ubiquitin Binding Domains

When a substrate undergoes a PTM, the chemical and physical properties of that protein are altered. Ubiquitination of a target protein not only changes the structural landscape of that target, but it may also act as a new binding domain for other molecules. Furthermore, ubiquitin can also block other PTMs as well as protein-protein interactions between the substrate and other targets. Throughout the human proteome, approximately 20 types of ubiquitin-binding-domains (UBDs) exist (*20*) (Figure 1.2).

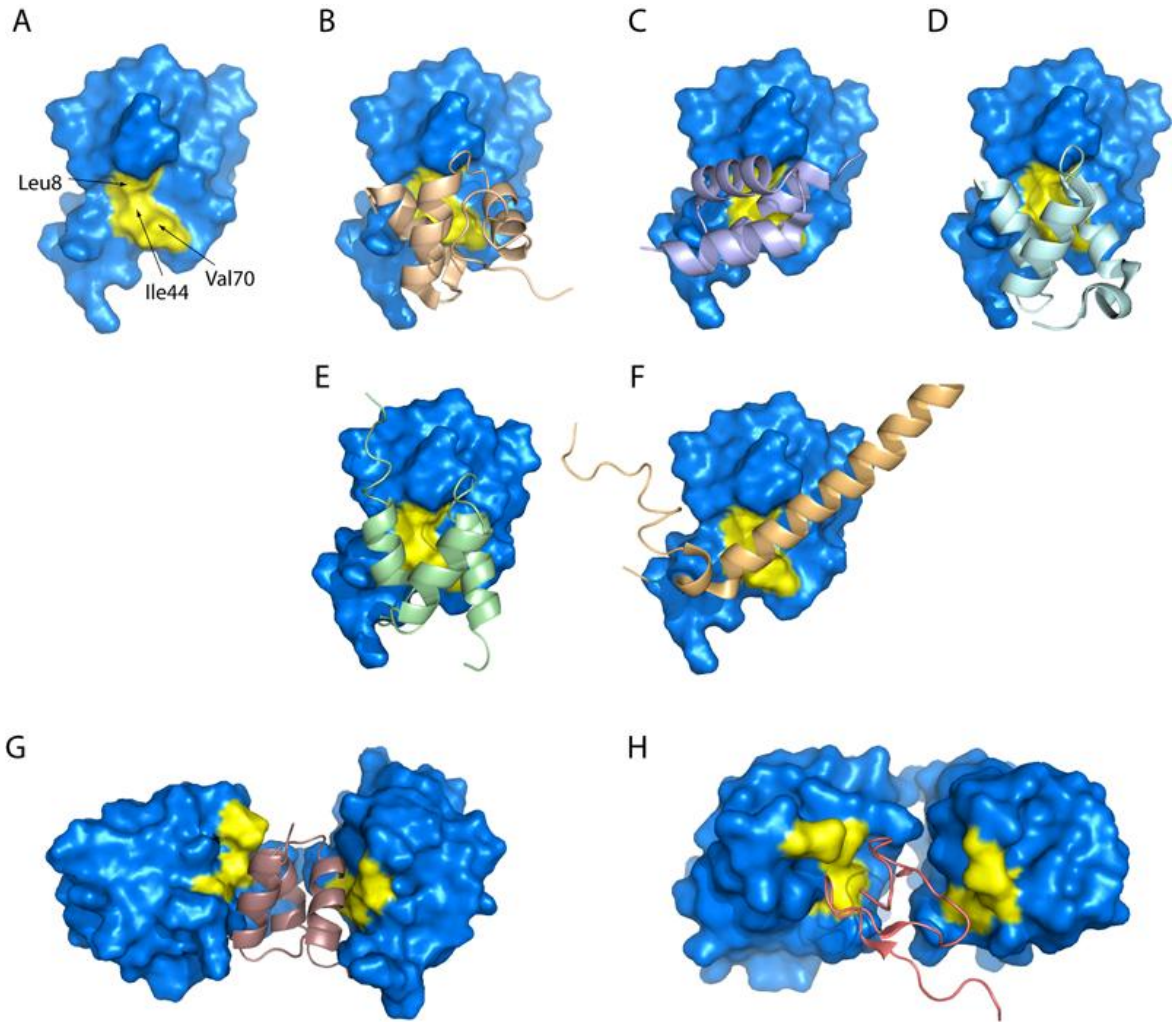

Figure 1.2. Ubiquitin Binding Domains A) Surface representation of Ubiquitin with essential hydrophobic patch highlighted in yellow. Ubiquitin bound to B) UBA domain of Ubiquilin (wheat, PDB ID: 2JY6 (*21*)); C) UBA domain of Cbl-b (violet, 2OOB (*22*)); D) UBA domain of Edd (teal, 2QHO (*23*)); E) CUE domain of Cue2 (light-green, 1OTR (*24*)); F) Miz domain of S5a (beige, 1YX5 (*25*)). G) Model of UBA domain of Rad23 bound to di-Ub^{K48} (purple, 1ZO6). F) NZF domain of TAB2 bound to di-Ub^{K63} (orange, 3A9J (*26*)).

Proteins containing these domains recognize and interact with Ub-modified substrates, depending on the type of Ub modification that has occured. Most often, UBDs utilize α-helical motifs to bind a critically conserved Ile44 residue exposed on Ub (*20*). For example, Ubiquitin-interaction motifs (UIM) and motifs interacting with Ubiquitin (MIU) use single α-helices for Ub binding whereas UBDs such as Ubiquitin-associated domains (UBA) and Coupling of Ubiquitin conjugation to Endoplasmic Reticulum Degradation domains (CUE) use split α-helices. (*20*).

Not all UBDs utilize α-helical structures to interact with Ub. Deubiquitinases (DUBs) are necessary to both reverse ubiquitination events as well as recycle Ub species at the cap of the proteasome. DUBs employ β-sheets to interact with the strictly conserved I44 of Ub (*27*). Moreover, a select number of E2 ligases also use β-sheets to bind Ub in a non-covalent manner, which is necessary for processive poly-ubiquitin chain elongation (*28*).

Ubiquitin Chain Topology

Perhaps the most unique structural quality of Ub is its ability to form polyubiquitin chains, i.e., after initial monoubiquitination of a target protein, the ubiquitination process may repeat. Ub can attach covalently to another Ub molecule through any one of its 7 lysine residues (K6, K11, K27, K29, K33, K48, and K63) or, to a lesser extent, its N-terminal methionine (*29*). Because the different lysine residues are located in different structural areas of Ub, radically different topologies can be formed allowing for different UBDs to bind to different chain linkages (Figure 1.3).

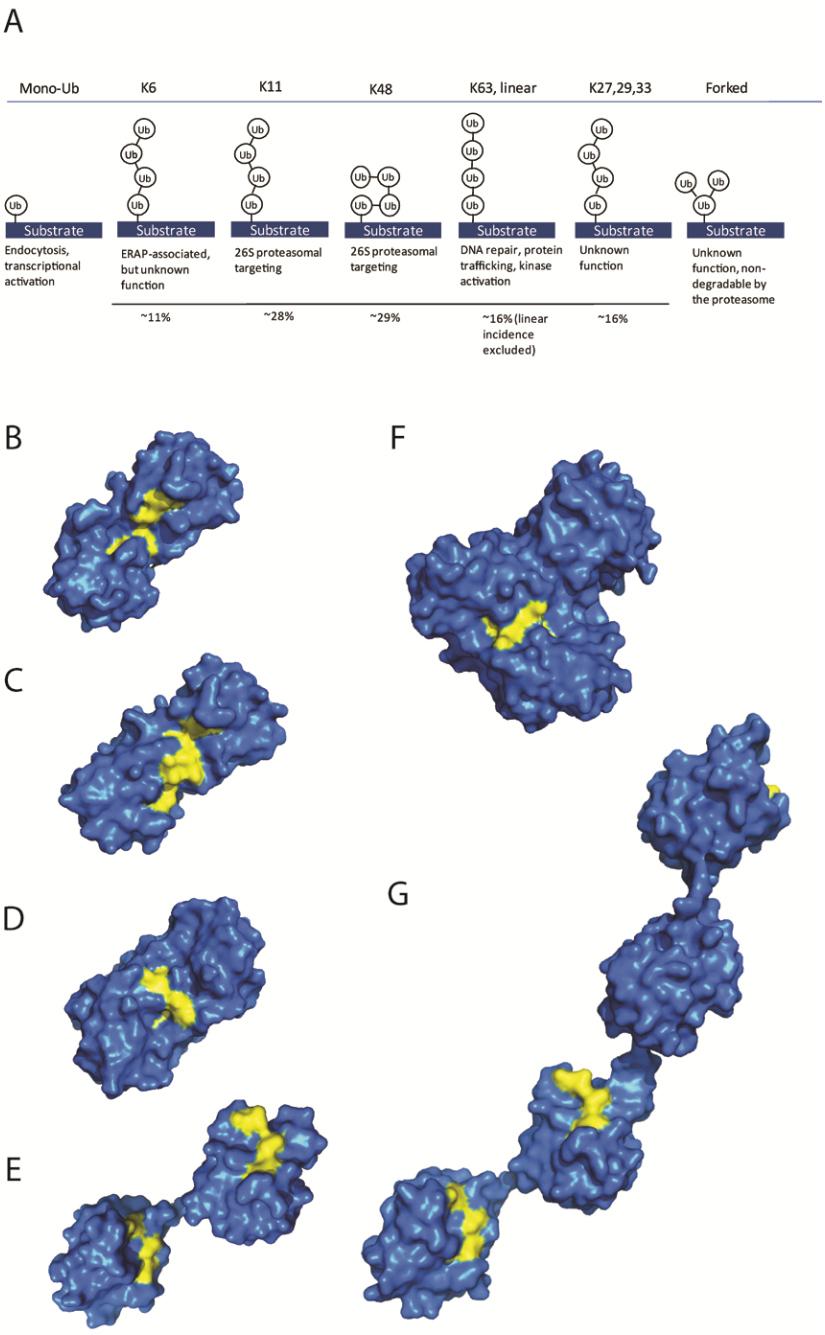

Figure 1.3. Ubiquitin chain topologies. Hydrophobic Leu8 – Ile44 – Val70 surface highlighted in yellow. A) Ubiquitin chain topologies and their known cellular functions and incidence. B) di-Ub^{K6} (PDB ID: 2XK5); C) di-Ub^{K11} (3NOB (*30*)); D) di-Ub^{K48} (2BGF (*31*)); E) di-Ub^{K63} (2W9N (*29*)); F) tetra-Ub^{K48} (2O6V (*32*)); G) tetra-Ub^{K63} (3HM3 (*33*)). Structural information not available for chains of linkages. K27, K29, K33, or forked linkages.

Formation of a poly-ubiquitin chain on a particular lysine residue has crucial implications for the resulting cellular signal. For instance, protein degradation by the 26S proteasome is the dénouement of K48 (Ub^{K48}) linkages, and is possibly the most widely studied aspect of polyubiquitination. When formed, Ub^{K48} chains adopt a distinct packed globular conformation, burying the conserved Leu8/Ile44/Val70 hydrophobic patch. Burying this patch inhibits the majority of UBDs from interaction with Ub. Therefore, this topology is only recognized by UBDs that bind to the specific motifs created between the modified K48 and G76 residues of the chain (*27*).

It has also been established that several Ub^{K48} binding proteins contain domains that mimic the β-grasp Ub-fold. These motifs, termed Ubiquitin Like Motifs (UBLs) bind to the protein subunit Npn10 located in the 19S cap of the proteasome (*34*). The ability of Rad23, Dsk2, and Ddi1 to concomitantly bind both Ub^{K48} motifs and proteasome subunits defines their central roles in protein degradation.

Ub^{K63} chains, on the other hand, adopt linear structures, exposing the majority of the Ub surface (*27*). Most notably, the Leu8/Ile44/Val70 hydrophobic patch is exposed allowing a plethora of UBD-containing proteins to recognize and bind Ub, propagating diverse cellular signals. Ub^{K63} chains play critical roles in activation of DNA repair checkpoints, most notably through tightly regulated modification of PCNA (*35*).

Despite only limited structural knowledge of polyubiquitin chain topologies, functional studies have recently demonstrated that Ub^{K11} chains also signal substrates for proteasomal degradation (*36*). Based on these cumulative studies, it is fascinating to note that the incidence of Ub^{K11} chains occurring in the cell is approximately 28.0±1.4% compared to the 29.1±1.9% and 16.3±0.2% of Ub^{K48} and Ub^{K63}, respectively. This

suggests a much wider role for Ub^{K11} chains in cellular signaling than previously thought. Ub^{K11} is known to modify many substrates, and is a crucial Ub chain topology associated with ERAD. At this time, complete structural information of Ub^{K11} chains is unknown, and it remains unclear whether unique proteasome shuttling factors exist for this Ub chain topology.

While *in vitro* studies in yeast have shown that all seven lysine residues of Ub can form an anchor for chains, specific cellular functions of all chain types have yet to be elucidated (*36*). Ub^{K6} chains have been shown to play an important role in BRCA1 mediated DNA damage response, but like Ub^{K11} chains, only limited structural information is currently available (*37*). More perplexing, substrates of Ub^{K6} in this process have yet to be identified, so it remains to be determined if BRCA1 (an E3 ubiquitin ligase) generates this chain topology or if other E3 ligases are necessary for the construction of this particular chain.

It has become increasingly evident that polyUb chains can form from the N-terminal methionine of Ub in addition to lysine residues as evidenced by the NEMO activation of NF-κB (*38*). In this pathway, the LUBAC E3 ligase complex will first ubiquitinate NEMO on specific lysine residues, and then form polyubiquitin chains using the N-terminal methionine residue of these ubiquitins. There is also evidence for forked Ub chains. These have been observed on multiple lysine residues on a substrate, but the cellular implications of this topology have not been determined except to note they do not signal for degradation by the proteasome (*39*). In summary, despite knowing the diverse effects of polyubiquitin signaling, the factors governing the elongation of Ub chains have yet to be elucidated.

The Ubiquitination Cascade

Ubiquitin modifies proteins through covalent attachment of its C-terminal glycine residue to the ε-amino group of an acceptor lysine on the target protein. In order for Ub modification to occur, the action of three enzymes must be precisely orchestrated to mediate the three steps of the overall reaction (*40*) (Figure 1.4). First, the E1 activating enzyme adenylates the C-terminal glycine of Ub in an ATP dependent manner. This step forms a thiolester bond between the catalytic cysteine residue of E1 and the C-terminal glycine of Ub. Second, the E1 hands Ub off to an E2 conjugating enzyme, which also binds Ub by a thiolester bond between a cysteine in the E2 and the C-terminal glycine in Ub. Finally, an E3 ligating enzyme recruits both the target substrate and the E2~Ub conjugate. Remarkably, the E2~Ub conjugate is activated by binding the E3 ligating enzyme which promotes Ub transfer to a lysine residue of the substrate. Chain building involves subsequent steps of E2~Ub activation and transfer to lysines on the Ub substrate. In some instances, an additional enzyme, termed an E4 ligase, may be necessary to build ubiquitin chains. While it has been shown that such enzymes catalyze elongation of Ub chains, it has not been established if the E4 has distinct enzymatic properties or if its function is merely a special case of E3 ligase activity.

Overall, the Ub cascade is a cumulative process that facilitates the recognition of, and enzymatic activity on, a large number of substrates. Notably, the number of enzymes able to catalyze the successive steps increases. For mammalian ubiquitination, only 2

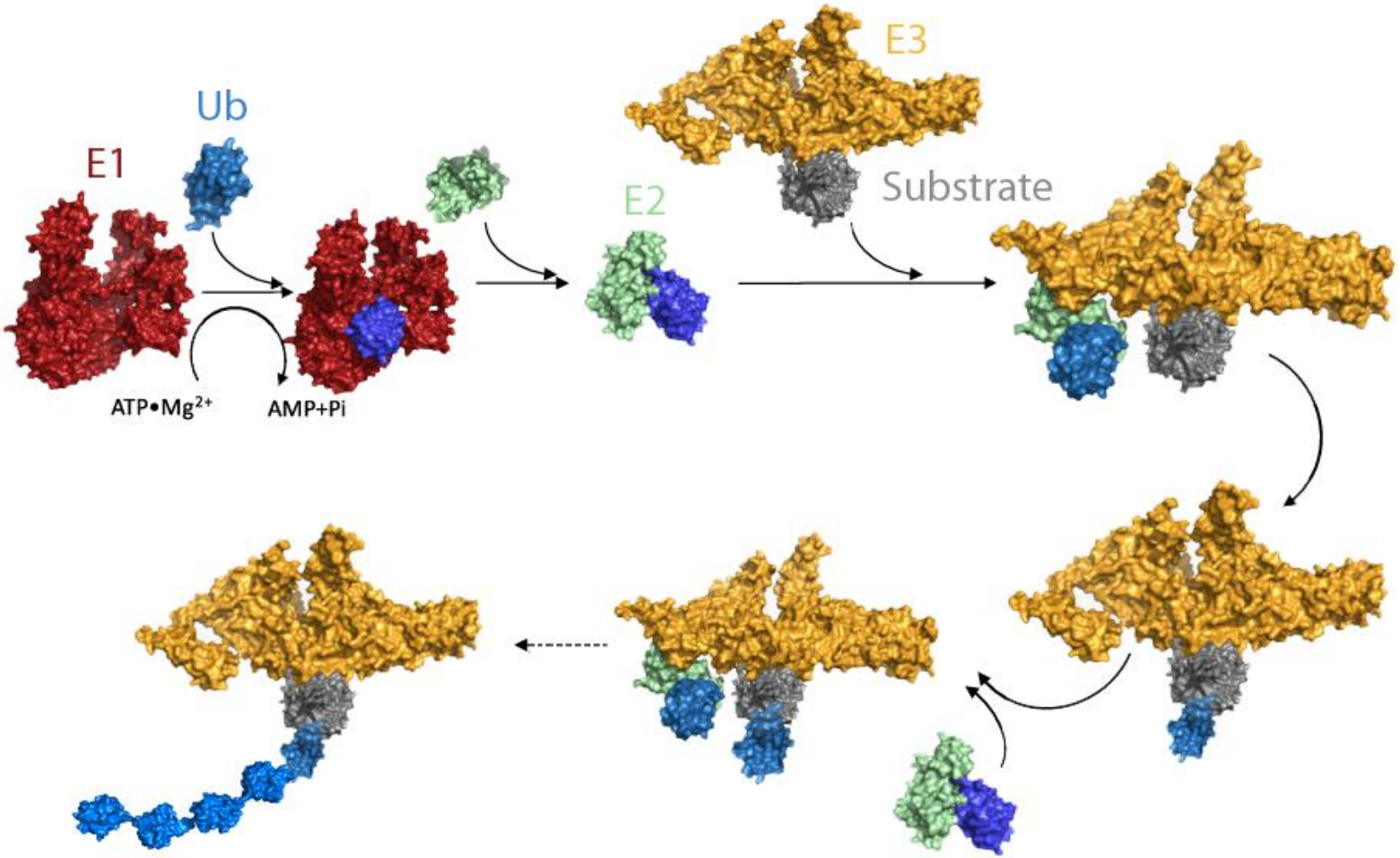

Figure 1.4. The Ubiquitination Cascade. Surface representations of proteins from the following PDB IDs: E1 and E1 bound to Ub (3CMM (*41*)), Ub (1UBQ (*15*)), E2 bound to Ub (3A33 (*42*)), and E3 (2QIZ (*43*)). For the purpose of this Figure, substrate and substrate bound to Ub are hypothetical models.

genes encode E1 activating enzymes, more than 30 encode E2 conjugating enzymes, and an ever expanding number of genes (currently above 600) code for E3 ligating enzymes (*5*). Along with the fact that E3 ligases can often bind multiple substrates, this hierarchy is critical to the complex and diverse nature of substrate modification through ubiquitination.

The E1 activating enzyme: E1 enzymes are large proteins comprised of three domains: an adenylation domain, catalytic cysteine domain (CCD), and an E2-interacting domain (*41*) (Figure 1.5). Two ThiF homology motifs form the adenylation domain, which bind both ATP•Mg^{2+} and Ub. The conserved cysteine located in the CCD forms the thiolester bond with Ub post-activation. A C-terminal Ubiqutiin fold domain (UFD, not to be confused with the *U*biquitin *F*usion *D*egradation family of proteins) recruits the E2 conjugating enzymes. Uba1 is the universal E1 enzyme in mammalian cells, which is reflected in the fact that the UFD of Uba1 recognizes nearly all E2 enzymes. In contrast to Uba1, Uba6 appears to only interact with the E2 enzyme, Use1 (*44*).

The E2 conjugating enzyme: There are currently 4 classes of E2 conjugating enzymes, though all share a common conserved central domain of approximately 150 residues consisting of an anti-parallel β-sheet flanked by pairs of α-helices (*45*) (Figure 1.6). Class I E2 enzymes only contain the core domain. Class II and III enzymes contain additional residues N- or C- terminal to the core domain, respectively. Class IV enzymes are rare, and contain additional residues both N- and C-terminal to the core domain.

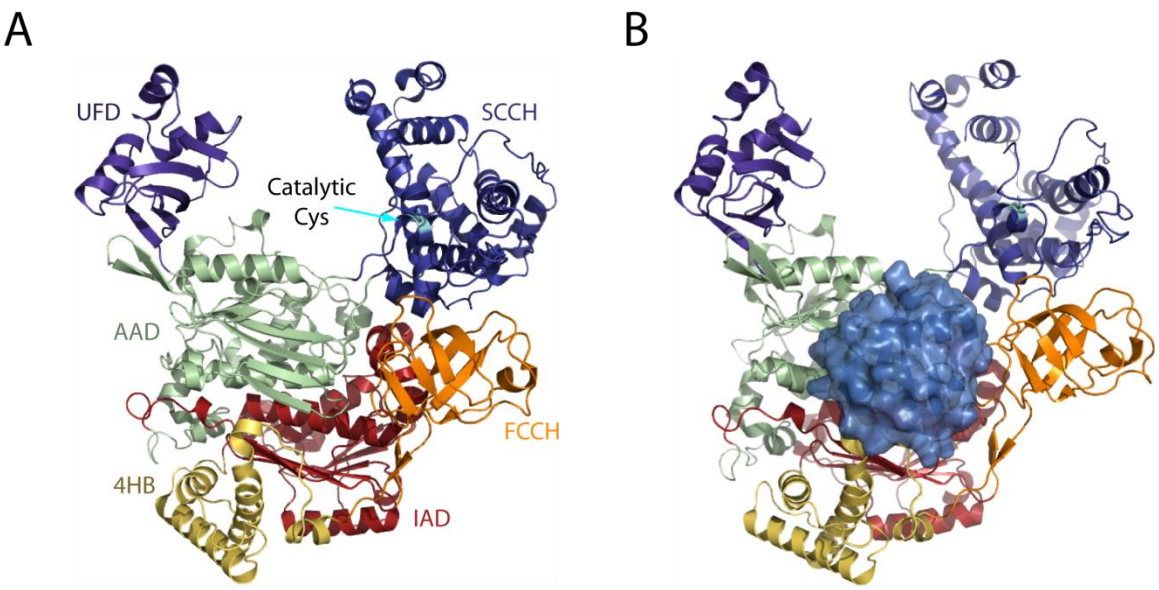

Figure 1.5. E1 activating enzyme. A) *Saccharomyces cerevisiae* Uba1 crystal structure in the absence of ubiquitin bound to the catalytic core. B) Same as (A) but bound to ubiquitin. Models generated from PDB ID 3CMM (*41*).

E2 enzymes contain a catalytic cysteine residue that is responsible for forming the thiolester bond with the C-terminal glycine of Ub. The most widely studied E2 enzyme, UbcH5 defines this group and is a class I E2 enzyme that has 4 mammalian isoforms, differing only minimally in amino acid sequence. UbcH5 has been described as a promiscuous enzyme because it interacts with the vast majority of E3 ligases and can both mono- and poly-ubiquitinate substrates without additional E2 enzymes (*46*). UbcH5 is also unique because its anti-parallel β-sheet can bind Ub in a non-covalent manner, and this characteristic is essential for the processivity of the enzyme in poly-ubiquitination (*28*). Despite evidence of UbcH5's interactions with a broad spectrum of E3 ligases, the mechanisms underlying UbcH5 function with regards to substrate modification

specificity (i.e., monoubiquitination versus lysine-specific polyubiquitination) remain unknown.

In contrast to UbcH5, Ubc13 appears to have a single function: building Ub^{K63} chain linkages. Like UbcH5, Ubc13 has been shown to interact with a significant number of E3 ligases, but its functions lie solely in non-proteolytic Ub pathways (*47*). This is mediated through a Ubc enzyme variant (Uev), Uev1a (yeast homologe Mms2), that interacts with Ubc13 in the cell (*48*). Formation of the Ubc13/Uev1a heterocomplex is necessary for Ubc13 to function in ubiquitination. Unlike Ubc enzymes, Uev1a does not contain a catalytic cysteine residue, and thus does not form thiolester bonds with Ub. However, like UbcH5, Uev1a binds Ub non-coavalently through its antiparallel β-sheet. Therefore, two molecules of Ub are bound in the Ubc13/Uev1a heterocomplex, and analogous to UbcH5, both binding sites are crucial for chain building processivity. Structural studies of Ubc13/Mms2 show that the orientation of Mms2 with respect to Ubc13 during substrate ubiquitination only allows Ub^{K63} construction (*49*). Specifically, Ile44 of ubiquitin interacts with Ile57 of Mms2, and this association is necessary for Ub^{K63} chain synthesis. Mutations of these residues to alanine completely abrogate chain synthesis of any type.

Cdc34 is a class II E2 enzyme that functions with Cul1-containing SCF type E3 Ubiquitin ligases. This particular enzyme contains a highly acidic C-terminal extension that is necessary to propagate a processive Ub^{K48} chain (*50*). Recently, it has been shown that this tail binds to a highly conserved basic cleft of the Cullin subunit within the SCF ligase (*51*). Remarkably, this interaction occurs with a dissociation constant in the nanomolar range, but at the same time somehow allows the

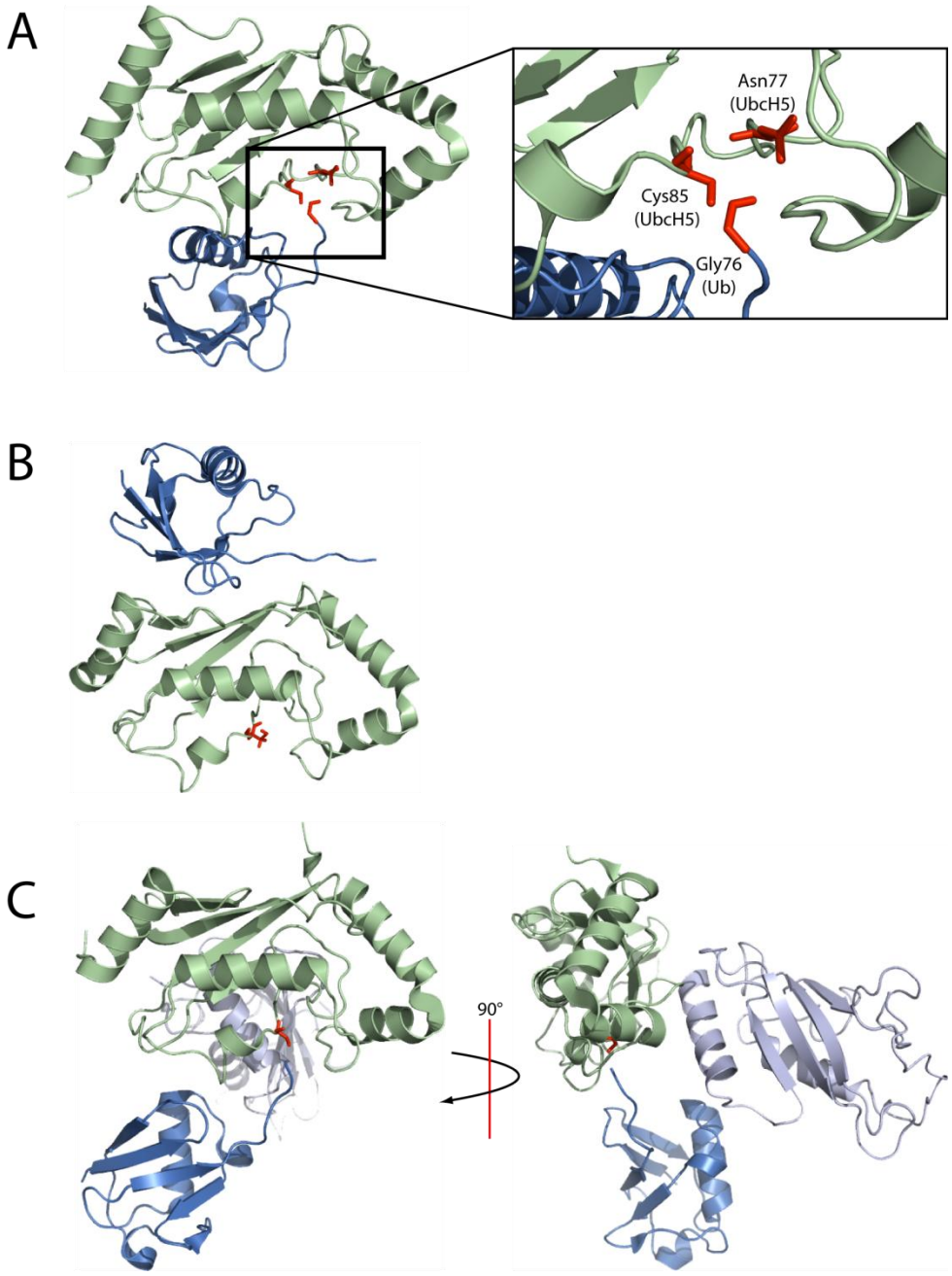

Figure 1.6. E2 conjugating enzymes. A) UbcH5 (green) bound covalently bound to ubiquitin (blue). Insert is close-up of the binding site, detailing interactions between UbcH5 active site Cys85 and Ub C-terminal Gly76. Also of interest is the position of UbcH5 Asn77 with respect to the bond. (3A33 (*42*)). B) UbcH5 non-covalently bound to ubiquitin. Active site Cys85 from UbcH5 highlighted in red to show distance between non-covalent site and active site (2FUH (*28*)). C) Ubc13(green)/Mms2(grey) heterodimer covalently bound to ubiquitin (blue). Active site Cys87 from Ubc13 highlighted in red (2GMI (*49*)).

complex to be extremely dynamic. To date, Cdc34 is the only observed *in vivo* partner of the Cul1-SCF Ub ligase, and it has been shown to play a major role in ubiquitination of β-catenin, a critical component of the Wnt-signaling pathway that controls embryogenesis.

The E3 ligating enzyme: Like E2 conjugating enzymes, there are multiple classes of E3 ligases. These include the U-box, RING, and HECT-type ligases, all named after the domain in the ligase that interacts with the E2 enzyme. In addition, E3 ligases can be simple (e.g. one protein chain recruits and interacts with both the E2 and the substrate) or complex (e.g. multiple proteins are required to facilitate the ligase function) (Figure 1.7). E3 ligases can facilitate their function either through one protein chain or as part of a protein complex. Some E3 ligases contain a substrate recruitment domain as well as a respective E2 binding domain and are completely functional as a single entity. However, in some cases, additional proteins are necessary to form a complex where one subunit will recruit the substrate and another will bind the E2. The primary example of complex E3s is the SCF family, which requires at least 4 different protein subunits for ligase function (*52*).

HECT domain containing E3 ligases only account for approximately 28 proteins out of the hundreds of mammalian E3 ligases. HECT (Homologous to E6 Carboxy-Terminus) domains were first identified in human papilloma virus (HPV) E6-Associate Protein (*40*). The HECT domain comprises approximately 350 conserved residues located in the C-termius of the host protein. Usually, substrate recruitment domains are located in the N-terminus of the protein. HECT domains are unique in that the Ub molecule is directly transferred from the E2~Ub conjugate, before

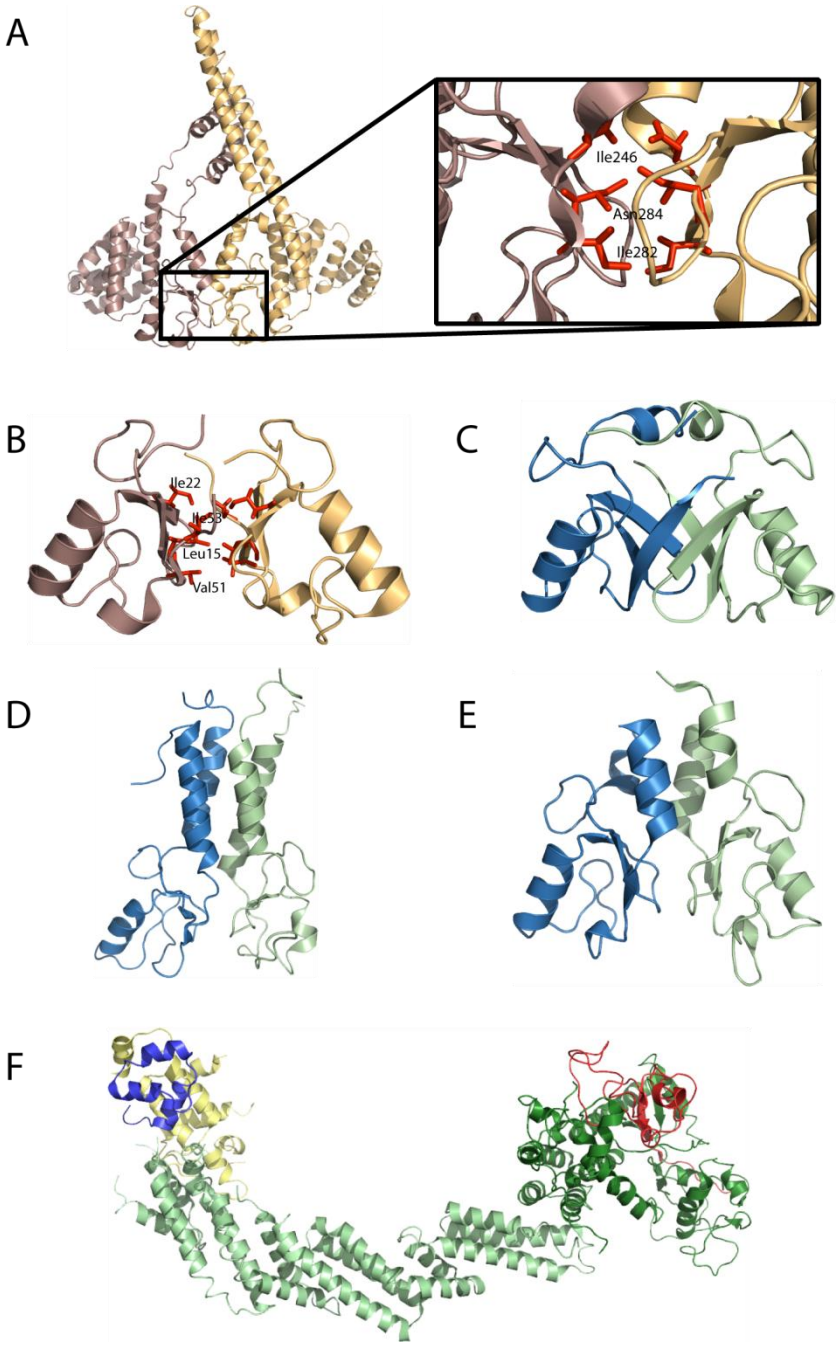

Figure 1.7. E3 ligating enzymes. A) CHIP asymmetric dimer with insert showing dimerization interface mediated by hydrophobic residues in the U-box domain. (2C2L (*53*)) B) Prp19 dimer with residues of the dimerization interface highlighted. (2BAY (*54*)) C) Heterodimer of the RING E3 ligases Mdm2 (blue) and MdmX (green). (2VJF (*55*)) D) Heterodimer of the RING E3 ligases BRCA1 (blue) and BARD1 (green) (1JM7 (*56*)). E) Heterodimer of the RING E3 ligases Ring1 (blue) and Bmi1 (green) (2CKL (*57*)). F) Complex SCF E3 ligase (1LDK (*58*)).

subsequent transfer to the substrate (*59*). Domain comparison of the various HECT proteins classifies them into 3 sub-families: Nedd4, HERC, and others. Despite their small number and unique mechanism, HECT members do not appear to exhibit specialized biological functions. They are involved in regulating a variety of both proteolytic and non-proteolytic signaling pathways (reviewed in (*59*)).

The majority of E3 ligases fall into the Really Interesting New Gene (RING) domain containing family. E3 ligases of this family act as scaffolding proteins, recruiting both the target substrate and the E2~Ub conjugate. RING E3 ligases catalyze the transfer of Ub from the E2 to the substrate in part by co-localization of the E2~Ub conjugate and substrate. However, the E3 ligase also is required for activation of the E2~Ub conjugate by an as yet unknown mechanism. RING domains comprise approximately 70 residues and are characterized by conserved zinc-chelating sites that are necessary for domain stabilization (reviewed in (*52*)).

The U-box domain is structurally similar to the RING domain, but does not contain Zn^{2+} binding sites. Instead, networks of hydrogen bonds and salt bridges provide structural stabilization to the 70 residues that make up the domain (*60*). The U-box family of E3 ligases is the smallest class of ligases with only 7 mammalian proteins identified to date. However, despite this low number, U-box containing E3 ligases have been shown to play diverse roles in RNA splicing, chaperone response, and ERAD (*61*).

Oligomerization is a critical feature for a number of different U-box and RING E3 ligases. For example, Prp19p is a U-box E3 ligase associated with the 19 complex of the spliceosome, necessary to both stabilize and activate the spliceosome machinery. Electron Microscopy (EM) studies revealed that Prp19p forms a tetramer (anti-parallel

dimer of dimers), which is required for Prp19p function (*62*). The central coiled-coil domain of Prp19 mediates anti-parallel oligomerization. The short linker between the coiled-coil domain and the U-box domain, combined with the close spatial proximity of the two U-boxes in the parallel dimer, promotes dimerization of the U-box domain. The U-box domain of the co-chaperone CHIP also oligomerizes forming a dimer through its coiled-coil domain (*53*). Remarkably, an x-ray crystal structure of CHIP revealed a unique asymmetric structure. This observation is consistent with "half-of-sites binding" stoichiometry (only one E2~Ub conjugate binds per CHIP dimer) (*63*).

Some RING domains directly mediate dimerization. For example, the RING ligase Rad18 forms a homodimer in solution (*64*). However, the most interesting aspect of RING protein oligomerization is the ability to form heterocomplexes. Several E3 ligases hetero-oligomerize via their RING domains with other members of the RING family. This is evident with BRCA1/BARD1, Ring1b/Bmi1, and Mdm2/MdmX (*65*). In each of these examples, however, only one of the RING domains is functional. As for CHIP, the purpose of "half-of-sites binding" remains unknown.

Because E3 ligases recruit and bind the target substrate, it was once thought that they were the enzymes responsible for conferring the Ub linkage specificity on the substrate. However, recent research shows that the E2 enzyme has a critical role in determining the type of ubiquitin chain formed. For example, only specific pairs of E2-E3 enzymes are able to efficiently elongate ubiquitin chains on specific substrates (*11, 12*).

The Ufd2/E4B U-box containing E3 ligase

"E4" enzyme activity

Ufd2 was first isolated in a screen by Varshavsky and colleagues to identify proteins associated with the N-end degradation pathway (*4*). Another protein identified in that screen, Ufd4, was shown to be a HECT E3 ligase that ubiquitinated a test substrate with low efficiency. These workers also showed that Ufd2 appears to add Ub moieties to the previously ubiquitinated substrate. These observations were later confirmed by *in vitro* ubiquitination assays in which reconstituted Ufd2 and Ufd4 were incubated together with E1, Ubc4 (yeast homologue of UbcH5), and ATP (*66*). It was found that while Ufd4 functioned as an E3 ligase, it did not efficiently ubiquitinate a Ub-Gal fusion substrate. Rather, Ufd2 was necessary to form the extended polyubiquitin chains required for proteasomal targeting. This work lead to the classification of Ufd2 as an "E4" ligating enzyme: one that efficiently catalyzes the elongation of pre-existing polyubiquitin chains.

Expanding on the E4 concept, another U-box containing protein, CHIP was found to bind to the RING E3 ligase Parkin and expedite the polyubiquitination of its substrate - unfolded Pael Receptor (Pael-R) (*67*). Parkin polyubiquitinates unfolded Pael-R independently, but the levels of poly-Ub species are limited, and only a portion is selected for degradation by the proteasome. However, when Parkin, CHIP, and the CHIP cochaperone, Hsp70 are all present, rapid and efficient polyubiquitination occurs (*67*). CHIP U-box mutants and deletion of the U-box showed this extended poly-Ub activity to be CHIP-dependent. Therefore, the E4 moniker was also applied to CHIP.

Shortly after the CHIP/Parkin observations were made, it was shown in *C. Elegans* that Ufd2 worked in concert with the E3 ligase CHN-1 (mammalian homologue

of CHIP) (*68*). While both Ufd2 and CHN-1 can catalyze monoubiquitination of the UNC-45 substrate on their own, efficient polyubiquitin chain elongation occurs only when both enzymes are present. It was shown that Ufd2 directly interacts with CHN-1 to promote this reaction. The N-terminal TPR and central coiled-coiled domains of CHN-1 are necessary for this enhanced ubiquitination activity. The domain of Ufd2 necessary for CHIP interaction remains unknown. The Ufd2-CHN-1 interaction is also observed for mammalian homologues, E4B and CHIP, as determined in yeast two-hybrid screening (*69*). These observations support the proposal that these ligases possess so-called "E4" activity. Notably, the mechanism behind E4 activity has yet to be elucidated. It is also interesting to note that this activity is only necessary for higher eukaryotes as a CHN-1 homologue does not exist in yeast.

Structural Organization of Ufd/E4B

The 961 residue Ufd2 is composed of three domains: an N-terminal "variable" domain, a central armadillo-like repeat core domain, and the C-terminal U-box (*43*) (Figure 1.8A). The core and U-box domains are highly conserved throughout higher order eukaryotes. Despite being essential for several protein-protein interactions, the N-terminus of Ufd2 is not conserved in mammalian species. Overall, the structure of Ufd2 is highly α-helical and spans ~146 Å in length, ~84 Å in height and ~70 Å in width

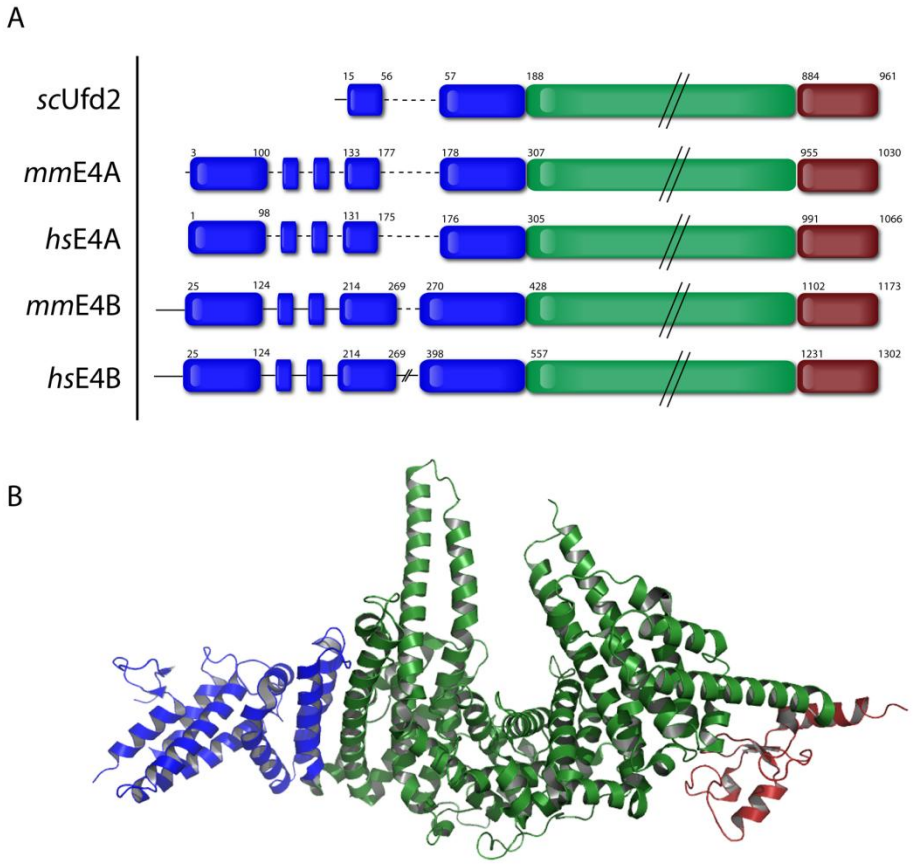

Figure 1.8. Domain organization of Ufd2/E4A/B. A) Sequence conservation of the different homologues of Ufd2/E4A/B between *Saccharomyces cerevisiae*, *Mus musculus*, and *Homo sapiens.* B) Cartoon representation of the crystal structure of *sc*Ufd2 (2QIZ (*43*)).

(Figure 1.8B). The N-terminus and core domains comprise a contiguous, elongated yet globular unit with a short linker to the U-box domain. The flexible tether presumably allows the U-box to move freely in solution with respect to the rest of the protein (*43*).

Two mammalian isoforms of Ufd2 are known, termed E4A and E4B. E4B appears to be the most physiologically relevant of the two, while little is known about

E4A. Genetic studies reveal the importance of E4B in pathways such as Wallerian degeneration, cardiac muscle development, and various instances of neurological regulation (*70-72*). The focus in this dissertation is on the E4B isoform.

The role of Ufd2 in ERAD

The N-terminus of Ufd2 is a protein-protein interaction module. One of its binding partners is the nucleotide excision repair protein, Rad23, which plays important roles in both DNA damage repair pathways as well as in proteasome shuttling (*73*). Rad23 is a multi-domain protein comprised of a Ubiquitin Like (UBL) domain and two UBA domains. As previously described, Rad23's UBA domains recognize and interact with the specific structural motif between K48 and G76 of a Ub^{K48} polyubiquitin chain (*74*). Through an unknown mechanism, Rad23 is recruited to the 26S proteasome where several subunits of the 19S cap bind to the UBL domain. The Rpn1, Rpn10, and Rpn13 subunits all appear to be involved in Rad23 and Dsk2 binding (*75, 76*). Ufd2 also interacts with Rad23's UBL domain, but the exact purpose of this association remains unknown. However, it is speculated that Ufd2 may act as a triage factor and polyubiquitinate proteins that cannot be refolded by the heat shock response (HSR). After functioning in polyubiquitin chain elongation, Ufd2 facilitates the transfer of the newly polyubiquitinated species to Rad23 for transport to the proteasome. Recently, structures of Ufd2 in complex with the UBL domains of both Rad23 and Dsk2 were reported (*73*). These structures highlight a novel UBL-interacting domain in the Ufd2 N-terminus. The UBL domain of Dsk2 contains only approximately 30% sequence identity to Rad23, but

both bind to the same face of Ufd2. The Rad23 and Dsk2 UBL domains utilize slightly different residues to interact with Ufd2.

Intriguingly, the N-terminus of Ufd2 is not conserved in its mammalian homologue, and it remains to be seen whether or not E4B interacts with the either of the mammalian Rad23 homologues, HR23A or HR23B. Another protein containing both UBL and UBA domains is Dsk2. Dsk2 functions in the proteasome shuttling pathway in the same fashion as Rad23. Like Rad23, the molecular mechanism underlying for handoff to the proteasome is unknown.

The substrate binding domains of neither Ufd2 nor E4B have been characterized. Approximately 50 residues immediately preceding the U-box domain of Ufd2 are necessary to interact with the homohexameric AAA+ ATPase Cdc48 (*77*). Ufd2 interaction with Cdc48 plays a vital role in ERAD of unfolded substrates, and thus Cdc48 may act as a scaffolding protein to recruit substrates for Ufd2 poly-ubiquitination. Cdc48 appears to function as a mediator between substrate recruiting and substrate processing at the ER-membrane level. A Ufd1/Npl4 heterocomplex ($Cdc48^{Ufd1/Npl4}$) stimulates Cdc48 binding to Ufd2, activating Ufd2 polyubiquitination chain elongation (*77*). When Ufd2 generates a Ub^{K48} chain of sufficient length (generally $\geq$ 4 Ub moieties), substrates are targeted for proteasomal degradation. Interestingly, when Cdc48 associates with Ufd2 it restricts the length of the polyubiquitin chains on the substrate. However, these shorter polyubiquitin chains do not inhibit substrate recognition

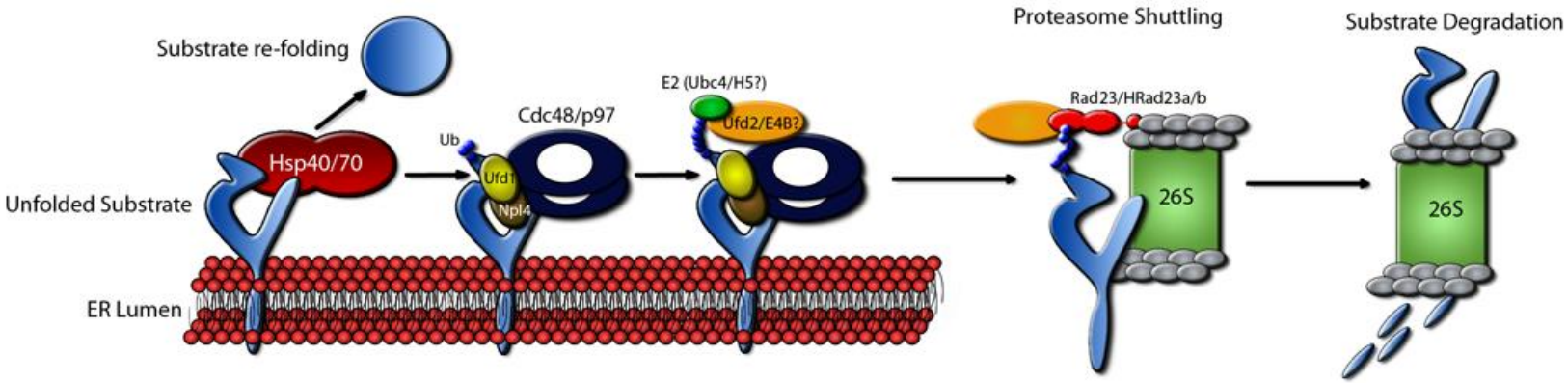

Figure 1.9. Endoplasmic Reticulum Associated Degradation Pathway

by Rad23 or Dsk2, and degradation by the proteasome is observed. Thus, the purpose of chain length restriction by Cdc48 on Ufd2 remains unknown.

Experimental Approach

Nuclear Magnetic Resonance (NMR) Spectroscopy

Much of the research accomplished in this dissertation involved NMR spectroscopy. NMR is a powerful tool for structural and biochemical analysis of biological macromolecules. NMR is commonly used for identification of inter-molecular interactions, determination and analysis of three-dimensional molecular structures, and measuring motional dynamics (*78, 79*). The NMR phenomenon arises from alignment of nuclear spins in a magnetic field. Electromagnetic pulses can then be applied to scan for resonance, and in turn measure certain physical properties of the observed atomic nuclei. The ability to simultaneously probe specific properties at the atomic level for multiple nuclei arguably makes NMR spectroscopy the most powerful method for biochemical analysis.

When performing NMR experiments on protein systems, ^{15}N and ^{13}C enrichment is used for observation of nitrogen and carbon nuclei, respectively. This is accomplished producing protein by overexpression in bacterial cells. Use of defined media with isotopically enriched nitrogen and/or carbon sources during growth provides a robust approach to produce isotopically enriched protein.

The most common NMR experiment used for this thesis work is the 2D ^{15}N-^{1}H Heteronuclear Single Quantum Correlation (HSQC) experiment. This experiment measures the spectral frequencies of ^{15}N-^{1}H amide pairs within the protein. A well-

dispersed spectrum enables the resonances of the entire amino acid backbone to be observed. Monitoring the backbone resonances is often sufficient for identification of molecular interactions as well as determining protein behavior under different experimental conditions. This chemical shift perturbation approach using ^{15}N-^{1}H HSQC experiments has been used extensively to measure the change in environment of a protein's backbone in the presence of a binding partner. In this scenario, the observed molecule is enriched with ^{15}N while the ligand of interest is not. Therefore, only information about the protein of interest is observed. The ligand is titrated in at various concentrations, and the perturbations in chemical shifts can be measured. Unlike other binding measurements such as fluorescence or isothermal titration calorimetry, information obtained from NMR chemical shift perturbation experiments can be correlated directly to specific atoms in the protein. This provides a powerful method to map the binding site on the structure of the protein (*80*). In the case of protein-protein interactions, reciprocal experiments may be able to be collected in which the ^{15}N-enrichment is switched from one protein to the other, thus allowing for binding site mapping on both species of the interaction, and generation of high quality models of the complex (see Chapter II).

In order to correlate information about which resonance peak corresponds to which residue of the protein, a series of experiments to assign the resonances are necessary. A standard set of five triple resonance experiments generally allow for the direct assignment of the backbone resonances of a protein: HNCACB, CBCACONH, HNCA, HNCO, and HNCOCA (*78*). These experiments require isotopic enrichment in both ^{15}N and ^{13}C.

For structural determination of a protein, complete side chain resonances must be identified and assigned as well. In regards to the structure determination in this thesis, this was accomplished through the double and triple resonance multi-nuclear experiments HBHACONH, (H)CC(CO)NH, and H(CC)(CO)NH, plus a homonuclear ^{1}H-DQF-COSY experiment to assign aromatic side-chain resonances. The information from the combined backbone and side chain assignments was then applied to 2D and 3D Nuclear Overhauser Effect (NOE) experiments. These NOE-experiments offer information about important dipolar (through space) couplings, which characterize the spatial relationships between the H atoms.

With NOE data in hand, structure calculations were executed by use of the automated computational software CYANA. Through a series of algorithms, CYANA calculates the 3D orientation of the H atoms in the protein (*81*). To ensure correct chemical geometry, additional refinement was performed using restrained molecular dynamics and simulations in the AMBER force field (*82*). Coupled together, these methods allow for generation of highly useful, atomic resolution protein structures.

Mass Spectrometry

Recent advances in Mass Spectrometry technology have enabled the study of multi-protein complexes. These studies are most commonly undertaken using Electrospray Ionization Mass Spectrometry (ESI-MS), this ionization technique can be adjusted to allow non-covalent protein interactions to remain intact (*83*). When non-covalent protein-protein interactions are left intact, ESI-MS is useful in verification of the stiochiometry of higher order protein complexes, including oligomerization. Despite the

transition from solution to the gas phase, these non-covalent complex species can be observed providing important biochemical data.

Analytical Ultracentrifugation

Analytical Ultracentrifugation has emerged as a powerful tool for the biochemical analysis of macromolecular molecules and molecular assemblies. AUC is most often used to study the hydrodynamic properties of a macromolecule in solution. Two distinct types of AUC experiments are routinely used: sedimentation velocity (SV) and sedimentation equilibrium (SE) (*84*). In the context of this thesis, SV experiments were applied to study the oligomeric state of the U-box domain of E4B.

SV experiments apply high centrifugal speeds to the sample, and optical dectors are used to monitor the redistribution of the sample from the top of the meniscus of the sample cell to the bottom of the sample cell over time (*85, 86*). The rate of movement of the molecule will vary based on its size and shape, and analysis of this data leads to determination of the overall oligomeric state or stoichiometric ratios of the molecule or molecular assembly under study. Determining oligomeric states as well as stoichiometric ratios in the context of protein assemblies provides valuable information on the biochemical basis for function.

Running an SV experiment requires a two-chambered sample cell. One chamber, or sector, contains the molecule of interest in solution, while the other sector contains the reference solvent. The concentration distribution of the sample as it moves through solution under centrifuge is monitored over time and fit by the software package, SedFit (*84, 86*). A sedimentation coefficient is determined from this concentration distribution

using the equation $s = v / (\omega 2r) = Mb / f$, where v is the volume of the particle, ω the rotor speed, r the distance from the center of the rotor, Mb the buoyant molar mass and f the frictional coefficient (*84*, *86*).

SV AUC experiments are used to determine the sedimentation coefficient for the sample of interest, whether it is a single protein species or a multi-protein complex. Two important parameters that can be determined from SV experiments are the frictional ratio and Stokes radius (f/f_o and R_s, respectively). The frictional ratio provides information about the shape of the molecule: globular proteins have low frictional ratios while those proteins with overall elongated structures have high frictional ratios. The Stokes radius is the radius of a hard sphere that diffuses at the same rate as that of the molecule of interest.

Research Summary

The research presented in this dissertation attempts to enhance our understanding of the cellular roles of E4B through both structural and functional analysis. The initial objective of E4B study was to identify its role in CHIP mediated polyubiquitination as well as to characterize interactions and analyze functions of its U-box domain with E2 enzymes. To assess CHIP-E4B binding, recombinant production of E4B was necessary. However, attempts to produce the full length, recombinant E4B from *Mus musculus* in *Eschericia coli* did not pan out, and thus shorter, stable U-box containing domains of this protein were sought to focus on E2 interactions.

To fully characterize E4BU-E2 enzyme interactions, structural determination of E4BU was necessary. After several attempts at E4BU crystallization, the structure was

determined via NMR spectroscopy and the interaction with the cognate E2 enzyme, UbcH5c was investigated using a combination of ^{15}N-^{1}H NMR chemical shift perturbation experiments and with the computational docking program, HADDOCK (*87*).

Oligomeric states of several U-box type E3 ligases play vital roles in their function (*53, 54*). Therefore, the oligmeric state of E4BU was assessed. Using a combination of Size Exclusion Chromatography coupled to Multi-Angle Light Scattering (SEC-MALS), Analytical Ultracentrifugation (AUC), and Electrospray Ionization Mass Spectrometry (ESI-MS), it was determined that E4BU is monomeric in solution. Based on this surprising information, the ligase function of E4BU was verified with an *in vitro* autoubiquitination assay. From these results, it was shown that monomeric E4BU is a fully functional E3 ligase. This work is discussed in depth in Chapter II.

The third chapter of this thesis describes research focused on understanding the dynamics of the E4BU-E2~Ubiquitin complex and indentifying how E4BU activates Ub release from the covalently bound E2. This portion of the research was completed as part of an extraordinary collaboration with Rachel Klevit's laboratory at the University of Washington. Our focus was on assessing the molecular motion of E2~Ub and how it is affected by E3 interactions and is discussed in Chapter III .

Using a stable UbcH5~Ub conjugate complex, chemical shift perturbation experiments were performed with E4BU. In addition, Ub resonances on the E2~Ub conjugate were monitored for changes induced by the binding of E4BU. From this information, several key residues were identified that appear to play a significant role in Ub activation by E4BU. Mutations were made on these residues and *in vitro* ubiquitination assays were performed to validate the presence of an allosteric network

necessary for Ub release from UbcH5 upon E4BU binding. Residual dipolar coupling NMR experiments were used to more accurately assess the flexibility of Ub in the presence and absence of E4BU. Interestingly, no change in Ub flexibility appears to be observed in the presence of E4BU, but more experiments must be completed to generate unambiguous conclusions.

Chapter IV summarizes the work completed on the U-box domain of E4B, discusses the impact and implications of these results, and assesses future studies necessary to further elucidate both structural and functional questions about E4B. Therapeutic implications of this research are also described. Taken together, this work provides critical insight into the structure and function of the monomeric U-box domain from E4B as well as its role in activating ubiquitin release from UbcH5.

CHAPTER II

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE MONOMERIC U-BOX DOMAIN FROM E4B[1]

Introduction

Ubiquitination is a critical post-translational modification shown to target proteins in a variety of cellular signaling contexts, including transcriptional activation, endocytosis, DNA repair, and proteasomal degradation (*1*, *88-90*). The molecular mechanism underlying the transfer of ubiquitin (Ub) to a substrate consists of three key enzymatic steps. First, ubiquitin itself is adenylated at its C-terminal glycine residue by an activating enzyme (E1). Second, the adenylated Ub forms a covalent linkage to a conjugating enzyme (E2). Finally, a ligating enzyme (E3) recruits both the Ub-charged E2 species and the target substrate protein. Co-localization of the charged E2 and the substrate catalyzes the transfer of Ub to an ε-amine group on a lysine residue of the target protein (*89*).

There are three classes of E3 enzymes: HECT, RING, and U-box, which are distinguished on the basis of their E2-recruiting domains. Whereas HECT and RING domain E3 ligases have been studied extensively, much less is known about the U-box proteins. The U-box and RING classes of E3 ligases act as scaffolding molecules that recruit and co-localize both a Ub-charged E2 and the substrate concomitantly. The

[1]The bulk of this chapter was published in Nordquist et al., (2010) Structural and functional characterization of the monomeric U-box domain from E4B. *Biochemistry*. Jan 19;49(2):347-55.

recruitment of substrate in these proteins involves protein interaction modules such as a WD-40 repeat, TPR and armadillo repeat domains (*62, 91, 92*).

In addition to a common organization, the architecture of U-box and RING domains are similar. Both contain a central alpha helix flanked by two surface exposed loops arranged in a cross-brace formation. The structure of RING domains is built around two zinc binding sites that are critical to its stability. In contrast, U-boxes do not bind zinc, but have evolved instead networks of hydrogen bonds and salt bridges in corresponding locations in the structure (*60*). Other similarities between these two domains include an antiparallel β-sheet type arrangement involving the first surface exposed loop and the central alpha helix. The β-sheet is stabilized by highly conserved hydrophobic residues responsible for the core packing and stability of the molecule. It has been reported that the L1 – α1 – L2 motif of U-box and RING domains are critical elements for E2 interactions (*12, 53, 93*). Most U-box and RING domain structures also contain an elongated C-terminal helix. The physical basis and physiological rationale for evolving distinct U-box and RING E3 ligases are not yet known.

U-box proteins constitute a unique class of E3 ubiquitin ligases that appear to function in association with heat shock proteins in protein folding, refolding or complex formation. Ufd2 was first recognized as a U-box containing protein by Koegl and colleagues who designated the protein as an E4 ubiquitin ligase because it was shown to catalyze the elongation of previously ubiquitinated substrates (*66*). Ufd2 plays a role in the Cdc48 chaperone pathway. Ufd2 ubiquitinates Cdc48 clients targeting them for proteasome mediated degradation. Interactions have been reported between Ufd2 and Cdc48, and this interaction is stimulated by Cdc48 binding with its cofactors Npl4 and

Ufd1 (*77*). Additionally, the proteasome shuttling protein, Rad23 binds Ufd2 through its UBL domain, implicating Ufd2 in proteasome shuttling (*94*). It has yet to be determined if these interactions are conserved for the mammalian homolog E4B.

In addition to the ability to interact with E2 enzymes, it was also found that Ufd2 could interact with other E3 proteins, including another member of the Ufd family, Ufd4. Ufd4 is a HECT E3 ligase that was shown to only monoubiquitinate the UNC-45 substrate. However, together with Ufd2, Ufd4 stimulates rapid and efficient substrate poly-ubiquitination. Ufd2, and its mammalian homologue, E4B, have also been reported to interact with another U-box containing protein, CHIP (*68, 69*). It appears that this interaction also promotes an increased rate of substrate poly-ubiquitination.

An X-ray crystal structure of Ufd2 has been determined (*43*). The structure reveals three domains: an N-terminal variable domain, a conserved central armadillo-like repeat, and a C-terminal U-box domain. The U-box and armadillo-like repeat domains likely serve for recruiting E2 and substrate, respectively. The N-terminal domain probably serves in regulation of function. Sequence alignment of the yeast Ufd2 sequence with the mammalian E4B homologue shows a very high degree of conservation within the armadillo-like repeat domain and U-box region; however, the N-terminal region is not conserved, and in fact, the yeast protein is 212 residues shorter than the mammalian protein. Differences in the putative regulatory domain are consistent with the more complex regulatory mechanisms in higher eukaryotes.

Our group has undertaken a comprehensive investigation of the structure and function of the U-box domain of mammalian Ufd2, E4B. We began by examining the state of oligomerization of the E4B U-box and show that it exists primarily as a monomer

in solution under a range of concentrations. The structure was determined in solution by NMR and chemical shift perturbation experiments were used to map the interaction with the E2 enzyme UbcH5c. This data was then used to generate a model of the complex. We also show the U-box alone is fully active in an *in vitro* autoubiquitination assay.

Results

To produce a stable E4B U-box domain, three C-terminal bacterial expression constructs were designed based on analysis of predicted secondary structure and comparison to the x-ray crystal structure of Ufd2: Ile1072-His1173, Val1082-His1173, and Ala1092-His1173. To determine the optimal construct for structural and biophysical analysis, all three constructs were expressed and purified with uniform ^{15}N enrichment then characterized by ^{15}N-^{1}H-HSQC NMR. The constructs exhibited similar spectra with the same group of well-dispersed signals, but with a progressively larger number of signals in the central region of the spectrum in proportion to the length of the construct (Figure 2.1). This observation indicated all constructs contained a common globular domain but that the region Ile1072-Ala1092 lacks stable structure. Moreover, the two longer constructs degraded over time, as evident from the appearance of heterogeneity in HSQC spectra acquired one week after the samples were first prepared. Consequently, all further NMR experiments were performed on the shortest construct, Ala1092-His1173, which will be referred to as E4BU.

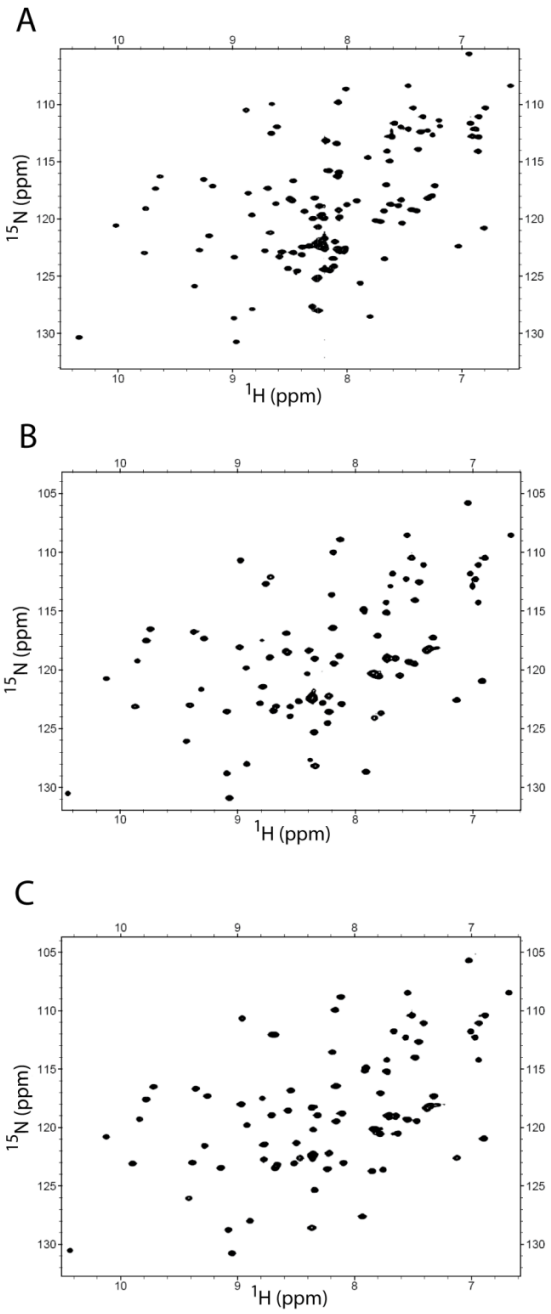

Figure 2.1. ^{15}N-^{1}H HSQC NMR spectra of three E4B U-box domain constructs: (A) E4B(1072-1173), (B) E4B(1082-1173), and (C) E4B(1092-1173). As the construct increases in length at the N-terminus, overlap in the 8.0-8.5 ^{1}H p.p.m. region of the spectra significantly increases but no additional dispersed peaks outside this region are observed, indicating that residues 1072-1092 are unstructured.

E4BU is a monomer in solution

All U-box E3 ligases characterized to date are single chain proteins that form homooligomers, except for the yeast homolog of E4B, Ufd2 (*43, 53, 54, 95, 96*). During the purification of E4BU, the protein was found to elute as a dimer with an apparent molecular weight of approximately 21 kDa according to size-exclusion chromatography (SEC) experiments. However, since molecular weight is notoriously difficult to quantify by SEC, additional experiments were performed to directly determine the oligomerization state of E4BU. The approaches included SEC-MALS, which indicated an ~10 kDa species in the primary peak (data not shown) and native mass spectrometry using electrospray ionization, which revealed a 10:1 ratio of monomer to dimer (Figure 2.2A). The predominance of monomer over dimer was further confirmed by analytical ultracentrifugation sedimentation velocity experiments, which also showed a 10:1 ratio of monomer over dimer (Figure 2.2B). Concentration-dependent NMR chemical shifts provide an additional sensitive indicator of self-association. ^{15}N-^{1}H HSQC spectra were acquired for 0.10 mM and 1.0 mM samples of ^{15}N-E4BU and no difference was observed between the two spectra (Figure 2.3), indicating no significant population of dimer or higher aggregates even at the higher protein concentration.

NMR Structure Determination of E4BU

A high resolution solution structure of E4BU was determined by multidimensional heteronuclear NMR spectroscopy using isotopically enriched protein samples. A total of 1565 NOE-derived distance constraints and 92 dihedral angle

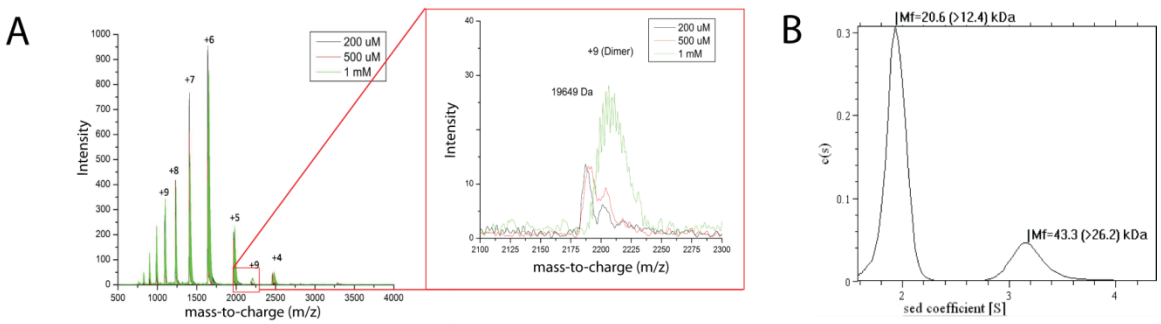

Figure 2.2. Characterization of self-association of E4BU. (A) Electrospray ionization mass spectrometry of E4BU. All ions detected correspond to monomeric species, except for a single additional (+9) charge state corresponding to a dimer (insert). All mass-to-charge (m/z) species correlate to a monomeric protein. (B) Analytical ultracentrifugation of a 100 µM solution of E4BU. The results of the sedimentation velocity experiment yielded ~87% of monomer and < 9% of dimer. The remaining protein appears to form higher order aggregates.

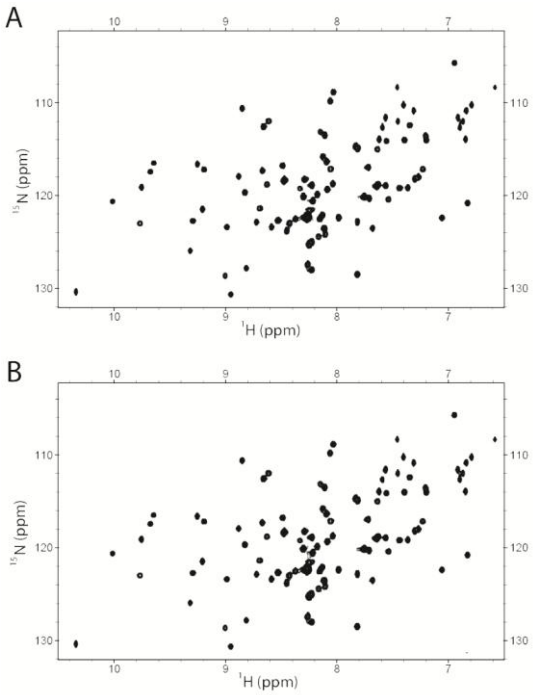

Figure 2.3. ^{15}N-^{1}H HSQC of E4BU collected at 100 µM (A) and 1 mM (B).

constraints were used as input for structure calculations. The input constraint list was refined by using iterative cycles of CYANA. The final ensemble of 50 CYANA starting structures was further refined by restrained molecular dynamics in AMBER. The final representative ensemble consists of the 20 conformers selected on the basis of lowest restraint violation energy (Figure 2.4). The structured region of the ensemble (residues 1096-1170) has a final rmsd versus the mean of 0.45 Å for the backbone and 1.13 Å for all atoms. The high quality of the structure is reflected in the results from PROCHECK analysis. For example, 99% of all backbone torsion angles occupy the most favored and favored regions of the Ramachandran plot. Further details, and structural statistics are reported in Table 2.1.

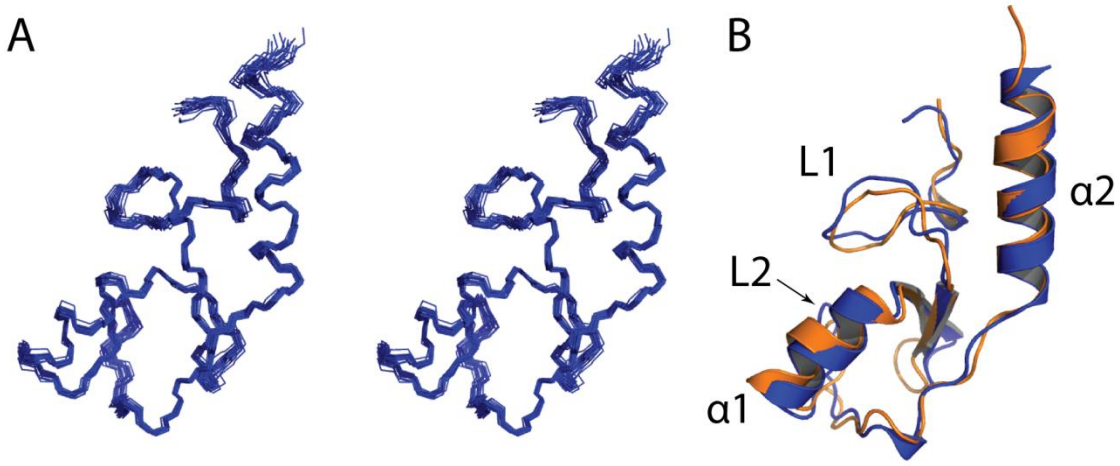

Figure 2.4. Solution NMR structure of E4BU. (A) Stereo view of the final ensemble of 20 conformers representing the structure. (B) Comparison of E4BU with the U-box domain of Ufd2 extracted from the crystal of the intact protein. Best-fit superposition of the two structures gives an rmsd over all backbone atoms of 1.15 Å.

Table 1: Structural Statistics for E4BU		
Restraints for Calculation		
Total NOE Restraints		1565
Short Range		780
Medium Range		370
Long Range		415
Dihedral Angle Restraint		92
Constraint Violations, mean ± S.D.		
Distance Violations		
0.1Å < d < 0.2 Å		0.20±.41
d > 0.2Å		0
Average maximum distance violations (Å)		0.08±.03
Torsion angle violations > 5.0°		0
Average maximum torsion angle violations (degree)		0
AMBER energies, mean ± S.D. (kcal mol^{-1})		
Restraint		1.1±0.1
van der Waals		-651.2±8.7
Total molecular		-3222.4±7.4
Precision, root mean square deviation from the mean (Å)		
	Ordered	All residues
Backbone	0.5±0.1	1.1±0.1
All heavy atoms	1.5±0.3	1.9±0.3
Ramachandran statistics (%)		
Most favored		87.8
Additionally allowed		10.5
Generously allowed		0.9
Disallowed		0.8

Table 2.1. Structural statistics for the solution NMR structure of the U-box domain of E4B.

The structure of E4BU consists of the ββαβα-fold typical of U-box and RING domains (Figure 2.4). The central α-helix is flanked by two prominent surface exposed loop regions, spanning residues Arg1104 to Thr1110 (L1) and Thr1138 to Gln1144 (L2). The characteristic network of hydrogen bonds within each loop stabilizes the overall structure. Critical polar contacts exist throughout L1. For example, the side chain of Arg1104 hydrogen bonds to the backbone carboxyl group of Asp1109. The side chain carboxyl group of Asp1105 is distinctive, forming a hydrogen bond to the backbone

amide of Leu1107 in 16 out of 20 structures of the ensemble. Of these 16, 10 show a bifurcated hydrogen bond to the amides of both Leu1107 and Met1108. L2 contributes to a network of interactions similar to that involving L1. Important stabilizing hydrogen bonds include two from the carboxyl side chain of Asp1139: one oxygen atom hydrogen bonds to the backbone amides of Phe1141 and Asn1142 and the other interacts with Gln1144. Both of these interactions are observed in all 20 structures of the ensemble. Asp1139 appears to play a central role in the stabilization network as its backbone amide also is involved in a hydrogen bond with the side chain of Arg1143. This stabilizing bond is also seen in all 20 structures of the ensemble.

Comparison of E4BU to other U-box and RING domains

Overall, the structure of E4BU has a high degree of structural similarity to other U-box and RING domains. An overlay of the backbone atoms of E4BU with the U-box domain of Ufd2 extracted from the crystal structure of the full protein (2QIZ) is shown in Figure 2.4B. The two structures superimpose with a backbone rmsd of 1.2 Å. Overlays with CHIP U-box, Prp19 U-box, and BRCA1 RING domains give rmsds of 1.3 Å, 1.4 Å, and 1.3 Å, respectively.

In order to obtain insights into the differences between monomeric E4BU and the dimeric U-box and RING domains, the sequence was analyzed by multiple sequence alignment. Two residues in the E4B U-box domain sequence were found to diverge significantly from the other U-boxes: Arg1117 and Glu1152, which are hydrophobic residues in all other U-boxes (Figure 2.5A, black arrows). In *sc*Prp19 and *mm*CHIP, the residues corresponding to Arg1117 (Leu15 and Ile246, respectively) make crucial

hydrophobic contacts across the dimer interface (*53*, *54*). The critical importance of this residue was shown for Prp19; mutation of Leu15 abrogated dimerization and abolished cell viability (*54*). The residues corresponding to Glu1152 in E4B are also hydrophobic in Prp19 and CHIP, and have been previously described as critical elements in the hydrophobic core and across the dimer interface. Analysis of the electrostatic surface potential of U-box domains reveals that E4BU is significantly more acidic than the Prp19 and CHIP U-boxes, especially in the region corresponding to the dimer interface (Figure 2.5B-D). In particular, E4BU Glu1152 is responsible for a highly negative patch at the core of what would be the dimer interface. Hence, electrostatic repulsion at the dimer interface due to substitution of hydrophobic residues with charged side chains appears to significantly destabilize the dimeric state of E4BU.

E4BU – UbcH5 interaction

U-box and RING domains serve as E2-recruiting modules in E3 ubiquitin ligases. To determine how E4BU interacts with E2 enzymes, the binding interface between the U-box domain of E4B and a cognate E2 enzyme UbcH5c was analyzed using NMR chemical shift perturbation experiments. UbcH5c was used for these experiments since the backbone NMR assignments and the crystal structure were already available (P. Brzovic and R. Klevit, unpublished; PDB ID 1X23). Titration of UbcH5c into ^{15}N-enriched E4BU was monitored in ^{15}N-^{1}H HSQC spectra acquired at molar ratios of 1:0.125, 1:0.25, 1:0.5, 1:1, and 1:2. Distinct chemical shift perturbations were observed for specific peaks even at the lowest titration point (Figure 2.6A). Over the course of the

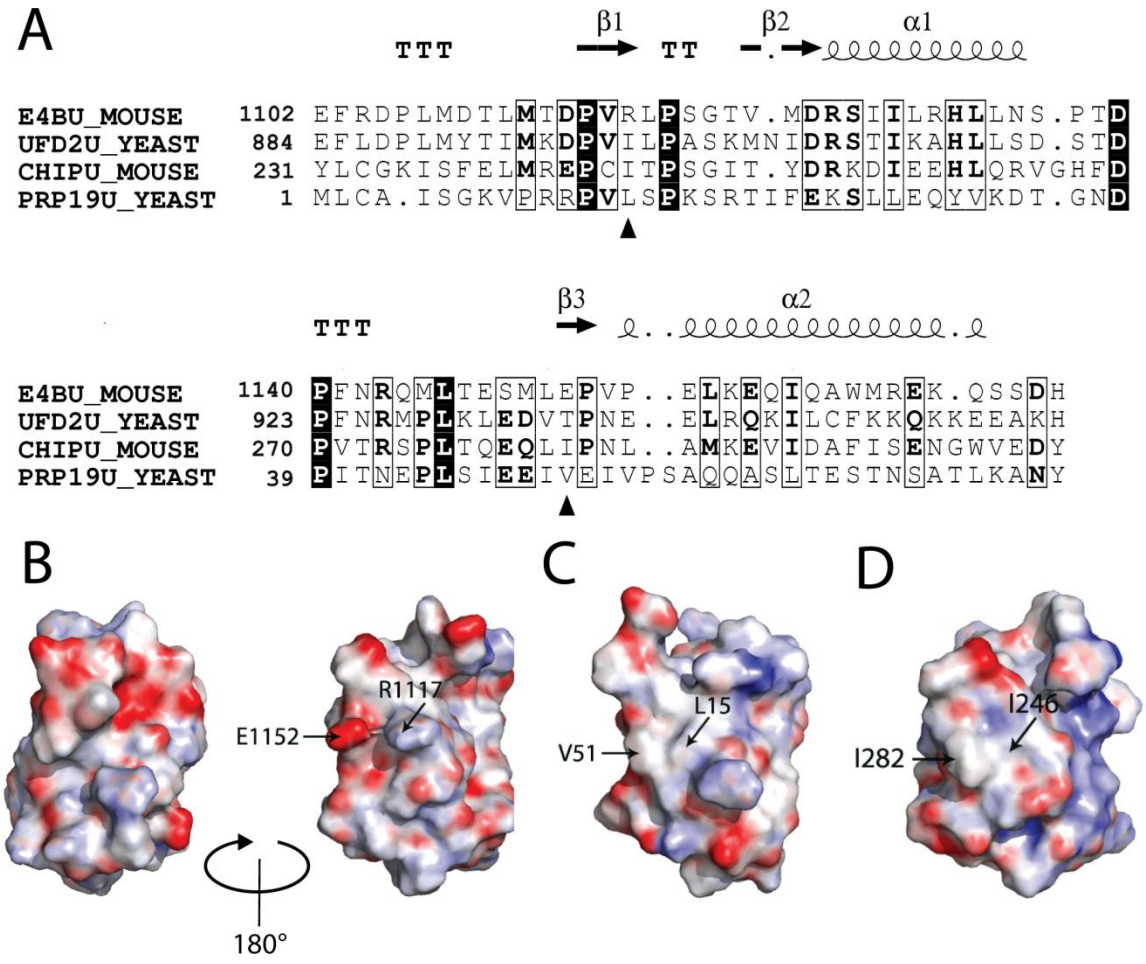

Figure 2.5. E4BU has a more negative surface potential than other U-box domains. (A) Multiple sequence alignment of E4BU with Ufd2, CHIP, and Prp19. Alignments made using the CLUSTALW server at EBI (*97*). The resulting alignment file was then processed through ESPript (*98*). (B) Electrostatic surface representation of the U-box dimer interface of E4BU. Blue coloring corresponds to positive charge and red corresponds to negative charge. (C) Electrostatic surface representation of the Prp19 U-box dimer interface. (D) Electrostatic surface representation of the CHIP U-box dimer interface. The electrostatic fields were calculated using the Adaptive Poisson-Boltzmann Solver algorithm for electrostatics (*99*).

titration a significant number of peaks broadened beyond detection as a consequence of formation of the 27 kDa complex, and/or effects associated with intermediate exchange on the NMR timescale. Overall, the resonances of 10 residues were significantly perturbed (mean + 1 standard deviation): Leu1107, Met1108, Asp1109, Leu1118, Thr1122, Ile1129, Leu1130, Arg1131, Asn1135, and Asn1142 (Figure 2.7). Using the backbone assignments for E4BU, the perturbed residues were then mapped onto the structure (Figure 2.6B). It has previously been reported that Ile235 in CHIP and the corresponding residue in BRCA1, Ile26, are crucial for interaction with E2 (*93*); the corresponding residue in E4BU, Leu1107, has one of the largest chemical shift perturbations, upon binding of UbcH5c. Overall, the perturbed residues are found to be located in and around the two surface exposed loops, including residues in the central α-helix, α1, consistent with the location of E2 binding sites in other RING and U-box domains.

Reciprocal titrations were also performed using unlabeled E4BU and ^{15}N-UbcH5c with spectra acquired at ratios of 1:0.063, 1:0.125, 1:0.25, 1:0.5, 1:1, and 1:2. In all, 14 residues were significantly perturbed including: Arg5, Ile6, Asn7, Lys8, and Ser11 in α1; Thr58, Asp59, and Phe62 in L4; Ser94, Ala96, Leu97, and Ile99, in L7; and Ile88 and Leu103 in helical regions that flank L7 (Figure 2.6C). When mapped out on the structure of UbcH5c, these residues form a contiguous surface on one face of the protein (Figure 2.6D). These results are consistent with previous reports of E2 interactions with CHIP (13, 14, 32).

In order to gain a more thorough understanding of the interface between E4BU and UbcH5c, the chemical shift perturbation data was used with HADDOCK to generate

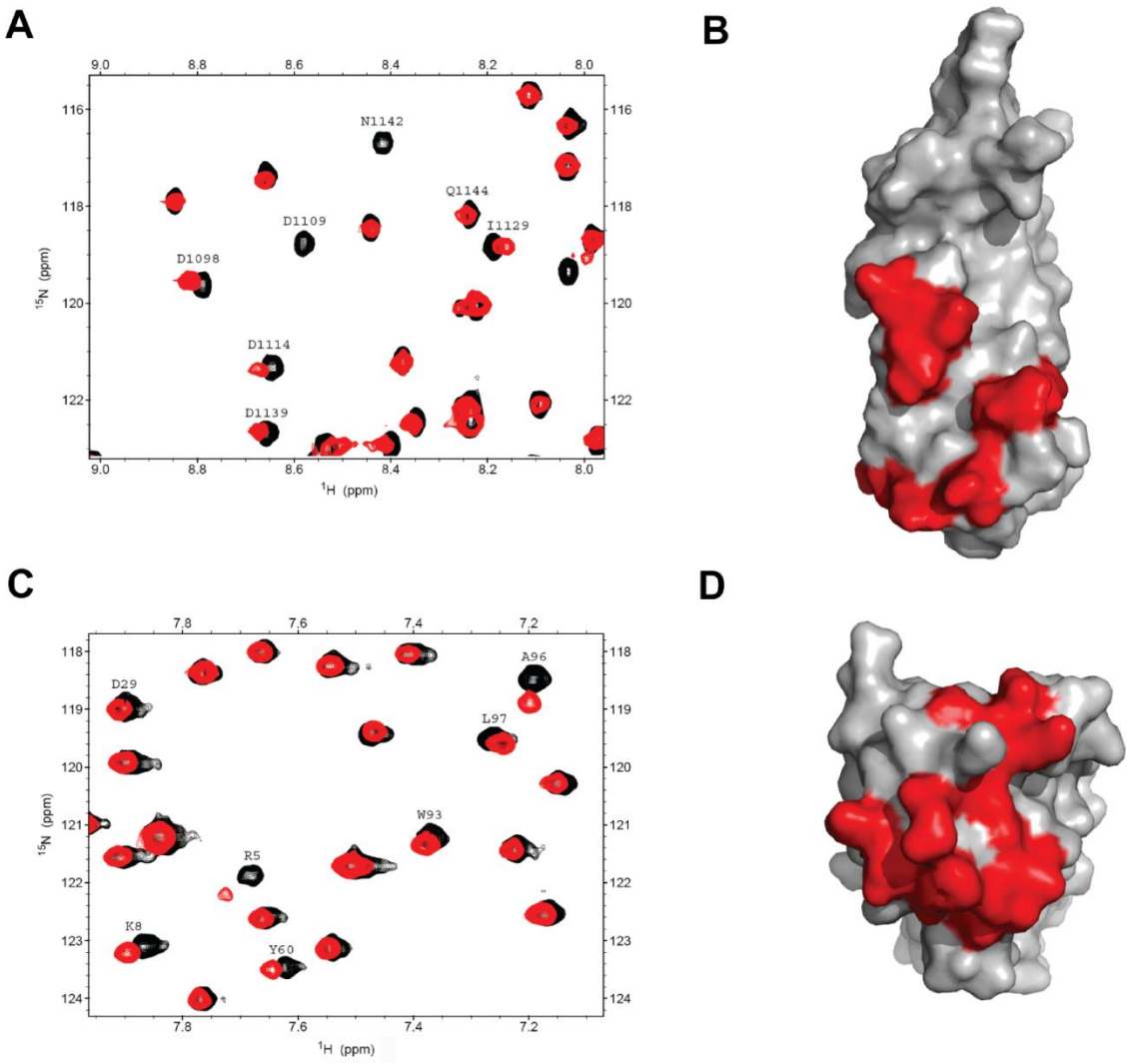

Figure 2.6. NMR chemical shift perturbations indicating interaction of E4BU and UbcH5c. (A) ^{15}N-^{1}H HSQC spectra of E4BU obtained in the absence (black) and presence (red) of 0.125 molar equivalent of UbcH5c. (B) Surface representation of E4B with residues that have significant chemical shift perturbations highlighted in red. (C) ^{15}N-^{1}H HSQC spectra of UbcH5c in the absence (black) and presence (red) of 0.125 molar equivalent of E4BU. (D) Surface representation of UbcH5c with residues that have significant chemical shift perturbations highlighted in red.

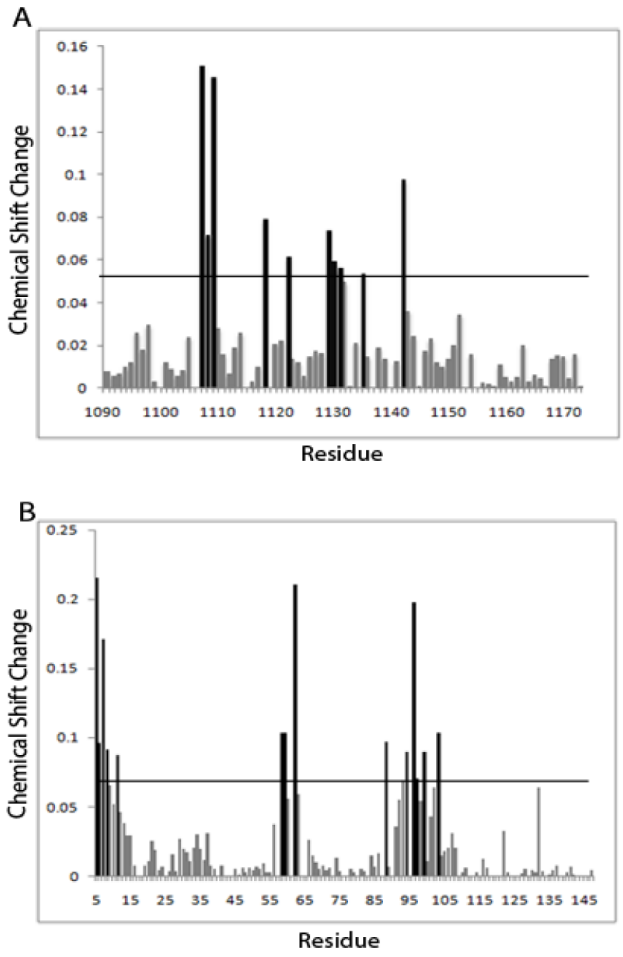

Figure 2.7. Plots of chemical shift perturbations showing interaction of E4BU with UbcH5c. (A) Unlabelled UbcH5c titrated into ^{15}N-enriched E4BU. Perturbations greater than the mean plus one standard deviation are represented by the black bars. (B) Unlabelled E4BU titrated into ^{15}N-enriched UbcH5c. Perturbations greater than the mean plus one standard deviation are represented by the black bars.

a molecular model of the complex. It has been previously demonstrated that using crystal structures in HADDOCK docking may not offer sufficient initial flexibility for proper docking (*100*). A molecular dynamics simulation was therefore carried out to generate an ensemble of 20 UbcH5c structures. Active residues for the HADDOCK ambiguous

interaction restraints (AIRs) were assigned to residues exhibiting significant chemical shift perturbations (mean+1SD) or broadening and backbone and side chain regions that occupied greater than or equal to 50% solvent accessibility (as calculated by NACCESS). Overall, 7 active and 4 passive residues were assigned for the E4BU ensemble, while 6 active and 7 passive residues were assigned for UbcH5c. Cluster analysis of the 200 models generated by HADDOCK generated 9 groups of 4 molecules or more. However, of these, two similar clusters stood out as being highly populated (Cluster 1- 54 conformers, Cluster 2- 16 conformers) and having the most favorable HADDOCK energies. The 20 lowest energy conformers of the most highly populated cluster were selected to represent the model of the complex. This model contains interactions between L1 of E4BU and α1 of UbcH5c, α1 of E4BU and L4 of UbcH5c, and L2 of E4BU and L7 of UbcH5c (Figure 2.8). Remarkably, this model aligns extremely well with the co-crystal structure of UbcH5a and the CHIP dimer, superimposing to an rmsd of only 0.81Å. Thus, this experimentally–guided model suggests monomeric E4BU binds E2 conjugating enzymes in the same manner as U-box and RING domain dimers.

E4BU function

We next asked whether or not the monomeric E4BU is functional, as interaction with an E2 is not sufficient to prove activity. Most E3 ligases are able to modify themselves through autoubiquitination, where the ligase itself serves as the substrate, a process that serves a useful role in auto-regulation through protein turnover. Because the E4B(1092-1173) construct encodes just the U-box domain, the longer 1072-1173

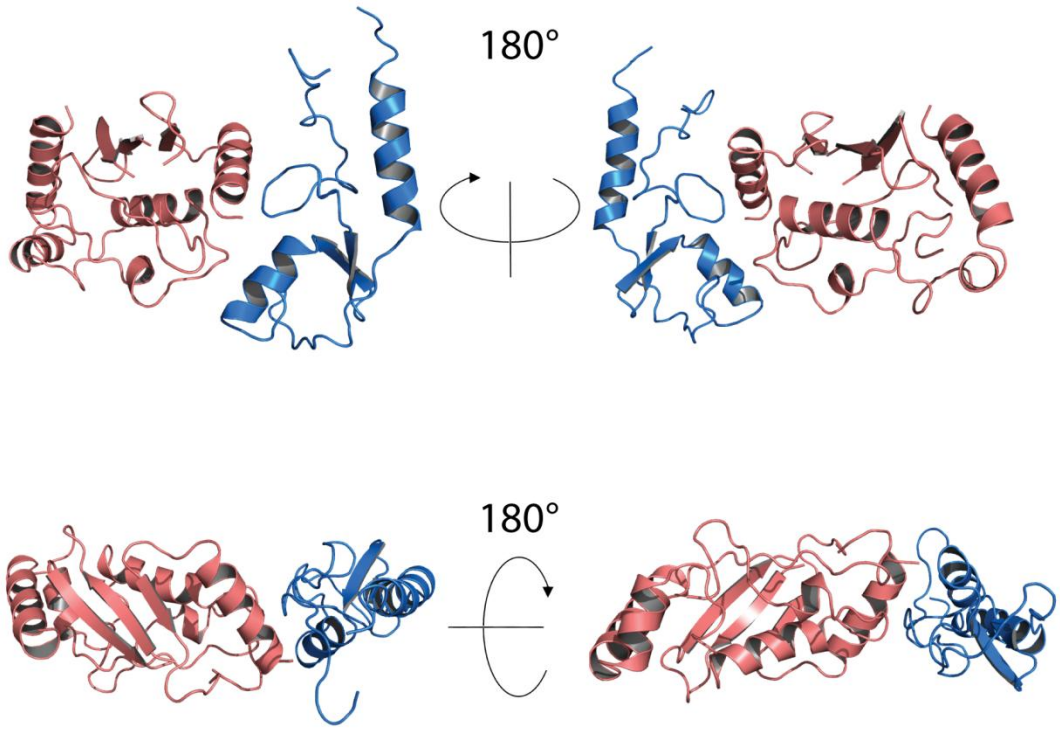

Figure 2.8. Model of the complex of E4BU and UbcH5c based on NMR chemical shift perturbations. Ribbon diagram of the mean structure from the 20 lowest energy conformers of the highest populated cluster from HADDOCK calculations with E4BU (blue) and UbcH5c (salmon).

construct was used for the autoubiquitination experiments. The additional 20 residues outside of the U-box domain, four of which are lysines, ensure there is an adequate substrate for ubiquitination. The E4BU *in vitro* autoubiquitination assay was developed based on our protocol [Y.N.D. and W.J.C., unpublished results] using recombinant bacterially expressed and purified proteins Uba1 (E1), UbcH5 (E2), E4B(1072-1173) (E3/substrate), and Ub. Analysis of the autoubiquitination reaction by SDS-PAGE

reveals time-dependent formation of higher molecular weight E4B species, showing that the U-box has robust autoubiquitination activity (Figure 2.9).

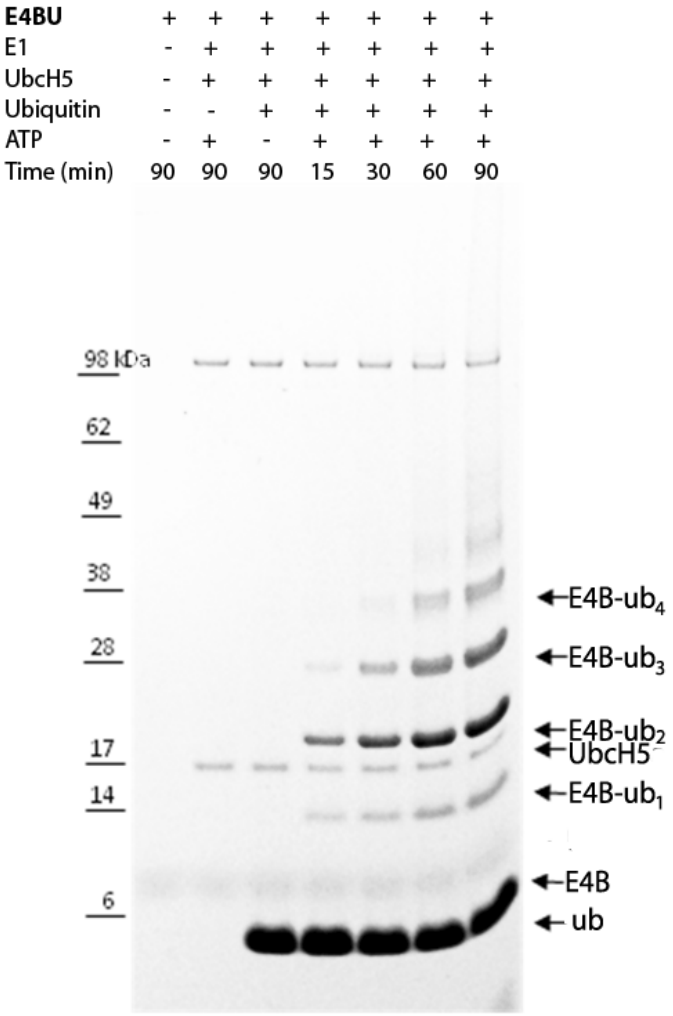

Figure 2.9. Ubiquitination activity of the monmeric E4B U-box in an *in vitro* ubiquitination assay. Time course of the assay using E3B(1072-1173) and the UbcH5c E2 conjugating enzyme.

Discussion

We have shown that unlike other U-box proteins characterized to date, the U-box domain of E4B exists and functions as a monomer in solution (*53, 54, 56*). It could be hypothesized that the availability of two or more E2 binding modules in oligomeric E3 ligases would increase the processivity of substrate ubiquitination through faster loading and unloading of E2 enzymes on the E3. However, this does not appear to be the case, as most oligomeric U-box and RING E3 ligases appear to possess only one active binding site. The clearest example of 'half-of-sites' binding is the U-box protein CHIP. The x-ray crystal structure of CHIP reveals a unique asymmetric dimer in which one of the U-box domains is blocked from interacting with E2 enzymes (*53*).

The intrinsic dimerization of isolated U-box domains appears to be weak. In the case of Prp19, analytical ultracentrifugation revealed that the Prp19 U-box (Prp19U) has only a weak propensity to dimerize. The solution NMR structure of Prp19U was determined in a monomeric state at a concentration of 1 mM, but a dimer was formed when the protein was crystallized from a solution at more than 2-fold higher concentration. A central coiled-coil domain adjacent to the U-box domain is the dominant oligomerization interface, leading to an overall tetrameric assembly (*62*). The short tether to the coiled-coil domain results in a very high local concentration of pairs of U-boxes in the tetramer, which drives U-box/U-box dimerization *in vivo* (*54*). The U-box domain of CHIP (CHIPU) exhibits similar properties. For example, isolated CHIPU crystallizes as a dimer (*93*). However, like Prp19, the U-box is not the dominant oligomerization domain, but rather, it is the central coiled-coil region of the protein (*53, 96*). Our studies showed E4BU has a very weak propensity to dimerize. Moreover, it has been resistant to

crystallization. It appears that oligomerizatin of E4BU would require an additional oligomerization domain in order to dimerize. However, dimerization of E4BU seems unlikely even in the context of oligomerization of the full-length protein due to the charged residues present in the common U-box domain dimerization interface. Thus, the E4B U-box E3 ligase stands out for its structure and apparent function in a monomeric state.

Previous reports have suggested that E4B interacts with CHIP in regulation of the myosin assembly pathway in *C. elegans* (*69*). The TPR domains of CHIP are required for interaction with E4B, but the corresponding region of E4B necessary for CHIP binding is not yet known (*68*). CHIP is a dimer and it is intriguing to consider the possibility that it forms a hetero-oligomer with E4B and whether one or both E3 ligases are active in the complex, since CHIP exhibits 'half of sites' binding, it would also be important to determine if U-box domains from one or both E3 ligases are able to bind E2 enzymes concomitantly. If a novel U-box hetero-oligomer were formed and it allowed E2 access to multiple U-box domains, the rates and relative processivity of substrate ubiquitination would likely increase.

Despite the information noted above, the function of E4B is not well characterized and further study is required. For example, only UbcH5 family E2 conjugating enzymes are known to function with the E4B U-box, but other E2 enzymes have not been tested. Evidence to date suggests E4B polyubiquitinates substrates leading to protein turnover (*101*), but more in-depth functional analysis is required to confirm this. Some clues can be obtained from characterization of the polyubiquitin chain topology of conjugates formed on E4B substrates, and such studies will be reported in

due course. The initial assignment of Ufd2 as an E4 ubiquitin ligase remains an open issue. Further analysis of E4B structure and biochemical properties will help and test and refine this hypothesis and generate a clearer understanding of the cellular activities of this unique monomeric ubiquitin ligase.

Methods

Expression Plasmids.

Mouse E4B DNA was amplified with overhanging 5' BamH1 and 3' Xho1 restriction sites for subcloning into in-house pBG vectors (L. Mizoue; Center for Structural Biology, Vanderbilt University). Residues 1072-1173, 1082-1173, and 1092-1173 were subcloned into the pBG102 vector, which produces an N-terminal 6xHis-SUMO tag fusion protein. Human UbcH5c was produced as an untagged construct from pET28 as described previously (*46*).

Protein Purification

The E4B U-box constructs as well as UbcH5c were overexpressed in BL21(DE3) *Escherichia coli* strains. The E4B proteins were purified by Nickel affinity chromatography and the His-SUMO tag was cleaved by H3C precision protease. A second purification step with the same column removed the affinity tag from the sample. The protein was further purified by anion-exchange chromatography using a SourceQ column and Superose6 gel filtration chromatography. UbcH5c was purified by cation exchange chromatography using SP resin, followed by size-exclusion chromatography with an S75 column. The proteins were dialyzed against 20 mM TRIS at pH 7.0 and 50

mM NaCl and concentrated for further studies. Protein concentrations were determined by using a calculated extinction coefficient for each protein. Protein samples for NMR experiments were expressed in minimal media with NH_4Cl and glucose as the sole nitrogen and carbon sources, respectively. $^{15}NH_4Cl$ and [$^{13}C_6$]glucose (Cambridge Isotope Laboratories) were used as needed to prepare the requisite isotopically enriched proteins.

Electrospray Mass Spectrometry

Experiments were conducted with an ESI-oaTOF (electrospray ionization orthogonal accelerating time-of-flight) mass spectrometer (micrOTOF, Bruker Daltonics, Inc., Billercia, MA), which has been modified for enhanced collisional cooling in the source for the analysis of noncovalent protein complexes. This modification and methods have been previously described (*102*). E4B residues 1092-1173 was prepared at concentrations of 250 µM, 500 µM, and 1 mM in 25 mM sodium acetate at pH 7.0.

Analytical Ultracentrifugation

The state of oligomerization of E4B was analyzed by sedimentation velocity analysis. Samples to be analyzed by AUC were first dialysed extensively against 20 mM TRIS buffer at pH 7.0 containing 50 mM NaCl, and dialysis buffer was saved to be used in the reference sector of each sample cell. AUC analysis was performed at 20 ^{o}C in a Beckman Optima XL-I analytical ultracentrifuge. Sample concentrations of 100 µM, 50 µM, and 25 µM were loaded into standard 12 mm 2-sector Epon-charcoal cells, and housed in an An-60 Ti rotor. Sedimentation behavior resulting from a rotor speed of

50,000 rpm was observed using the optical absorbance of each sample at a wavelength of 280 nm. The initial 200 scans obtained from each cell were fitted according to the continuous c(S) Lamm equation model in the SEDFIT software package (version 9.4) in order to obtain the relative prevalence of each oligomeric species (*103*).

NMR Spectroscopy

NMR experiments were conducted using Bruker DRX 600 and 800 MHz spectrometers equipped with z-axis gradient TXI cryoprobes. The ^{15}N-^{1}H HSQC experiments to compare different E4B U-box constructs were acquired at 25 °C with 750 μM protein in 25 mM NaH_2PO_4 at pH 6.5, 25 mM NaCl, and 5% D_2O. To search for oligomerization-dependent changes in linewidth, ^{15}N-^{1}H HSQC experiments were acquired at concentrations of 100 μM and 1 mM ^{15}N-E4B(1092-1173) under identical conditions.

To obtain backbone resonance assignments, 750 μM ^{13}C,^{15}N-E4B(1092-1173) was prepared in 25 mM NaH_2PO_4 at pH 6.5, 25 mM NaCl, and 5% D_2O. 2D ^{15}N-^{1}H HSQC and 3D HNCACB, CBCA(CO)NH, HNCA, and HNCO experiments were conducted. To obtain the side chain resonance assignments of E4B (1092-1173), the same sample was used to acquire 2D ^{13}C-^{1}H HSQC and 3D HCCH-TOCSY and HCCH-COSY experiments. The final resonance assignments were deposited at the Biological Magnetic Resonance Data Bank under entry 16623. 2D ^{1}H-NOESY and 3D ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC experiments were recorded to assign intramolecular NOEs. All NMR data were processed using NMRPipe (*104*) and analyzed with Sparky (*105*).

The NMR chemical shift perturbation assays employed either ^{15}N-E4BU and unlabeled UbcH5c or ^{15}N-UbcH5c and unlabeled E4BU. The NMR samples contained 100 μM ^{15}N-protein in 25 mM NaH_2PO_4 at pH 7.0, 150 mM NaCl, and 5% D_2O. Unlabeled protein in the same buffer was added to the labeled sample until the molar ratio reached 1:2 (^{15}N labeled:unlabeled). The normalized chemical shift change was calculated using the equation $\Delta\sigma = \sqrt{[(\Delta H)^2 + (\Delta N/5)^2]}$, where ΔH and ΔN are chemical shift changes in proton and nitrogen, respectively (*106*).

Structure Calculations

Upper bounds for distance restraints were generated using the CALIBA module in CYANA (*107*). The reference volume determined by CALIBA was increased 5 times before conversion in order to loosen the distance restraints. Backbone torsion angle restraints for E4B (1092-1173) were estimated from backbone chemical shifts using TALOS (*108*).

The structural analysis involved multiple iterations of CYANA calculations and inspection of the restraints and the spectra. NOEs were identified through a combination of manual and automatic assignments using CYANA. At each iteration, 100 structures were generated using CYANA. In the later stages, the 50 structures with lowest target function were selected for further refinement by restrained molecular dynamics (rMD) calculations in AMBER 9 with implicit treatment of the water solvent (*82*). After 1 ps of energy minimization to regularize CYANA structures in the AMBER force field, the temperature of the system was rapidly increased to 1200 K over 5 ps, then was slowly cooled to 0 K over 15 ps. NMR restraints were slowly turned on during the first 3 ps and

kept for the rest of the simulation. Force constants for distance and torsion angle constraints were set to 32 and 50 kcal/mol, respectively. To select the final representative ensemble of conformers, the structures were placed in order of increasing restraint violation energy and the top 20 were selected. The quality of the final ensemble was assessed using the Procheck-NMR and MOLPROBITY programs (*109, 110*). The final ensemble of 20 conformers was deposited at the Protein Data Bank under code 2KR4.

Free Molecular Dynamics Simulation of UbcH5c

Using chain A of PDB ID 1X23, the first N-terminal 6 residues from the cloning vector were removed so as to not interfere with modeling. Initially, 500 steps of minimization over a period of 2 fs were run to allow relaxation of the initial structure. A 1000 fs molecular dynamic simulation was run in which coordinates were extracted every 50 fs, generating an ensemble of 20 structures. The molecular dynamics simulations were run in AMBER9 with the ff03 force field at a constant temperature of 300 K. The structure was solvated in a cubic box of TIP3P water using a minimum distance of 14 Å between the protein and solvent edges.

E4BU-UbcH5c docking using HADDOCK

A model for the E4BU-UbcH5c complex was generated using HADDOCK2 (*87*). The four unstructured N-terminal residues of the E4BU construct were removed from the coordinate file so as to not interfere with docking of UbcH5c. The average solvent accessibilities per residue in the ensemble of E4BU and UbcH5c were calculated using NACCESS (*111*). Residues with a solvent accessible surface higher than 50% and

chemical shift perturbations more than 1 S.D. above the mean were designated as active for HADDOCK calculations. The adjacent residues with solvent accessibility greater than 50% were designated as passive residues. In the first iteration of the calculation, an initial ensemble of 1000 rigid body docking models was generated. The 200 lowest energy models were selected for a second iteration with semi-flexible simulated annealing and these were further refined in explicit water. The ensemble of structures was generated using the final ensemble of 20 E4BU structures and 20 UbcH5c structures from a room temperature MD simulation in AMBER starting from the crystal structure of UbcH5c, as described above. The 20 lowest energy models were selected from the most populated cluster with the lowest HADDOCK score and used for analysis.

In vitro autoubiquitination assay

Ubiquitination experiments were carried out at a final volume of 20 µl with E1 (BostonBiochem) at 52 nM, bacterially expressed His_6-E2-UbcH5a at 0.6 µM, ubiquitin (BostonBiochem) at 50 µM, and E4B 1072-1173 at 1 µM. The assay was performed in buffer containing 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$ and 25 mM Tris-Cl at pH 7.5. The reactions were activated with 5 mM ATP and incubated at 30 ^{o}C for the different times indicated. To stop the ubiquitination reaction, the samples were incubated for 15 minutes at 90 ^{o}C after the addition of 5 µl SDS-loading buffer. Reactions were resolved on a NuPAGE 4-12% Bis-Tris gradient gel (Invitrogen) and detected by a SimplyBlue SafeStain (Invitrogen).

CHAPTER III

ADDRESSING ALLOSTERIC ACTIVATION OF THE E2~UB CONJUGATE USING THE E4B U-BOX DOMAIN

Introduction

Post-translational modifications propagate molecular signals critical for cellular viability. The small protein ubiquitin is one such modifier, which functions in a large number of signaling pathways. Ubiquitination proceedes via a series of three enzymes to target and attach to a substrate (*40*). While this process has been extensively studied, the full molecular details of how ubiquitin is activated and transferred to specific residues on the target substrate remain uknown (*112*). Full characterization of this mechanism is essential to understanding how the different enzymes act in concert to generate the diverse range of ubiquitin modifications and functional outcomes.

Even though ubiquitin is a small and unremarkably globular protein, it has several unique characteristics (*113*). This in turn allows substrates to be modified by ubiquitin in several different ways and consequently, downstream signals can be quite diverse (*6, 27, 114, 115*). The variety of ubiquitin signals arises in part because extended poly-ubiquitin chains can be built on any one of its seven lysine residues. The three-dimensional architecture of each of these chains is unique based on which lysine residues serve as the linker between molecules in the chain. The differences in architecture generate specificity for its function. Thus, different chain topologies have been linked to DNA repair

pathways (Ub^{K63}), protein recruitment (Ub^{K6}), and the benchmark signaling pathway - proteasome mediated degradation (Ub^{K11}, Ub^{K48}) (*10, 27*).

Ubiquitination occurs via the action of three different enzymes, termed E1, E2, and E3. E1 activating enzymes begin the process by adenylating the C-terminal glycine of ubiquitin, forming a thiolester bond between the C-terminus of ubiquitin and a cysteine located in the active site of the E1. Once E1 is covalently linked to Ub, it recruits an E2 conjugating enzyme via an E2 binding domain and subsequently transfers Ub to the E2 (*41*). E2 conjugating enzymes also utilize a cysteine residue to form a thiolester bond with the C-terminus of Ub. E3 ligases bind the E2~Ub conjugate as well as the substrate of interest, and catalyze the transfer of Ub from the E2 to the substrate.

Two E1 enzymes modulate all ubiquitination activity in the mammalian system, and Uba1 appears to be responsible for an overwhelming majority of these events (*44*). Unlike the limited number of activating enzymes, there are currently 37 classified E2 conjugating enzymes and over 600 genes thought to encode E3 ligases (*27*). While all E2 enzymes contain a conserved catalytic fold domain, E3 enzymes are divided into three classes: U-box, RING, and HECT. HECT domain E3 ligases utilize a unique mechanism for Ub ligation, in which a cysteine residue accepts Ub from the E2~Ub conjugate. (*59*). RING and U-box type E3 ligases are structurally similar and make up the overwhelming majority of all E3 ligases in the proteome. U-box and RING E3s do not directly bind ubiquitin, and thus function as scaffolding-type catalysts that co-localize both the substrate and the E2~Ub conjugate to promote ubiquitination (*116, 117*). Hence, the Ub is transferred directly from the E2 enzyme to the substrate.

The mechanism by which Ub is transferred from RING or U-box bound conjugates to substrates remains a veritable black box: neither the mechanism for how Ub is released from the E2 nor how the activated Ub reaches the E3 bound substrate are known (*118*). Knowledge of these mechanisms will greatly enhance understanding of how different Ub~E2-E3 interactions propagate different Ub modifications. This in turn will inform the underlying basis for the pathophysiology of defective ubiquitination function.

In this study, we have investigated the structure and dynamics of the E2~Ub conjugate formed by the monomeric U-box domain of E4B and the E2 enzyme, UbcH5c. NMR chemical shift perturbation experiments revealed differences in the E4BU structural interface with UbcH5c versus the interface with E4BU-UbcH5c~Ub. By monitoring NMR chemical shifts, the residues whose structural interface is altered have been identified. Mutation of these residues either reduces or abrogates ubiquitin activity in *in vitro* ubiquitination assays. Together, this information provides new insight into the molecular mechanisms of E2~Ub activation, suggesting the existence of an allosteric network of residues essential to activation of the E2~Ub conjugate.

Results

Optimization of the stability of a model UbcH5~Ub conjugate

To date, the only E2 enzymes shown to function with E4B have been from the promiscuous UbcH5 family. UbcH5a, UbcH5b, and UbcH5c all function with Ufd2 or E4B, and for the ease of these experiments, UbcH5c was chosen for this study (*43, 66, 119*). An inherent problem in studying the characteristics of E2~Ub conjugates is the

labile nature of the native thiolester bond. To overcome this obstacle, we utilized a thiol conjugate analog that is produced when substituting the active site Cys85 with serine. As a result, reaction of Ub with the E2 produces a more stable oxyester bond (*120*). For example, the charged UbcH5c mutant (UbcH5m) has a long enough lifetime to be visualized by SDS-PAGE, whereas the charged native UbcH5c does not. When UbcH5m~Ub was incubated with increasing amounts of purified E4BU at pH 7.0 for 1 hour at 25 °C, E4BU catalyzes the hydrolysis of Ub even with this more stable oxyester bond (Figure 3.1A). The hydrolysis of Ub from UbcH5m is a base catalyzed reaction, and thus the pH was optimized to prevent unwanted hydrolysis in the presence of E4BU. To this end, UbcH5m~Ub was incubated with equimolar amounts of E4BU at varying pH levels, and the extent of hydrolysis over the course of 66 hours was analyzed. pHs in the 4-5 range effectively inhibit E4BU catalyzed UbcH5m~Ub hydrolysis during this period while a pH of 6 or greater does not (Figure 3.1B). For future experiments involving study of the UbcH5m~Ub conjugate, a pH of 5 was used to prevent ubiquitin release in the presence of E4BU.

Ubiquitin conjugated to UbcH5 affects E4BU binding

Previously, NMR chemical shift perturbation analysis was used to map the interaction between free (unconjugated) UbcH5c and E4BU (*119*). Based on the observation that E4BU catalyzes the activation of Ub~UbcH5m, we examined the differences in chemical shifts of E4BU complexes with UbcH5c and the UbcH5m~Ub conjugate to determine if the covalent attachment of ubiquitin alters the way in which the two proteins interact.

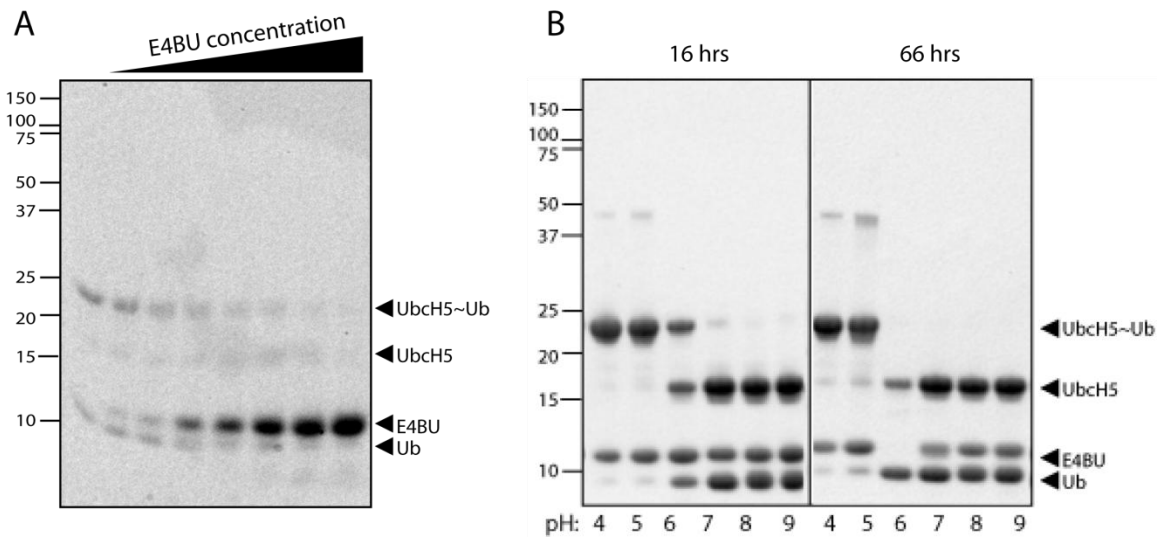

Figure 3.1. E4BU catalyzes Ub release from UbcH5m~Ub, but is inhibited at pHs below 6.0. A) UbcH5~Ub was incubated with increasing amounts of E4BU at pH 7.0 for 60 minutes, and the extent of UbcH5 – Ub hydrolysis was monitored by SDS-PAGE. B) UbcH5~Ub at varying pHs was incubated with an equimolar amount of E4BU, and the rate of hydrolysis was monitored at 16 and 66 hours by SDS-PAGE.

Comparisons of the chemical shifts of E4BU in complex with UbcH5c alone and with the UbcH5m~Ub conjugate show there are subtle differences (Figure 3.2). The residues of E4BU that are perturbed upon binding to UbcH5c are the same as those perturbed by the UbcH5m~Ub conjugate, indicating that the binding interface on E4BU does not appear to be altered by conjugation of ubiquitin. However, several of the E4BU residues are perturbed in a different manner when titrated with UbcH5c versus UbcH5m~Ub. Notably, the chemical shifts of Phe1141, Asn1142, and Arg1143 of E4BU are significantly different. For example, the chemical shifts of Phe1141 and Arg1143 of E4BU shift in different directions when titrated with UbcH5m~Ub versus UbcH5c alone. Additionally, the chemical shift perturbation of Asn1142 shows a decrease in intensity

when titrated with UbcH5m~Ub versus UbcH5c alone. Remarkably, these residues are located in the binding interface between E4BU and UbcH5, remote from the site of Ub attachment. Nevertheless, since their chemical shifts are different, these data suggest they play some role in the activation of the UbcH5m~Ub conjugate.

The next step in our analysis was to determine if there are any unique effects on UbcH5m~Ub upon binding to E4BU. This involved analyzing the extent of chemical shift perturbations of UbcH5c and Ub upon binding to

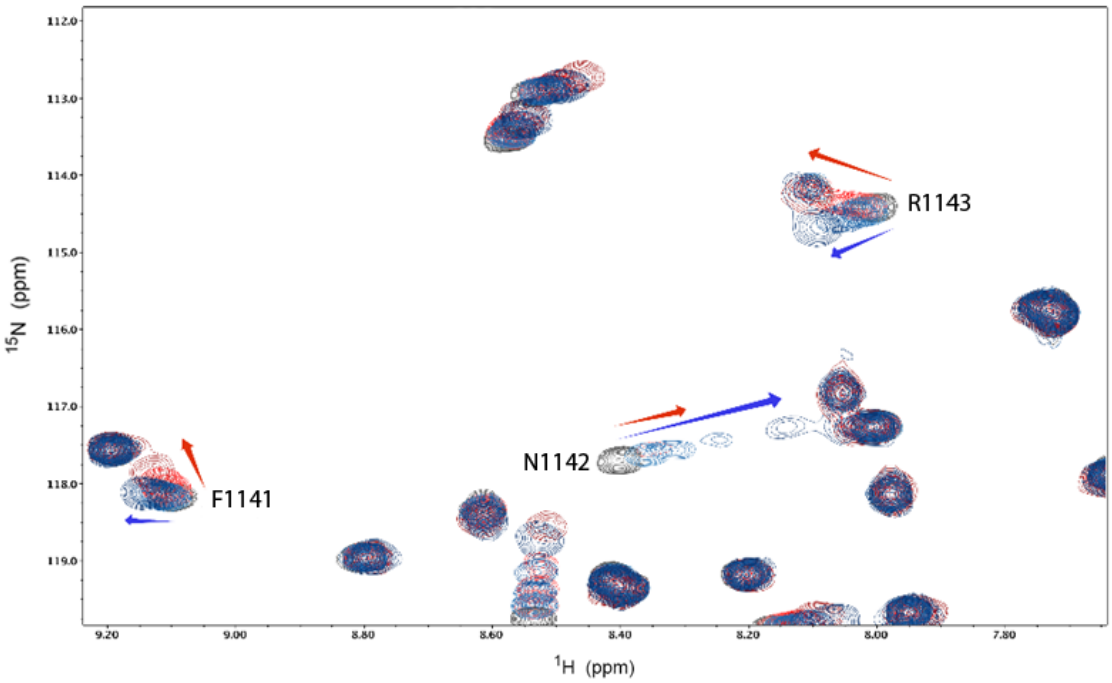

Figure 3.2. The presence of Ub conjugated to UbcH5 causes changes in the interaction with E4BU. ^{15}N-^{1}H HSQCs of ^{15}N-E4BU were collected in the absence (black) and presence of UbcH5 alone (blue) and UbcH5 conjugated to Ub (red). Changes in chemical shifts along the E4BU:UbcH5 binding interface are noted by arrows. Data courtesy of Jonathan Pruneda.

E4BU. Hence both subunits of the conjugate were isotopically enriched with ^{15}N and analyzed. As expected, the residues of UbcH5c previously known to have chemical shift perturbations when bound to E4BU, are also perturbed in the context of UbcH5m~Ub bound to E4BU, indicating that UbcH5m~Ub interacts with E4BU. Remarkably, however, several new residues exhibit small perturbations, indicating that there is a difference in binding between UbcH5c alone versus UbcH5m~Ub to E4BU (Figure 3.3). Remarkably, these new residues are not situated directly in the binding interface, but rather are located in what is referred to as the "cross-over" helix of the E2 – spanning residues 98 to 112. Effects on these residues are intriguing because of their proximity to the active site cysteine residue. Val102, Leu104, and Cys107 of UbcH5 all exhibit differences in chemical shift perturbations for the free E2 enzyme versus the E2~Ub conjguate. This observation suggests that E4BU binding to UbcH5m~Ub senses the presence of ubiquitin, at residues that are remote from the site of covalent Ub attachment.

The effects on ubiquitin were also be analyzed by this approach. Surprisingly, several residues of ubiquitin in the UbcH5m~Ub conjugate undergo significant perturbations upon binding to E4BU, even though the binding site for the E3 ligase is remote from the site of covalent attachment. Of particular interest, Leu8 and Ile44 are both perturbed, and these constitute part of an essential Ub binding site. Therefore, it appears that interaction between E2~Ub and E4BU results in an allosteric change in ubiquitin. This finding is highly intriguing as such an allosteric effect could conceivably contribute to puzzling activation of the UbcH5m~Ub conjugate.

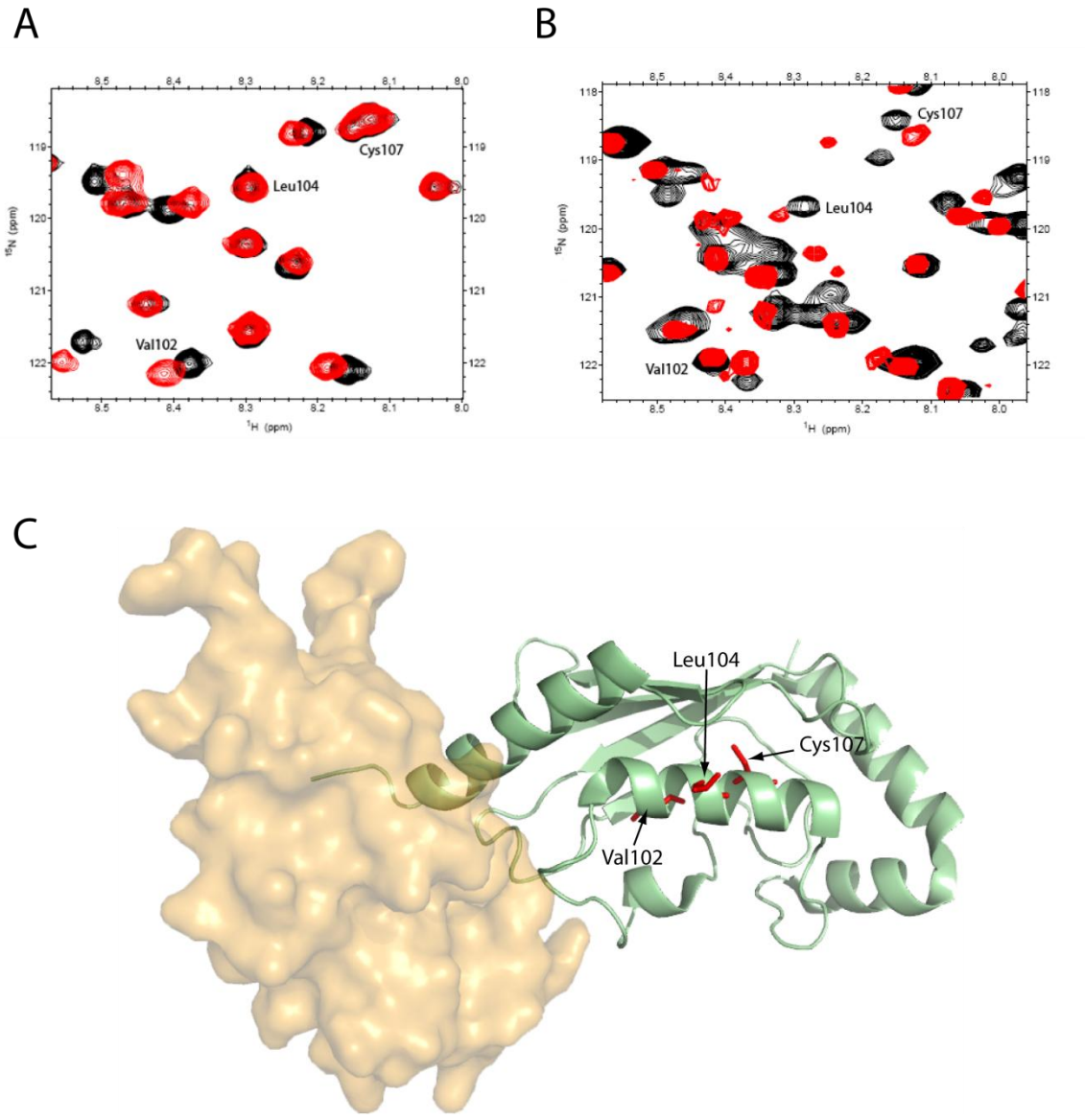

Figure 3.3. Effects of E4BU binding on ^{15}N-UbcH5~^{15}N-Ub conjugate. A) ^{15}N-^{1}H HSQC of free ^{15}N-UbcH5c (black) titrated with an equimolar ratio of E4BU (red). B) ^{15}N-^{1}H HSQC of ^{15}N-UbcH5m~^{15}N-Ub (black) titrated with an equimolar ratio of E4BU (red). C) Residues of interest on UbcH5c (green) highlighted with sidechains in red. E4BU surface shown in orange.

Testing the putative allosteric network

The chemical shift perturbation data suggest the existence of an allosteric network coupling binding of the E2 to the E3 ligase and activation of E2~Ub conjugate. In order to test the functional relevance of this putative allosteric network, we designed a series of mutations in Ub, UbcH5m, and E4BU and assayed their activity in *in vitro* autoubiquitination assays. Since the isolated U-box domain of E4B cannot undergo autoubiquitination, a construct with an additional 20 residues at the N-terminus was used for these reactions (E4BU20). Among these 20 residues are three lysine residues that have previously been shown to undergo autoubiquitination in a standard *in vitro* assay (Figure 2.9). For the mutational analysis, a total of eight mutations were made on six different residues in each of the three components of the complex (Table 3.1). E4B residues Phe1141 and Arg1143 have previously been shown to play a part in UbcH5

PROTEIN	MUTATION
E4BU	F1141A
E4BU	R1143A
UbcH5m	L104Q
UbcH5m	L104V
Ubiquitin	L8A
Ubiquitin	I44A
Ubiquitin	V70A
Ubiquitin	V70I

Table 3.1. Mutations made in E4BU, UbcH5m, and Ubiquitin to characterize the function of the residues in ubiquitination activation.

binding, so these were the first tested to determine if UbcH5 binding was affected. Samples of the mutant proteins were prepared with ^{15}N enrichment and NMR chemical shift perturbation experiments were performed. Upon titration of UbcH5 into either F1141A or R1143A E4BU20, the chemical shift perturbations of other residues in the binding interface are still observed. Hence, it was concluded that neither F1141A nor R1143A adversely affect UbcH5c binding (Figure 3.4).

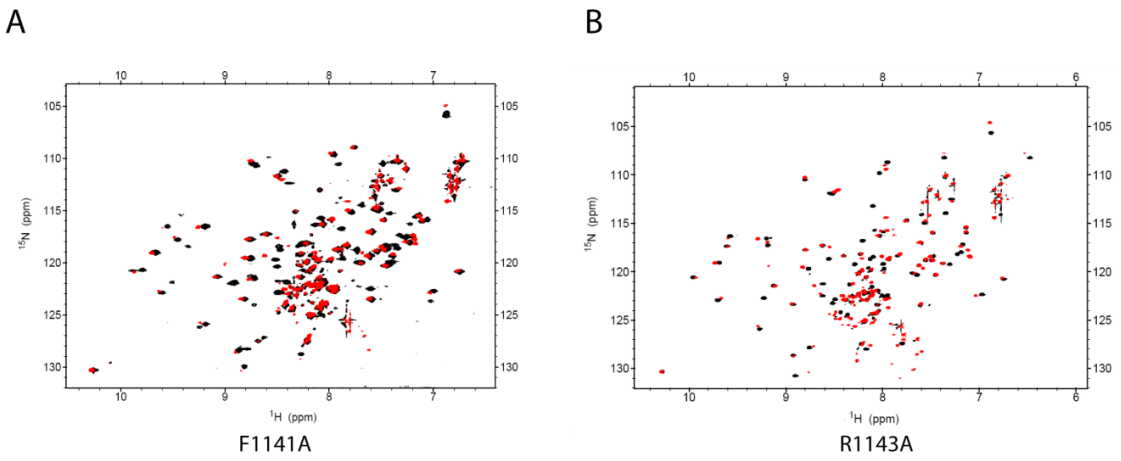

Figure 3.4. Mutations along the E4BU:UbcH5 binding interface do not adversely affect E4BU:UbcH5 binding. A) ^{15}N-^{1}H HSQC of ^{15}N-E4BU^{F1141A} in the absence (black) and presence (red) of UbcH5. B) Same as (A) but with ^{15}N-E4BU^{R1143A}

To investigate the importance of the cross-over helix of UbcH5c on ubiquitin activation, Leu104 was selected for mutation because its side chain is exposed to solvent. Residues in the hydrophobic patch of ubiquitin previously shown to be critical for non-covalent interactions (Leu8, Ile44, and Val70) were also selected for mutation. Finally, a mutation of Val70 to Ile was also made to analyze steric effects in the hydrophobic interface.

E4BU autoubiquitination assays were analyzed by coomassie stained SDS-PAGE as well as Western Blotting with anti-ubiquitin antibody. Strikingly, all 8 mutations are defective in autoubiquitination (Figure 3.5A, compare lane 6 to lanes 7-14). Wild-type UbcH5c functions normally with wtE4BU20 under these conditions, generating several autoubiquitination species. However, performing this assay with any of the eight mutants inhibits autoubiquitination activity, as clearly seen in the western blot analysis (Figure 3.5A, bottom panel).

In examining the results of these experiments, we recognized that E4BU20 autoubiquitination does not produce high molecular weight polyubiquitination species. To determine if this observation is strictly a limitation of this particular small construct, we sought to produce a more complete U-box ligase since all efforts to produce full-length E4B, or larger constructs were unsuccessful, we produced a Ufd2-E4BU chimera fusing the U-box domain of *Mus musculus* E4B to the substrate binding domain of *Saccharomyces Cerevisiae* Ufd2 ($Ufd2^{E4BU}$). This chimera allowed us to test functional questions regarding the E4B U-box, but with the ease of production characteristic of Ufd2.

Autoubiquitination of $Ufd2^{E4BU}$ produces the characteristic high molecular weight smear of highly polyubiquitined substrate (Figure 3.5B Lanes 6-8). Consequently, this chimera was used for probing the activity of the six Ub and UbcH5c mutants. These experiments showed some slight differences in autoubiquitination of $Ufd2^{E4BU}$ versus E4BU. With regards to UbcH5c, L104Q completely abrogates ubiquitination activity, whereas L104V severely restricts, but does not completely inhibit ubiquitination (Figure 3.5B, compare lanes 9 and 10). Similar observations are made with the ubiquitin mutants:

L8A and I44A completely disrupt ubiquitination activity whereas V70A and V70I appear to retain a small amount of activity (Figure 3.5B, lanes 11-14). Clearly, Leu104 and the three residues composing the hydrophobic patch of ubiquitin are essential for proper autoubiquitination of both E4BU and Ufd2^{E4BU}.

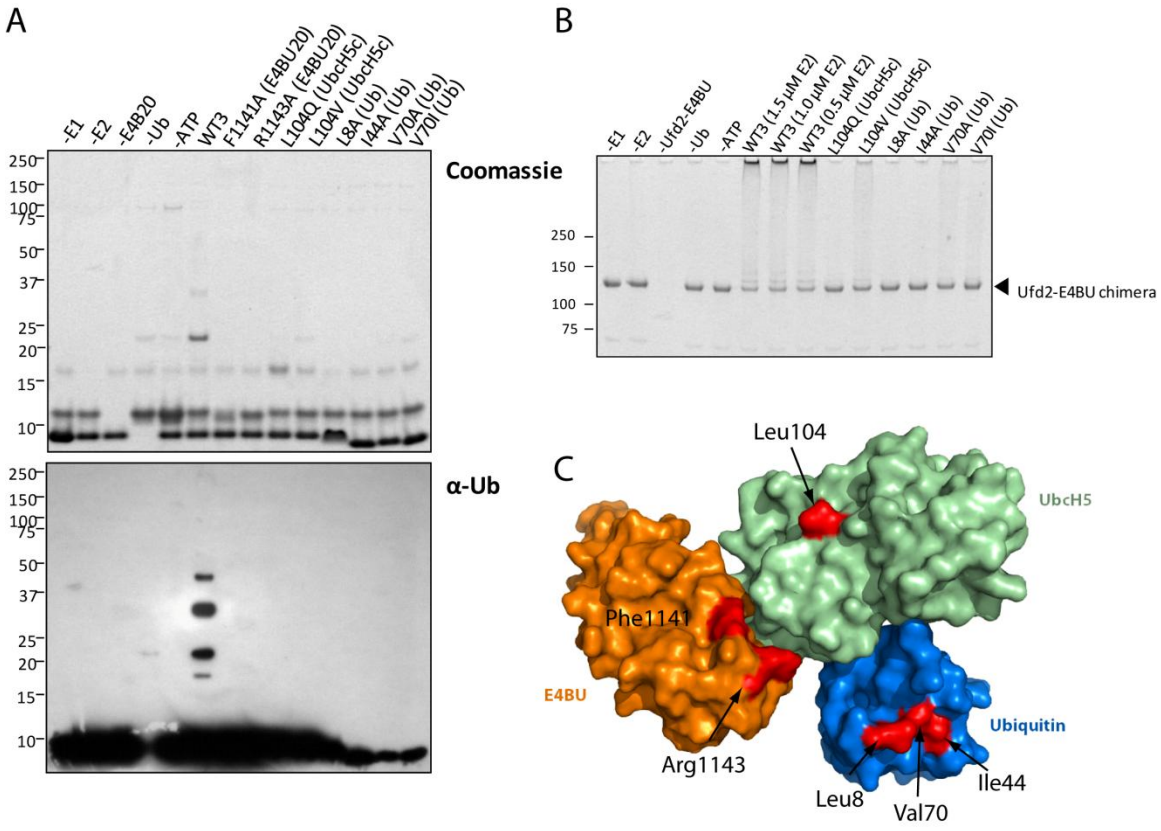

Figure 3.5. E4BU activates an allosteric network of residues in the ubiquitination machinery. A) Autoubiquitination of E4BU 1072-1173 was analyzed by SDS-PAGE. Negative controls (lanes 1-5), wild type reaction (lane 6), and mutant reactions (lanes 7-14). Gels analyzed by both Coomassie stain (top panel) and Western Blotting with α-ubiquitin antibody (bottom panel). B) Autoubiquitination of Ufd2^{E4BU} shows more high molecular weight ubiquitinated species. Controls (lanes 1-5), wild type reactions (lanes 6-8), and mutant reactions (lanes 9-14). Analyzed by Coomassie stain. C) Model surface representation of the E4BU:UbcH5~Ub complex with mutations of note represented in red. Generated by overlaying the HADDOCK model of E4BU:UbcH5 with the conjugate structure taken from NEDD4:UbcH5~Ub (PDB ID: 3JWO).

Dynamics of Ubiquitin Activation

Although three-dimensional structures of UbcH5~Ub conjugates, UbcH5-Ub non-covalent complexes, and UbcH5-E3 complexes are available, only minimal structural information has been reported for the ternary E3-E2~Ub complexes (*28*, *42*, *53*, *93*). Schulman and colleagues determined the crystal structure of a $HECT^{NEDD4L}$-UbcH5~Ub complex, and provided excellent insight into the transfer mechanism of Ub to the HECT E3 ligase (*118*). Despite offering glimpses into a ubiquitin transfer, this single structure provides only limited insight in ubiquitin activation. Moreover, the overwhelming majority of ubiquitin E3 ligases are of the RING or U-box family and do not directly interact with Ub as do HECT E3s. Consequently, many questions remain concerning how ubiquitin is activated for release from the E2 for U-box/RING E3 ligases. Here, we have applied an NMR approach to investigate the structure and dynamics of the E4BU-UbcH5m~Ub complex.

To investigate the relative dynamics of the UbcH5m~Ub conjugate in the presence of E4BU, we measured residual dipolar couplings. For these experiments, ^{15}N-enriched ubiquitin was used to form the UbcH5m~Ub conjugate, and experiments were performed in the presence and absence of E4BU. Reduction in the degree of flexibility of ubiquitin in the conjugate should result in significant differences in the RDCs of the bound versus the unbound state using C12E5:hexanol as an alignment medium (*121*). Both proteins remained soluble upon addition of C12E5 and hexanol, and partial alignment was verified through deuterium splitting. Once these proper conditions for partial alignment were identified, a production run was performed. Surprisingly, no significant differences in RDCs of ubiquitin for the free conjugate versus the E4BU-

bound conjugate were observed (Figure 3.6). Since there are a number of optimization steps involved in obtaining reliable data, it is not yet clear if these observations are fully representative of the actual complex. At the very least, the samples need to be prepared again and the data re-measured. The fact that the conjugate is hydrolyzing over the course of the experiment poses the most severe problem and the extent of Ub release needs to be carefully monitored. Nevertheless, it is possible that there are no changes in the dynamics of Ub or perhaps that the changes fall outside the range of sensitivity of the method.

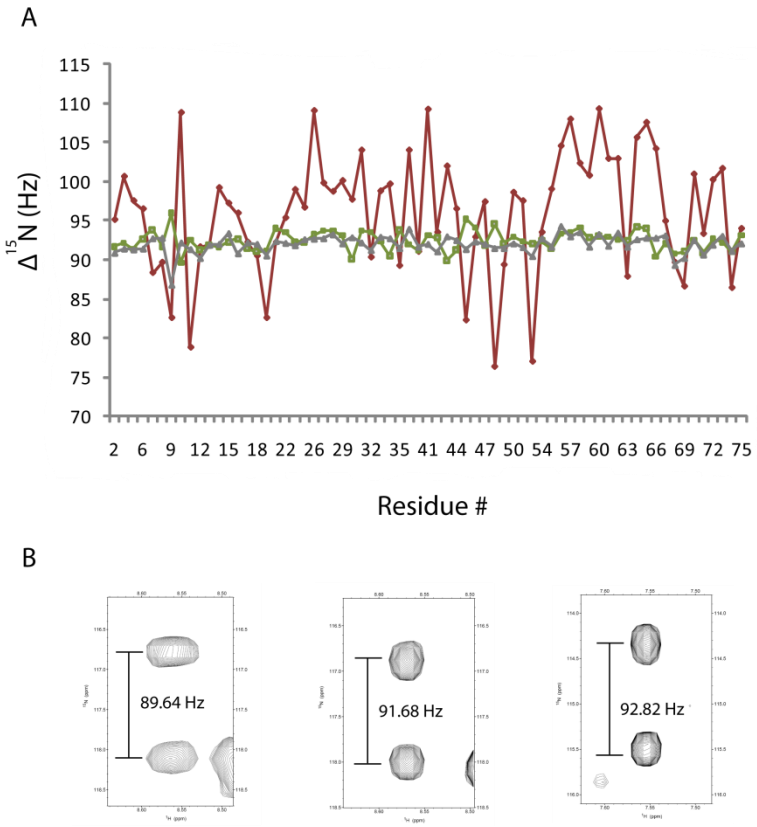

Figure 3.6. Ubiquitin motion is not significantly altered by UbcH5~Ub:E4BU binding. A) Ubiquitin RDCs in the free state (red), conjugated to UbcH5 (green) and conjugated to UbcH5 in the presence of E4BU (grey). B) Comparison of Leu8 of ubiquitin from (L-R) free, conjugated to UbcH5c, and conjugated to UbcH5c bound to E4B.

Discussion

Here we show that binding of the U-box domain of E4B catalyzes the hydrolysis of the covalent bond of the UbcH5c~Ub conjugate. NMR chemical shift perturbation experiments were used to investigate the underlying structural basis of this effect. These experiments reveal differences in binding of E4BU to the UbcH5~Ub conjugate versus free UbcH5. The residues that appear to be affected in these interactions were identified and a few key mutations were made to test their role in the activation of UbcH5~Ub. These mutations severely restricted activity in *in vitro* autoubiquitination assays, suggesting the presence of an allosteric network within the E3-E2~Ub ternary complex. Preliminary residual dipolar coupling experiments were performed to examine the dynamics of ubiquitin in the ternary complex, but no significant differences were observed. The data presented here provides a framework with which to further investigate the mechanism of ubiquitin activation.

Recent studies have begun to provide insights into the activation and transfer of ubiquitin from the E2~Ub conjugate to substrate. However, many questions remain. Here we provide evidence that binding to the E3 ligase triggers Ub activation via an allosteric mechanism. But how do these effects facilitate the hydrolysis of the UbcH5~Ub thiolester and subsequent transfer to substrate? Thorough characterization of the structural dynamics of the E3-E2~Ub complex will be required to understand the molecular basis for this key step in ubiquitination.

Investigation of available structures of E2~Ub conjugates, as well as personal communication with the Klevit lab at the University of Washington, suggests a high degree of flexibility in the E2~Ub conjugate (*42, 118, 120*). Comparing structures of free

UbcH5~Ub conjugate (PDB ID: 3A33) with that of $HECT^{NEDD4L}$-UbcH5~Ub (PDB ID: 3JVZ), reveals remarkably different dispositions of ubiquitin with respect to the E2, spanning approximately 50-60 Å of conformational space (Figure 3.7). Therefore, we hypothesize that E2-E3 binding restricts the flexibility of ubiquitin, facilitating its release to the substrate targeted by the ligase. Evidence has been obtained that this restriction in flexibility favors a "closed" state conformation of the E3-E2~Ub complex. Apparently, these closed conformations are more readily hydrolyzed, thereby leading to accelerated release of Ub.

Additional dynamic measurements will be necessary to generate a full understanding of the effect of E3 on the E2~Ub conjugate. RDC measurements and/or NMR relaxation data have the potential to provide insights regarding the flexibility of ubiquitin in the presence and absence of the E3 ligase (*122*). RDCs can be used to characterize the dynamics of protein or protein complexes in addition to providing information on the folding of the protein or complex. NMR relaxation experiments measuring the backbone dynamics of a protein can be used to characterize its rotational diffusion. Together, complete sets of RDC measurements in conjunction with relaxation data should offer a detailed understanding of the relative structural dynamics of the E2~Ub conjugate, to enhance understanding of the mechanism for activation by the E3 ligase. Characterization of the ubiquitin activation process will provide insight into ubiquitination substrate selectivity and the propagation of cellular signals.

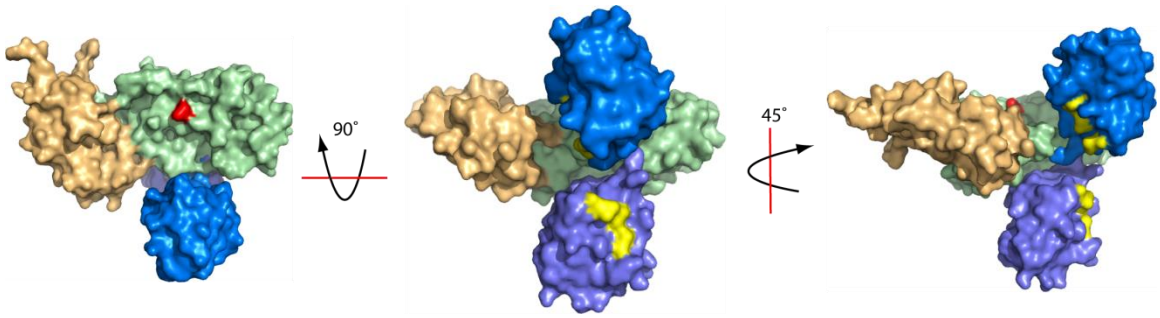

Figure 3.7. Flexibility in the UbcH5c~Ub conjugate. The HADDOCK derived complex structure (Chapter II) of E4BU (beige) bound to UbcH5c (green) was used to demonstrate the flexibility of ubiquitin. Crystal structures of UbcH5~Ub conjugates (slate, 3A33 (*42*); blue, 3JWO (*118*)) were aligned with the E4BU-UbcH5c HADDOCK model using UbcH5c as the basis for alignment. Hydrophobic patch of ubiquitin is highlighted in yellow and Leu104 of UbcH5 is shown in red.

Many hurdles exist in obtaining structural information for the E4BU-UbcH5c~Ub complex. The overall size of the complex is ~35 kDa and significantly moreso if a substrate is involved. The apparent flexibility in the E2~Ub conjugate and the transient nature of the bond between E2 and Ub, even if it is stabilized through an oxyester bond, will pose significant challenges because stability of many hours (at a minimum) is required for structural analysis. However, advances in solution methods such as Small Angle X-Ray Scattering (SAXS) offer a promising potential way forward (*123*, *124*). Instead of pursuing static information of what is intrinsically a dynamic complex, SAXS and NMR have the potential to offer insight into movement and flexibility, and enable generation of dynamic models for this system. Moreover, these approaches could be used to build structurally dynamic models for multiple complexes - (1) isolated E2~Ub, (2) E2~Ub bound to E3, and (3) E2~Ub bound to E3 with substrate.

Function in diverse signaling pathways is made possible because ubiquitin is able to form a number of distinct chains. It is important to understand the molecular processes that govern how Ub modifies its substrates and the factors responsible for variability and substrate selectivity. Continued study into the dynamic mechanism of U-box/RING mediated ubiquitin activation and transfer will provide essential insight into ubiquitin specificity and has the potential to impact treatment of a number of ubiquitination-associated diseases (*5, 125*).

Methods

Expression Plasmids.

Mouse E4B DNA was amplified with overhanging 5' BamH1 and 3' Xho1 restriction sites for subcloning into in-house pBG vectors (L. Mizoue; Center for Structural Biology, Vanderbilt University). Residues 1092-1173 were subcloned into the pBG102 vector, which produces an N-terminal 6xHis-SUMO tag fusion protein. Human UbcH5c and UbcH5c (S22R/C85S) were produced as untagged constructs from pET28 as described previously (*28*). Human ubiquitin was produced with an N-terminal 6x-His tag from pET28a as described previously (*119*).

Mutations

Mutations in E4B 1072-1173 were generated by site-directed mutagenesis. An initial stage of 2 cycles of amplification was used to generate requisite excess of product by amplifying the gene of interest with only forward or reverse primers in each reaction. Following this step, forward and reverse primer reactions are combined, and

amplification proceeds for an additional 15 cycles. Sample products were transformed directly into XL1-BLUE cells. Resulting colonies were picked and grown in LB^{Kan} media, and plasmid DNA was isolated and purified (Qiagen). Mutations were confirmed through DNA sequencing by the Vanderbilt University DNA sequencing core.

Protein Purification

Mus Musculus E4BU (1072-1173, 1072-1173^{F1141A}, 1072-1173^{R1143A}, and 1092-1173), *Homo Sapiens* UbcH5c and UbcH5c (S22R/C85S) as well as *Saccharomyces Cerevisiae* Ufd2 (1-961) were overexpressed in the BL21(DE3) STAR strain of *Escherichia coli*. Ubiquitin was overexpressed in BL21(DE3) cell line. E4BU constructs were purified by Nickel affinity chromatography and the His-SUMO tag was cleaved by H3C precision protease. A second purification step with the same column removed the affinity tag from the sample. The proteins were further purified by anion-exchange chromatography using a SourceQ column and Superose6 gel filtration chromatography. UbcH5c constructs were purified by cation exchange chromatography using SP resin, followed by size-exclusion chromatography with an S75 column. Ufd2 and ubiquitin were purified by Nickel chromatography followed by size exclusion chromatography with an S75 column. Protein concentrations were determined by using a calculated extinction coefficient for each protein. The E4BU, UbcH5, and Ufd2 extinction coefficients were calculated at 280 nm, whereas the value for ubiquitin was calculated at 277 nm. Protein samples for NMR experiments were expressed in minimal media with NH_4Cl and glucose as the sole nitrogen and carbon sources, respectively. $^{15}NH_4Cl$ and

[$^{13}C_6$]glucose (Cambridge Isotope Laboratories) were used as needed to prepare the requisite isotopically enriched proteins.

UbcH5~Ub conjugation

Purified UbcH5 (S22R/C85S) in 25 mM Na_2HPO_4 and 150 mM NaCl at pH 7.0 was incubated with purified ubiquitin in the same buffer in a 1:2 molar ratio, respectively. 5 μM Uba1, 5 mM ATP, and 10 mM $MgCl_2$ were added, and the reaction mixture was incubated for 4-5 hours at 30 °C. Reaction components were separated by an additional step of size exclusion chromatography using an S75 column equilibrated in reaction buffer or Sodium Acetate at pH 5.0 with 150 mM NaCl. Conjugated UbcH5~Ub samples were used immediately after purification.

NMR chemical shift perturbations

NMR experiments were conducted using a Bruker DRX 600 MHz spectrometer equipped with a z-axis gradient TXI cryoprobe. The NMR chemical shift perturbation assays employed either ^{15}N-E4BU and unlabeled UbcH5c or ^{15}N-UbcH5c and unlabeled E4BU. The NMR samples contained 100 μM ^{15}N-protein in 25 mM NaH_2PO_4 at pH 7.0, 150 mM NaCl, and 5% D_2O. Unlabeled protein in the same buffer was added to the labeled sample until the molar ratio reached 1:2 (^{15}N labeled:unlabeled).

Residual Dipolar Couplings

All proteins were aligned in 5% pentaethylene glycol monododecyl ether (C12E5) brought into nematic phase with addition of hexanol. 500 μl ^{15}N-enriched protein in 25

mM Na_2HPO_4 at pH 7.0 (or 25 mM $NaC_2H_3O_2$ at pH 5.0), 150 mM NaCl, and 10% D_2O was added to 25 µl C12E5 and vortexed thoroughly. Hexanol was added in 1 µl increments until solution turns from milky and turbid to translucent. Alignment of samples was checked by deuterium splitting where acceptable values range from 23-26 Hz (*126*). All IPAP ^{15}N-^{1}H HSQC experiments were collected on a Bruker DRX 600 spectrometer equipped with a TCI cryoprobe (*127*).

In vitro autoubiquitination assays

Ubiquitination experiments were carried out at a final volume of 20 µl with E1 (BostonBiochem) at 52 nM, UbcH5c at 0.6 µM, ubiquitin (BostonBiochem) at 50 µM, and either E4B 1072-1173 or Ufd2 at 1 µM. The assay was performed in buffer containing 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$ and 25 mM Tris-Cl at pH 7.5. The reactions were activated with 5 mM ATP and incubated at 30 ^{o}C for 60 minutes. To quench the ubiquitination reaction, the samples were incubated for 15 minutes at 90 ^{o}C. Gels were run on 4-12% Bis-Tris gradient polyacrylamide (Invitrogen) and detected by a SimplyBlue SafeStain (Invitrogen). Western blot analysis was completed using an enhanced chemiluminescence reagent (Thermo Scientific). Polyclonal α-ubiquitin primary antibody at a 1:1000 ratio and secondary α-mouse antibody at a 1:7000 ratio were used for detection. A 60 second exposure time was used for film development.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

Summary of this work

Structure of the E4B U-box domain

The structure of the U-box domain from *Mus musculus* E4B was determined via NMR spectrometry. As anticipated, it is structurally homologous to the U-box domain of the yeast homologue, Ufd2, previously determined by X-ray crystallography, and aligns with an *rmsd* of 1.15 Å over all backbone atoms. Analogous to other U-box and RING structures, E4BU contains an elongated α-helix at its C-terminus. In previously studied U-box and RING-containing ligases, this C-terminal helix acts as part of a dimerization motif, either creating homodimers such as the case for the U-box containing proteins Prp19 and CHIP, or heterodimers in the case of RING protein complexes BRCA1/BARD1, Mdm2/MdmX, and Ring1/Bmi1 (*53-57*).

Unlike all other U-box and RING domain E3 ligases, E4B was found to be monomeric in solution. Although initial qualitative analysis by size-exclusion chromatography suggested E4BU was a dimer, quantitative analysis by SEC-MALS, AUC, and ESI-MS all confirmed the predominance of a monomeric state in an approximate 10:1 ratio versus dimer. We went to great lengths to be sure of this because all other structurally characterized U-box domains are dimers and it has been shown that inhibition of dimerization compromises ubiquitination function (*53, 54, 95*). In the case of CHIP, oligomerization has been proposed to provide flexibility to the ligase,

increasing processivity of ubiquitination while enabling accommodation of substrates of various sizes (*63*). Clearly, additional research is needed to understand why E4B exists as a monomeric E3 ligase.

The structural information obtained about the E4B U-box domain has been related to E2 interactions. NMR chemical shift perturbation analysis was used to map the binding interface of E4BU and UbcH5c, and this residue specific data was used to generate a computation model of the UbcH5c-E4BU complex. The model reveals a distinct binding interface centered on the central α-helix of E4BU in addition to the two flanking loop regions. These three structural elements interact with helix 1, and loops 4 and 7 of UbcH5c. This binding interface is very similar to that observed for other U-box-E2 complexes (*53, 93*). Based on our analysis to date, E4B appears to be unique in that it only interacts with UbcH5 E2 enzymes. Additional study to fully characterize the E2 selectivity of E4B would be of great interest with respect to gaining better understanding of ubiquitination specificity resulting from pairing of E2 conjugating and E3 ligating enzymes.

Function of E4B

Despite its unique monomeric state, E4BU functions in *in vitro* autoubiquitination assays. E4BU binds and ubiquitinates using the UbcH5 family of E2 conjugating enzymes, and is readily capable of catalyzing poly-ubiquitin chain synthesis without the need for any additional E2 (or E3) enzymes (Figure 2.9 and 3.5). Surprisingly, our preliminary E2 screening showed that E4BU functions only with the UbcH5 family (Figure 4.1-2). This observation coincides with UbcH5 enzymes being used in all

published studies of E4B and Ufd2 function (*66, 68, 72, 101*). However, it should be noted that a systematic and complete survey of all E2 conjugating enzymes has yet to be performed. This would require additional experiments exploring for example optimization of conditions for the assay, and examination of ubiquitination activity on substrates.

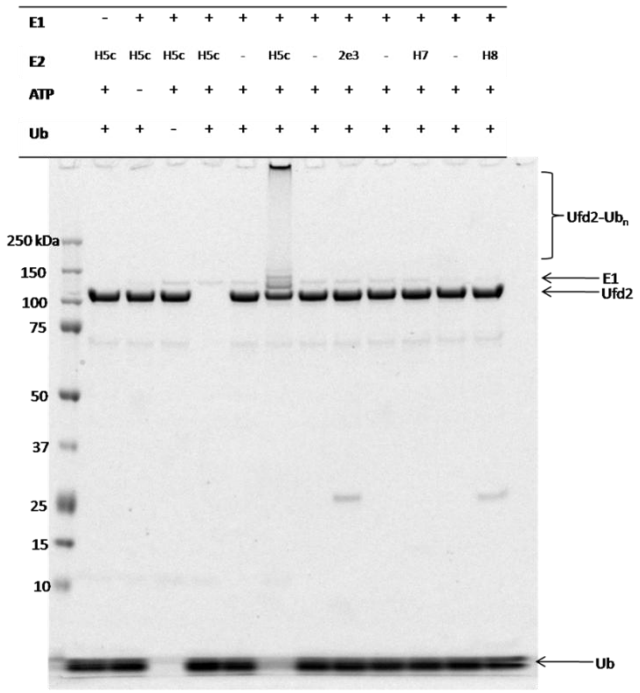

Figure 4.1. Ufd2-E2 screening. Ufd2 was subject to *in vitro* autoubiquitination assays using a subset of different E2 enzymes.

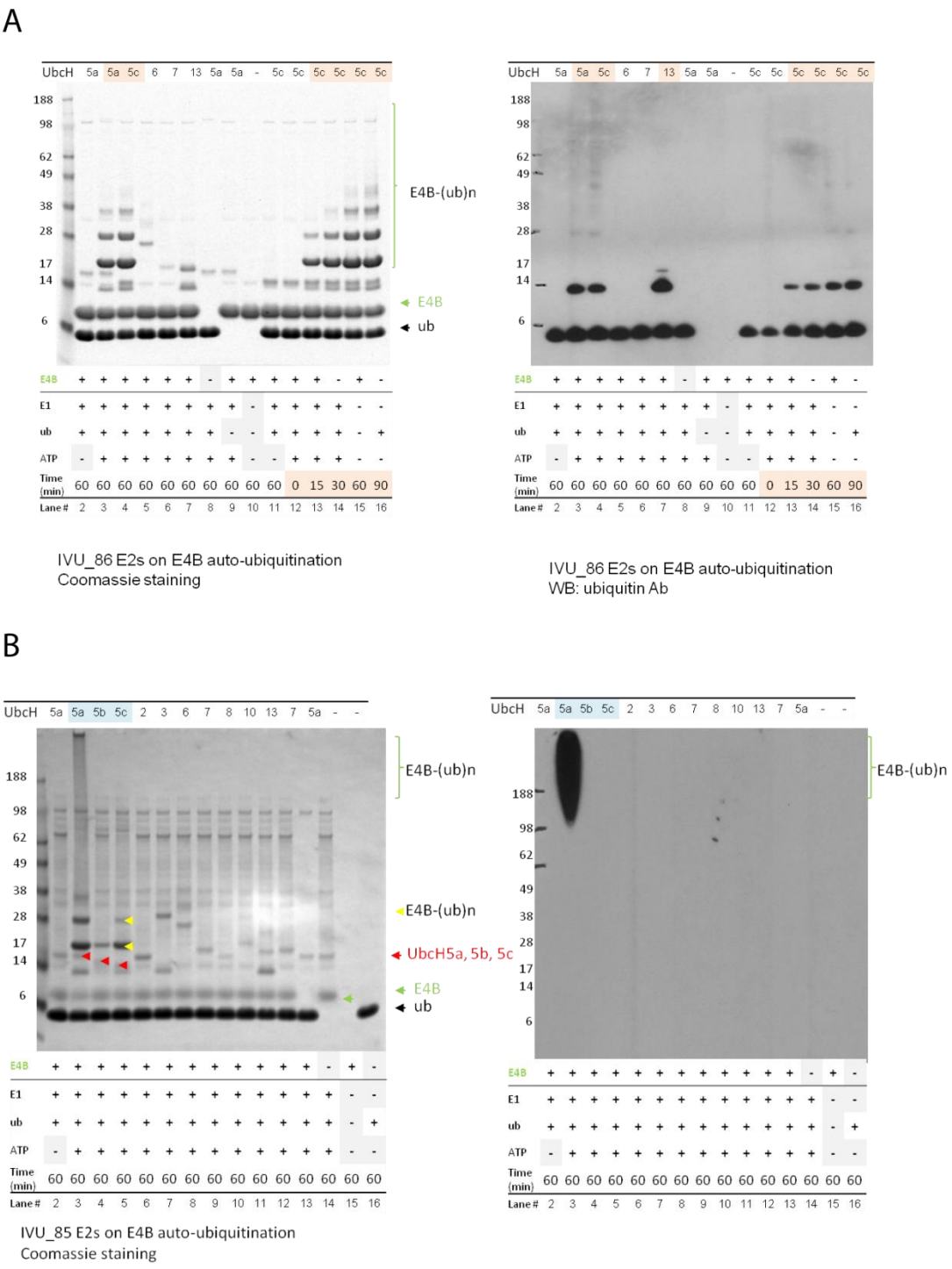

Figure 4.2. E4B was tested for function with a variety of E2 enzymes in *in vitro* autoubiquitination assays. A) Enzymes UbcH5a, UbcH5c, Ube2e1, Ube2e2, and Ubc13/Uev1. Detection via Coommassie stain (left) and α-Ub antibody (right). B) Enzymes UbcH2, Ube2r1, Ube2e1, Ube2l3, Ube2e2, Ube2c1, and Ubc13/Uev1. Coomassie stain (left) and α-Ub antibody (right). Figure courtesy of Yoana Dimitrova.

The UbcH5 E2 enzymes are best known for their promiscuity. Nevertheless, function with only these E2s would be intriguing (*46*). In particular, one may be led to ask what cellular role E4B has if it only requires broad, unspecific ubiquitination of substrates as generally observed with UbcH5. E4B/Ufd2 has primarily been implicated in ubiquitin-mediated proteasomal degradation of substrates such as UNC-45, ataxin-3, and Ste6p* (*68, 101, 128*). In addition to these roles in proteasome-dependent degradation, Ufd2 has also been shown to polyubiquitinate FEZ1, a regulator of neuritogenesis, via Ub^{K27} linkages that do not promote degradation, but rather function in an as of yet unidentified pathway (*72*). E4B has also been shown to polyubiquitinate securin, a factor in sister chromatid exchange separation, but it remains to be seen whether or not this results in degradation via the proteasome (*129*). Clearly, E4B ubiquitinates many different substrates, but further study is necessary to generate a more general understanding of its role in cellular function.

A critical role for the E3 ligase in ubiquitin activation and transfer to substrates has been previously demonstrated for RING type E3 ligases (*130*). The work in this thesis suggests a similar role for the U-box domain of E4B. A few key residues in ubiquitin, UbcH5c, and E4BU were identified and shown to be essential for triggering activation of the UbcH5c~Ub conjugate, suggesting the existence of an allosteric mechanism for this process. To further characterize the role of E4B in the cell, it will be necessary to identify any additional E2 enzymes with which it interacts and functions. It will then be imperative to expand upon these findings and investigate potential allosteric networks for other E2-E4B pairs and determine whether or not these networks are conserved from E2 to E2. Interestingly, reports from the Klevit laboratory indicate that a

similar allosteric network exists for the RING E3 heterodimer BRCA1/BARD1. In particular, they found that L63A and K65A mutations in BRCA1 abrogate ubiquitin transfer, just as mutations in the analogous E4BU residues F1141A and R1143A. The mutations in UbcH5c and ubiquitin used in our study also inhibited ubiquitination activity, as we observed for E4BU. Together, the data on E4BU and BRCA1/BARD1 suggests that this allosteric network may be a common mechanism used to promote the activation of the E2~Ub conjugate. Moreover, these results suggest E4BU is an excellent model system for future study of ubiquitin activation and transfer. E4B is highly attractive as a model system because as a monomer it has clear advantages over the RING and U-box protein dimers for structural analysis.

Future Directions

Production of full-length mammalian E4B

Full-length protein is required n order to fully understand the function of E4B. Attempts to express full-length E4B in *E.coli* cells were unsuccessful, most likely due to the sheer size of this protein, since bacteria do not typically produce such large proteins. While sequence analysis suggests conservation of the coiled-coil motifs, and the structural studies contained in Chapter II validate the homology of the U-box domain, the structure of the core amino acids of E4B is not known. During the course of this thesis work, various constructs of the N-terminal domain of E4B were designed and tested for expression and solubility, but all proved to be prone to rapid degradation or insoluble. Shifting recombinant expression of these constructs, as well as the full-length protein, to

either insect or mammalian cells is the next logical step toward achieving the goal of producing full-length E4B and large N-terminal fragments.

One of the pressing questions about mammalian E4B is whether or not it functions in proteasome shuttling like its yeast counterpart, Ufd2. Because other aspects of Ufd2 function have been validated in E4B (such as p97 and CHIP interactions), it is anticipated that interactions with Rad23 and Dsk2 homologues would also be preserved. If E4B does not share these properties, it would most likely mean that additional factors play a part in mammalian proteasome shuttling, and in order to effectively characterize the proteasome response, these factors would need to be identified.

E4B function with E2 enzymes

The UbcH5 family (Ubc4 in lower eukaryotes) are the only E2 conjugating enzymes currently known to function with E4B. Functioning exclusively with the H5 family would be intriguing since most other E3 ligases appear to be able to interact with multiple E2 conjugating enzymes, a factor that is key in generating the diversity of polyubiquitin chain types (*12, 53*). Furthermore, research shows that multiple E2 enzymes, and even multiple E2-E3 pairs are necessary to catalyze polyubiquitin chain synthesis (*11-13, 39*). While UbcH5 enzymes alone are effective in synthesizing chain elongation, it is important to identify if any other E2 enzymes function with E4B as part of the effort to understand E4B function. My initial screens of E4BU and Ufd2 with seven additional E2 enzymes did not reveal autoubiquitination activity function for any of these, but it will be necessary to perform this search more systematically as described above. A yeast two-hybrid approach has previously been employed to explore E2-E3

interactions for BRCA1/BARD1 and this could be quite valuable as well for E4BU (*12*). Once E2 conjugating enzymes that interact are identified, binding to E4BU can be characterized in more detail by NMR chemical shift perturbation analysis and function assessed by *in vitro* autoubiquitination assays, as demonstrated in Chapter II of this thesis.

E4B substrate identification

Neither the substrate binding domain, nor the substrates of E4B are known. Production of full-length E4B, or a ΔU-box fragment is necessary to identify its substrates. Yeast two-hybrid or protein microarray screening provide potential avenues to E4B binding partners, and thus possible substrates. This information is not only essential to fully understand the purpose of E4B in the cell, but would also provide the substrate required to further elucidate the detailed mechanism of ubiquitin transfer. Current ubiquitination activity assays rely on autoubiquitination of E4B. Testing ubiquitination assays on a substrate may reveal E2-E4B interactions and functional nuances that are not evident from study of autoubiquitination alone. Consequently, comparison to substrate ubiquitination is absolutely essential.

The central armadillo-like repeat domain of Ufd2 has a high degree of sequence similarity with the analogous region of E4B. The crystal structure of Ufd2 shows an elongated structure, spanning approximately 146 Å in total length (*43*). Because of this large distance, it is believed that the substrate binding region of Ufd2/E4B is located proximal to the C-terminal U-box. If this were not the case, ubiquitin transfer from the U-

box bound conjugate would potentially become inefficient since there would be a large gap in atomic space between the bound substrate and U-box bound E2~Ub conjugate.

Understanding ubiquitination dynamics

Many questions still remain concerning the molecular basis for ubiquitin transfer to a substrate. While we have established the presence of an allosteric network, structural understanding of the mechanism at the atomic level is lacking. The two main hurdles in characterizing the structural basis of ubiquitin activation and transfer are (1) utilizing biophysical techniques that characterize the dynamics of the complex, and (2) developing the requisite methods for these large molecular weight complexes. For example, the studies carried out in Chapter III were initially proposed for the study of the RING E3 BRCA1/BARD1 heterodimer. However, the high molecular weight of the dimeric complex severely limits the quality of the NMR data. Similar problems have also been encountered for NMR studies of the CHIP U-box homodimer (S.E. Soss, personal communication). The unique monomeric nature of E4BU makes it an excellent structural model for such studies.

X-ray crystallography is the most common technique employed to study multi-protein complexes since it is not limited by high molecular weights like NMR. However, the downside to this method is the static nature of the structural information, which can be limiting if dynamics are critical to the mechanism of action. Since ubiquitin transfer is a dynamic process, more than a crystallographic snapshot is required. NMR is uniquely well suited to address this problem. In particular, the application of new NMR techniques such as TROSY, CRINEPT, and CRIPT on deuterated proteins allow for study of higher

molecular weight systems (*131*). However, because these techniques are relatively new and require substantial optimization, it is not yet clear if they provide a universal solution for studying the mechanism of activation of the E2~Ub conjugate and Ub transfer. Using E4BU as a model substrate has allowed us to construct an E3-E2~Ub model without the need for specialized techniques, and the potential for using full-length E4B for further analysis appears promising.

Preliminary residual dipolar coupling data were acquired for E4BU-UbcH5c~Ub complex. These results, while inconclusive, suggest the method can be used to obtain information about the dynamics of ubiquitin and how they change when E4BU binds to the UbcH5c~Ub conjugate. The data presented in Chapter III shows RDCs measured from one type of alignment media, but it is likely that different alignment media will be needed to complete the analysis since evidence for interaction between UbcH5c and the alignment media was obtained. NMR relaxation data offers an alternative approach to obtain information about the dynamics of the system. We firmly believe that the dynamic flexibility of Ub in the conjugate is shifted when the conjugate is bound to E4B, and that this effect catalyzes Ub transfer to the substrate.

SAXS may also prove useful for generating a working molecular model of the dynamic process of ubiquitin transfer. Despite the low resolution of the structural data obtained, SAXS is not limited by molecular weight and it has potential to provide insight into dynamic systems (*123*). The approach would involve building up a series of experiments on the isolated proteins, sub-complex, and the full ubiquitination machinery. Initial assessment of the E2~Ub conjugate alone should reflect a molecular envelope with a high occupancy of molecular space. Upon addition of the U-box domain (and

eventually full-length E4B), the change in ubiquitin dynamics should be directly observed in a reduction in the conformational envelope. Once E4B substrates are identified, inclusion into the SAXS studies will provide crucial insight relating how ubiquitin is positioned for transfer to substrates. Such a dynamic structural model will offer unprecedented insight into the mechanism for protein ubiquitination.

Implications of this work

The proteome may encode more enzymes (activating, conjugating, ligating, and deubiquitinating) associated with the ubiquitination pathway than any other signaling pathway in the cell (*132*). Understanding how these enzymes interact to elicit distinct and specific responses is of critical importance for understanding the various disease processes associated with faulty ubiquitination. Miscommunication or defects in the ubiquitination process can result in developmental defects as well as a vast array of diseases such as cancer and neurodegenerative diseases (reviewed in (*5*)). By identifying the defects in ubiquitination, it is hoped that new therapeutic strategies for treatment of disease could be developed. For example, if it is found that ubiquitin is not being transferred to the substrate correctly, the ubiquitin activation process can be investigated and potentially targeted by small molecules to increase or decrease substrate affinity. In addition to analyzing the mechanism of ubiquitination, insights into proteasome function also offer multiple avenues for therapeutic development (*133*).

Currently, the ubiquitination pathway is a hotbed of activity for disease identification and treatment. Proteasome targeting is of high interest in drug development. Proteasome inhibitors are now being developed as potential cancer

chemotherapeutics (*134-136*). However, global proteasome inhibition creates its own set of problems because it may not be able to effectively differentiate between those substrates that require degradation and those that should not be degraded. Because E4B is hypothesized to factor in proteasome shuttling, it could be a likely target for therapeutics upstream of the proteasome itself. This potential can be assessed once a more complete set of E4B substrates is identified.

Structural study of ubiquitination associated proteins and their complexes is leading to a detailed understanding of the ubiquitination cascade at the atomic level. This atomic level information provides a solid foundation for identification of potential drug targets. E3 ligases such as E4B are a prime target. For example, since the ligase appears to be responsible for recruiting specific substrates, targeting of these interactions has the potential for selective inhibition of a ubiquitination process. In instances of unregulated ubiquitination leading to unwanted substrate degradation, such as that seen for CFTR, this strategy offers a logical therapeutic direction (*137*). E4B and the proteins with which it interacts should be readily amenable to this approach, once the explicit functional roles and any associated pathologies are unambiguously identified.

E4B and the ERAD pathway

The most studied functional role for E4B, as well as its *Saccharomyces cerevisiae* and *Caenorhabditis elegans* homologue Ufd2, is in the ERAD pathway. During nascent polypeptide secretion to the cytosol via the ER, checkpoints are employed to verify proper folding in order to preserve native cellular function. In the event of a misfolded protein, the homohexameric AAA+ ATPase p97 (Cdc48 in yeast) extracts the misfolded

substrate and begins the process of signaling for degradation. It has been proposed that E4B binds to p97 and functions as a chaperone in response to misfolded proteins (*77*). Despite the conservation between Cdc48 and p97, and between Ufd2 and E4B, the interactions of these proteins differ. In the lower eukaryotes, Ufd2 interacts with the C-terminal tail region of Cdc48 (*138*). In constrast, in the mammalian system, E4B interacts with the N-terminal domain of p97 (*139*). It is not yet known why these two homologs differ.

Ufd2 is also known to associate with the proteasome shuttling factors, Rad23 and Dsk2, but these interactions have yet to be characterized for E4B in higher order eukaryotes (*138, 140*). Rad23 and Dsk2 are UBA-UBL proteins, containing carboxy terminal ubiquitin binding domains and an amino terminal ubiquitin-like domain. It has been proposed that Ufd2 will function as a scaffolding protein, bridging Cdc48 and the UBL domain of Rad23/Dsk2 (*140*). Therefore, a misfolded substrate identified by Cdc48 could then be passed onto Rad23/Dsk2, and trafficked to the proteasome for removal. This is a likely scenario, since the binding sites of Cdc48 and Rad23 do not overlap on Ufd2. However, further study is warranted to establish if this is a valid model. It should be noted that there are significant differences between E4B and Ufd2. For example, the N-terminus of Ufd2 is not conserved in E4B. This difference may explain why Ufd2 interacts with Rad23, but no interaction has been identified between E4B and hHR23.

When initially characterized, Ufd2 was classified as an "E4" enzyme (*66*). It was observed that an engineered substrate was polyubiquitinated only in the presence of both the HECT E3 ligase Ufd4 and Ufd2. Separately, Ufd2 and Ufd4 could only catalyze monoubiquitination on the substrate, which is not sufficient for recognition by the

proteasome. This observation has implications in the context of the ERAD pathway as well. In the absence of Ufd2, a substrate may only be mono- or di-ubiquitinated. Once Ufd2 is recruited by Cdc48, sufficient chain elongation may take place to target the substrate for degradation by the proteasome. This extra step in the pathway may serve as another checkpoint before the substrate is removed from the ER.

E4B and interaction with CHIP

Another study reported a similar phenomenon of E4 activity: the *C. Elegans* protein UNC-45 was only minimally ubiquitinated in the presence of either Ufd2 or another U-box E3 ligase, CHN-1 (homologue of CHIP). Ufd2 was found to interact with CHN-1, and this interaction was necessary to generate sufficient polyubiquitin chains on the substrate (*68*). There is evidence that E4B interacts with CHIP as well, but identification of the respective binding regions and structural characterization of this complex has yet to be reported (*69*). It is intriguing to consider what might be the overall oligomerization state of this assembly. For example, does CHIP still adopt a dimeric architecture when bound to E4B/Ufd2? Is CHIP dimerization somehow disrupted by the presence of E4B/Ufd2? Studies of CHIP reveal that its unique asymmetric dimeric form increases the processivity of the ligase (*63*). Perhaps a similar effect would be observed in the context of an E4B/CHIP complex.

I attempted to narrow down the binding region on E4B of the CHIP interaction using pull-down assays. Preliminary data suggests that E4B does, in fact, bind CHIP. The C-terminal half of E4B, spanning residues 742-1173 was found to effectively pull-down full-length CHIP from CNBr activated resin (Figure 4.3). Further studies are necessary to

confirm this observation and narrow down the binding region on E4B, and ultimately identify the residues on CHIP that mediate this interaction. NMR chemical shift perturbation analysis may prove to be a useful tool for this analysis, as it can specifically identify the residues on the protein in the binding interface. This data could then be used to design appropriate constructs for crystallographic-based structural studies of the complex. Characterizing a unique heteromeric E4B/CHIP E3 ligase would provide new insight into the molecular basis for the putative E4 ligase activity, and possibly more generally impact understanding of polyubiquitin chain elongation.

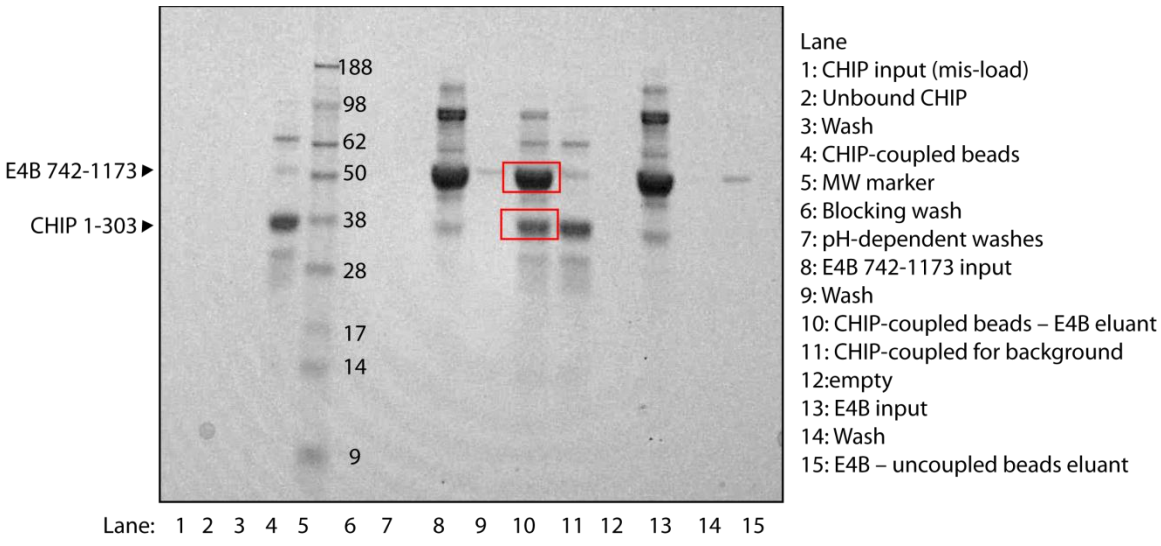

Figure 4.3. The C-terminal half of E4B binds to CHIP. Full-length CHIP (1-303) was recombinantly purified and bound to CNBr-activated resin. Input E4B (742-1173) was added to the beads in molar excess. The beads were thoroughly washed and sample retention was analyzed by boiling a sample of the beads and analyzing via SDS-PAGE. Figure courtesy of Dario Gutierrez.

Significance

Structural and functional characterization of the U-box domain of the E3 ubiquitin ligase E4B provides an excellent system for elucidation of the mechanism of ubiquitin activation. While much information is now known both structurally and functionally about the ubiquitination process, critical gaps in knowledge still remain concerning how ubiquitin is activated in an E2 conjugate and how it is then transferred to a substrate. Because ubiquitin is able to elicit different cellular signals through the location or type of polyubiquitin chain on a substrate, elucidating the mechanism of ubiquitination activation and transfer is of utmost importance to fully understand the complete function of ubiquitination and how defects in this pathway relate to disease.

The E4B U-box domain structural studies presented in this thesis suggest that like Ufd2, E4B is a monomeric protein, unlike other members of the U-box family. A monomeric protein is highly advantageous for structural analysis. Based on the similarities of E4BU-UbcH5 binding compared with other RING and U-box E3 ligases, it is implicated that mechanistic insights of ubiquitin activation generated from E4BU studies will apply to other E3 ligase systems as well.

Once further understanding of ubiquitin activation and transfer is obtained, investigation of defects within this pathway can be made. Many diseases are thought to result from problems of the ubiquitin system, but precise knowledge of where these problems exist along the pathway remains to be determined. If it is known that ubiquitin is not being transferred to a substrate correctly, or if association between the E3 ligase and substrate is negatively affected, then proper therapeutic measures can be pursued.

The work completed here using the U-box domain of the E4B ubiquitin ligase offers exciting new insight into the ubiquitin activation process.

APPENDIX A

Table A.1. E4BU backbone chemical shifts

Residue	Atomic Chemical Shift			
	CA	CB	H	N
P1089	63.53	32.12	-	-
G1090	45.33	-	8.647	110.3
S1091	58.45	63.97	8.136	115.6
A1092	52.61	19.34	8.38	125.7
E1093	56.41	30.25	8.245	120
I1094	60.64	38.7	7.988	122.7
D1095	53.42	41.74	8.23	125.1
Y1096	57.3	38.15	8.546	122.9
S1097	61.61	63.13	8.683	117.3
D1098	52.98	39.26	8.814	119.6
A1099	50.88	17.83	7.655	123.4
P1100	62.34	31.83	-	-
D1101	58.07	40.44	8.364	122.4
E1102	58.19	28.49	9.237	116.4
F1103	53.47	37.53	7.622	115
R1104	54.96	32.33	7.546	118.8
D1105	53.41	46.98	9.312	125.9
P1106	64.59	31.58	-	-
L1107	56.57	44.06	9.281	122.7
M1108	54.05	33.37	9.622	116.3
D1109	56.14	40.89	8.609	118.7
T1110	59.01	70.32	6.921	105.5
L1111	56.06	42.39	8.411	122.9
M1112	57.04	35.41	8.791	127.7
T1113	62.45	70.24	8.022	108.6
D1114	51.78	41.22	8.677	121.3
P1115	62.99	33.19	-	-
V1116	58.53	35.83	9.68	117.4
R1117	54.84	32.85	9.794	122.9
L1118	54.95	40.06	8.939	130.5
P1119	65.94	31.6	-	-
S1120	59.05	63.53	8.058	109.7
G1121	44.71	-	8.601	111.8
T1122	63.96	69.62	7.353	119.1
V1123	61.16	32.79	8.26	128.3
M1124	54.87	38.52	9.041	123.3

Table A.1. Continued

D1125	55.18	43.87	10.02	120.6
R1126	60.73	30.42	8.979	128.6
S1127	61.7	62.33	8.853	110.5
I1128	64.02	37.15	7.026	122.4
I1129	60.19	37	8.06	119.3
L1130	58.66	40.51	8.58	123.3
R1131	59.5	29.71	7.284	118.1
H1132	60.64	30.63	7.683	120.2
L1133	56.33	42.05	8.465	118.4
L1134	57.36	42.16	7.247	117.9
N1135	53.95	40.7	7.374	113.8
S1136	54.7	64	7.798	114.6
P1137	62.36	27.93	-	-
T1138	58.88	73.6	8.579	111.9
D1139	51.35	43.04	8.686	122.6
P1140	63.79	30.9	-	-
F1141	59.14	38.81	9.165	117.1
N1142	52.16	39.14	8.441	116.6
R1143	58.28	27.3	8.09	113.3
Q1144	55.79	29.53	8.204	118.7
M1145	57.55	32.2	8.395	121.1
L1146	54.33	46.26	7.833	127.4
T1147	58.95	71.82	7.349	112.5
E1148	59.53	29.19	9.179	121.3
S1149	60.29	62.75	8.148	113.2
M1150	56.77	33.92	7.607	118.8
L1151	54.72	41.89	7.22	117.2
E1152	52.81	32.42	8.125	123.4
P1153	63.71	33.04	-	-
V1154	59.16	30.81	7.446	119.1
P1155	65.41	31.69	-	-
E1156	59.76	28.19	9.729	119
L1157	56.69	42.28	6.791	120.7
K1158	60.88	32.66	7.734	120
E1159	59.36	29.26	8.055	116.3
Q1160	59.36	28.99	7.535	120.3
I1161	66.03	38.02	8.258	122.3
Q1162	59.53	28.07	8.264	118.1
A1163	55.44	18.06	8.259	122.1
W1164	62.2	29.17	8.122	122
M1165	59.75	33.54	8.871	117.8

Table A.1. Continued

R1166	59.53	30.13	8.011	118.6
E1167	58.79	29.14	7.732	120
K1168	56.91	31.42	7.614	118.9
Q1169	56.67	29.04	7.687	116.8
S1170	58.9	64.06	7.798	114.7
S1171	58.6	64.02	8.058	117.1
D1172	54.54	41.06	8.257	122.3
H1173	57.02	29.92	7.814	122.7

APPENDIX B

Additional Figures

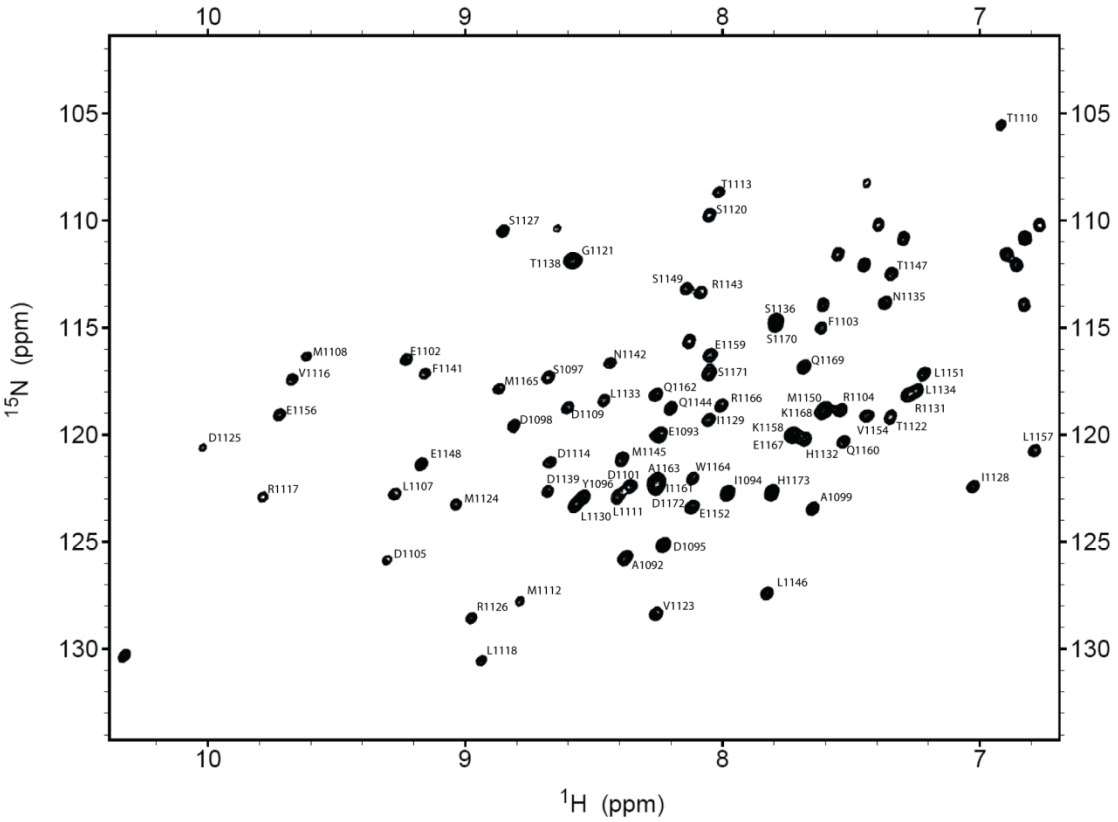

Figure B.1. Assigned ^{15}N-^{1}H HSQC spectrum of E4B residues 1092-1173.

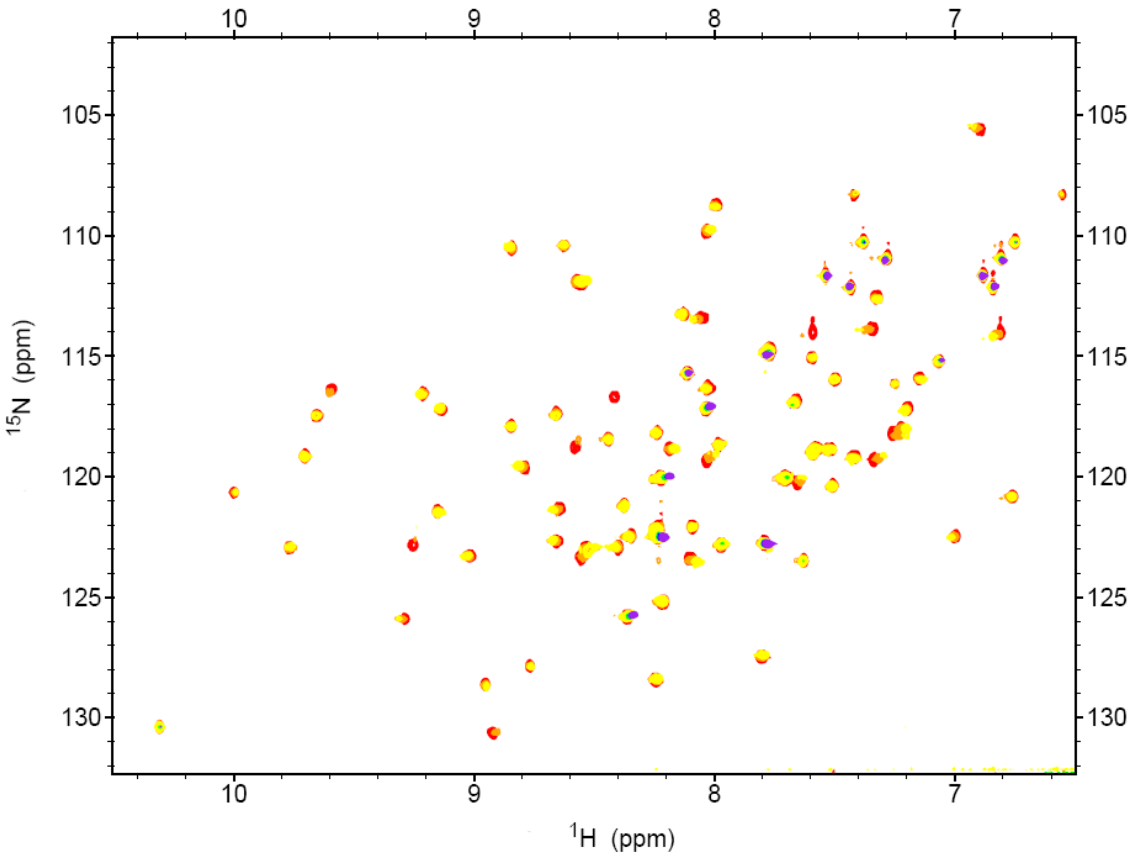

Figure B.2. ^{15}N-enriched E4BU titrated with UbcH5 at molar equivalents of 1:0.13 (orange), 1:0.25 (yellow), 1:0.5 (green), 1:1 (violet), and 1:2 (blue). Baseline spectrum shown in red.

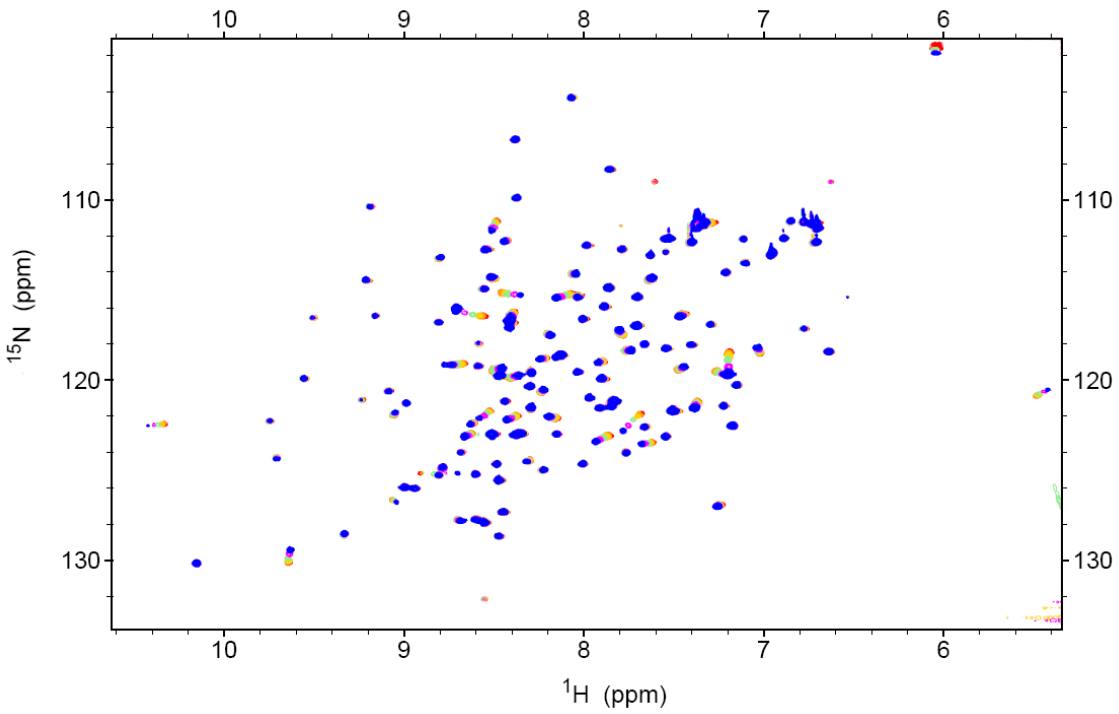

Figure B.3. ^{15}N-enriched UbcH5 titrated with E4BU at molar equivalents of 1:0.13 (orange), 1:0.25 (yellow), 1:0.5 (green), 1:1 (violet), and 1:2 (blue). Baseline spectrum shown in red.

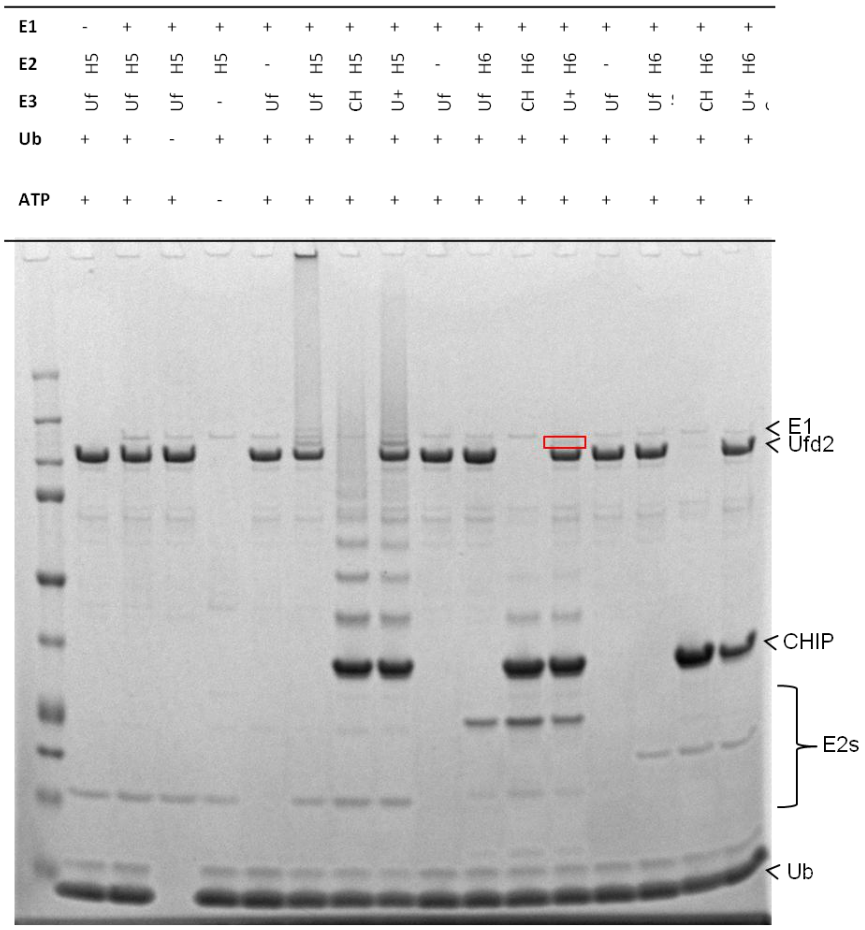

Figure B.4. Ufd2 was tested for E4 ubiquitination activity in the presence of CHIP against a variety of E2 enzymes. Ufd2 and CHIP were incubated together during an *in vitro* autoubiquitination assay with various E2 conjugating enzymes and the level of autoubiquitination on Ufd2 and CHIP was detected by Coommassie stain.

REFERENCES

1. Mukhopadhyay, D., and Riezman, H. (2007) Proteasome-independent functions of ubiquitin in endocytosis and signaling, *Science 315*, 201-205.

2. Bergink, S., and Jentsch, S. (2009) Principles of ubiquitin and SUMO modifications in DNA repair, *Nature 458*, 461-467.

3. Freiman, R. N., and Tjian, R. (2003) Regulating the regulators: lysine modifications make their mark, *Cell 112*, 11-17.

4. Johnson, E. S., Ma, P. C., Ota, I. M., and Varshavsky, A. (1995) A proteolytic pathway that recognizes ubiquitin as a degradation signal, *J Biol Chem 270*, 17442-17456.

5. Petroski, M. D. (2008) The ubiquitin system, disease, and drug discovery, *BMC Biochem 9 Suppl 1*, S7.

6. Ravid, T., and Hochstrasser, M. (2008) Diversity of degradation signals in the ubiquitin-proteasome system, *Nat Rev Mol Cell Biol 9*, 679-690.

7. Wickliffe, K., Williamson, A., Jin, L., and Rape, M. (2009) The multiple layers of ubiquitin-dependent cell cycle control, *Chem Rev 109*, 1537-1548.

8. Yoshida, K., and Miki, Y. (2010) The cell death machinery governed by the p53 tumor suppressor in response to DNA damage, *Cancer Sci 101*, 831-835.

9. Osley, M. A. (2006) Regulation of histone H2A and H2B ubiquitylation, *Brief Funct Genomic Proteomic 5*, 179-189.

10. Messick, T. E., and Greenberg, R. A. (2009) The ubiquitin landscape at DNA double-strand breaks, *J Cell Biol 187*, 319-326.

11. Windheim, M., Peggie, M., and Cohen, P. (2008) Two different classes of E2 ubiquitin-conjugating enzymes are required for the mono-ubiquitination of proteins and elongation by polyubiquitin chains with a specific topology, *Biochem J 409*, 723-729.

12. Christensen, D. E., Brzovic, P. S., and Klevit, R. E. (2007) E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages, *Nat Struct Mol Biol 14*, 941-948.

13. Rodrigo-Brenni, M. C., and Morgan, D. O. (2007) Sequential E2s drive polyubiquitin chain assembly on APC targets, *Cell 130*, 127-139.

14. Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D., and Boyse, E. A. (1975) Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells, *Proc Natl Acad Sci U S A 72*, 11-15.

15. Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987) Structure of ubiquitin refined at 1.8 A resolution, *J Mol Biol 194*, 531-544.

16. Ciechanover, A., Heller, H., Elias, S., Haas, A. L., and Hershko, A. (1980) ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation, *Proc Natl Acad Sci U S A 77*, 1365-1368.

17. Ciechanover, A., Elias, S., Heller, H., Ferber, S., and Hershko, A. (1980) Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes, *J Biol Chem 255*, 7525-7528.

18. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., and Rose, I. A. (1980) Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis, *Proc Natl Acad Sci U S A 77*, 1783-1786.

19. McGuire, M. J., Reckelhoff, J. F., Croall, D. E., and DeMartino, G. N. (1988) An enzyme related to the high molecular weight multicatalytic proteinase, macropain, participates in a ubiquitin-mediated, ATP-stimulated proteolytic pathway in soluble extracts of BHK 21/C13 fibroblasts, *Biochim Biophys Acta 967*, 195-203.

20. Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains - from structures to functions, *Nat Rev Mol Cell Biol 10*, 659-671.

21. Zhang, D., Raasi, S., and Fushman, D. (2008) Affinity makes the difference: nonselective interaction of the UBA domain of Ubiquilin-1 with monomeric ubiquitin and polyubiquitin chains, *J Mol Biol 377*, 162-180.

22. Peschard, P., Kozlov, G., Lin, T., Mirza, I. A., Berghuis, A. M., Lipkowitz, S., Park, M., and Gehring, K. (2007) Structural basis for ubiquitin-mediated dimerization and activation of the ubiquitin protein ligase Cbl-b, *Mol Cell 27*, 474-485.

23. Kozlov, G., Nguyen, L., Lin, T., De Crescenzo, G., Park, M., and Gehring, K. (2007) Structural basis of ubiquitin recognition by the ubiquitin-associated (UBA) domain of the ubiquitin ligase EDD, *J Biol Chem 282*, 35787-35795.

24. Kang, R. S., Daniels, C. M., Francis, S. A., Shih, S. C., Salerno, W. J., Hicke, L., and Radhakrishnan, I. (2003) Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding, *Cell 113*, 621-630.

25. Wang, Q., Young, P., and Walters, K. J. (2005) Structure of S5a bound to monoubiquitin provides a model for polyubiquitin recognition, *J Mol Biol 348*, 727-739.

26. Sato, Y., Yoshikawa, A., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF domains of TAB2 and TAB3, *Embo J 28*, 3903-3909.

27. Komander, D. (2009) The emerging complexity of protein ubiquitination, *Biochem Soc Trans 37*, 937-953.

28. Brzovic, P. S., Lissounov, A., Christensen, D. E., Hoyt, D. W., and Klevit, R. E. (2006) A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination, *Mol Cell 21*, 873-880.

29. Komander, D., Reyes-Turcu, F., Licchesi, J. D., Odenwaelder, P., Wilkinson, K. D., and Barford, D. (2009) Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains, *EMBO Rep 10*, 466-473.

30. Matsumoto, M. L., Wickliffe, K. E., Dong, K. C., Yu, C., Bosanac, I., Bustos, D., Phu, L., Kirkpatrick, D. S., Hymowitz, S. G., Rape, M., Kelley, R. F., and Dixit, V. M. (2010) K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody, *Mol Cell 39*, 477-484.

31. van Dijk, A. D., Fushman, D., and Bonvin, A. M. (2005) Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to

Lys48-linked di-ubiquitin and validation against 15N-relaxation data, *Proteins 60*, 367-381.

32. Eddins, M. J., Varadan, R., Fushman, D., Pickart, C. M., and Wolberger, C. (2007) Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH, *J Mol Biol 367*, 204-211.

33. Datta, A. B., Hura, G. L., and Wolberger, C. (2009) The structure and conformation of Lys63-linked tetraubiquitin, *J Mol Biol 392*, 1117-1124.

34. Madura, K. (2004) Rad23 and Rpn10: perennial wallflowers join the melee, *Trends Biochem Sci 29*, 637-640.

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

36. Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D., and Peng, J. (2009) Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation, *Cell 137*, 133-145.

37. Nishikawa, H., Ooka, S., Sato, K., Arima, K., Okamoto, J., Klevit, R. E., Fukuda, M., and Ohta, T. (2004) Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase, *J Biol Chem 279*, 3916-3924.

38. Tokunaga, F., Sakata, S., Saeki, Y., Satomi, Y., Kirisako, T., Kamei, K., Nakagawa, T., Kato, M., Murata, S., Yamaoka, S., Yamamoto, M., Akira, S., Takao, T., Tanaka, K., and Iwai, K. (2009) Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation, *Nat Cell Biol 11*, 123-132.

39. Kim, H. T., Kim, K. P., Lledias, F., Kisselev, A. F., Scaglione, K. M., Skowyra, D., Gygi, S. P., and Goldberg, A. L. (2007) Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages, *J Biol Chem 282*, 17375-17386.

40. Scheffner, M., Nuber, U., and Huibregtse, J. M. (1995) Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade, *Nature 373*, 81-83.

41. Lee, I., and Schindelin, H. (2008) Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes, *Cell 134*, 268-278.

42. Sakata, E., Satoh, T., Yamamoto, S., Yamaguchi, Y., Yagi-Utsumi, M., Kurimoto, E., Tanaka, K., Wakatsuki, S., and Kato, K. (2010) Crystal structure of UbcH5b~ubiquitin intermediate: insight into the formation of the self-assembled E2~Ub conjugates, *Structure 18*, 138-147.

43. Tu, D., Li, W., Ye, Y., and Brunger, A. T. (2007) Inaugural Article: Structure and function of the yeast U-box-containing ubiquitin ligase Ufd2p, *Proc Natl Acad Sci U S A 104*, 15599-15606.

44. Jin, J., Li, X., Gygi, S. P., and Harper, J. W. (2007) Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging, *Nature 447*, 1135-1138.

45. van Wijk, S. J., and Timmers, H. T. (2010) The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins, *Faseb J 24*, 981-993.

46. Brzovic, P. S., and Klevit, R. E. (2006) Ubiquitin transfer from the E2 perspective: why is UbcH5 so promiscuous?, *Cell Cycle 5*, 2867-2873.

47. Petroski, M. D., Zhou, X., Dong, G., Daniel-Issakani, S., Payan, D. G., and Huang, J. (2007) Substrate modification with lysine 63-linked ubiquitin chains through the UBC13-UEV1A ubiquitin-conjugating enzyme, *J Biol Chem 282*, 29936-29945.

48. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M., and Wolberger, C. (2001) Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer, *Cell 105*, 711-720.

49. Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M., and Wolberger, C. (2006) Mms2-Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation, *Nat Struct Mol Biol 13*, 915-920.

50. Petroski, M. D., and Deshaies, R. J. (2005) Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34, *Cell 123*, 1107-1120.

51. Kleiger, G., Hao, B., Mohl, D. A., and Deshaies, R. J. (2009) The acidic tail of the Cdc34 ubiquitin-conjugating enzyme functions in both binding to and catalysis with ubiquitin ligase SCFCdc4, *J Biol Chem 284*, 36012-36023.

52. Deshaies, R. J., and Joazeiro, C. A. (2009) RING domain E3 ubiquitin ligases, *Annu Rev Biochem 78*, 399-434.

53. Zhang, M., Windheim, M., Roe, S. M., Peggie, M., Cohen, P., Prodromou, C., and Pearl, L. H. (2005) Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex, *Mol Cell 20*, 525-538.

54. Vander Kooi, C. W., Ohi, M. D., Rosenberg, J. A., Oldham, M. L., Newcomer, M. E., Gould, K. L., and Chazin, W. J. (2006) The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases, *Biochemistry 45*, 121-130.

55. Linke, K., Mace, P. D., Smith, C. A., Vaux, D. L., Silke, J., and Day, C. L. (2008) Structure of the MDM2/MDMX RING domain heterodimer reveals dimerization is required for their ubiquitylation in trans, *Cell Death Differ 15*, 841-848.

56. Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C., and Klevit, R. E. (2001) Structure of a BRCA1-BARD1 heterodimeric RING-RING complex, *Nat Struct Biol 8*, 833-837.

57. Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T. K. (2006) Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b, *Embo J 25*, 2465-2474.

58. 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) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex, *Nature 416*, 703-709.

59. Rotin, D., and Kumar, S. (2009) Physiological functions of the HECT family of ubiquitin ligases, *Nat Rev Mol Cell Biol 10*, 398-409.

60. Ohi, M. D., Vander Kooi, C. W., Rosenberg, J. A., Chazin, W. J., and Gould, K. L. (2003) Structural insights into the U-box, a domain associated with multi-ubiquitination, *Nat Struct Biol 10*, 250-255.

61. Hatakeyama, S., and Nakayama, K. I. (2003) Ubiquitylation as a quality control system for intracellular proteins, *J Biochem 134*, 1-8.

62. Ohi, M. D., Vander Kooi, C. W., Rosenberg, J. A., Ren, L., Hirsch, J. P., Chazin, W. J., Walz, T., and Gould, K. L. (2005) Structural and functional analysis of essential pre-mRNA splicing factor Prp19p, *Mol Cell Biol 25*, 451-460.

63. Qian, S. B., Waldron, L., Choudhary, N., Klevit, R. E., Chazin, W. J., and Patterson, C. (2009) Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer, *J Biol Chem 284*, 26797-26802.

64. Notenboom, V., Hibbert, R. G., van Rossum-Fikkert, S. E., Olsen, J. V., Mann, M., and Sixma, T. K. (2007) Functional characterization of Rad18 domains for Rad6, ubiquitin, DNA binding and PCNA modification, *Nucleic Acids Res 35*, 5819-5830.

65. Hibbert, R. G., Mattiroli, F., and Sixma, T. K. (2009) Structural aspects of multi-domain RING/Ubox E3 ligases in DNA repair, *DNA Repair (Amst) 8*, 525-535.

66. Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H. D., Mayer, T. U., and Jentsch, S. (1999) A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly, *Cell 96*, 635-644.

67. Imai, Y., Soda, M., Hatakeyama, S., Akagi, T., Hashikawa, T., Nakayama, K. I., and Takahashi, R. (2002) CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity, *Mol Cell 10*, 55-67.

68. Hoppe, T., Cassata, G., Barral, J. M., Springer, W., Hutagalung, A. H., Epstein, H. F., and Baumeister, R. (2004) Regulation of the myosin-directed chaperone UNC-45 by a novel E3/E4-multiubiquitylation complex in C. elegans, *Cell 118*, 337-349.

69. Janiesch, P. C., Kim, J., Mouysset, J., Barikbin, R., Lochmuller, H., Cassata, G., Krause, S., and Hoppe, T. (2007) The ubiquitin-selective chaperone CDC-48/p97 links myosin assembly to human myopathy, *Nat Cell Biol 9*, 379-390.

70. Kaneko-Oshikawa, C., Nakagawa, T., Yamada, M., Yoshikawa, H., Matsumoto, M., Yada, M., Hatakeyama, S., Nakayama, K., and Nakayama, K. I. (2005) Mammalian E4 is required for cardiac development and maintenance of the nervous system, *Mol Cell Biol 25*, 10953-10964.

71. Mack, T. G., Reiner, M., Beirowski, B., Mi, W., Emanuelli, M., Wagner, D., Thomson, D., Gillingwater, T., Court, F., Conforti, L., Fernando, F. S., Tarlton, A., Andressen, C., Addicks, K., Magni, G., Ribchester, R. R., Perry, V. H., and Coleman, M. P. (2001) Wallerian degeneration of injured axons and synapses is delayed by a Ube4b/Nmnat chimeric gene, *Nat Neurosci 4*, 1199-1206.

72. Okumura, F., Hatakeyama, S., Matsumoto, M., Kamura, T., and Nakayama, K. I. (2004) Functional regulation of FEZ1 by the U-box-type ubiquitin ligase E4B contributes to neuritogenesis, *J Biol Chem 279*, 53533-53543.

73. Hanzelmann, P., Stingele, J., Hofmann, K., Schindelin, H., and Raasi, S. The yeast E4 ubiquitin ligase Ufd2 interacts with the ubiquitin-like domains of Rad23 and Dsk2 via a novel and distinct ubiquitin-like binding domain, *J Biol Chem 285*, 20390-20398.

74. Varadan, R., Assfalg, M., Raasi, S., Pickart, C., and Fushman, D. (2005) Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain, *Mol Cell 18*, 687-698.

75. Husnjak, K., Elsasser, S., Zhang, N., Chen, X., Randles, L., Shi, Y., Hofmann, K., Walters, K. J., Finley, D., and Dikic, I. (2008) Proteasome subunit Rpn13 is a novel ubiquitin receptor, *Nature 453*, 481-488.

76. Walters, K. J., and Zhang, N. (2008) Rpn10 protects the proteasome from Dsk2, *Mol Cell 32*, 459-460.

77. Richly, H., Rape, M., Braun, S., Rumpf, S., Hoege, C., and Jentsch, S. (2005) A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting, *Cell 120*, 73-84.

78. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego.

79. Wuthrich, K. (1990) Protein structure determination in solution by NMR spectroscopy, *J Biol Chem 265*, 22059-22062.

80. Takeuchi, K., and Wagner, G. (2006) NMR studies of protein interactions, *Curr Opin Struct Biol 16*, 109-117.

81. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J Mol Biol 319*, 209-227.

82. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *J Comput Chem 26*, 1668-1688.

83. Benesch, J. L., Ruotolo, B. T., Simmons, D. A., and Robinson, C. V. (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics, *Chem Rev 107*, 3544-3567.

84. Cole, J. L., Lary, J. W., T, P. M., and Laue, T. M. (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium, *Methods Cell Biol 84*, 143-179.

85. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling, *Biophys J 78*, 1606-1619.

86. Brown, P. H., and Schuck, P. (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation, *Biophys J 90*, 4651-4661.

87. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information, *J Am Chem Soc 125*, 1731-1737.

88. Hochstrasser, M. (2004) Ubiquitin signalling: what's in a chain?, *Nat Cell Biol 6*, 571-572.

89. Pickart, C. M., and Eddins, M. J. (2004) Ubiquitin: structures, functions, mechanisms, *Biochim Biophys Acta 1695*, 55-72.

90. Kim, I., and Rao, H. (2006) What's Ub chain linkage got to do with it?, *Sci STKE 2006*, pe18.

91. Ballinger, C. A., Connell, P., Wu, Y., Hu, Z., Thompson, L. J., Yin, L. Y., and Patterson, C. (1999) Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions, *Mol Cell Biol 19*, 4535-4545.

92. Mudgil, Y., Shiu, S. H., Stone, S. L., Salt, J. N., and Goring, D. R. (2004) A large complement of the predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase family, *Plant Physiol 134*, 59-66.

93. Xu, Z., Kohli, E., Devlin, K. I., Bold, M., Nix, J. C., and Misra, S. (2008) Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes, *BMC Struct Biol 8*, 26.

94. Kim, I., Mi, K., and Rao, H. (2004) Multiple interactions of rad23 suggest a mechanism for ubiquitylated substrate delivery important in proteolysis, *Mol Biol Cell 15*, 3357-3365.

95. Andersen, P., Kragelund, B. B., Olsen, A. N., Larsen, F. H., Chua, N. H., Poulsen, F. M., and Skriver, K. (2004) Structure and biochemical function of a prototypical Arabidopsis U-box domain, *J Biol Chem 279*, 40053-40061.

96. Xu, Z., Devlin, K. I., Ford, M. G., Nix, J. C., Qin, J., and Misra, S. (2006) Structure and interactions of the helical and U-box domains of CHIP, the C terminus of HSP70 interacting protein, *Biochemistry 45*, 4749-4759.

97. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics 23*, 2947-2948.

98. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics 15*, 305-308.

99. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc Natl Acad Sci U S A 98*, 10037-10041.

100. Dominguez, C., Bonvin, A. M., Winkler, G. S., van Schaik, F. M., Timmers, H. T., and Boelens, R. (2004) Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches, *Structure 12*, 633-644.

101. Matsumoto, M., Yada, M., Hatakeyama, S., Ishimoto, H., Tanimura, T., Tsuji, S., Kakizuka, A., Kitagawa, M., and Nakayama, K. I. (2004) Molecular clearance of ataxin-3 is regulated by a mammalian E4, *Embo J 23*, 659-669.

102. Lee, Y. T., Dimitrova, Y. N., Schneider, G., Ridenour, W. B., Bhattacharya, S., Soss, S. E., Caprioli, R. M., Filipek, A., and Chazin, W. J. (2008) Structure of the S100A6 complex with a fragment from the C-terminal domain of Siah-1 interacting protein: a novel mode for S100 protein target recognition, *Biochemistry 47*, 10921-10932.

103. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J., and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems, *Biophys J 82*, 1096-1111.

104. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J Biomol NMR 6*, 277-293.

105. Goddard, T. D., and Kneller, D. G. (2006) SPARKY 3, University of California, San Francisco.

106. Mulder, F. A., Schipper, D., Bott, R., and Boelens, R. (1999) Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins, *J Mol Biol 292*, 111-123.

107. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J Mol Biol 273*, 283-298.

108. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR 13*, 289-302.

109. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res 35*, W375-383.

110. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J Biomol NMR 8*, 477-486.

111. Hubbard, S., and Thornton, J. (1992-96) NACCESS, University of Manchester, U.K.

112. Hochstrasser, M. (2006) Lingering mysteries of ubiquitin-chain assembly, *Cell 124*, 27-34.

113. Winget, J. M., and Mayor, T. (2010) The diversity of ubiquitin recognition: hot spots and varied specificity, *Mol Cell 38*, 627-635.

114. Welchman, R. L., Gordon, C., and Mayer, R. J. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals, *Nat Rev Mol Cell Biol 6*, 599-609.

115. Li, W., and Ye, Y. (2008) Polyubiquitin chains: functions, structures, and mechanisms, *Cell Mol Life Sci 65*, 2397-2406.

116. Aravind, L., and Koonin, E. V. (2000) The U box is a modified RING finger - a common domain in ubiquitination, *Curr Biol 10*, R132-134.

117. Hatakeyama, S., and Nakayama, K. I. (2003) U-box proteins as a new family of ubiquitin ligases, *Biochem Biophys Res Commun 302*, 635-645.

118. Kamadurai, H. B., Souphron, J., Scott, D. C., Duda, D. M., Miller, D. J., Stringer, D., Piper, R. C., and Schulman, B. A. (2009) Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex, *Mol Cell 36*, 1095-1102.

119. Nordquist, K. A., Dimitrova, Y. N., Brzovic, P. S., Ridenour, W. B., Munro, K. A., Soss, S. E., Caprioli, R. M., Klevit, R. E., and Chazin, W. J. (2010) Structural and functional characterization of the monomeric U-box domain from E4B, *Biochemistry 49*, 347-355.

120. Pruneda, J. N., Stoll, K. E., Bolton, L. J., Brzovic, P. S., and Klevit, R. E. (2011) Ubiquitin in Motion: Structural studies of the E2~Ub conjugate, *Biochemistry 50*, 1624-1633.

121. Rückert, M., and Otting, G. (2000) Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments, *J. Am. Chem. Soc. 122*, 7793–7797.

122. Bruschweiler, R. (2003) New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins, *Curr Opin Struct Biol 13*, 175-183.

123. Svergun, D. I. (2007) Small-angle scattering studies of macromolecular solutions., *J. Appl. Cryst. 40*, s10-s17.

124. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, *Q Rev Biophys 40*, 191-285.

125. Ardley, H. C., and Robinson, P. A. (2004) The role of ubiquitin-protein ligases in neurodegenerative disease, *Neurodegener Dis 1*, 71-87.

126. Otting, G., Ruckert, M., Levitt, M. H., and Moshref, A. (2000) NMR experiments for the sign determination of homonuclear scalar and residual dipolar couplings, *J Biomol NMR 16*, 343-346.

127. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra, *J Magn Reson 131*, 373-378.

128. Nakatsukasa, K., Huyer, G., Michaelis, S., and Brodsky, J. L. (2008) Dissecting the ER-associated degradation of a misfolded polytopic membrane protein, *Cell 132*, 101-112.

129. Spinette, S., Lengauer, C., Mahoney, J. A., Jallepalli, P. V., Wang, Z., Casciola-Rosen, L., and Rosen, A. (2004) Ufd2, a novel autoantigen in scleroderma, regulates sister chromatid separation, *Cell Cycle 3*, 1638-1644.

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

131. Wider, G. (2005) NMR techniques used with very large biological macromolecules in solution, *Methods Enzymol 394*, 382-398.

132. Kleiger, G., Saha, A., Lewis, S., Kuhlman, B., and Deshaies, R. J. (2009) Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates, *Cell 139*, 957-968.

133. Driscoll, J. J., and Dechowdhury, R. (2010) Therapeutically targeting the SUMOylation, Ubiquitination and Proteasome pathways as a novel anticancer strategy, *Target Oncol 5*, 281-289.

134. Chou, T. F., Brown, S. J., Minond, D., Nordin, B. E., Li, K., Jones, A. C., Chase, P., Porubsky, P. R., Stoltz, B. M., Schoenen, F. J., Patricelli, M. P., Hodder, P., Rosen, H., and Deshaies, R. J. (2011) Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways, *Proc Natl Acad Sci U S A*.

135. Chou, T. F., and Deshaies, R. J. (2011) Quantitative cell-based protein degradation assays to identify and classify drugs that target the ubiquitin-proteasome system, *J Biol Chem*.

136. Wu, W. K., Cho, C. H., Lee, C. W., Wu, K., Fan, D., Yu, J., and Sung, J. J. (2010) Proteasome inhibition: a new therapeutic strategy to cancer treatment, *Cancer Lett 293*, 15-22.

137. Shi, D., and Grossman, S. R. (2010) Ubiquitin becomes ubiquitous in cancer: Emerging roles of ubiquitin ligases and deubiquitinases in tumorigenesis and as therapeutic targets, *Cancer Biol Ther 10*.

138. Bohm, S., Lamberti, G., Fernandez-Saiz, V., Stapf, C., and Buchberger, A. (2011) Cellular functions of Ufd2 and Ufd3 in proteasomal protein degradation depend on Cdc48 binding, *Mol Cell Biol 31*, 1528-1539.

139. Fernandez-Saiz, V., and Buchberger, A. (2010) Imbalances in p97 co-factor interactions in human proteinopathy, *EMBO Rep 11*, 479-485.

140. Hanzelmann, P., Stingele, J., Hofmann, K., Schindelin, H., and Raasi, S. (2010) The yeast E4 ubiquitin ligase Ufd2 interacts with the ubiquitin-like domains of Rad23 and Dsk2 via a novel and distinct ubiquitin-like binding domain, *J Biol Chem 285*, 20390-20398.