ABSTRACT

Title of Dissertation: PRINCIPLES OF COMPLEX POLYUBIQUITIN SIGNALING, THE STRUCTURAL BASIS FOR UBIQUITIN-UBISTATIN INTERACTIONS, AND NOVEL ASSAYS FOR THE CHARACTERIZATION OF DEUBIQUITINASES

Mark A. Nakasone, Doctor of Philosophy, 2013

Directed By: Professor David Fushman
Department of Chemistry and Biochemistry
Center for Biomolecular Structure and Organization

Ubiquitination is the most versatile and is certainly one of the most difficult post-translation modifications to understand in eukaryotic life. In the process of ubiquitination the C-terminus of ubiquitin (Ub), a small 8.65 kDa protein is covalently attached to εNH_2 groups of lysine side chains on target proteins. Once attached, additional Ubs can be added to the original Ub at eight unique linkage sites (M1, K6, K11, K27, K29, K33, K48, or K63) to form polyUb chains. This internal Ub-Ub linkage dictates the structural conformation of the polyUb chain, which in turn governs the receptors that can recognize a given chain. PolyUb chains were thought to be homogeneously linked until very recently when mixed linkage polyUb chains were detected on several cellular pathways. This observation implied that instead of having just eight distinct polyUb signals, there were now potentially quadrillions of

unique chains. The results presented within represent the first in depth studies of mixed linkage polyUb chains, focusing on the structural impact of linkage mixing. For mixed K48 and K63 linked chains the findings support that their individual linkage properties are preserved regardless of linkage mixing. However, simulations for mixed linkage chains containing different linkages imply that many novel polyUb signals are possible.

The ubiquitin-proteasome pathway is the primary mechanism to degrade short lived proteins in the cell and has also emerged as a top therapeutic target. Ubistatins, a class of small molecules bring about the same effects as existing proteasome inhibitor drugs by directly binding the polyUb chain. However, virtually nothing is known about the structural properties for any ubistatin/Ub complex. Here is provided the first structure of a ubistatin/Ub complex along with data that overwhelmingly validates the structure. Other important factors regarding the ubistatin/Ub interaction including the stoichiometry and dual hydrophobic / electrostatic binding mechanism are also uncovered.

Proteomic analysis of polyUb conjugates has been limited to determining which linkage types are present. A novel method for K63 linked polyUb conjugates, which can measure consecutive K63 linkages is described here. This method allows the proteomics community to gain unprecedented information on cellular pathways utilizing K63 linkages.

PRINCIPLES OF COMPLEX POLYUBIQUITIN SIGNALING, THE STRUCTURAL BASIS FOR UBIQUITIN-UBISTATIN INTERACTIONS, AND NOVEL ASSAYS FOR THE CHARACTERIZATION OF DEUBIQUITINASES

By

Mark A. Nakasone

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor David Fushman, Chair
Professor Dorothy Beckett
Professor Steven Rokita
Professor Vitali Tugarinov
Professor Zvi Kelman - Dean's Representative

Preface

The expanding field of ubiquitin biology has passed many milestones since its discoverers conducted their first experiments in the late 1970's. Decades later, at this present time, the significance of the ubiquitin-proteasome system and other ubiquitin mediated signaling pathways have been noted by researchers the world over. The sheer number of cellular pathways and processes that utilize ubiquitin are simply astounding, however this has triggered many more researchers to enter the field and blaze trails to new frontiers. Moreover, basic research focused on understanding the molecular details of how different components on ubiquitin signaling pathways function has led to numerous therapeutic targets for an array of human diseases. Specifically, Velcade and subsequent second-generation small molecule proteasome inhibitors have turned the tide in several different malignancies. Therapeutics for many other targets on ubiquitin pathways are currently in testing and expected to have a great impact across a range of diseases. Many others outside of the scientific community, such as investors of all stripes have also recognized the potential ubiquitin signaling holds. As a result much capital has poured into small pharmaceutical companies focused on therapeutics targeting ubiquitin pathways and larger companies are creating whole divisions for development of similar therapeutics or simply placing high bids to buy up smaller companies that hold promising technologies (e.g. the 2009 acquisition of Proteolix, Inc by Onyx Pharmaceuticals).

Despite the high profile of ubiquitin research, the most significant outcome is not the large amount of money invested, jobs created, the ever increasing number of

research grants, or even the therapeutics, it is the intellectual revolution which has led to an exponential increase in publications spanning across many disciplines for the past decades. Since the initial discovery of ubiquitin there have been around 37,000 published research articles and 2,200 structures and models deposited in the Protein Data Bank (roughly 14% of all known biomolecular structures) pertaining to some aspect of ubiquitin biology. Each year brings more publications than the last, a trend that has continued exponentially for over three decades with 2012 yielding 3,511 publications on ubiquitin biology, compared to 3,167 in 2011, 982 in 2000, and just 35 in 1985. It is an undeniable fact that ubiquitin biology is an established field, however there is an argument as to whether the field has matured. Given the increasing number of publications and consistent stream of new methodology, most would conclude that there is still much more room for growth. On the other hand, the large number of researchers and technical experts that have entered the field can make publishing on certain areas extremely competitive. Yet, overall this is not a drawback for conducting research in the ubiquitin field since there are more than enough areas that remain undiscovered or which are poorly understood for anyone, veteran or newcomer to explore. More and more is being discovered about the importance and versatility of ubiquitin signaling every day and the future seems almost limitless.

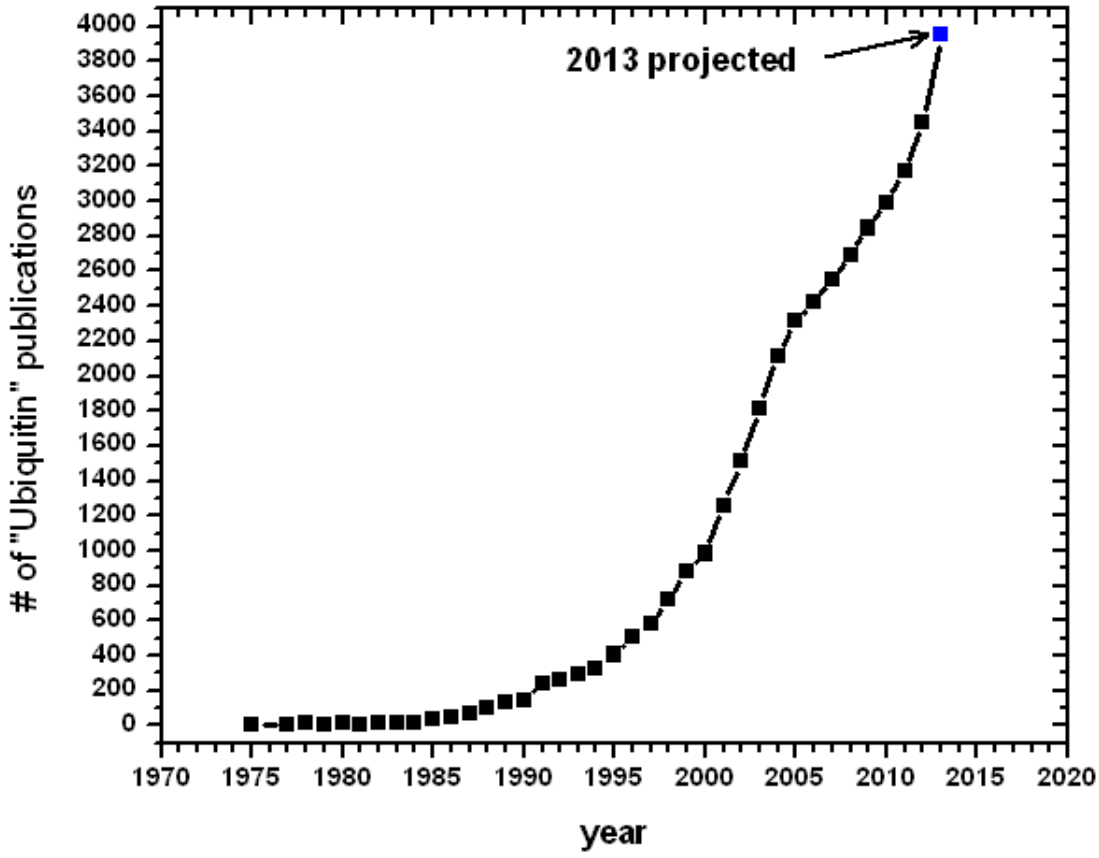


Figure 0.1 – Number of publications on ubiquitin vs. time.
Since the discovery of ubiquitin, there have been over 37,000 publications relating to it. The above plots the number of publications on ubiquitin for each year since 1975. The data were obtained by using "ubiquitin" as keyword search in the PubMed, a service from the National Library of Medicine (NLM), a division of the National Institutes of Health (NIH).

Although cell biologists with basic biochemical skills initially set the ubiquitin field in motion, other more sophisticated methods have at times eclipsed some of the more traditional biochemical methods in respect to the excitement they spark. Most notable are the advances in structural biology, synthetic chemistry, and proteomics. Solution NMR has always used monomeric ubiquitin as a model protein, due to its favorable spectral properties and extremely long shelf life, however NMR methods have leaped up a few steps. Lewis Kay and Vitali Tugarinov previously recorded

NMR spectra for malate synthase G (MSG), an 82-kDa protein giving them the distinction of largest molecule every approached with NMR. Shortly after in 2008, Lewis Kay attempted the 20S proteasome core with a similar NMR approach used for MSG. This not only allowed him to surpass his previous record of studying the largest molecule by NMR, but also allowed him to use the method to determine how substrates enter the proteasome core and how small molecule drugs interact with individual subunits of the proteasome. Several leading labs specializing in cryoEM then approached the whole 26S proteasome and independently reported near atomic resolution models for the whole 2.5 MDa proteasome in late 2012. This finding attracted much attention and although it was not the largest structure solved with cryoEM it was a great advance towards determining the structure of the 26S proteasome, which is also known as "the most complicated protease in nature." Aside from the structural biologists, many other leading experts from around the world have devoted their efforts to ubiquitin research. The bioorganic and synthetic chemists are relatively new to the field, but the reagents available through their expertise have allowed others to make unprecedented advances in characterizing ubiquitin-signaling pathways. One of the leading synthetic chemists, Ashraf Brik holds the distinction of synthesizing the longest peptide in addition to all of the other technologies he has developed for the ubiquitin field. Many other synthetic chemists are entering the field and their unique technologies will ultimately benefit many researchers in addition to bringing notoriety to their methods. Several mass spectrometry leaders studying the ubiquitin modification have arguably conducted many landmark studies in proteomics over the last decade. Utilizing experts from

bioinformatics, researchers specializing in mass spectrometry have been able to report the details of ubiquitin modification in many systems and reveal that ubiquitin modification is in fact ubiquitous for a majority of proteins in eukaryotic cells. This has induced other world-class proteomics labs to enter the ubiquitin field, a trend that is welcome given the complexity of ubiquitin modification.

In 1993 I had my first laboratory experience at the University of Maryland in the lab of Prof. Scott Angle. Over a decade later, in 2007 I returned to the University of Maryland as a doctoral student in biochemistry. After several laboratory rotations, I selected the laboratory of Prof. David Fushman to conduct my doctoral research. When I first started, 2D NMR spectra were confusing dots I could not interpret, I was completely dependent on others to perform basic laboratory functions, and I understood virtually nothing about ubiquitin biology let alone its structural properties. Over my doctoral career with the careful guidance of Prof. David Fushman I developed to a point where I could independently carry out NMR experiments, design my own projects, and where I was frequently consulted as an expert on various aspects of ubiquitin biology. Now in the present year, 2013 my doctoral career at the University of Maryland is concluding and I am leaving behind a body of work in the shadow of ~37,000 ubiquitin focused publications. The field has become increasingly more technical and at times I have been dazed by the sheer amount of information. However, through much effort I was able to design and carry out experiments that answered many pressing questions and opened new insights to ubiquitin biology.

During my doctoral career I was constantly advancing many diverse projects simultaneously and was fortunate to learn from them all. My main projects included the first structural studies of mixed linkage polyubiquitin chains, determining the structural properties of ubiquitin/ubistatin (a class of small molecules) interactions, and the characterization of proteasomal deubiquitinases. In addition to these projects I created a NMR method to measure the activity of deubiquitinases, used the F45W ubiquitin mutant for several fluorescence assays, and developed new methods to make novel polyubiquitin chains. In collaboration with Prof. Catherine Fenselau and her graduate student Joseph R. Cannon, I was able to demonstrate a technique to measure the number of consecutive K63-linkages in polyubiquitin chains, a feat that has never been accomplished before, but provides critical information for understanding several signaling pathways. My collaboration with Prof. TingTing Yao and Prof. Robert E. Cohen allowed me to determine the structural properties of ubiquitin modified histones. Work with Ming-Yih Lai on H68 ubiquitin mutants and Carlos A. Castaneda on K11-linked polyubiquitin allowed me to establish how deubiquitinases select catalytically favorable conformations of polyubiquitin chains. A study carried out with Prof. Michael Glickman and Dasha Krutauz where I provided polyubiquitinated Ubb(+1) established that inhibiting the activity of deubiquitinases contributes to neurodegenerative diseases. With the work of my collaborators and myself, I present my dissertation, which details how my doctoral work expands our understanding of ubiquitin biology and the true significance our work has among the galaxy of publications in the field.

Acknowledgements

The dissertation herein is the culmination of my major findings since I joined the lab of Prof. David Fushman in January of 2008. The writing requirements for this work strictly constrain me to discuss technical scientific topics, however some of my most important findings were the experiences and professional relationships I had with members of David Fushman's lab, the University of Maryland community, researchers outside of our institution, and also those of late. In this section I will briefly describe how many individuals from diverse scientific backgrounds facilitated the path to this dissertation and also how they helped me survive as a researcher.

My greatest acknowledgment is to my advisor, Prof. David Fushman who trained me as an NMR spectroscopist, mentored me through five years of intense research, and who has routinely made extraordinary efforts to ensure my scientific development. Utilizing his talents as a world-class NMR spectroscopist and as an authority on ubiquitin signaling, Prof. Fushman created the perfect environment for conducting cutting edge research in several interdisciplinary fields. Over my career Prof. Fushman's lab continued to grow and I was honored to provide data for our grants and also help shape the lab with my input as we evolved to accommodate larger projects. The diversity of the technical skills perfected in Prof. Fushman's lab is simply astounding and spans biomolecular NMR, quantum physics, high performance computing, solution scattering, organic synthesis, X-ray crystallography, and traditional biochemical methods. My only regret is that I did not have enough time to master all of the skills and I am still in awe of Prof. Fushman's seemingly superhuman ability to acquire so many advanced technical skills. I am also grateful

that Prof. Fushman spent the time to carefully edit and comment my written work, which dramatically improved my scientific writing, but cost him many hours of his time over the years. Aside from the research training Prof. Fushman has also aided me greatly in developing networking skills and generously sent me to four international conferences to present my findings, allowing me to create a web of collaborators from around the world in the process. Outside of the lab, Prof. Fushman also encouraged me to take on leadership roles and to organize events for the CBSO and biochemistry, which exposed me to other biochemical techniques and experts in different subfields of biochemistry. One of the major reasons I am in Prof. Fushman's debt is that he took a major gamble and allowed me to pursue my own original projects simply based on my own scientific curiosity. These projects include the structural characterization of mixed-linkage ubiquitin chains, a method to determine deubiquitinase kinetics by NMR, fluorescence assays with the F45W ubiquitin mutant, decoding the activity of proteasomal deubiquitinases, and designing model substrates for deubiquitinases. Without Prof. Fushman taking a major risk and granting me free rein to independently carry out my projects I would have never developed my critical thinking skills.

Since my 2007 rotation in the Fushman lab as a first year graduate student I had the opportunity to interact with over twenty lab members. Ming-Yih Lai was officially the first lab member I would work with and he helped guide me through my rotation focused on ubiquitin/ubistatin interactions. Ming-Yih and I have worked alongside each other since 2008 and he is also a collaborator on one publication. Daoning Zhang has been with me every step of the way, guiding me in the lab,

working as a collaborator, and most importantly his great help with the acquisition and analysis of NMR data. Without the expertise of Daoning, I would not be the researcher I am today. Tony Chen and Shirley Lee were two very talented undergraduate students in the lab before I joined. Together they aided me in designing assays, preparing reagents, optimizing protocols, and ultimately helped me get my start in the lab. Although never formal collaborators Adithya Sundar and Dr. Rajesh Singh who shared similar projects offered me some of the most stimulating scientific discussion and were also extremely helpful for obtaining many reagents. Over the years Raquel Godoy-Ruiz and Tanuja R. Kashyap would also help me acquire reagents and equipment even though we never shared any related projects. I had the opportunity to mentor Andrew Timmons, an undergraduate who used his new skills to help me prepare important reagents and pioneer pull down assays with polyubiquitin chains. Working with Dr. Carlos A. Castaneda as a collaborator for the K11-linked ubiquitin project was a rewarding experience and exposed me to new techniques. Urszula K. Nowicka, who joined the lab at the same time as me, helped me analyze NMR experiments and always was open to discuss topics regarding our research. Konstantin Berlin or just "Dr. Berlin" as he is known, is a computer scientist that taught me the value of scientific computing, the best ways to implement my own programs, and the importance of selecting the proper methods for error analysis. Overall, the personnel Prof. Fushman selected for his lab complemented each other well and created a productive environment with satisfying day-to-day experiences. We are all in debt to each other for our combined efforts to advance our

research and I acknowledge all lab members past and current for aiding me all of these years.

The University of Maryland has proven to be an excellent place to conduct collaborative interdisciplinary research and I have been very fortunate to work with Prof. Catherine Fenselau's lab on several projects. In 2011 Prof. Catherine Fenselau and her graduate student Joseph R. Cannon, also a personal friend suggested we attempt microwave assisted acid cleavage of polyubiquitin chains. I recognized that this method or any ASP specific cleavage could be used to measure consecutive K63 linkages and we were able to publish this exciting result in 2012. In addition, I was also able to learn much about several mass spectrometry methods and how to analyze the data. The two resident biomolecular crystallographers Prof. Nicole LaRonde-LeBlanc and Prof. Paul J. Paukstelis have been extremely helpful in my X-ray crystallography pursuits, lending their time and also their resources. Prof. Paukstelis kindly mentored me through every step of my attempts to crystallize ubiquitin/ubistatin complexes and Ub mutants, which took up a significant amount of his valuable time and I am very grateful for everything I got to learn from him. Prof. Dorothy Beckett also opened her lab to me and allowed me to investigate binding interactions in early stages of some of my projects. Prof. Beckett was also extremely supportive in graduate student initiatives and excellent for discussing topics in biophysics. Prof. George H. Lorimer generously opened his lab to me and allowed to explore fluorescence assays, specifically for the F45W ubiquitin project. Inside of biochemistry all of the faculty were extremely supportive of graduate students and always willing to provide their advice on conducting experiments and how to

interrupt data. Our biomolecular NMR program has expanded greatly since I joined and I was fortunate to have Prof. Vitali Tugarinov who provided me with much advice for NMR experiments, in particular ^{13}C based experiments. Teaching faculty like Dr. Daniel Steffek did their best to facilitate the TA assignments of graduate students and I am very grateful for his efforts during my tenure as a TA.

I would also like to acknowledge my former biochemistry instructor Prof. Melanie R. Nilsson at McDaniel College for her rigorous course work and mentoring my undergraduate research, which prepared me for graduate school. If my late undergraduate advisor Prof. David W. Herlocker were still alive to see me graduate I would thank him for guiding me to graduate school and all of his helpful advice. Additionally, I'd also like to acknowledge Prof. Michael Rosenthal, a former physical chemistry instructor for introducing me to his friend Prof. Michael P. Doyle who was kind enough to take the time and recruit me on his visit to McDaniel College in 2007.

The collaborations I have had with other institutions have been invaluable and greatly increased the significance of my research. Collaborators Prof. Robert E. Cohen (Colorado State) and Prof. Tingting Yao (Colorado State) aided me greatly with the study of mixed linkage ubiquitin chains. They also sent me key plasmids and we also carried out a small project focused on the properties of ubiquitinated histones. I am in a great debt to Prof. Olivier Walker (Université Claude Bernard Lyon 1, France) who took some of his vacation time during the summers of 2010 and 2011 to teach me how to perform molecular docking, convert PREs to distances, and teach me computational methods for protein/small molecule interactions. Prof. Walker is also a major collaborator for the ubistatin project and without him we

would not have our structural data. Medicinal chemist Dr. Tim Lewis (Broad Institute of Harvard and MIT) was vital for synthesizing ubistatin compounds and kindly distributing the compounds to other collaborators. Prof. Raymond Deshaies (California Institute of Technology) and Prof. Jeffery Brodsky (University of Pittsburgh) carried out the critical structural activity relation and *in vivo* assays for ubistatin compounds on the project. Prof. Cynthia Wolberger (Johns Hopkins) with her postdoctoral fellows Reuven Wiener and Christopher Berndsen were extremely helpful in optimizing the enzymatic synthesis of our polyubiquitin chains, providing reagents, and also in discussions we had at conferences. Prof. Brenda Schulman (St. Jude Children's Research Hospital) and Michael J. Eddins (Progenra, Inc) allowed me to advance in my early attempts to enzymatically synthesize polyubiquitin chains. A major competitor, Prof. David Komander (MRC Cambridge, UK) was always open to technical discussion at conferences and useful for exchanging ideas. Additionally, Prof. Gai Prag (Tel Aviv University, Israel) and Prof. Aaron Ciechanover (Technion, Israel) generously made their reagents available to the research community.

Prof. Fushman's internationally renowned research has indeed attracted many collaborators. Perhaps the most productive collaboration over the past years has been with the lab of Prof. Michael H. Glickman (Technion, Israel). Prof. Glickman has become a trusted friend and was instrumental for many of my studies regarding the ubiquitin-proteasome pathway. There were several collaborations with Prof. Glickman's lab and I would like to acknowledge Nurit Livnat-Levanon for her role in the mixed linkage ubiquitin chain project and Dasha Krutauz for allowing me to be involved with the Ubb(+1) project. Noa Reis, the lab manager was always helpful in

sending plasmids and writing concise protocols. Out of all of the lab members, I spent the most time with Wissam Mansour working on polyubiquitin processing by the proteasome. Through much adversity, Wissam visited the Fushman lab in the summer of 2012 were we discovered a new deubiquitinase in the proteasome and developed novel assays to decrypt the function of isolated proteasomal deubiquitinases. On my visit to the Technion in November of 2012, we expanded our assay to the whole proteasome and are on track to have several publications from this collaboration.

After extensively reviewing the literature and also the reagents in our lab I have come to realize that the late Prof. Cecile M. Pickart (Johns Hopkins) deserves a very special acknowledgment. Without the plasmids she created and early collaborations with Prof. Fushman, our lab would not be able to synthesize polyubiquitin chains, let alone monomeric ubiquitin. However, just mentioning Prof. Pickart's contribution to my projects is a bit demeaning. In 1982 Prof. Pickart was working under Irwin Rose at the Fox Chase Cancer Center, at the same time Ciechanover, Hershko, and Rose were credited with discovering the ubiquitin proteasome pathway. By any standard she is considered the mother of the ubiquitin field and Prof. Pickart's seminal work, which established that different ubiquitin linkages initiate distinct signaling outcomes has provided the foundation for much of my own work. In addition, many critical reagents used for my projects were personally prepared by Prof. Pickart herself and I am truly grateful she was able to share them with Prof. Fushman, before she passed away at the age of 51 on April 5,

2006. Without the decades of work from Prof. Pickart and her generosity in providing materials none of my projects would have been possible.

In February of 2013 the presidentially appointed Foreign Scholarship Board selected me for the prestigious J. William Fulbright post-doctoral research grant for study in Israel. This highly competitive award was the most significant honor of my doctoral career and I could not have achieved this without the efforts of Prof. Lorimer, Prof. Glickman, and my advisor Prof. Fushman. I would like to acknowledge them for persuading me to apply, writing excellent letters of support, and carefully reviewing my proposal.

Last, I would like to acknowledge my friends and family. Although they did absolutely nothing to contribute to my research and at times kept me from research with events such as vacations and Christmas, their moral support impacted me greatly. I never was able fully explain my research to them, but my friends and family certainly improved the quality of my life and my character.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

µg	micro gram
µL	micro liter
µM	micro molar
ADP	adenosine diphosphate
AMP	adenosine monophosphate
AMPK	AMP-activated protein kinase
AMSH	associated molecule with the SH3 domain of STAM
AMSH-LP	AMSH like protease
ANS	8-anilino-1-naphthalenesulfonic acid
ATP	adenosine-5'-triphosphate
CID	collision-induced dissociation
CP	core particle
CSN5	COP9 signalosome complex subunit 5
CSP	chemical shift pertibation
CUE	coupling of ubiquitin conjugation to ER degradation
CYLD	cylindromatosis
D_2O	deuterium oxide (heavy water)
Da	dalton (mass unit)
Ddi1	DNA damage inducible protein 1
DNA	deoxyribonucleic acid
DOSY	diffusion ordered spectroscopy
ds	double strand
DSK2	dominate suppressor of kar1 mutant-2
DTT	dithiothreitol
DUB	deubiquitinase / deubiquitinating enzyme
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
ESCRT	endosomal sorting complex required for transport
FRET	Förster resonance energy transfer
GEF	guanine nucleotide exchange factor
GPCR	G protein-coupled receptors
H_2O	water
HECT E3	homologous to the E6-AP carboxyl terminus E3 ligase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hHR23A	human homolog of RAD23
HMQC	heteronuclear multiple quantum coherence
HOIL-1	haem-oxidized iron-regulatory protein 2 ubiquitin ligase-1

HOIP	HOIL-1 interacting protein
HSQC	heteronuclear single quantum coherence
IAEDANS	5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid
INEPT	insensitive nuclei enhanced by polarization transfer
ISG15	interferon-induced 15 kDa protein
JAMM/MPN+	JAB1/MPN/MOV34 metalloenzymes
K	lysine
kD	equilibrium dissociation constant
kDa	kilodalton
LB	Luria broth
LUBAC	linear ubiquitin chain assembly complex
Lys	lysine
MAAH	microwave assisted acid hydrolysis
MALDI	matrix-assisted laser desorption/ionization
MBP	maltose-binding protein
MDa	megadalton
Mdm2	mouse double minute 2 homolog
MJD	Machado-Joseph disease protease
mM	millimolar
mRNA	messenger RNA
MS	mass spectrometry
MTSL	S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
MWCO	molecular weight cut off
Nedd8	neural precursor cell expressed, developmentally down-regulated 8
NEMO	NF-kappa-B essential modulator / IKK-γ
NF-κB	nuclear factor-κB
nm	nano meter
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhancement spectroscopy
OTUB1	OTU domain, ubiquitin aldehyde binding
OUT	ovarian tumor protease
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCI	proteasome, COP9 signalosome, initiation factor 3
PDB	protein data bank
PEG	polyethylene glycol
PMSF	phenylmethanesulfonylfluoride or phenylmethylsulfonyl fluoride
ppm	parts per million
PRE	paramagnetic relaxation enhancement
Pru	Pleckstrin-like receptor for ubiquitin

RAP80	receptor associated protein 80
RING E3	really interesting new gene E3 ligase
RNA	ribonucleic acid
RP	regulatory particle
rpm	revolutions per minute
RTK	receptor-tyrosine kinase
SANS	small angle neutron scattering
SAXS	small angle X-ray scattering
SDS	sodium dodecyl sulfate
smFRET	single molecule FRET
sofast	selective optimized-flip-angle short transient
SUMO	SMT3 suppressor of mif two 3 homolog
TCEP	tris(2-carboxyethyl)phosphine
TLCK	tosyl lysine chloromethyl ketone hydrochloride
TOCSY	total correlation spectroscopy
TRAF6	TNF receptor associated factor-6
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
tRNA	transfer RNA
Ub	ubiquitin
UBA	ubiquitin-associated domain
UBD	ubiquitin binding domain
UBL	ubiquitin like
UCH	ubiquitin C-terminal hydrolase
UIM	ubiquitin interacting motif
UPP	ubiquitin-proteasome pathway
USP	ubiquitin specific protease
VWA	von Willebrand factor A
WT	wild type
β-grasp	beta grasp

Chapter 1: Introduction to polyubiquitin signaling

1.1 Introduction to ubiquitin biology

Early studies in the biological sciences focused on understanding how proteins were created in the cell and there was little attention directed at determining how proteins were removed at the end of their lifetimes. Decades later, studies of protein turnover would fill this void in knowledge, when the issue was starting to be addressed in the mid 1950's. Although protein turnover was probed with a range of techniques and in a variety of different organisms, seemingly unrelated studies reported the presence of a highly conserved polypeptide (*1-3*). The universal presence of this polypeptide in eukaryotes led to its name, "ubiquitin" (Ub). In a series of papers Ciechanover, Hershko and Rose established the foundation for a degradation pathway of cellular proteins that was dependent on both ATP and the Ub molecule (originally called APF-1) (*1, 3-11*). This work was the first to clearly layout what would later be named the ubiquitin-proteasome pathway (UPP), but more importantly provided a mechanism for explaining protein turnover in the cell (*12*). For this contribution Ciechanover, Hershko and Rose were awarded the 2004 Nobel Prize in chemistry, "discovery of ubiquitin-mediated protein degradation." Evidently, the philosophy of Hershko and Ciechanover, "to work where no one else was working" paid off. Since the initial discovery of the UPP, Ub mediated signaling has also been found to be involved with regulating many other diverse cellular processes.

In the cell Ub is present in a multitude of oligomeric forms and is commonly post-translationally added to other proteins as a monomeric unit or as a polyUb chain.

Although the same unit (Ub) is added in the case of polyubiquitination, the signaling outcomes that result from modification with polyUb are extremely broad due to the number of distinct signals that can be created by specific Ub-Ub linkages in a given polyUb chain. Typically the E1, E3, E3 enzyme cascade catalyzes the formation of an isopeptide bond between the C-terminus of an Ub and the ε-amino group of a lysine on a target protein, a process known as ubiquitination. Once the first Ub is added to a target protein, subsequent Ubs can be added to the previously ligated Ub at eight different positions; K6, K11, K27, K29, K33, K48, K63 or the N-terminus "M1." The general principle of how Ub is linked to create isopeptide bonds is represented in **(Figure 1.1)**. It is thought that once a particular linkage is formed subsequent Ubs will maintain this linkage in the polyUb chain (*13*). Once attached to a target protein, the Ub or polyUb signal can be recognized by different receptors which ultimately determine the fate of the ubiquitinated protein (*14*). For example, attachment of a single Ub to a target protein generally signals for its transport across cell membranes, while attachment of four or more Ubs internally linked through K48 to the same protein would signal for its destruction by the 26S proteasome (*15, 16*). Modification with Ub is reversible and a special class of enzymes termed deubiquitinases (DUBs) can remove Ub from a target protein and also fully reduce polyUb to its component monomeric units (*17, 18*). Under normal conditions the ubiquitination state of a given protein is tightly regulated, but dynamic due to the opposing actions of the ubiquitinating enzymes (E1, E2, E3) against DUBs.

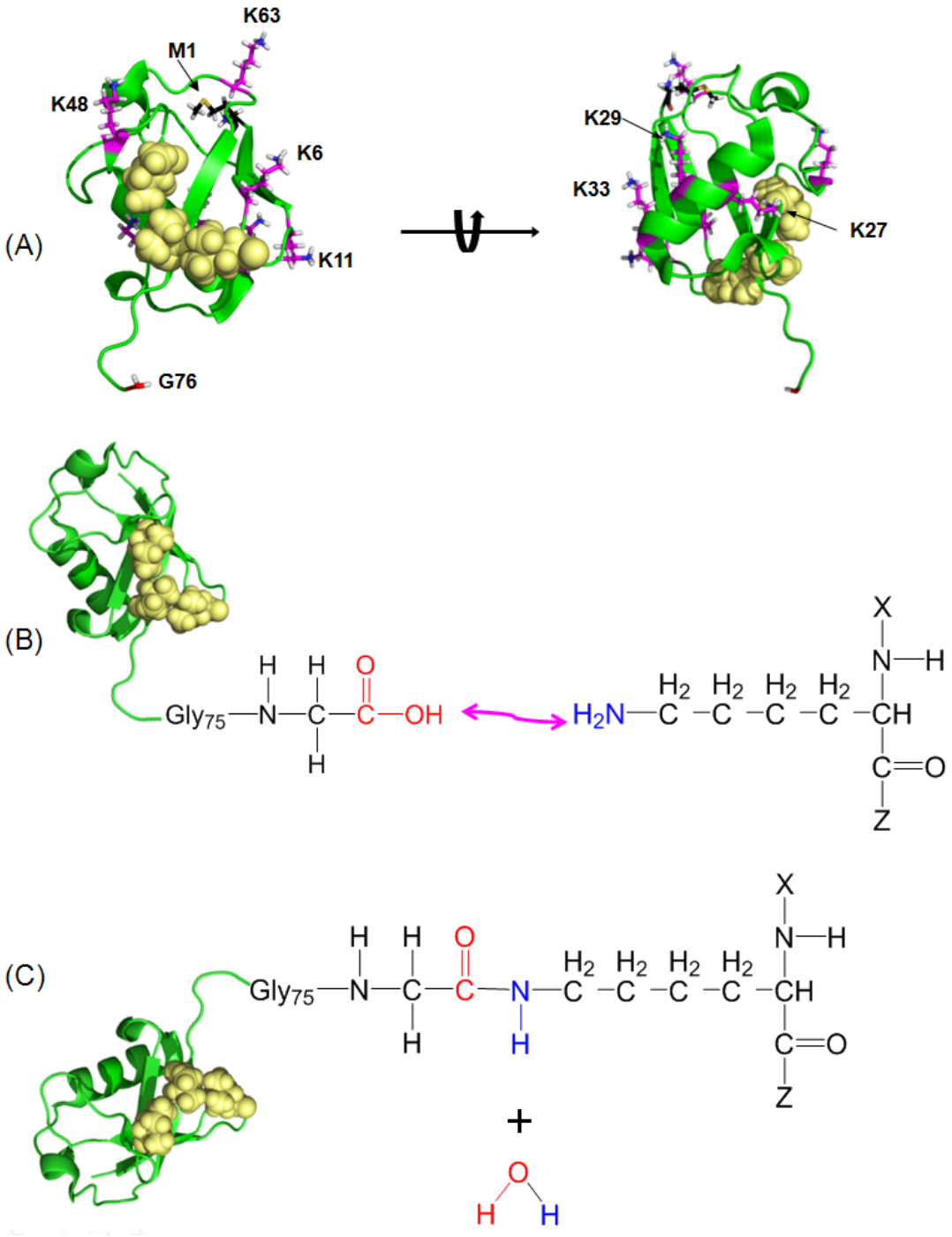


Figure 1.1 – Critical residues of the Ub molecule involved in isopeptide bond formation.
(**A**) Cartoon representation of Ub with lysine residues (magenta sticks), M1 (black sticks), and the L8,I44,V70 hydrophobic patch (yellow spheres). (**B**) The free C-terminus of G76 in Ub is used to form an isopeptide bond with the ε-amine of a lysine side chain. (**C**) Formation of an isopeptide bond between G76 of Ub and a lysine side chain

1.2 Enzymatic modification with ubiquitin

Ubiquitination of a target protein is accomplished by the actions of three enzymes; E1 activating, E2 conjugating, and E3 ligase. In humans there is only one E1 activating enzyme, around thirty-five E2 conjugating enzymes and hundreds (>630) of E3 ligases (*19*). Generally E1s can directly interact with virtually all E2s, while E2/E3 interactions are slightly more selective, but still very promiscuous (*20*). Ubiquitination of a specific "substrate" protein is achieved by its corresponding E3 ligase which is extremely selective for that particular substrate protein (*21*). The structural details of how an individual E3 ubiquitinates its target protein are unknown, however the mechanisms by which an E3 attaches Ub to a target have been classified in two sub-groups, "homologous to the E6-AP carboxyl terminus" (HECT) and "really interesting new gene" (RING) based on sequence homology. Regardless of the E3 type, in the first step of ubiquitination E1 hydrolyzes ATP to generate the energy needed to form a high energy thioester bond between its active site cysteine and the C-terminus of Ub. Next, the Ub "loaded" E1 binds an E2 which allows the Ub from its active site to be transferred to an active site cysteine on the E2 resulting in the same thioester bond, now on the E2. Next, if a HECT E3 is involved, the HECT E3 will simultaneously bind its substrate and the Ub loaded E2, transfer the Ub to an active site cysteine on itself, and finally catalyze the formation of an isopeptide bond between the C-terminus of Ub and a free amine group on the substrate protein. While a RING E3 will produce the same outcome for the substrate protein, but instead this is accomplished by the RING E3 positioning the Ub loaded E2 close enough to the ubiquitination site on the substrate protein which catalyzes the

formation of an isopeptide bond without the RING E3 directly transferring Ub to itself. Inside the cell ubiquitination of a target protein happens in a large complex which can contain many other proteins depending on the E3 ligase (*22*). Interestingly, it has been shown that proteins in ubiquitinating complexes which lack any E1, E2 or E3 activity still have the ability to influence ubiquitination, a property that has been mainly attributed to their scaffolding roles (*23*). One clear example is a group of proteins termed "E4s" which have no E1, E2, or E3 activity and little sequence homology, but nevertheless are essential for elongating polyUb chains on certain substrates (*24, 25*). The schematics of ubiquitination are shown in (**Figure 1.2**).

Recent *in vitro* studies demonstrated that heterogeneous Ub-Ub linkages form with certain E2/E3 combinations, but the addition of an ubiquitin binding domain (UBD) induces the E2/E3 pair to generate polyUb chains with homogeneous Ub-Ub linkages (*26-29*). This suggests that entities other than E1, E2s, or E3s direct the formation of Ub-Ub linkages during ubiquitination in some cases. This is in contrast to other known E1, E2, E3 combinations which form homogenous Ub-Ub linkages without the need for external entities, such as the E3s E6AP and TRAF6, which ubiquitinate their substrates with K48 and K63 Ub-Ub linkages, respectively. After the action the E1, E2, E3/(E4) enzyme cascade, the newly ubiquitinated protein is ready to embark on a specific pathway, dictated by which Ub-Ub linkages are present in the polyUb modification.

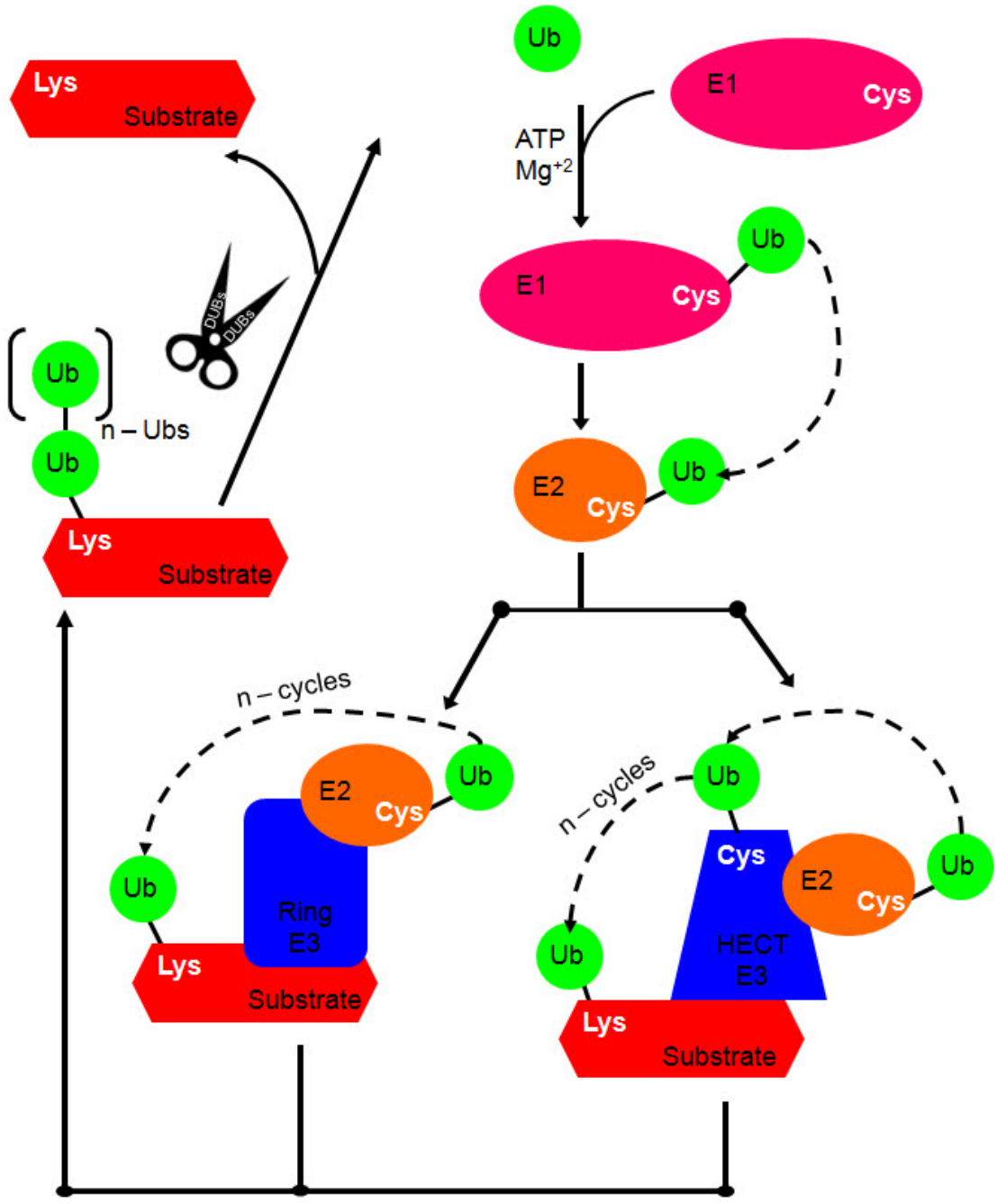


Figure 1.2 – The E1, E2, E3 ubiquitination cascade.
The attachment of Ub (green) to a substrate begins with a thioester bond between the C-terminus of Ub and the active site cysteine of E1 (magenta), next Ub is transferred to an E2 (orange). Then the E3(blue) binds the loaded E2 and substrate (red) and catalyses the transfer of Ub to a lysine on the substrate. The E3 can keep adding Ubs to create a chain or just attach one Ub to the substrate. Once modified with Ub, the substrate can embark on signaling pathway or DUBs can remove the Ub modification.

1.3 Deubiquitinases

Deubiquitinases (DUBs) are a diverse class of enzymes, which have one common property, the ability to recognize Ub and remove a modification beyond Ub's C-terminus. Enzymatically DUBs are all classified as (EC 3.1.2.15), however they can employ different catalytic mechanisms, centered around a single cysteine residue or a coordinated zinc (Zn^{+2}) ion. DUBs in conjunction with the E1,E2,E3 enzymes are responsible for maintaining the free Ub pool in the cell by creating a balance between the opposing processes of ubiquitination and deubiquitination. Aside from having a domain with DUB activity, a majority of the proteins classified as DUBs also contain other domains, mainly used for regulatory functions (*18*). Around 100 DUBs have been identified in humans, while there is direct evidence for just 20 DUBs in yeast (*30*). Schematics for all known human DUBs are show in (**Figure 1.3**). Pinpointing similarities between DUBs is difficult since they vary greatly in size, cellular location, and even in the structures of their catalytic DUB domains. Given that we know Ub is present in virtually all parts of the cell and ubiquitination occurs in many diverse locations as well, it is reasonable that proteins which interact with Ub exhibit a high degree of diversity to function in so many different environments.

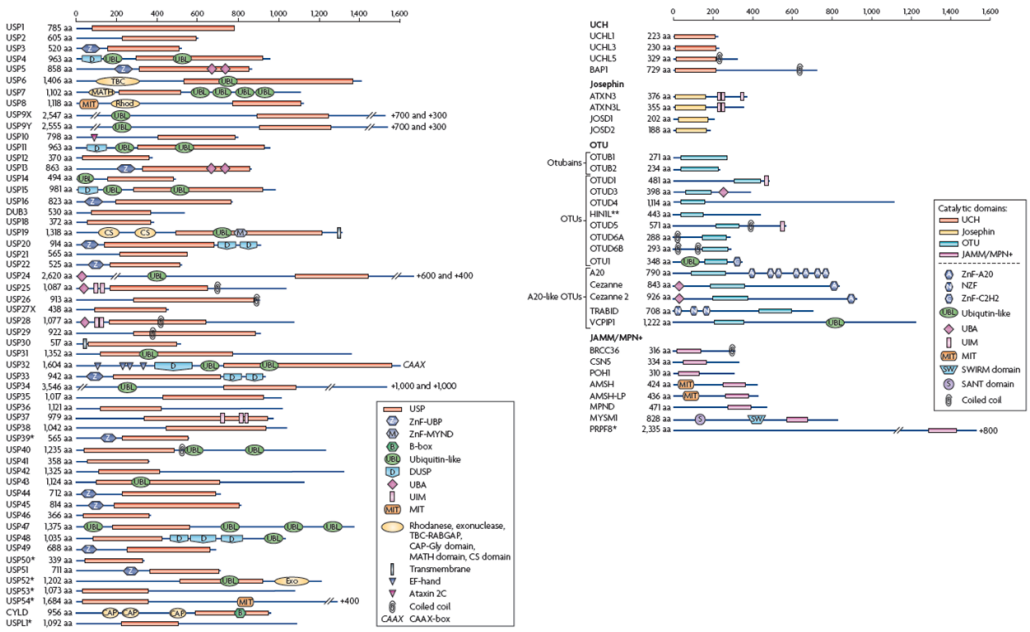


Figure 1.3 – Schematic representation of all known human deubiquitinases. Individual DUBs are clustered into the five families USP, UCH, OTU, MJD, and JAMM/MPN+. Note the diversity and prevalence of non-catalytic domains. adopted from (*18*).

1.3.1 The five families of deubiquitinases

Although there are many differences across DUBs, structural and bioinformatics data has shown that some DUBs share similar 3D folds and primary sequences in their catalytic domains (*31-34*). Given these properties are some of the only clear similarities between DUBs, they have been classified into five families which reflect the fold of their catalytic domains: ubiquitin C-terminal hydrolases (UCH), ubiquitin specific proteases (USP), ovarian tumor proteases (OTU), Machado-Joseph disease proteases (MJD), and JAB1/MPN/MOV34 metalloenzymes ($JAMM/MPN^+$). UCHs, USPs, OTUs, and MJDs are all cysteine proteases, while the $JAMM/MPN^+$ family of DUBs are zinc dependent metalloproteases. Of the five families, the USP family is the best characterized by a large margin with over sixty

members identified in humans. In addition, the size of the USP family is also significantly larger compared to the other four families which each have no more than ten members each.

DUBs in the UCH family contain a catalytic domain ~200 residues with an active site cysteine shielded by a crossover loop which governs what can be cleaved (*35-40*). Upon Ub binding there are significant conformational changes in the active site crossover loop, however only small adducts or unfolded proteins attached to the C-terminus of Ub can enter. This mechanism allows UCHs to have activity for Ub attached to small adducts, but not the Ub-Ub bond.

The USP family of DUBs share a distinct hand-like fold (thumb, fingers, and palm) and are generally large proteins ~1,000 residues. Aside from their catalytic domain, many USPs contain a variety of other domains including ubiquitin-binding domains (UBDs). As with the UCH family, there are significant conformational changes upon Ub binding, however USPs readily cleave Ub-Ub bonds which makes determining the catalytic structural conformations difficult. One of the more intriguing properties of USPs is their cross reactivity or ability to hydrolyze bonds between ubiquitin-like (UBL) polymers, such as the cleavage of ISG15 (UBL) by the DUB USP21 (*33, 41*). Another important functional property of select USP DUBs is their ability to autoregulate using UBL domains contained in their sequence as reported with USP4 and USP7 (*42*). The activity of USPs in respect to linkage preference and how they cleave polyUb chains is broad and varies depending on the individual USP.

The OTU family exhibits a wider variation between members in respect to the size of the catalytic domain and its three dimensional fold. Similar to USPs, OTUs also typically contain other UBDs in their full sequences. The OTUs are unique in that they exhibit some of the greatest Ub-Ub linkage specificity among DUBs and do so across a wide variety of linkages (*41, 43-46*). It has been established that OTUB1 has an almost antibody-like selectivity for K48 linkages, OTUD5 (DUBA) has a preference for K63 linkages, TRABID has its greatest activity for K29 and K33 linkages, and Cezanne is highly selective for K11 linkages (*44-48*). In addition to their remarkable linkage selectivity, OTUs are considered true Ub-isopeptidases and preferentially act on only Ub-Ub isopeptide bonds, unlike other families which can cleave Ub-substrate bonds or native peptide bonds (e.g. M1 linkages).

The MJD family of DUBs is the least studied by far. Ataxin-3 has been well studied due to its relevance in Machado-Joseph disease, however much remains unclear. For instance, it has been suggested that Ataxin-3 functions optimally when it is polyubiquitinated and although UIM domains in Ataxin-3 have a preference to bind K48 linked polyUb, it has increased DUB activity for K63 linkages. MJDs do share a similar cysteine active site with UCH, USP, and OTU family DUBs and also undergo significant conformational changes upon Ub binding. In the MJD family, a long α-helix controls access to the active site which changes upon Ub binding, similar to the crossover loop in the UCH family.

JAMM/MPN^{+} are the only family of DUBs that utilize a coordinated Zn^{+2} ion in their active sites. Although their catalytic domains differ slightly, virtually all JAMM/MPN^{+} DUBs are found as components in multi-subunit complexes. A

handful of studies have isolated JAMM/MPN$^+$ DUBs and demonstrated that they can function alone *in vitro* (*49-51*). To demonstrate the utility of this DUB family consider that POH1 (Rpn11 in yeast) is a critical component of the proteasome, AMSH and AMSH-LP are essential in the ESCRT complexes, CSN5 has a major role in the COP9 signalosome, and a number of JAMM/MPN$^+$ members (BRCC3, MYSM1, and PRPF8) are involved with multi-subunit complexes that interact with DNA (*52*). The JAMM/MPN$^+$ family of DUBs has clearly evolved to take on essential roles in many parts of the cell. Similar to the USP family, JAMM/MPN$^+$ DUBs have a wide range of properties e.g. AMSH is highly specific for K63 linkages, POH1 is reported to cleave the Ub-substrate isopeptide bond, and CSN5 can cleave heterologous Nedd8-Ub chains (*34, 53, 54*).

1.3.1 Deubiquitinases and disease

As previously stated, abnormal or irregular ubiquitination patterns can lead to many adverse outcomes in the cell. In the most general terms, an overactive E3 which destroys a disease suppressing protein or an E3 which is underactive against a pathogenic protein are both scenarios that have been documented to trigger disease (*25, 55*). However, since DUBs can also directly influence the ubiquitination state of a specific protein, it has also been shown that a DUB which excessively removes Ub from a disease causing protein targeted for degradation by the UPP or a situation where a DUB that does not efficiently remove Ub from a disease suppressing protein cause a disturbance in ubiquitination that leads to disease just as the E3 examples (*56, 57*). Given that several DUBs and DUB families are named for disease which they are involved (recall MJD for Machado-Joseph disease and OTU for ovarian tumor

proteases) I will highlight several less obvious DUBs with functions directly related to disease. Proteasome associated DUBs, to be discussed (**section 1.4.6**) are extremely important across a range of diseases.

The level of p53, "guardian of the genome" is chiefly regulated through ubiquitination, which is controlled by the actions of the E3 ligase Mdm2 and the DUB USP7. Several studies have shown that the delicate balance between p53, Mdm2, and USP7 is perturbed in a variety of malignancies and that USP7 activity promotes oncogenesis (*58*). In the case of p53, the DUB activity of USP7 saves proteins that ubiquitinate p53 from their own destruction by the UPP, allowing p53 to be excessively destroyed by the UPP (*59*). This observation has triggered many to develop USP7 inhibitors, which have shown highly promising results in the early stages and at the present time P005091 (Progenra, Inc) is set to be the first approved DUB inhibitor for cancer treatment (*60*).

In contrast to USP7, another USP family DUB cylindromatosis (CYLD) has been shown to be a tumor suppressor due to its ability to modulate NF-κB (*61*). Thus pathology is caused by a loss of function in CYLD and several rare genetic diseases arising from mutations (mostly truncations) in the USP domain of CLYD have also be reported, which further highlights the importance of CYLD's DUB activity (*62*). Unlike other characterized USP DUBs, which show a preference for K48 linkages, CLYD preferentially cleaves K63 linkages due to subtle structural arrangements in the USP domain (*41*). The unique structural features of CYLD compared to other USP family DUBs is shown (**Figure 1.4**). Logically no CLYD inhibitors have been

proposed, but gene therapy to restore CLYD function is currently being explored (*63*).

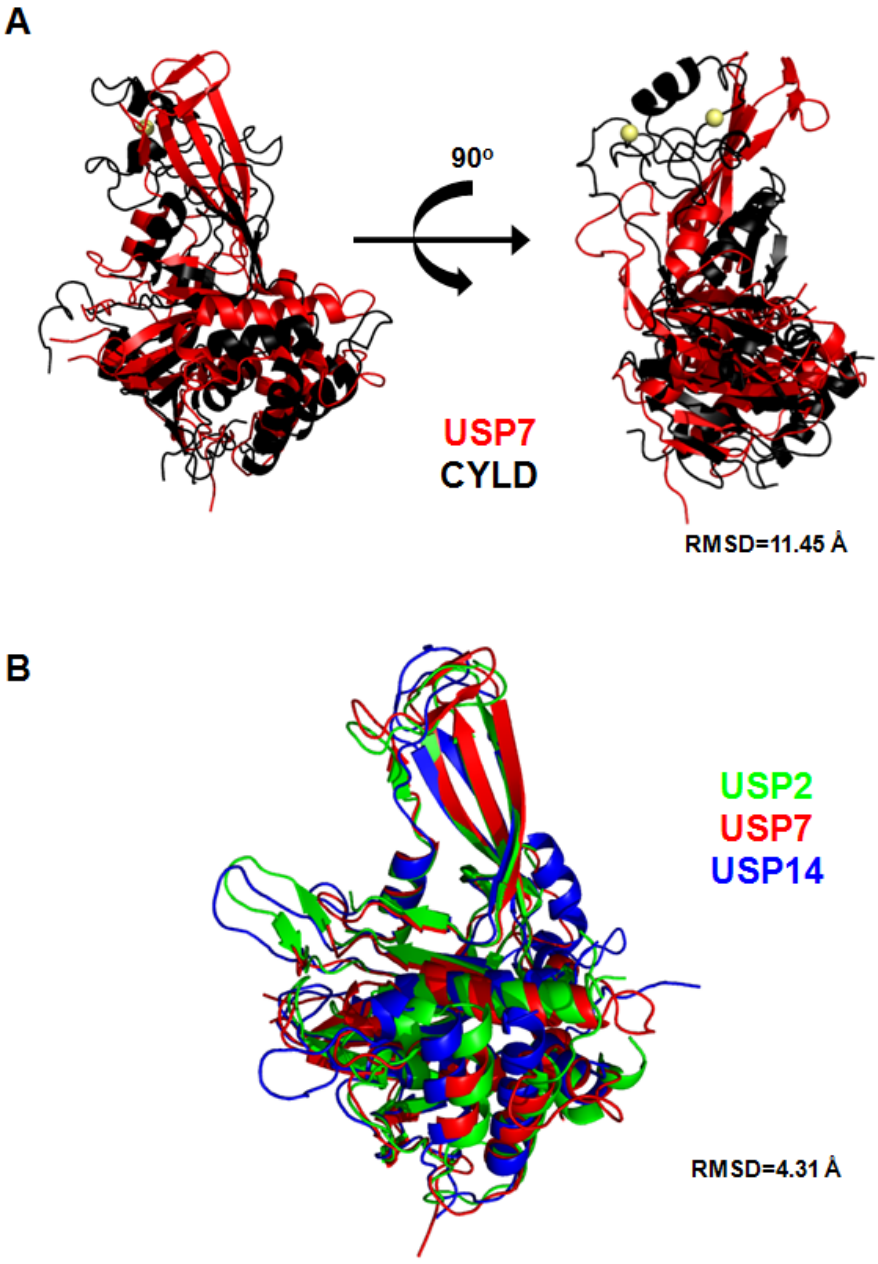


Figure 1.4 – Structural comparison of K48 and K63 linkage selective USP DUBs. (**A**) Alignment of USP7 (PDB-1NB8) in red with CYLD (PDB-2VHF) in black revels major structural differences in the general fold of the USP domain. CYLD also contains two Zn^{+2} ions shown as yellow spheres. (**B**) Alignment of K48 selective USP DUBs USP2 (PDB-2HD5) in green, USP7 (PDB-1NB8) in red, and USP14 (PDB-2AYN) in blue revels conserved structural features.

AMSH, a JAMM/MPN$^+$ DUB which regulates ESCRT machinery has been found to have both beneficial and pathological roles. The best understood function of AMSH is to cleave K63 linked Ub modifications on endosomal cargo when it is recruited to membranes by ESCRT complexes (*64*). On one hand, AMSH function can be extremely beneficial since it has been reported to reduce the accumulation of disease associated proteins in the central nervous system (*65*). On the other hand, AMSH function is also responsible for the maturation of several oncogenic G protein-coupled receptors (GPCRs) and receptor-tyrosine kinases (RTKs) (*66*). Thus it has been proposed and also demonstrated in several cancer cell lines that with proper delivery systems, AMSH inhibitors are highly effective against rapidly growing cancers (*67*). In some cases EGFR is also deubiquitinated by the OTU K11 specific DUB Cezanne-1, which also contributes to cancer progression (*68*).

1.4 The ubiquitin proteasome pathway

1.4.1 Overview of protein homoeostasis

Just as protein production is critical to the cell, so is the opposite process of protein destruction. Many proteins are required for optimal cell function, however proteins generally have relatively short lifetimes and must be disposed of properly. All proteins in the cell originate from DNA starting with transcription to mRNA and ending with ribosomal translation. Yet, the cell has several distinct mechanisms for the destruction of proteins at the end of their lifetimes. Ultimately proteins reach their end in lysosomes through autophagic pathways, which also can be signaled by

polyUb or they are destroyed by the proteasome with the aid of Ub signaling (*69*). It has been suggested that short lived proteins are generally destroyed by the UPP, while proteins with longer life times are targeted to the lysosomes (*70*). Although many mysteries surround both, it is clear that the proper function of lysosomes and the proteasome are essential for cell survival.

1.4.2 Architecture of the 26S proteasome

The full, "doubly capped" 26S proteasome is a 2.5 MDa complex composed from thirty-three distinct subunits (*71, 72*). The proteasome, just like the Ub molecule is highly conserved across eukaryotes and many subunits are nearly identical in structure and sequence between yeast and humans (*73*). There are three distinct components of the proteasome, the 20S core and the 19S regulatory particle (RP) that consists of the base (RP-base) and lid (RP-lid) (*30, 74, 75*). The 20S core particle resembles a barrel and is composed of four heptameric rings in a $\alpha_7\beta_7\beta_7\alpha_7$ arrangement. The outer α rings of the CP regulate access to the inner chamber where proteolysis occurs (*76*). The 20S CP can simultaneously associate with two 19S RPs on either end to form a “doubly capped” proteasome or accept just one 19S RP to form a “singly capped” proteasome (*74*). A native gel assay which separates the various states of the CP in combination with a fluorogenic substrate is commonly used to gauge the distribution of the various oligomeric states (see **Figure 1.5**) (*77*) .

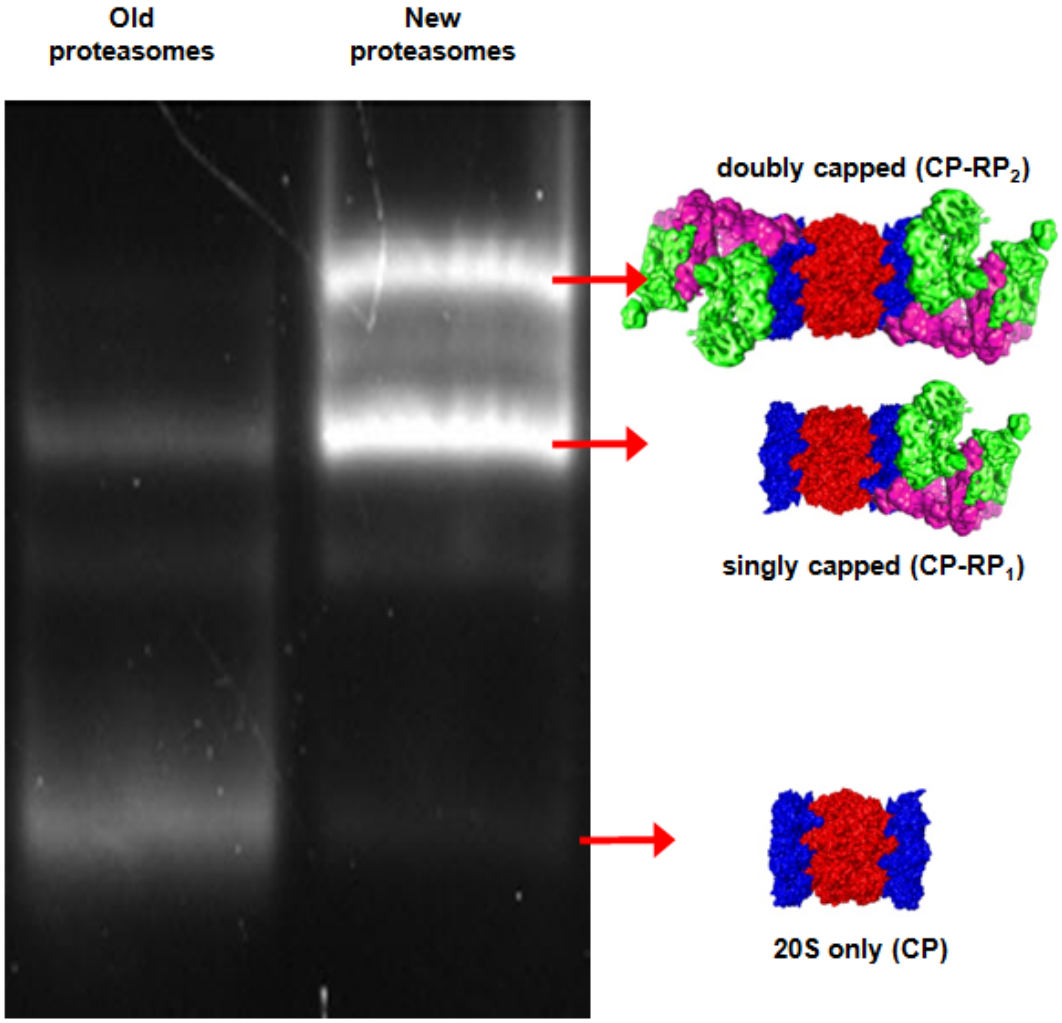


Figure 1.5 – Native gel assay to characterize the 20S CP.
Both lanes show purified yeast proteasomes run on a 4% gel, then assayed by a fluorogenic peptide substrate. The left lane shows that a four year old proteasome sample looses most RPs. The weaker signal intensity also indicates a decreased proteolytic activity in the aged proteasomes. The right lane contains freshly prepared proteasomes with both singly and doubly capped CPs and virtually no disassociation from the RP. The strong intensity of bands for the singly and doubly capped proteasomes indicates their high proteolytic activity.

The main feature of the RP base is a hexameric ring of AAA-ATPases (Rpt1-Rpt6) which inserts directly to the alpha subunits of the 20S. Rpn1 and the structurally similar Rpn2 subunit are the largest subunits in the proteasome and are

also components of the RP-base. It is proposed that Rpn1 and Rpn2 serve as scaffolds, but studies have shown that Rpn1 takes on other roles such as serving as a receptor for UBL containing proteasome associated proteins (*78, 79*). Rpn10 and Rpn13 are also found in the RP-base and contain the only known polyUb receptors in the proteasome, VWA and UIM domains of Rpn10 and the Pru domain of Rpn13 (*80, 81*). A majority of the subunits that form the RP-lid (Rpn3, Rpn5, Rpn6, Rpn7, Rpn9, and Rpn12) contain PCI domains and are critical for holding the RP-lid to the RP-base. Rpn11 and Rpn8 are structurally similar and contain the JAMM/MPN^{+} fold. Rpn11 is the only known subunit of the proteasome to have DUB activity and is positioned directly over the hexameric ring of ATPases.

The assembly of the proteasome has been intensely studied. The consensus is that many cofactors are involved, which direct the step wise assembly of the 26S proteasome by parts, such that the CP, RP-base, and RP-lid are preassembled before combining to form the 26S proteasome (*82*). The 20S CP, which is usually studied as a whole can be crystallized or selectively labeled for solution NMR studies (*83, 84*). High resolution structures of individual subunits of the RP have been studied in isolation since the RP is not suitable for crystallization by conventional means (*85*). Interestingly, the structure of the whole 26S proteasome has eluded everyone who has attempted. Low resolution cryoEM models often lacking many key subunits have been the closest success produced by modern structural biology (*86-89*). There has been no success in determining where a polyubiquitinated substrate binds the proteasome, yet this information is critical for our understanding of how substrates are processed.

1.4.3 Substrate targeting to the 26S proteasome and polyUb processing

A polyUb chain of at least four units internally linked through K48 was the first definitive signal for proteasomal degradation discovered (*15, 16*). Since that particular study it has also been found that some proteins, such as ornithine decarboxylase (ODC) can be degraded by the 26S proteasome in an Ub independent manor, while others simply require monoubiquitination (*90, 91*). In fact, many different ubiquitin modifications including virtually all linkage types can interact with purified proteasomes *in vitro* and K63-linkages have been show to be sufficient for degradation (*92, 93*). However, in the cell many other proteins interact with the Ub signal and generally ensure that only K48 signals reach the proteasome (*94*). Hence the caveat: in the absence of other cellular proteins, the proteasome can interact with any Ub-Ub linkage, but in the cell this does not happen because other proteins intercept the Ub signal upstream of the proteasome. K48-linkages are also intercepted upstream of the proteasome by UBL/UBA shuttle proteins, which facilitate the transport of ubiquitinated substrates to the proteasome. The shuttling proteins recognize the polyUb signal with a UBA domain and the Rpn1 subunit of the proteasome with their UBL domain (*95-98*). In yeast, Ddi1, Rad23, and DSK2 are the known UBL/UBA shuttling proteins (*30*). Once a polyubiquitinated substrate is docked to the proteasome by a shuttle protein, little is known about how the Ub signal is processed. It is assumed that the Ub signal is transferred from the UBA of the shuttle to one of two Ub receptors, the UIM of Rpn10 or the Pru-domain of Rpn13 on the proteasome via competitive binding interactions (*86, 95, 99*).

After the proteasome receptors accept the polyubiquitinated substrate, the polyUb signal can be removed by the Rpn11 subunit. Alternatively other proteasome associated DUBs such as Ubp6 can be recruited to facilitate processing of the polyUb signal. For a substrate, recognition by the proteasome does not necessarily ensure its immediate destruction. In some cases proteasomal DUBs prematurely remove the polyUb signal, sparing the substrate from degradation (*100*). While in other cases Hul5 an E3/E4 Ub ligase is recruited to speed up the degradation of a substrate by adding more K48 linkages or slow down the degradation of a substrate by adding K63 linkages to the existing polyUb signal (*101*). The substrate is committed for proteasomal degradation when the hexameric AAA-ATPase ring initiates its unfolding and it begins to enter the chamber of the 20S (*102*). From *in silico* simulations it is thought that attachment of Ub helps destabilize substrates, facilitating their unfolding at the proteasome, but this theory has yet to be proven (*103*). Also, given that the proteasome functions in protein quality control i.e. to degrade misfolded proteins it is thought that some components can recognize abnormal proteins (*104*). During its destruction a substrate will enter one end of the 20S CP and exit from the other in the form of peptides ranging from 3-22 residues in length (*105*). This is accomplished by three proteolytic subunits β1, β2, and β5 that have activity similar to trypsin (basic residues), caspase (acidic residues), and chymotrypsin (hydrophobic residues) (*106*). Interestingly these peptides released from the 20S may hold significant biological functions in both the immune system and as general regulators of apoptosis (*107, 108*).

1.4.5 Therapeutic intervention on the UPP

The proteasome itself became a drug target after some very astute observations by Alfred Goldberg made over the course of several decades. Early studies noted that proteasome inhibition gave a brief protective effect to many cells (ranging from yeast to human), but if the proteasome was inhibited long enough, abnormal proteins would build up in the cytosol and ER, which led to JNK-kinase induced apoptosis (*109*). Initially small molecule proteasome inhibitors were developed by a biotech company Goldberg helped start (MyoGenics/ProScript) for studies in cell culture, not therapeutics. Other observations established that a majority of the cell's proteins, both short and long lived are degraded by the UPP as opposed to autophagic (lysosomal) pathways (*107*). A little know fact is that initially the intended therapeutic use of proteasome inhibitors was for muscle atrophy and to modulate MHC class I antigen presentation (*107*). However, advances in oncology along with many of the pieces coming together led to a new hypothesis: if cancer cells (particularly the blood cancers) over express proteins which need to be degraded by the proteasome, then inhibiting the proteasome's ability to degrade substrates will causes extensive ER stress and induce apoptosis (*110*). At the present this now validated hypothesis represents a great success story of translational research that paved the way for Velcade® (Millennium Pharmaceuticals) and other small molecule proteasome inhibitors.

Generally, proteasome inhibitors are peptide-like small molecules that reversibly bind the proteolytic subunits located in the proteasome's 20S core. The first generation proteasome inhibitors were designed based on peptide mimics without

any structural knowledge of the actual target (*111*). The basic science behind proteasome inhibitors has come a long way and second generation molecules are still in various stages of testing (*110, 112*). The architecture of the 20S core has been well characterized and many structures of inhibited 20S CPs have been solved by X-ray crystallography (**Figure 1.6)**. The therapeutic effectiveness of proteasome inhibitors is proven, but first generation drugs (Velcade®) produced many severe off target effects and some cancers are developing resistance with mutations in the proteolytic β5-subunit where Velcade® binds (*112-115*). Structural data along with special chemistry that changes the binding mechanism from reversible to irreversible has allowed subsequent proteasome inhibitors (e.g. Carfilzomib, Onyx Pharmaceuticals) to be much more effective by overcoming resistance and minimizing off target effects (*110, 112, 116-118*).

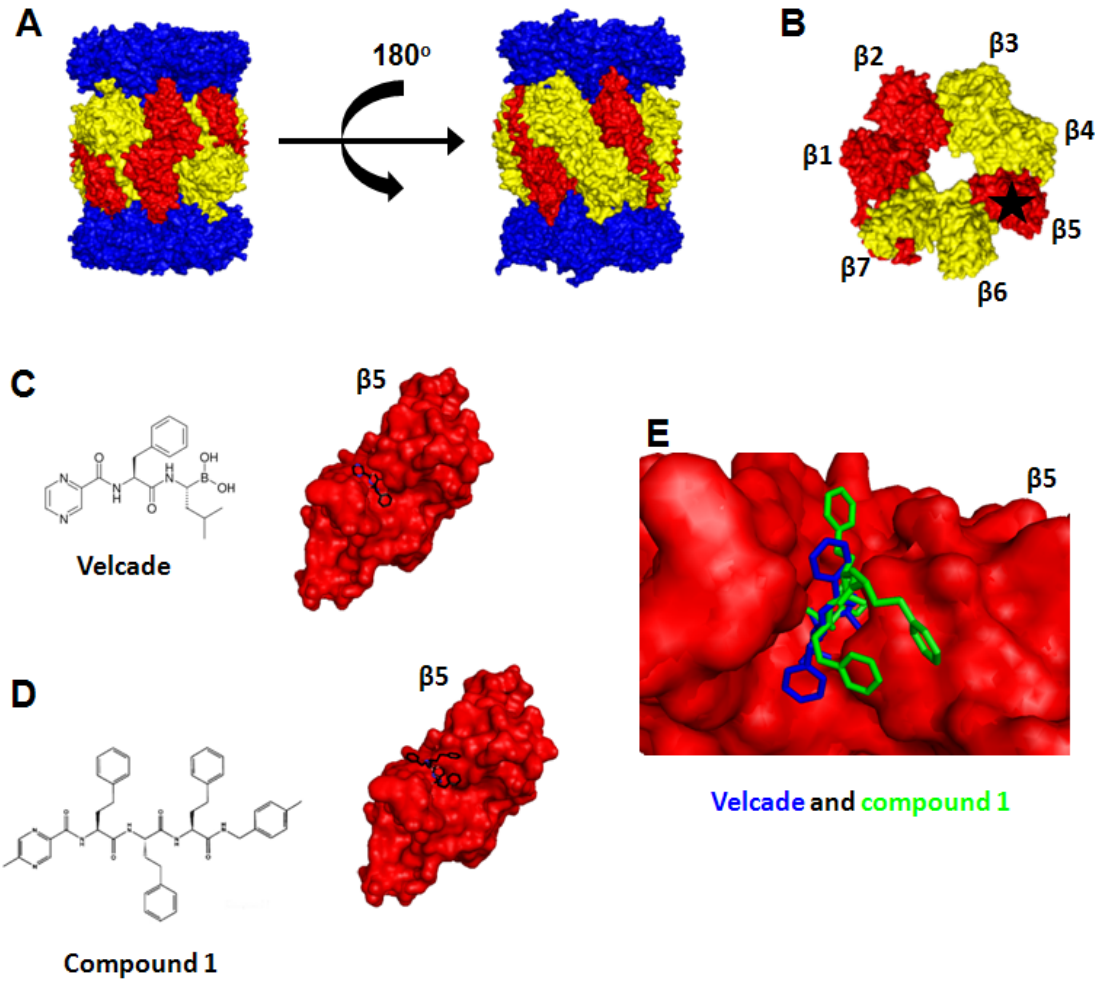


Figure 1.6 – Overview of proteasome inhibitors.
(A) 20S CP from yeast proteasome (PDB-1RYP) with outer α_7 rings colored blue, inner β rings colored yellow, and proteolytic β1, β2, and β5 subunits colored red. **(B)** Arrangement of heptameric (β_7) ring of β subunits with proteolytic units shown in red. **(C)** Structure of the proteasome inhibitor Velcade and in complex with the β5 subunit (red) from (PDB-2F16). **(D)** Structure of a lead second generation proteasome inhibitor "compound-1" in complex with β5 subunit (red) from (PDB-3MG4). **(E)** Alignment of Velcade (blue) and compound 1 (green), note the substituent groups of compound 1 extend to different surfaces of the β5 subunit (red) which allows this compound to overcome resistance.

1.4.6 Proteasome associated DUBs

The proteasome harbors a single DUB, Rpn11 (POH1 in humans, also known as PSMD14, MPR1, S13, or CeP1) from the JAMM/MPN$^+$ family. Of the many other proteins that can associate with the proteasome there are two DUBs, Ubp6 (USP14 in humans) a UBL containing DUB that binds the proteasome via Rpn1 and UCH37 (also called UCHL5) that is recognized by the Rpn13 subunit. These three DUBs all play an important role, recycling the Ub signal back into the cell, which maintains the free pool of Ub. Rpn11 deletions are lethal to yeast since the subunit is essential for proteasome assembly, while mutations that eliminate DUB activity have profound phenotypes marked by the accumulation of polyUb conjugate (*49*). The function of Ubp6 has been extensively pioneered by Finley and coworkers who have also found that Ubp6 itself can modulate substrate processing on the proteasome non-catalytically (*119*). Finley has generally concluded that Ubp6 prematurely removes the polyUb signal from substrates, delaying their destruction by the proteasome (*120*). The role of UCH37 is less studied, reports show that it has functions similar to Ubp6 (*54*). Interestingly, these proteasome associated DUBs have also been found to have roles outside of the proteasome. For example, Ubp6 functions in telomeric silencing by modulating histone modifications (*121*).

Much has been learned about substrate processing from yeast and simple *in vitro* studies, however the actual sequence of events in which a substrate goes through on the proteasome is not clear. It is accepted that Rpn11 only acts after a substrate is completely committed to degradation (*51*). Yet a body of work by Finley and coworkers demonstrates there is indeed many events including chain remodeling that

can happen before a substrate is degraded. Regardless of the details, the DUB function of both Rpn11 and Ubp6 have been proposed as therapeutic targets in a range of diseases (*122, 123*). To date, there is a strong effort underway to develop inhibitors for both Ubp6 and Rpn11, but it seems that Ubp6 inhibitors are the furthest in development with IU1 showing many promising results (*124*). The vital role of proteasome associated DUBs has been well established and the development of specific inhibitors is the next logical step for this field. Fully understanding substrate processing requires techniques and capabilities that have yet to be developed.

1.5 Non-degradative ubiquitin signaling outcomes

1.5.1 DNA damage repair

DNA damage repair pathways were the first non-degradative Ub signaling pathways discovered. In contrast to the UPP, DNA repair employs the K63-linkage as opposed to the K48 linkage. Initially these non-degradative pathways were found to recruit a novel K63 forming E2 heterodimer (Ubc13:Mms2), in UV sensitive yeast (*125, 126*). Once the K63 signal is formed DNA repair is initiated by localization of RAP80 at double strand DNA breaks, which recruits the BRCA1-BARD1 heterodimer (*127*). Many E3 Ub ligases that form K63 linkages have also been found to have other roles in DNA repair and transcription. Aside from their roles in double strand break repair, TRAF6 (tumor necrosis factor receptor-associated factor 6) and RNF4 have also been found to have important roles for genomic stability and the progression of cancers (*128-130*).

1.5.2 Histone modification

Ub also has other roles inside the nucleus aside from DNA repair. Histones, the most fundamental unit of chromatin structure allow transcription to be regulated through several post-translational modifications on C-terminal lysine residues. Ub and the UBL (SUMO) are the only protein modifiers known to be directly attached to histones and it is still not fully understood how the post-translational modification of histones alters transcription and chromatin structure (*131*). The Ub modification has also been found to be essential for development and is key to ensuring the correct genes are expressed at the exact time they are required (*132*). Specialized E3 Ub ligases, Ring1B and E6-AP (E6 associated protein) which directly modify histones with Ub have been implicated in Angelman syndrome and there are other examples for other developmental diseases (*132*).

1.5.3 Membrane trafficking

One of the most vital non-degradative functions of Ub signaling is membrane trafficking. Typically monoubiquitination, multiple monoubiquitination, or K63 polyubiquitination are required for proteins to cross several eukaryotic membranes and also for endosomal trafficking (*133*). In fact, many transmembrane proteins are known to carry UBDs or have E3 ligase activity. The Ub modification can also help control the sequence of events in complex systems. For example the endosome transport factor Vps9 (Rab5 in humans) is composed of a CUE domain (a type of UBD) and GEF (guanine nucleotide exchange factor) domain that is only active after the CUE domain binds Ub (*134*). Since the Ub or polyUb signal must be removed, DUBs also play key roles in membrane trafficking. Several USP family DUBs

(USP27X and USP29) have evolved transmembrane domains or are able to be recruited to membrane complexes like the JAMM/MPN$^+$ family DUB AMSH.

Chapter 2: Structural and conformational review of ubiquitin signaling

2.1 Introduction

Although Ub is just one protein its ability to be conjugated into distinct polymeric chains allows it to function on divergent signaling pathways. Proteins that interact with Ub have evolved several varieties of binding domains and recognition surfaces to differentiate between Ub and numerous UBL proteins in the cell. Typically the affinity of Ub for its receptors is moderate with equilibrium dissociation constants (K_d) ranging between 10-120 μM. It is important to note that not all polyUb chains are the same and the K_d of a particular receptor changes based on the Ub-Ub linkage and number of Ubs in the chain. For example, the K63 selective tandem UIM (tUIM) from RAP80 binds K63-Ub_2 with a $K_d = 21.6 \pm 0.8$ μM, K63-Ub_4 with a $K_d = 3.6$ μM, and K48-Ub_2 with a $K_d = 157 \pm 8$ μM (*135*). The structures of M1, K48, and K63 linked free polyUb chains and in complex with various UBDs represent a majority of our structural knowledge (*136-139*). However, there is still a void of structural data for atypical K27, K29, and K33 Ub linkages. Recent advances in bioorganic chemistry have led to several structures of Ub in complex with various E1, E2, E3 enzymes (*140*). Similar methods have also been employed to covalently attach Ub to active site cysteine residues of several DUBs (*31*). Structures of Ub modified substrates have eluded structural biologists and there are several pressing questions regarding the outcome of Ub modification. In general, there is much data for Ub interacting with individual domains, however in the cell this only represents a tiny piece of much larger complexes. Determining the complete layout of such complexes is regarded as a "Holy Grail" by many structural biologists and there is an

intense effort to tackle SCF E3 complexes, spliceosomes, the proteasome, the COP9 signalosome, mitotic complexes, and many other multisubunit complexes involved with Ub signaling.

2.2 Structural properties of monomeric Ub and UBL molecules

2.2.1 The Ub molecule

Ub is highly conserved across eukaryotes with just three subtle differences (P19S, E24D, A28S) in primary structure going from humans to yeast. The structure of monomeric Ub has been studied extensively using many techniques for a wide range of purposes. Aside from its biological significance, Ub has become a model protein for method development due to its many favorable properties, most notably its ease of production, small size, incredible stability (T_m~100^{o}C), and long shelf life. In addition to the seven lysine residues discussed earlier, Ub from all species retains the beta grasp (β-grasp) fold created by five antiparallel β-sheets, packed against a long α-helix. Many bulky hydrophobic residues (I3, V5, I13, L15, V26, I30, L43, L50, L67 and L69) are essential for maintaining the packing of Ub's hydrophobic core (**Figure 2.1**). This was highlighted in several studies which demonstrated that single point mutations (e.g. L67S or L69S) can dramatically impact the three dimensional structure of Ub (*141*). Interestingly the eight linkage sites are not restricted to one face on Ub, but are distributed about the surface of the molecule, with K27 being slightly buried. Surface exposed hydrophobic residues L8, I44, and V70 located on the β-sheets form the hydrophobic patch, which is essential for Ub to be recognized by receptors. Other notable structural features on Ub include a small 3(10) helix, several loops, and a highly flexible C-terminus or "tail" (residues 72-76). Although

the tail residues typically do not directly bind UBDs, they are essential for recognition by E1, E2, E3 enzymes and DUBs (*142*).

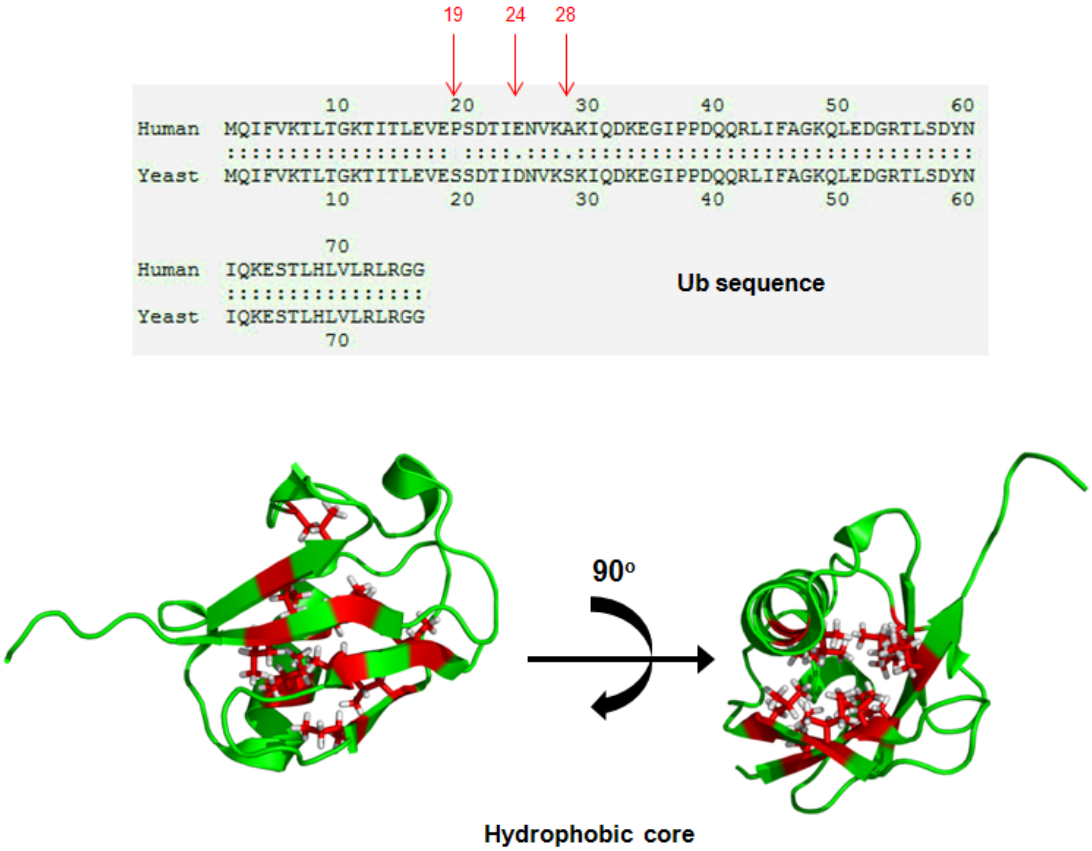


Figure 2.1 – Comparison of yeast and human Ub
Ub is highly conserved across all species. The primary sequence (top) shows three subtle mutations (red arrows) between yeast and human Ub. All known structures for monomeric Ub align nearly perfectly and the β-grasp fold is always conserved. The hydrophobic core residues (red sticks) are mainly responsible for the packing of the front five β-sheets against the long α-helix. Note both the β-sheets and α-helix have a network of hydrophobic residues with long side chains oriented inward.

2.2.2 UBL molecules

UBL molecules, which are "like Ub" in respect to their three dimensional structures also contain the hallmark β-grasp fold and can be found as monomeric units, polymers, or as domains in a variety of proteins. A number of UBLs (ISG15,

FAT10, Nedd8, and SUMO) are commonly conjugated to themselves or other proteins via isopeptide bond with a dedicated set of E1, E2, E3 enzymes and deconjugation is possible with isopeptidases just as with Ub. UBLs that exist as domains, such as the UBLs of DUBs or shuttling proteins have been show to have defined regulatory functions and remain confined to the proteins they originate i.e. the UBL of USP14 cannot be attached to another protein. The linkable UBLs (listed above) can modify a variety of proteins just as Ub, and in some cases these UBLs are attached to the same site as Ub (*143*). Recent studies have shown that there is much crosstalk between UBLs and Ub signaling pathways, and heterologous mixed Ub-UBL modifications exist in the cell (*144-147*). Of the numerous UBLs, Rub1(Nedd8 in humans) is the closest related to Ub with a nearly exact three-dimensional structure, 77% sequence similarity, and also the L8, I44,V70 hydrophobic patch residues (*148*) see (**Figure 2.2**). Many linkable UBLs are also cross-reactive with E1, E2, E3 enzymes and DUBs. For example, the E1 for Nedd8 can catalyze ubiquitination and the DUB USP21 can cleave ISG15 modifications as well as polyUb (*32, 149*).

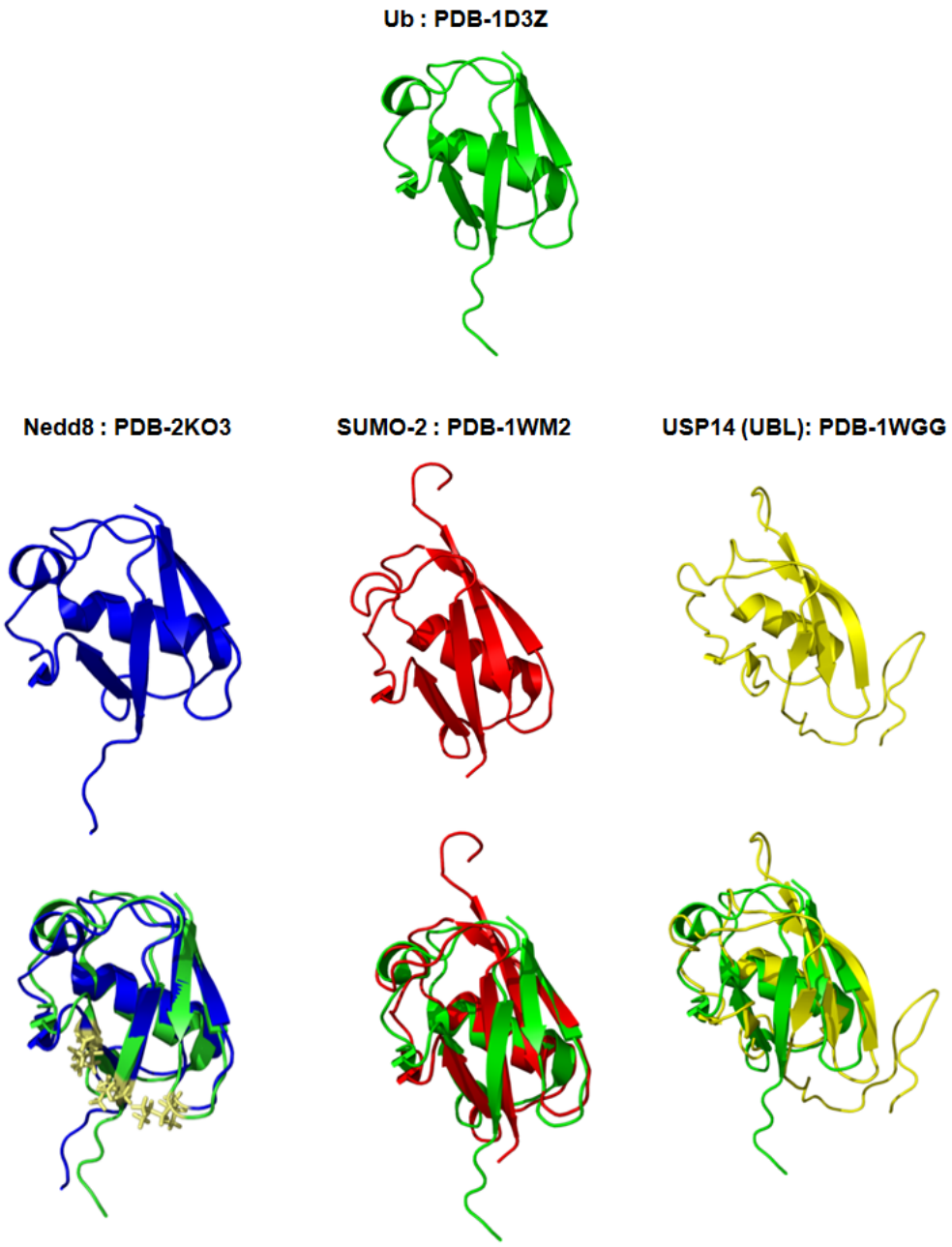


Figure 2.2 – UBLs share a similar fold to Ub
The structure of human Ub (green) is shown alone in the top row. The middle row shows the structure of Nedd8 (blue), SUMO-2 (red), and the UBL of USP14 (yellow). Below in the bottom row, Ub (green) is aligned to each UBL molecule. Note the alignment of Nedd8 (blue) and Ub (green), the structures are nearly identical even down to the position of the L8,I44,V70 hydrophobic patch residues. SUMO-2 (red), which is slightly larger also shows good alignment with Ub (green). The UBL of USP14 (yellow) has a β-grasp fold, however it differs in the arrangement of β-sheets and has more unstructured loop regions.

2.3 Structure and conformations of polyUb chains

2.3.1 Different as night vs. day: K48 and K63 polyUb chains

K48 and K63 polyUb chains represent essentially opposites in structure and signaling outcomes. These linkages account for just over 67% of all Ub-Ub linkages in the cell and are the best studies by far (*150*). It is generally accepted that K48 linkages signal for proteasomal degradation while K63 linkages signal for regulatory pathways, such as DNA repair. Until the landmark studies of Fushman and Pickart, little attention was given to the actual structure of the polyUb signal itself. In a series of papers published from 2001-2006, it was established that in solution K48 Ub linkages adopted compact conformations with extensive interdomain contacts, while the K63 Ub linkage resulted in an extended conformation with virtually no interdomain contacts (**Figure 2.3**) (*151-155*). The surface exposed L8, I44,V70 hydrophobic patch was found to be key in forming interdomain contacts between K48 linked Ubs and the structural arrangement of K63 linked chains prevented this interaction. This observation would become the main principle in explaining how a different Ub-Ub linkage resulted in a different signal.

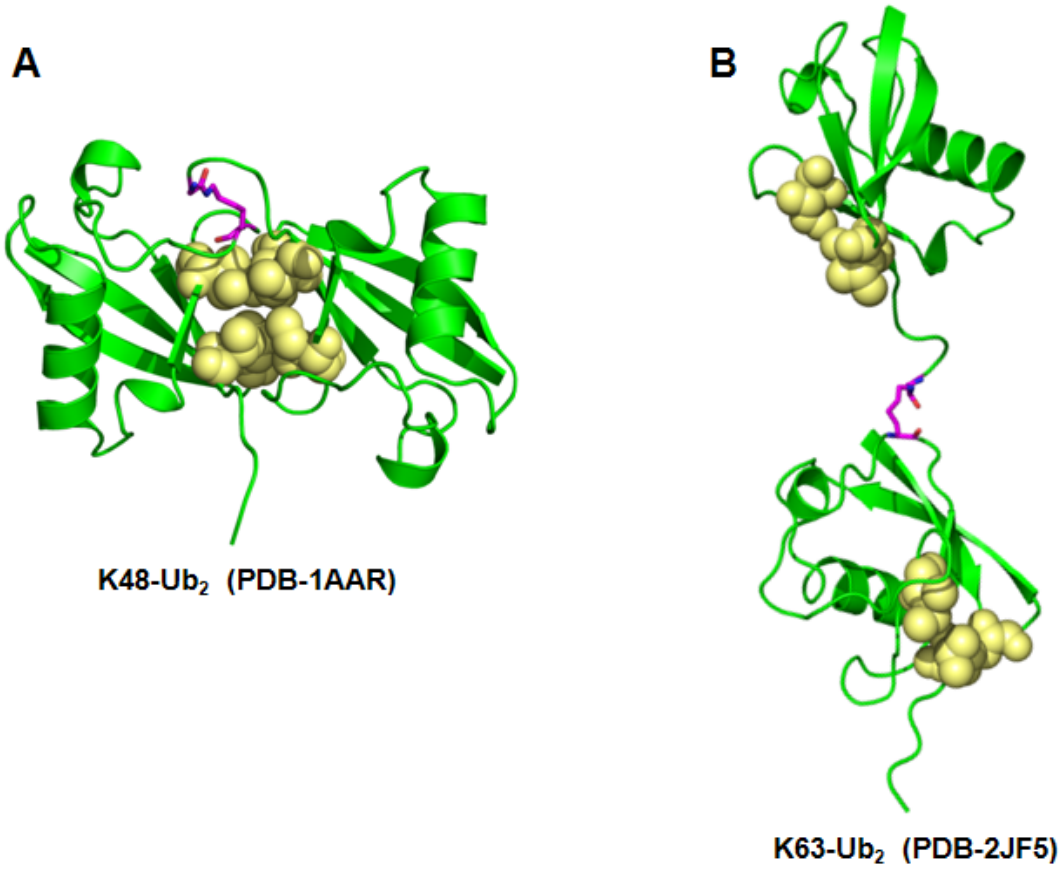


Figure 2.3 – The major structural conformations of K48 and K63 linkages. (**A**) The Structure of K48-Ub_2 has extensive interdomain contacts created through hydrophobic patch residues L8,I44,V70 (yellow spheres), which contributes to the compact conformation. (**B**) Conversely, K63-Ub_2 adopts an extended conformation with essentially no interdomain contacts.

Perhaps more importantly, these studies also established that polyUb chains were dynamic and took on distinct conformations. Later X-ray derived structures would also corroborate this observation and this led to the hypothesis that specific Ub-Ub linkages resulted in distinct signaling outcomes due to the ability of the polyUb chain to interact with specific receptors. After these initial studies much attention was directed at determining the exact populations of each conformation for a given linkage. Prof. David Fushman with his cutting edge solution NMR techniques was

the first to determine the intricacies of the conformational equilibria that governed K48 and K63 linkage (*151, 155*). Years later other solution methods such as small angle X-ray and neutron scattering, (SAXS) and single molecule FRET (smFRET) studies confirmed Prof. Fushman's early observations that K48 linkages had two predominant conformations, "open" and "closed" with populations of 15% open and 85% closed (**Figure 2.4**), while K63 linkages had two predominate conformations 86% extended and 14% compact (**Figure 2.5**) (*156, 157*). The protonation state of the H68 side chain was shown to dictate the conformational equilibrium of K48 linked chains in an unique electrostatic phenomenon (*158*). At neutral pH, above the pKa of histidine, H68 has a net charge of 0 that allows for extensive hydrophobic interaction favoring the closed conformation. While below the pKa at acidic pH, H68 has a +1 charge that causes repulsion between the two Ubs favoring the open conformation (*158*). Interestingly the crystal and solution structures conflicted in this regard for K48 linkages, with the crystal structures showing an open conformation of K48-Ub_2 at neutral pH and the closed conformation at acidic pH. This discrepancy is explained by crystal lattice forces dictating the conformation of K48 linked polyUb (*159*).

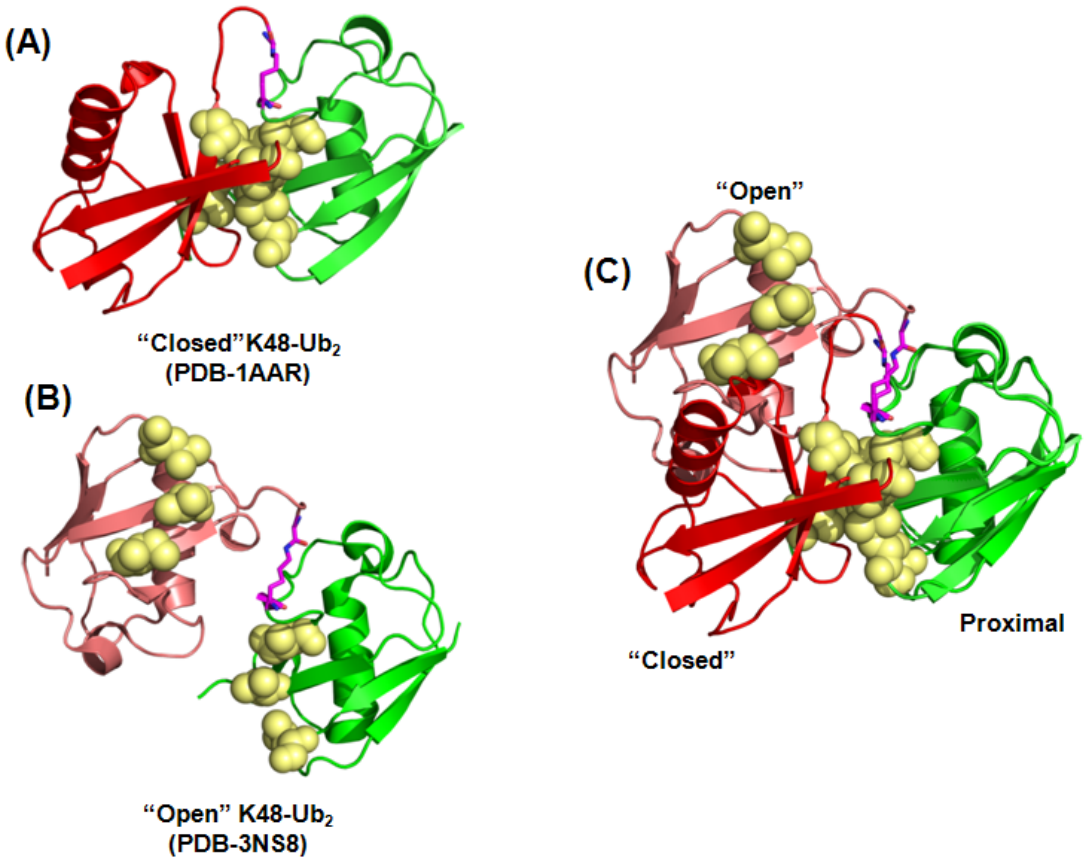


Figure 2.4 – Open and closed conformations of K48-Ub_2
(**A**) The closed conformation of K48-Ub_2 is represented by PDB-1AAR with the distal Ub (red), proximal Ub (green), and the L8,I44,V70 hydrophobic patch residues as (yellow spheres). (**B**) An open conformation of K48-Ub_2 showing the hydrophobic patch residues do not interact, distal Ub (salmon) and proximal Ub (green). (**C**) Alignment by proximal Ub for the open and closed structures of K48-Ub_2.

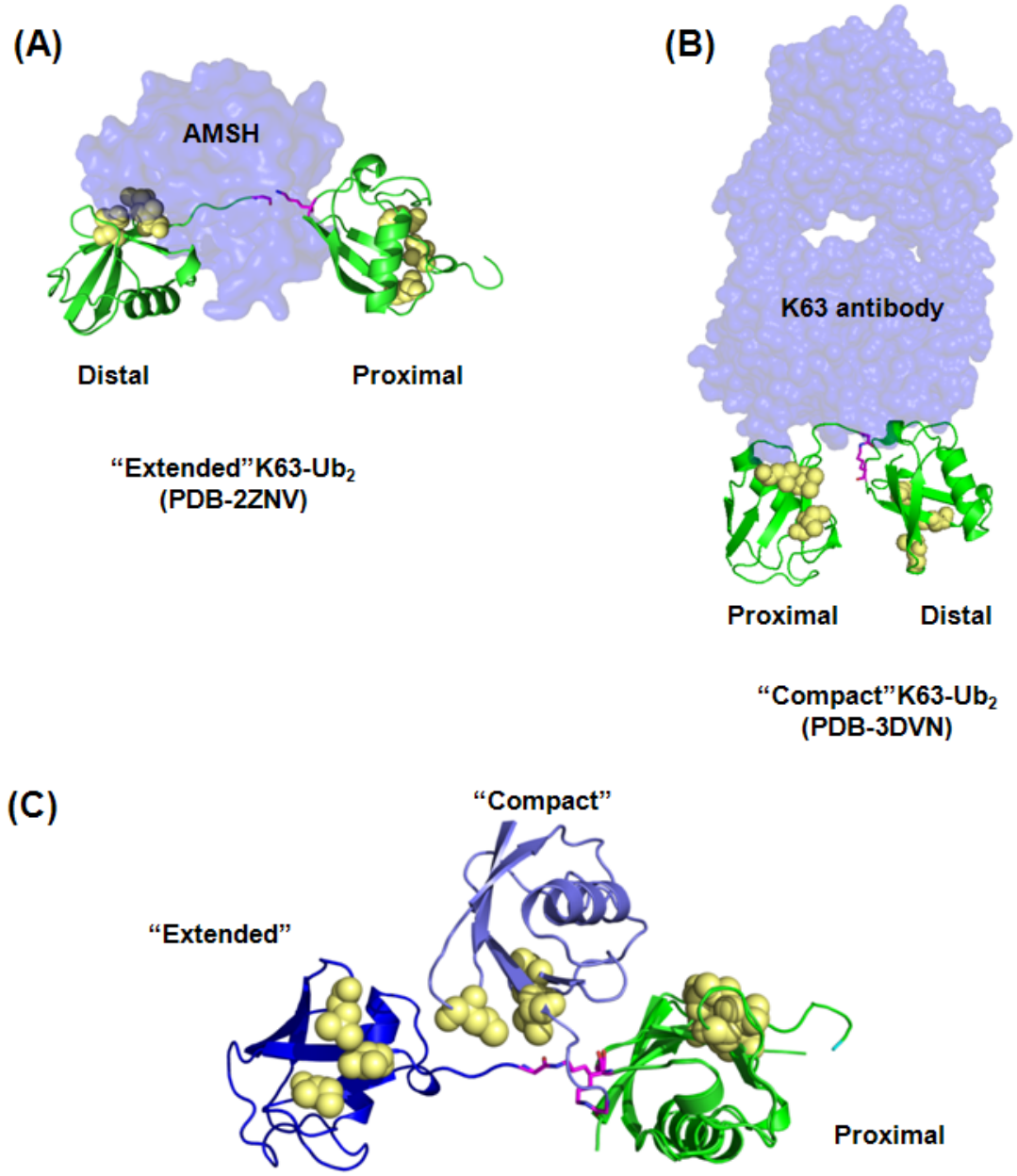


Figure 2.5 – Extended and compact structural conformations for K63-Ub_2
(**A**) The DUB AMSH-LP (blue) binds K63-Ub_2 extending the chain in the process. (**B**) FAB fragment of the K63 linkage specific antibody (blue) creates a compact conformation in the K63-Ub_2 chain. (**C**) Alignment of the extended and compact structures of K63-Ub_2 shows a dramatic difference, however both conformations lack interdomain contacts.

Interestingly, the distinct conformations for each linkage appear to be independent of the number of Ubs in the chain i.e. Ubs linked through K48 will remain compact and Ubs linked through K63 will remain extended in both di and tetra Ub chains (**Figure 2.6**). For technical reasons, longer chains (greater than four Ubs) have not been explored structurally, but modeling suggests that the conformation of the shorter chains will manifest themselves in the longer chains for both linkages.

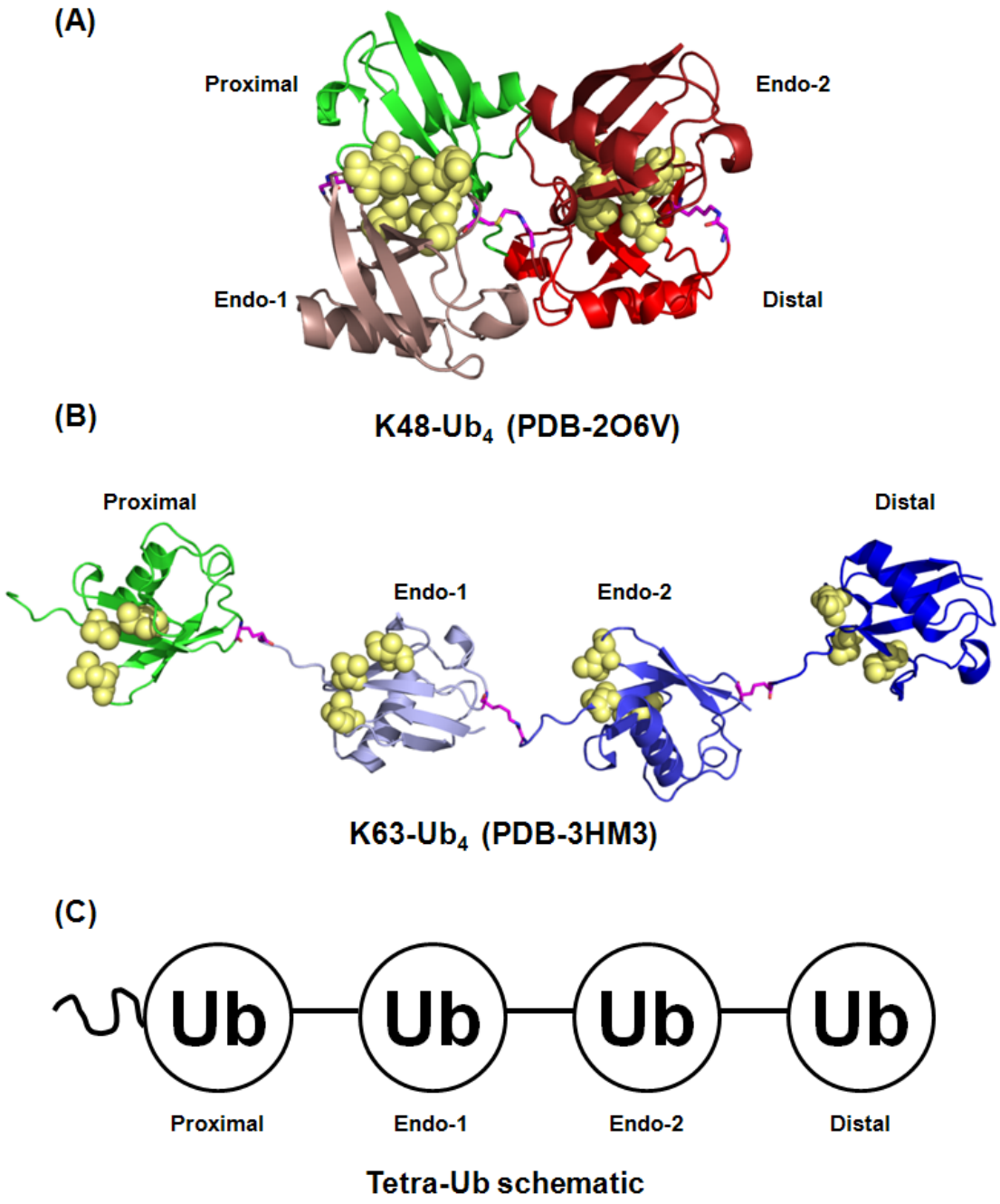


Figure 2.6 – The unique conformations of K48 and K63 linkages are retained in longer chains.
(**A**) Crystal structure (PDB-2O6V) of K48-Ub_4 shows that the interdomain contacts between the L8,I44,V70 hydrophobic patch residues (yellow spheres) is maintained in the tetrameric chain. (**B**) Crystal structure of K63-Ub_4 (PDB-3HM3) shows the Ubs retain their extended conformation with no interdomain contacts. (**C**) Schematic of the linkage pattern for tetra Ub with the only free C-terminus on the proximal Ub.

2.3.2 K11-linked polyUb chains

Following K48 and K63 linkages, K11 is the next most abundant Ub-Ub linkage in the cell. Defined pathways such as ERAD, the UPP, and also mitotic events have been shown to involve K11 linkages (*46, 160*). In addition, the cell has developed specific machinery for K11 linkages such as the DUB Cezanne (OTUD7B) and the E2 Ube2S which are highly selectively for cleaving and synthesizing K11 linkages, respectively (*161*). X-ray derived structures for K11 linked chains (PDB-2XEW and PDB-3NOB) certainly illustrate that these chains adopt different conformations from K48 and K63 linked chains (*160, 161*). However, these "snap shots" of K11-Ub_2 completely ignore dynamics and are dramatically different between each other, shown in (**Figure 2.7**), thus it is hard to accept that either PDB-2XEW or PDB-3NOB provides an accurate representation of K11-Ub_2. A recent solution study by Prof. Fushman (unpublished) has established that K11 chains mainly populate conformations that fall between the compact K48 and extended K63 linkages. Although K11-Ub_2 is slightly compact, it does not appear that the L8,I44,V70 hydrophobic patch forms strong interdomain contacts as it does with the K48 linkage. Furthermore, the K11 linkage has only recently begun to be studied and there are no structures showing any UBD in complex with K11 linked Ub, making it impossible to predict the functional conformations of the K11 linkage.

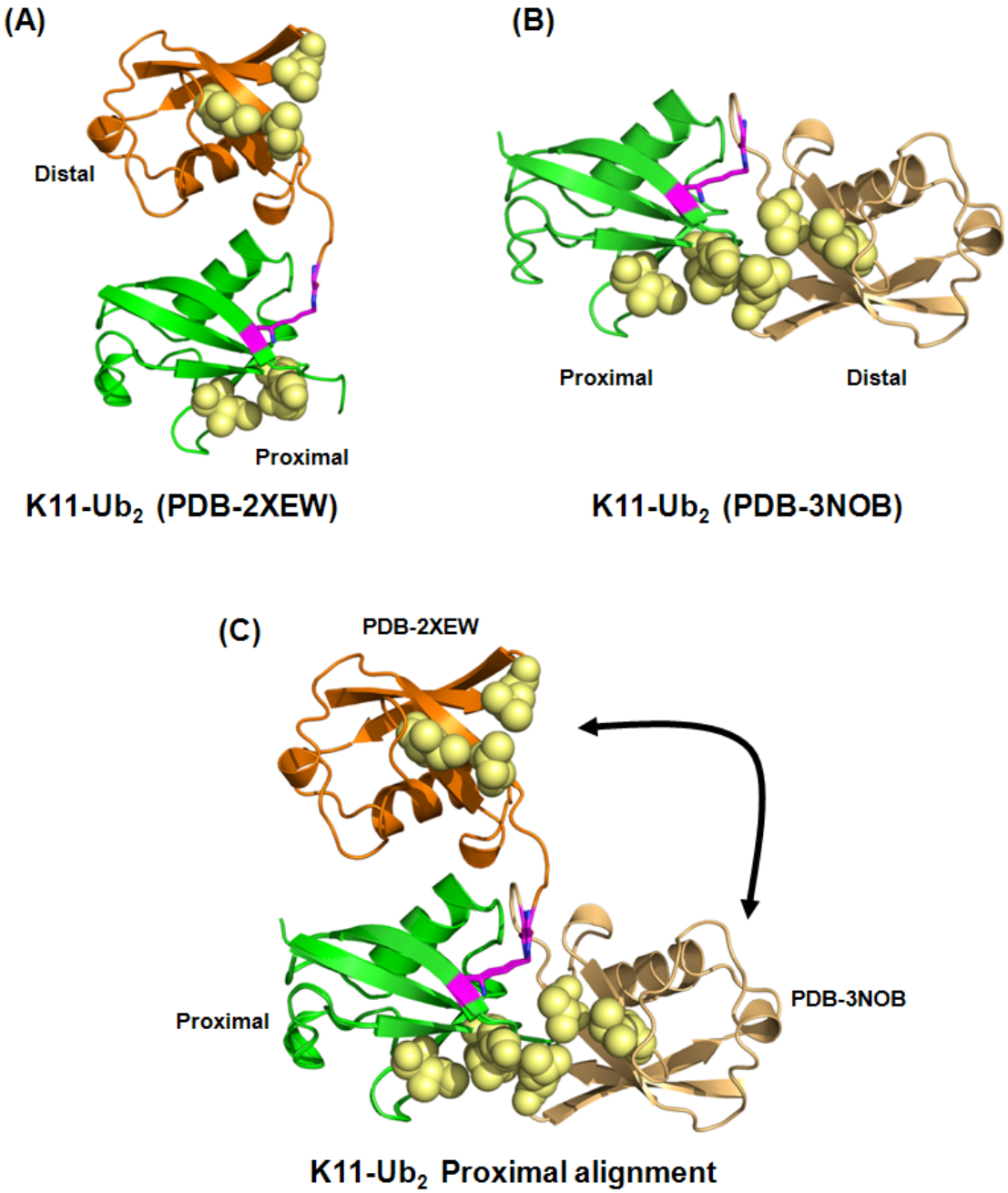


Figure 2.7 – Radically different crystal structures of K11-Ub$_2$
(**A**) PDB-2XEW shows a slightly extended conformation of K11-Ub$_2$ with proximal Ub (green), distal Ub (orange), and hydrophobic patch residues (yellow spheres). (**B**) Contact between hydrophobic patch residues (yellow spheres) in PDB-3NOB, with proximal Ub (green) and distal Ub (beige). (**C**) Alignment of K11-Ub$_2$ structures by proximal Ub (green) shows the distal domain must undergo a large conformational change to satisfy either structure.

2.3.3 K6-linked polyUb chains

The initial discovery of K6-linked polyUb chains sparked great interest since the linkage was found to be created by the BRCA1/BARD1 heterodimer (RING E3 Ub ligase), a key tumor suppressor in breast cancer (*162*). Since then only a handful of other studies have identified pathways or enzymes that create K6 linkages and this is most-likely due to their low (<4%) abundance in the cell (*163, 164*). Interestingly, a HECT-like E3 Ub ligase from bacteria (NLEL) can create mixed K48 and K6 linkages with mammalian E2s (*163*). K6 linked chains have several notable features including an asymmetric interdomain interface. Aside from the K48 linkage, K6 is the only other linkage known that forms significant interdomain contacts between adjacent Ubs in the same chain (**Figure 2.8**). Unlike the structures reported for K11-Ub_2, the K6-Ub_2 structure is in good agreement with data from solution NMR and the conformation creating the new interface accounts for a major proportion of all conformations. Given the asymmetric interface of K6 there is much speculation regarding how longer K6 linked chains would appear. For K48 linkages tetrameric chains retain their same interaction, however it may be possible for longer K6 linked chains to take one some unique conformations. There are no known K6 specific receptors and the K6 linkage has also been found to be resistant to several DUBs (*20, 45*).

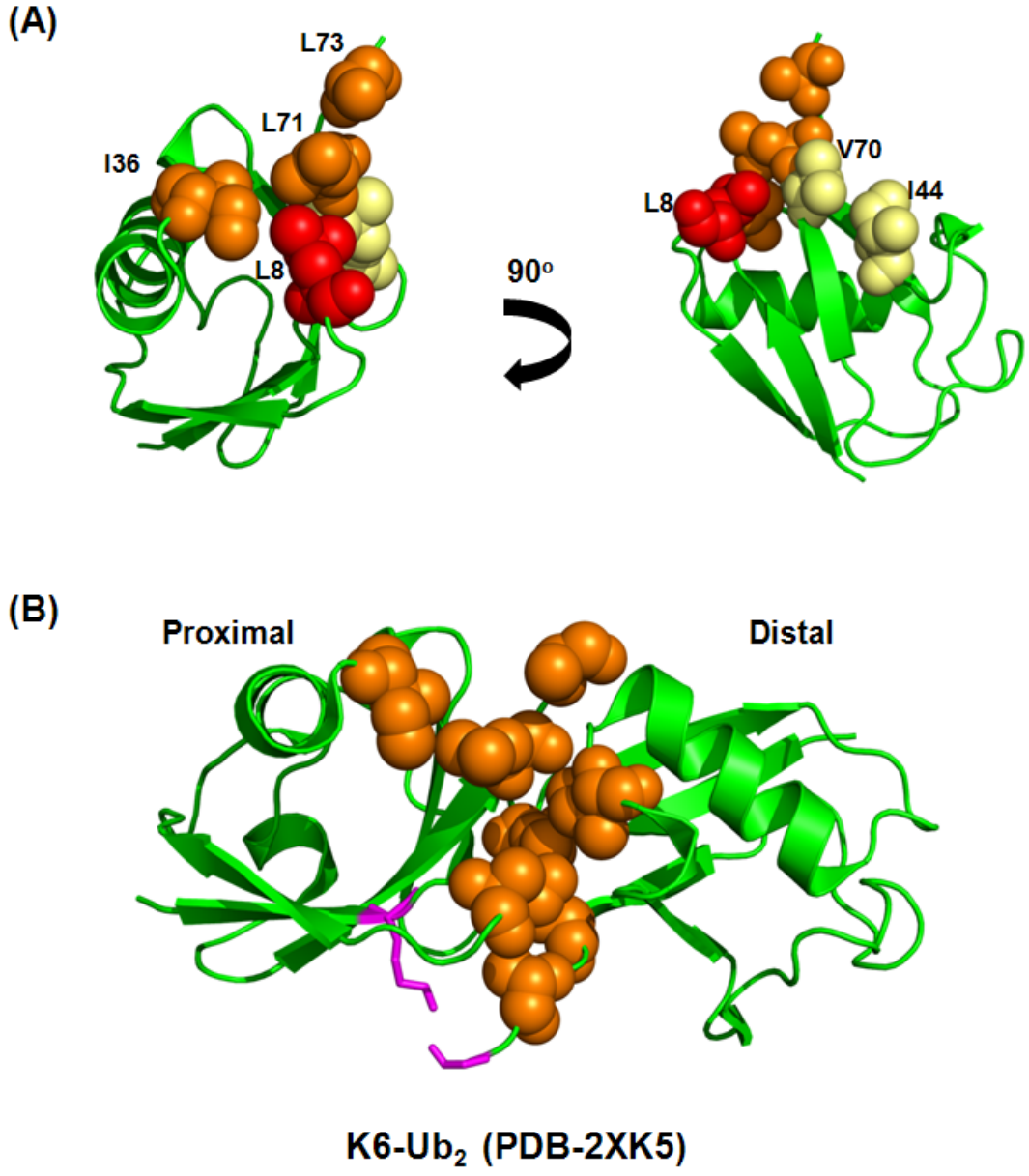


Figure 2.8 – Crystal structure of K6-Ub_2 revels a new interface
(**A**) Two distinct hydrophobic patches on Ub, the L8, I36, L71, L73 patch (orange spheres) and L8, I44, V70 patch (yellow spheres) with L8 as red spheres. (**B**) Crystal structure (PDB-2XK5) of K6-Ub_2. As opposed to the symmetric K48 interface formed by the L8,I44,V70 patch, the asymmetric K6 interface is formed by the L8,I36,L71,L73 patch (orange spheres).

2.3.4 M1-linked polyUb chains

M1 linked chains are also referred to as "head-to-tail" or "linear chains," which can be confusing. For simplicity I will use the unambiguous M1 designation to refer to such chains. In the cell, all Ub starts as M1-linked chains when it is expressed from the Ub gene, that encodes a fusion of four M1-linked Ubs (*165*). DUBs from the UCH family help process the M1 linked Ub gene product into monomeric Ub units. However if M1 linked Ub chains are needed they can be reassembled by the linear ubiquitin chain assembly complex (LUBAC), which contains two RING E3's haem-oxidized iron-regulatory protein 2 ubiquitin ligase-1 (HOIL-1) and HOIL-1 interacting protein (HOIP) (*166*). Similar to K63 linked chains, M1 linked chains adopt an extended conformation with minimal interdomain contacts (**Figure 2.9**). Although M1 linked chains are similar to K63 linked chains in structure, they appear to be used for very different pathways *in vivo*. NF-κB activation is one of the only characterized pathways for M1 linkages (*167*). One this pathway M1 linkages are recognized by the ubiquitin binding in ABIN and NEMO (UBAN) domain in the IKK subunit NF-kappa-B essential modulator (NEMO). Structures of M1 linked dimers in complex with a linkage specific antibody and UBAN of NEMO are shown in (**Figure 2.9 D,E**). These examples show another similarity between K63 and M1 linked chains, in that upon binding, the conformation of the chain is subject to much change from the unbound state.

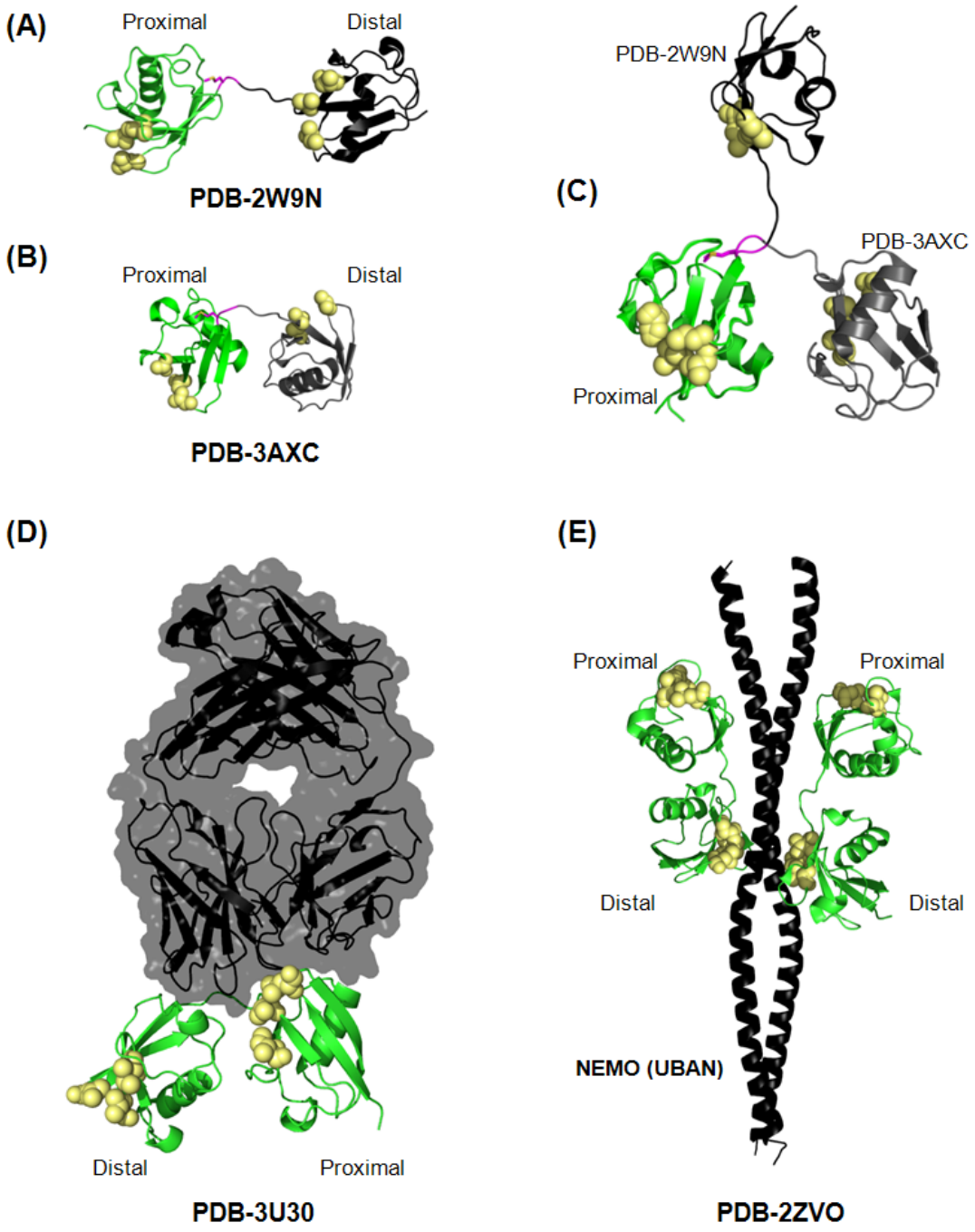


Figure 2.9 – Representative structures of M1 linked polyUb chains.
(**A**) Crystal structure (PDB-2W9N) shows and extended conformation of M1-Ub_2 with proximal Ub (green), distal Ub (black), the L8,I44 hydrophobic patch (yellow spheres), and M1-G76 linkage (magenta sticks). (**B**) Compact crystal structure (PDB-3AXC) of M1-Ub_2, same coloring as (**A**) with the exception of the distal Ub (gray). (**C**) Alignment of the extended (black) and compact (gray) M1-Ub_2 structures by the proximal Ub (green). (**D**) M1 linkage specific antibody (black) recognizes a compact form of M1-Ub_2 (**E**) The UBAN motif from NEMO in complex with M1-Ub_2

2.3.5 Cyclized polyUb

The possibility of cyclized polyUb chains was first proposed when such chains were discovered as a product from the *in vitro* enzymatic synthesis of K48 chains by E2-25K (*168*). To date, K48 linkages have been the only linkage found to allow for cyclization of the Ub chain. However, synthetic chemists have reported that other linkages with artificial modifications can be cyclized (*169*). A recent study has eluded that such chains exist in the cell and may even play specific roles (*170*). DUBs designed to process cyclized chains also support an *in vivo* function for cyclized chains (*170*). Based on the crystal structure (PDB-3ALB) of cyclic K48-Ub_4 and several solution studies, it appears that cyclization of K48 linked Ub results in a similar conformational ensemble to non-cyclized K48 linked chains (*171, 172*). The tight locking of the L8, I44,V70 hydrophobic patch between the Ub units is the most pronounced feature, however the interdomain interaction and conformation are nearly identical to non-cyclized K48 linked chains in that the overall compact conformation is preserved (**Figure 2.10**). Essentially cyclization of K48 linked chains locks the closed conformation and does not impact the overall structure. Even when cyclized it appears the L8,I44,V70 hydrophobic patch is still slightly accessible, which hints that cyclized K48 chains can still participate in binding interactions (*172*). This theory remains to be tested, but there certainly will be some difference in binding given the effects of cyclization on the conformational equilibrium of K48 linked chains.

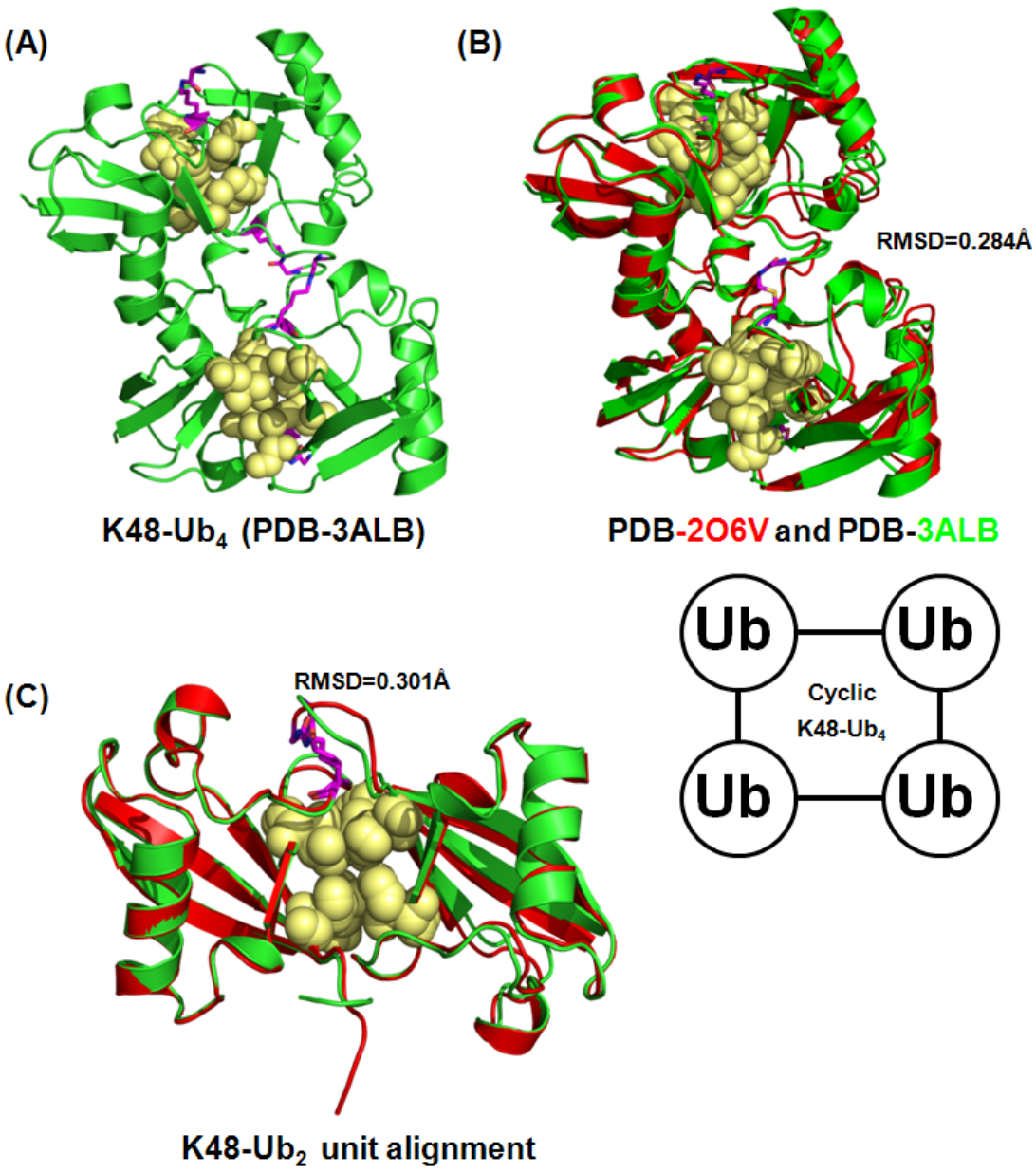


Figure 2.10 – Structural similarities between cyclized and non-cyclized K48 polyUb chains.
(**A**) Crystal structure (PDB-3ALB) of cyclized K48-Ub_4 with all Ubs (green), G76-K48 linkage (magenta sticks), and the L8,I44,V70 hydrophobic patch (yellow spheres). (**B**) Alignment of cyclized (PDB-3ALB) in green and non-cyclized (PDB-2O6V) in red K48-Ub_4 shows similarity, with a RMSD=0.284 Å. (**C**) The K48-Ub_2 unit is taken from each K48-Ub_4 structure and aligned showing the hydrophobic interface in the non-cyclized form (red) is well maintained in the cyclized form (green).

2.3.6 K27, K29, K33, and other linkages

The true non-canonical linkages are K27, K29, and K33 which are all located on Ub's long α-helix. No structural data exists for these linkages, however it is proposed that they form extended conformations similar to M1 or K63 linkages, while it is also possible a few could form transient interdomain contacts (*173*). Together these three linkages account for slightly over 10% of all linkages in yeast and have been reported to exist in low abundance for higher eukaryotes (*174*). The close proximity of the linkages has made it hard to create antibodies for their detection. Unlike M1, K6, K11, K48, and K63 linkages which have specific antibodies, no attempt has been successful in creating a reliable K27, K29, or K33 linkage specific antibody. K27 linkages have been detected *in vivo* by MS studies, but there has yet to be a defined role attributed to this linkage type. Some enzymes, such as Ubch5b (an E2), the E3 ligase Ring1b, and the E4 protein Ufd2 can create K27 linkages as well as other linkages, resulting in mixed linkage chains (*20, 164, 175*).

More studies have elucidated the roles of K29 and K33 linkages. The best characterized role for K29 linkages is signaling lysosomal degradation (*176*). K29 and K33 linkages have been found to modify several AMPKs (AMP-activated protein kinase), which are essential for cellular proliferation. Interestingly, USP9x is one of a few DUBs that can efficiently remove K29 and K33 modifications (*177*). A more extensive study on the linkage preferences of DUBs has revealed the OTU DUB TRABID preferentially cleaves K29 and K33 linkages suggesting the linkages do indeed have distinct signaling pathways (*178*). It is also speculated that there are

selective UBDs for either K29 or K33 linkages, but this is still to be determined. The fact that K27, K29 and K33 are in such close proximity hints that they could be recognized similarly. Studies by Goldberg and coworkers demonstrated that it was possible to generate "forked" i.e. simultaneously attach two Ubs at K27&K29 or K27&K33 on a proximal Ub with Ubch5b, an E2 conjugating enzyme (*29*). This work supports that the α-helix of Ub can potentially carry many different signals.

Ubiquitination does not always occur at the N-terminus, lysine side chain, or the active cysteine of enzymes. The mouse gamma-herpesvirus protein mK3 encodes a special E3 Ub ligase that can attach Ub to serine or threonine residues, as well as the traditional lysine residues of its substrate, major histocompatibility complex I (MHC-1) heavy chain (*179, 180*). The advantage of ligating an Ub to serine or threonine residues is not fully understood, but this does add another dimension to Ub proteomics. Attachment of Ub to cysteine residues results in a high energy bond that has favorable energetics for Ub transfer between E1, E2, and E3 enzymes. By definition ubiquitination with a HECT-E3 requires the transfer of Ub to flow from the E1, then the E2, next to the E3 and finally to the substrate, resulting in three Ub-cysteine thioester linkages. As an autoregulatory mechanism, E2's such as Ubc7 can even become polyubiquitinated at active site cysteine residues, targeting themselves for degradation by the proteasome (*181*).

Chapter 3: Materials and Methods

3.1 Protein expression and purification

3.1.1 Plasmid construction and design

All double stranded (ds) DNA plasmids contained a commercially available vector and a single gene sequence for expression, optimized for *E. Coli.* When a plasmid was needed it was grown and then isolated from chemically competent DH5α cells. To express gene products several variations of BL-21(DE3) cells were used. For large proteins with established protocols, expression was carried out in BL-21(DE3) pLysS cells. BL-21(DE3) Rosetta™ II cells were used to express a majority of human proteins under the T7 promoter with rare codons for *E. Coli.* Noted by Prof. CM Pickart it was absolutely essential to express Ub mutants containing KtoR mutations with rare arginine codons AGA and AGG in Rosetta II cells (*182*). In addition, Rosetta II cells also have a library of tRNAs to ensure other rare codons are properly introduced into the expressed protein. A subset of proteins under the T5 promoter in the pQE30 vector (Qiagen) were expressed in M15 *E. coli* cells.

With just one exception, all gene sequences were already contained in plasmids for expression and required no additional cloning or ligations. Certain plasmids were obtained through collaborators or via the addgene data base. Many point mutations were introduced on various plasmids for different purposes. This was accomplished using mutagenic primers designed with primerX software and PCR with KOD high fidelity hot start DNA polymerase (Novagen). Truncations were also

introduced with the insertion of a stop codon. Up to eight residues were inserted to particular constructs (e.g. 6xHis tag or strep-tag), again using similar PCR conditions in 50 µL reactions. Following PCR all template DNA was digested with 15 U of Dpn1 (New England Biolabs) for 3 hours at 37^{o}C following the manufacturer's instructions. PCR products were checked on a 1% agarose gel with ethidium bromide staining, then transformed into DH5α cells. The plasmid was then extracted and the presence of the desired product was confirmed by sequencing with the appropriate primer.

3.1.2 Bacterial growth conditions and expression

Detailed growth and expression conditions for 87 different proteins are outlined in (**Table 3.1**). There were several different media employed based on the need for autoinduction and isotope labeling. Highly stable proteins such as Ub and GST constructs could be expressed in autoinducing media which required just a single temperature for incubation. 6xHis tagged proteins generally required induction with IPTG, however IPTG concentration and the expression temperature needed to be optimized for each protein. Antibiotics were added from 1,000x stock solutions to a final concentration in culture: chloramphenicol (50 mg/L), ampicillin (100 mg/L), and kanamycin (50 mg/L).

	Protein	plasmid	promoter	auto inducing	Induce [IPTG]	OD600 Induction	Expression temperature (C)	Expression time (hr)	Affinity tag
1	Ub	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
2	Ub-F45W	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
3	Ub-K48R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
4	Ub-K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
5	Ub-D77	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
6	Ub74	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
7	Ub-G76C	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
8	6xHis-Ub	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
9	Ub-T12C	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
10	Ub-I36C	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
11	Ub-K48C	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
12	Ub-T55C	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
13	Ub-K63D	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
14	Ub-6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
15	Ub-strep	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
16	Ub-6xHis/K48R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
17	Ub-K11R/K48R/K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
18	Ub-K48R/K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
19	Ub74/K48R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
20	Ub74/K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
21	Ub-D77/K48R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
22	Ub-D77/K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	none
23	Ub-K0	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
24	Ub-K6	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
25	Ub-K11	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
26	Ub-K27	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
27	Ub-K29	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
28	Ub-K33	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
29	Ub-K48	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
30	Ub-K63	pET14b	T7	YES	1mM	0.7-1.0	37	4	none
31	Ub-K0/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
32	Ub-K6/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
33	Ub-K11/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
34	Ub-K27/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
35	Ub-K29/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
36	Ub-K33/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
37	Ub-K48/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
38	Ub-K63/6xHis	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
39	MMS2	pET16b	T7	YES	.8mM	0.7	37	4	6xHis
40	GST-Ubc13	pGEX4T2	T7	YES	1mM	0.7-1.0	37	4	GST
41	E2-25K	pGEX4T2	T7	YES	1mM	0.7-1.0	37	4	GST
42	Ube2S	pMAL	T7	NO	1mM	0.7-1.0	37	4	MBP/ 6xHis
43	human E1	pet15b	T7	NO	0.5mM	0.7	16	16	6xHis
44	Ubch5b	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
45	Ubch5b-C85A	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
46	Ubch5b-C85K	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
47	Ubch5b-K4R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
48	Ubch5b-K8R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
49	Ubch5b-K63R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
50	Ubch5b-K66R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis

Table 3.1 – Proteins and expression conditions.

	Protein	plasmid	promoter	auto inducing	Induce [IPTG]	OD600 Induction	Expression temperature (C)	Expression time (hr)	Affinity tag
51	Ubch5b-K101R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
52	Ubch5b-K128R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
53	Ubch5b-K133R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
54	Ubch5b-K144R	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
55	Ubch5b-K4	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
56	Ubch5b-K8	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
57	Ubch5b-K63	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
58	Ubch5b-K66	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
59	Ubch5b-K101	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
60	Ubch5b-K128	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
61	Ubch5b-K133	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
62	Ubch5b-K144	pET3a	T7	YES	1mM	0.7-1.0	37	4	6xHis
63	RAP80-tUIMy	pET28a	T7	NO	1mM	0.8	37	5	6xHis
64	RAP80-v7-tUIMy	pET28a	T7	NO	1mM	0.8	37	5	6xHis
65	hHR23a-UBA(2)	pGEX	T7	YES	1mM	0.7	37	5	GST
66	UBQ1-UBA	pGEX	T7	YES	1mM	0.7	37	5	GST
67	RPN1	pQE30	T5	NO	0.1mM	0.8	16	16	6xHis
68	DSK2	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
69	DSK2-Sti1	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
70	DSK2-UBA	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
71	DSK2-UBL	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
72	DSK2-ΔUBL	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
73	DSK2-ΔUBA	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
74	Rpn10	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
75	Rpn10-UIM	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
76	Rpn10-VWA	pQE30	T5	NO	0.5mM	0.7	20	16	6xHis
77	Ubp6	pQE30	T5	NO	0.5mM	0.8	18	18	6xHis
78	Ubp6ΔUBL	pQE30	T5	NO	0.5mM	0.8	18	18	6xHis
79	Ubp6-C129A	pQE30	T5	NO	0.5mM	0.8	18	18	6xHis
80	YUH1	pET21	T7	NO	0.75mM	0.7	37	4	none
81	AMSH	pGEX-6p-1	T7	NO	0.35mM	0.8	16	16	GST
82	OTUB1	pProEx	T7	NO	0.5mM	1mM	16	16	6xHis
83	Rpn11	pQE30	T5	NO	.3mM	0.8	16	16	6xHis
84	Rpn11 D122A	pQE30		NO	.3mM	0.8	16	16	6xHis
85	Rpn8	pQE30	T5	NO	.3mM	0.8	16	16	6xHis
86	Rpn8 (1-186)	pQE30	T5	NO	.3mM	0.8	16	16	6xHis
87	Rpn9	smt3-X	T5	NO	0.5mM	0.8	20	12	SUMO / 6xHis

Table 3.1 (continued) – Proteins and expression conditions

1L of the autoinducing media for unlabeled proteins was prepared by combining 10 g NaCl, 10 g tryptone, 5 g yeast extract, 50 mL of 20x NPS, and 1 mL of 1M $MgSO_4$ in a final volume of 1L. 20x NPS contains 90 mL of H_2O, 6.6 g $(NH_4)_2SO_4$, 13.6 g KH_2PO_4, and 14.2 g Na_2HPO_4. After the 1L culture was sterilized, 25 mL of 50x ZYP5052 was syringe filtered (0.45 μm) into the culture. 50x ZYP5052 is composed of 25 g glycerol, 73 mL H_2O, 2.5 g glucose, and 10 g lactose in a final volume of 100 mL. The antibiotic(s) and starter culture were added last. Cells were harvested 16-22 hours after allowing growth at a constant temperature.

For isotope labeling in autoinducing media 1 mL of 1M $MgSO_4$ was added to 50 mL of 20x (^{15}N)NPS in a final volume of 1L then autoclaved. The 20x (^{15}N)NPS contains (90 mL H_2O, 2.84 g Na_2SO_4, 13.6 g KH_2PO_4, 14.2 g Na_2HPO_4). After sterilization 25 mL of 50x ZYP5052 was syringe filtered (0.45 μm) into the culture, followed by a solution containing 1 g $^{15}NH_4Cl$ dissolved in 10 mL of H_2O. The antibiotics were then added and the starter culture was pelleted, then resuspended in the 1L culture to avoid any incorporation of ^{14}N. Once prepared the cells were allowed to grow for 16-22 hours.

Standard Luria broth (LB) expression was achieved by mixing 10 g NaCl, 10 g tryptone, 5 g yeast extract, and 1 mL of 1M $MgSO_4$ in a final volume of 1L and autoclaving. Starter cultures and antibiotics were then added. The OD_{600} was monitored every 30 minutes after an initial growth period of 3 hrs. When the culture reached the desired OD_{600} the appropriate amount of IPTG was added. If needed for expression, the temperature of the incubator was also changed. For growths at 37^o C a time of 4 hrs was sufficient for expression, but for low temperature expression times of 16-24 hours were necessary, (see **Table 3.1**).

$^{13}C/^{15}N$, ^{13}C, and ^{15}N labeled proteins that could not express with autoinducing media were expressed in M9 media. 6 g Na_2HPO_4, 3 g KH_2PO_4, and 0.5 g NaCl were added in a final volume of 1L, then autoclaved. Next all nutrients were combined in a 50 mL conical and vortexed into solution. If ^{15}N labeling was desired, 1 g of $^{15}NH_4Cl$ was added in place of $^{14}NH_4Cl$. For ^{13}C labeling 3 g of ^{13}C glucose was added, while 5 g of standard glucose was added for all other proteins as the carbon source. The final for concentration for the rest of the nutrients was 1 mg/L

of biotin, 1 mg/L of thiamin, 50 μM $CaCl_2$, and 0.5 mM $MgSO_4$. After these nutrients were syringe filtered into the culture, the antibiotics and pelleted starter culture were added. The OD_{600} was checked frequently after an initial 5 hour growth period due to the slow growth in M9 media. The proper amount of IPTG was added when the optimal OD_{600} was reached.

After induction the cultures were harvested at 4,000 rpm. The cell pellet was immediately frozen at -80°C or prepared for lysis. The buffer used for lysis varied depending on the affinity tag and next steps for purification, GST fusion proteins were suspended in PBS pH 7.4, the buffer for 6xHis tagged proteins was 20mM phosphate, 0.5M NaCl, pH 7.4, and for untagged proteins (mostly Ub variants) 50mM Tris, pH 8.0. The lysis solution contained the cell pellet resuspended in ~25 mL of the appropriate buffer along with 0.02% Triton X-100, lysozyme (0.4 mg/mL), DNase1 (20 μg/mL), 10 mM $MgCl_2$, and a homemade protease inhibitor cocktail with 1 mM PMSF, 50 μM TLCK, soybean trypsin inhibitor (5 μg/mL), and leupeptin (2.5 μg/mL). This lysis mixture was lysed on ice by three rounds of a two minute sonication followed by a two minute recovery. Due to the excess heating and stress from sonication, some proteins were lysed using the French press method. After lysis the debris was cleared by spinning the mixture in an ultracentrifuge at 22,000 rpm on a Ti-45 rotor for 20 minutes. From lysis to every subsequent step, the protein and buffer solutions were kept with 0.02% (v/v) NaN_3 to inhibit the growth of unwanted microbes. Following lysis proteins were purified using affinity columns or using other methods for untagged proteins, see below.

3.1.3 Purification of highly stable Ub and Ub mutants

Directly after lysis, the cleared supernatant containing a stable Ub variant was transferred to a 50 mL beaker with a stir bar on ice. With constant stirring, 70% perchloric acid ($HClO_4$) was added drop-by-drop to a final concentration of 1% (v/v), usually around 300 µL for 30 mL of lysis supernatant. This step is unique in that the addition of $HClO_4$ drops the pH to 1.85 where many proteins precipitate, however Ub remains stable. If slightly unstable Ub mutants were being purified a variation of this technique substituting 3% (v/v) glacial acetic acid in place of 1% (v/v) $HClO_4$ could be used, which dramatically increased yields. Following the acid precipitation step, the milky white precipitate was cleared by ultracentrifugation, 22,000 rpm for 15 minutes. The supernatant was then transferred into 3 kDa MWCO dialysis tubing and dialyzed against 2L of 50 mM ammonium acetate, pH 4.5 at 4^{o}C for 8-12 hours. The Ub solution was then moved from the dialysis tubing, syringe filtered (0.45 µm), and loaded on to a pre-equilibrated 5 mL cation column (GE Life Sciences, SP FF) at a flow rate of 1.5 mL/min with buffer A (50 mM ammonium acetate, pH 4.5). Ub was eluted using a program that went from 0% to 35% buffer B (50 mM ammonium acetate, 1M NaCl, pH 4.5) over 20 cv and the major Ub peak was detected a 16% buffer B. The major peak was then pooled, concentrated, and exchanged into a desired buffer (50 mM Tris pH 8.0 for chain synthesis, 20 mM sodium phosphate pH 6.8 for NMR studies, or PBS pH 7.4 for assays) using a centrifugal unit with a 3,500 kDa MWCO. The purity of Ub was checked by 15% SDS-PAGE and if needed the Ub was further purified on a superdex 75 120 mL size exclusion column (GE Life Sciences) with a flow rate of 0.5 mL/min in PBS pH 7.4 buffer. After acid

precipitation and the cation exchange step most Ub preparations were highly pure and did not require the size exclusion step. This protocol is virtually identical to those established by Prof. CM Pickart (*183, 184*). Wild type Ub, KtoC mutants, and a majority of the KtoR Ub mutants were purified following this method.

3.1.4 Purification of marginally stable Ub mutants

For Ub mutants unable to withstand the harsh acid precipitation, I developed another purification method. Eventually a 6xHis tag was added to these mutants as an alternative, but for many the lack of an affinity tag was advantageous. After lysis, the supernatant was added to a 50 mL conical suspended in a 65^{o}C water bath. With the top of the 50 mL conical perforated to release pressure. The solution was allowed to incubate at 65^{o}C for 15 minutes and a white precipitate was observed. Immediately after heating, the 50 mL conical was placed on ice for 12 minutes to trap unfolded proteins before they could refold. The debris was cleared by ultracentrifugation, 22,000 rpm for 20 minutes. The supernatant was then syringe filtered into a fresh 50mL conical which was used to load it on to a pre-equilibrated 5mL anion exchange column (GE Life Sciences Q FF) at 1 mL/min in buffer A (50 mM Tris, pH 8.0). In the anion step, Ub flows through and does not bind while many other proteins bind the anion column. The Ub flow through was pooled and dialyzed against 2L of cation buffer A (50 mM ammonium acetate, pH 4.5) for 8-12 hours at 4^{o}C. Following dialysis the Ub solution was loaded onto a pre-equilibrated 5 mL cation column at flow rate of 1.0 mL/min. Marginally stable Ub mutants were eluted just as in **section 3.1.3** above, but usually required a size exclusion step for high

purity. K0-Ub, single lysine Ubs (e.g. K63), and the F45W mutants were all originally purified with this protocol.

3.1.5 YUH1 purification

YUH1, an important DUB for removing D77 in chain synthesis was purified without a tag using a modified protocol from Cohen and coworkers (*40*). Following lysis, YUH1 was syringe filtered onto a 10 mL anion column in buffer A (50mM Tris, pH 7.6). Then YUH1 was eluted with 5 cv steps of 15%, 30%, 50%, and 100% buffer B (50mM Tris, 1M NaCl, pH 7.6). The 30% buffer B fractions containing the 25 kDa YUH1 were pooled, exchanged in to PBS pH 7.4 and further purified with a superdex 75 size exclusion column at 0.4 mL/min. This method, initially developed by Dr. Daoning Zhang efficiently isolated highly active YUH1 without the need for salt precipitation or reverse-phase chromatography. To remove D77 from monomeric Ub or polyUb chains containing D77 at the proximal end, the Ub species were incubated with a 1% molar ratio of YUH1 at 37^{o}C for 2 hours in a buffer containing 50 mM Tris pH 7.6 and 1mM EDTA. The concentration of the D77 Ub was kept less than 1 mM for all reactions. After the incubation, 8 mL of anion buffer A (50mM Tris, pH 7.6) was added and the mixture, then slowly hand injected on to a 1 mL anion column (GE Life Sciences Q FF) column. An additional 5 mL of anion buffer A was slowly injected to remove the unbound proteins and both Ub containing flow through portions were collected. The YUH1 that bound the anion column was eluted in buffer B (50mM Tris, 1M NaCl, pH 7.6) while the Ub that did not bind flowed through effectively separating YUH1 from Ub.

3.1.6 Purification of standard proteins (6xHis and GST)

Many proteins simply required lysis and affinity purification with the only caveat being expression. After lysis 6xHis and GST fusion proteins were purified on a 5 mL His-Trap column (GE Life Sciences) or 10 mL GST column (GE Life Sciences), respectively. Once ideal expression conditions were determined, all of the GST proteins listed in (**Table 1.3**) were easily purified following the manufacturer's GST protocol, loading in buffer A (PBS pH 7.4) and elution with buffer B (10 mM glutathione, 50 mM Tris, pH 7.6). 6xHis tagged proteins were loaded in buffer A (20 mM sodium phosphate, 0.5 M NaCl, pH 7.4) and eluted over an 8 cv gradient from 10-100% buffer B (20 mM sodium phosphate, 0.5 M imidazole, 0.5 M NaCl, pH 7.4). Following purification, many proteins were found to function fine with the affinity tag left on so no effort was wasted removing it. For example, the 6xHis tag on Mms2 was left on for K63 chain synthesis, the 6xHis tag was on tUIM of RAP80 for binding assays, and GST-fusions GST-Ubc13 and GST-E2-25K were used in chain synthesis. However, GST did need to be removed from GST-UBA fusions and this was accomplished by a thrombin cleavage site and separating the 5 kDa UBA from 23 kDa GST on the size exclusion column.

3.2 Enzymatic Synthesis and purification of polyUb chains and conjugates

3.2.1 Enzymatic synthesis of Ub_2 chains

With the exception of M1 linked chains that could be produced as linear fusions, all Ub-Ub linkages involving a lysine residue were created enzymatically from purified monomeric units. Standard reactions were in volumes of 2 mL

primarily containing 50 mM Tris pH 8.0 and incubated in a 30^{o}C water bath for 24 hours. Each reaction contained 4 mM TCEP in place of DTT. The reason being was that at high concentrations DTT could disrupt the E1 or E2 thioester bonds with Ub, however TCEP selectively reduces only disulfides. TCEP contains three carboxylic acid groups and the pH of the 0.1 M stock solution was adjusted from pH 4 to pH 7 with NaOH. 20 mM ATP was enough to keep the E1 activating enzyme highly functional. To regenerate ADP to ATP and provide Mg^{+2} for E1, a 5x regeneration system resulting in a final concentration of 15 mM $MgCl_2$, 20 mM creatine phosphate, 1.2 U/mL inorganic yeast pyrophosphate, and 1.2 U/mL creatine phosphokinase was used. Due to the high activity of our human E1, concentrations of 500 nM were found to be ideal. The rest of the 2 mL reaction was made up of the Ub monomers and the E2 conjugating enzyme. For dimers, 15 mg of each monomer was added and for trimers (dimer + monomer) a slight 1.25 molar ratio excess of the desired monomer was added to ensure the best yield of the trimer. Reactions designed to produce a distribution of chain lengths with wild type Ub used 30 mg. In these particular reactions, the E2 dictated the linkage formed. For K11 linkages MBP-Ube2s was used at a final concentration of 20 μM, K48 linkages made with GST-E2-25K were also at 20 μM, while the Ubc13:Mms2 heterodimer used both proteins at 10 μM to create K63 linkages, see (**Figure 3.1**) for chain synthesis schemes.

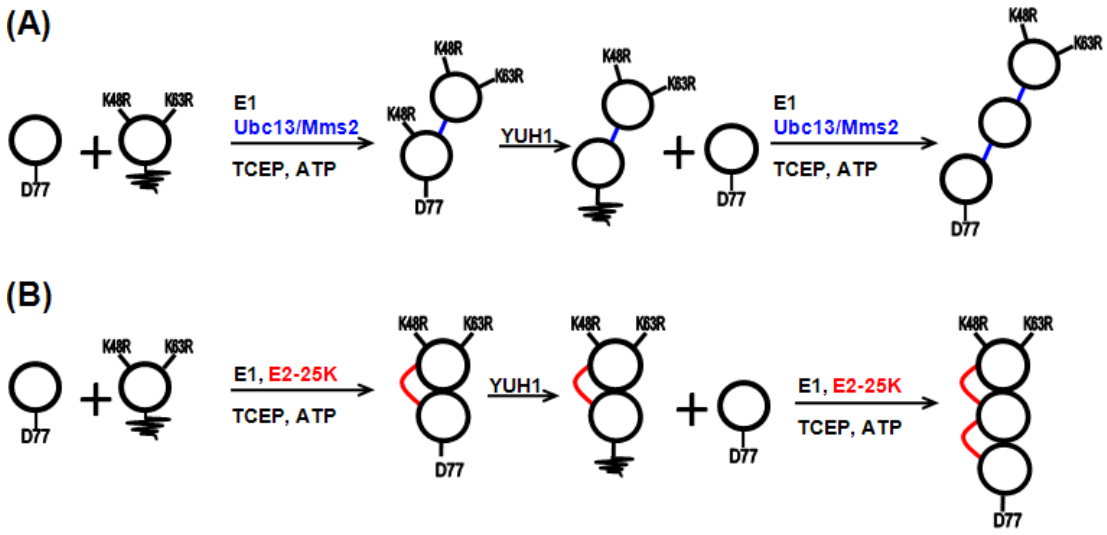


Figure 3.1 – Synthesis schemes for K48 and K63 linked di and tri Ub chains
The simple enzymatic synthesis scheme can be used to make (**A**) K48-Ub_2 and K48-Ub_3 or (**B**) K63-Ub_2 and K63-Ub_3. For each linkage the D77 blocking extension from the proximal end can be removed with YUH1. The K48R/K63R double Ub mutant prevents extension from the distal end.

After the 24 hour incubation period, the reactions were either stored at -20^{o}C or 5 mL of cation buffer A (50 mM ammonium acetate, pH 4.5) and 100 µL of glacial acetic acid was added for immediate purification. 2 minutes after the addition of the acetic acid the reaction was spun down at 13,000 rpms on a microcentrifuge to remove precipitated proteins. The supernatant was then added to a 5 mL loop and injected to a pre-equilibrated 5 mL cation exchange column at 0.25 mL/min in buffer A (50 mM ammonium acetate, pH 4.5). Other proteins flowed through and were monitored by UV. After the UV baseline stabilized, the polyUb chains bound to the cation column were eluted using 100% buffer B (50 mM ammonium acetate, 1M NaCl, pH 4.5). The eluted polyUb was collected then exchanged into PBS pH 7.4 for size exclusion (0.35 mL/min flow rate) to separate the polymeric chains. The size exclusion superdex 75 resin failed to resolve polyUb chains longer than four Ubs, but could separate distributions of significantly longer chains from shorter chains (e.g.

Ub_{1-4} from Ub_{7-10}). Fractions containing pure dimers or trimers were verified with 15% SDS-PAGE, concentrated in a 3,500 MWCO centrifugal unit, and exchanged into a different buffer is necessary. In some cases triUbs eluted such that they overlapped with the dimer on the size exclusion column. To take the burden off of the size exclusion column a long cation separation which can isolate di and triUbs was preformed prior to size exclusion. For this long cation step polyUb reactions were transferred into 5 mL of cation buffer A and slowly loaded onto a 5 mL cation column at 0.20 mL/min. After the UV base line stabilized, polyUb was eluted using a shallow gradient from 0% to 30% buffer B over 80 cv. Fractions containing the desired chain length were pooled and exchanged into PBS pH 7.4 for size exclusion. This scheme allowed for the synthesis and purification of virtually any polyUb dimer or trimer. Variations of enzymatic synthesis for mixed linkage polyUb chains (see chapter 4) were also successful.

3.2.2 Autoubiquitinated E2-Ub conjugates

Several E2 conjugating enzymes have been reported to autoubiquitinate on their active site cysteine as well as lysine residues. Given that the Ub thioester bond to cysteine residues is marginally stable and of less physiological interest, I attempted to isolate E2-Ub conjugates with Ub linked via isopeptide bond to lysine residues of E2 (see **section 6.3.1** for more details). Enzymatic reactions were identical to the ones described in **section 3.2.1**, with two exceptions. The E2, Ubch5b containing a C-terminal 6xHis tag was used in 2 mL reactions at a concentration of 100 μM and 3 mM DTT was used in place of TCEP to ensure polyUb chains did not accumulate on cysteine residues. The Ub monomer was kept at a concentration of 1 mM. If mono

ubiquitination was desired Ub(K0), a mutant with all lysine residues mutated to arginine was used. For a specific linkage the corresponding single lysine mutant was used e.g. Ub(K63) to make only K63 linkages. Following the reaction 8 mL of His-Trap buffer A (20 mM sodium phosphate, 0.5 M NaCl, pH 7.4) was added and this mixture was slowly injected by with a 10 mL syringe onto a 1 mL His-Trap column. Then an additional 5 mL of buffer A was injected to wash unbound proteins form the column. Next, 8 mL of buffer B (20 mM sodium phosphate, 0.5 M imidazole, 0.5 M NaCl, pH 7.4) was used to elute the 6xHis tagged Ubch5b and ubiquitinated conjugates. The 6xHis tag on Ubch5b does not contain a cleavage site, however the 6xHis tag on human E1 does. Given the molecular weight difference between the E1 (~110 kDa) and Ubch5b conjugates (less than 50 kDa) I found it unnecessary to remove the 6xHis tag from E1 since this issue could easily be resolved with a size exclusion step. Once the Ubch5b conjugates were eluted from the 1 mL His-Trap column they were exchanged into PBS, pH 7.4 and DTT was added to a concentration of 50 mM. The DTT mixture was incubated at 30°C for 2 hours and exchanged back into PBS pH7.4 as it was concentrated in a 10,000 kDa MWCO centrifugal unit for size exclusion. Unligated Ubch5b was separated from other forms modified with varying amounts of Ub on a superdex 75 120 mL size exclusion column with a flow rate of 0.3 mL/min in PBS pH 7.4 buffer. Fractions were checked using 15% SDS-PAGE gels and desired fractions were concentrated and stored at -20°C for later use.

3.3 Standard biochemical assays

3.3.1 DUB digests and analysis

Digests of polyUb conjugates with DUBs was carried out in PBS pH7.4 buffer. After several controls studies, I found it unnecessary to use a reducing agent for any DUB, even the cysteine proteases as their activity is largely dependent on pH (*185*). The digest reactions were kept simple only containing the DUB, polyUb conjugate, and PBS pH 7.4 buffer in a total volume of 50 µL incubated in a 30^{o}C water bath. The concentration of the polyUb to be digested was kept at 25-50 µM and the concentration of the DUBs were as follows: OTUB1 500 nm, AMSH, 1 µM, Ubp6 5-10 µM, Rpn11 10 µM. Note for some assays, these concentrations changed, but this was found to be an ideal working concentration for each DUB. 4µL samples for each time point were taken over the digest a stored in 4µL of 4x SDS sample buffer at -20^{o}C. The samples were run on 15% SDS-PAGE gels, stained with Coomassie brilliant blue, photographed, and also scanned on densitometer. Gel bands from images were analyzed with ImageJ software and Adobe Photoshop, then plotted against time (*186*). Sections of the actual gels containing the polyUb bands are typically displayed next to these plots to provide a different perspective.

3.3.2 F45W tryptophan emission titrations

The buffer system for these fluorescence experiments was modular and could accommodate a variety of different buffers. To allow for direct comparison to NMR measurements, the NMR buffer (20 mM sodium phosphate pH 6.8) was used. A concentration of 50 µM F45W Ub was selected since it was above the reported K_d for many ligands tested and it also allowed for a robust emission signal from

tryptophan. Spectra were acquired by excitation at 280 nm with 5 nm slit widths, over an emission range spanning 300 to 600 nm in 0.5 nm increments. For each point in titration three spectra were collected and averaged. The averaged spectra were used for analysis and the emission maximum was used for fitting.

3.3.3 F45W-AEDANS FRET assays

For K48-Ub_2 chains used for FRET assays consisted of F45W on a proximal Ub serving as a FRET donor for 5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid (IAEDANS) located on T12C of a distal Ub, this construct is formally named Ub($T12C^{AEDANS}$) –^{48}Ub(F45W). Note since the iodine leaves upon attachment, IAEDANS is commonly referred to as AEDANS throughout. To create the di-Ub, Ub-74/F45W, or Ub-D77/F45W was enzymatically reacted with Ub-T12C. The K48 dimer was purified as described above and exchanged into the desired phosphate buffer with TCEP kept in a fivefold molar excess of T12C. Labeling with IAEDANS was performed on 200 μM of K48-Ub_2 in a 30^{o}C water bath for 3 hours in the presence of 1.2 mM IAEDANS and 600 μM TCEP. Excess IAEDANS was buffer exchanged out and the K48-Ub_2 was run through the size exclusion column to ensure no residual IAEDANS was present. All ANS steps were performed in the dark or properly wrapped in foil to preserve the lifetime of the fluorophore. Once created Ub($T12C^{AEDANS}$) –^{48}Ub(F45W) was tested for a characteristic AEDANS FRET emission signal at 485 nm using an excitation at 285 nm. After the expected FRET signals were observed, 50 μM Ub($T12C^{AEDANS}$) –^{48}Ub(F45W) was titrated with UBA(2) from hHR23A. The spectra were acquired as described above, with excitation at 285 nm and monitoring every 0.5 nm in the emission spectra range 300-

700 nm. The average of three spectra for each point was taken for analysis and the emission maxima for the tryptophan and AEDANS were used.

3.4 Solution NMR studies

3.4.1 Sample preparation for solution NMR

Isotope enriched protein samples for NMR were transferred to NMR buffer (20 mM sodium phosphate, pH 6.8) for traditional solution NMR studies, PBS pH 7.4 for select DUB assays, or acidic buffer (20 mM sodium acetate, pH 4.5) for low pH studies. All samples contained 5% D_2O for a reliable lock signal and for titration experiments in which the sample was diluted from the addition of ligand, extra D_2O was added to compensate for this effect. All samples were measured in standard 5 mm NMR tubes with sample volumes of 450 µL or 5 mm Shigemi tubes with sample volumes of 250 µL. With the exception of the experiments for DUB kinetics carried out at 303.0 K, all other experiments were acquired at 298 K.

3.4.2 Chemical shift perturbation (CSP) mapping and titration analysis

Differences between signals in ^{1}H-^{15}N NMR spectra for two species (A and B) were quantified as chemical shift perturbations, defined as follows:

$$CSP = [(\delta_{HA} - \delta_{HB})^2 + ((\delta_{NA} - \delta_{NB})/5)^2]^{1/2} \quad (1)$$

where δ_H and δ_N are chemical shifts of ^{1}H and ^{15}N, respectively, for a given backbone N-H group. The scaling factor of 5 for ^{15}N of is commonly accepted and adopted

from (*187*). An in-house MATLAB based software package PICK was used to ensure the maximum intensity for each 2D peak was properly picked. For titration analysis the same equation was used to quantify spectral perturbations upon titration; in this case, A refers to the unbound species, and B corresponds to spectra from subsequent points in the titration. Titrations were performed until there were no detectable shifts in $^{1}H,^{15}N$-HSQC. For cases of tight (sub μM K_d) binding i.e. slow off rates that resulted in signal attenuation, the effect was noted and residues exhibiting this behavior had to be excluded from CSP analysis due to signal loss.

To determine the K_d for 1:1 binding of a ligand to a protein (the ^{15}N enriched species) Eq. 2 was used to fit residue specific CSPs.

$$p_B = \left(\frac{[P_T]+[L_T]+K_D - \sqrt{(([P_T]+[L_T]+K_D)^2 - 4[P_B][L_T])}}{2[P_T]}\right) \quad (2)$$

Where p_B is the population of bound protein, $[P_T]$ represents the total concentration of protein, $[L_T]$ is the total concentration of ligand, and K_D is the variable to be fit. The CSP titration data were fit using an in-house MATLAB software package, KDFIT. The trajectories of each peak were also plotted using the same software package to ensure they were linear. To describe a protein with two equivalent binding sites and a ligand that can only occupy one at a time, Eq. 3 was used.

$$p_B = \left(\frac{[P_T]+[L_T]+\frac{1}{2}K_D - \sqrt{\left(\left([P_T]+[L_T]+\frac{1}{2}K_D\right)^2 - 4[P_B][L_T]\right)}}{4[P_T]}\right) \quad (3)$$

In the case where a protein presents two independent, but equivalent binding sites Eq. 4 was used to fit the data.

$$p_B = \left(\frac{2[P_T]+[L_T]+K_D - \sqrt{\left(\left(2[P_T]+[L_T]+\frac{1}{2}K_D\right)^2 - 8[P_B][L_T]\right)}}{4[P_T]} \right) \quad (4)$$

After determining the residue specific K_d values, the overall K_d reported was obtained by averaging residue specific K_d values that showed an excellent agreement to the binding model. Eq.2-4 are taken from (*153*); these equations are incorporated into MATLAB program KDFIT (Prof. Fushman).

3.4.3 ^{15}N relaxation rates

^{15}N longitudinal relaxation rates (T_1) were used to measure the overall size of polyUb chains and Ub/ligand complexes. Experiments were run in the pseudo 3D format as a series of 2D ^{1}H,^{15}N-HSQC planes with only two delays (short and long) repeated in multiple 2D planes for each delay. Relaxation data was analyzed with in-house software, ROTDIF. The development of this software is described in (*188*). Following peak picking and integration of noise, the T_1 for each residue was fit using the RELAXFIT module set for a single exponential decay, two parameter fit. The RELAXFIT module was also used to determine the error in each T_1 measurement, which was mainly dictated by the noise in each spectrum, see (*188*).

3.4.4 Translational diffusion measurements

Diffusion ordered spectroscopy (DOSY) experiments were acquired in a pseudo 2D format as a series of 1D increments that differed by the strength of the gradient applied. To distinguish between signals from ^{15}N labeled Ub and those of the unlabeled ligands, a ^{15}N inept filter was used to suppress Ub signals in the amide region (7-10 ppm). The data for each point was fit using the in-house TRANSDIF module.

3.4.5 Site specific paramagnetic spin labeling and analysis

Ub cysteine mutants were labeled with MTSL (S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate) at several different positions (T12C, I36C, K48C, and K63C). Purified Ub cysteine mutants were exchanged into NMR buffer (20 mM phosphate pH 6.8) containing no reducing agent. A threefold molar excess of MTSL was added and allowed to incubate with Ub for 12 hours at ambient temperature. The excess MTSL was removed by buffer exchanging three times into NMR buffer. MTSL Oxidized NMR spectra were recorded as $^{1}H,^{15}N$-HSQC and 1D ^{1}H-NOESY with ^{15}N filtering. After allowing the NMR sample to incubate with a threefold molar excess of sodium ascorbate (added form a stock solution, 140 mM, pH 7.0) for one hour, the reduced spectra were acquired using settings identical to the oxidized spectra. All NMR experiments were run with a high number of scans to provide for more reliable downstream analysis. The exact position of the paramagnetic center in MTSL was fit from the intensities between the oxidized and reduced $^{1}H,^{15}N$-HSQC spectra using SLFIT, an in-house MATLAB package. PDB-1D3Z was taken as the reference structure for Ub. Once the position of the

paramagnetic center was determined, the distance to individual protons on the ubistatin observable 1D ^{1}H-NOESY spectra was calculated. The peak intensities and input parameters for SLFIT were generated in Bruker Topspin. An unpaired election in the nitroxide group of MTSL represents the paramagnetic center. When NMR active nuclei are in close proximity to this center, there is a distance dependent effect on their transverse (T_2) relaxation rates, which results in a decrease of signal intensity. This paramagnetic relaxation enhancement (PRE) effect for a spin S is described by Eq. 5.

$$\Delta R_{2para} = K\left[4\tau_c + 3\,\frac{\tau_c}{(1+\omega_H{}^2\tau_c{}^2)}\right] / r^6 \qquad (5)$$

Where ΔR_{2para} is the effect on T_2 relaxation of a ^{1}H nucleus, r is the distance to the paramagnetic center, τ_c is the T_2 rotational correlation time, and K= $\frac{1}{15}S(S+1)\gamma_H{}^2\beta_e{}^2g_e{}^2$ (γ is the gyromagnetic ratio of ^{1}H, β_e is the Bohr magneton, and g_e is the electron g-factor). Between the oxidized and reduced states of MTSL Eq. 6 applies, and Eq. 5 allows the distance between the paramagnetic center and each nucleus to be determined using the SLFIT software package. For practical reasons we used signal intensities as opposed to volumes.

$$\frac{I_{ox}}{I_{red}} = \left(e^{-\Delta R_{2\,para} * t_H}\right) X \frac{R_{2\,red}}{R_{2\,red} - R_{2\,ox}} \qquad (6)$$

In Eq. 6 I_{ox} and I_{red} are the intensities of either 1D or 2D peaks in the oxidized or reduced form of MTSL. Note that oxidized is the paramagnetic state of MTSL, while the reduced state is diamagnetic: in this state MTSL does not have any paramagnetic effect on signal intensity. t_H represents the time that magnetization is on ^1H during the INEPT elements of the NMR experiment.

3.4.6 Detection of intermolecular NOEs in protein/small molecule complexes

Intermolecular NOEs between Ub and ubistatins were detected using double labeled $^{13}C/^{15}N$-Ub_1 in a series of 3D TOCSY and NOESY experiments. Assignment of Ub and the ubistatin in the bound state was critical. The $^1H,^{15}N$-HSQC spectra served as a starting point for assignment of residues in Ub. 3D ^1H-CCCONH TOCSY and 2D $^{13}C,^1H$-HSQC were used to assign the chemical shifts for residues in the bound state of Ub. Intermolecular NOEs between Ub and ubistatins were detected in 2D NOESY experiments with ^{15}N filtering and also with a 3D NOESY (Bruker hsqcgpnowgx33d) with ^1H and ^{13}C INEPT transfer that shows NOEs between ^1H's in the ubistatin and both ^1H's and ^{13}C's in Ub (*189*). Using the chemical shifts for the bound state, NOEs were assigned between individual ^1H's in the ubistatin and $^1H/^{13}C$ groups in Ub. Intensities for these NOEs were integrated from the 2D NOESY-inept spectrum and converted into distance using Eq. 7. For calibration purposes, known distances from intramolecular NOEs were used for determining the constant A in Eq. 7. If an NOE was between a methyl group which contained three indistinguishable protons, this was accounted for in the calculation.

$$volume = \left(\frac{A}{r^6}\right) \quad (7)$$

3.4.7 DUB kinetics by $^{1}H,^{15}N$-HSQC

The kinetics of cleavage in ^{15}N-Ub conjugates was determined using sequential series of identical sofast $^{1}H,^{15}N$-HSQC experiments covering a broad time range. Each experiment was run as a normal $^{1}H,^{15}N$- HSQC with 128 ^{15}N points and 16 scans. Time was carefully monitored manually and critical points such as when the DUB was added, when the first experiment was started, and when the last experiment finished were recorded. The intensity corresponding to the free G76 signal was picked in each experiment and any other peaks relevant to the linkage such as the peak corresponding to the ligated G76 were also used if available. Given the amount of resources needed for such experiments, it was difficult to assess multiple samples under varying amounts of DUB or polyUb concentrations. This led me to fit the apparent rates of cleavage as opposed traditional steady state parameters. Using a modified version of RELAXFIT, the intensity of free G76 was fit to the single exponential decay model to determine the apparent rate of cleavage.

3.5 Modeling and structural calculation with HADDOCK v2.1

Out of the many commercial and open source biomolecular docking programs available Prof. A.M.J.J. Bonvin's (Utrecht University) HADDOCK (High Ambiguity Driven biomolecular DOCKing) was chosen (*190*). Unlike other docking programs HADDOCK allows accurate docking with CSPs (see **section 4.3**), as well as other ambiguous restraints. The software has also been constantly developed for over a decade and used for solving many complexes of biomolecules. In fact, K48-Ub_2 was used as a test system in the past for design of the optimal strategy for incorporation of

CSPs together with residual dipolar couplings in HADDOCK (*191*). In addition HADDOCK has also gained high scores in CAPRI (*192*) and its incorporation of CNS (*193*) is fine tuned for proteins and nucleic acids. With the proper use of custom CNS files HADDOCK is also ideal for small molecule-biomolecule docking (*194*).

3.5.1 Modeling of complex polyUb chains

Models for theoretical complex (mixed linkage) triUbs were generated on the HADDOCK web server (*195*). All Ub units were from PDB-1D3Z and the unambiguous distance restraints creating the isopeptide bond between the Ubs were adopted from Prof. Walker's study (*173*). The coefficients for the distance energy penalty were E_{dist}=0.02, E_{dist}=0.4, E_{dist}=0.7 for rigid body energy minimization, simulated annealing, and solvent refinement in water, respectively. Residues 72-76 were defined as fully flexible for all Ubs. As an ambiguous restraint, the L8, I44,V70 hydrophobic patch residues were defined as active which allowed each Ub to explore many potential interactions with the other two Ubs. For each theoretical tri-Ub chain 3,000 structures were generated for rigid body docking, 300 structures were taken for simulated annealing, and 300 structures were take for refinement and analysis in water. The CNS topology and parameter files along with the other energy settings were optimized for proteins by default and left unchanged. Cluster analysis was performed by the HADDOCK web server with a RMSD cut off of 7.5 Å and minimum cluster size of 10 structures. Clusters were sorted by HADDOCK score which is a summation of many energy parameters. The clusters were then analyzed in PyMol to define distinct conformations (see **section 4.3.2**).

3.5.2 Ubistatin/Ub structure calculation

PDB-1D3Z was used to model the c112/Ub complex. The HADDOCK web server only has CNS topology and parameter files available for proteins and nucleic acids. In order to run HADDOCK with small molecule ligands (i.e. ubistatins), custom CNS topology and parameter files were created for each molecule using the Dundee PRODRG2 server (*196*). Since the web server could not accept these custom files, calculations were run on a local 64 node Intel based cluster. To guide the docking residues I44, V70, and H68 were defined as active for the AIR restraints. The distances determined from intermolecular NOEs and paramagnetic relaxation enhancement with MTSL were used as unambiguous restraints. Docking was run with the custom CNS files for the ubistatin, the temperature for the early rigid body docking steps was increased to 3000 K, and the number of MD steps was reduced to 0. 3,000 structures were generated during rigid body docking, 300 structures were taken for simulated annealing, and 300 structures were take for refinement and analysis in water. Clusters were very tight and showed little deviation. The output was checked extensively for any distance violations. To ensure our structure conformed to the real experimental distances the penalty for the unambiguous distance restraint (E_{dist}) was increased for each step. Many runs were performed varying the E_{dist} penalties and also excluding certain distances. After numerous runs a consensus was reached and a single cluster was chosen as the structure of the c112/Ub_1 complex. A similar approach lacking extensive experimental data was used to model other ubistatin/Ub complexes.

Chapter 4: Laying the foundation for structural studies of complex polyubiquitin chains

4.1 Background and research aims

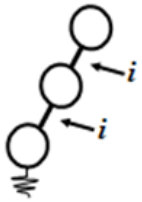
Since the discovery of Ub it was typically assumed that once formed a polyUb chain would retain a high degree of linkage homogeneity (*13*). However, more recent studies found numerous examples where the polyUb chains exhibit linkage heterogeneity to accomplish specific tasks in the cell (*164*). In addition, studies on isolated E2/E3 enzymes and model substrates have also demonstrated that heterogeneous polyUb chains are routinely formed (*20, 26, 28*). Although intriguing it is almost impossible to gauge the extent that heterogeneous polyUb chains are present in the cell since they are virtually undetectable with current methods and the polyUb chain itself is constantly being remodeled (*144*). Nevertheless, when heterogeneous polyUb chains are formed they could potentially provide for novel signaling outcomes based on the unique three-dimensional conformations they could present for receptors.

In 2009 we introduced the term "complex polyUb chains" to describe any polyUb chain with linkage heterogeneity. To limit confusion it is important to understand that heterogeneous, mixed linkage, and complex are all synonyms and used interchangeably to refer to polyUb chains containing more than one type of Ub-Ub linkage. To clearly communicate my findings a formal nomenclature system with standardized terms had to be devised in order to report on this new frontier. Our original nomenclature system was adequate for specific cases, however after careful evaluation we ultimately would settle on the Cohen-Nakasone-Fushman convention

for complex chains, which preserved the same principals from our original system, but employed different symbols. The objective of this study is rather straightforward: determine if linkage branching or mixing in the same polyUb chain results in new conformations and signaling properties. However, the different polyUb chains that must be explored are nearly infinite. Often it is assumed that there are just eight polyUb signals resulting from the different linkages (K6, K11, K27, K29, K33, K48, K63, and M1). This assumption has led the field to study individual homogeneous linkage types without considering complex chains. To illustrate exactly what the field is missing by myopically studying homogeneous chains and discounting complex chains, allow for just three Ubs and only two Ub-Ub linkages. When complex chains are allowed there are now 92 unique polyUb signals of which only 8 are homogeneous, and exponentially more depending on the number of Ubs introduced (**Figure 4.1**). In fact, my initial plunge into this concept identified 23,190,029,720 complex polyUb chains with just 10 Ubs and the possibility for many more depending on the number of linkages and Ubs allowed. I stopped naming complex polyUb chains after sixteen Ubs since that number is generally accepted as the most Ubs found in a polyUb chain from living cells and naming 8.26×10^{17} chains is computationally restrictive (*197*). Although not clearly stated in the literature, when complex polyUb chains are accounted for, Ub becomes one of the most sophisticated naturally occurring biopolymers. In this chapter I will use molecular modeling to explore the possibility that linkage branching or mixing in the same polyUb chain will create new signals. Then I present my findings on the structural outcomes of K48 and K63 linkage mixing.

A. Homogeneous-linkage tri-Ub chains

	K6	K11	K27	K29	K33	K48	K63	M1	
K6	1								
K11		1							
K27			1						
K29				1					
K33					1				
K48						1			
K63							1		
M1								1	total
#str	1	1	1	1	1	1	1	1	8



B. Mixed-linkage: branched tri-Ub chains

	K6	K11	K27	K29	K33	K48	K63	M1	
K6		A	B	C	G	K	P	V	
K11	A		C	E	H	L	Q	W	
K27	B	C		F	I	M	R	X	
K29	D	E	F		J	N	S	Y	
K33	G	H	I	J		O	T	Z	
K48	K	L	M	N	O		U	Ä	
K63	P	Q	R	S	T	U		Ë	
M1	V	W	X	Y	Z	Ä	Ë		total
#str	7	6	5	4	3	2	1		28

C. Mixed-linkage: unbranched tri-Ub chains

	K6	K11	K27	K29	K33	K48	K63	M1	
K6		pA	pB	pC	pD	pE	pF	pG	
K11	dA		pH	pI	pJ	pK	pL	pM	
K27	dB	dC		pN	pO	pP	pQ	pR	
K29	dD	dE	dF		pS	pT	pU	pV	
K33	dG	dH	dI	dJ		pW	pX	pY	
K48	dK	dL	dM	dN	dO		pZ	pÄ	
K63	dP	dQ	dR	dS	dT	dU		pË	
M1	dV	dW	dX	dY	dZ	dÄ	dË		total
#str	7	7	7	7	7	7	7	7	56

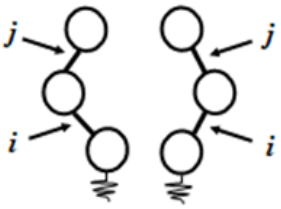


Figure 4.1 – All possible chains from three Ubs and two linkages.
If we account for all possible linkages (7 lysines and M1) between three Ubs we get 92 distinct tri-Ub chains, of which only 8 are homogeneous-linkage chains **(A)** and the rest are mixed-linkage chains. Of the 84 mixed-linkage tri-Ub chains, 28 are branched **(B)** and 56 are unbranched chains **(C).** In the schematic representations on the right, ***i*** and ***j*** indicate the two linkages in tri-Ub: $i = j$ for homogeneous-linkage chains, and $i \neq j$ for mixed-linkage chains. The order of the linkages (***i*** and ***j***) is important for the unbranched mixed-linkage chains **(C)**, where $i \,\&\, j \neq j \,\&\, i$. If we continue adding Ubs we find that with eight possible linkage sites per Ub, the total number of possible distinct chains is 1,240 for tetra-Ub, 18,276 for penta-Ub, 285,384 for hexa-Ub, … , 2.3×10^{10} for deca-Ub and ever increasing.

4.2 A historical perspective: Original Nomenclature

4.2.1 Describing the topology of the chain: Linear, mixed, branched

By default homogeneously linked polyUb chains have a linear topology where each Ub is linked to another Ub through the same bond with no mixing or branching. Mixed linkage chains also have linear topologies provided that each Ub in the chain has one and only one linkage site (K6, K11, K27, K29, K33, K48, K63, or M1) per individual Ub in the chain. Branched chains arise from more than one linkage site on any individual Ub in the chain. The term "forked" was introduced by Alfred Goldberg to describe branched chains where the Ub-Ub linkages occur at adjacent positions on a Ub e.g. K6&K11, K27&K29 or K29&K33 (*20*), however these forked chains still fall under the definition of branched. In practice mixed- and branched-linkage chains are different sub-groups of complex chains and great care should be taken to report as specifically as possible to limit confusion. For shorthand I designate linear-mixed linkages chains with (L) and branched chains with (B).

4.2.2 Pinpointing an individual Ub in a complex chain

Communicating the interactions of individual Ubs in a given polyUb chain is critical for understanding cellular processes. For example explaining a polyUb/protein complex or how polyUb is cleaved by a DUB requires that the exact position of an individual Ub in a chain is clearly identifiable. Traditionally chains studied have been very simple, often dimers, and an individual Ub could be referred to as "distal" or proximal" in the case of a dimer. These terms were popularized by the late Prof. C.M Pickart to describe the Ub unit closest to the substrate (proximal) and the furthest Ub (distal). For di-Ubs distal and proximal can be used

unambiguously to referrer to the specific Ub, however for longer chains we must introduce other terms. If the chain is of linear topology, regardless if it is homogeneous or mixed-linkage, the proximal Ub is still the proximal Ub, the terminating Ub remains the distal Ub, and any Ubs between connecting the two are sequentially referred to as "endo-X" where X indicates the Ub counting from the proximal Ub. For example, say we have two hexameric Ub chains (Ub_6), one all K63-linked and another with five different linkages (K6, K11, K27, K29, and K33 in order from the proximal Ub) (**Figure 4.2)**. Given that the two chains are linear we can describe the domains as follows; proximal, endo-1, endo-2, endo-3, endo-4, and distal. The chains would be named K63-Ub_6 and (L)-[K6-K11-K27-K29-K33]-Ub_6 respectively. If we reverse the order of linkages in (L)-[K6-K11-K27-K29-K33]-Ub_6 the name changes since the chain is named according to the order of linkages starting from the proximal end **(Figure 4.2)**. In this original nomenclature system all chains have the format "(A)-[B1-B2…-Bn]-Ub_x." The first part "(A)" refers to any special topology the chain may have, if it is purely homogeneous this entry is left blank and if it is a linear-mixed linkage or branched the section will have "(L)" or "(B)," respectively. The next portion of the name "[B1-B2…-Bn] indicates the linkages present in the chain written as they appear from the proximal Ub. If a branched chain is being described the linkages are written with the lowest residue number first e.g. [M1K11K48]-Ub_4 not [K48M1K11]-Ub_4. The last part of the chain name "Ub_x" refers to the total number of Ubs in a given chain.

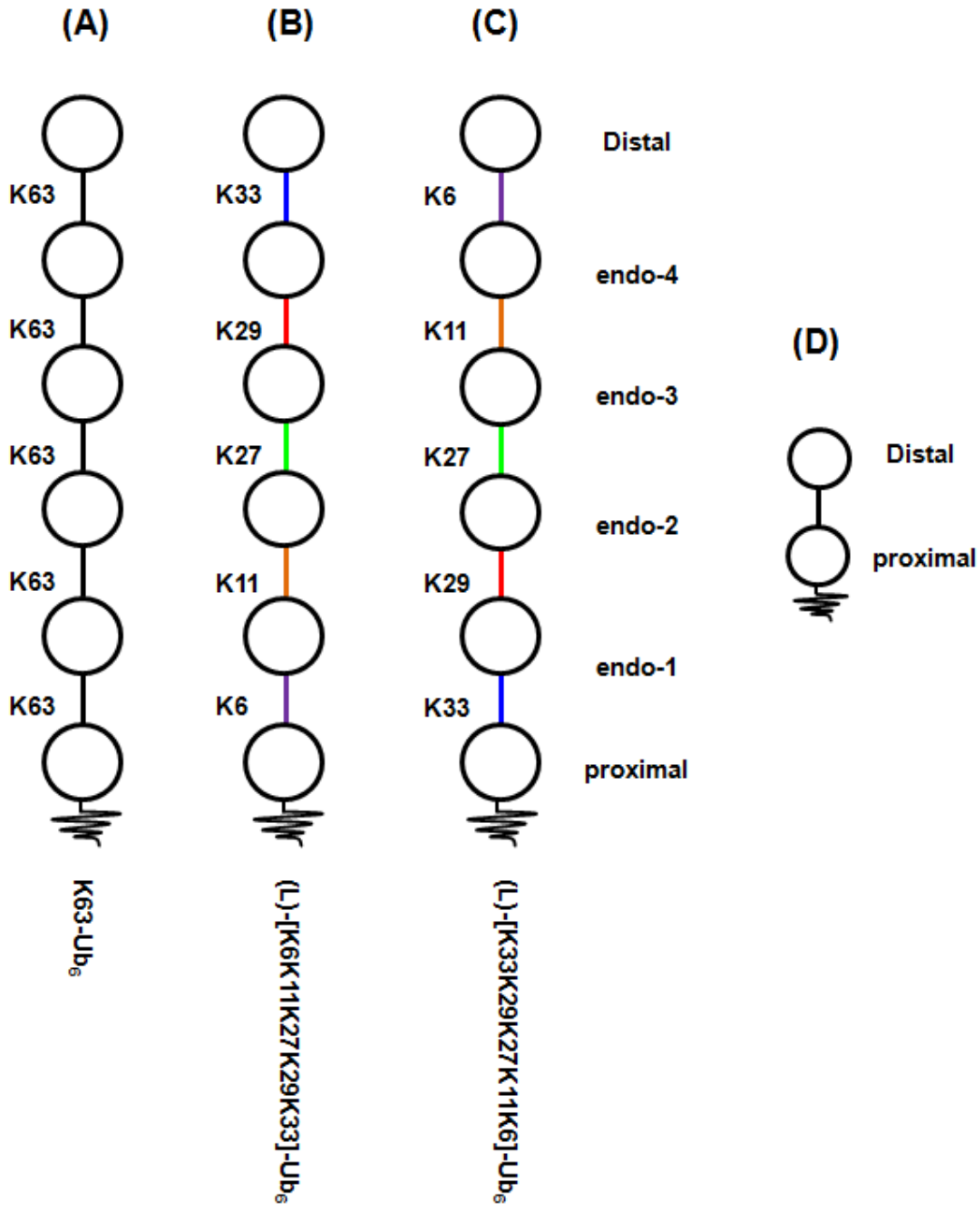


Figure 4.2 – Nomenclature in practice for hexa-polyUb chains.
(A) homogeneous Ub_6 linked exclusively through K63 is simply termed K63-Ub_6, while the other linear-mixed linkage chains are named according to the order of their linkages from the proximal Ub **(B)** (L)-[K6K11K27K29K33]-Ub_6 and (L)-[K33K29K27K11K6]-Ub_6 **(C)**. The proximal Ub initiates the chain, while the distal Ub terminates the chain and anything in between is termed "endo-X", where X increases from the proximal Ub. **(D)** Ub dimers can only contain distal and proximal domains.

4.2.3 Dimensions of complexity

The signaling possibility of complex polyUb chains is mainly due to the ability of one Ub to contain multiple linkages. For homogenous chains each Ub can create a linkage with its C-terminus and through one of eight linkage sites, while the simple complex polyUb chains (e.g. (B)-[K48K63]-Ub_3) use only two of the eight sites. The real intricacy in the polyUb signal arises when complex chains contain multiple linkages on a single Ub. Experientially, I have shown that a single Ub can be ligated with five other Ubs using different combinations of linkage specific E2s. My results from *in silico* modeling clearly show that eight Ubs can be ligated to a single Ub without any steric conflicts and there is even room for the C-terminus to be ligated to a ninth Ub (**Figure 4.3**).

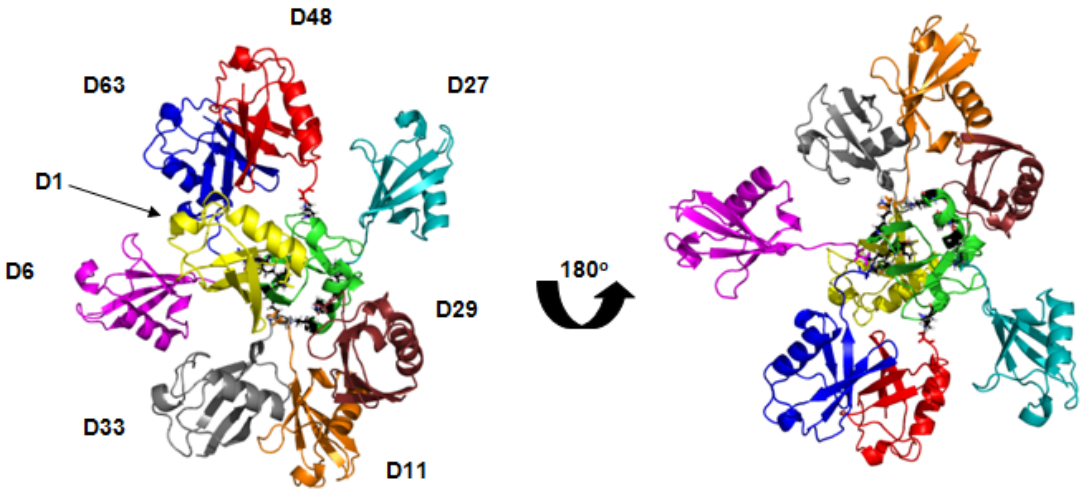


Figure 4.3 – Molecular modeling of an 8D branched Ub_9.
The HADDOCK v2.1 generated model suggests that eight Ubs can easily be attached to all eight linkage sites on a single Ub (green). The distal "D" Ubs are colored as followed; M1 Ub D1 (yellow), K6 Ub D6 (magenta), K11 Ub D11 (orange), K27 Ub D27 (cyan), K29 Ub D29 (rust), K33 Ub D33 (gray), K48 Ub D48 (red), and K63 Ub D63 (blue). The chain represented, (B)-[M1K6K11K27K29K33K48K63]-Ub_9 has eight dimensions of complexity.

Given that the attachment of a single Ub can potentially serve as an anchor for any other linkage we refer to this as adding a dimension such that every Ub added to a given Ub accounts for a single dimension. For example homogeneous chains are all 1D as well as linear-mixed linkage chains, while a branched chain with two Ubs attached to K48 and K63 on single Ub is a 2D chain and eight Ubs attached to a single Ub in a chain results in an 8D chain. Not every Ub in the chain may be branched equally, some will contain more linkages while others have less. To determine the dimension of complexity for a chain we simply take the highest dimension for any part of the chain. For example, if the 8D (B)-[M1K6K11K27K29K33K48K63]-Ub_9 chain had one of its distal Ubs modified with a K11+K6 branch, one part of the chain would be 8D while another part is only 2D, however the chain as a whole would be 8D since that represents the most dimensions of complexity.

4.3 The standardized Cohen-Nakasone-Fushman nomenclature system for complex polyUb chains

4.3.1 The new standard in poly Ub chain nomenclature

After extensive consideration we revised the nomenclature system in the previous **section 4.2** to be more intuitive and easier to conceptualize. The (B) and (L) designations are abandoned and we have switched to a schematic system, which allows readers to visualize the Ub chains from just their printed names. This system also can accurately describe heterotypic (also called heterologous) chains that contain both Ub and UBL molecules. We formally introduced this system in 2013 to the

scientific community as the Cohen-Nakasone-Fushman nomenclature system for complex polyUb chains.

4.3.2 Guidelines

The common abbreviations, Ub and UBLs (e.g. SUMO2, Nedd8) designate that specific protein. Point mutations or other special modifications on these proteins can be noted within parenthesis (e.g. the K63R mutant of Ub is written as Ub(K63R), Ub(K0) for all lysine residues mutated to arginine, if a particular Ub is ^{15}N-labeled enriched Ub(^{15}N)). This can be expanded to Ub or UBL modified substrates and their modifications can be noted in the chain also using their common abbreviation. In the formal name the distal Ub unit within a polyUb chain is written to the left, whereas the proximal Ub (or the substrate if present) is to the right. Accordingly, internal "endo" Ub units are listed from left to right transitioning from the distal to proximal direction. The Ub-Ub, Ub-UBL, or UBL-UBL linkage is represented with an *en dash* (–). Specific linkage sites of the distal component at each linkage are indicated as *superscripts*. If there is a traditional isopeptide linkage to a lysine ε-amine, a residue number is sufficient, however this notation can include any amino acid type or side chain atom.

For two simple dimers, one containing a single K63 linkage is written as Ub–^{63}Ub, another dimer linked “head-to-tail” through M1 would be Ub–^{M1}Ub, and the substrate p53 with mono Ub attached to lysine 101 is represented as Ub–^{101}p53. Uncertainties in linkage length and multiple sites of modification can be accommodated as well. If the p53 now becomes modified with K48 linked polyUb at two sites K101 and K124, but the chain length is uncertain it would be named Ub(–

$^{48}Ub)_n$–$^{101/124}$p53 to indicate K48 linked polyUb of undetermined length (n) attached to K101 orK124 of p53. If the site of modification is unknown, the superscripts can be omitted from the substrate. The notation for polyUb (also polyUBL) chains comprised of identical monomers and one linkage type can employ parentheses to indicate repeated units. Thus, the K63-Ub_6 example from the previous section is $(Ub)_5$–^{63}Ub and a mixture of polymers $3 \leq n \leq 10$ of Nedd8 is described as Nedd8(–^{48}Nedd8)$_{1–8}$–^{48}Nedd8.

Mixed linkage (complex) chains containing more than one type of Ub-Ub or UBL-UBL linkage, are specified with the appropriate superscripts to indicate linkage. For a linear mixed-linkage chain with a distal Ub linked to K11 of a middle (endo-1) Ub that in turn is linked to K27 of a proximal Ub the proper name would be, Ub–^{11}Ub–^{27}Ub. Heterotypic Ub-UBL chains can be written following the pattern above, a K48 linked di-Ub attached to K11 of SUMO2 would be Ub–^{48}Ub–^{11}SUMO2.

For branched chains containing more than one linkage per unit the point of the branch is highlighted by the use of brackets (brackets = branching). Ub[Ub]–^{6,48}Ub or $[Ub]_2$–^{6,48}Ub indicates two distal Ub units linked to K6 and K48 of a proximal Ub which now has two superscript indicating linkage sites. Different extensions of branches are assigned according to the order that they are written e.g., [Ub–^{29}Ub][Ub]–^{29,63}Ub is tetra-Ub in which the proximal Ub of K29 linked trimer is also modified with mono Ub at K63. After implementing these nomenclature rules in a simple in-house program, the names for all 92 possible chains with three Ubs and two linkages are shown (**Table 4.1**).

Branched	
A	$[Ub]_2$–^{6,11}Ub
B	$[Ub]_2$–^{6,27}Ub
C	$[Ub]_2$–^{11,27}Ub
D	$[Ub]_2$–^{6,29}Ub
E	$[Ub]_2$–^{11,29}Ub
F	$[Ub]_2$–^{27,29}Ub
G	$[Ub]_2$–^{6,33}Ub
H	$[Ub]_2$–^{11,33}Ub
I	$[Ub]_2$–^{27,33}Ub
J	$[Ub]_2$–^{29,33}Ub
K	$[Ub]_2$–^{6,48}Ub
L	$[Ub]_2$–^{11,48}Ub
M	$[Ub]_2$–^{27,48}Ub
N	$[Ub]_2$–^{29,48}Ub
O	$[Ub]_2$–^{33,48}Ub
P	$[Ub]_2$–^{6,63}Ub
Q	$[Ub]_2$–^{11,63}Ub
R	$[Ub]_2$–^{27,63}Ub
S	$[Ub]_2$–^{29,63}Ub
T	$[Ub]_2$–^{33,63}Ub
U	$[Ub]_2$–^{48,63}Ub
V	$[Ub]_2$–^{6,1}Ub
W	$[Ub]_2$–^{11,1}Ub
X	$[Ub]_2$–^{27,1}Ub
Y	$[Ub]_2$–^{29,1}Ub
Z	$[Ub]_2$–^{33,1}Ub
Ä	$[Ub]_2$–^{48,1}Ub
Ê	$[Ub]_2$–^{63,1}Ub

Mixed-linkage unbranched	
pA	Ub–^{6}Ub–^{11}Ub
pB	Ub–^{6}Ub–^{27}Ub
pC	Ub–^{6}Ub–^{29}Ub
pD	Ub–^{6}Ub–^{33}Ub
pE	Ub–^{6}Ub–^{48}Ub
pF	Ub–^{6}Ub–^{63}Ub
pG	Ub–^{6}Ub–^{1}Ub
pH	Ub–^{11}Ub–^{27}Ub
pI	Ub–^{11}Ub–^{29}Ub
pJ	Ub–^{11}Ub–^{33}Ub
pK	Ub–^{11}Ub–^{48}Ub
pL	Ub–^{11}Ub–^{63}Ub
pM	Ub–^{11}Ub–^{1}Ub
pN	Ub–^{27}Ub–^{29}Ub
pO	Ub–^{27}Ub–^{33}Ub
pP	Ub–^{27}Ub–^{48}Ub
pQ	Ub–^{27}Ub–^{63}Ub
pR	Ub–^{27}Ub–^{1}Ub
pS	Ub–^{29}Ub–^{33}Ub
pT	Ub–^{29}Ub–^{48}Ub
pU	Ub–^{29}Ub–^{63}Ub
pV	Ub–^{1}Ub–^{29}Ub
pW	Ub–^{33}Ub–^{48}Ub
pX	Ub–^{33}Ub–^{63}Ub
pY	Ub–^{33}Ub–^{1}Ub
pZ	Ub–^{48}Ub–^{63}Ub
pÄ	Ub–^{48}Ub–^{1}Ub
pÊ	Ub–^{63}Ub–^{1}Ub

Mixed-linkage unbranched	
dA	Ub–^{11}Ub–^{6}Ub
dB	Ub–^{27}Ub–^{6}Ub
dC	Ub–^{27}Ub–^{11}Ub
dD	Ub–^{29}Ub–^{6}Ub
dE	Ub–^{29}Ub–^{11}Ub
dF	Ub–^{29}Ub–^{27}Ub
dG	Ub–^{33}Ub–^{6}Ub
dH	Ub–^{33}Ub–^{11}Ub
dI	Ub–^{33}Ub–^{27}Ub
dJ	Ub–^{33}Ub–^{29}Ub
dK	Ub–^{48}Ub–^{6}Ub
dL	Ub–^{48}Ub–^{11}Ub
dM	Ub–^{48}Ub–^{27}Ub
dN	Ub–^{48}Ub–^{29}Ub
dO	Ub–^{48}Ub–^{33}Ub
dP	Ub–^{63}Ub–^{6}Ub
dQ	Ub–^{63}Ub–^{11}Ub
dR	Ub–^{63}Ub–^{27}Ub
dS	Ub–^{63}Ub–^{29}Ub
dT	Ub–^{63}Ub–^{33}Ub
dU	Ub–^{63}Ub–^{48}Ub
dV	Ub–^{1}Ub–^{6}Ub
dW	Ub–^{1}Ub–^{11}Ub
dX	Ub–^{1}Ub–^{27}Ub
dY	Ub–^{1}Ub–^{29}Ub
dZ	Ub–^{1}Ub–^{33}Ub
dÄ	Ub–^{1}Ub–^{48}Ub
dÊ	Ub–^{1}Ub–^{63}Ub

Table 4.1 – Cohen-Nakasone-Fushman nomenclature system applied to mixed linkage Ub trimers.
The formal names for the 92 possible Ub trimers discussed in section 4.2 are written in full. Note the brackets for branched trimers and how the order of linkages going from left to right results in a unique chain, such that Ub–^{48}Ub–^{63}Ub ≠ Ub–^{63}Ub–^{48}Ub.

It is possible to cyclize a chain using the K48 linkage and this may also be possible with other UBLs. A cyclized chain is represented with an unattached *en dash* on the left and right, representing its circular conformation with no beginning or end. To indicate the number of the Ub units in the chain a subscript can be written outside of the parenthesis of the units with the Ub or the UBL name. $(-^{48}Ub-)_3$ and $(-^{48}Ub-^{48}Ub-^{48}Ub-)$ would both indicated a cyclized K48 linked Ub trimer. If the similar UBL, Need8 cyclized it would be $(-^{48}Nedd8-)_3$.

4.3 Theoretical structural outcomes of branched and unbranched mixed linkage chains

Only three of the ninety-two complex tri-Ub chains identified in (**Figure 4.1**) have been addressed experimentally, while the remaining have yet to be synthesized. The tedious work involved with assembling the chains, their sheer numbers, and also the time needed for data collection and analysis combine to make the study of complex chains prohibitive. However, given that we know which surfaces Ub uses to form interdomain contacts, several *in silico* approaches have been used to predict the possible three dimensional conformations of polyUb chains (*173*). Using the same *in silico* approach I analyzed the remaining complex trimeric Ub's to explore which structural ensembles were possible.

4.3.1 Validity polyUb models from *in silico* docking software

Regardless of whether a chain is complex or simply homogeneous, the structural ensembles a chain can adopt are greatly dependent on the ability of

individual Ubs in the chain to interact with other Ubs. This type of interaction is mainly achieved by hydrophobic contact between different Ubs in the chain via two solvent exposed hydrophobic patches, the well known L8, I44, V70 patch (for K48 linkages) or another created by L8, I36, L71, L73 that creates the interface between K6 linkages (**Figure 4.4)**. As show in (**Figure 4.4**) there are other important solvent exposed regions on Ub, however the L8, I44, I70 and L8, I36, L71, L73 hydrophobic patches are the only ones that have been observed to create interdomain interactions in polyUb chains, while the other surfaces are used for receptor recognition.

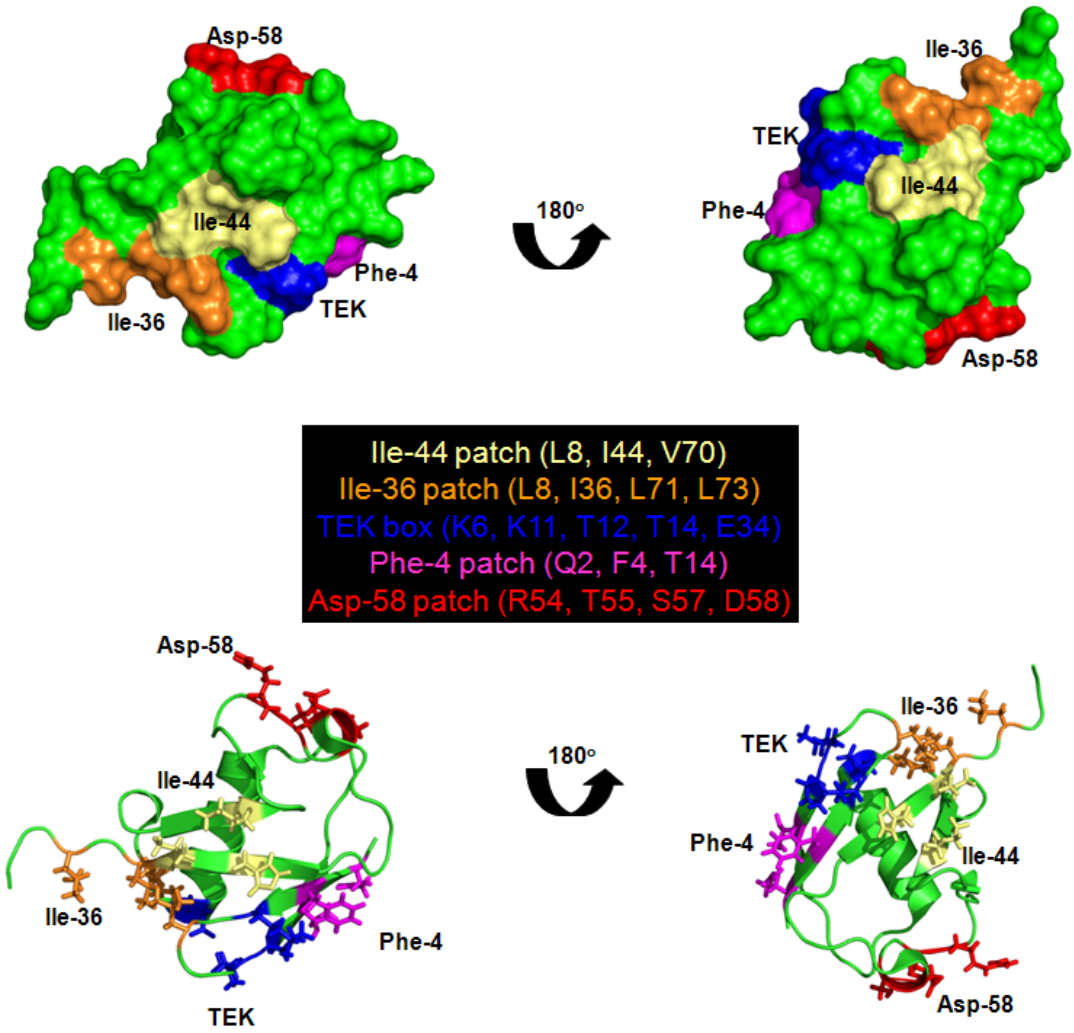


Figure 4.4 – Important surfaces on Ub.
Ub has several solvent exposed patches which are important for interactions with both UBDs and maintaining the proper conformation of polyUb chains through interdomain contacts. Surface representation (top) and cartoon representation (bottom).

Given the existing knowledge of how Ubs interact within the same chain, one can use molecular modeling to determine if certain interactions are possible. HADDOCK v2.1 is tailored for biomolecules and can properly maintain critical properties of proteins such as backbone torsion angles (ω, Φ, and ψ) while allowing for multiple domains to sample different conformations, which are governed by user-

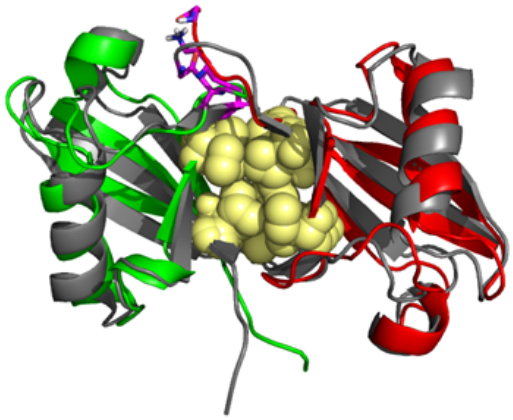
defined parameters (*173*). In the case of two or more Ubs, the resulting structures will reflect a majority of the possible conformations for a given chain. Using just the L8, I44,V70 hydrophobic patch with no other restraints, HADDOCK can accurately predict the structure of K48-Ub_2 with a Cα RMSD=1.72Å (**Figure 4.5)**. With this approach I determined which structural conformations were possible in complex polyUb chains and formulate several new theories on the signaling properties of these unique chains.

Restraints used in Haddock simulations

Distal Ub		
	Active residues	L8, I44, V70
	Passive residues	G10, L71
	Fully felexible segments	R72-G76
Proximal Ub		
	Active residues	L8, I44, V70
	Passive residues	G10, L71
	Fully felexible segments	R72-G76

Isopeptide bond (G76 of distal to εNH2 of lysine on proximal Ub)

Atoms (pdb name)	Unambiguos restraint (Å)
O-NZ	2.25±0.05
C-NZ	1.35±0.05
C-CE	2.45±0.05
CA-NZ	2.45±0.05



PDB-1AAR

Figure 4.5 – HADDOCK accurately predicts the K48 interface from minimal input.
The HADDOCK generated K48-Ub_2 (proximal=green and distal=red) results in a structure nearly identical to the X-ray derived PDB-1AAR (gray). The key interface between the L8,I44,V70 hydrophobic patches is shown as yellow spheres. Restraints for docking are shown on the left.

4.3.2 Structural ensembles from linkage branching

With branched linkages it is possible that interactions between distal domains could form an entirely novel Ub signal, which would not be possible with homogeneously linked chains. To determine which branched Ub trimers could create this type of interaction, I modeled all twenty-eight of them using HADDOCK v2.1 following an established protocol which creates Ub-Ub linkages and simulates all possible interdomain interactions (*173*). The actual structures calculated for each run varied greatly, however based on generalized observations a conformation for a branched chain can be classified in one of four sub groups: (i) "free form" with no inter domain interactions, (ii) "distal hug" where branching causes the two distal Ubs to form an interaction, (iii) "clump" where all Ubs appear to interact, and (iv) "classic interface" that is defined by a K6 or K48 linkage retaining the same interface it would in a homogeneously linked chain. For a visual representation of each possible outcome see (**Figure 4.6**).

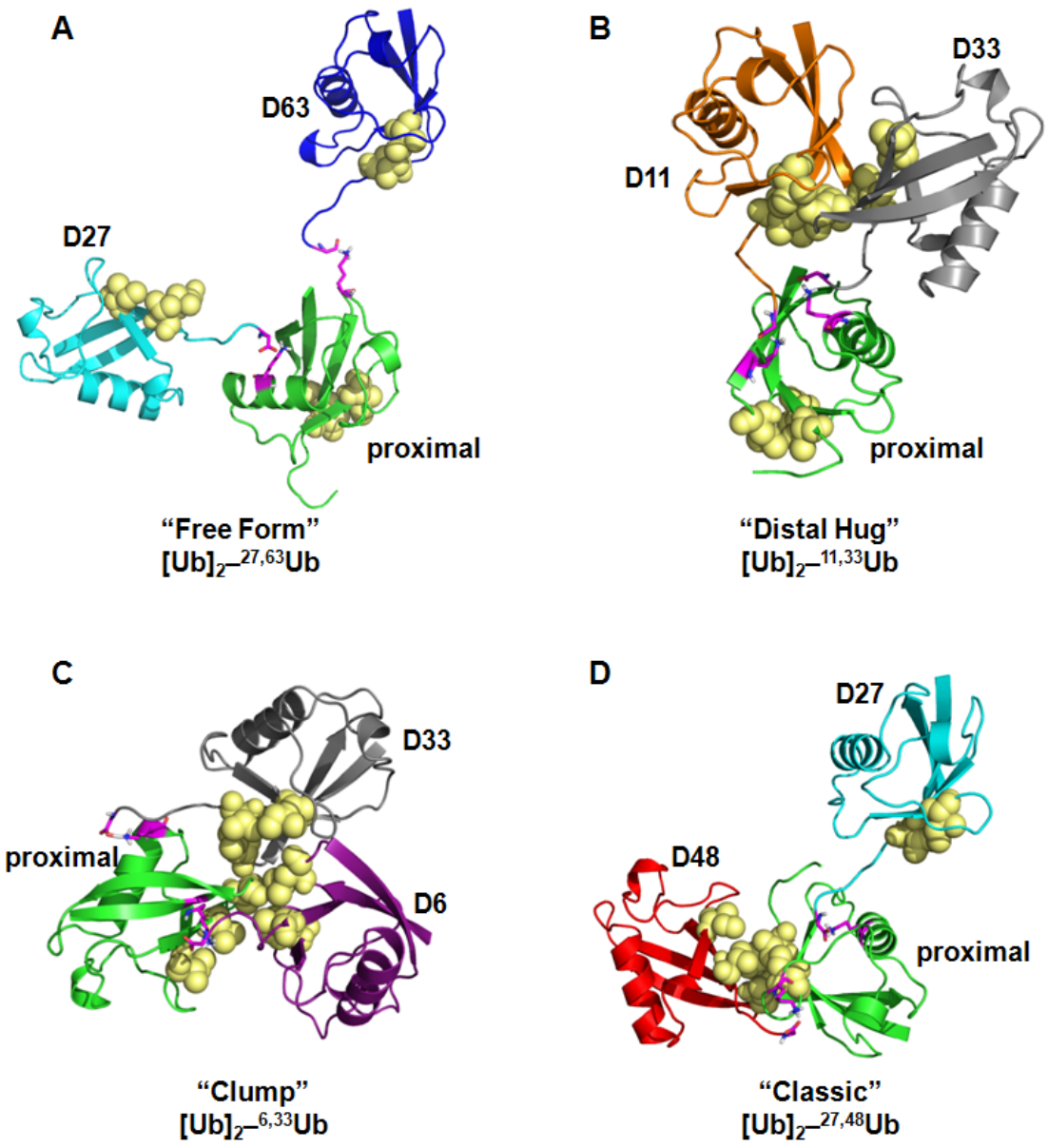


Figure 4.6 – Representation of possible outcomes from linkage branching with tri-Ub chains

The four possible outcomes were assigned to each cluster after analysis of HADDOCK output with every possible branched tri-Ub. (**A**) Depicts the "free form" in which no Ub has any interdomain contacts, $[Ub]_2$–^{27,63}Ub. (**B**) The novel "distal hug" is represented by $[Ub]_2$–^{11,33}Ub where the D11(orange) and D33 (gray) Ubs create a hydrophobic contact. (**C**) $[Ub]_2$–^{6,33}Ub exemplifies a "clump" where all hydrophobic patches from each Ub interact. (**D**) The "classic interface" is retained in $[Ub]_2$–^{27,48}Ub where D48 (red) interacts with the proximal Ub (green) using the same mechanism as the homogeneous chain, while the D27 Ub is excluded and not disruptive to the K48 signal.

4.3.3 Theoretical outcomes of linkage branching with Ub trimers

After analyzing the lowest energy clusters from HADDOCK, it was apparent that the theoretical structures did indeed represent all four possibilities that could result from linkage branching. Of particular interest were the distal hugs, which are the most likely to present new signals for receptors since the moiety formed by the distal Ubs can only result from linkage mixing. Perhaps even more interesting was the fact that linkages such as K27 and K63, which do not form any interdomain contacts in homogeneous chains, were able to create distal hugs in several branched chains. It is important to note that many of the tri-Ubs predicted to form distal hugs had varying degrees of hydrophobic interaction (buried surface area) and different interfaces even though they were classified in the same group. Several examples of distal hugs are depicted (**Figure 4.7**). Analysis of all clusters suggests that distal hugs could be more prevalent than thought. An interesting possibility is that linkages that do not promote interdomain contacts (e.g. K63) appear to promote distal hugs, suggesting that extended polyUb conformations may be susceptible to these interactions. This implies that although a homogenous chain is extended, once it is branched, new signals would arise from the distal hug interaction. This theory needs to be tested more extensively, but if true it would add much more versatility to Ub signaling.

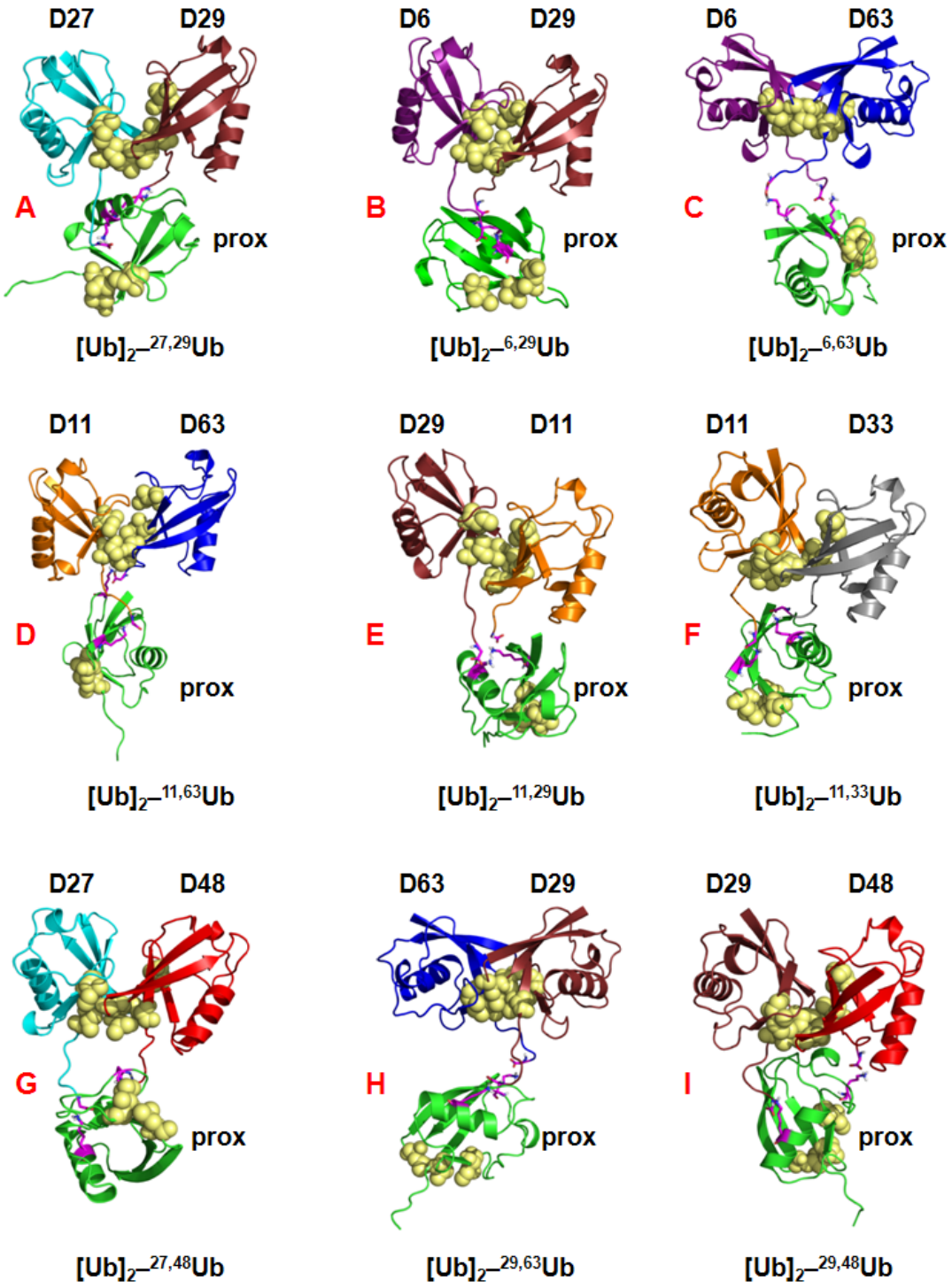


Figure 4.7 – Theoretical models of the distal hug conformation
(**A-I**) Nine, selected models from clusters classified as distal hugs. The proximal Ub (green) is free from interactions with the distal Ubs (colored by linkage). Formal names are listed below and the L8,I44,V70 hydrophobic patch is shown as yellow spheres.

4.3.4 A failed prediction $[Ub]_2$–^{11,63}Ub

Simulating the conformations of branched trimers is relatively straight forward, but proving these conformations is both difficult and time consuming. The K63 linkage is the easiest to test with our solution NMR methods. The fact that the distal Ub in K63-Ub_2 does not experience any interdomain contacts makes the linkage an ideal candidate to search for distal hugs. If a distal hug is formed in a branched trimer, the distal K63 Ub will show a characteristic pattern in $^{1}H,^{15}N$-HSQC, resulting from the new interdomain interactions in the conformation. Therefore, as an initial search synthesizing any 2D branched trimers with the distal 63 Ub ^{15}N labeled would provide an easy means for determining if a distal hug can be formed. The following unique branched trimers were candidates for investigation: Ub[Ub(^{15}N)]–^{1,63}Ub, Ub[Ub(^{15}N)]–^{6,63}Ub, Ub[Ub(^{15}N)]–^{11,63}Ub, Ub[Ub(^{15}N)]–^{27,63}Ub, Ub[Ub(^{15}N)]–^{29,63}Ub, Ub[Ub(^{15}N)]–^{33,63}Ub, and Ub[Ub(^{15}N)]–^{48,63}Ub. Given that $[Ub]_2$–^{11,63}Ub was predicted to form a distal hug, I decided to test if Ub[Ub(^{15}N)]–^{11,63}Ub showed any evidence of this conformation. As a control I recorded the spectrum of the corresponding Ub monomer and Ub(^{15}N)–^{63}Ub, (^{15}N enriched distal Ub in K63-Ub_2). Alignment by the proximal Ub of existing structures PDB-2XEW for K11-Ub_2 and PDB-2JF5 for K63-Ub_2 hints that the distal hug in $[Ub]_2$–^{11,63}Ub would require the D11 and D63 Ub to undergo significant rearrangement to adopt the conformation (**Figure 4.8**). The overlay of the $^{1}H,^{15}N$-HSQC show no significant difference between Ub_1, Ub(^{15}N)–^{63}Ub, and Ub[Ub(^{15}N)]–^{11,63}Ub and there are barley any CSPs between Ub(^{15}N)–^{63}Ub and Ub[Ub(^{15}N)]–^{11,63}Ub (**Figure 4.8**). Taken together this suggest that the distal hug conformation of $[Ub]_2$–^{11,63}Ub is invalid and the theoretical

alignment of the component dimers provides for a more accurate representation. However, determining the true structure requires a more extensive investigation.

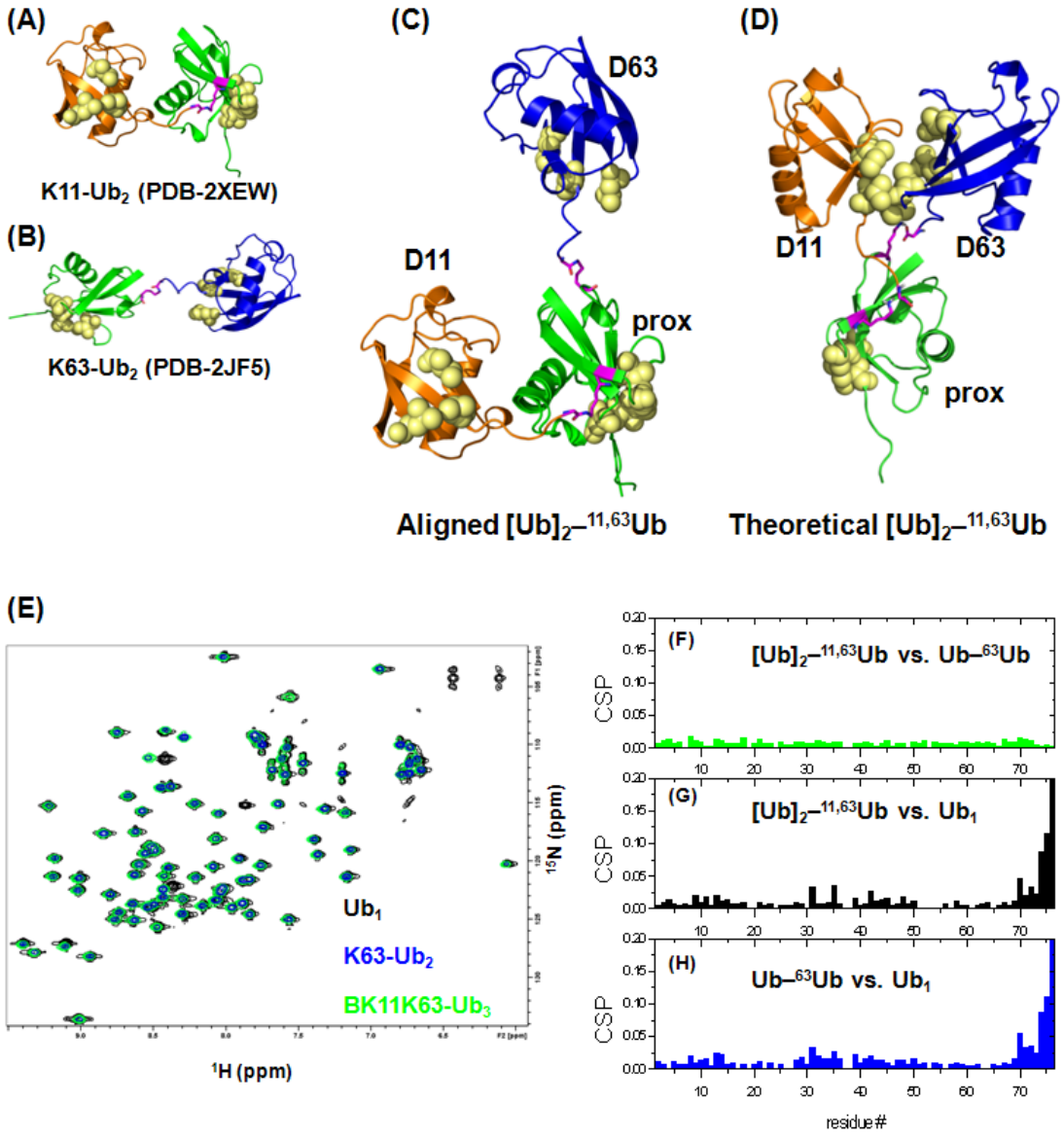


Figure 4.8 – Testing for the $[Ub]_2$–^{11,63}Ub distal hug.
(**A**) Structure of K11-Ub_2 (PDB-2XEW) and (**B**) K63-Ub_2 (PDB-2JF5). (**C**) Proximal alignment of model for $[Ub]_2$–^{11,63}Ub based on existing X-ray structures with no distal hug. (**D**) Theoretical HADDOCK model for $[Ub]_2$–^{11,63}Ub showing the distal hug between D11 and D63. (**E**) ^{1}H,^{15}N-HSQC overlay of Ub_1 in black, Ub(^{15}N)–^{63}Ub in blue, and Ub[Ub(^{15}N)]–^{11,63}Ub in green. (**F**) Residue specific CSPs between Ub(^{15}N)–^{63}Ub and Ub[Ub(^{15}N)]–^{11,63}Ub, (**G**) Ub_1 and Ub[Ub(^{15}N)]–^{11,63}Ub, and (**H**) Ub_1 and Ub(^{15}N)–^{63}Ub.

4.5 Structural studies of K48 and K63 mixed linkage polyUb chains

4.5.1 Precedent and research aims

Although many homogeneous linkages remain uncharacterized, there is a mounting body of evidence that polyUb chains contain mixed linkages in the cell (*28, 29, 144, 164, 198-201*). Mass spectrometry studies have confirmed that all eight Ub linkages are present *in vivo* and quantitative analyses has established that K48 and K63 linkages are the most abundant by a large margin (*202, 203*). Recent advances have led to the development of linkage-specific antibodies (for K11, K48, K63, and M1) which have revealed that the polyUb signal is remodeled on several substrates, and also instances where K48 and K63 linkages co-localized in the cell (*160, 201, 202, 204, 205*). However, it is still unclear if linkage mixing or branching creates a new polyUb signaling property unavailable to homogeneous chains, or if the signaling properties of the individual linkages in mixed-linkage chains would be preserved.

Precedent for a functional mixed-linkage chain comes from Ring1b, an E3 Ub ligase that requires autoubiquitination with a mixed polyUb chain containing K6, K27, and K48 linkages (*164*). For this example, our understanding of the structural and signaling properties is limited because the topology and sequence of Ub–Ub linkages in this novel polyUb chain are unknown. In addition, because the solution structures of homogenous K6 and K27 polyUb chains are unavailable, comparisons of the conformation of the chains from Ring1b chain with component homogenous chains are not possible. A mixed linkages chain containing K11 and K63 linkages has also been shown to facilitate internalization of MHC I membrane proteins (*200*). At

some point during substrate processing on the 26S proteasome, the polyUb chain can be remodeled by proteasome associated DUBs and E3 Ub ligases and in yeast, the E3/E4 Ub ligase Hul5 has been shown to make linkages primarily through K63, but also K11 and K48 (*199*). Although these and other studies indicate that complex polyUb chains occur *in vivo*, they do not provide clues about the structures of the chains or how they could be recognized differently by receptors. Thus, our understanding of the roles and signaling properties of branched or unbranched mixed linkage chains is quite limited, in part because structural information is lacking.

4.6 Design and synthesis of unbranched and branched mixed linkage polyUb chains

4.6.1 Justification for K48 and K63 mixed linkages as a starting point

One challenge to studying complex polyUb chains is that, depending on the number of individual Ub units, there are theoretically quadrillions of unique chains (discussed in **section 4.1**). Even discussion of this problem is confounded by the absence of a standardized nomenclature system for such chains, which required me to design such a system to clearly communicate my findings (see **section 4.3.1**). To initiate the first structural study of complex polyUb chains I focused on branched and unbranched mixed-linkage chains containing K48 and K63 linkages as a logical starting point. The fact that the two linkages are essentially "orthogonal" with respect to (i) their location on the surface of Ub, (ii) the resulting structural conformations (compact versus extended (*153, 155*)), and (iii) the signaling properties (proteolytic versus regulatory (*154*)) of the corresponding homogeneously-linked chains suggests that combinations of K48 and K63 linkages could provide the most extreme example

of linkage mixing and branching. Additionally, the high relative abundances of both K48 and K63 linkages in the cell suggest that these linkages would predominate in randomly assembled mixed-linkage chains. Although a cellular process that requires both K48 and K63 linkages has yet to be identified, it has been reported that K48 and K63 linkages co-localize in the cell, both linkages were detected on the same substrate, and at least one DUB, ataxin-3, preferentially cleaves mixed K48 and K63-linkage chains *in vitro* (*201, 202, 206*). These studies hint that polyUb chains containing both K48 and K63 linkages form in the cell, but whether they serve a specific function or are simply a mistake that is later "edited" remains to be seen. Thus the main objective of this study is to determine the structural and functional outcomes of K48 and K63 linkage mixing in polyUb chains.

4.6.2 Fundamental unbranched and branched chains

Given that this is the first structural study of such chains, I found it logical to focus on the simplest possible model system by limiting chains to just three Ubs with combinations of K48 and K63 linkages. With these constraints, the resulting set of Ub timers included a single branched chain, $[Ub]_2$–^{48,63}Ub, and two unbranched mixed-linkage chains, Ub–^{63}Ub–^{48}Ub and Ub–^{48}Ub–^{63}Ub (**Figure 4.9**). To have an unprecedented look at the solution properties of these chains, I had to create each chain with only one Ub ^{15}N enriched resulting in a total of nine chains for analysis, see (**Figure 4.9**).

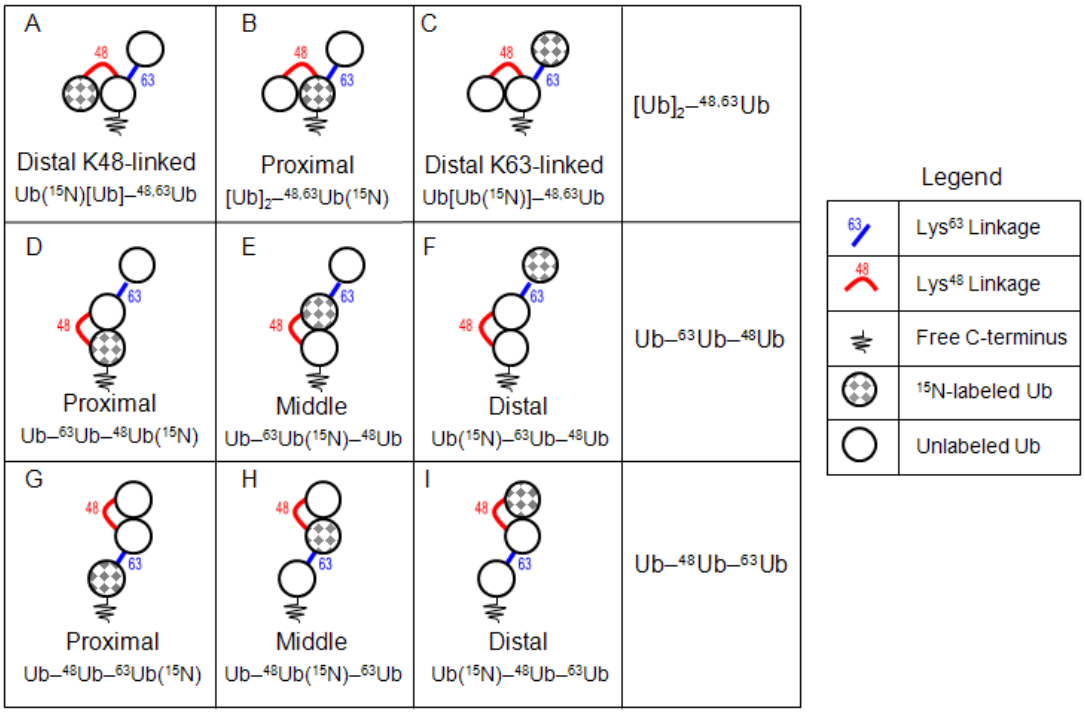


Figure 4.9 – Fundamental unbranched and branched trimers.
Chain schematics, nomenclature, and unit-specific ^{15}N-enrichment (gray checker pattern) for each of the nine chains studied. Rows depict isotope labeling schemes for an individual chain and the ^{15}N labeled Ub is written out (distal-48, distal-63, distal, proximal, or middle). The formal name of each chain, including labeling is shown under each schematic. **(A-C)** the branched chain $[Ub]_2$–^{48,63}Ub, **(D-F)** unbranched mixed linkage chain Ub–^{63}Ub–^{48}Ub, **(G-I)** unbranched mixed linkage chain Ub–^{48}Ub–^{63}Ub.

4.6.3 Synthesis of chains

To assemble the nine polyUb chains for this study I utilized the well established enzymatic synthesis approach (*207*), but had to make several modifications. The linkage specific E2s, E2-25K forK48 linkages and Ubc13:Mms2 for K63 linkages efficiently created the desired linkages, however I needed to ensure no unwanted modifications were introduced. This required several new KtoR Ub mutants as well as a carefully thought out stepwise synthesis scheme. The need to selectively label specific Ubs in each chain also dictated the use of the stepwise

approach. The necessary steps for the synthesis of each chain are show in (**Figure 4.10**). The finding that E2-25K could easily create K48 linkages on an existing K63 chain and Ubc13:Mms2 could introduce K63 linkages to a K48 chain was encouraging and suggested that these naturally occurring E2s could do the same in the cell.

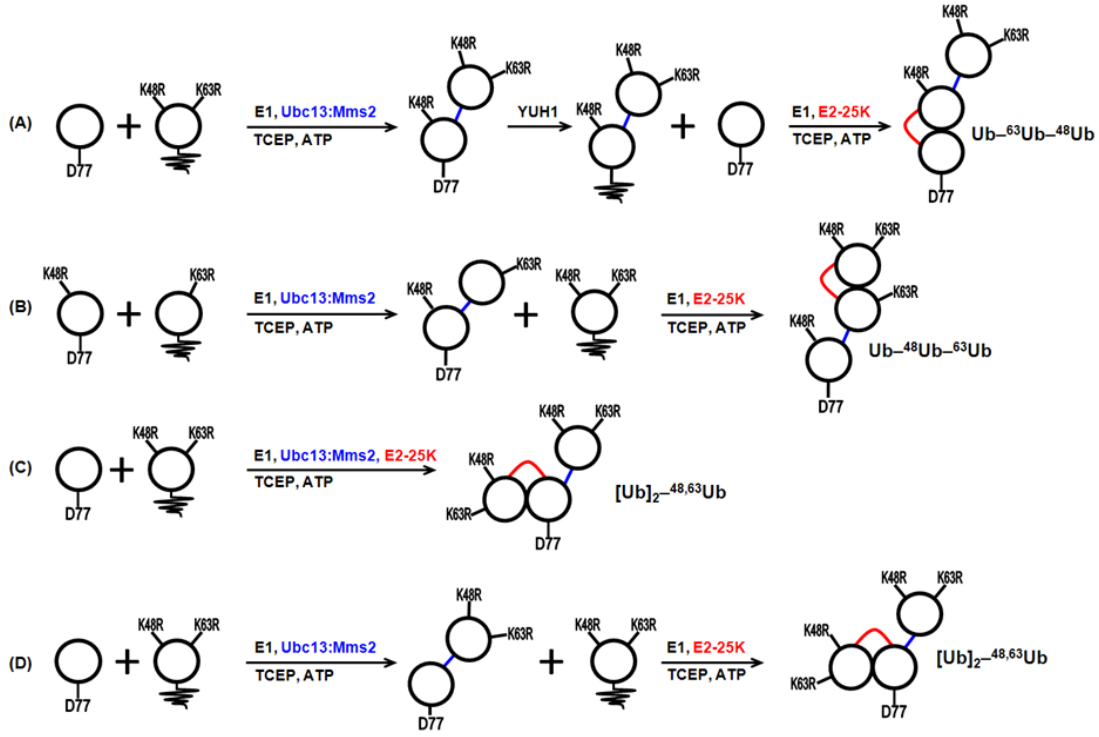


Figure 4.10 – Stepwise enzymatic synthesis of trimers for study.
D77 terminates chain elongation from the proximal end while a K-to-R mutation terminates chain elongation from that lysine on the distal end. D77 can be removed after treatment with YUH1, exposing a free G76 terminus to be used for chain formation. Alternatively, a truncated (Ub-R74) variant of Ub that lacks G75 and G76 can be used for the proximal-to-be unit in these chains. **(A)** Steps used to create Ub–^{63}Ub–^{48}Ub with full control of which Ub unit is ^{15}N enriched. **(B)** Assembly of Ub–^{48}Ub–^{63}Ub in two separate steps. **(C)** Two linkage-specific E2s are used simultaneously to form [Ub]$_2$–^{48,63}Ub. **(D)** Alternative method to assemble the branched [Ub]$_2$–^{48,63}Ub with full control of the placement of the ^{15}N enriched Ub.

In the one step synthesis of the branched chain $[Ub]_2$–^{48,63}Ub, E2-25K and Ubc13-Mms2 simultaneously add each linkage in the same reaction. I explored this concept further by adding Ube2s, an E2 that forms K11 linkages to the mixture which yielded $[Ub]_3$–$^{11, 48,63}$Ub (**Figure 4.11**). The results demonstrate that these naturally occurring linkage-specific E2s could readily form branched or unbranched polyUb chains working together and on existing chains, suggesting this was a common property of all E2's and E3's.

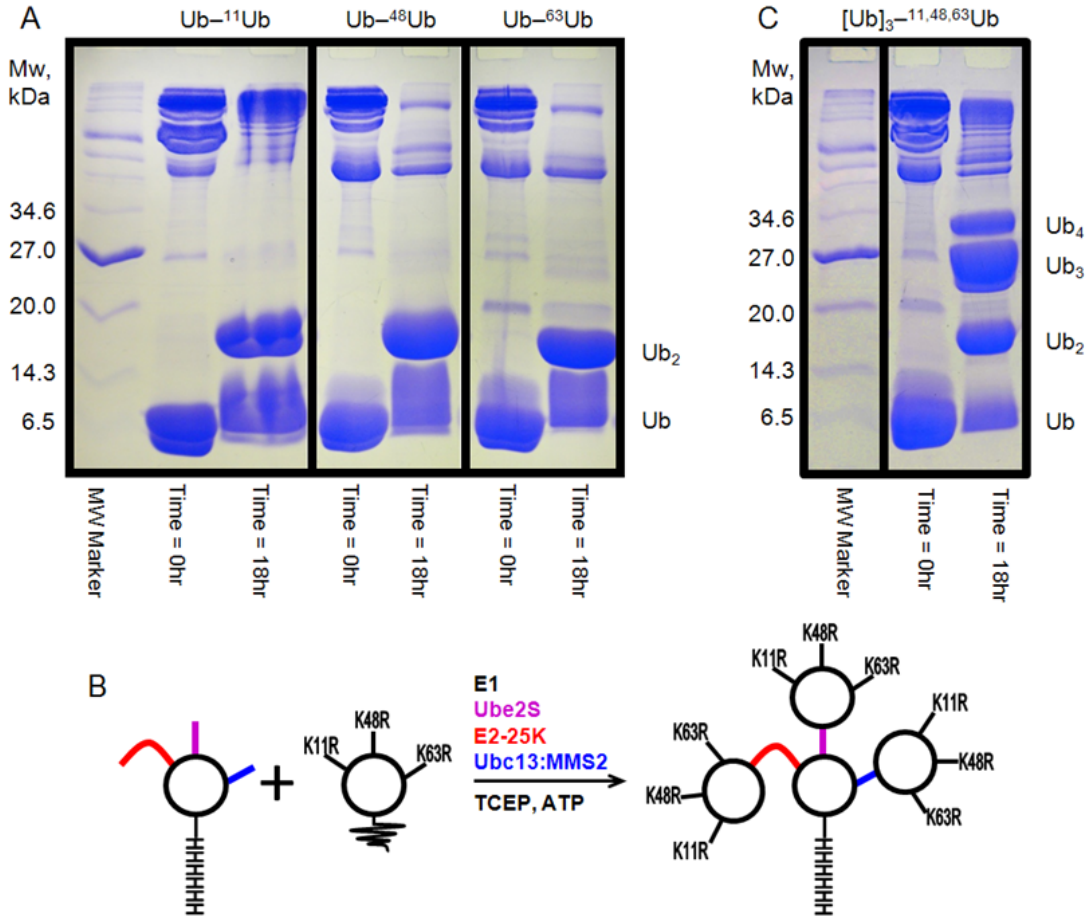


Figure 4.11 –Assembly of a branched $[Ub]_3$–^{11,48,63}Ub tetra-Ub by simultaneous action of three linkage specific E2s.
Various linkage-specific E2 ubiquitin-conjugating enzymes can be used simultaneously to create highly branched polyUb chains in which a single Ub unit is ubiquitinated with multiple Ubs attached through different lysines. Here we demonstrate the formation of a branched tetra-Ub, $[Ub]_3$–^{11,48,63}Ub, containing a single proximal Ub and three distal Ubs, linked to it via K11, K48, and K63. **(A)** 15% SDS PAGE gel showing that enzymatic reactions with individual linkage-specific E2s Ube2S, E2-25K, and Ubc13:Mms2 efficiently create only di-Ub chains. **(B)** Synthesis scheme to assemble $[Ub]_3$–^{11,48,63}Ub from Ub monomers using the simultaneous action of all three linkage specific E2s. **(C)** 15% SDS PAGE gel demonstrates that $[Ub]_3$–^{11,48,63}Ub is formed (the Ub_4 band) using the reaction shown in **(B)**.

4.6.4 NMR validation of isopeptide linkages

Initially, quality control was one of the most pressing concerns after the synthesis of each mixed linkage chain. Knowing the linkage type (K48 or K63) and its exact position in the chain were essential to confirming that the exact chain was synthesized. Simple gel assays could easily confirm the presence of tri-Ub and also contaminants di-Ub and mono-Ub. While more sensitive techniques such as tryptic digestion or linkage-specific antibodies could detect the linkages present, but the methods provided no information about the sequence of the linkages in the chain. To determine that the exact linkage(s) were present on the desired Ub, I used $^{1}H,^{15}N$-HSQC to observe diagnostic isopeptide signals in each of the nine chains. Unlike any other method, a single Ub in each timer can be noninvasively monitored to revel which linkages it contains. If the Ub under observation is a distal or middle (endo-1), then the C-terminus should reflect a ligated G76, while a proximal Ub should have a completely free G76 signal. Each of the nine trimers adhered to these rules as expected. To determine the linkage type, I looked for the diagnostic signals that result from the newly formed N-H group in the ligated lysine side chain. Provided the ligated lysine is ^{15}N enriched, the transition from a free εNH_2 group to the NH isopeptide will create a new signal in $^{1}H,^{15}N$-HSQC. This new signal has a specific chemical shift, whose position depends on the lysine forming the linkage. For example, the signal resulting from K48 has a different position from the same εNH_2 group in K63 (**see Figure 4.12**). This unconventional, but robust $^{1}H,^{15}N$-HSQC assay confirmed each of the nine chains were assembled as desired.

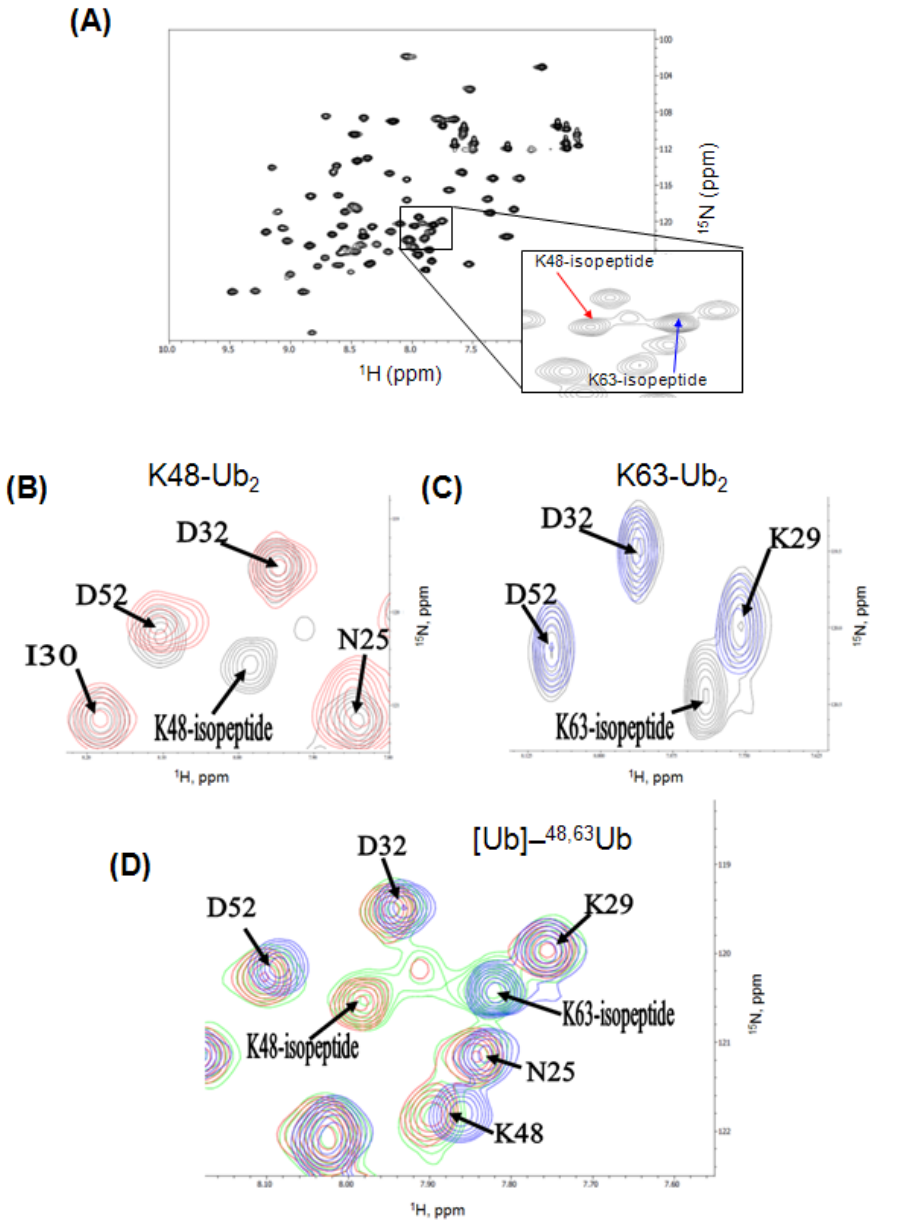


Figure 4.12 – Linkage specific diagnostic isopeptide signals in ^{1}H,^{15}N-HSQC. (**A**) ^{1}H,^{15}N-HSQC spectrum of [Ub]$_2$–^{48,63}Ub(^{15}N) with isopeptide zoomed in. (**B-C**) Overlay of the ^{1}H-^{15}N-HSQC spectra of the proximal Ub in **(B)** Ub–^{48}Ub(^{15}N) or **(C)** Ub–^{63}Ub(^{15}N) (black contours) with the corresponding spectra of the distal Ub in the distal-labeled constructs: Ub(^{15}N)–^{48}Ub (red contours) and Ub(^{15}N)–^{63}Ub (blue contours), reveal extra signals corresponding to the isopeptide bonds. **(D)** If we then overlay the ^{1}H-^{15}N-HSQC spectrum of the proximal Ub of [Ub]$_2$–^{48,63}Ub(^{15}N) (green) with the spectra of the proximal Ubs of Ub–^{48}Ub(^{15}N) (red) and Ub–^{63}Ub(^{15}N) (blue) it is obvious that the spectrum of [Ub]$_2$–^{48,63}Ub(^{15}N) contains isopeptide signals from both the K48 and K63 linkages, while the spectra of Ub–^{48}Ub(^{15}N) and Ub–^{63}Ub(^{15}N) alone each contains only one isopeptide signal. This allows for NMR detection of both K48 and K63 linkages in [Ub]$_2$–^{48,63}Ub as well as the other chains.

4.7 Structural properties of component linkages are preserved in K48, K63 mixed linkage chains

4.7.1 Chemical shift perturbation mapping of inter Ub contacts

NMR chemical shift perturbation (CSP) mapping was used to identify interactions between the Ub units in the mixed linkage chains. In order to unravel monomer-specific contacts, each Ub unit in $[Ub]_2$–^{48,63}Ub, Ub–^{63}Ub–^{48}Ub, and Ub–^{48}Ub–^{63}Ub (see **Figure 4.9** for chain notations) was individually ^{15}N enriched (*i.e.*, one ^{15}N-Ub per chain) resulting in nine distinct ^{15}N-labeled tri-Ub constructs. ^{1}H,^{15}N-NMR spectra of each chain were acquired under identical conditions and compared to those of monomeric Ub and of the corresponding Ub unit in Ub–^{48}Ub and Ub–^{63}Ub. Consistent with the absence of detectable non-covalent contacts between Ub units in Ub–^{63}Ub, the only CSPs (versus mono-Ub) observed in that chain were in the C-terminal residues of the distal Ub and those immediately surrounding K63 of the proximal Ub, reflecting the residues involved in the isopeptide linkage between these two units (**Figure 4.13 A,C**). By contrast, CSP mapping of Ub–^{48}Ub revealed that both the distal and the proximal Ubs exhibited highly specific spectral perturbations in and around the hydrophobic surface patch residues (L8, I44,V70). These CSPs, observed in addition to those in the vicinity of the isopeptide linkage between the C-terminal G76 of the distal Ub and K48 of the proximal Ub, are a clear indicator of the hydrophobic interface between the two Ubs in Ub–^{48}Ub (**Figure 4.13 B,D)**. These distinctive features of the NMR spectra of Ub–^{63}Ub and Ub–^{48}Ub serve as hallmarks of the corresponding linkages and the resulting Ub–Ub contacts (*152, 153, 155*).

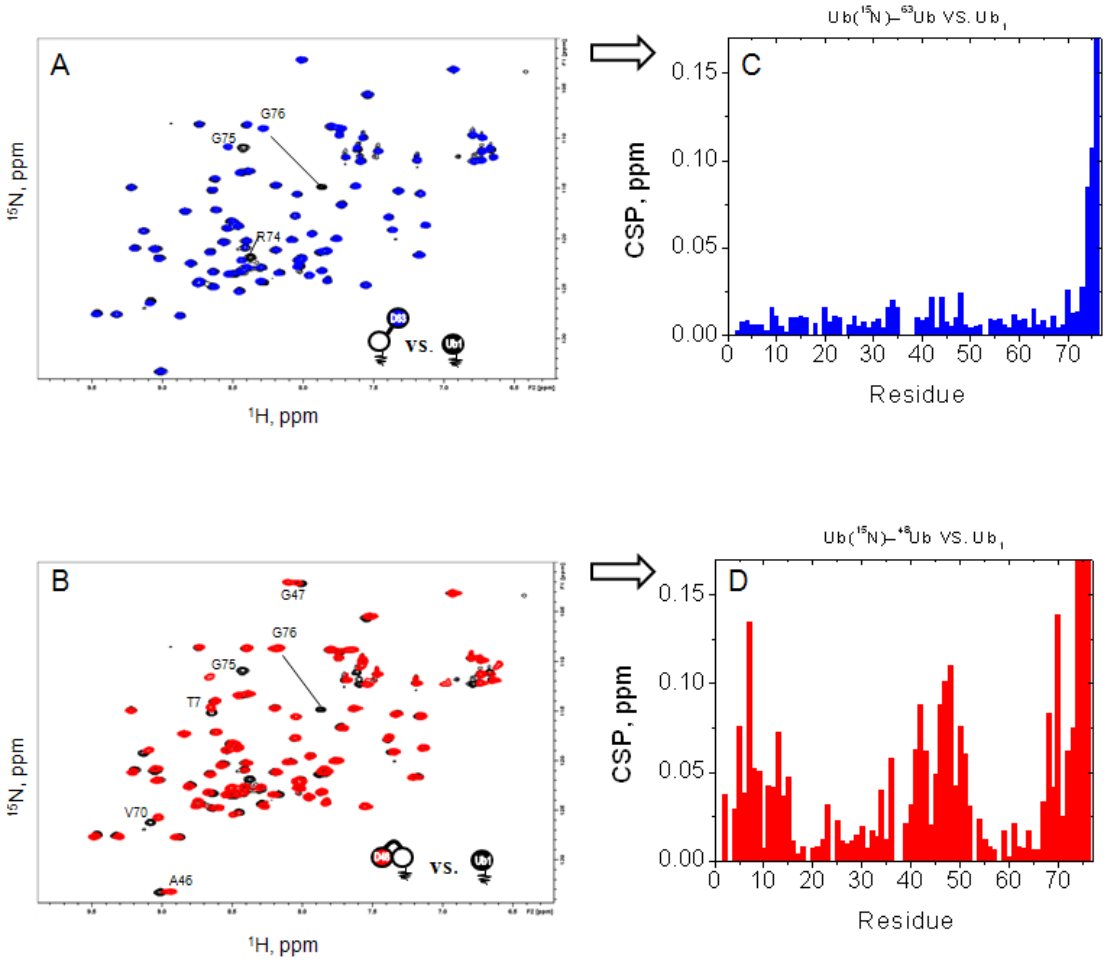


Figure 4.13 – CSP analysis from ^{1}H,^{15}N-HSQC spectra
(**A**) ^{1}H,^{15}N-HSQC overlay of Ub$_1$(black) and Ub(^{15}N)–^{63}Ub (blue) shows similar residues specific chemical shifts with the exception of the C-terminus (**B**) ^{1}H,^{15}N-HSQC overlay of Ub$_1$(black) and Ub(^{15}N)–^{48}Ub (red) shows different residue specific peaks for the C-terminus and hydrophobic patch residues. (**C**) Calculated residue specific CSPs between spectra for Ub$_1$(black) and Ub(^{15}N)–^{63}Ub (blue). (**D**) Calculated residue specific CSPs between spectra for Ub$_1$(black) and Ub(^{15}N)–^{48}Ub (red).

NMR spectra of [Ub]$_2$–^{48,63}Ub show that the amide resonances in the distal K63-linked Ub are almost identical to those in monomeric Ub with the exception of the C-terminal residues 74-76, where the observed CSPs are caused directly by ligation to K63 of the proximal Ub (**Figure 4.14 C**). Consistent with this observation,

there are virtually no spectral differences between distal K63-linked Ub and the distal Ub in Ub–^{63}Ub, indicating their close structural similarity and the absence of non-covalent contacts with the proximal Ub in the corresponding chains (**Figure 4.14 I**).

A similar comparison of the spectra of the distal K48-linked Ub in $[Ub]_2$–^{48,63}Ub revealed a strikingly different picture. Here I detected large site-specific CSPs between the distal K48-linked Ub and mono-Ub (**Figure 4.14 B**) and almost negligible spectral differences between the distal K48-linked Ub and the distal Ub in Ub–^{48}Ub (**Figure 4.14 E**). These results show that the distal K48-linked Ub in $[Ub]_2$–^{48,63}Ub makes essentially the same interdomain contacts (*i.e.*, the hydrophobic interface with the proximal Ub and the isopeptide linkage through G76) as the distal Ub in Ub–^{48}Ub. Predictably, these contacts result in large site-specific spectral differences between the distal K48-linked Ub and the distal Ub of Ub–^{63}Ub (**Figure 4.14 H**), which are also similar to the CSPs between the distal K48-linked Ub and monomeric Ub (**Figure 4.14 B**).

Careful analysis of the spectra for the proximal Ub of $[Ub]_2$–^{48,63}Ub revealed two signals originating from the isopeptide εNH groups of K48 and K63 (**Figure 4.12**). These signals are diagnostic for the isopeptide linkage through the ε-amino group of the corresponding lysine. That they are at the same resonance frequencies as in the respective homogeneous di-Ub (K48- or K63-linked) supports that there is little difference in the linkage between the branched and component homogenous di-Ub chains. For both unbranched mixed-linkage tri-Ub chains (see below) the isopeptide εNH signals appeared at nearly identical positions in ^{1}H,^{15}N-HSQC.

Because the proximal Ub of $[Ub]_2$–^{48,63}Ub is ubiquitinated at two lysines simultaneously, its NMR spectrum could differ from the proximal-Ub spectrum in the respective isolated di-Ubs. Nevertheless, a clear picture emerged from comparison of the corresponding spectra. For example, a comparison to mono-Ub indicates that the hydrophobic-patch residues of the proximal Ub form an interface with another Ub unit, while its C-terminus is unligated (as in monomeric Ub) unlike those of the distal K48- and K63-linked Ubs (**Figure 2.14 A**). The spectral differences between the proximal Ubs of $[Ub]_2$–^{48,63}Ub and Ub–^{48}Ub are minimal except for the region around K63 which is linked to another Ub only in the branched trimer (**Figure 2.14 D**). As with the distal K48-linked Ub (see above), the strong site-specific CSPs with respect to the proximal Ub in Ub–^{63}Ub reflect the hydrophobic-patch contacts between the two K48-linked Ubs (**Figure 2.14 G**). Based on these and the abovementioned data, I conclude that the two K48-linked Ubs (the distal-48 and proximal) in $[Ub]_2$–^{48,63}Ub form essentially the same hydrophobic contact as in Ub–^{48}Ub whereas only the distal K63-linked Ub behaves as in Ub–^{63}Ub, where non-covalent contact were not detected.

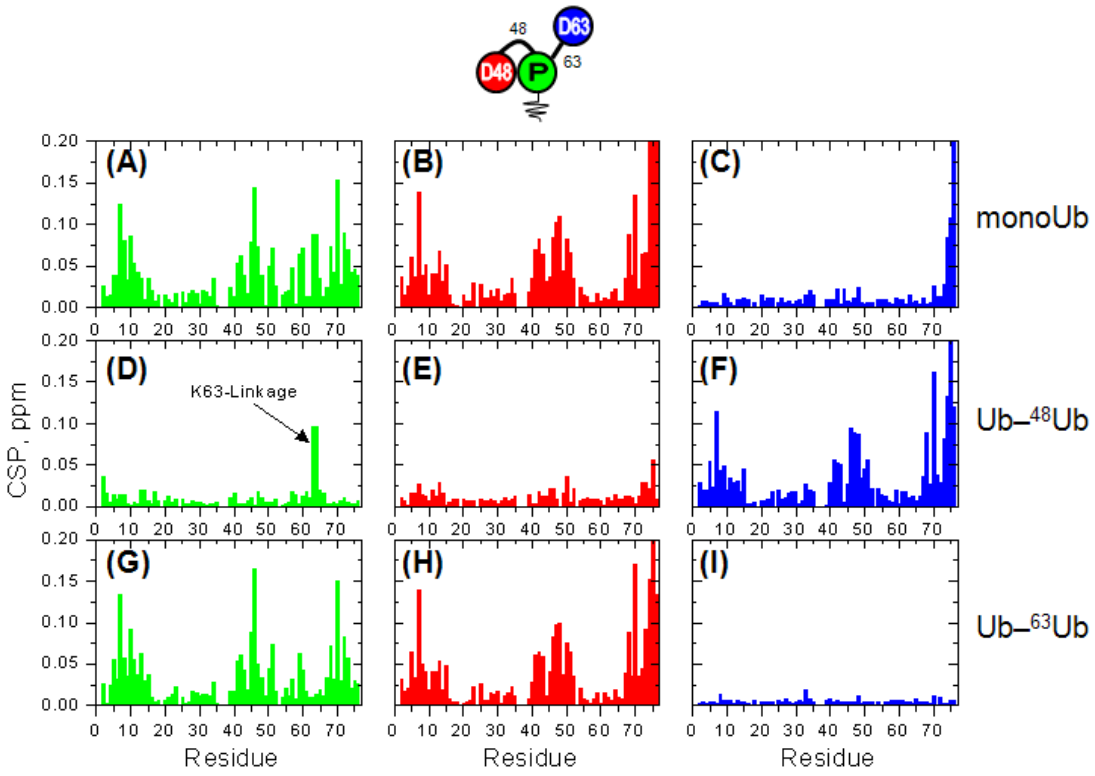


Figure 4.14 – ^{1}H,^{15}N-HSQC CSP characterization of $[Ub]_2$–^{48,63}Ub.
Data for the proximal Ub are shown in green, for distal K48-linked Ub in red, and for distal K63-linked Ub in blue (see pictograph, *top*). (**A-I**) Spectral differences (quantified as CSPs) between each Ub in $[Ub]_2$–^{48,63}Ub and monomeric Ub (top row), the corresponding Ub units in Ub–^{48}Ub (middle row), and Ub–^{63}Ub (bottom row). Left column: CSPs between the proximal Ub in $[Ub]_2$–^{48,63}Ub and (**A**) Ub_1, (**D**) proximal Ub in Ub–^{48}Ub, and (**G**) proximal Ub in Ub–^{63}Ub. Middle column: CSPs between distal K48-linked Ub in $[Ub]_2$–^{48,63}Ub and (**B**) Ub_1, (**E**) distal Ub in Ub–^{48}Ub, and (**H**) distal Ub in Ub–^{63}Ub. Right column: CSPs between the distal K63-linked Ub in $[Ub]_2$–^{48,63}Ub and (**C**) Ub_1, (**F**) distal Ub in Ub–^{48}Ub, and (**I**) distal Ub in Ub–^{63}Ub.

When we rearranged the linkages to form an unbranched mixed-linkage chain, Ub–^{63}Ub–^{48}Ub, the Ub units maintained their expected contacts and spectral properties. The proximal Ub of Ub–^{63}Ub–^{48}Ub has its K48 ligated, and a hydrophobic interface indicated from the large site-specific CSPs relative to mono-Ub and Ub–^{63}Ub (**Figure 15 A,G**) was detected. The spectral differences are minimal between the proximal Ub in Ub–^{48}Ub and that in Ub–^{63}Ub–^{48}Ub, which indicates close similarity between the two units ligated via K48 (**Figure 15 D**). The next (*i.e.,* middle) Ub of Ub–^{63}Ub–^{48}Ub has features of both a proximal unit (linked through its K63) and a distal unit linked through its C-terminal G76. The large site-specific CSPs versus mono-Ub (**Figure 15 B**) clearly support that the middle Ub participates in a hydrophobic interface. Based on the spectral similarity with the distal Ub of Ub–^{48}Ub (**Figure 15 E**), I conclude that the K48-linked proximal and middle Ubs in Ub–^{63}Ub–^{48}Ub retain the hydrophobic interface characteristic of Ub–^{48}Ub. Note also that the large CSPs observed around K63 in the middle Ub reflects the linkage at K63 to the distal Ub.

The NMR spectra of the distal Ub in Ub–^{63}Ub–^{48}Ub demonstrate that this unit is excluded from the hydrophobic interface formed between the middle and proximal Ubs. Indeed, the CSPs in distal versus monomeric Ub are localized to the (ligated) C-terminal region (**Figure 15 C**), whereas the spectral differences between the distal Ubs of Ub–^{63}Ub–^{48}Ub and Ub–^{63}Ub are negligible (**Figure 15 I**). These observations strongly support that the distal Ub of Ub–^{63}Ub–^{48}Ub has the same structure and contacts (or absence thereof) as the distal Ub in Ub–^{63}Ub.

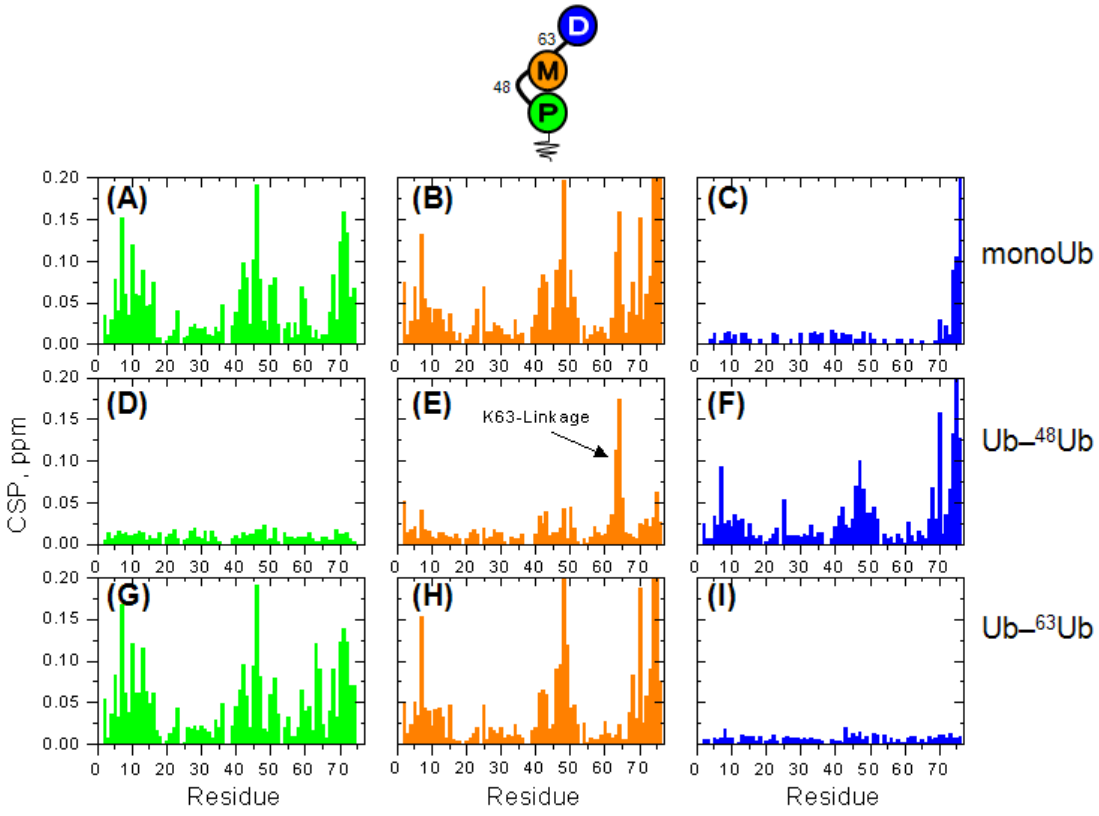


Figure 4.15 – ^{1}H,^{15}N-HSQC CSP characterization of Ub–^{63}Ub–^{48}Ub
Data for the proximal Ub are shown in green, for middle Ub in orange, and for the distal K63-linked Ub in blue (see pictograph, *top*). (**A-I**) Spectral differences (quantified as CSPs) between each Ub in Ub–^{63}Ub–^{48}Ub and monomeric Ub (top row), the corresponding Ub units in Ub–^{48}Ub (middle row), and Ub–^{63}Ub (bottom row). Left column: CSPs between the proximal Ub in Ub–^{63}Ub–^{48}Ub and (**A**) Ub_1, (**D**) proximal Ub in Ub–^{48}Ub, and (**G**) proximal Ub in Ub–^{63}Ub. Middle column: CSPs between middle Ub in Ub–^{63}Ub–^{48}Ub and (**B**) Ub_1, (**E**) distal Ub in Ub–^{48}Ub, and (**H**) distal Ub in Ub–^{63}Ub. Right column: CSPs between the distal Ub in Ub–^{63}Ub–^{48}Ub and (**C**) Ub_1, (**F**) distal Ub in Ub–^{48}Ub, and (**I**) distal Ub in Ub–^{63}Ub.

With a similar pattern emerging from the other two chains I expected it to repeat for the Ub–^{48}Ub–^{63}Ub chain. In Ub–^{48}Ub–^{63}Ub, the proximal Ub should be the only unit that does not form a hydrophobic interface. Indeed, this is evident from the strong spectral similarity of that proximal Ub with the proximal Ub of Ub–^{63}Ub (**Figure 4.16 G**) as well as with mono-Ub (**Figure 4.16 A**). Comparison of the proximal Ub in Ub–^{48}Ub–^{63}Ub to that of Ub–^{48}Ub clearly indicates the absence of the hydrophobic interface (**Figure 4.16 D**). In contrast to the proximal Ub which exhibits no non-covalent contacts, the middle and distal units of Ub–^{48}Ub–^{63}Ub form a characteristic K48-linked Ub–Ub interface. This is particularly apparent from the minimal CSPs (except for the ligated C-terminus) between the middle Ub of Ub–^{48}Ub–^{63}Ub and the proximal Ub of Ub–^{48}Ub (**Figure 4.16 E**), and the virtual absence of CSPs between the distal Ub and Ub–^{48}Ub (**Figure 4.16 F**). As expected from previous observations, the middle and distal Ubs in Ub–^{48}Ub–^{63}Ub show significant spectral differences when compared to monomeric Ub and Ub–^{63}Ub (**Figure 4.16 B,C,H,I**).

CSP analysis of all of all nine tri-Ubs clearly demonstrated that K48 linked elements retain their structural properties and could override K63 linkages in some cases such as the branched chain $[Ub]_2$–^{48,63}Ub where the proximal is also K63 linked, but fully takes the structural characteristics of the K48 linkage.

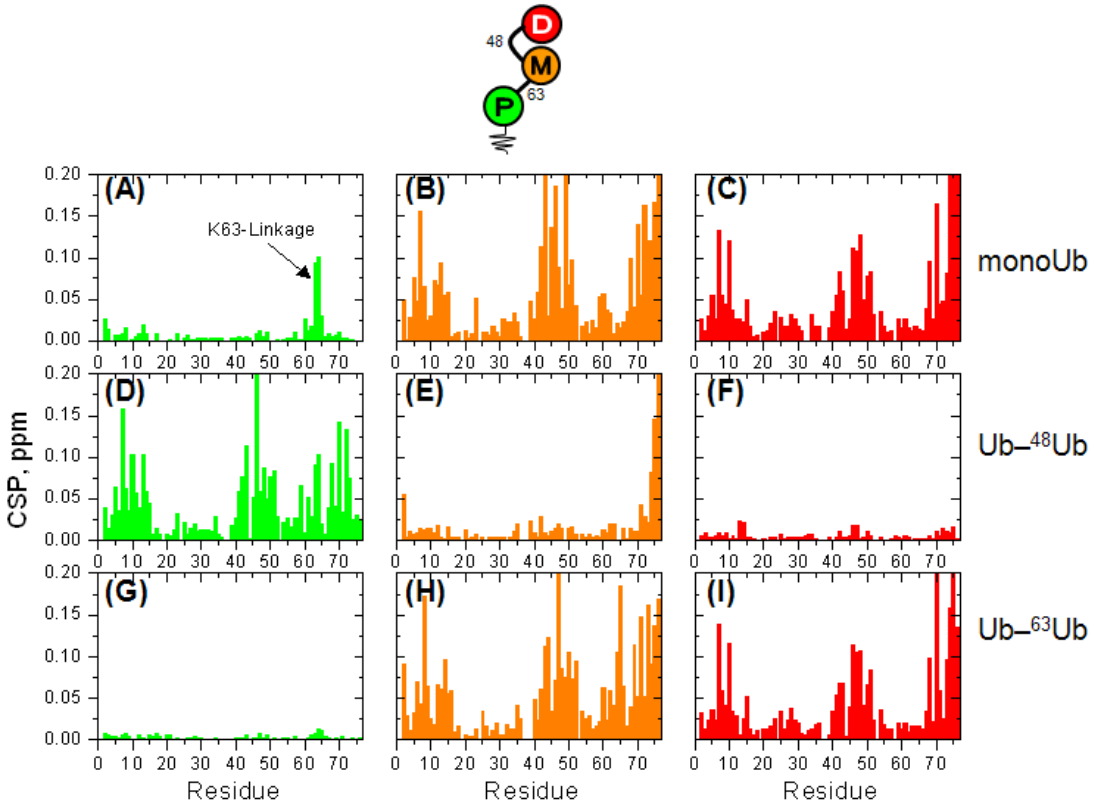


Figure 4.16 – ^{1}H,^{15}N-HSQC CSP characterization of Ub–^{48}Ub–^{63}Ub
Data for the proximal Ub are shown in green, for middle Ub in orange, and for the distal K48-linked Ub in red (see pictograph, *top*). (**A-I**) Spectral differences (quantified as CSPs) between each Ub in Ub–^{48}Ub–^{63}Ub and monomeric Ub (top row), the corresponding Ub units in Ub–^{48}Ub (middle row), and Ub–^{63}Ub (bottom row). Left column: CSPs between the proximal Ub in Ub–^{48}Ub–^{63}Ub and (**A**) Ub_1, (**D**) proximal Ub in Ub–^{48}Ub, and (**G**) proximal Ub in Ub–^{63}Ub. Middle column: CSPs between middle Ub in Ub–^{48}Ub–^{63}Ub and (**B**) Ub_1, (**E**) distal Ub in Ub–^{48}Ub, and (**H**) distal Ub in Ub–^{63}Ub. Right column: CSPs between the distal Ub in Ub–^{48}Ub–^{63}Ub and (**C**) Ub_1, (**F**) distal Ub in Ub–^{48}Ub, and (**I**) distal Ub in Ub–^{63}Ub.

4.7.2 Conformational mobility observed by T_1 relaxation

The CSP analysis from the previous section clearly indicates that the two K48-linked units in all three tri-Ub chains form a specific interface while the third unit, linked via K63, shows no contacts except for being tethered (via its flexible C-terminus or via K63) and therefore is expected to be more mobile than the other two Ubs. To test this assumption, ^{15}N longitudinal relaxation times, T_1, for each Ub unit in all nine chains were measured and compared. Generally, T_1 senses the overall tumbling of the chain (reflecting its size and shape), the relative intra-chain mobility (on a ns time scale) of the Ub unit under observation (which could depend on its location and intra-chain contacts), and the local polypeptide backbone dynamics within the Ub protomer (which are expected to be similar for all Ubs). Whereas the overall shape could differ between chains, the T_1 values within the same chain are expected to reflect the relative mobility of each Ub unit. For example, if a particular Ub forms a hydrophobic interdomain interface it becomes less mobile and this should result in an increase in T_1. Furthermore, a Ub that is linked to two other Ubs should also show an increase in T_1 as its reorientation will partially depend on the movement of the other two Ubs. Therefore, one would expect the proximal Ub of $[Ub]_2$–^{48,63}Ub and the middle Ubs of Ub–^{63}Ub–^{48}Ub and Ub–^{48}Ub–^{63}Ub to have the largest T_1 values, while the K63 linked Ubs would have the smallest T_1 values.

Indeed, in $[Ub]_2$–^{48,63}Ub the proximal Ub has the longest T_1, whereas of the two distal (singly-linked) Ubs, the distal K63-linked Ub shows the shortest T_1 (i.e., the least restricted mobility), and the distal K48-linked Ub has longer T_1, consistent

with its contact with the proximal Ub (**Figure 4.17 A**). The T_1 values for each Ub in $[Ub]_2$–^{48,63}Ub fully agree with the CSP analysis and structural models (**section 4.7.4**).

Just as the CSPs followed a predictable pattern, the T_1 data for each Ub in the two unbranched chains followed a pattern. Expectedly, the middle Ubs in Ub–^{63}Ub–^{48}Ub and Ub–^{48}Ub–^{63}Ub showed the greatest T_1 (most restricted mobility). While the K63-linked Ubs, the distal Ub in Ub–^{63}Ub–^{48}Ub and the proximal Ub in Ub–^{48}Ub–^{63}Ub are consistently the most mobile (i.e., smallest T_1) regardless of whether they are at the distal or the proximal end of the chain (**Figure 4.17 B,C**). The K48-linked Ubs show intermediate T_1 values, as its tumbling is slowed by the hydrophobic interaction with the middle Ub, again regardless if the K48 linked Ub is at the distal or proximal end. However, in all of these chains the two K48-linked units have slightly different T_1 values; thus, despite forming a hydrophobic contact, they are not rigidly locked. This is consistent with previous findings for Ub–^{48}Ub (*155*) and indicates dynamic opening and closing of the hydrophobic interface, which is critical for the ability of Ub-receptors to bind to this chain.

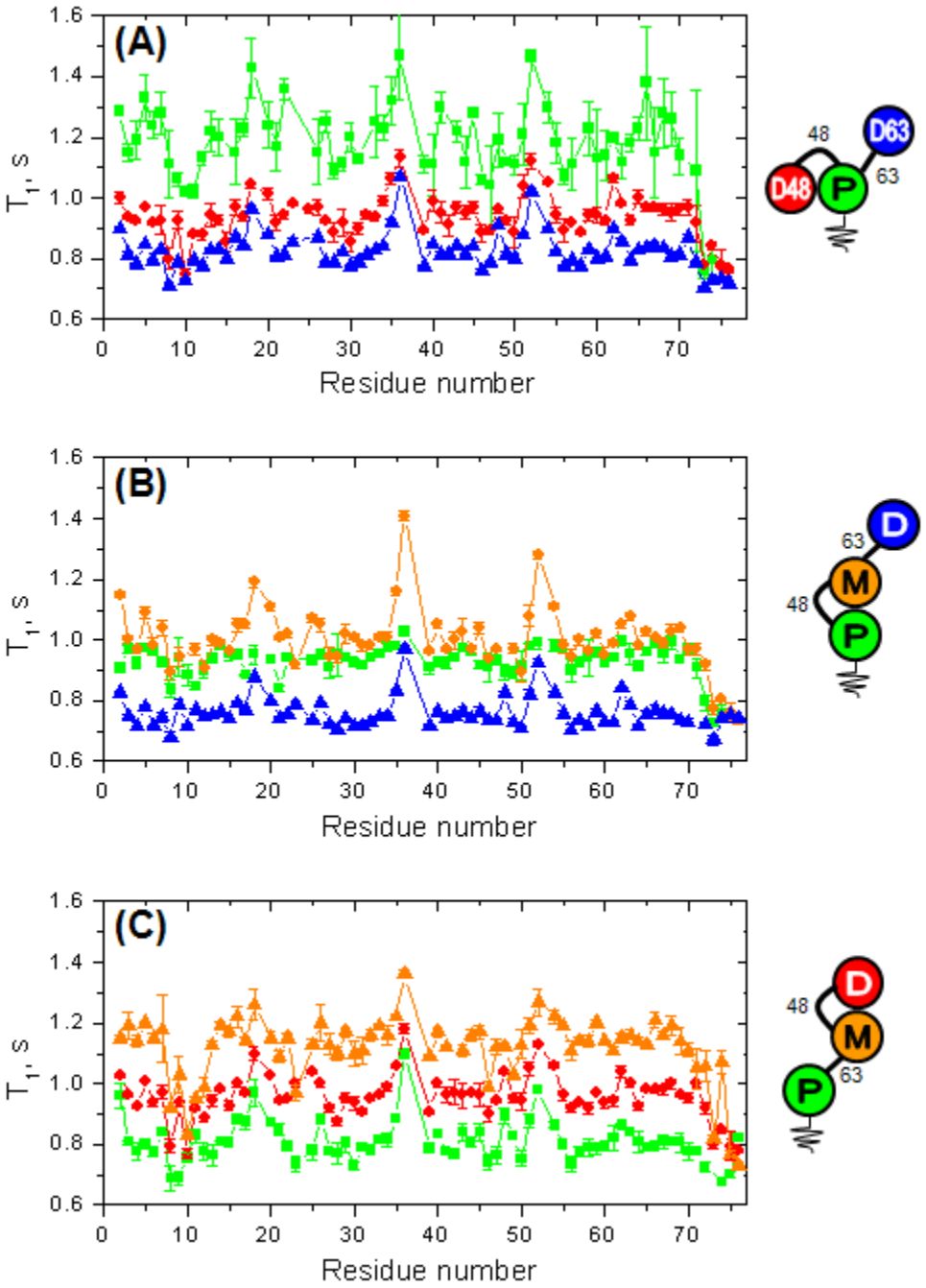


Figure 4.17 – Unit specific ^{15}N-T_1 relaxation in each mixed linkage trimer
(**A**) T_1 relaxation rates for each Ub in $[Ub]_2$–^{48,63}Ub with proximal Ub (green), distal-48 linked Ub (red), and distal-63 linked Ub (blue), (see pictograph, *right*). (**B**) T_1 relaxation rates for each Ub in Ub–^{63}Ub–^{48}Ub with proximal Ub (green), middle Ub (orange), and distal Ub (blue), (see pictograph, *right*). (**C**) T_1 relaxation rates for each Ub in Ub–^{48}Ub–^{63}Ub with proximal Ub (green), middle Ub (orange), and distal Ub (red), (see pictograph, *right*).

4.7.3 The classic pH dependent hydrophobic interface is present

One hallmark of K48-linked di-Ub is that its conformation is pH-dependent: lowering the pH shifts the equilibrium from the predominantly closed conformation at neutral pH to predominantly open. This effect is caused by a +1 net charge from protonation of the δ^1 nitrogen in the H68 side chain. At neutral pH above the pKa of H68, there is no proton on the δ^1 nitrogen and the net charge is 0. This simple acid base chemistry controls the K48 interface. Since the H68 side chains from the distal and proximal Ubs are in such close proximity with a net charge of 0 at neutral pH their hydrophobic interaction is strong, but at acidic pH the electrostatic repulsion from the +1 charges on each H68 heavily favors the open conformation, breaking the K48 interface. With our solution NMR method this manifests itself as nearly complete disappearance of the CSPs at pH 4.5 (*155*).

For the branched chain the CSP data (**Figure 4.18**) indicate the pH-dependent behavior of $[Ub]_2$–^{48,63}Ub is nearly identical to that reported for the corresponding Ubs in homogeneous K48- and K63-linked polyUb. This supports the conclusion that interdomain interactions in the branched tri-Ub are governed by the same forces as in the respective homogeneous-linkage chains. Thus the same K48 interface is maintained with linkage mixing.

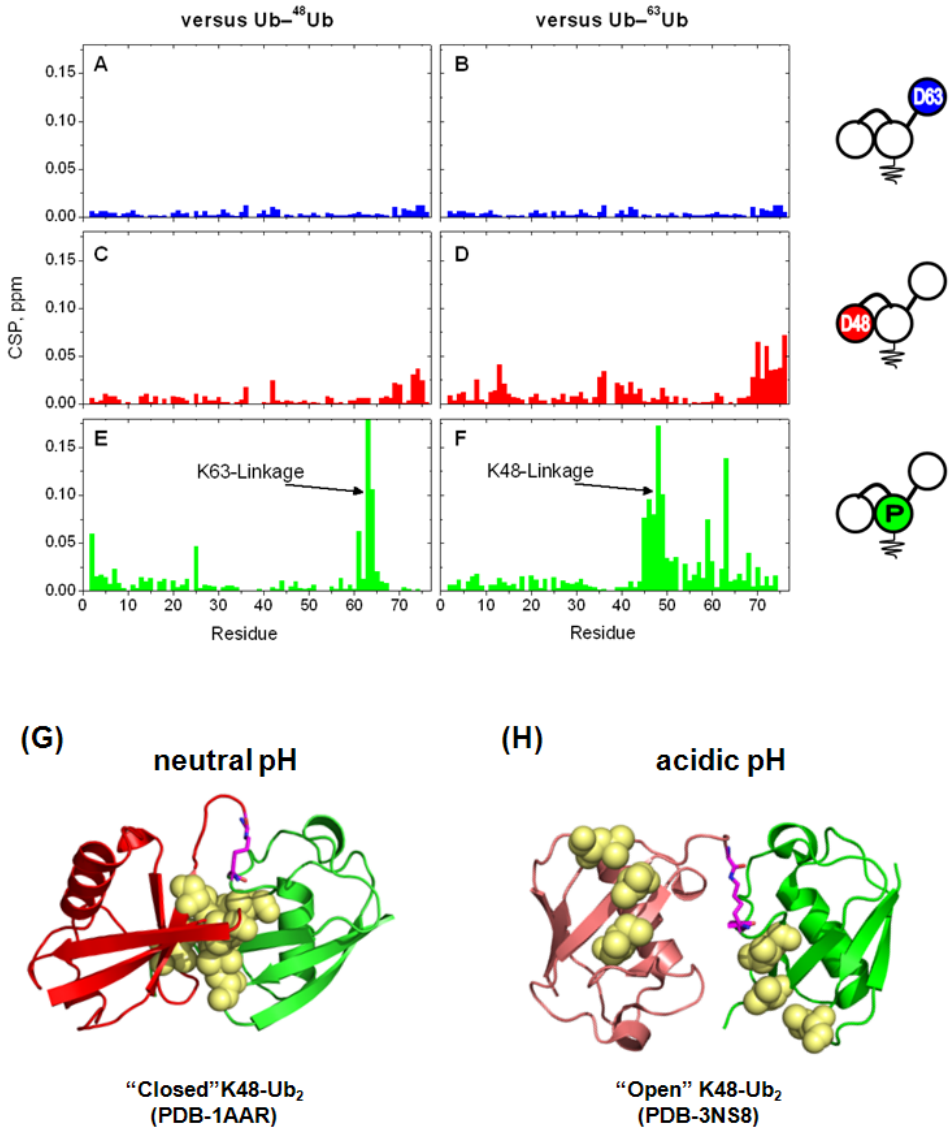


Figure 4.18 – CSP analysis at acidic pH reveals $[Ub]_2$–^{48,63}Ub retains the classical K48 interface.

CSPs between each Ub unit of $[Ub]_2$–^{48,63}Ub and its respective Ubs in Ub–^{48}Ub (left column) and Ub–^{63}Ub (right column) in 20 mM sodium acetate buffer, pH 4.5. (**A,B**) residue specific CSPs for the distal-63 Ub shows that it remains free of any interdomain contacts at pH 4.5. (**C**) Comparison of the distal-48 Ub in $[Ub]_2$–^{48,63}Ub and Ub–^{48}Ub are nearly identical indicating an open conformation. (**E,F**) CSP data shows the proximal Ub also adopts an open conformation at pH 4.5. There are some minor discrepancies in the CSPs attributable to the isopeptide bond that the Ub_2 controls lack, i.e. the large CSP observed in **(E)** around residue 63 is due to the K63 linkage present in the proximal Ub of $[Ub]_2$–^{48,63}Ub but absent in the Ub–^{48}Ub chain which it is being compared to. The other minor deviations in (**F**) CSPs match those reported for similar analysis of K48 and K63-linked polyUb. (**G**) neutral pH structure of Ub–^{48}Ub representing the closed conformation and (**H**) acidic pH structure of Ub–^{48}Ub representing the open conformation with no interdomain contacts.

4.7.4 Structural models reveal preserved homogenous features

With a consensus in the NMR CSP and ^{15}N T_1 relaxation data, structural models were generated for each of the three tri-Ubs studied using the biomolecular docking program HADDOCK, with site-specific CSPs to guide the docking. Although there were slightly different overall structures in the top clusters for these chains, all featured a distinct arrangement of the interdomain contacts in which the two K48-linked Ubs formed a well-defined contact mediated by their hydrophobic-patch residues whereas the K63-linked Ub was represented in essentially random positions. For example, the generated structures for $[Ub]_2$–^{48,63}Ub (**Figure 4.19 A**) clearly shows that the proximal and distal K48-linked Ubs form the "canonical" L8, I44,V70 hydrophobic interface, whereas the distal K63-linked Ub samples several different orientations, depending on the HADDOCK cluster (**Figure 4.19 D**). This variation in the K63-linked Ub is not unexpected given that (i) the CSP data show no close non-covalent interactions involving this Ub unit and (ii) the ^{15}N T_1 relaxation data indicate that this unit has greater mobility compared to the other two K48-linked Ubs. Moreover, the two K48-linked Ubs overlay quite well (RMSD = 1.72 Å) with the crystal structure (PDB: 1AAR) of Ub–^{48}Ub. For the unbranched mixed linkage chains Ub–^{63}Ub–^{48}Ub and Ub–^{48}Ub–^{63}Ub the structural models followed the same predictable pattern. The two K48 linked Ubs, middle and proximal in Ub–^{63}Ub–^{48}Ub and distal and middle in Ub–^{48}Ub–^{63}Ub created an interface, while the K63 linked Ub was represented in a variety of positions just as the ^{15}N T_1 relaxation data suggested (**Figure 4.19 B,C**). The overall structural models for all three chains are almost superimposable with their two K48 linked Ubs creating an interface and the K63

linked Ub with the exception of the linkage pattern. Taken together these structural models suggest that the same outcome is reached regardless of the order of linkages and branching.

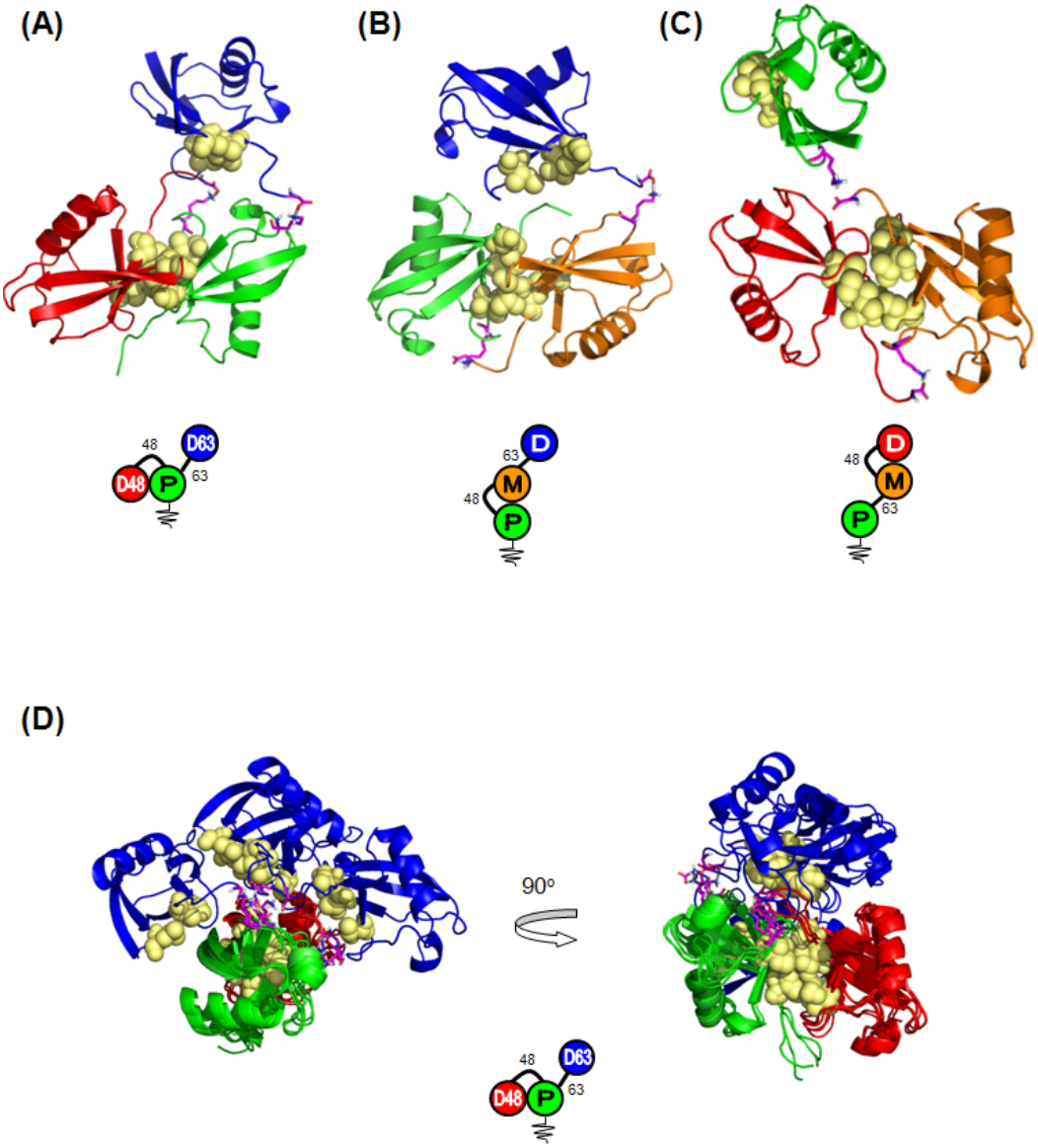


Figure 4.19 – Structural models of mixed linkage tri-Ub chains.
HADDOCK generated models using CSP contacts, (*see chain schematics below*), the L8,I44,V70 hydrophobic patch shown as yellow spheres (**A**) $[Ub]_2$–^{48,63}Ub with proximal Ub (green), distal-48 Ub (red), and distal-63 Ub (blue) (**B**) Ub–^{63}Ub–^{48}Ub with proximal Ub(green), middle Ub (orange), and distal Ub (blue) (**C**) Ub–^{48}Ub–^{63}Ub with proximal Ub (green), middle Ub (orange), and distal Ub (red). (**D**) multiple cluster alignment of $[Ub]_2$–^{48,63}Ub showing the same K48 interface (green and red), but a variation in the position of the distal-63 Ub (blue).

4.8 Signaling properties of component linkages are preserved in K48, K63 mixed linkage chains

4.8.1 K48 linkage selectivity of UBA(2) in $[Ub]_2$–^{48,63}Ub

Having established that the branched and unbranched mixed-linkage chains combine the structural properties characteristic of both linkages, I set to examine whether these chains can be recognized by linkage-selective receptors. To determine if the K48-linkage in $[Ub]_2$–^{48,63}Ub retains its receptor selectivity, separate titrations of $[Ub]_2$–^{48,63}Ub (for each Ub unit ^{15}N-enriched; see **Figure 4.9**) with the K48-linkage selective receptor UBA(2) from the proteasomal shuttle protein hHR23A were carried out. UBA(2) binds Ub–^{48}Ub in a "sandwich" mode and significantly more tightly (K_d=8 ± 7 μM) than to Ub–^{63}Ub (K_d =180 ± 80 μM) or monomeric Ub (K_d = 400 ± 100 μM) (*153, 208*). Strong residue-specific CSPs in both the proximal and the distal K48-linked Ubs were observed (**Figure 4.20 C-F**); these perturbations center around the hydrophobic-patch residues and the isopeptide linkage, consistent with UBA(2) insertion into the hydrophobic pocket formed by the K48-linked Ubs. The extent of binding can be gauged from titration curves of residue-specific CSPs or by comparing the CSPs for all residues at various points in titration. The CSPs in the proximal and distal K48-linked Ubs of $[Ub]_2$–^{48,63}Ub show little change between [UBA(2)]:[$[Ub]_2$–^{48,63}Ub] = 1 and 3 **(Figure 4.20 C-F)**, suggesting that UBA(2) binding to both Ubs reached saturation at approximately 1:1 molar ratio. Additional changes occurring in these Ubs above the 1:1 molar ratio could reflect binding of a second UBA(2) molecule, as observed for Ub–^{48}Ub. Interestingly, the distal K63-linked Ub exhibited an entirely different behavior and showed spectral perturbations only after the distal K48-linked and proximal Ubs were saturated with UBA(2). Even

at [UBA(2)]:[[Ub]$_2$–^{48,63}Ub] = 1, there are virtually no CSPs in the distal K63-linked Ub, and it was not until after this point that we finally observed UBA(2) binding to this Ub (**Figure 4.20 G,H**).

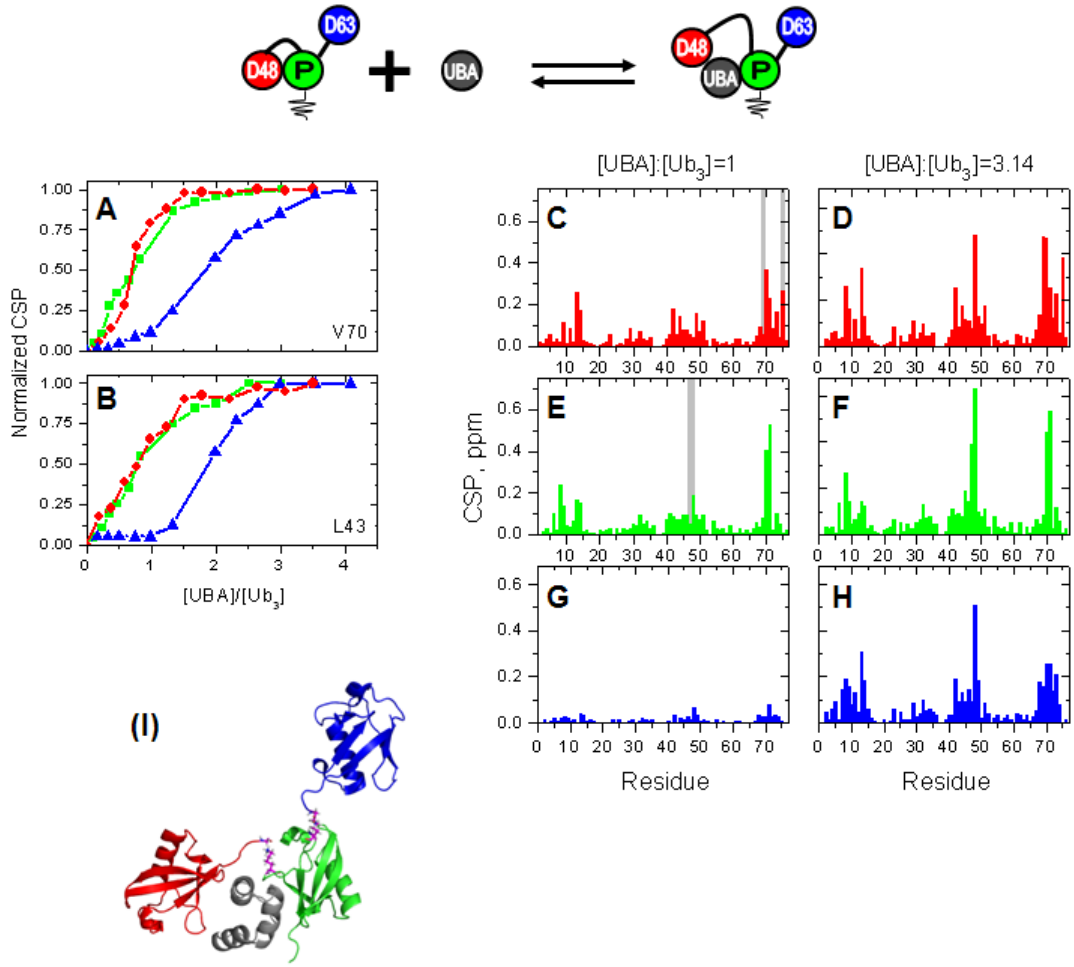


Figure 4.20 – UBA(2) K48 linkage selectivity in [Ub]–^{48,63}Ub by NMR titration. Each Ub in [Ub]$_2$–^{48,63}Ub was individually titrated with UBA(2). The results from each domain are color coded according to the pictograph on top: proximal green, distal-48 red, and distal-63 blue. Normalized residue specific titration curves for (**A**) V70 and (**B**) L43. Notice the lag phase for the distal-63 Ub (blue) in each. (**C-H**) Residue-specific CSPs in individual Ub units at two points in titration with UBA(2): at the 1:1 molar ratio (left column) and at saturation (right column). Gray bars indicate residues with signals broadened beyond detection at the 1:1 molar ratio. Note that the distal K63-linked Ub exhibits virtually no binding until [UBA(2)]:[[Ub]$_2$–^{48,63}Ub] > 1. (**I**) A model of the [Ub]$_2$–^{48,63}Ub/UBA(2) complex at the 1:1 molar ratio, based on the observed CSPs and the structure of the Ub–^{48}Ub/UBA(2) complex (PDB-1ZO6)

The titration curves for individual residues illustrate this trend from another angle by showing "standard" binding behavior for the distal K48-linked and proximal Ubs, but a "lag phase" for the distal K63-linked Ub (**Figure 4.20 A,B)**. This strong binding preference for K48-linked Ubs is also supported by our relaxation data that showed a larger increase in ^{15}N T_1 of the distal K48-linked Ub compared to distal K63-linked Ub at saturation (**Figure 4.21**). This difference in T_1 relaxation rates between the UBA(2) bound forms of distal-48 and distal-63 is attributable to tighter binding of the distal-48 Ub. These results demonstrate that a K48-linkage receptor can selectively recognize this linkage in a branched polyUb chain using the same binding mechanism as for homogeneous-linkage chains.

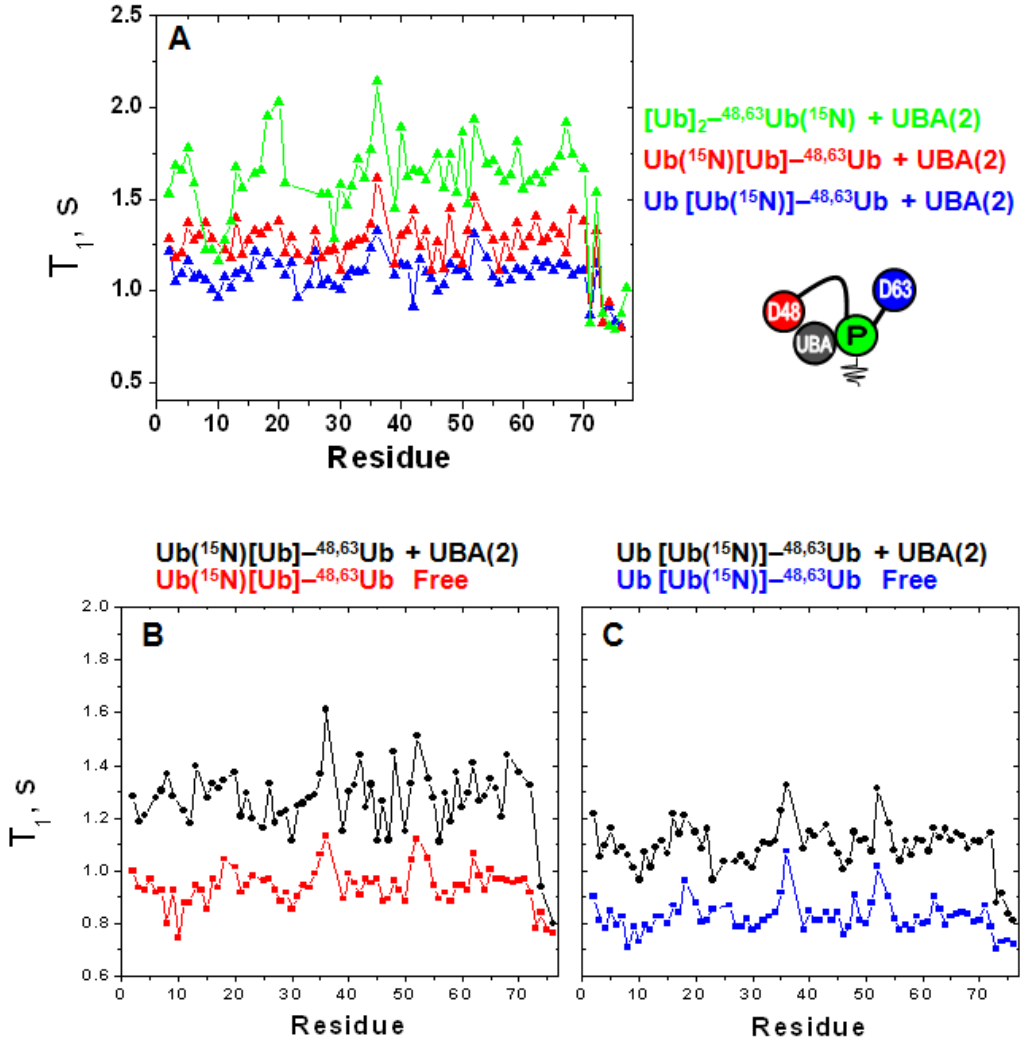


Figure 4.21 – ^{15}N T1 relaxation analysis of the UBA(2):[Ub]–^{48,63}Ub complex. T_1 relaxation rates were measured for $[Ub]_2$–^{48,63}Ub in complex with UBA(2) at saturation. (**A**) T_1 data for all Ubs in $[Ub]_2$–^{48,63}Ub show a systematic increase due to the increased size of the complex. (**B**) The distal-48 Ub in $[Ub]_2$–^{48,63}Ub has much lower T_1 values in the unbound form (red lines) vs. the UBA(2) bound form (black lines). (**C**) The distal-63 Ub in $[Ub]_2$–^{48,63}Ub shows a slight increase in T_1 relaxation rates between the unbound form (blue lines) and the UBA(2) bound form (black lines). This indicates that the distl-63 domain is still less restricted than any Ub in $[Ub]_2$–^{48,63}Ub due to its weaker binding.

4.8.2 K63 linkage selectivity of Rap80 in $[Ub]_2$–^{48,63}Ub

Rap80 contains a tandem Ubiquitin Interacting Motif (tUIM) that shows a strong binding preference for Ub–^{63}Ub versus Ub–^{48}Ub (*135*). The tUIM contains a flexible linker region between its two helical UIMs which adopts an alpha-helical conformation upon binding to Ub–^{63}Ub and perfectly aligns the two UIMs' interacting surfaces for binding to adjacent Ubs in a K63 linkage (*135, 209*). To examine how the tUIM interacts with branched chains, $[Ub]_2$–^{48,63}Ub, as well as Ub–^{48}Ub and Ub–^{63}Ub as controls were titrated separately using several NMR methods to monitor the interactions for each ^{15}N-enriched Ub unit. As shown in (**Figure 4.22**) signals for all three Ubs in $[Ub]_2$–^{48,63}Ub were altered from the tUIM binding interaction. The observed perturbations were site-specific and centered around the hydrophobic-patch residues as well as the Ub C-termini. The distal K48-linked Ub showed noticeably smaller CSPs at saturation (suggesting weaker binding) compared to the other two Ubs. Importantly, during these titrations some residues in the proximal and distal K63-linked Ubs, but not in the distal K48-linked Ub, showed strong attenuations or even disappearance of the NMR signals. This is caused by slow exchange (*i.e.*, slow off-rates) on the NMR chemical shift time scale and is indicative of tighter binding to the K63-linked Ubs. Comparison of the domain-specific CSPs in $[Ub]_2$–^{48,63}Ub with those in the control di-Ub samples (**Figure 4.22**), as well as comparison of the signal trajectories upon titration (**Figure 4.23**), led me to the following conclusions. First, the distal K63-linked and proximal Ubs of $[Ub]_2$–^{48,63}Ub bind the tUIM in the same mode as their respective units in Ub–^{63}Ub (**Figure 4.22 D,E**). This is supported by the fact that at saturation with the tUIM virtually all

signals in the $^{1}H,^{15}N$-HSQC NMR spectra of the distal K63-linked and proximal Ubs of $[Ub]_2$–^{48,63}Ub overlay perfectly with the corresponding signals of Ub–^{63}Ub. Also, the signal trajectories (in the course of the titration) for the same residue in the K63-linked components of the branched and control chains were nearly superimposable (**Figure 4.23**), consistent with identical binding interactions. Second, there is a stark difference between tUIM binding to the distal K48-linked Ub of $[Ub]_2$–^{48,63}Ub and the Ubs of Ub–^{48}Ub. Indeed, our data show spectral differences (ΔCSP) at saturation between the distal K48-linked Ub and either Ub of Ub–^{48}Ub (and even mono-Ub) that are much larger than the differences observed for the K63-linked components (**Figure 4.22**).

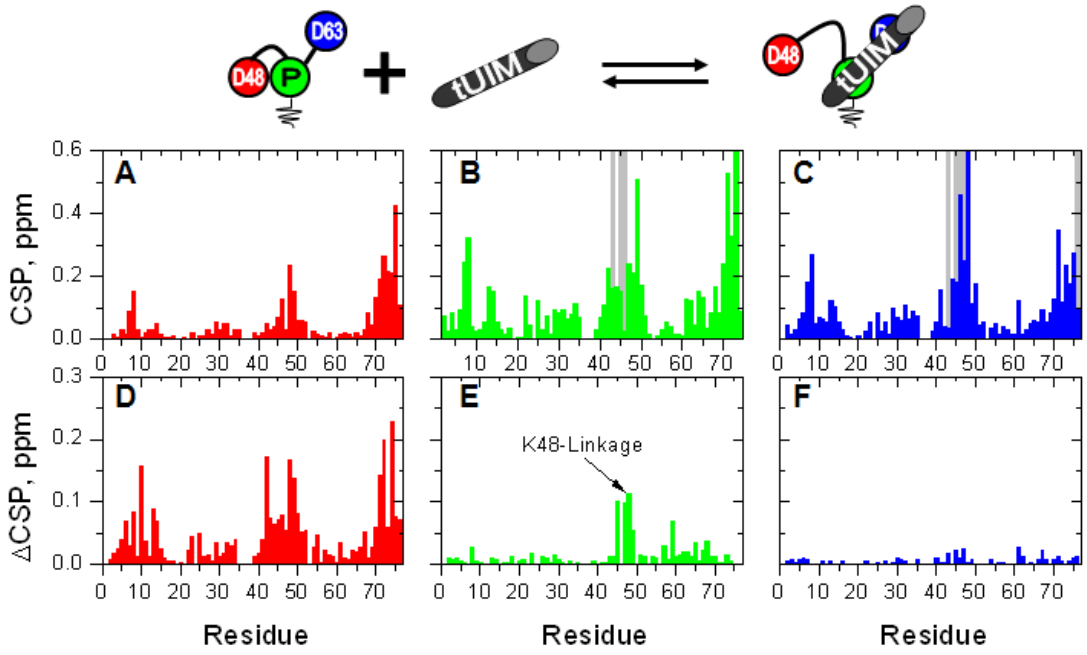


Figure 4.22 – CSP and ΔCSP analysis of Rap80 titration of $[Ub]_2$–^{48,63}Ub. All data are color coded as indicated in the pictograph on the top: proximal Ub data are shown in green, distal K48-linked Ub in red, and distal K63-linked Ub in blue. (**A-C**) CSP plots for each Ub unit in $[Ub]_2$–^{48,63}Ub at saturation: (**A**) distal K48-linked Ub at [tUIM]:[$[Ub]_2$–^{48,63}Ub] = 2.43, (**B**) distal K63-linked Ub at [tUIM]:[$[Ub]_2$–^{48,63}Ub] = 1.52, and (**C**) proximal Ub at [tUIM]:[$[Ub]_2$–^{48,63}Ub] = 2.00. Residues that showed strong signal attenuation during the titration are represented with gray bars. (**D-E**) Difference in the CSPs (ΔCSP) for each Ub of $[Ub]_2$–^{48,63}Ub and the corresponding Ub in the respective di-Ub controls at saturation with tUIM: (**D**) distal K48-linked Ub of $[Ub]_2$–^{48,63}Ub vs. distal Ub of Ub–^{48}Ub, (**E**) proximal Ub of $[Ub]_2$–^{48,63}Ub vs. proximal Ub of Ub–^{63}Ub, and (**F**) distal K63-linked of $[Ub]_2$–^{48,63}Ub vs. distal Ub of Ub–^{63}Ub. In complex with tUIM, the distal K63-linked Ub shows virtually no difference from Ub–^{63}Ub, indicating a similar bound state. A similar conclusion holds for the proximal Ub, except for the CSP differences around K48 resulting from the attachment of the distal K48-linked Ub. The distal K48-linked Ub in the branched trimer compared to Ub–^{48}Ub shows large systematic ΔCSPs which suggests that this Ub does not bind the tUIM as its counterpart in Ub–^{48}Ub.

When the tUIM binds to Ub–^{48}Ub it has to disrupt the hydrophobic interface between the two Ubs in order to access the Ub hydrophobic surface. Even though avid tUIM binding to K48-linked di-Ub has not been observed, the reported interaction (K_d ~ 160 μM) was still relatively strong (*135*). In the branched chain,

simultaneous tUIM binding to the proximal and distal K63-linked Ubs should displace the distal K48-linked Ub from its hydrophobic contact with the proximal Ub, thus freeing it and making it available for binding to excess tUIM molecules. To verify this model, the ^{15}N T_1 for each Ub of $[Ub]_2$–$^{48,63}Ub$ at saturation with the tUIM was measured (**Figure 4.23**) and compared them with those of unbound $[Ub]_2$–$^{48,63}Ub$ (**Figure 4.17 A**). The tUIM binding significantly increased the T_1 values (reflecting slower tumbling) for both the distal K63-linked and the proximal Ub but caused almost no change in the T_1 values of the distal K48-linked Ub, rendering it the lowest-T_1 unit in the chain (**Figure 4.23 A**). This confirms that the Rap80 tUIM binds simultaneously to both the proximal and (formerly "free" and the most mobile) distal K63-linked Ubs, thereby restricting their conformational freedom. At the same time, the tUIM binding displaced and freed the distal K48-linked Ub from its contact with the proximal Ub, making it the most mobile (shortest T_1 value) Ub in the chain. Thus, I conclude that the tUIM-induced perturbations detected in the distal K48-linked Ub are primarily due to this secondary effect rather than direct interaction with the tUIM. However, we cannot completely rule out that some small fraction of this distal Ub interacts with the excess Rap80 tUIM present at the saturation conditions.

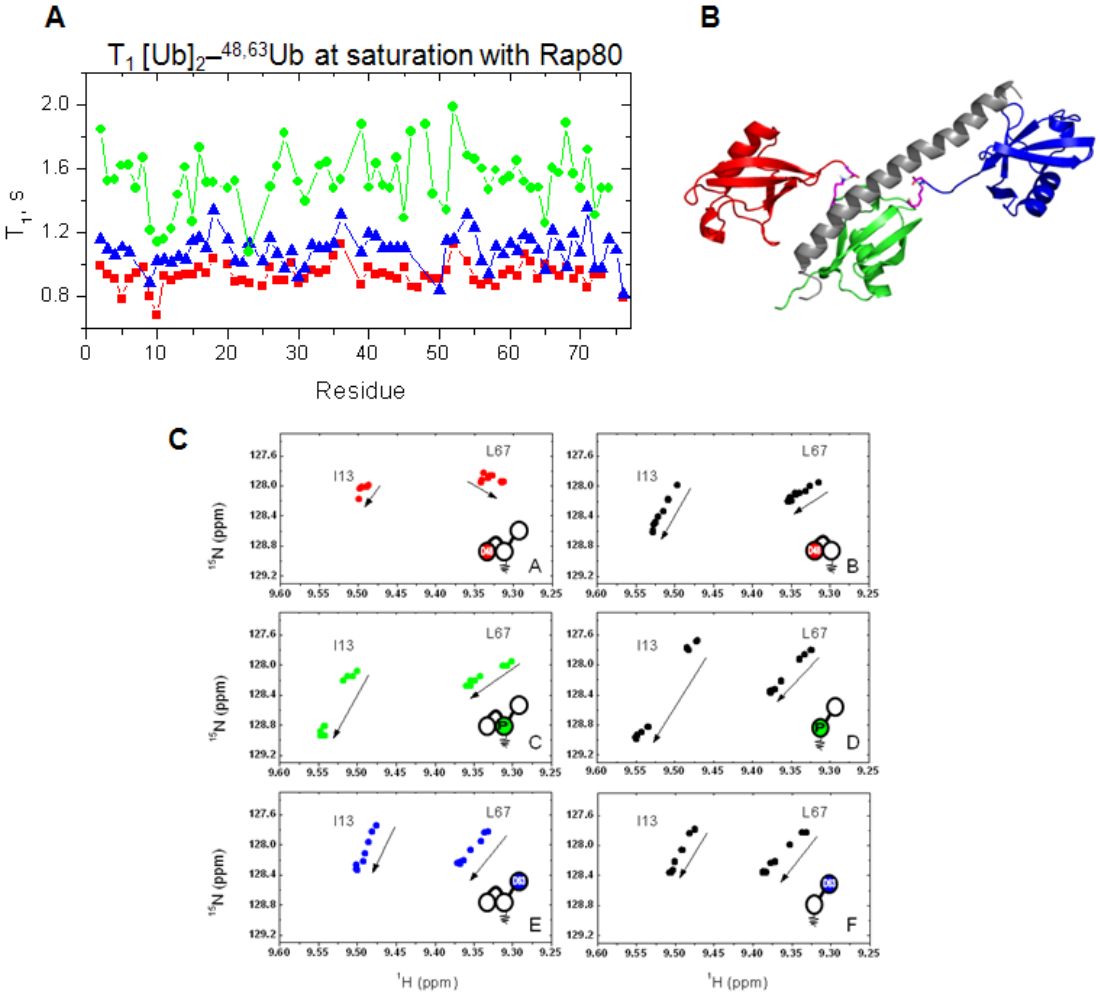


Figure 4.23 – Additional binding characterization of tUIM and $[Ub]_2$–^{48,63}Ub. All data are color coded as indicated in (**B**): proximal Ub data are shown in green, distal K48-linked Ub in red, and distal K63-linked Ub in blue. (**A**) ^{15}N T_1 relaxation rates for proximal, distal K48-linked, and distal K63-linked Ubs in $[Ub]_2$–^{48,63}Ub at saturation with Rap80 tUIM. Binding of the tUIM significantly increased T_1 for distal K63-linked and proximal Ubs of $[Ub]_2$–^{48,63}Ub, but had little effect on the T_1 of distal K48-linked Ub although the molar ratio of tUIM to the (^{15}N) distal-48 sample was the largest. (**B**) A model of the $[Ub]_2$–^{48,63}Ub/tUIM complex at 1:1 molar ratio, based on the observed CSPs and ^{15}N T_1 values, and the Ub–^{63}Ub/tUIM complex structure (PDB: 3A1Q). (**C**)Residue specific ^{1}H,^{15}N-HSQC trajectories show that the distal-63 and proximal Ubs undergo a similar binding interaction whether they are in a homogeneous (right column) or branched (left column) chain. The trajectories for the distal-48 domain show no similarities indicating that the interaction of the distal-48 Ub in $[Ub]_2$–^{48,63}Ub deviates from its homogeneous chain.

4.8.3 Linkage selectivity of DUBs

Given that DUBs are critical for maintaining the pool of free Ub and regulating conjugate levels, linkage-selective DUBs were tested to determine if they would still retain their ability to recognize and cleave their cognate Ub–Ub linkage in the branched and unbranched mixed-linkage chains. For this purpose OTUB1, which specifically cleaves K48 linkages (*210*), and AMSH, which is specific for K63 linkages (*64*) were selected. As evident in (**Figure 4.24**), OTUB1 readily reduced $[Ub]_2$–^{48,63}Ub, Ub–^{63}Ub–^{48}Ub, and Ub–^{48}Ub–^{63}Ub to di- and mono-Ub. However, unlike the case with the homogeneously linked Ub–^{48}Ub–^{48}Ub substrate, OTUB1 alone could not further process the di-Ub species generated from any of the mixed-linkage chains (**Figure 4.24**). Only when both OTUB1 and AMSH were present was complete chain disassembly achieved, indicating that the uncleaved di-Ub was linked via K63. An essentially identical behavior was observed for AMSH, which readily cleaved the K63-linkage in all tri-Ub chains while leaving the K48-linkage intact in the three mixed linkage chains (**Figure 4.24**). Taken together, these results demonstrate that the K48 and K63 linkages in branched and unbranched mixed-linkage polyUb can each individually be processed by linkage-selective DUBs. Recently, a similar finding using OTUB1 and AMSH was also reached with unnatural K48 and K63 mixed linkage chain analogs formed via a Nε-Gly-L-homothiaLys mimic (which is one C–C bond longer and contains a sulfur atom instead of carbon) of an isopeptide bond (*211*). This work demonstrates linkage selectivity for fully natural isopeptide bonds and both unbranched and branched mixed linkages.

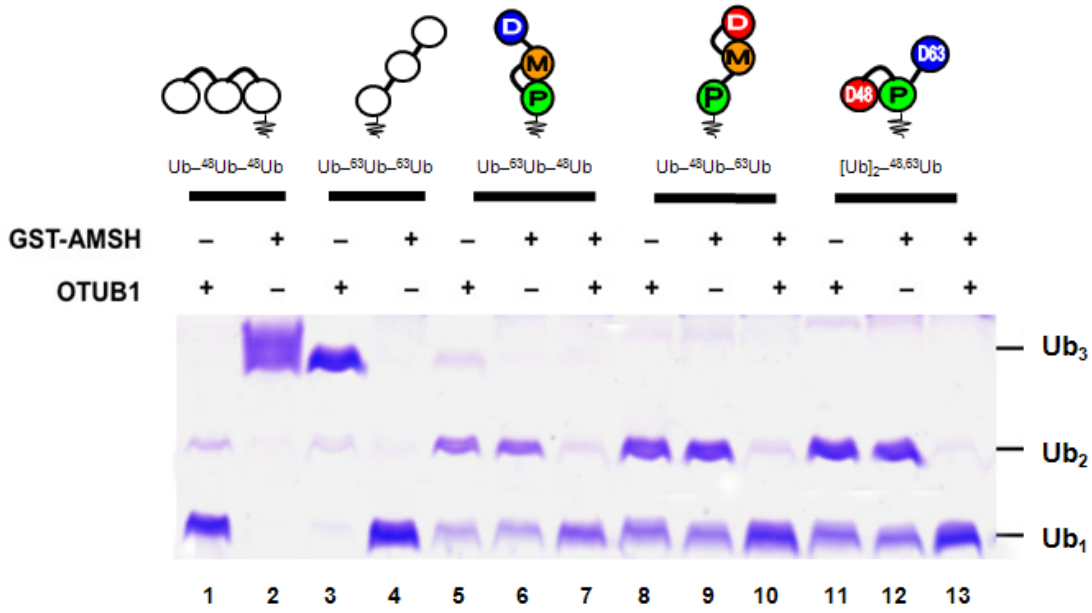


Figure 4.24 – Linkage selective DUBs efficiently cleave their preferred linkages in mixed linkage chains.
Linkage-selective DUBs cleave their cognate Ub–Ub linkages in homogeneous or mixed-linkage polyUb. The indicated chains were incubated with OTUB1, GST-AMSH, or both and the products evaluated by 15% SDS-PAGE followed by Coomassie staining. Lanes 1-4 show reaction products of homogeneous Ub–^{48}Ub–^{48}Ub and Ub–^{63}Ub–^{63}Ub. Lanes 5 and 6, 8 and 9, and 11 and 12 show that alone each DUB can only process one linkage in the unbranched or branched mixed-linkage chains. Only with both DUBs present were the mixed-linkage chains hydrolyzed fully to monomeric Ub (lanes 7, 10 and 13).
* Assay performed by Prof. Robert E. Cohen

4.8.4 Linkage selective antibodies and proteasomal processing

The use of linkage specific antibodies led to the hypothesis that mixed linkage polyUb chains were formed *in vivo* mainly due to the observation that specific gel bands would blot with more than one linkage specific antibody. However even with analysis by tryptic digest, it was not clear if the linkages were contained in the same chain, and assuming that the linkages were in the same chain it was undetermined if they were of branched or unbranched topology. Given that I could create excellent model chains with a controlled *in vitro* method, I found it absolutely necessary to test

if the K48 and K63 linkage specific antibodies could detect their respective linkages in my system. Antibodies were provided as a kind gift to Prof. Michael H. Glickman from Genentech. First it was established that both the K48 and K63 linkage-selective antibodies recognized their respective Ub–Ub linkages in $[Ub]_2$–^{48,63}Ub (**Figure 4.25**). This result in itself clearly demonstrates that chain branching preserves both Ub-specific and linkage-specific epitopes for antibody recognition, and also shows that linkage-specific antibodies can be used to probe if a particular polyUb chain contains multiple linkage types.

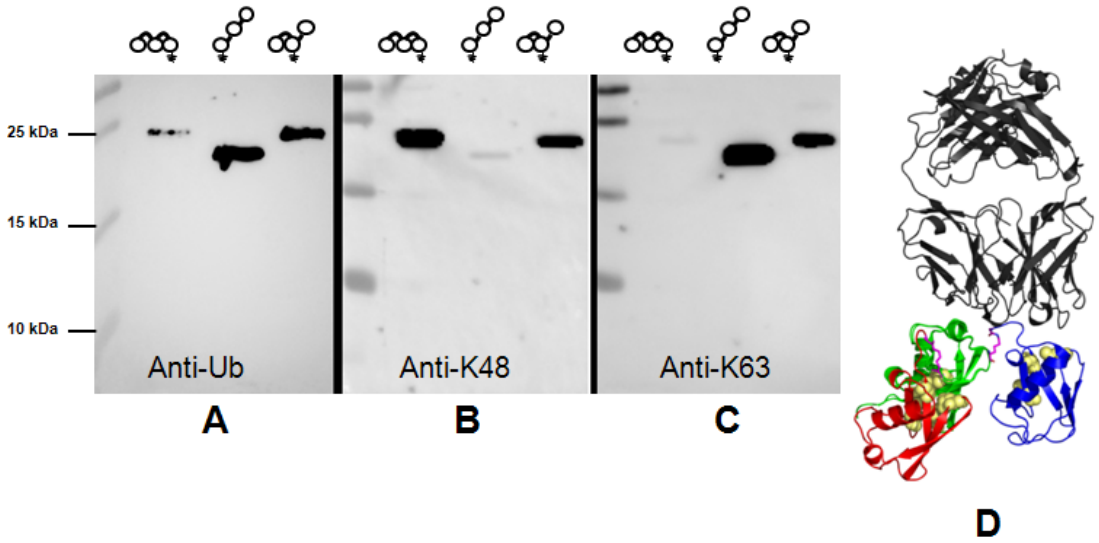


Figure 4.25 – Linkage selective antibodies detect $[Ub]_2$–^{48,63}Ub.
Western blots of controls Ub–^{48}Ub–^{48}Ub, Ub–^{63}Ub–^{63}Ub and the branched $[Ub]_2$–^{48,63}Ub. (**A**) The non-linkage specific anti-Ub detects all trimers. (**B**) anti-K48 detects only Ub–^{48}Ub–^{48}Ub and $[Ub]_2$–^{48,63}Ub, and (**C**) anti-K63 detects Ub–^{63}Ub–^{63}Ub and $[Ub]_2$–^{48,63}Ub. The controls work as expected and $[Ub]_2$–^{48,63}Ub is the only trimer to blot against both linkage specific antibodies. (**D**) Model of the K63 linkage specific antibody (gray) binding $[Ub]_2$–^{48,63}Ub across the proximal Ub (green) and distl-63 Ub (blue). The distal-48 Ub is far removed from the interaction and does not interfere with recognition. However, in highly branched (5D) chains it could be possible to disrupt antibody recognition.
* Western blots performed by Dr. Nurit Livnat-Levanon

One of the most studied outcomes of Ub modification is targeting to the proteasome. In a poorly understood process, proteasome-associated DUBs mainly function to modify or remove the polyUb tag from conjugated substrates (*212, 213*). The three major DUBs associated with the proteasomes are Rpn11 (POH1 in humans), Ubp6 (USP14 in humans), and Uch37 (only in humans). Whether and how branched polyUb chains are processed by the proteasome is an open question.

A series of studies has shown that the so-called "forked" polyUb chains that are branched on two closely positioned lysine residues (*i.e.,* K6 & K11, K27 & K29, or K29 & K33) are degraded slowly by purified proteasomes and inhibit proteasomal DUB activity and substrate degradation (*20, 26, 28*). Whether this observation holds for other chains such as those containing the more abundant K48 and K63 linkages was not reported. To evaluate the efficiency of proteasome processing of Ub–^{48}Ub–^{48}Ub, Ub–^{63}Ub–^{63}Ub, or [Ub]$_2$–^{48,63}Ub signals, these tri-Ubs were incubated with purified yeast 26S proteasomes and the cleavage was assessed by blotting with anti-Ub and linkage-selective antibodies.

Processing of the homogenous Ub–^{63}Ub–^{63}Ub chains was faster than that of Ub–^{48}Ub–^{48}Ub (**Figure 4.26**); this result is consistent with previous reports (*92, 214*). Accordingly, the branched [Ub]$_2$–^{48,63}Ub chain was disassembled faster than Ub–^{48}Ub–^{48}Ub, and generation of di-Ub (by removal of the single K63-linked Ub) was as efficient as from the Ub–^{63}Ub–^{63}Ub chain. These observations suggest that (i) the branching did not hinder proteasomal cleavage, and (ii) the presence of a K63 linkage in this chain facilitates its conversion to di-Ub. The latter was verified using a K63 linkage-specific antibody, which revealed that only a minor fraction of the di-Ub

products generated from $[Ub]_2$–^{48,63}Ub were K63-linked (**Figure 4.26**). Taken together, these results demonstrate that the 26S proteasome can systematically process K48 and K63 mixed linkages polyUb chains.

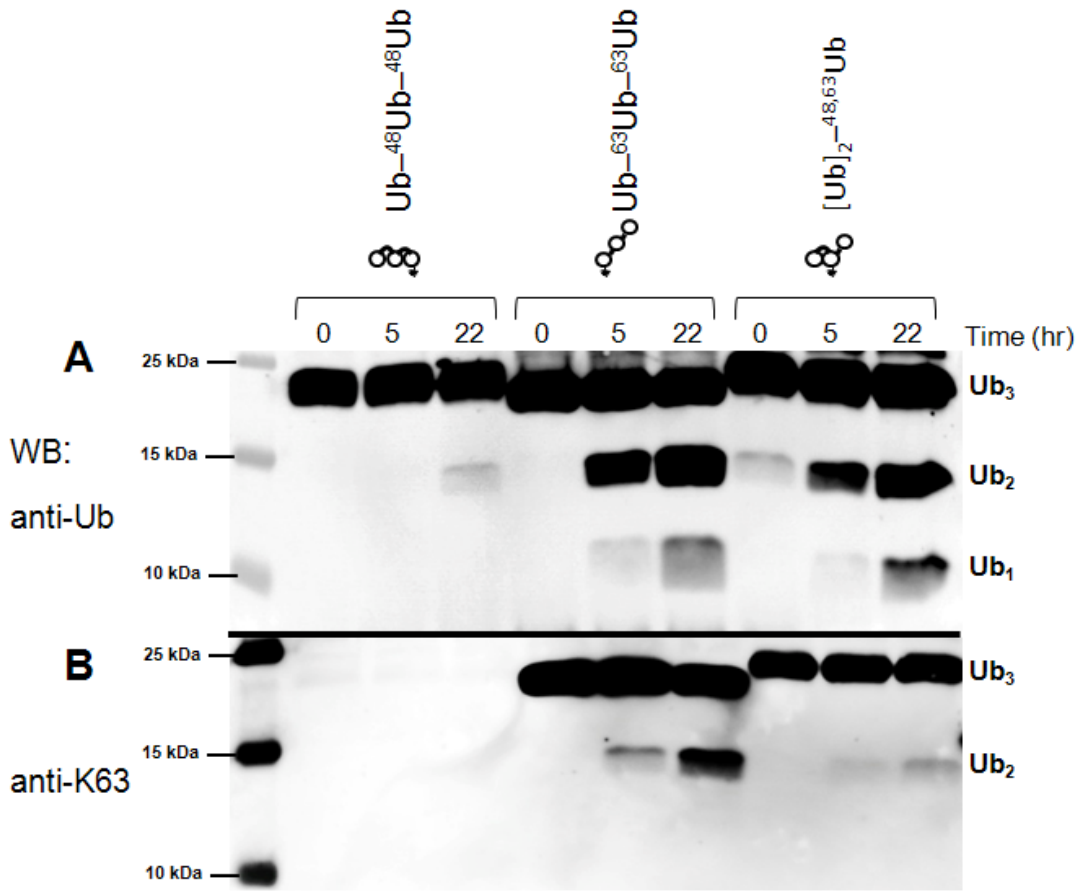


Figure 4.26 – Proteasomal disassembly of $[Ub]_2$–^{48,63}Ub.
Time course of disassembly for homogeneous and branched chains by 26S proteasomes. Ub–^{48}Ub–^{48}Ub, Ub–^{63}Ub–^{63}Ub, or $[Ub]_2$–^{48,63}Ub were incubated with purified yeast 26S proteasomes and the digests were monitored at 0, 5, and 22 hrs by SDS-urea PAGE and blotting with (**A**) anti-Ub and (**B**) anti-K63 antibodies. The trace di-Ub band for $[Ub]_2$–^{48,63}Ub in right lanes of (**B**) indicate that the K63 linkage is cleaved much faster than the K48 linkage.
* Assay performed by Dr. Nurit Livnat-Levanon

4.9 Discussion and Summary

4.9.1 Theoretical complex chains

As outlined, there are theoretically quadrillions of unique polyUb chains if linkage mixing is allowed. Attempting to study all of these chains individually would be both impractical and produce data that is meaningless to the scientific community due to the overwhelming amount of information. It is sterically possible for a single Ub to be simultaneously ligated at all eight linkages sites, but there is much to be learned from less complex systems and there currently is not a clear approach for the study of such chains. For my theoretical study, I focused on a very simple system of just three branched Ubs containing just two linkages. The four distinct conformations that could result from such chains (i) "free form" (ii) "distal hug" (iii) "clump" and (iv) "classic interface" are simple to define and give us meaningful insight on the structural properties of such chains. For example, the classic interface conformation supports that individual linkage properties are preserved and the chain would be expected to act similar to its homogenous components, while a distal hug suggests that branching allows the chain to present a new signal for receptors. These basic categories also help to classify the vast amount of complex chains and also simplify the field from a vast sea of individual complex chains. With improved computational methods for docking and analysis, it would be logical to investigate more chains and there certainly is no shortage of them to study. The predictions for the branched chains here should also be tested experimentally and it is highly possible a significant percentage of the predicted structures are valid. The failure to experimentally detect the distal hug in $[Ub]_2$–^{11,63}Ub was disappointing, but this does not rule out the

possibility of the distal hug conformation in other branched tri-Ubs. Furthermore, the distal hug conformation may in fact be over played considering a receptor carrying two different linkage selective UBDs also provides a unique mechanism to recognize complex polyUb chains. In this case separate UBDs of the receptor recognizing two different linkages in close proximity on a mixed linkage chain would be the new signal, without the need for a novel conformation such as the distal hug. Ultimately, I would expect the quadrillions of theoretical polyUb chains to have much degeneracy in their signaling properties, but future work could uncover particular combination of linkages that do present new signals for receptors. Much remains to be explored in this new field and my study represents a good first step.

4.9.2 K48 and K63 mixed linkage polyUb chains

Ubiquitinating machinery has been shown to be possessive in generating homogeneously-linked chains. I demonstrated here that, when acting together simultaneously or sequentially linkage-specific E2 enzymes (K11 Ube2s, K48 E2-25K, and K63 Ubc13:Mms2) can form both branched and unbranched mixed-linkage forms of polyUb. The extent of the branching is yet to be determined, but the example of $[Ub]_3$–^{11,48,63}Ub, indicates that exposure of Ub (or polyUb) to more than one E2 enzyme can result in highly branched chains carrying many linkage types. Thus, depending on the factors involved (which are currently unclear), the resulting polyUb landscape would include chains of unbranched and branched topologies. This raises the question of why such chains have not been observed more often. Possible reasons are that intracellular DUBs rapidly edit mixed-linkage chains, or that assembly of mixed-linkage chains largely is prevented by regulation (e.g., via

compartmentalization) of the various E2 and E3 enzymes in cells. However, because current methods are not suited for the detection of mixed-linkage chains, such chains may in fact be quite abundant but remain unseen.

This study revealed that mixing of K48 and K63 linkages in the same polyUb chain retains structural the features of each individual linkage type. The NMR data show unambiguously that the hydrophobic interface, characteristic of the K48-linkage is formed across the Ub–^{48}Ub unit in both the branched and unbranched mixed-linkage trimers, and that this interface exhibits the same pH dependence as the "classical" interface in homogeneous K48-linked polyUb. Furthermore, the ^{15}N T_1 relaxation data revealed that in the mixed-linkage chains the K63-linked Ub is highly mobile compared to the other two Ubs, which are restricted by their contacts in the K48-dimer.

A major finding was that highly selective receptors for each of the two linkage types bind the branched $[Ub]_2$–^{48,63}Ub chain in a linkage-specific manner; i.e., binding to the cognate Ub–Ub unit was essentially as if the other linkage was not present. It is noteworthy that the proximal Ub in $[Ub]_2$–^{48,63}Ub is highly adaptable and can bind in both K48 and K63 linkage-selective modes. Furthermore, chain branching also preserves both Ub-specific and linkage-specific epitopes for antibody recognition. Thus, the K48 and K63 linkages within the same polyUb chain can retain their characteristic signaling properties. This conclusion is further strengthened by the fact that linkage-selective DUBs can efficiently process their cognate linkages in both branched and unbranched mixed-linkage chains. The finding that the 26S proteasome recognizes and cleaves the branched chain suggests that *in vivo* polyUb

containing both K48 and K63 linkages can be disassembled by the proteasome essentially as a homogeneous chain. Together, these observations demonstrate that the K48 and K63 signaling properties can be encoded into the same polyUb via linkage mixing or branching, thus allowing the chain to carry two distinct signals.

Although this somewhat abstract study is the first attempt to investigate the structural properties of mixed linkage polyUb chains the mounting physiological data strongly supports roles for such chains and that this work paves the way into a new frontier of biology (*144*). Another possible mechanism explaining the formation of mixed linkage polyUb chains is that the polyUb signal is constantly subject to editing or remodeling, and some branched or other mixed-linkage chains could simply be accidents that, in analogy to mismatched bases in DNA or misfolded proteins, are eventually corrected by cellular machinery. Nonetheless, as I have demonstrated here, linkage mixing or branching could enhance the signaling capability of polyUb. That mixed-linkage chains can carry multiple recognition signals is an exciting concept opening a new perspective on Ub signaling.

Chapter 5: Decrypting ubistatin-ubiquitin interactions

5.1 Background and research aims

5.1.1 History and prior work for ubistatins

Ubistatin is a term given to class of naphthyl-based molecules introduced in 2004 (*215*). This stems from the observation that these molecules inhibited proteasomal degradation of polyubiquitinated substrates by directly binding the polyUb signal, preventing its recognition with receptors. Dyes such as Congo Red, trypan red, and suramin were predecessors to ubistatins and well before their association with Ub, ubistatins were being pursued as lead compounds to treat HIV and other life threatening viruses, see US Patent 5,681,832 (1997 Haugwitz *et al.*).

The interaction of ubistatins with Ub was discovered in 2004 when 109,113 compounds from the NCI open library were screened in an assay designed to find molecules that inhibited cell cycle progression (*215*). After the field of compounds was narrowed, several were found to block mitotic entry by disrupting either the UPP or APC/C activation. However, only two compounds, c59 and c92 were found to inhibit the destruction of polyubiquitinated substrates, but they did not interfere with proteasome function directly. Further testing revealed that both c59 and c92 could outcompete proteasomal polyUb receptor Rpn10 and the UBL/UBA shuttle Rad23 for polyUb chains (*215*). In the same study, solution data from NMR mapped CSPs to the L8,I44,V70 hydrophobic patches in polyUb establishing that c92 directly bound the polyUb signal (*215*). With the mechanism of ubistatins established, another study in 2008 used c92 to show that receptor/Ub interactions play a key role in the spliceosome assembly pathway and that factors such as Prp8 are polyubiquitinated

(*216*). The use of ubistatins to probe receptor/Ub interactions on other pathways aside from the UPP is an interesting concept, but has yet to become common practice.

5.1.2 Ubiquitin small molecule interactions and objectives

Given the success of small molecule proteasome inhibitors, which essentially produce the same outcome as ubistatins, there has been an interest in the further development of these molecules as potential therapeutics. Historically ubistatins are unique in that they were the first small molecules shown to bind Ub, however since Ub itself is a model protein, many others have characterized small molecule interactions with the protein. There has recently been an interest in the interaction of Ub with gold nanoparticles and there is experimental data to support specific interactions (*217, 218*), yet there is no structural data other than MD simulations. To further complicate matters the actual structure of the gold nanoparticles themselves is hard to determine. Studies by Whitesides and coworkers have shown that at certain concentrations, the anionic detergent SDS binds specific regions of Ub and they have also characterized the stoichiometry of several Ub/SDS complexes (*219, 220*). In the future others may discover more small molecules that interact with Ub, but at this present point in time ubistatins are the most promising therapeutics and the only small molecules shown to directly modulate defined cellular pathways.

To further advance ubistatins, the greatest obstacle to overcome is the lack of any structural information. The *in vivo* capabilities of ubistatins are well characterized and promising, but for development we will need to know exactly how these molecules interact with Ub and polyUb. On a special note, the original 2004 ubistatin study incorrectly reported the structure of c59 and it was not until late 2008

that this error was formally corrected (*215*). As a result those attempting to further investigate c59 were actually working with the wrong molecule. Thus my main goals will be to determine the structures of ubistatin/Ub complexes, their stoichiometry, and also determine if other ubistatin derivatives directly bind to Ub. My secondary objectives are focused on exploring new properties of ubistatins to determine if they have any other potential applications.

5.2 Physical properties of ubistatins

5.2.1Structural classification of ubistatins and derivatives

The original ubistatin molecules (c59 and c92) were highly conjugated symmetric ring systems with electronegative sulfonate substituent groups. Since 2004 many other ubistatins derivatives have been synthesized in an effort to improve affinity for Ub conjugates. With the recommendation of Prof. David Fushman ubistatins were also synthesized as half compounds containing only one of the ring systems, which allowed unambiguous NMR studies. Logically these compounds containing just one ring system are termed "half ubistatins" and the original ubistatins are referred to as "full ubistatins." The chemical structures of full ubistatins are presented in (**Figure 5.1**) and half ubistatins are presented in (**Figure 5.2**). All ubistatins compounds used for study were designed and synthesized by our collaborator Dr. Tim Lewis.

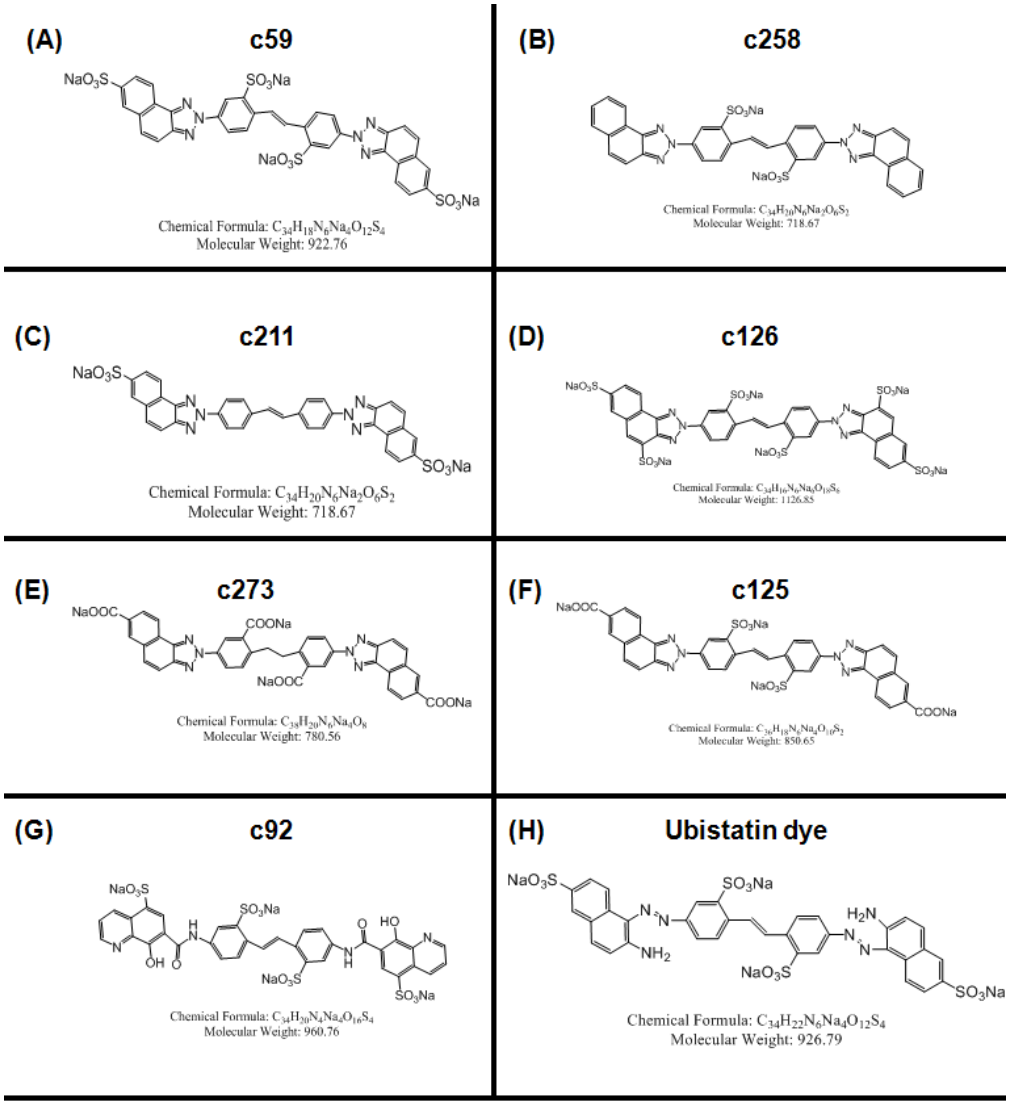


Figure 5.1 – Chemical structures of full ubistatin compounds.
(**A-H**) full ubistatins contain a common ring system and are symmetric. The main differences are in the substituent groups.
* Structures and compounds provided by Dr. Tim Lewis

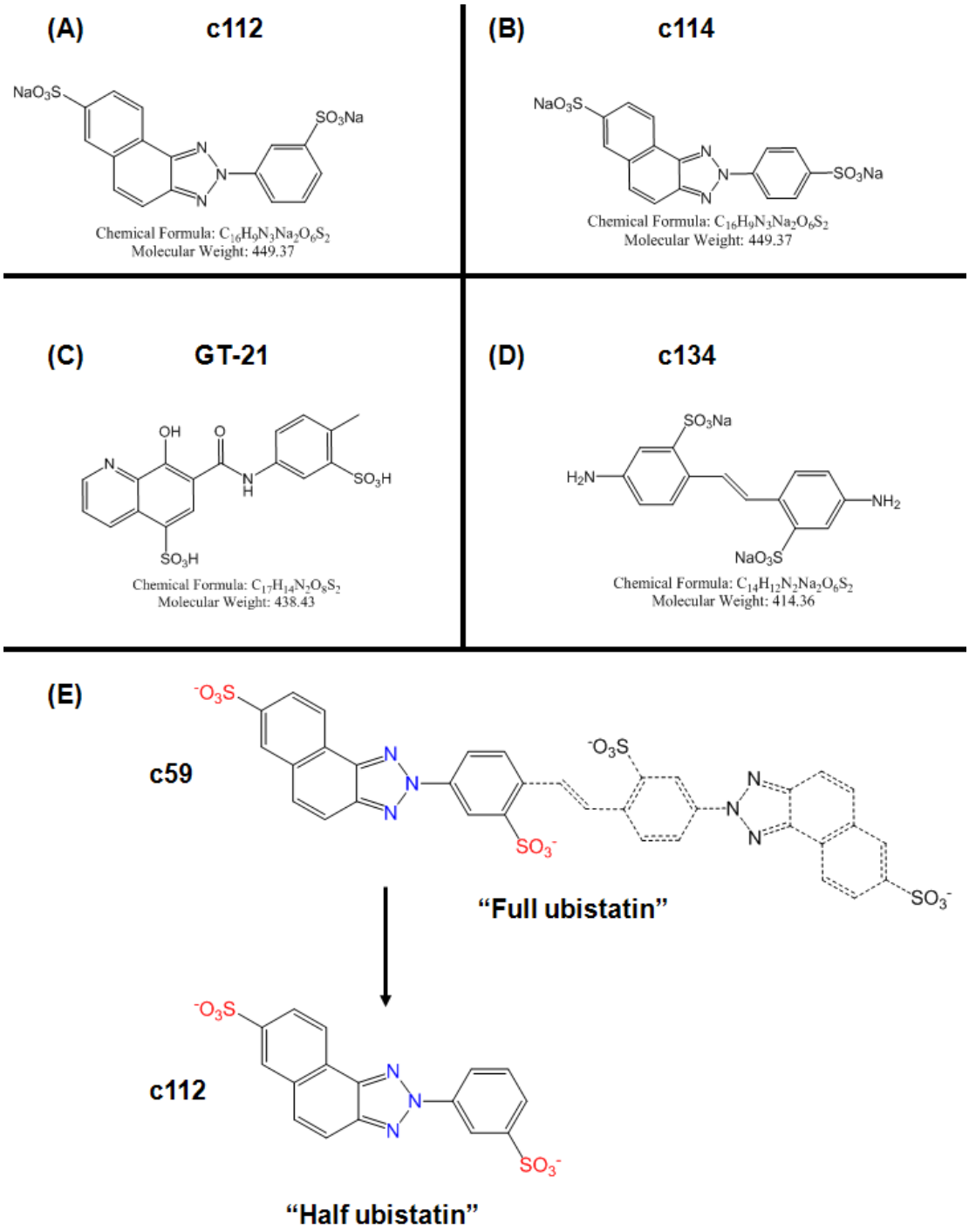


Figure 5.2 – Chemical structures of half ubistatin compounds.
(**A-D**) Chemical structures of half ubistatins maintain the same ring system and lack the alkene linker. (**E**) Schematic showing how the corresponding half ubistatin of c59 results in c112.
* Structures and compounds provided by Dr. Tim Lewis

Ubistatin compounds although highly conjugated do contain at least one single bond which allows for several conformers. In particular, the terms *cis* and *trans* are used to describe the relation of substituent groups to each other. MD simulation and energy minimization performed on the PRODRG server suggests that the *trans* conformation is favored in ubistatins with multiple electro negative groups. A simple example is the half ubistatin c112, where the ring system containing H6-H9 can flip resulting in *cis* and *trans* conformations, however the larger ring system remain rigidly locked (**Figure 5.3**).

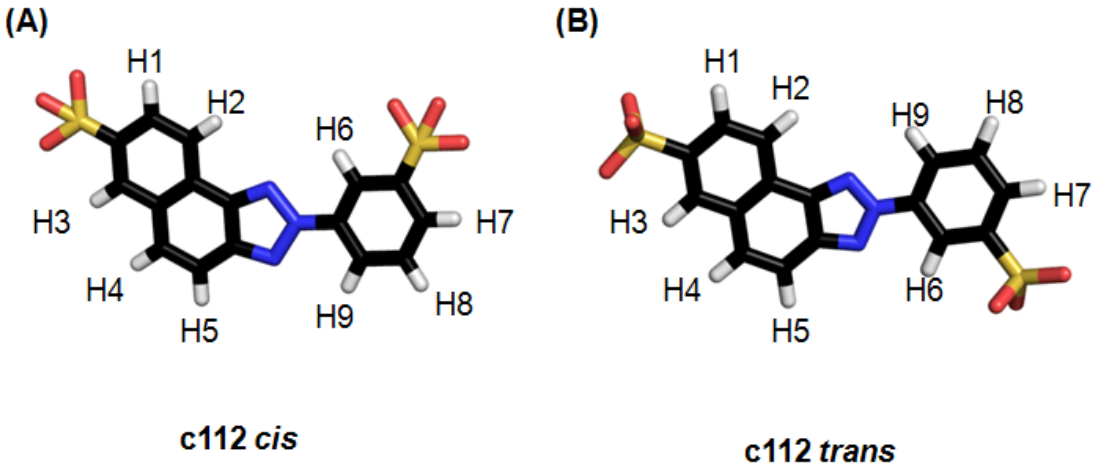


Figure 5.3 – *cis* and *trans* conformations of c112.
(**A**) c112 can take on a *cis* conformation by rotation of the H6,H7,H8,H9 ring positioning the two sulfonate groups (red and yellow sticks) on the same side of the compound. (**B**) The energetically favored *trans* conformation of c112 has the two sulfonate groups on opposite sides.

5.2.2 Spectroscopic properties of ubistatins

Due to the symmetry of full ubistatins, the chemical shifts corresponding to their ^{1}H signals are indistinguishable from either ring system in the compound. Additionally, full ubistatins appear to have a propensity to aggregate at high

concentrations. This aggregation effect is not surprising and has been reported for similar dyes such as Congo Red (*221-223*). For NMR purposes half ubistatins are much more suitable with distinct resonance frequencies for ^{1}H signals (**Figure 5.4**).

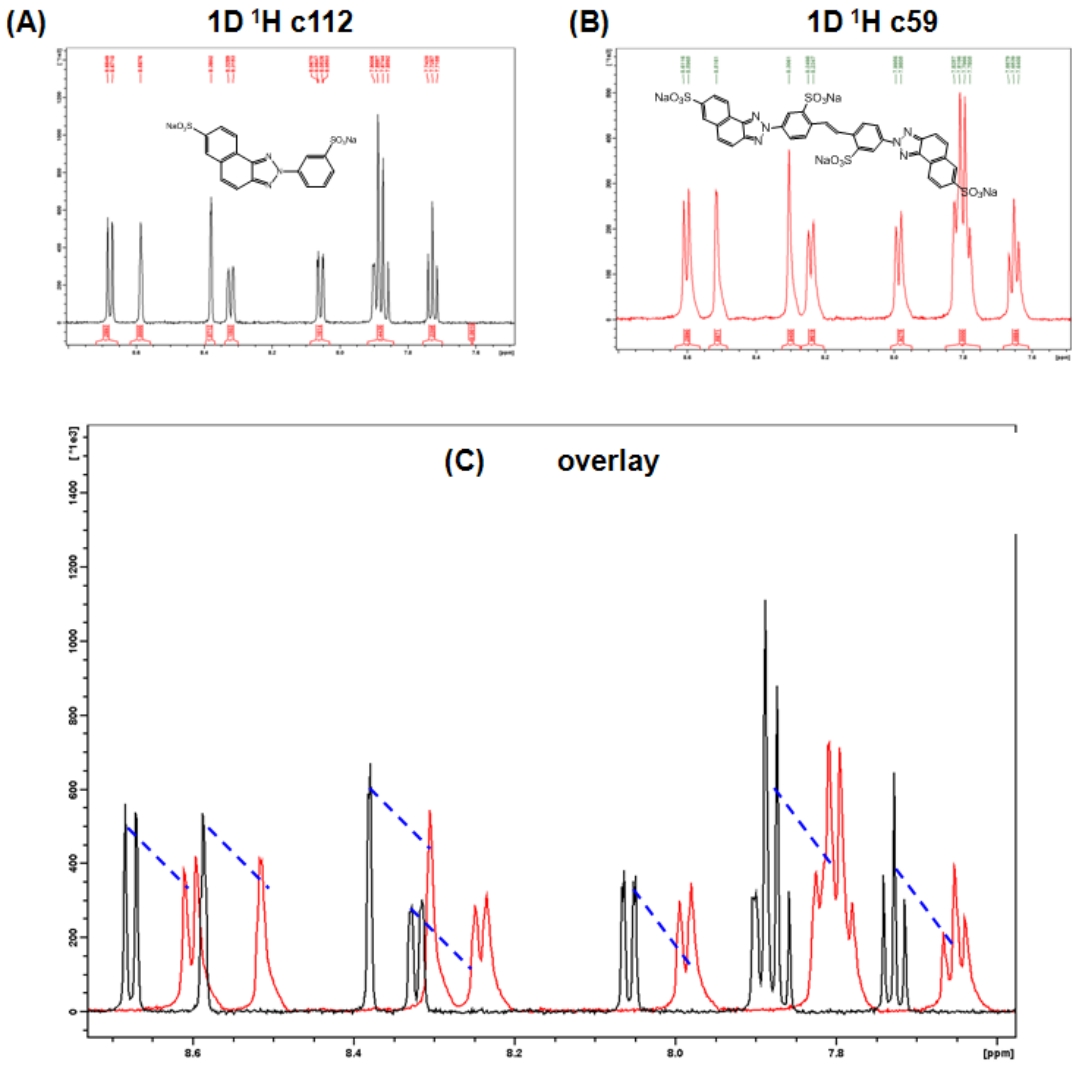


Figure 5.4 – 1D ^{1}H-NMR spectra of c112 and c59
(**A**) ^{1}H-NMR spectrum of the half ubistatin c112 (black) shows narrow line shapes. (**B**) ^{1}H-NMR spectrum of full ubistatin c59 shows (red) a broad line shape suggesting a larger molecular weight due to aggregation. (**C**) Overlay of 1D spectra from c112 (black) and c59 (red) have a nearly identical pattern (traced by blue lines), however there are 9 protons in c112 and 18 in c59 suggesting that protons in c59 have degenerate chemical shifts.

More importantly, the ^{1}H signals from half ubistatins allow for intermolecular NOE's to be assigned to individual protons. Using analysis from 2D ^{1}H, ^{1}H-TOCSY experiments which show three bond ^{1}H J-couplings and 2D ^{1}H, ^{1}H-NOESY, which shows NOEs between ^{1}H nuclei within 5Å, all nine protons in c112 can be assigned (see **Figure 5.5**). Integration of peaks in the 1D spectrum provides an estimate as to how many protons each peak contains. All protons in c112 are isolated except for H4, H5, and H9 which overlap. Nevertheless, useful TOCSY and NOESY signals arise from their positions.

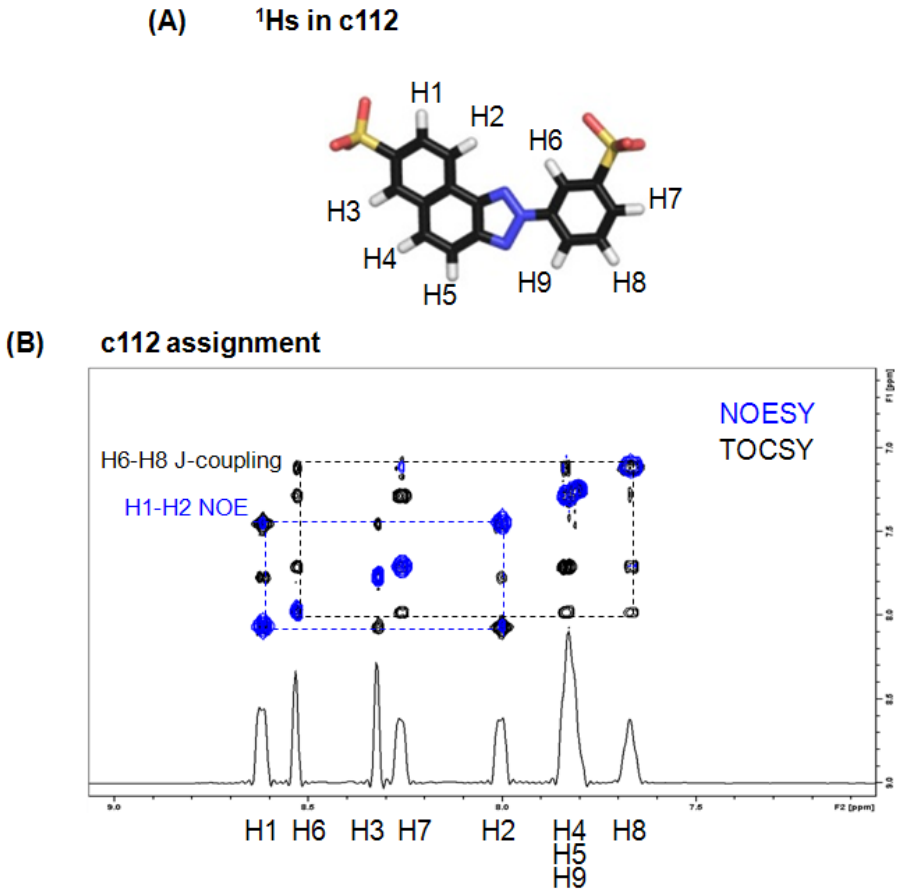


Figure 5.5 – Assignment of protons in c112.
(**A**) Structure of c112 with individual protons labeled 1 through 9. (**B**) Overlay of 2D ^{1}H,^{1}H -TOCSY (black), 2D ^{1}H,^{1}H -NOESY (blue), and 1D ^{1}H (bottom). A trace showing an NOE cross peak between H1 and H2 is highlighted (blue dashed line) and a trace connecting H6 and H8 in TOCSY is shown (black dashed line).

To expand the capabilities of ubistatin compounds I also assessed their fluorescent properties. For quantification purposes the maximum absorbance for each compound was determined. All ubistatins showed strong absorbance in the short UV range (220-300 nm) and had slightly different emission profiles. Given that ubistatins strongly absorbed at 280 nm, I went to test their fluorescent properties at 345 nm, a wave length well above the absorbance of any protein. These studies were performed to assess whether tryptophan emission and FRET assays could be used with ubistatins. Interestingly, the absorbance range of the ubistatin molecules was very broad ranging from 220-360nm for the compounds tested. Conveniently all ubistatins gave a detectable emission with excitation at 308 nm and a transilluminator could be used to confirm their presence in samples (**Figure 5.6**). Furthermore, the wave length corresponding to the emission maxima for each compound deviated highly between each (see **Figure 5.7**).

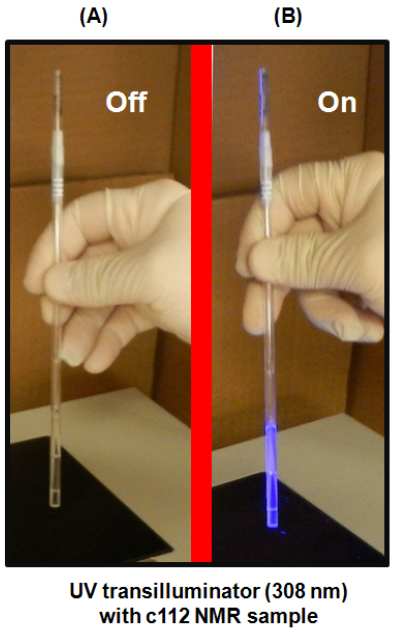


Figure 5.6 – Detection of a colorless ubistatin sample with a UV transilluminator.
(**A**) The 5 mm Shigemi tube containing c112 and Ub_1 appears colorless, but when the transilluminator is turned on (**B**) c112 produced a diagnostic emission in the presence of the 308nm lamp.

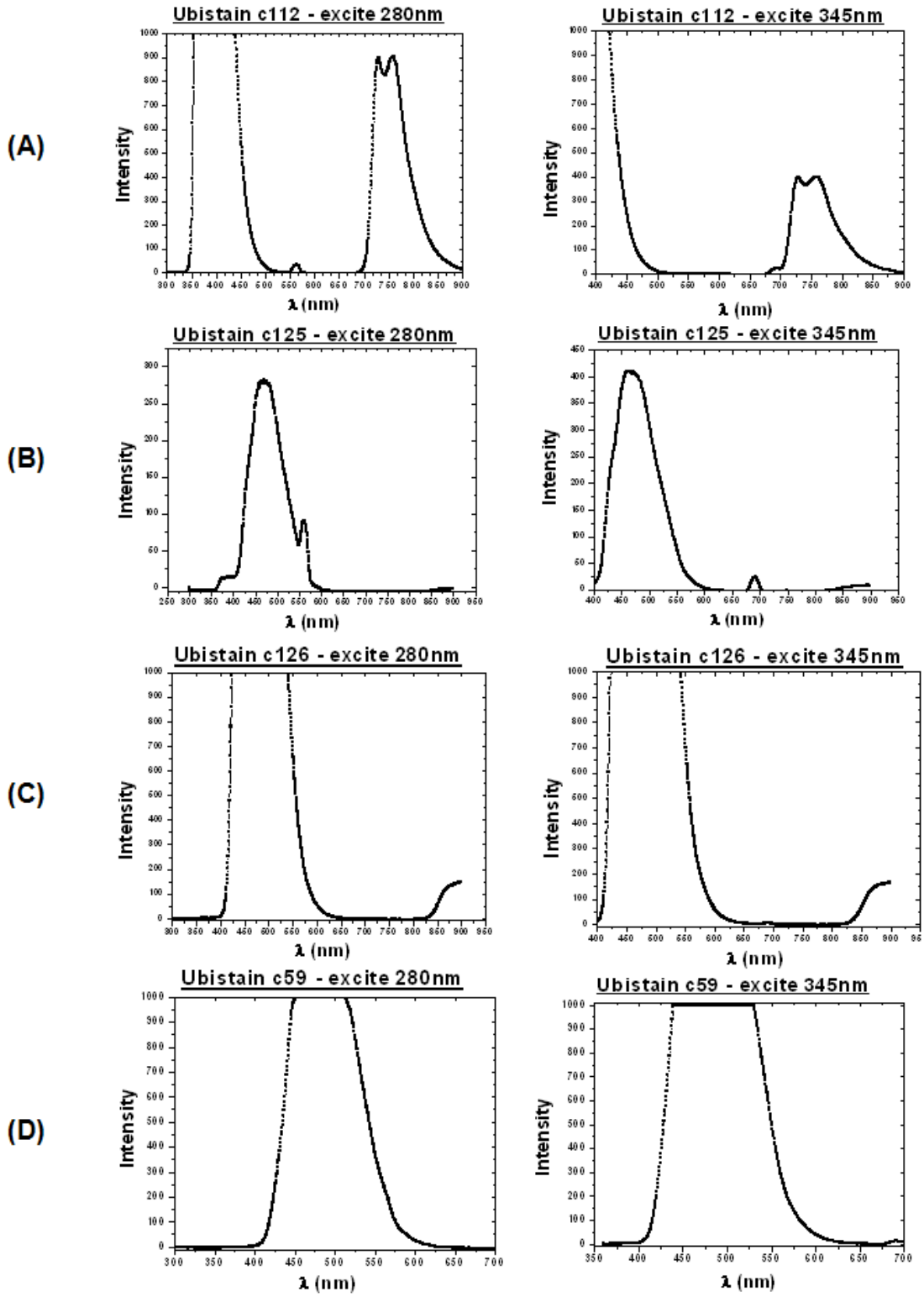


Figure 5.7 – Emission profiles of ubistatin compounds with excitation at 280 nm and 345 nm.

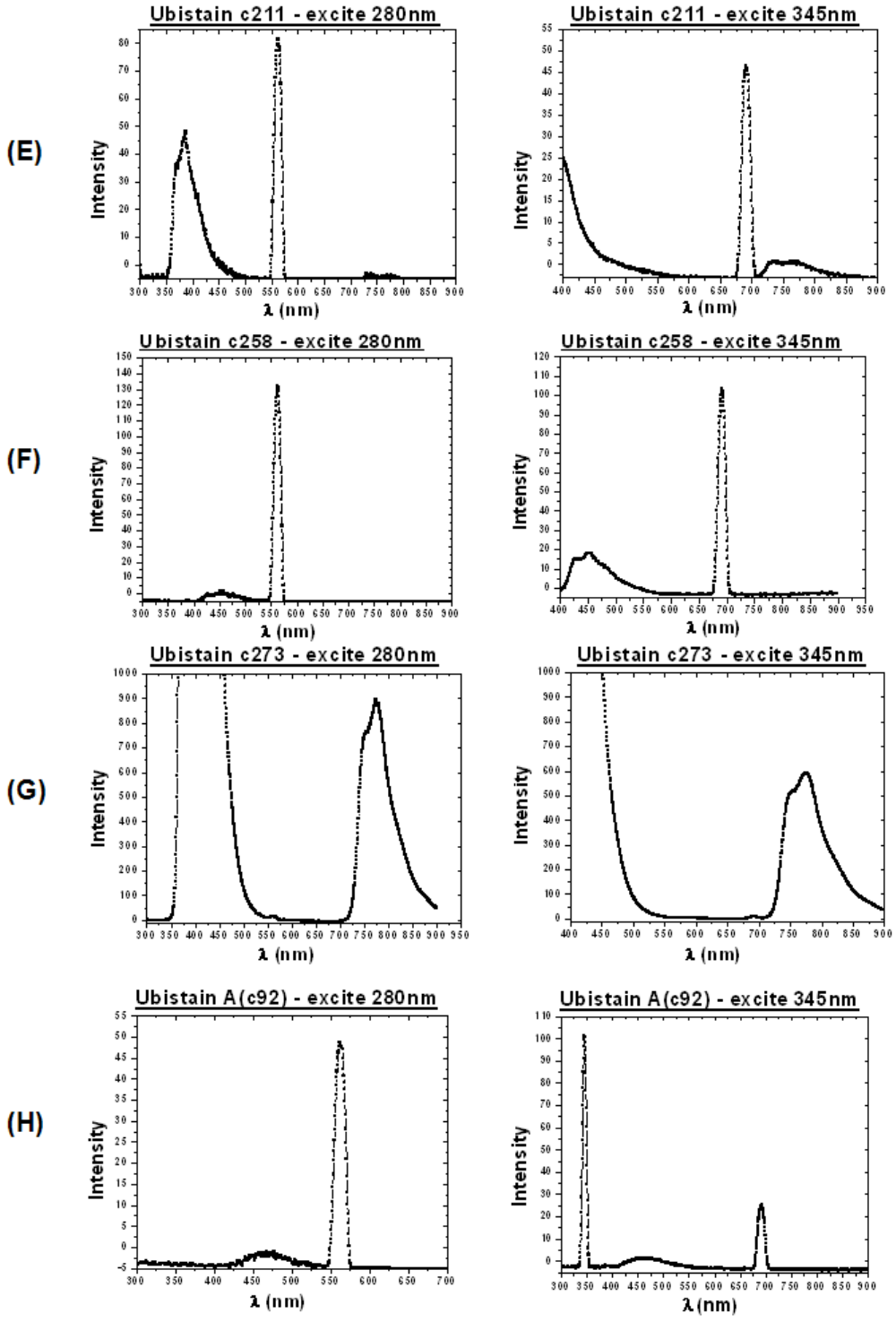


Figure 5.7 (continued) – Emission profiles of ubistatin compounds with excitation at 280 nm and 345 nm.

5.3 Screening for Ub binding ubistatin compounds

5.3.1 $^{1}H,^{15}N$-HSQC titration of Ub_1

Initially $^{1}H,^{15}N$-HSQC titrations were used to determine whether a particular ubistatin compound was able to bind to Ub. The standard *in vivo* assay used previously detected activity, not binding and it is possible that some ubistatins could weakly bind Ub, but fail to show the expected outcome in the assay. For future development of ubistatins it is also essential to understand how certain derivatives bind and the affinity of the interaction. ^{15}N-Ub_1 was titrated with ubistatins c59, c125, c126, c211, c258, c273, half ubistatin c112, half ubistatin c114, and ANS as a control. Originally DMSO was added to help dissolve the ubistatins, however even at low concentrations ~5% (v/v) DMSO appeared to interfere with Ub as evident by large global CSPs in $^{1}H,^{15}N$-HSQC. A buffer containing 20 mM sodium phosphate pH 7.5 greatly improved the solubility of the ubistatins and there was no detectable effect on Ub, which was in 20 mM sodium phosphate pH 6.8 buffer. Furthermore, the ability to dissolve a higher concentration of ubistatins meant a smaller volume would need to be added to the ^{15}N-Ub_1 sample. This system was ideal for many ubistatins, however c258 and c273 were poorly soluble and precipitated early in the titration, while c211 precipitated at a molar ratio of [c211]:[Ub_1]=1.0. The CSP plots (**Figure 5.8**) demonstrate that c59 and its corresponding half ubistatin c112 effectively recognize the L8,I44,V70 hydrophobic patch of Ub.

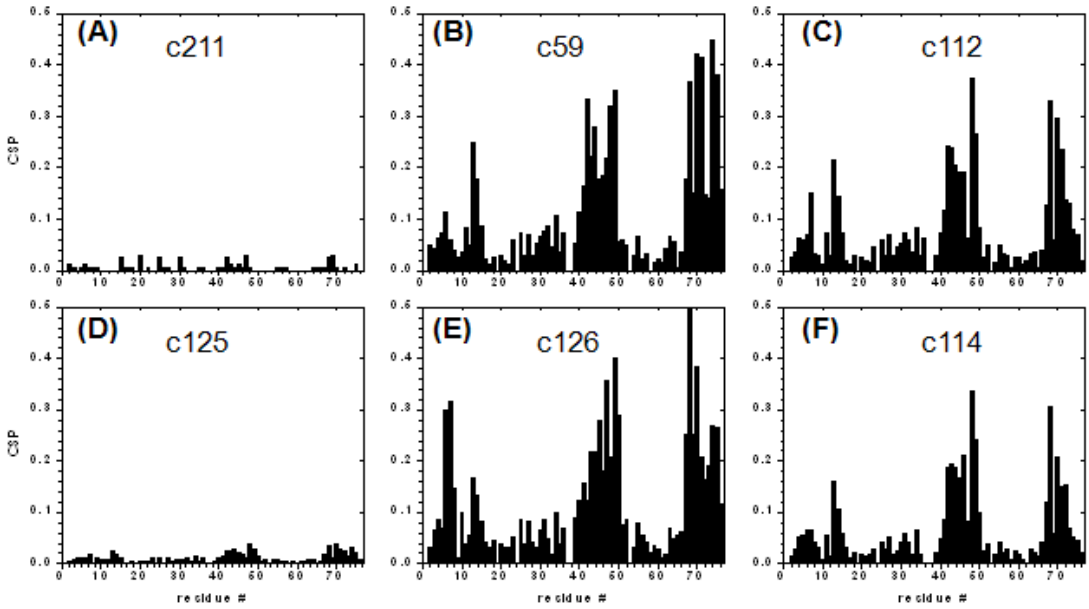


Figure 5.8 – End point CSPs for ubistatin/Ub_1 titrations.
Initial screening by $^{1}H,^{15}N$-HSQC titration with ^{15}N-Ub_1 aimed to characterize binding by the appearance of CSPs around Ub's L8,I44,V70 hydrophobic patch. (**A**) c211 precipitated at [c211]:[Ub_1]=1.0, but did not produce any CSPs on Ub. (**B**) c59, produced large CSPs consistent with binding to the hydrophobic patch. The half ubistatins, (**C**) c112 and (**F**) c114 each create similar CSPs upon binding to Ub. The full ubistatin (**E**) c126 binds Ub strongly, while a similar molecule (**D**) c125 does not bind to Ub.

As a control experiment to assess if related molecules could create specific interactions with Ub, ANS was also titrated into ^{15}N-Ub_1. ANS was chosen because its similar structure to ubistatins and its ability to strongly bind solvent exposed hydrophobic residues suggests that Ub could nonspecifically interact with similar molecules. At a molar ratio of [ANS]:[Ub_1]=2.0, there where small (less than 0.1 ppm) CSPs in the hydrophobic patch residues, which supports that ANS forms weak transient interactions with Ub (**Figure 5.9**). The lack of detectable intermolecular NOEs supports the ANS/Ub interaction is very weak, and stems from a primarily transient effect.

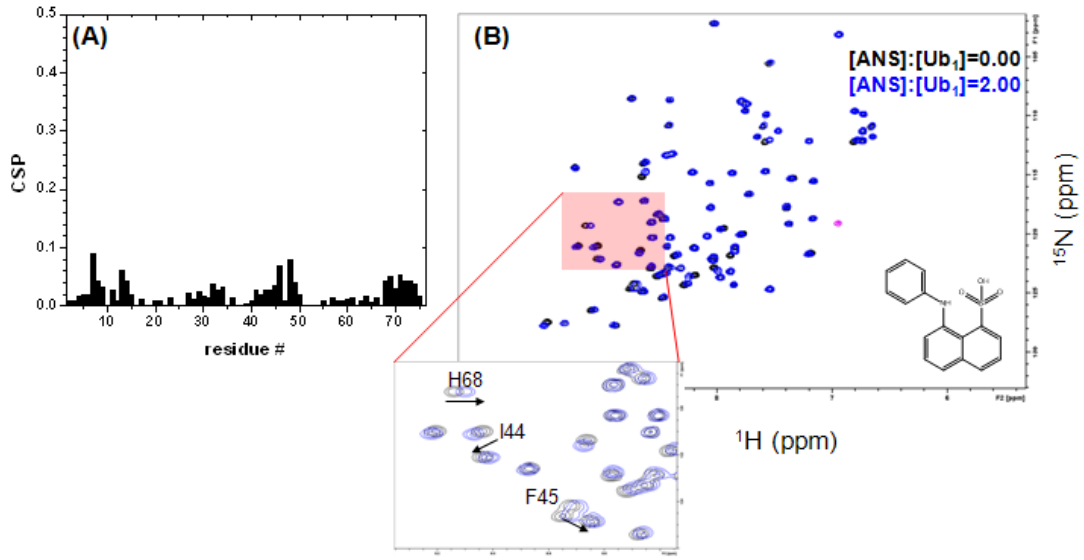


Figure 5.9 – ANS control titration
(**A**) Residue specific CSPs in Ub at a molar ratio of [ANS]:[Ub_1]=2.0. There are slight interactions around residues L8 and I44. (**B**) $^{1}H,^{15}N$-HSQC overlay with no ANS (black) and at [ANS]:[Ub_1]=2.0 (blue). A zoomed in section shows small shifts in H68, I44 and F45. The structure of ANS is at the bottom right of the spectrum.

Interestingly the ubistatins that bound Ub all contained at least two sulfonate groups per ring system. Compounds that preserved the ring system, but lacked substituents were not soluble, while compounds that contained less electronegative groups (e.g. carboxylate groups) failed to bind Ub. At the same time these observations were made, our collaborators also noticed the same trend for their *in vivo* assays, i.e. ubistatin activity was only observed with sulfonate containing compounds. The collaborators also noticed that half ubistatins had a reduced activity. Presumably this was due to different binding properties or affinities of Ub between half and full ubistatins. A comparison $^{1}H,^{15}N$-HSQC spectra shows that both c112 and c59 interact with the same residues on Ub and the fact that trajectories for many

signals in the spectra are identical further supports similar binding between c112 and c59 (**Figure 5.10**).

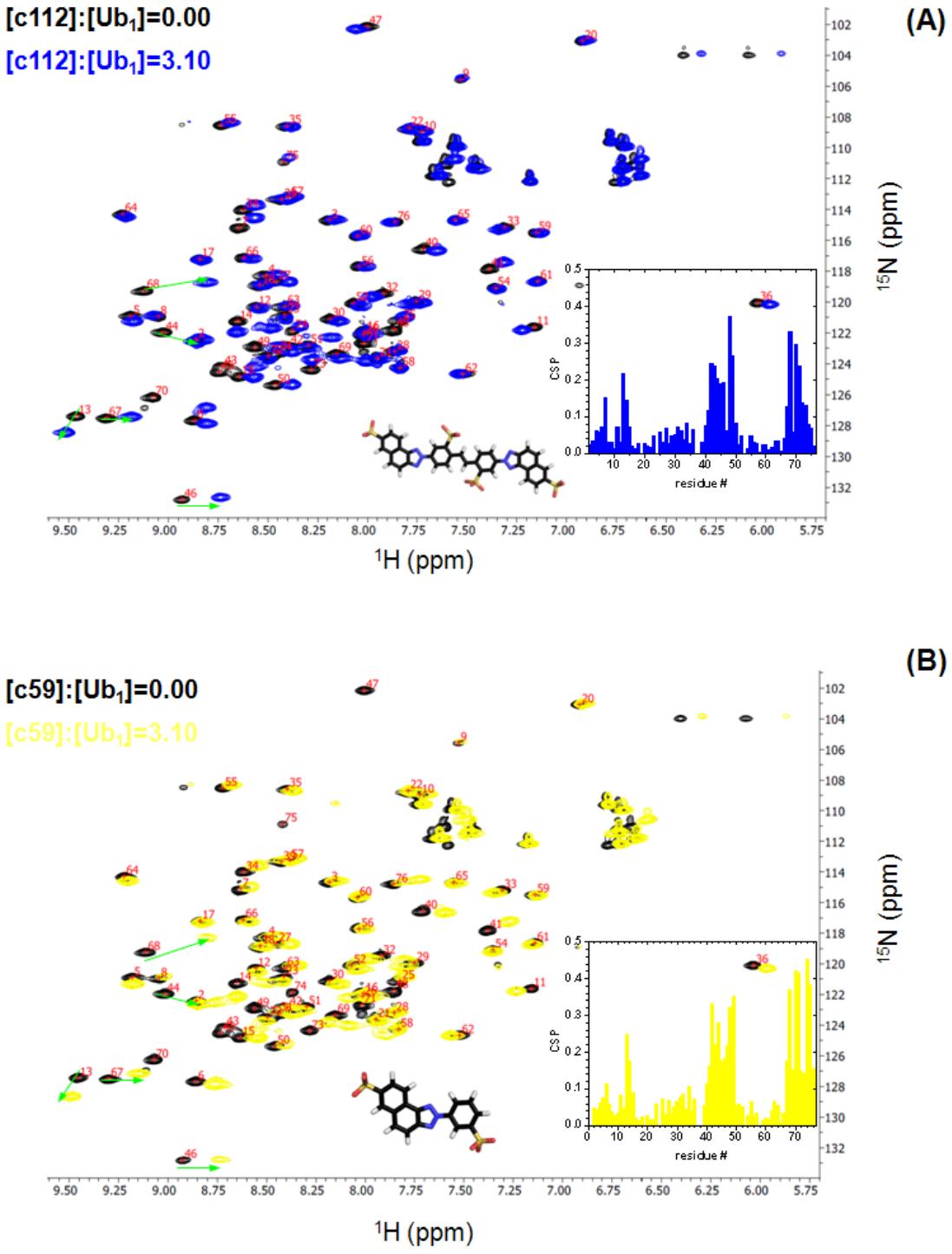


Figure 5.10 – Spectral comparison of c112 and c59 titration with Ub_1.
Residue specific CSP plots at the end point molar ratio (**A**) [c112]:[Ub_1]=3.10 (blue) and (**B**) [c59]:[Ub_1]=3.10 (yellow) are inlayed into their corresponding ^{1}H,^{15}N-HSQC overly. (**A**) Spectral overlay showing starting point [c112]:[Ub_1]=0.00 (black) and [c112]:[Ub_1]=0.00 (blue). (**B**) [c59]:[Ub_1]=0.00 (black) and [c59]:[Ub_1]=0.00 (yellow).

One particular difference between c112 and c59 binding of Ub, was that c59 appeared to cause a loss in signal intensity in $^{1}H,^{15}N$-HSQC, while c112 did not. This could be due to an increase in size, strong sub-micromolar binding, or both. Experiments to address the size of the c59/Ub_1 complex are described in the next section. The residue specific titration curves for c112 and c59 with Ub_1 showed a very different story, the binding of c59 appeared much stronger than the binding of c112 (**Figure 5.11**).

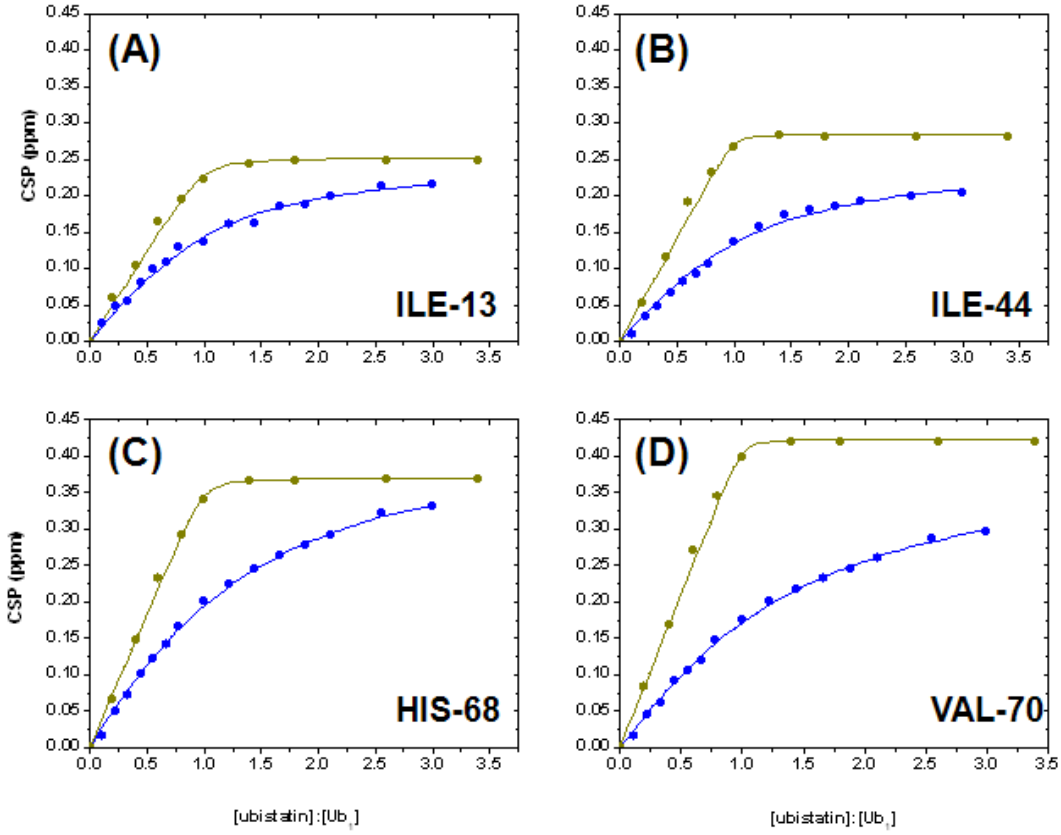


Figure 5.11 – Comparison of residue specific titration curves for c112 and c59. The CSPs for select residues in $^{1}H,^{15}N$-HSQC titration are plotted vs. the molar ratio of c112 and c59 over the course of the titration for residue (**A**) I13, (**B**) I44, (**C**) H68, and (**D**) V70, curves for c59 are in dark-yellow circles and curves for c112 are in blue circles. The line connecting the experimental points is the simulated 1:1 binding curve. Note that the kink in the curves for c59 indicating strong stoichiometric binding. 0.9 mM Ub_1 was titrated and reached a final concentration of ~0.83 mM.

With a 1:1 binding model, the K_d for the c112 for Ub_1 interaction determined by NMR $^{1}H,^{15}N$-HSQC titration was calculated to a $K_d = 212 \pm 44$ μM. Analysis of residuals and a simple visual inspection made it obvious that residue specific curves for the c59/Ub_1 titration exhibited tight stoichiometric binding, preventing me from determining a K_d. To determine the K_d for the c59/Ub_1 interaction, a different method preferably with the concentration of Ub_1 kept in the micromolar range would be needed. Using the F45W mutant of Ub (see **section 7.2**) the K_d of the c59/Ub_1 interaction was measured by tryptophan emission quenching. The spectroscopic properties of c59 allowed for the tryptophan emission to appear undisturbed in the emission spectra, as well as emission from c59 with excitation at 280 nm. A major concern was that overlap from the highly fluorescent c59 compound would disturb the tryptophan signal by increasing intensity in that region. However, this was not the case and upon addition of c59, the emission for tryptophan showed a dramatic decrease in intensity (**Figure 5.12**). The change in intensity at 335 nm was calculated between each titration point and fit to a 2:1 binding model resulting in a $K_d = 3.4 \pm 2.8$ μM for the c59/Ub_1 interaction. If each of the two ring systems of c59 represents a separate Ub binding motif, then the difference in K_d between c112 and c59 is expected. It is also possible that weak affinity of c112 for Ub_1 could dramatically change for polyUb chains, just as the K_d of UBA(2) from hHR23A changes from hundreds of μM for mono-Ub to just 18.7 ± 7 μM for K48-Ub_2.

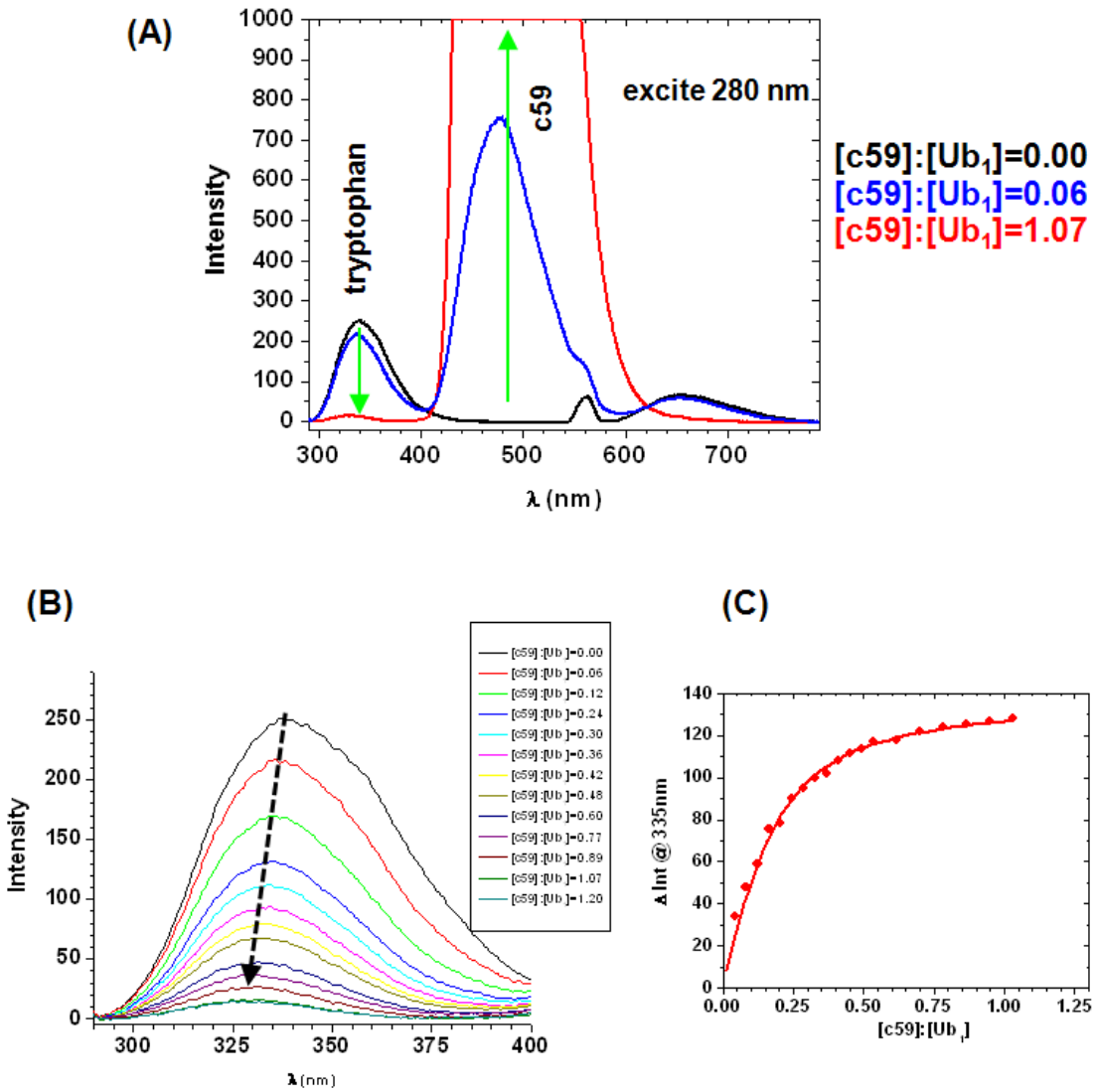


Figure 5.12 – Fluorescence quenching using the Ub F45W mutant to monitor c59 interaction.

(**A**) The tryptophan at reside 45 in Ub_1 and c59 both have detectable emission with excitation at 280 nm. In the absence of c59 (black line) the emission of tryptophan ~340nm is strong, however as c59 is added (blue line) the intensity of tryptophan emission decreases while the emission intensity of c59 spikes. At saturation, there is barely any tryptophan emission (red line), however the large intensity of the c59 signal exceeds the detection limit using these settings. (**B**) Curves representing the spectra for each titration point focused on wave lengths that capture tryptophan emission. (**C**) Fitting the Δ intensity at 335 nm to a 2:1 binding model.

5.3.2 Gauging the specificity of ubistatins

To test if c59, the compound with the highest affinity for Ub could still form strong interactions in a cellular environment I performed a titration in whole cell lysate. ^{15}N-Ub$_1$ was spiked into freshly lysed MCF-7 cells, then c59 was titrated in and the interaction was observed by ^{1}H,^{15}N-HSQC. Upon addition of Ub$_1$ many CSPs were observed without any c59 present along with some decreases in signal intensity for certain residues on Ub. This suggested that in addition to the change in solvent and viscosity, changes in Ub were caused by specific interactions with factors in the MCF-7 lysate. Nevertheless, when c59 was added, nearly an identical CSP pattern and spectral changes were produced (**Figure 5.13**). This result directly proves that ubistatin compounds selectively bind Ub, discriminating from all other proteins in the cell. Furthermore, this represents a straight forward method to monitor the interactions of a specific protein in a realistic cellular environment.

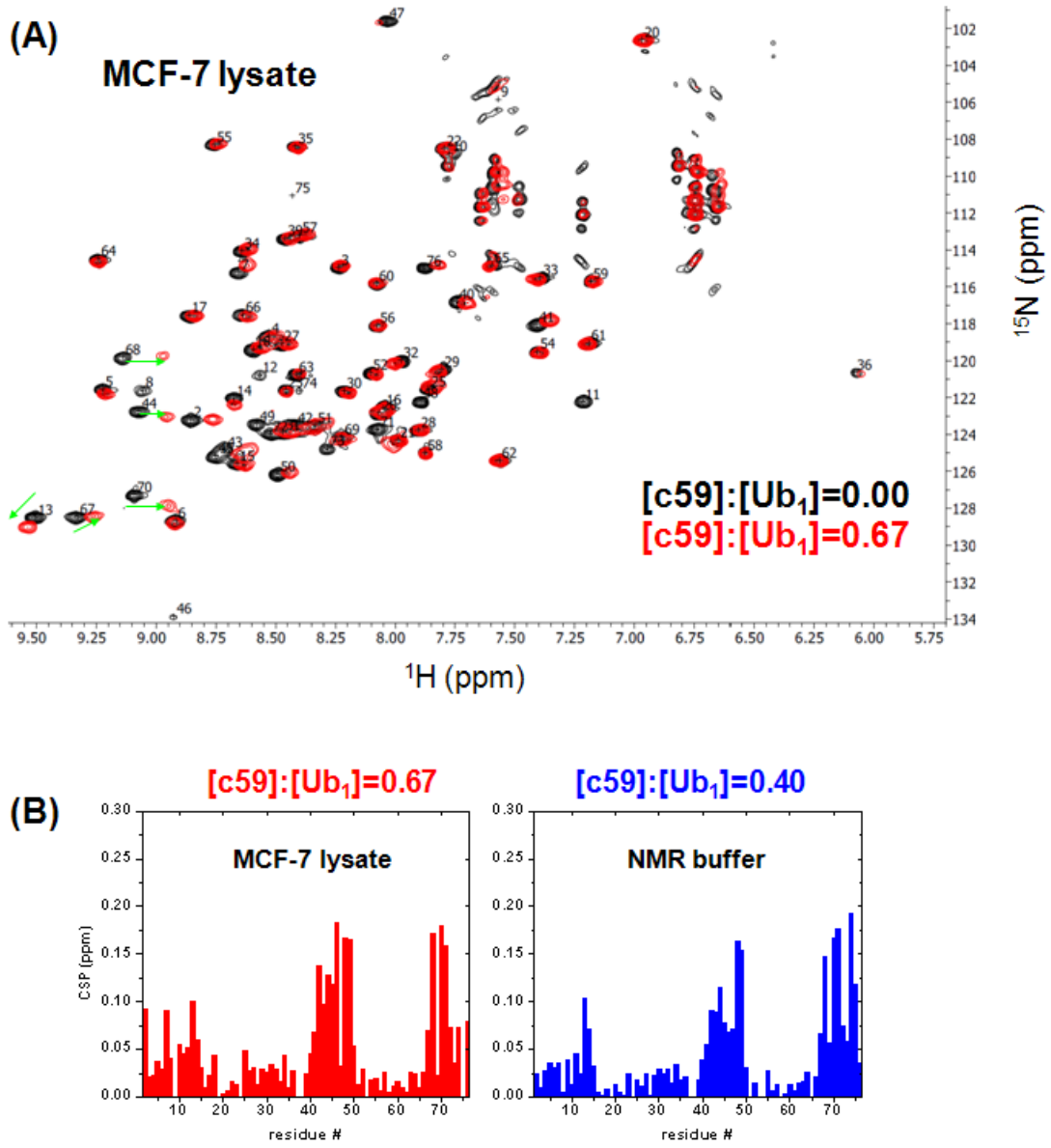


Figure 5.13 – Selectivity of c59 for Ub_1 in MCF-7 lysate.
(**A**) $^{1}H,^{15}N$-HSQC overlay of ^{15}N-Ub_1 in MCF-7 lysate with no c59 (black) and at molar ratio [c59]:[Ub_1]=0.67 (red). (**B**) Residue specific CSPs of from the MCF-7 sample at [c59]:[Ub_1]=0.67 (red) and in standard NMR buffer at molar ratio [c59]:[Ub_1]=0.40 (blue). The MCF-7 sample has smaller CSPs due to the additional factors c59 must overcome to bind Ub.

To determine if ubistatins could discriminate between UBLs and Ub, I titrated Rub1, the closest related UBL to Ub with c112. Interestingly, Rub1 is also 76 residues in length and contains the same L8,I44,V70 hydrophobic patch in the exact same location as Ub. However, some of the residues in Rub1 surrounding the hydrophobic patch differ from those in Ub. If c112 does bind Rub1 there should be some detectable binding interaction at a moderate molar ratio. When ^{15}N-Rub1 was titrated with c112, very small CSPs were observed in ^{1}H,^{15}N-HSQC early in the titration. Still at molar ratio [c112]:[Rub1]=1.45 there were hardly any noticeable CSPs, however the small CSPs that were quantifiable resulted from residues I44, H68, and V70 (see **Figure 5.14**). This suggested that c112 formed weak transient interactions with Rub1, analogous to those attributed to the ANS/Ub interaction. In addition to the titration data for c112 with Rub1 (**Figure 5.14**), the absence of intermolecular NOEs and ^{15}N T_1 measurements also supported that there was hardly any interaction of c112 with Rub1.

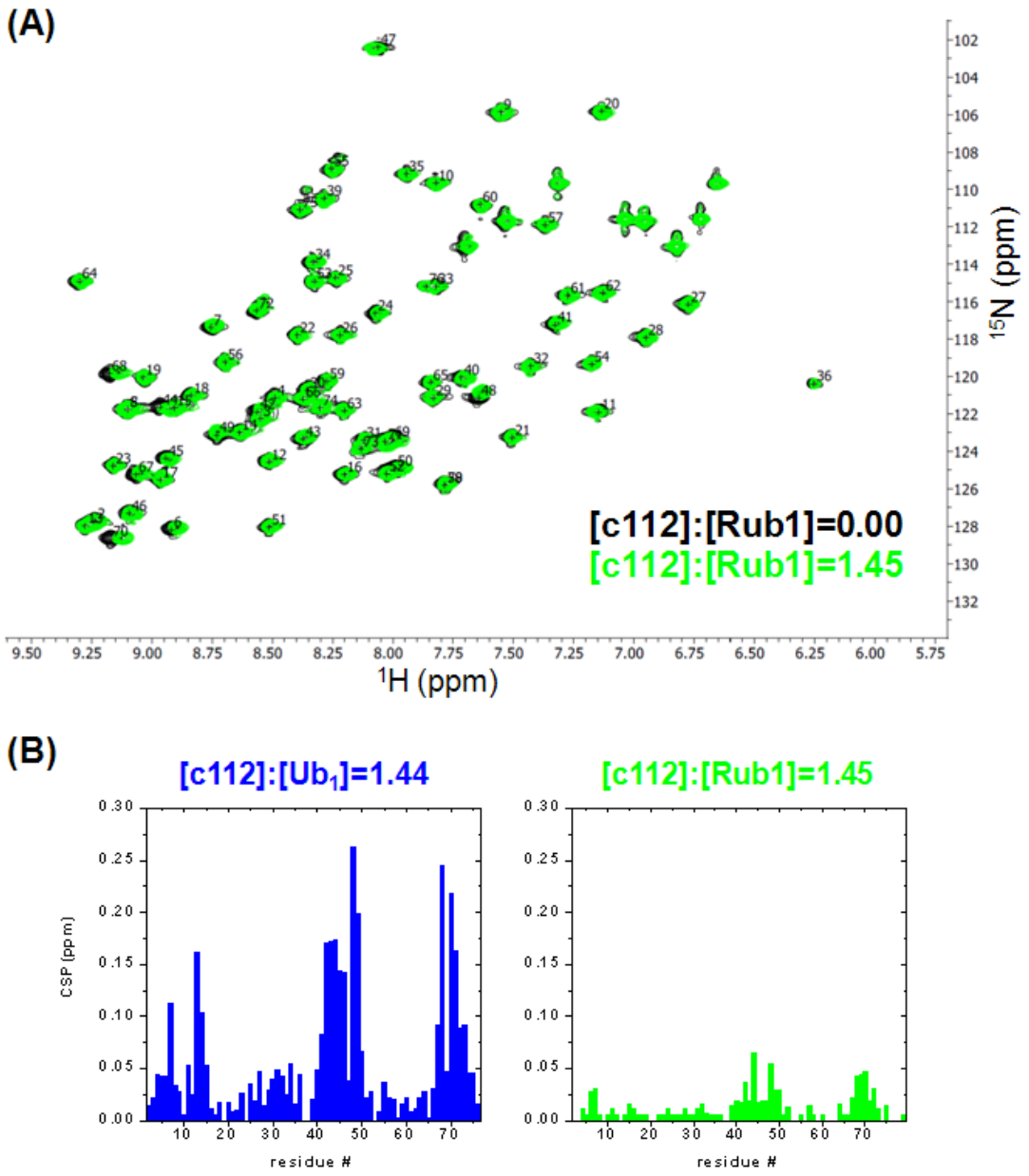


Figure 5.14 – The UBL Rub1 forms weak interactions with c112.
(**A**) Titration of ^{15}N-Rub1 with c112 produces small CSPs in ^{1}H,^{15}N-HSQC. Signals corresponding to [c112]:[Rub1]=0.00 (black) show little deviation from [c112]:[Rub1]=1.45 (green), with slight shifts in I44, H68, and V70. (**B**) Residue specific CSPs for [c112]:[Rub1]=1.45 (green) and for comparison [c112]:[Ub$_1$]=1.44 (blue). The large difference in magnitude supports that c112 binds Ub$_1$ with a much greater affinity than Rub1.

5.3.3 Proteasomal DUBs are inhibited by c112 and c59

Given that it is established ubistatins bind the polyUb chain, I wanted to determine if this interaction would interfere with the activity of DUBs. In the 2004 study they reported polyUb conjugates increased or stabilized in the cell, however the proteasomal DUB Rpn11 still retained some activity (*215*). Another observation was that ubistatins had a greater affinity for K48 linkages compared to K63 linkages. Therefore it is possible that Rpn11 was active because it was cleaving linkages other than K48 that reached the proteasome or since Rpn11 is reported to cleave the Ub-substrate bond, the ubistatins bound to the polyUb signal did not affect Rpn11 activity (*51, 212*). The effects of ubistatins on DUB activity has yet to be tested and I found it necessary to investigate this question. Ubp6, a proteasome associated USP DUB is accepted as the second major DUB associated with the proteasome and prefers to cleave Ub-Ub bonds from the distal end. This is in contrast to Rpn11 which can cleave the Ub-substrate bond and also endo Ub-Ub bonds (*213*). Similar to other USP DUBs, Ubp6 can cleave several different Ub-Ub linkages, but preferentially cleaves K48 over K63 (*31*). Using a simple gel assay, I set to answer if (i) ubistatins generally inhibit DUB activity and (ii) if ubistatins do inhibit DUBs, is this only for certain Ub-Ub linkages. The molar ratio of Ubp6 to K48 or K63 linked Ub_2 was always kept at 0.2. The molar ratio of c112 and c59 were varied in respect to the Ub_2, testing ratios of 0.2, 0.5, 1.0, and 3.0.

The data for K48-Ub_2 with c112 and c59 presented in (**Figure 5.15**) show that both compounds are capable of shielding the polyUb chain from Ubp6. Even at concentrations bellow a 1:1 molar ratio (ubistatin:K48-Ub_2) there is still a noticeable

inhibitory effect. Quantitatively we see that c59 is a more potent inhibitor of Ubp6 than c112, as 80% of K48-Ub_2 remains after 3 hours, while only 40% remains for c112 at the same molar ratio, [ubistatin]:[K48-Ub_2]=3.0. This observation is consistent with the previously reported data.

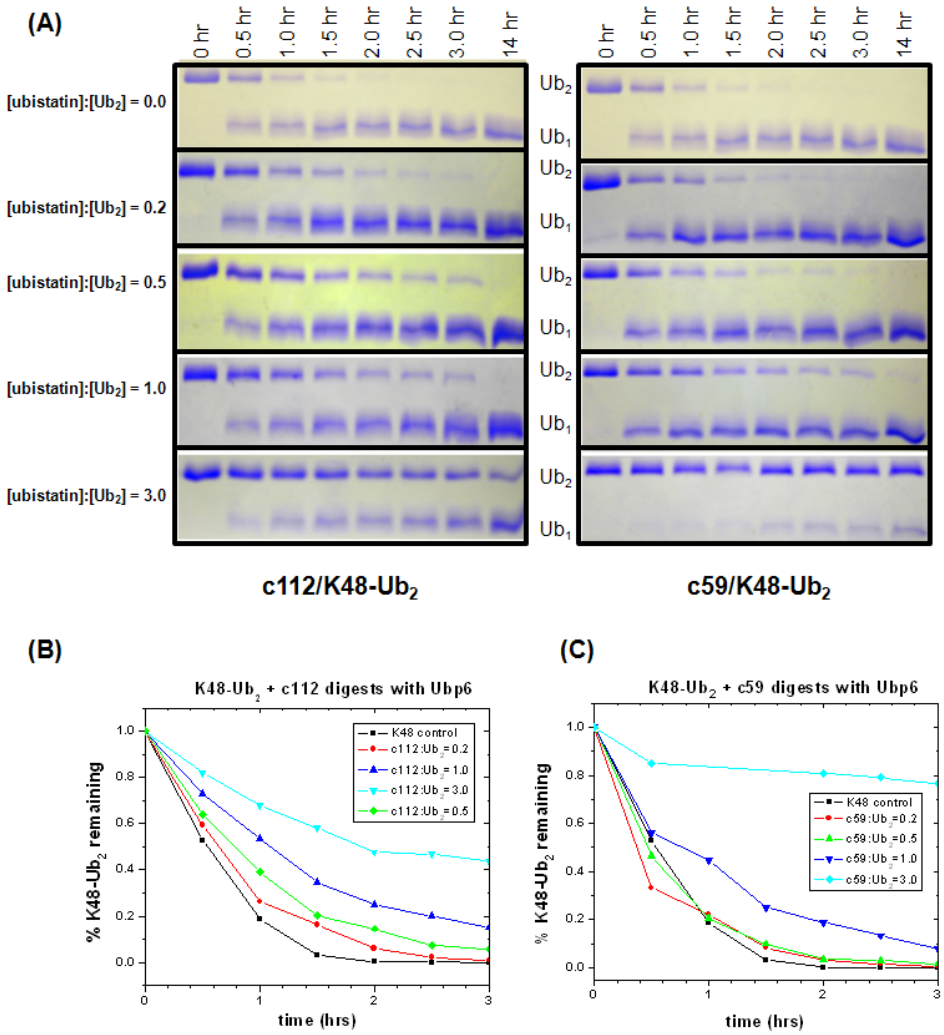


Figure 5.15 – c112 and c59 shield K48 linked polyUb from Ubp6.
(**A**) The mono- and di-Ub regions of 15% SDS-PAGE gels with time points of 0, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 hrs. In the control (top gels) K48-Ub_2 is reduced to mono-Ub quickly, however at sufficient concentrations of c112 (left panel) or c59 (right panel) K48-Ub_2 is preserved. (**B, C**) For each time point, the band corresponding to K48-Ub_2 was integrated and then normalized. (**B**) Inside of the 3 hr time window all concentrations of c112 retain a greater amount of K48-Ub_2 than the control (**C**) For c59 and there is as similar effect and at a molar ratio [c59]:[K48-Ub_2]=3.0 a staggering 80% of the K48-Ub_2 remains.

With a noticeable inhibitory effect on K48 linkages, I tested K63-Ub_2 using the same assay. The results (**Figure 5.16**) show a slightly different story. K63 linkages are cleaved at slower rate than K48 linkages by Ubp6 without the ubistatin, however it does not appear that moderate concentrations of either c112 or c59 were helpful in shielding K63-Ub_2 from Ubp6. For K48-Ub_2 Ubp6 inhibition was observed even at low molar ratio [ubistatin]:[Ub_2]=0.5 (**Figure 5.15 B,C)**, however for K63-Ub_2 inhibition of Ubp6 does not occur until the relatively high ratio, [ubistatin]:[Ub_2]=3.0. Furthermore, ubistatins appear to facilitate the cleavage of K63-Ub_2 by Ubp6. Neither with K48 nor K63 linked polyUb chains is the half ubistatin c112 more effective than its corresponding full ubistatin c59. Taken together, this result suggests that when ubistatins are introduced to the cell they also prevent K48 polyUb signals from being cleaved by DUBs, in addition to blocking their interactions with receptors. If the polyUb signal must be removed by the proteasome this observation provides an additional mechanism as to how ubistatins inhibit the UPP as we understand.

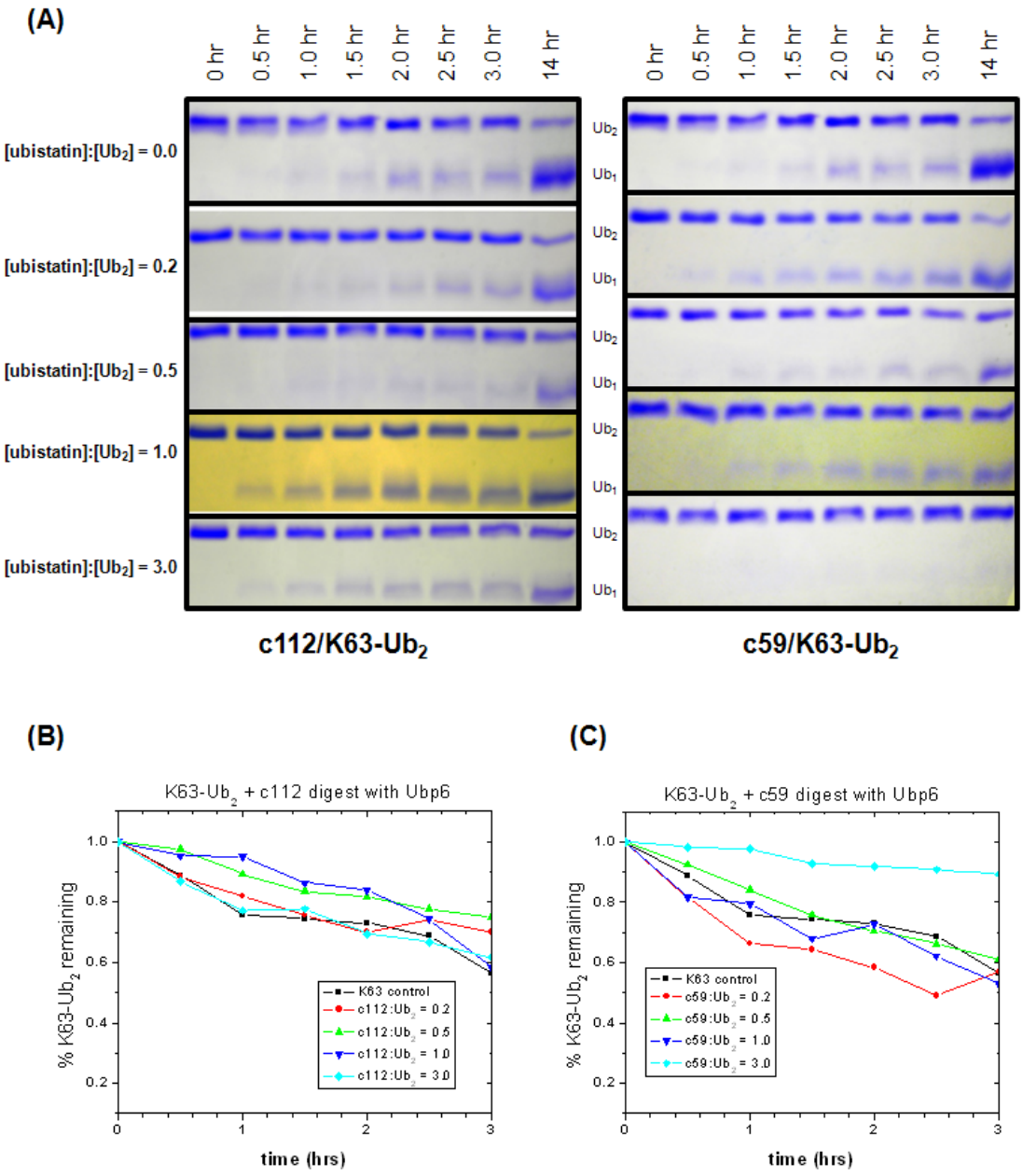


Figure 5.16 – Minimal effect of c112 and c59 on Ubp6 activity for K63 linkages. (**A**) The mono- and di-Ub regions of 15% SDS-PAGE gels with time points of 0, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 hrs. The control (top gels) show that K63-Ub_2 begins being cleaved after about 1.5 hrs. Several concentrations of c112 (left panel) and c59 (right panel) show ubistatins can inhibit Ubp6 from cleaving K63-Ub_2. (**B, C**) For each time point, the band corresponding to K63-Ub_2 was integrated and then normalized. (**B**) Very little deviation from the control is seen for c112 inside of the 3 hr time window (**C**) For c59 all other conditions besides [c59]:[K63-Ub_2]=3.0 shows little to no inhibition.

5.4 Comparison of c112 and c59 interactions

5.4.1 Molecular weight by translational diffusion and ^{15}N-T_1 relaxation

The strong signal attenuations in ^{1}H,^{15}N-HSQC titration experiments described in (**section 5.3**) indicated that c59 bound Ub_1 tightly, but this alone could not be used to determine if the c59/Ub_1 complex oligomerized multiple Ubs. To determine the average molecular weight of the c112/Ub_1 and c59/Ub_1 complexes DOSY experiments which measure translational diffusion, and ^{15}N T_1 relaxation experiments which are sensitive to both the size and shape of the molecule under observation were performed for each complex. Over the course of the titration it was possible to measure the translational diffusion properties of both the ubistatin and ^{15}N-Ub_1, however due to practical reasons ^{15}N-T_1 relaxation was only measured for the beginning and endpoints of each titration for Ub.

Translational diffusion is a less sensitive method for measuring molecular weight compared to rotational diffusion experiments (T_1 and T_2 relaxation), however it is one of the surest NMR experiments to gauge the oligomeric state of ubistatin/Ub complexes in a relatively short acquisition time. For select points along each titration the translational diffusion coefficient was measured using a series of 1D experiments that differed in z-gradient strengths and also had ^{15}N filtering. The ^{15}N filter was used to segregate signals from the ubistatin and those from Ub. Conveniently the proton signals from c59 and c112 fall in the amide region of Ub, but do not overlap with aromatic protons (^{12}C-^{1}H) in Ub. The translational diffusion coefficient was calculated using ^{1}H signals from the methyl containing 3-0 ppm region of ^{1}H spectra. The data over each titration (**Figure 5.17**) show that the translational diffusion

coefficient of Ub in the presence of c59 decreases, while c112 causes hardly any change in Ub.

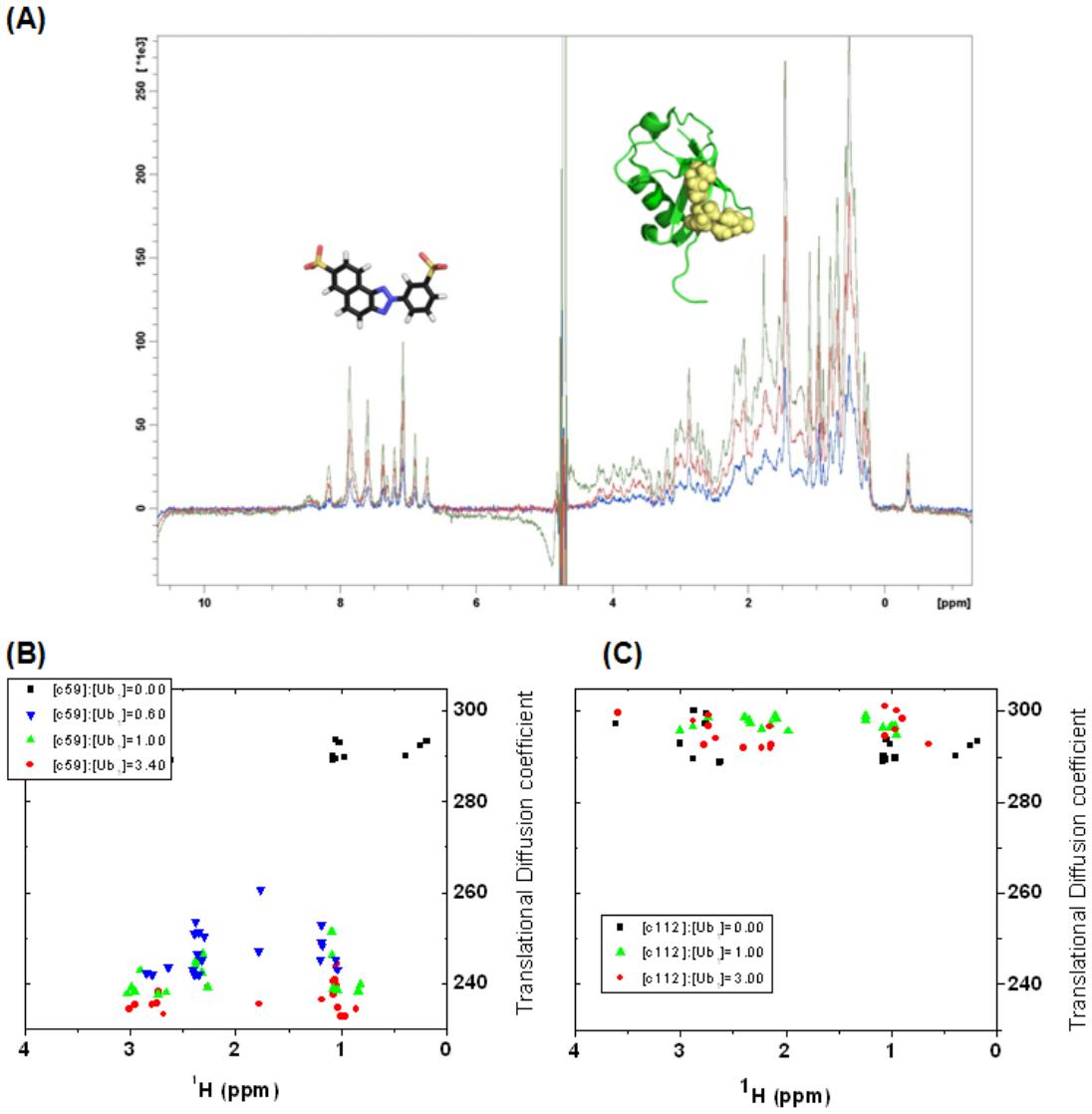


Figure 5.17 – Translational diffusion measurements of Ub with c59 and c112. (**A**) A series of 1D ^{1}H spectra collected under varying z-gradient strengths produce different intensities that can be used to calculate the translational diffusion coefficient. Signals from ubistatin only appear between 7-9ppm, while the whole right side of the spectrum contains signals from Ub. With no ubistatin, signals from Ub_1 have a translational diffusion coefficient between 280-300 (black squares). (**B**) As c59 is titrated in, the translational diffusion efficient decreases early [c59]:[Ub_1]=0.60 (blue triangles), but does not change significantly when c59 reaches [c59]:[Ub_1]=3.4 (red circles). (**C**) For Ub_1 with c112 there is no significant change between free Ub_1 (black squares) or when [c112]:[Ub_1]=3.0 (red circles). This suggests that c59 creates a complex with more than one Ub.

The findings from translational diffusion measurements painted a stark contrast between c112 and c59, even though each bound the same surface of Ub, c59 appeared to form at least a 1:2 (c59:Ub_1) complex with Ub while c112 formed a well behaved 1:1 complex.

To understand each complex better, I also measured ^{15}N T_1 for each titration end point and used existing T_1 data for monomeric Ub, K48-Ub_2, and K63-Ub_2 to estimate how large the complex was. The results (**Figure 5.18**) certainly show an increase in size for c59, but also a small ~80 ms increase in ^{15}N T_1 for c112 that was not as pronounced in the translational diffusion data. K63-Ub_2 lacking any interdomain contacts has a shorter T_1 compared to K48-Ub_2, yet the T_1 for the c112/Ub_1 complex still has a shorter T_1 compared to both di-Ubs suggesting that it does not oligomerize Ub. T_1 data for the c59/Ub_1 complex gives a much clearer picture about the size, which fall just below T_1 values of K48-Ub_2. This assertion agrees well with the translation diffusion data and supports a 1:2 c59/Ub_1 binding interaction. Interestingly, the c59/Ub_1 complex (more accurately c59/2Ub_1) does not appear to reach a larger size at any point in the titration. The T_1 data would support a model in which two Ubs tightly enclose one molecule of c59, given that the T_1 values are closer to K48-Ub_2 than K63-Ub_2.

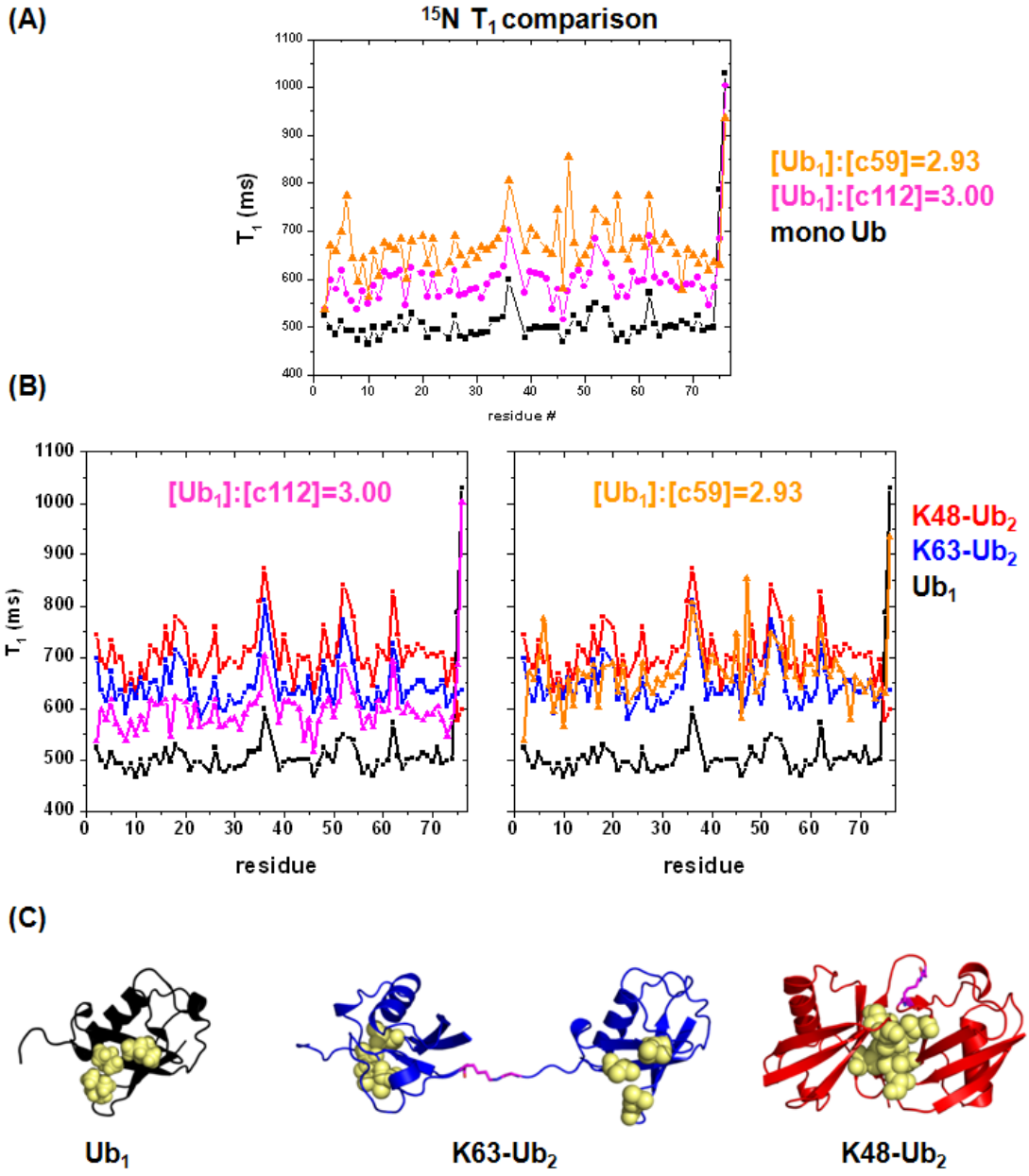


Figure 5.18 – ^{15}N T$_1$ relaxation reveals the stoichiometry of ubistatin/Ub complexes.
(**A**) At the end of the c112 and c59 titrations, the T$_1$ was measured. In respect to free monomeric Ub (black line), the c112/Ub complex (magenta line) shows a slight increase in T$_1$ for Ub, however, the c59/Ub$_1$ complex (orange line) has the largest increase in T$_1$. (**B**) To estimate the size of each complex, the T$_1$ from (**A**) was plotted against monomeric Ub (black), K63-Ub$_2$ (blue), and K48-Ub$_2$ (red). For c112 (magenta), the T$_1$ values do not reach those of K63-Ub$_2$, however the T$_1$ values for c59 (orange) essentially fall right between K63-Ub$_2$ and K48-Ub$_2$ indicating the c59/Ub$_1$ complex behaves as a dimer. (**C**) Corresponding structures of the Ub species used for comparison color coded as in the T$_1$ data.

5.4.2 Detection of intermolecular NOEs

To confirm the close proximity of c112 and c59 to Ub, a 2D $^{1}H,^{1}H$-NOESY with ^{15}N filtering was acquired at each titration endpoint as well. NOE's are detectable when two nuclei are within a distance of 5Å and since the ^{1}H signals of both ubistatins fall within the amide range (see **Figure 5.4**), the ^{15}N filter would suppress intermolecular NOEs in Ub and only allow for NOE's between the ubistatin and Ub to be isolated. Each ubistatin compound would be expected to show some intramolecular NOEs, however if the compound is in close proximity to the L8,I44,V70 hydrophobic patch there should be some NOEs to the methyl region as well. The c59/Ub_1 complex is essentially twice the size of the c112/Ub_1 complex and this increases the line shape for all signals, even NOEs. When the 2D $^{1}H,^{1}H$-NOESY was performed on a 600 MHz instrument NOEs were not detected for c59/Ub_1. This is attributed to both the size of the complex and also the tight binding of c59 at residues that would produce NOEs. As shown in $^{1}H,^{15}N$-HSQC, the tight binding of c59 can broaden signals beyond detection. Many NOEs were observed between protons in c112 and side chains of Ub (**Figure 5.19**). These NOEs are consistent with binding to the methyl side chains of the L8,I44,V70 hydrophobic patch and there also appears to be NOEs to non-methyl groups on Ub as well.

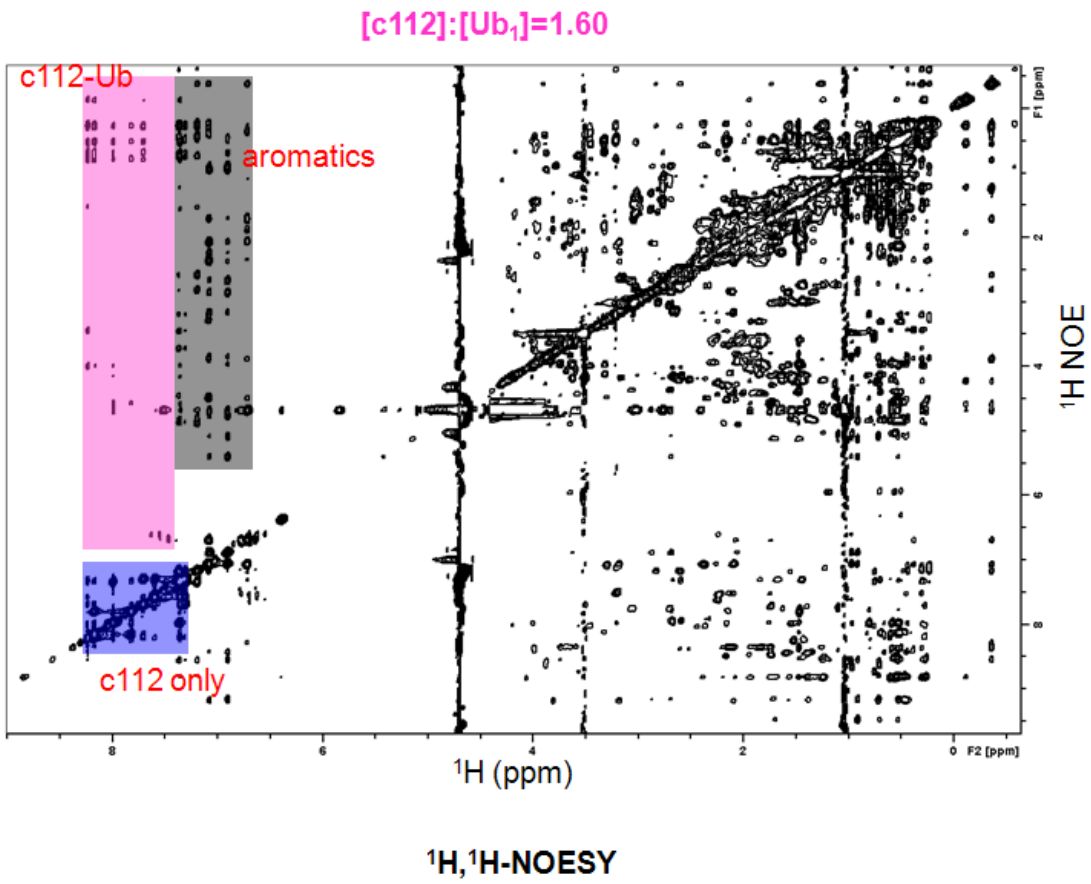


Figure 5.19 – Detection of intermolecular NOEs between c112 and Ub.
In this spectrum all NOEs are cross peaks formed off of the diagonal line. ^{1}H,^{1}H-NOESY with ^{15}N filtering shows that the protons in c112 (blue region) have NOEs to residues in Ub (magenta region). Many of the NOEs between c112 and Ub fall within the methyl range. Aromatic protons in Ub (black region) also show many NOEs, but signals do not interfere with those from protons in c112.

5.4.3 Interactions with K48 and K63 polyUb chains

To address how both ubistatins interacted with different linkages, K48-Ub_2 and K63-Ub_2 were titrated with each ubistatin. It has been reported that c59 is K48 selective, but this conclusion was determined in a relatively crude gel shift assay. Using ^{1}H,^{15}N-HSQC titration in combination with ^{15}N T_1 relaxation, I show that both c59 and c112 form essentially the same interactions as they do with monomeric Ub. Again, c112 and c59 both utilize the same binding surface on Ub, however the characteristics of binding are very different. From just a simple overlay of the ^{1}H, ^{15}N-HSQC spectra it is apparent from the pattern of signal attenuations that c59 strongly binds both K48 and K63 linkages (**Figure 5.20 and 5.21**). Spectra for c112 demonstrate that it also interacts with both linkages, but does not produce the attenuations observed for c59.

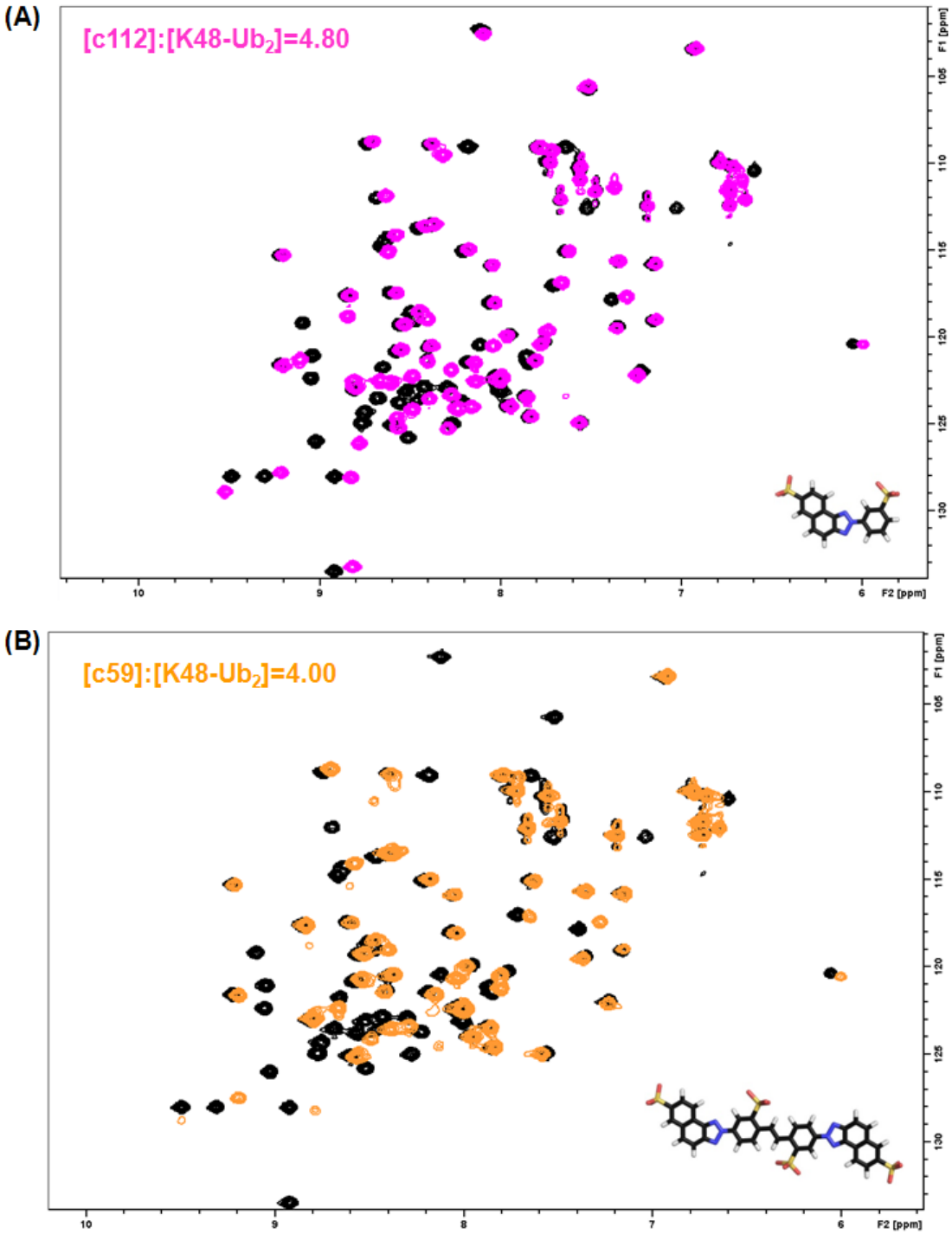


Figure 5.20 – $^{1}H,^{15}N$-HSQC titration of K48-Ub_2
Titration of K48-Ub_2 with both c59 and c112 produces residue specific shifts from binding. (**A**) Overlay of unbound K48-Ub_2 (black) with titration point [c112]:[K48-Ub_2]=4.8 (magenta) shows that signals retain sharp line shapes at saturation (**B**) The overlay of unbound K48-Ub_2 (black) with titration point [c59]:[K48-Ub_2]=4.00 (orange) shows many attenuated peaks indicative of tight binding.

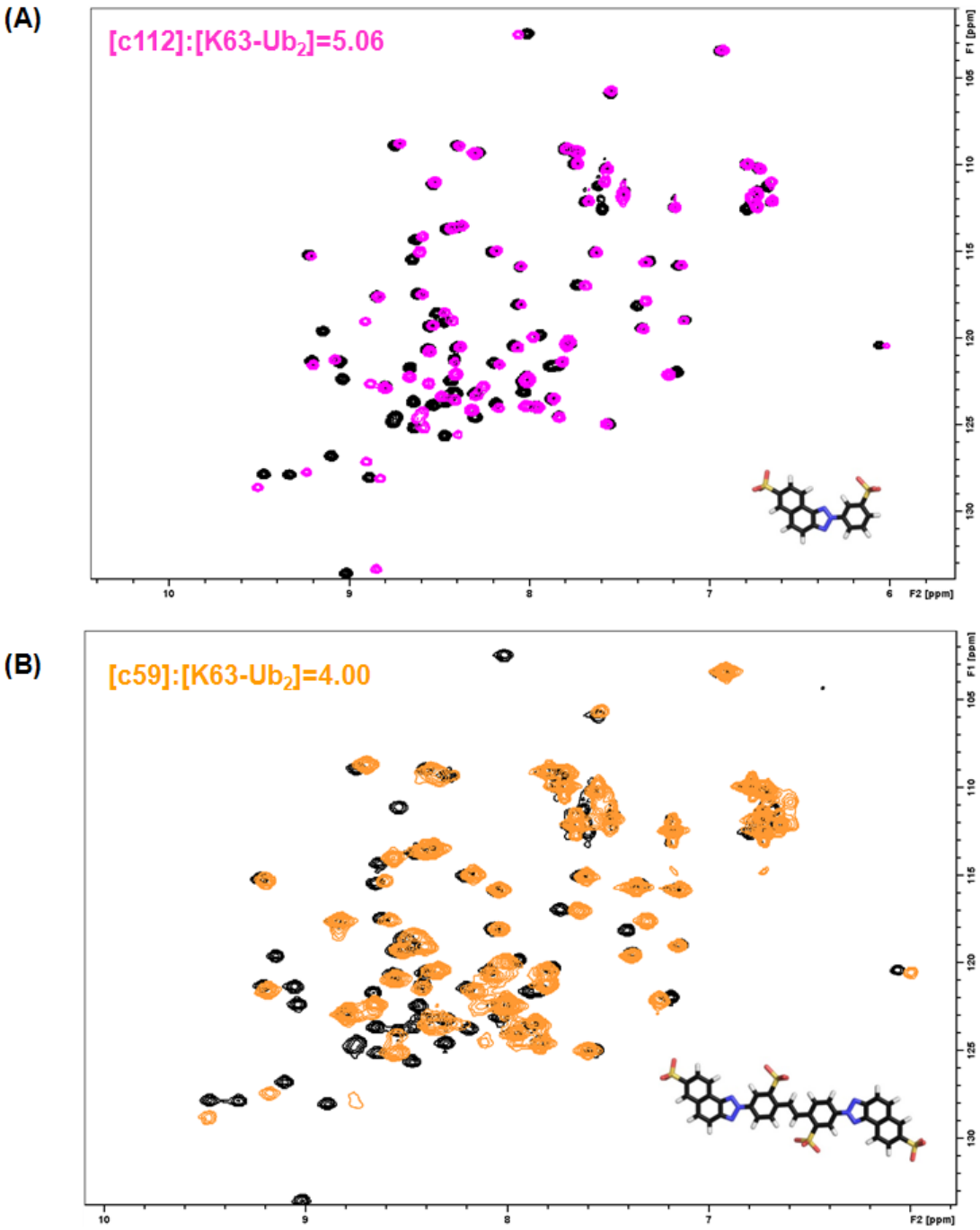


Figure 5.21 – $^{1}H,^{15}N$-HSQC titration of K63-Ub_2
Titration of K63-Ub_2 with both c59 and c112 produces residue specific shifts from binding. (**A**) Overlay of unbound K63-Ub_2 (black) with titration point [c112]:[K63-Ub_2]=4.8 (magenta) shows that signals retain sharp line shapes at saturation (**B**) The overlay of unbound K63-Ub_2 (black) with titration point [c59]:[K63-Ub_2]=4.00 (orange) shows many attenuated peaks indicative of tight binding.

To investigate the size of each complex ^{15}N T_1 was measured at the end point for each titration. The results (**Figure 5.22**) show trends similar to monomeric Ub. Interestingly, the complexes for K48-Ub_2 and K63-Ub_2 appear to be dramatically different for c59, but not with c112 or in the unbound form. This large difference could be due to c59 oligomerizing K63-Ub_2 such that one molecule creates an essentially tetrameric Ub (2 x K63-Ub_2) complex. This is interesting considering that c59 was reported to bind K48 linkages with a great affinity. The observation of the apparent difference in stoichiometry for the K48 vs. K63 linked di-Ub complexes, suggests a different binding mechanism for each linkage. This may explain how ubistatins discriminate between K48 and K63 linkages. I attribute this difference to the conformations each linkage presents for binding. As the CSPs show (**Figure 5.22**) ubistatins will bind the L8,I44,V70 hydrophobic patch, but the resulting complex could take on many forms depending on the conformation of the chain.

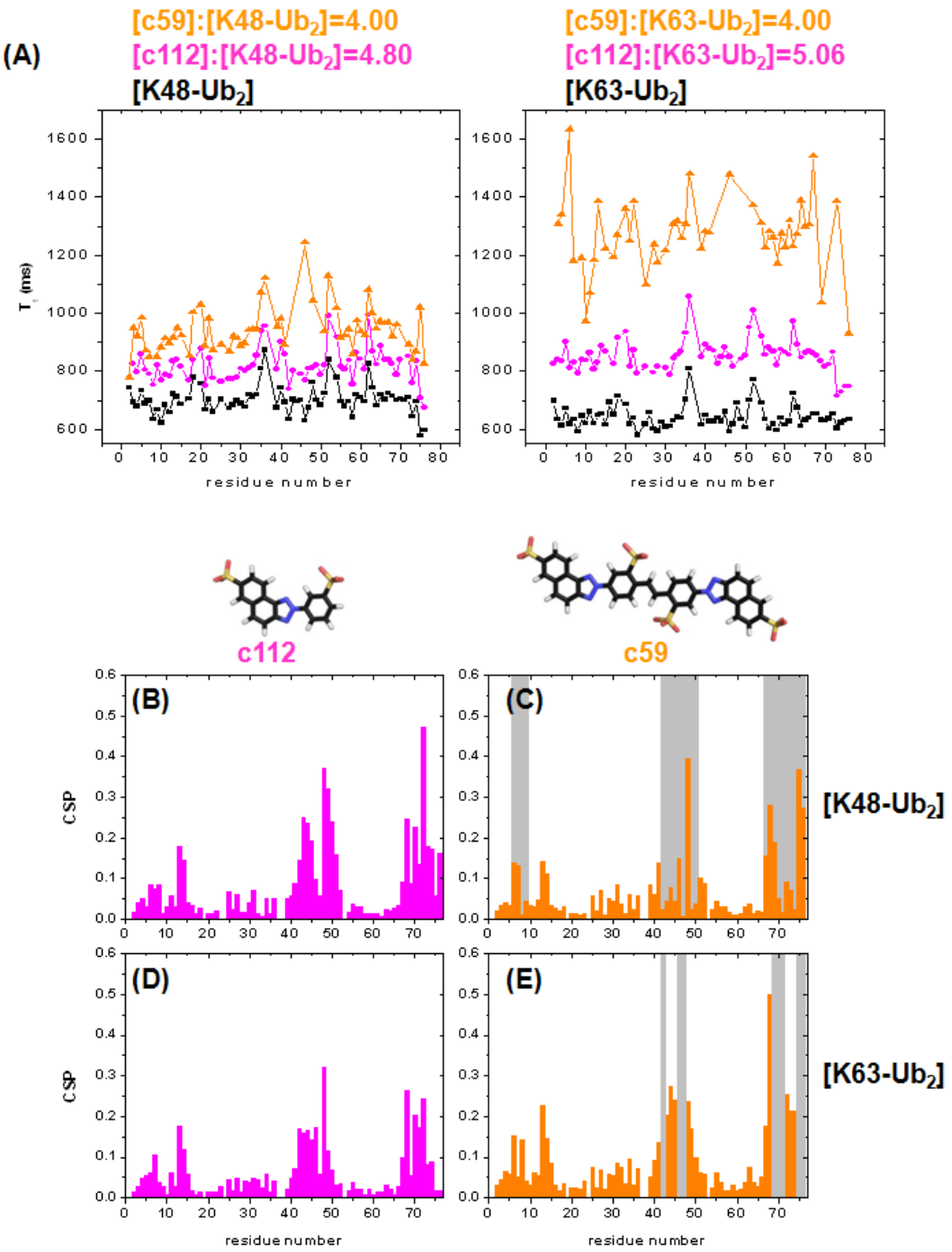


Figure 5.22 – Analysis of K48 and K63 interactions with ubistatins
(**A**) ^{15}N T_1 data measured at the endpoint of each titration is reported for the K48 linkage (left) and K63 linkage (right). T_1 values for unbound di-Ubs are in black, c112/di-Ub complexes are magenta, and c59/di-Ub complexes are in orange. (**B-E**) Reside specific CSPs for c112 (magenta) and c59 (orange) titration with K48-Ub_2 (top row), and K63-Ub_2 (bottom row). Signal attenuations are designated by gray bars.

5.4.4 Anionic Ub mutants disrupt ubistatin binding

Based on the CSP, NOESY, and T_1 relaxation data, there was no question that the L8,I44,V70 hydrophobic patch in Ub was necessary for the binding of ubistatins. However, the Rub1 titration which essentially showed no interaction, even though Rub1 contained the same L8,I44,V70 hydrophobic patch led me to search for other factors essential for ubistatin interactions. Neighboring residues were a logical place to search in addition to considering the properties of the ubistatins themselves. In c112 and c59, the protons in the ring systems must account for a majority of the hydrophobic interaction, however the substituent groups, all electronegative sulfonates were also shown to be important (see **Figure 5.8**). Understandably, I searched for cationic residues surrounding the hydrophobic patch and noticed that R42 and R74 are in very close proximity. Their significantly larger CSPs and attenuations compared to other cationic residues also supported their involvement. However, there was a pitfall in the detection of this ionic interaction that would presumably involve a sulfonate group from the ubistatin and a guanidinium group of either R42 or R72. The nuclei involved are arranged in such a way that they are essential invisible my NMR methods and the large size of these groups puts ^{1}H-^{1}H pairs outside of the 5Å range of NOE detection. With limited options, I created several Ub mutants to test the importance of R42 and R72 for ubistatin interaction. Ub(R42A) and Ub(R72A) were designed to determine if either hydrophobics or electrostatics were more important for binding, Ub(R42E) and Ub(R72E) were introduced to see if was possible to overcome the electrostatic interaction, and Ub(K63D) served as a control to ensure that changing the bulk charge of Ub did not

disrupt binding. Electrostatic potential maps are shown for several mutants (**Figure 5.23**).

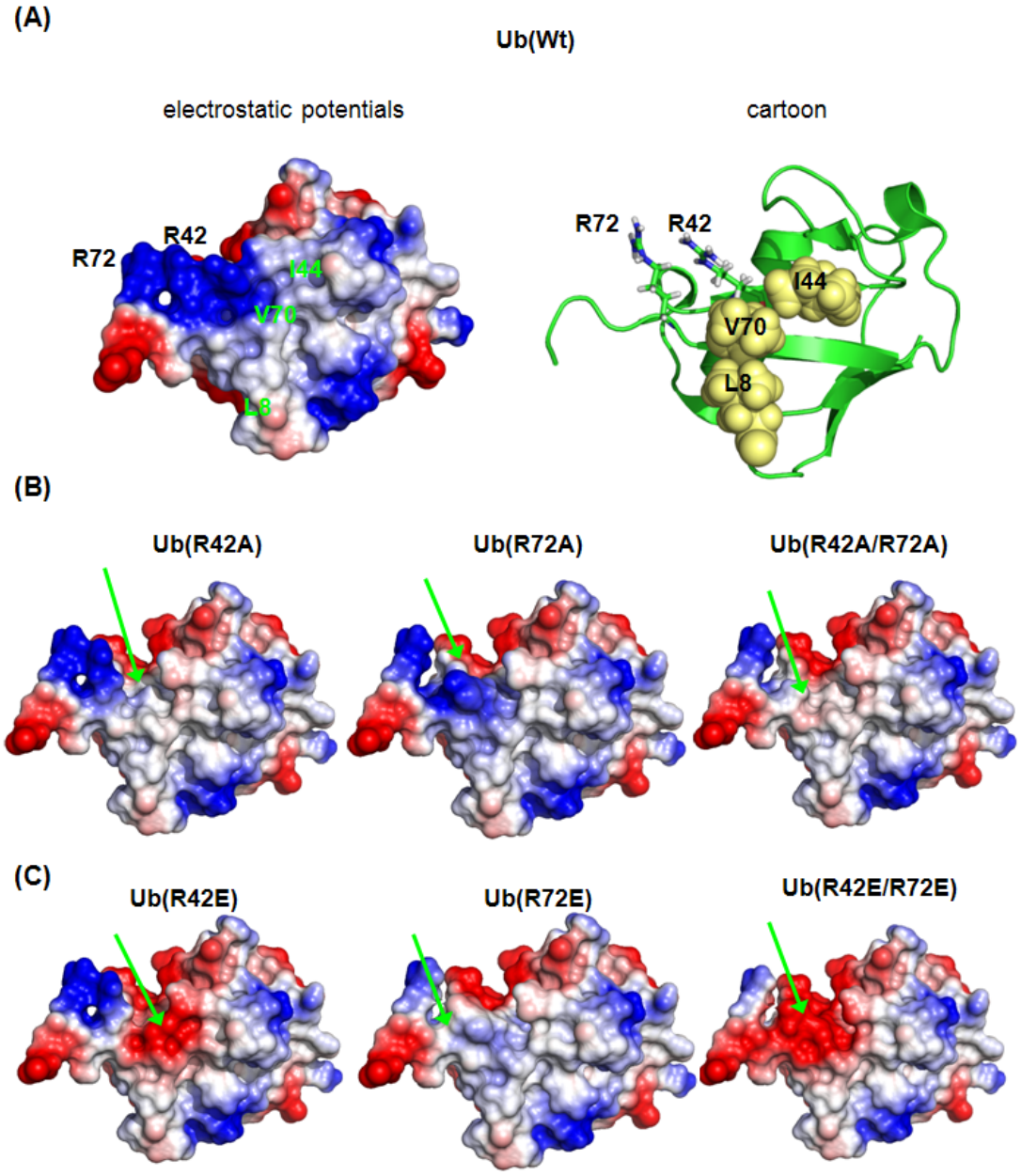


Figure 5.23 – Electrostatic potentials of Ub R42 and R72 mutants.
(**A**) Electrostatic potentials (blue=basic, white =hydrophobic, and red=acidic) shown on wild type Ub and cartoon representation with relevant features. (**B**) Alanine mutants of Ub used, note the change from basic to hydrophobic in the area indicated by the green arrow. (**C**) Glutamic acid mutants of Ub. The green arrows indicates a complete reversal in potential from basic to acidic in the double mutant, Ub(R42E/R72E).

$^{1}H,^{15}N$-HSQC titration of several alanine and glutamic acid Ub mutants revealed that c112 utilized significant hydrophobic contacts to bind, however electrostatic interactions were also important for binding Ub. The end point CSPs for Ub(R42A) and (R72A) are still follow the same profile as wild type Ub although they are smaller suggesting weaker binding (**Figure 5.24**). The Effect of glutamic acid mutants, Ub(R42E) and Ub(R72E) is more pronounced for each mutant, with Ub(R42E) showing a more notable impact due to its closer proximity to the L8,I44,V70 hydrophobic patch. When a the double mutant, Ub(R42E/R72E) is titrated with c112, binding is all but abolished (**Figure 5.24**). It appears that R42 and R72 are both in close proximity to c112 and that they form important electrostatic interactions as demonstrated in these titrations. This effect is localized and changing the bulk charge of Ub as shown with Ub(K63D) does not disrupt binding of c112. Over the course of the titration the effects of each mutant are shown for residue specific titration curves (**Figure 5.25**). In agreement with the endpoint CSPs, the titration curves show that the Ub(K63D) binds just as wild type, the alanine mutant still retains some binding, while the glutamic acid mutants abolish binding.

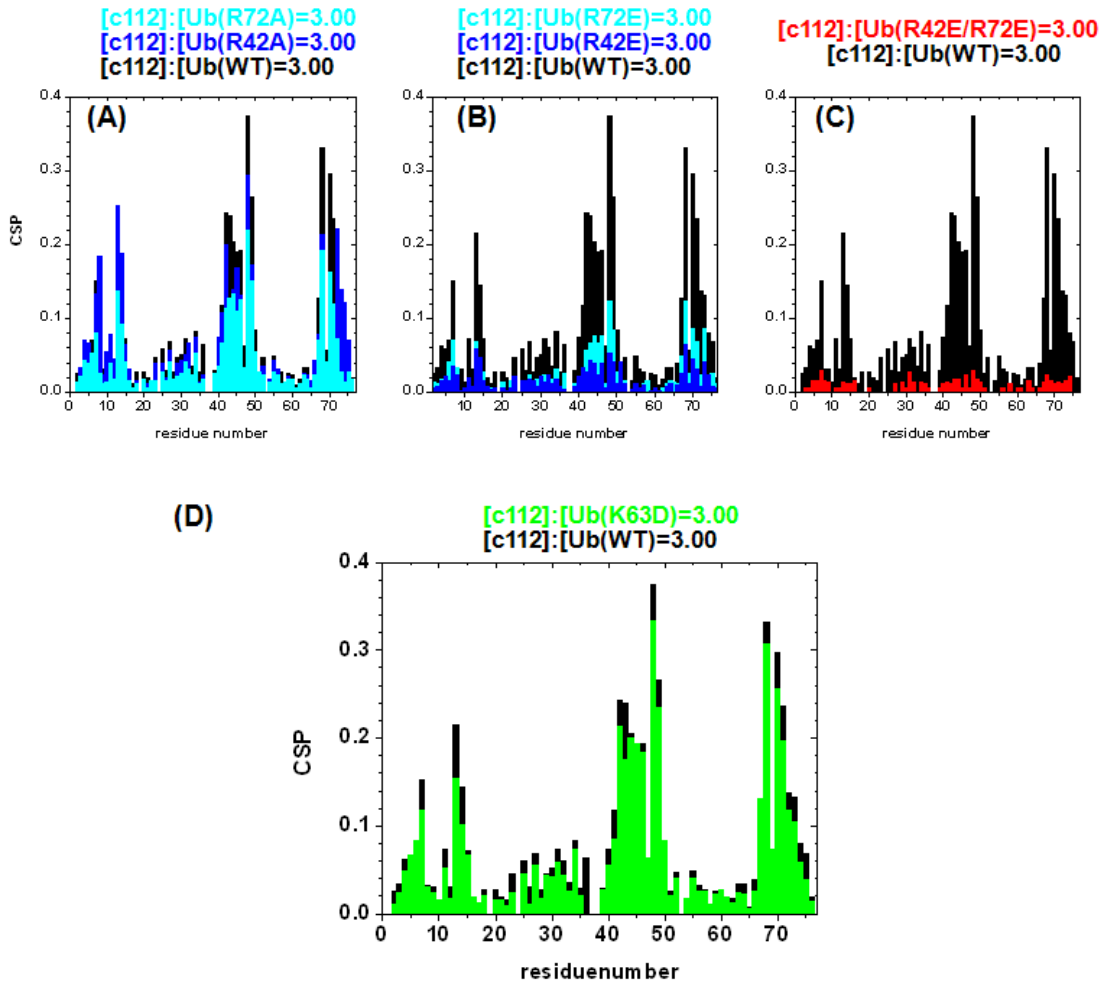


Figure 5.24 – CSPs demonstrate electrostatics interactions from R42 and R72 facilitate c112 binding.
(**A**) End point CSPs for Ub wild type (black) compared to alanine mutants Ub(R42A) (cyan), and Ub(R72A) (blue). (**B**) The glutamic acid mutants Ub(R42E) (cyan) and Ub(R72E) (blue) show diminished CSPs compared to wild type Ub (black). (**C**) The double mutant Ub(R42E/R72E) (red) has virtually no CSPs indicating that binding of c112 is abolished. (**D**) Ub(K63D) (green) containing the same bulk charge as the other glutamic acid mutants still retains the ability to bind c112.

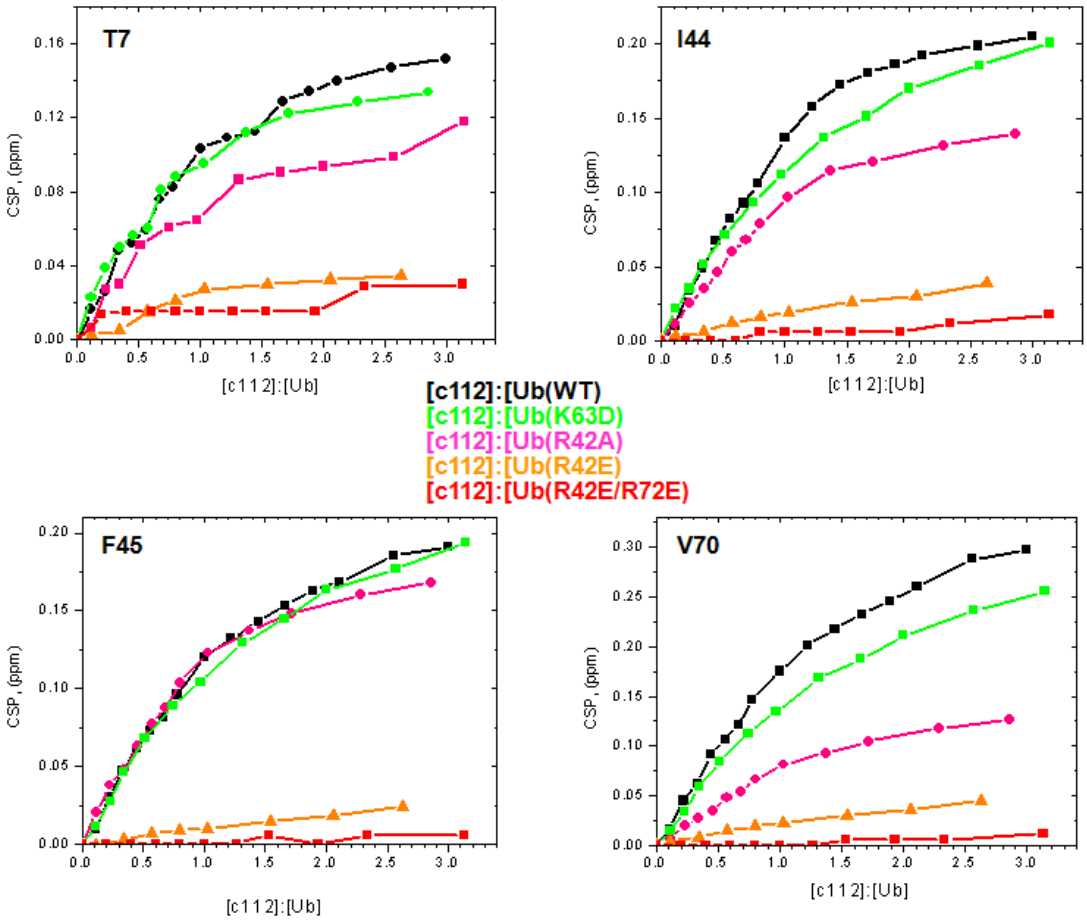


Figure 5.25 – Residue specific titrations curves monitor c112 binding.
The titration curves for residues T7, I44, F45, and V70 in Ub are plotted for each Ub variant: Ub(WT) black, Ub(K63D) green, Ub(R42A) pink, Ub(R42E) orange, and Ub(R42E/R72E) red.

With these results I tested if these same mutants would prohibit binding to c59, the full ubistatin. The results (**Figure 5.26**) show virtually no CSPs from binding and also, the c59/2Ub$_1$ complex does not form as the T$_1$ data shows at saturation shows, the values essentially match those of monomeric Ub.

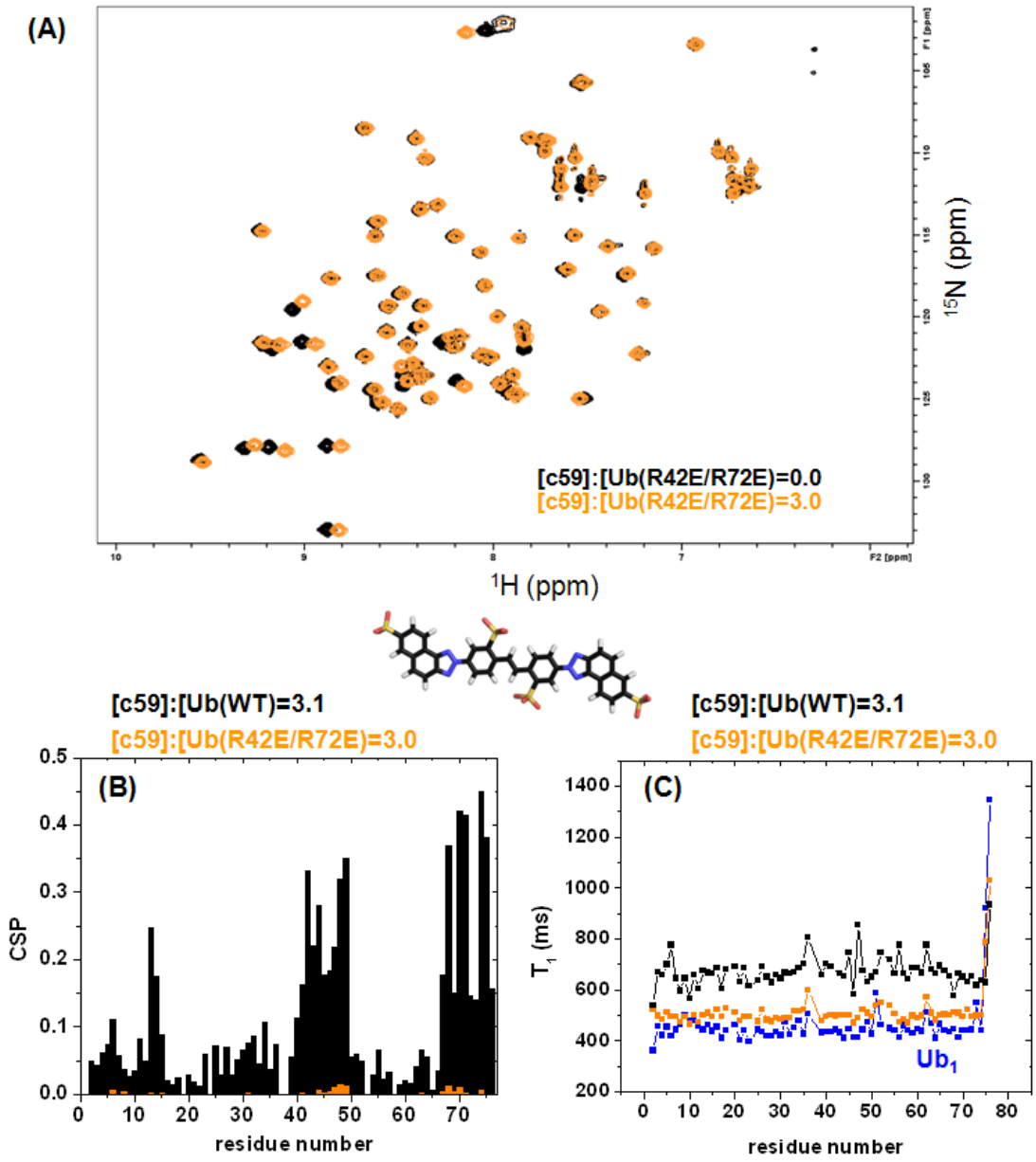


Figure 5.26 – Glutamic acid mutants abolish c59 binding to monomeric Ub. (**A**) ^{1}H,^{15}N-HSQC overlay of titration endpoint of [c59]:[Ub(R42E/R72E)]=3.0 in orange with unbound Ub(R42/R72E) in black. There are minimal CSPs and the peaks are narrow, unlike titration with wild type Ub$_1$. (**B**) Residue specific CSPs of [c59]:[Ub(R42E/R72E)]=3.0 in orange are barely visible compared to the CSPs from [c59]:[Ub(WT)]=3.1. (**C**) T_1 data for unbound Ub$_1$ (blue), [c59]:[Ub(R42E/R72E)]=3.0 (orange), and [c59]:[Ub(WT)]=3.1 (black). Note that Ub(R42E/R72E) does not experience a large increase in T_1 upon addition of c59.

5.4.5 Failure: crystallization of c59 and c112 Ub complexes

The extensive solution NMR studies provided much useful information on several ubistatins. However, the nature of the c59 compound with symmetric unassignable protons and its ability to form large complexes with Ub resulting in broadened NMR line widths made structural determination impossible with NMR methods. In a desperate attempt to determine a structure, crystal screening was performed on c112 and c59 in complex with monomeric Ub, K48-Ub_2, and K63-Ub_2. Six different 96-condition screens were attempted for each of the six complexes, however none produced the desired crystals. In the case of K48-Ub_2, the dimer often crystallized without the ubistatin, which was discovered upon diffraction.

5.5 Structural determination and validation of the c112/Ub_1 complex

5.5.1 Assignment and analysis of intermolecular NOEs

Using a simple ^{15}N filtered 2D $^{1}H,^{1}H$-NOESY, NOEs between c112 and mono-Ub were detected. To assign these intermolecular NOEs between ^{1}H-^{1}H pairs, a battery of experiments on $^{13}C/^{15}N$-Ub in complex with c112 were performed. First, the individual signals for protons in bound c112 were assigned in an approach similar to the assignment of free c112 (**see Figure 5.5**). For these experiments, the $^{1}H,^{1}H$-TOCSY and $^{1}H,^{1}H$-NOESY experiments filtered any ^{1}H attached to ^{13}C or ^{15}N, leaving only signals from c112. To ensure only NOEs between c112 and Ub were detected; a variation of NOESY that only allows for NOEs between ^{12}C-^{1}H and ^{13}C-^{1}H pairs was used. This experiment resulted in the ^{1}H and ^{13}C chemical shifts for atoms in Ub that had NOEs to c112 (**Figure 5.27**).

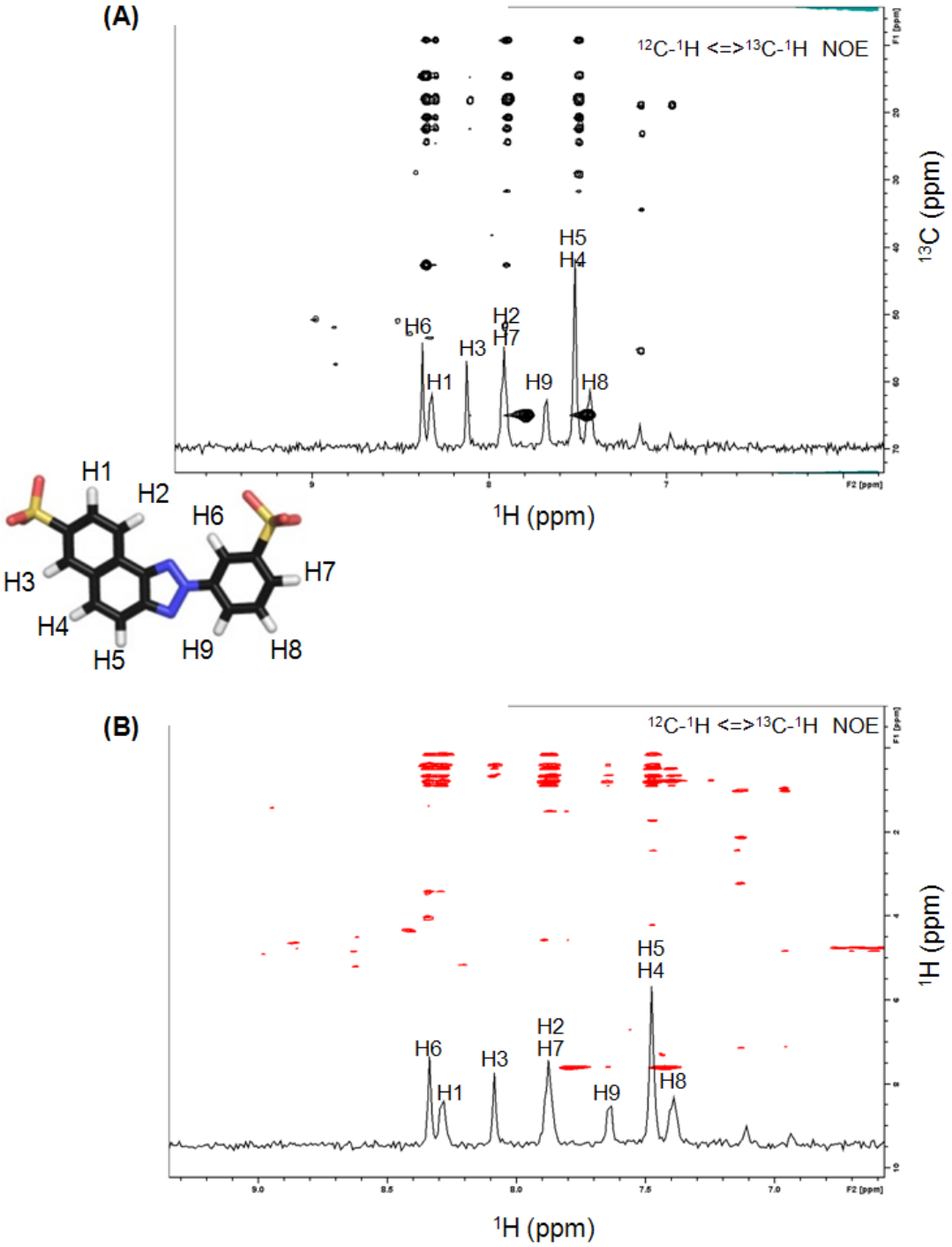


Figure 5.27 – Intermolecular NOEs between c112 and Ub.
Variation of 3D-NOESY with filter to only allow NOEs between ^{12}C-^{1}H and ^{13}C-^{1}H. The 1D ^{1}H spectra for c112 in the bound state are shown to illustrate how the signals observed in the NOESY towers for each of the protons in c112 are assigned to protons in Ub (**A**) The 2D ^{1}H,^{1}H plane is shown and (**C**) 2D ^{1}H,^{13}C plane is shown. Collectively each plane gives the spectral coordinates for an H-C bond in ^{1}H,^{13}C-HSQC.

Using a combination of 3D ^{1}H-CCCONH TOCSY and 2D ^{13}C, ^{1}H-HSQC signals for Ub were assigned in the c112 bound state of Ub. With proper assignment, NOEs could be quantified in the 2D ^{1}H, ^{1}H-NOESY. Intramolecular NOEs in c112 and also Ub of known distances served for calibration. There were some concerns about spectral overlap for protons in c112, but with much help from Prof. Olivier Walker we determined the following distances from NOEs (**Table 5.1**).

NOE distances			
Ub		c112	Distance*
residue	atom	atom	
I44	Hδ	H2	3.5 Å
I44	Hδ	H3	3.6 Å
I44	Hδ	H1	3.0 Å
I44	Hδ	H6	2.8 Å
I44	Hγ	H6	2.4 Å
I44	Hγ	H1	2.9 Å
I44	Hγ	H3	3.2 Å
I44	Hγ	H9	4.0 Å
V70	Hγ	H2	3.4 Å
V70	Hγ	H3	3.4 Å
V70	Hγ	H1	2.8 Å
V70	Hγ	H6	2.8 Å
R42	Hγ	H6	5.5 Å
H68	Hβ	H5	3.3 Å
G47	$H\alpha_1$	H6	3.1 Å
G47	$H\alpha_1$	H7	5.0 Å
G47	$H\alpha_2$	H6	2.9 Å
G47	$H\alpha_2$	H7	2.9 Å

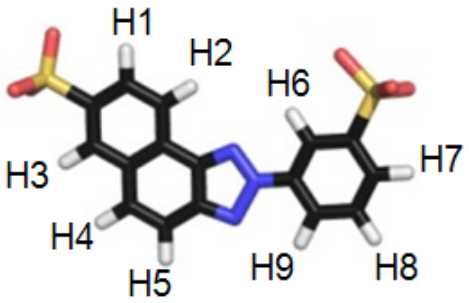


Table 5.1 – Intermolecular NOEs between Ub and c112.
NOE distance calculated from internal standards.
*For each distance an error of ±2.0 Å was assumed for docking.

5.5.2 Long range distance constraints from PREs

The NOE data were encouraging and certainly gave us a starting point for structural calculations. However, in such short distances there could be error and also it is possible that c112 can be oriented several different ways on Ub's hydrophobic patch. To better understand how c112 was oriented we designed several experiments for use with MTSL as paramagnetic probe. With several NOE based structural models of the c112/Ub complex, we simulated how different sites of attachment for MTSL on Ub would affect individual protons in c112 (**Figure 5.28**). We proceeded with several Ub cysteine mutants. The position of the paramagnetic center was fit using changes in intensity of Ub signals in $^{1}H,^{15}N$-HSQC and distances to individual proton in c112 were detected with 1D ^{1}-NOESY with a ^{15}N-filter. Several positions on Ub (T12, I36, K48 and K63) were tested for MSTSL attachment (**see Figure 5.29**). As expected, the control with MTSL at residue 63 did not produce any quantifiable attenuations in c112 due to the large distance between the unpaired electron in MTSL and nuclei in c112. The data for MTSL at the other positions was valuable for determining exactly how c112 was oriented on Ub's hydrophobic patch.

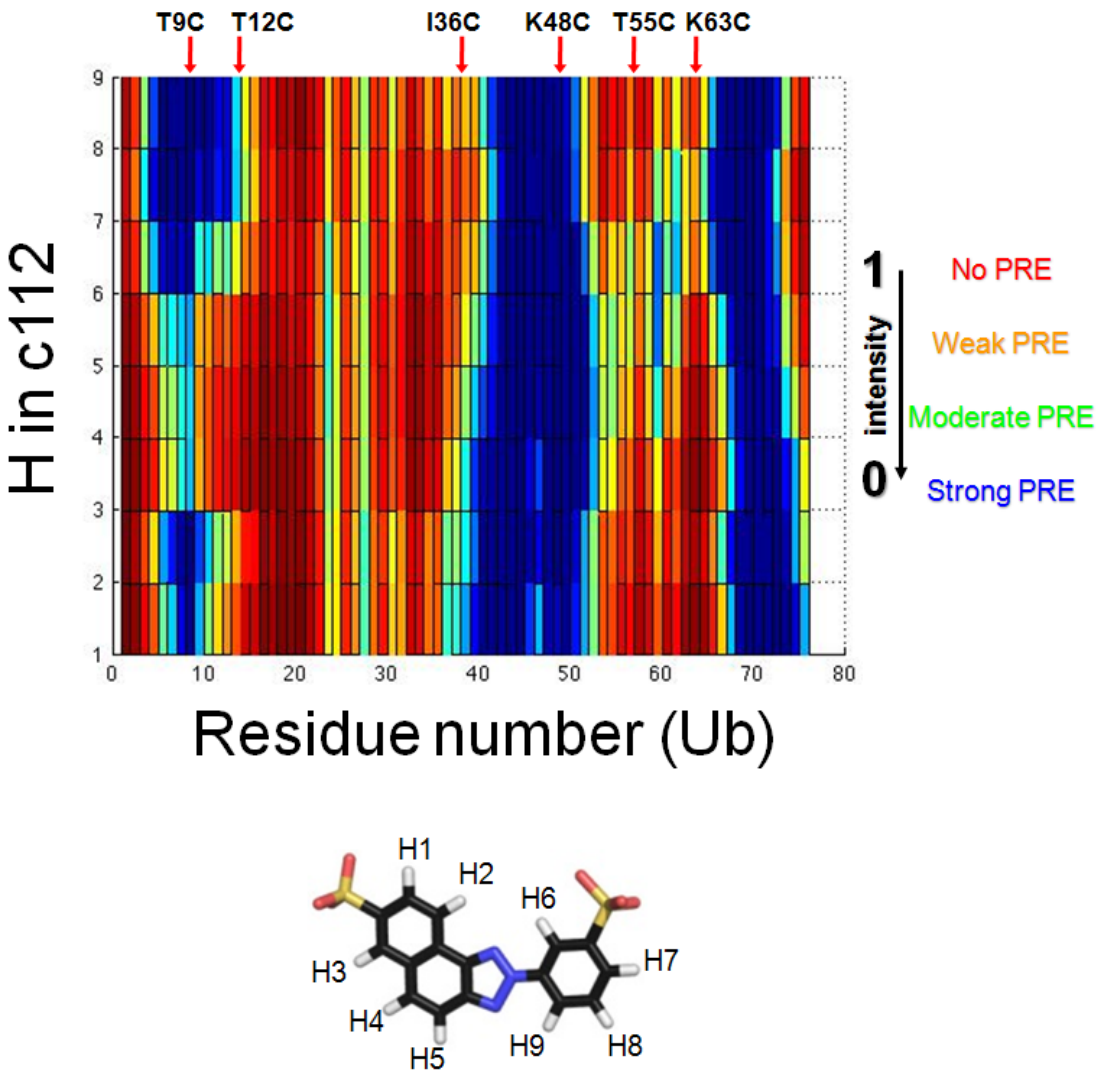


Figure 5.28 – Simulation of MTSL attachment at every residue in Ub
To determine the best position for MTSL attachment, the effect of MTSL positioned on every residue of Ub was simulated for every proton in c112. Sections in blue such as at residue number 48 indicate all signals from c112 would be attenuated, while all red regions e.g. residue number 20 are too far away from c112 to have an effect.

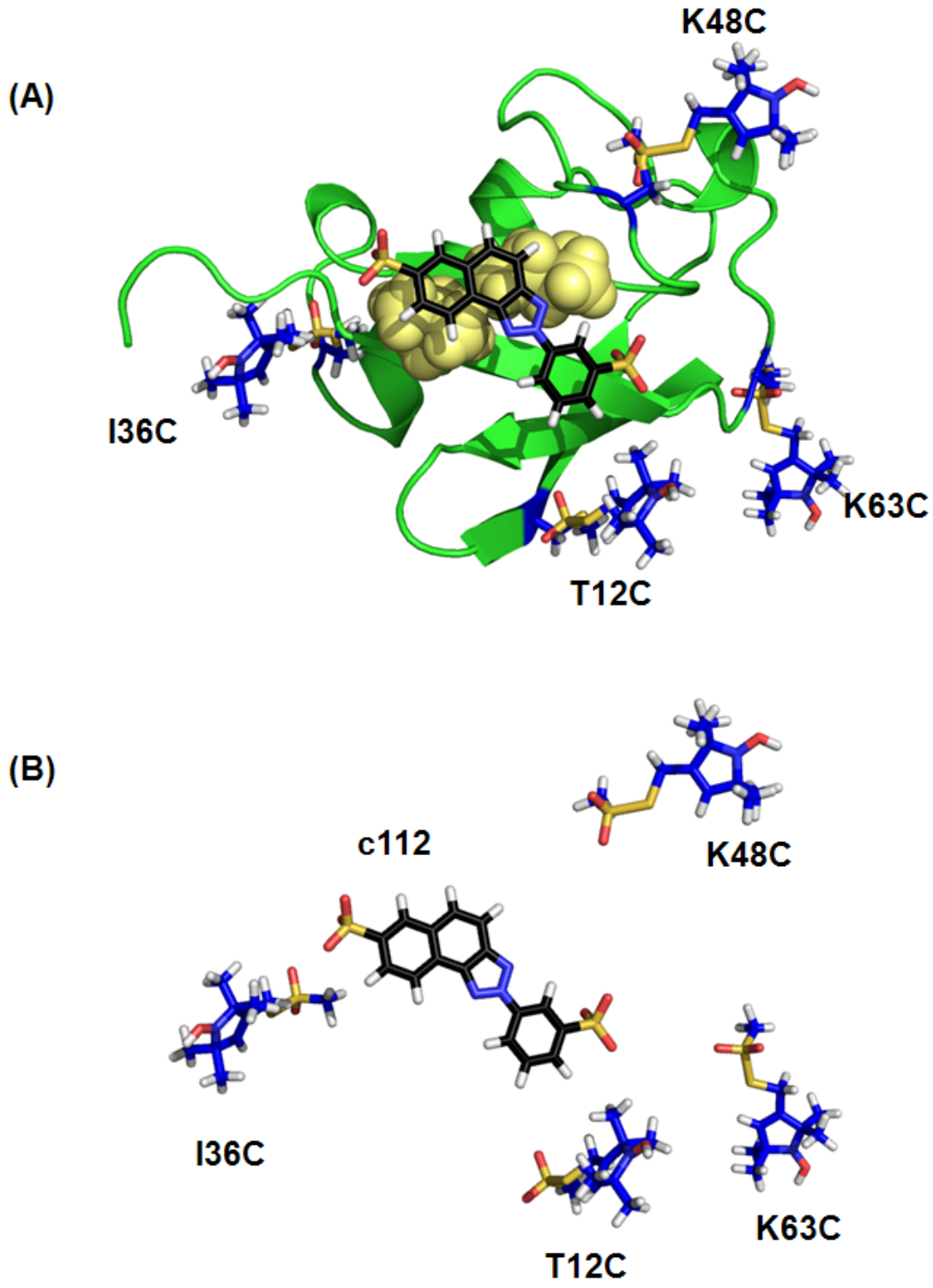


Figure 5.29 – MTSL positions on Ub used
(**A**) The attachment of MTSL to Ub mutants T12C, I36C, K48C, and K63C is shown on the surface of Ub with c112 modeled using NOEs from I44 and V70. (**B**) The Ub is removed and just the position of MTSL and c112 are shown. These positions were chosen deliberately to probe c112 from different angles. K63, far removed is the negative control and did not produce a detectable effect on protons in c112.

The details of the PRE experiments are described in (**section 3.4.5**). After evaluating the MTSL data, the following distances were taken for docking (**Table 5.2**)

Long Range PRE Distances			
Ub		c112	
MTSL Residue	atom	atom	Distance*
48	O (MTSL)	H1	13.3 Å
48	O (MTSL)	H2	12.0 Å
48	O (MTSL)	H3	11.9 Å
48	O (MTSL)	H4	10.5 Å
48	O (MTSL)	H5	9.5 Å
48	O (MTSL)	H6	10.0 Å
48	O (MTSL)	H7	12.1 Å
48	O (MTSL)	H8	12.3 Å
48	O (MTSL)	H9	11.4 Å
36	O (MTSL)	H1	14.0 Å
36	O (MTSL)	H2	14.8 Å
36	O (MTSL)	H3	15.7 Å
36	O (MTSL)	H4	16.8 Å
36	O (MTSL)	H5	17.7 Å
36	O (MTSL)	H6	18.3 Å
36	O (MTSL)	H7	18.9 Å
36	O (MTSL)	H8	17.5 Å
36	O (MTSL)	H9	16.5 Å

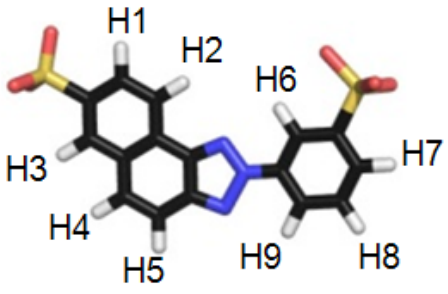


Table 5.2 – PRE derived distances to protons in c112
*For docking an error of ±4.0 Å was assumed to account for flexibility in MTSL and represented an appropriate tolerance range.

5.5.3 Structural calculation of c112

With the NOE and PRE derived distances, we initiated structural calculations for the c112/Ub_1 complex. The additional knowledge of the stoichiometry and along with the importance of cationic residues R42 and R72 also help to guide docking. For this docking step and many of the previous steps in distance calculation, Prof. Olivier Walker's expertise was invaluable. After performing many runs with HADDOCK (see **section 3.5.2** for details), the results were evaluated and for the first time a high resolution structure of an ubistatin/Ub complex was determined (**Figure 5.30**). Our structure satisfied all of the distances that were calculated from NOEs and PREs, in addition to including the electrostatic interaction between R42 and a sulfonate of c112. With this structure as a template we were also able to model how the full ubistatin, c59 would look if superimposed on the c112/Ub_1 complex. We find that if just bound to one Ub, one half of c59 is left hanging off the binding surface, allowing a second Ub to bind. This explains how c59 forms a c59/$2Ub_1$ complex with monomeric Ub, but the actual structure may deviate slightly from our models (**Figure 5.31**).

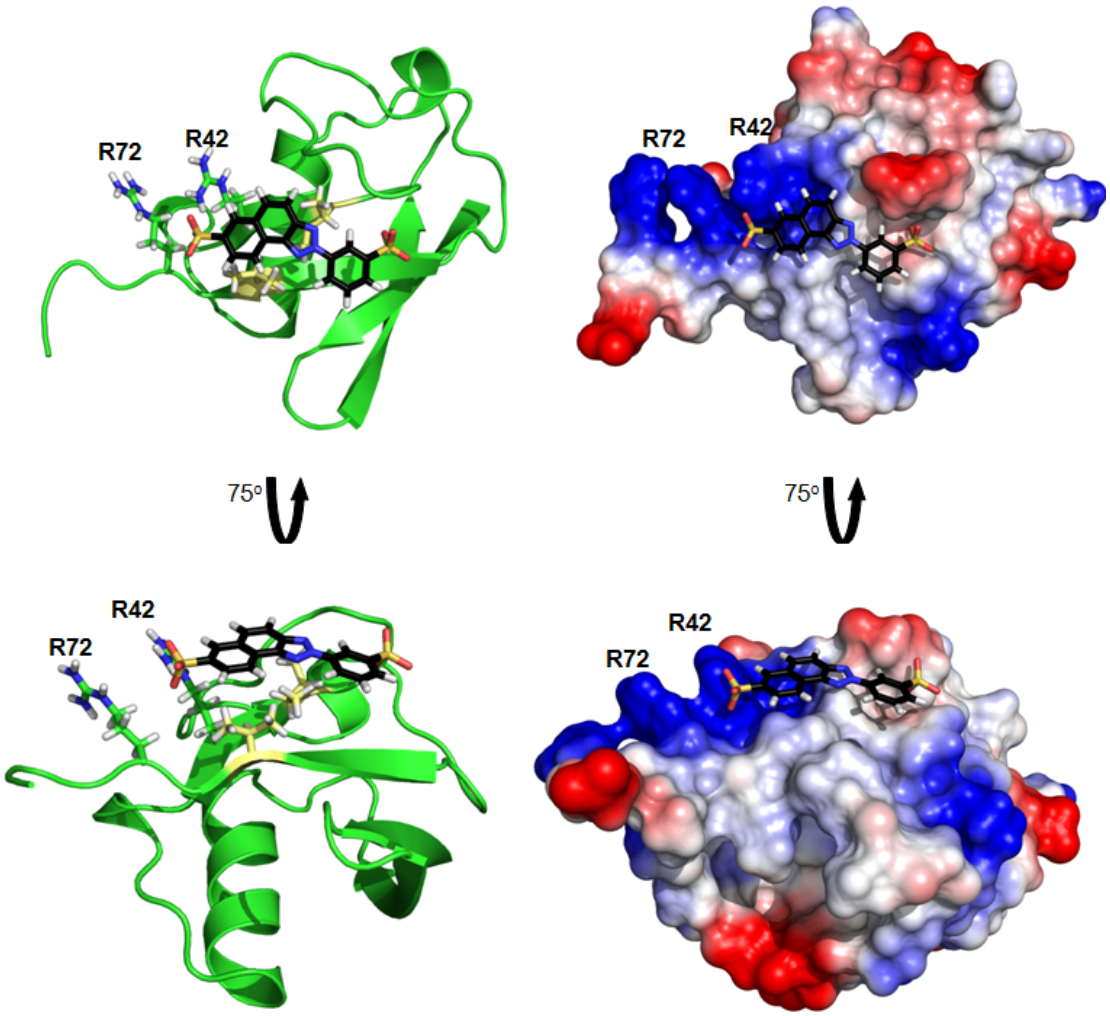


Figure 5.30 – Solution NMR structure of c112/Ub$_1$
Cartoon and electrostatic representations of Ub with c112 (black sticks). Cationic R42 and R72 are in close proximity to a sulfonate of c112. As the NOEs show, the main ring of c112 is positioned directly over I44 and V70 of the hydrophobic patch in Ub. The bottom view shows c112 fits in nicely to a cleft on the surface of Ub.
* Structure calculated in collaboration with Prof. Olivier Walker

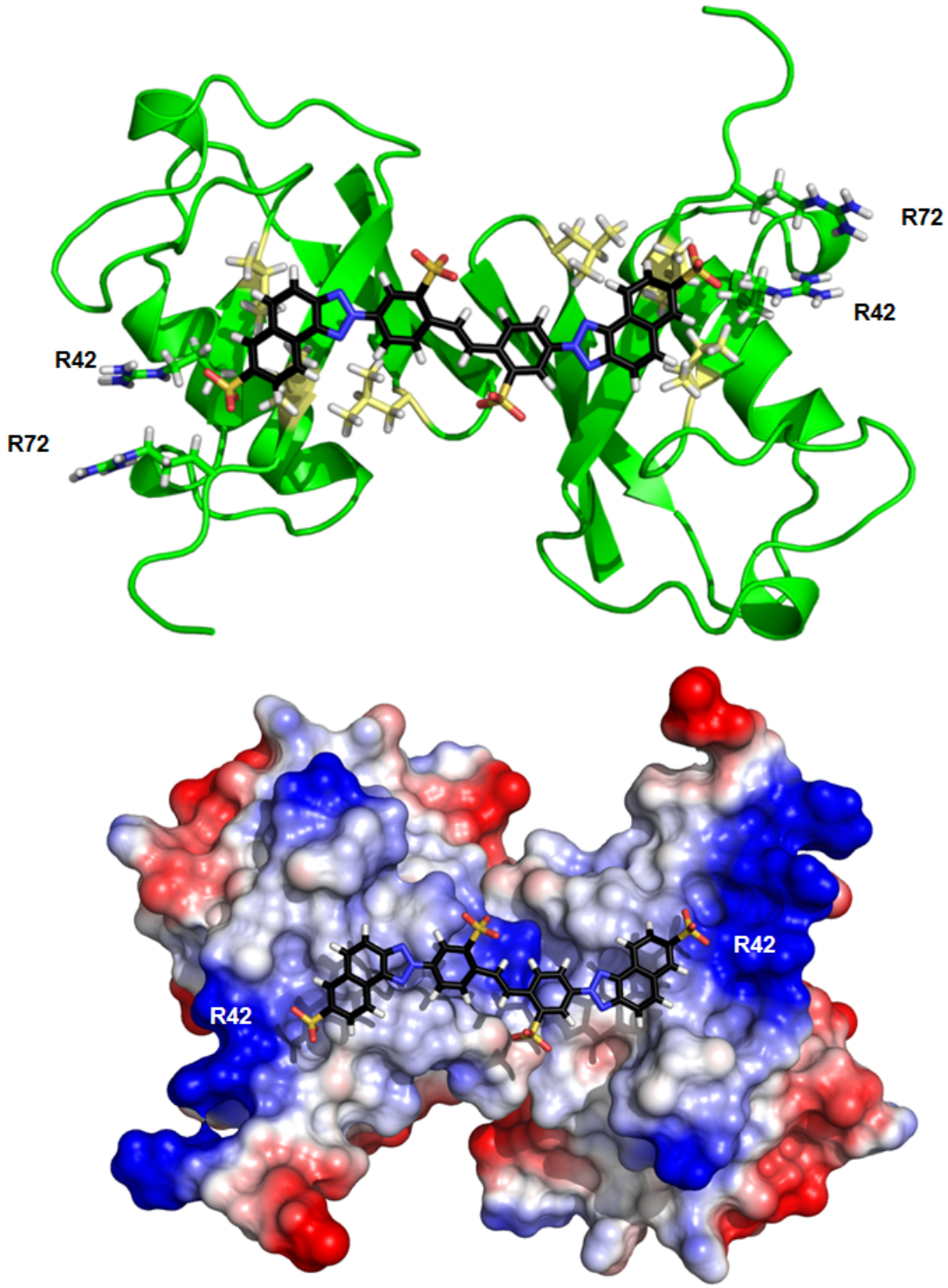


Figure 5.31 – Structural model of the c59/Ub$_1$ complex
Using what was learned from the titration data and c112/Ub$_1$ structure, the proposed model for the c59/Ub$_1$ complex is shown with at c59:Ub$_1$ stoichiometry, 1:2. The size of c59 allows two Ubs to simultaneously bind.
* Structure calculated in collaboration with Prof. Olivier Walker

5.6 Discussion and summary

The idea of designing small molecules to bind to Ub is an exciting concept. Targeting the surface of a non-catalytic protein is also an interesting strategy and the *in vivo* data supports that ubistatin compounds are already highly effective, thus validating the strategy. Until this work virtually nothing was known about how ubistatins bound Ub. Through extensive NMR experiments, I have demonstrated that ubistatins always bind the L8,I44,V70 hydrophobic patch of Ub, regardless of the linkage type. In collaboration with others we conclude that the sulfonate groups are necessary for proper Ub binding. The observations that c112 can discriminate between Rub1 and Ub, while ANS does not bind Ub are particularly encouraging and highlights the specificity ubistatins have achieved. The solution structure obtained for the c112/Ub_1 complex gives many paths for further development of these molecules. Even if c59 is an already potent drug, it does have an inherent problem of membrane permeability (*224*). However, our structural models would allow for the chemical addition of specialized groups to c59 and c112 to overcome the limitation. This study represents a first step into the structural investigation of ubistatins, yet it was disappointing that neither a solution or crystal structure could be solved for any c59 or polyUb complexes. It may be that the nature of ubistatin/Ub complexes simply do not favor crystallization. Nevertheless much was learned from ubistatin titrations of polyUb chains using solution NMR. The finding that ubistatins also inhibit DUBs provides a new explanation of how these molecules interfere with cellular machinery. There is much more work to be done on ubistatin compounds and also for my study to be expanded on, especially the structural aspects. Perhaps others

will design different molecules (e.g. gold nanoparticles) to bind Ub, but at the present time ubistatins are the best characterized Ub binding small molecules and certainly the most proven. With the proper structural information, the same concept of ubistatins targeting Ub could be attempted for small molecules targeting UBLs. From a drug design point of view, ubistatins may produce an unacceptable number of off target effects, considering they inhibit all polyUb signaling. However, proteasome inhibitors arrest the UPP and also disrupt a large number of cellular pathways. Given that ubistatins have been shown to arrest the cell cycle, producing effects similar to proteasome inhibitors they are certainly worth pursuing. If ubistatins do fail as drugs then they could simply be used as reagents for *in vitro* studies as they have been used before (*216*). With the proper design, it is not likely that ubistatins will fail as drugs and the lessons learned from our structural studies with ubistatins could be applied to inhibit many other signaling proteins.

Chapter 6: A breakthrough: proteomic methods for the analysis of K63-linked polyubiquitin conjugates

6.1 Background and research aims

Tryptic digestion of Ub conjugates followed by mass spectrometry (MS) analysis has been the traditional approach to Ub proteomics for the past few decades (*225*). The hallmark of this approach is that cleavage at R74, leaves a very distinct di-glycine (GG) motif attached to Ub modified proteins. Conveniently, tryptic cleavage isolates all Ub-Ub linkage sites on Ub, which has enabled researchers to quantify the abundance of individual Ub-Ub linkages from whole cell extracts and also pinpoint ubiquitination sites in the proteome (*226*). However, even with the success of the tryptic cleavage approach, critical information is still unattainable with the method.

Traditional tryptic analysis of the Ub modification can be rather ambiguous due to the fact that the GG tag on a lysine is not definitive evidence of ubiquitination. This modification could easily be the result of UBL proteins e.g. Rub1 (Nedd8 in humans) and ISG15 that would also leave the same GG tag following tryptic cleavage. Thus, it is never possible to state with certainty if the GG tag originated from Ub or a number of UBL proteins. Furthermore, trypsin cleaves Ub in many different places producing many small fragments, which could be unfavorable for some MS applications. When other methods such as, Ub-Ub linkage specific antibodies are used in collaboration with tryptic cleavage it is possible to unambiguously determine sites of ubiquitination and identify the Ub-Ub linkages (e.g. K48 and K33) present. Without extensive supporting experiments, tryptic

cleavage alone cannot prove polyubiquitination at a given site or differentiate between mixed and branched Ub-Ub linkages. However, using an alternate cleavage method specific for aspartic acid residues, I have demonstrated that it is possible to measure the number of consecutive K63-linkages in free or substrate attached polyUb conjugates (*227*). Notably, my method is the first MS base approach, which is feasible for proteomic analysis. In addition it is also the first instance where a polyUb chain greater than two Ubs has been "measured." We have demonstrated this approach on both unanchored (free) K63 linked polyUb chains and K63 modified substrates (*228*). The main aim of this alternate cleavage method is to overcome the limitations of tryptic cleavage and advance what we can learn about Ub signaling using proteomics. This method does indeed give us an unprecedented view of K63-linked Ub conjugates. All MS experiments and analysis presented in this chapter were carried out by Joseph R. Cannon in the lab of Prof. Catherine Fenselau.

6.2 Residue specific cleavage of ubiquitin

6.2.1 Cleavage of Ub by known proteases

The serine protease trypsin will selectively cleave bulky cationic residues, lysine and arginine at eleven cleavage sites approximately evenly distributed across the sequence of monomeric Ub (K6, K11, K27, K29, K33, R42, K48, R54, K63, R72, and R74). A more selective serine protease, Lys-C is highly specific for lysine residues will only cleave seven times at each of the seven lysines in Ub. However, both trypsin and Lys-C have virtually no activity for modified lysines, whether the modification is acetylation, glycosylation, or ubiquitination. The arginine specific protease, Arg-C produces cleavage products of little interest in Ub which are mainly

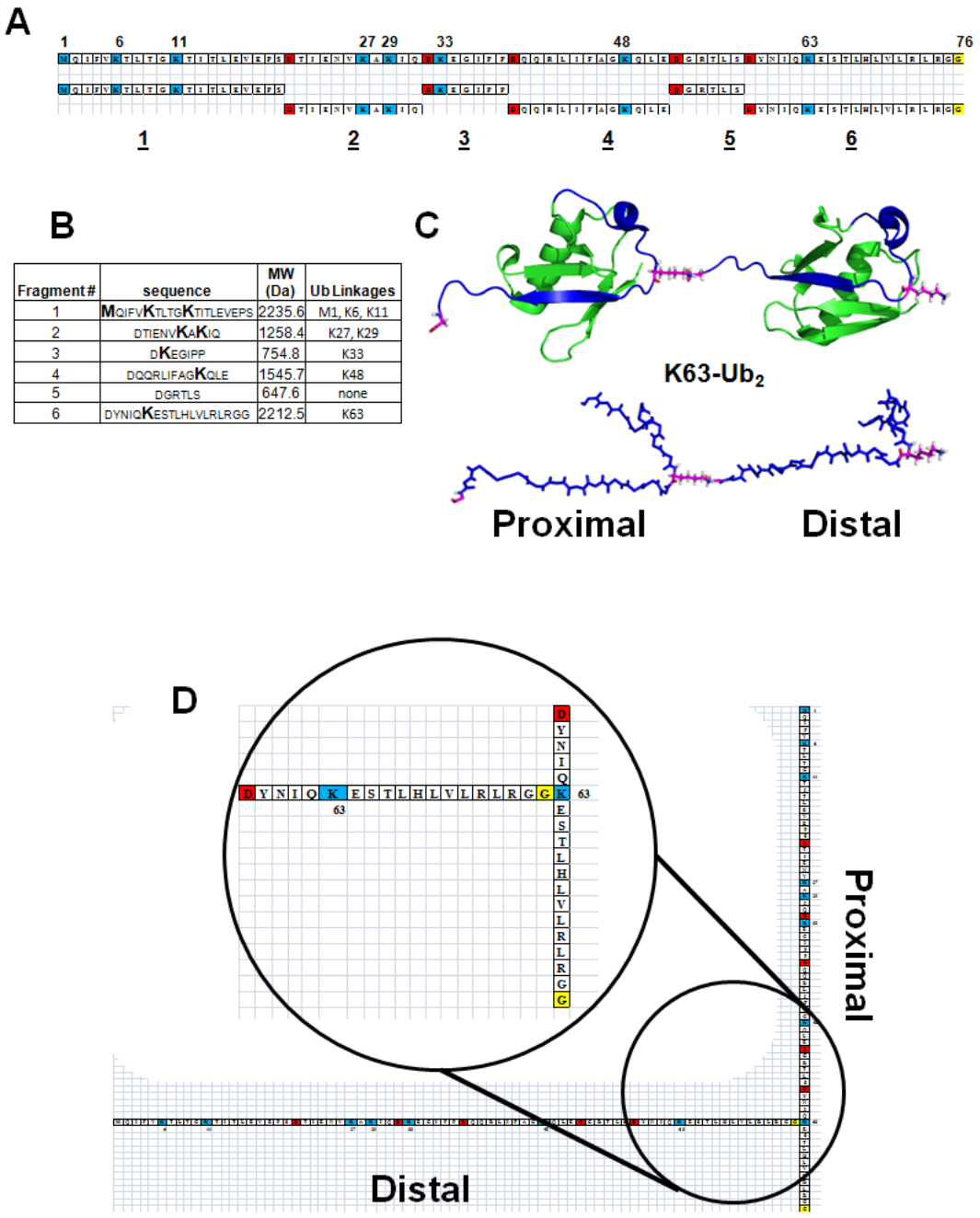
located near the end of the sequence. Tryptic and Lys-C cleavage are very effective in that they both help separate the lysine residues in Ub's sequence. However, lysine residues are prone to modification and this will more than likely reduce the usefulness of both trypsin and Lys-C. Alternate cleavage at residues, which are generally never modified can help get around this obstacle. Aspartic acid specific cleavage with Asp-N cleaves the five Asp residues in the sequence of Ub, which aids in separating the eight linkage sites on Ub regardless of how they are modified. The Asp-N digestion fragments of Ub will contain between one and three linkage sites making downstream analysis more manageable. As an alternative to Asp-N, chemical methods such as microwave assisted acid hydrolysis (MAAH) provide another means to achieve aspartic acid specific cleavage. In MAAH, an acid such as acetic acid or citric acid is placed in solution with the protein (~12% v/v), then heating in a specialized microwave catalyzes the selective cleavage of Asp residues resulting in the desired cleavage products. The resulting cleavage products of Ub with common proteases are shown (**Table 6.1**).

Name of Enzyme	# of Cleavages	Cleave sites in Ub
CNBr	1	1
Chymotrypsin-high specificity (C-term to [FYW], not before P)	**3**	**4 45 59**
Arg-C proteinase	4	42 54 72 74
Clostripain	4	42 54 72 74
Asp-N endopeptidase	**5**	**20 31 38 51 57**
Formic acid	5	21 32 39 52 58
Glutamyl endopeptidase	6	16 18 24 34 51 64
Staphylococcal peptidase I	6	16 18 24 34 51 64
LysC	7	6 11 27 29 33 48 63
LysN	7	5 10 26 28 32 47 62
Asp-N endopeptidase + N-terminal Glu	11	15 17 20 23 31 33 38 50 51 57 63
Trypsin	**11**	**6 11 27 29 33 42 48 54 63 72 74**
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	14	1 4 8 15 43 45 50 56 59 67 68 69 71 73
Pepsin (pH>2)	15	3 4 7 14 15 43 45 49 55 66 67 68 69 71 73
Pepsin (pH1.3)	17	3 4 7 14 15 43 45 49 55 58 59 66 67 68 69 71 73
Thermolysin	22	2 3 4 7 12 14 22 25 27 29 42 43 44 45 49 55 60 66 68 69 70 72
Proteinase K	38	3 4 5 7 8 9 12 13 14 15 16 17 18 22 23 24 26 28 30 34 36 43 44 45 46 50 51 55 56 59 61 64 66 67 69 70 71 73

Table 6.1 – Cleavage products of Ub by known proteases.
Note the number of cleave sites and their position in Ub. This initial analysis allowed me to select the appropriate protease for further MS studies with Ub.

6.2.2 Making the proper cuts in Ub

One of the greatest disadvantages of proteolysis of polyUb chains is that the exact sequence of the chain is destroyed such that it yields little information. The two Asp specific cleavage methods above present a special case for the K63 linkage. In principle, since Ub is always liked via the C-terminus to any of eight linkage sites on another Ub, cleaving such that the C-terminal residues and linkage site are connected will preserve the chain. Conveniently, D58 is the last residue in Ub's sequence to be cleaved by Asp specific methods. This cleavage product contains the K63 linkage site as well as the C-terminus of Ub. With this cleavage method any K63 linked polyUb conjugate (unanchored or substrate attached) will retain every consecutive K63 linkage. If a chain is anchored to a substrate, the proximal Ub will also include a fragment of the substrate. Another notable outcome from Asp specific cleavage is that other linkage sites are isolated; M1, K6, K11 are on the first fragment, K27 and K29 are on the second, K33 is isolated on the third, K48 is isolated on the fourth, and K63 is isolated on the sixth (the fifth fragment contains no linkage site). See (**Figure 6.1**). In addition to the applications with K63 linkages this also allows Asp cleavage to reveal the nature of K33 modifications. For Asp fragments that contain more than one linkage site, these masses can be used to determine if there is any branching.



Fragment #	sequence	MW (Da)	Ub Linkages
1	MQIFVKTLTGKTITLEVEPS	2235.6	M1, K6, K11
2	DTIENVKAKIQ	1258.4	K27, K29
3	DKEGIPP	754.8	K33
4	DQQRLIFAGKQLE	1545.7	K48
5	DGRTLS	647.6	none
6	DYNIQKESTLHLVLRLRGG	2212.5	K63

Figure 6.1 - Asp cleavage of K63-Ub$_2$ creates a signature fragment.
(A) Primary sequence of mono Ub shown with six fragments after Asp digestion bottom. Linkage sites are show in blue, Asp residues are shown in red, and G76 is yellow **(B)** The six Asp fragments of mono Ub are shown with their mass (in Da) and linkage sites in bold. **(C)** Structure of K63-Ub$_2$ with diagnostic Asp cleavage fragment in blue and linkage sites K63 and G76 show in magenta sticks. Bottom shows only the diagnostic peptide from K63-Ub$_2$ in blue. **(D)** Sequence of K63-Ub$_2$ with an enlarged view of the diagnostic peptide, note that K63 and G76 are perfectly contained in the Asp cleavage fragment.

6.3 Validation of the Asp specific cleavage method

6.3.1 Optimization of cleavage

Although Asp-N had previously been used by Goldberg to cleave Ub chains for a different purpose, the chemical MAAH method was never tested. This prompted us to perform both Asp-N and MAAH digestion on mono Ub. Conditions such as exposure time to Asp-N and microwave times for MAAH were optimized. With both methods the expected cleavage products were observed using MALDI-TOF. We were able to produce nearly identical spectra from both stock solutions and 15% SDS-PAGE gels, demonstrating that both digest methods would be highly effective in proteomics. Interestingly, there was a very pronounced missed cleavage product at D58 of Ub with MAAH digestion (**Figure 6.2**). Initially we overlooked this since it is generally assumed that the conditions for MAAH, (12.5% (v/v) acetic acid, pH 2.24) fully denature proteins, however Ub is extremely stable with a T_m~90^{o}C and Ub is also purified in ~2% (v/v) perchloric acid reaching a pH 1.5 (*229, 230*). To determine how folded Ub was under MAAH conditions we used ^{1}H,^{15}N-TROSY (variation of HSQC) to observe Ub in varying concentrations of acetic acid (**Figure 6.2 D-F**). The ^{1}H,^{15}N-TROSY spectra clearly indicate that Ub is well folded in 12.5% (v/v) acetic acid. There are extra peaks corresponding to alternate conformations, which also may indicate small percentage of the Ub is beginning to unfold, but the residue specific signals indicate that most Ub retains its overall native structure. D52 which is located in a flexible loop of Ub with virtually no intramolecular contacts is readily cleaved by MAAH. The 3_{10}-helix containing D58 allows for D58 to be involved in an extensive hydrogen bonding network, making it

highly resistant to MAAH which requires both the Asp side chain and backbone for efficient cleavage (*231*).

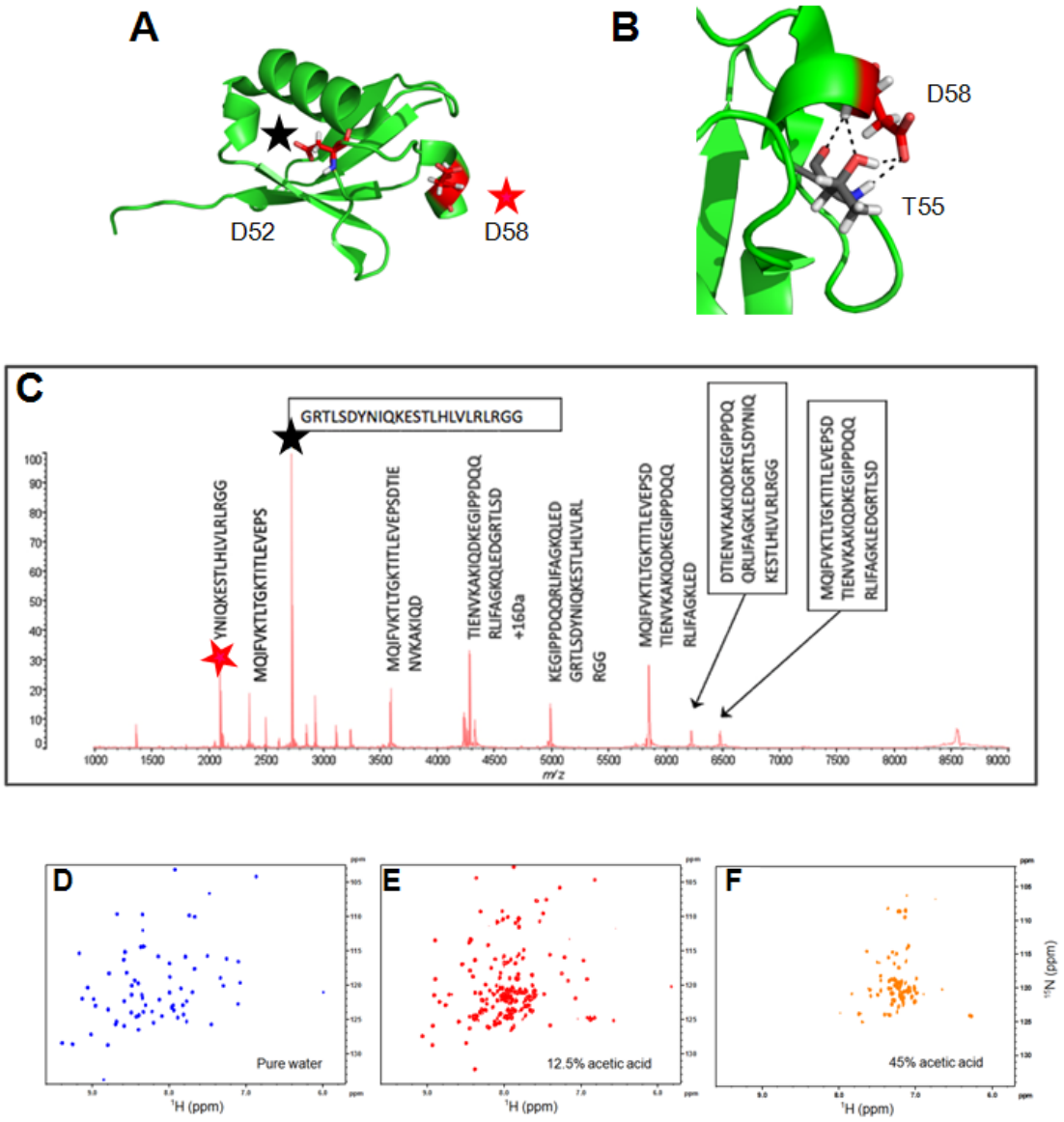


Figure 6.2 – Asp specific digestion of monomeric Ub.
(A) Structure of Ub with red sticks show the position of D52 (black star) and D58 (red star) on the surface of Ub. D52 is located in an unstructured loop, while D58 is in a 3_{10} helix. **(B)** PyMol assigned close polar contacts of D58 and T55 are shown as dashed lines. **(C)** MALDI-TOF spectrum of MAAH digested monomeric Ub. The peak corresponding to cleavage at D52, but a missed cleavage at D58 is designated with a black star. The peak representing cleavage at D58 is indicated with a red star. **(D-F)** ^{1}H,^{15}N-TROSY spectra of ^{15}N-Ub$_1$ in different concentrations of acetic acid: **(D)** in pure water, **(E)** in 12.5% (v/v) acetic acid, and **(F)** in 45% (v/v) acetic acid. The residue-specific NMR signals corresponding to backbone amides are well spread in water (blue) and still retain the spread in 12.5% acetic acid, pH 2.24 (red), indicating that ubiquitin remains well folded under both conditions. The spectrum of 45% acetic acid, pH 1.85 (orange), is characteristic of a nearly unfolded protein.

We attribute the near-native structure of Ub to be responsible for partially inhibiting MAAH cleavage. With properly optimized conditions we found there was cleavage of a significant portion of D58. Asp-N cleavage proved to be very efficient, but required significant time periods for the digest ranging from 18-24 hours. In many cases we still observed some missed cleavages in MALDI-TOF.

6.3.2 The benefits of a missed cleavage for the analysis of unanchored K63-linked polyubiquitin chains

As shown in (**Figure 6.2**) a missed cleavage frequently happens at D58 during MAAH, but D58 is also cleaved. For the analysis of K63 Ub conjugates this would produce multiple diagnostic peptides for identification, in addition to the theoretical peptides in (**Figure 6.1**). After careful consideration we determined that these alternate cleavage products resulting from a missed cleavage at D58 would actually be beneficial and allow us to analyze MS data with greater confidence. The reason being is that the D58 and D52 cleavage products resulted in two distinct fragments that would always be separated by MW=629.7 Da using this cleavage method. This concept is demonstrated excellently with unanchored K63-Ub_2 chains (**Figure 6.3**).

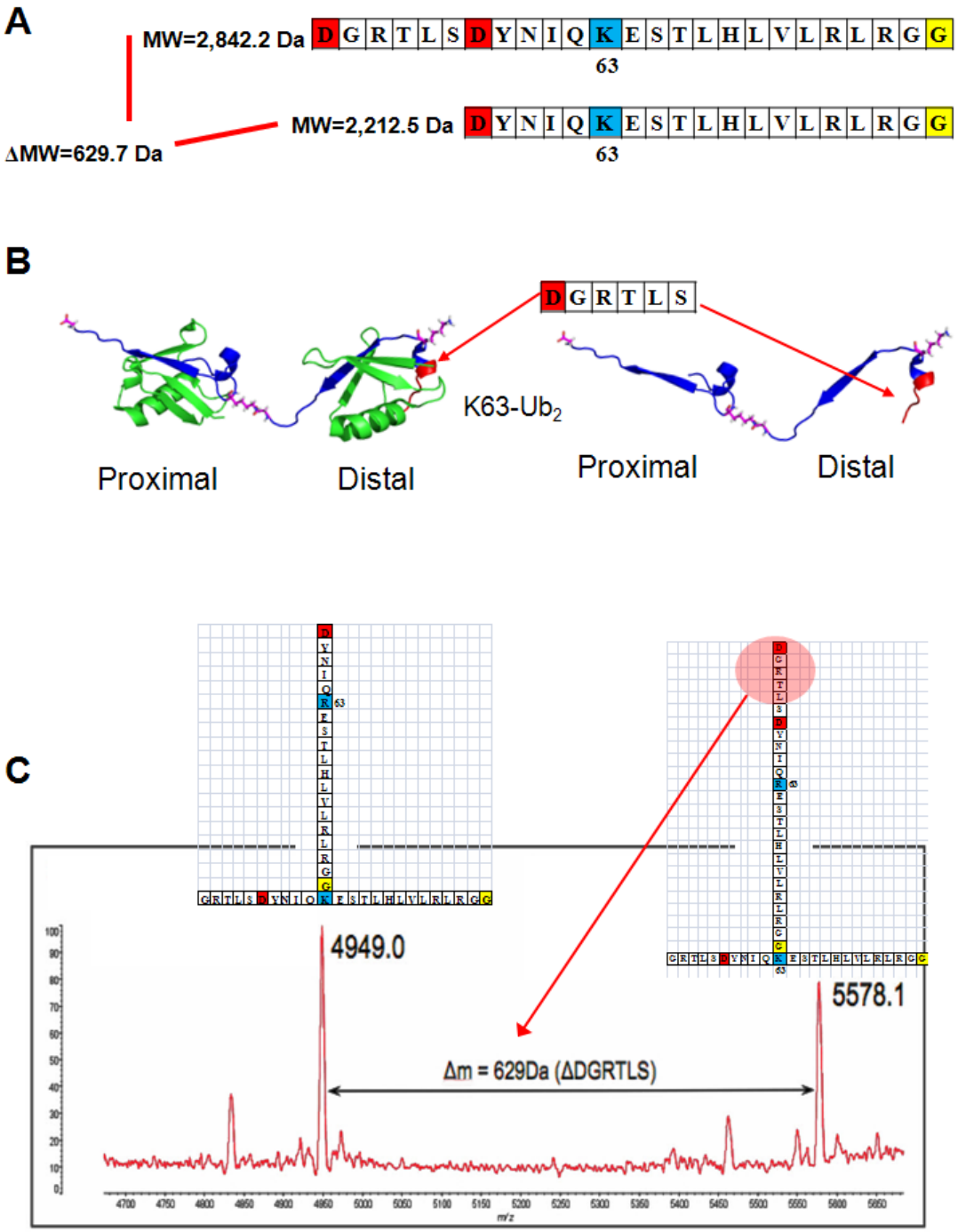


Figure 6.3 – MAAH digestion of K63-Ub$_2$ with Δ629Da diagnostic peptides
(**A**) D52 fragment of Ub with missed cleavage at D58 (top) and D58 fragment (bottom). (**B**) Structure of K63-Ub$_2$ analyzed showing diagnostic peptide (blue) and missed cleavage at D52 (red) on distal domain. (**C**) MALDI-TOF spectrum shows two distinct diagnostic MAAH peptides form K63-Ub$_2$. The missed cleavage at D52 on the distal Ub produces an increase in 629 Da.

After we established that both Asp-N and MAAH cleavage were effective for detecting unanchored K63-Ub_2 we set to extend this methodology to longer K63 linked Ub chains. If there is 100% cleavage, the mass of the Asp cleavage fragments will be the same except for the last, which contains K63. As the number of consecutive K63 linkages increases the mass of the K63 fragment will increase in increments of 2,194.49 Da with no missed cleavages and 2,824.08 Da if D52 is missed. The next heaviest fragment in Ub is the first Asp cleavage fragment containing M1, K6 and K11 at 2,235.6 Da. Since all of the Asp fragments except the one containing K63 maintain the same mass, the K63 fragment will easily be the largest as the number of consecutive K63 linkages increases. As a practical matter, the presence of excess non-K63 containing fragments severely complicates MS analysis and for most K63 linked chains the number of fragments containing K63 are greatly outnumbered by those not containing K63. To separate the K63 containing fragment, we used LC-MS/MS to isolate a K63-Ub_2 fragment and then assigned many diagnostic b ions from collision-induced dissociation (CID). Although LC-MS/MS analysis was successful for K63-Ub_2 the method would prove impractical for longer K63 linked chains that have fragments with different elution profiles and produce many more diagnostic b ions. A detailed search of existing reagents led us to the K63 linkage specific mouse monoclonal antibody (clone HWA4C4). Most commercially available K63 specific antibodies do not recognize monomeric Ub, but instead the K63 linkage. In the case of the HWA4C4 clone, it was created with a very small antigen containing residues 71-76 (RLRGG) of the distal Ub, linked to residues 58-63 (DYNIQKEST) of the proximal Ub. This suggested that this antibody

would be highly effective in capturing the Asp fragment whose sequence overlaps the antigen (**Figure 6.4 A-C**). Structures of similar K63 antibodies (Apu.3A8 and Apu2.16) bound to K63-Ub_2 (PDB-3DVG and 3DVN) supported this assertion (*201*). After we successfully immunoprecipitated the K63 containing fragment from the digest of an octameric K63 chain we set to test if this could be applied to a mixture of chain lengths. In the cell, K63 linkages do not have defined lengths, and can range from range from 2-9 linkages for certain substrates (*201*). Following MAAH immunoprecipitation by the K63 antibody allowed us to successfully isolate the K63 fragment for each chain length in the K63-Ub_{2-5} mixture. Each chain was assigned based on pairs of peaks spaced 629 Da apart (**Figure 6.4**). This was a clear demonstration that Ub chain length could be measured and also is the longest Ub chain measured by any method.

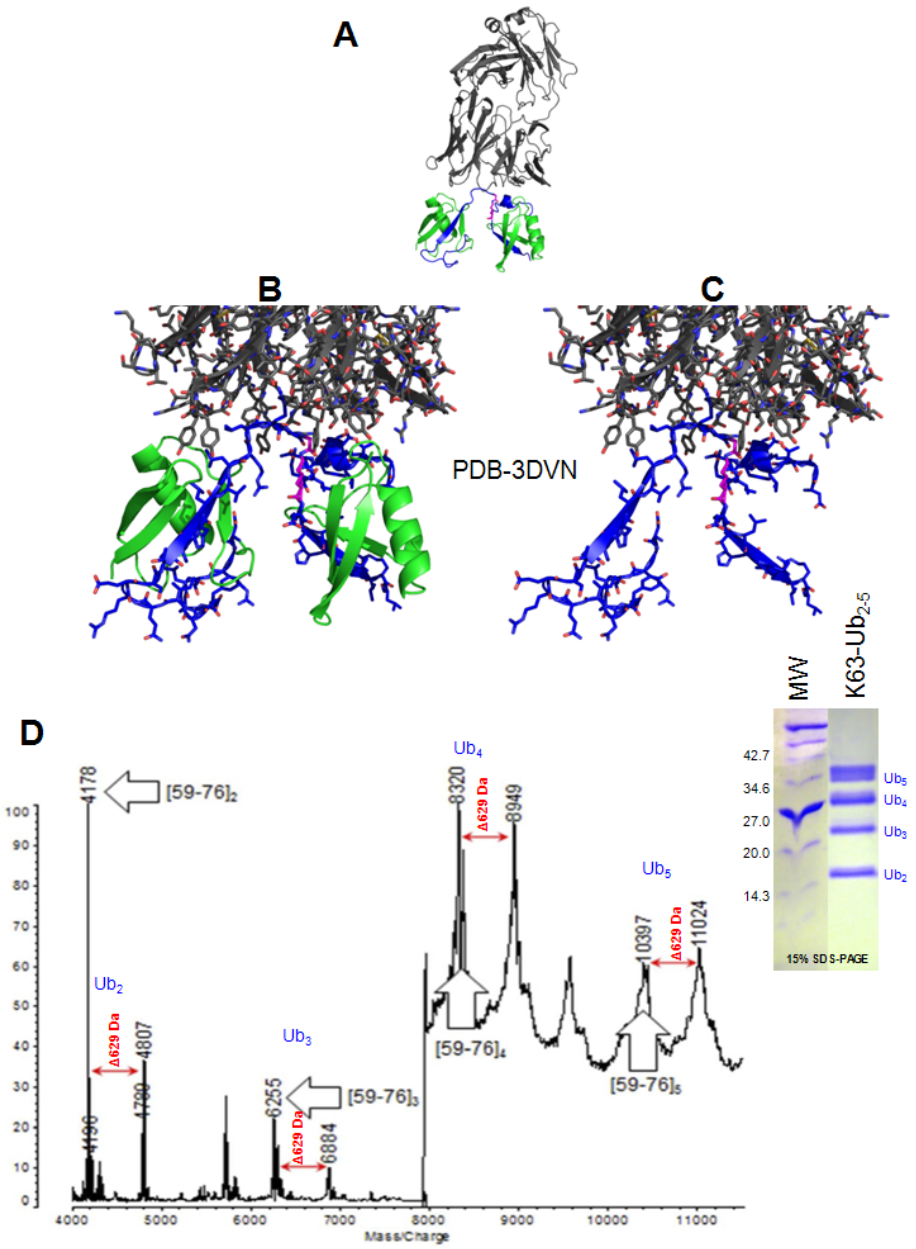


Figure 6.4 – Asp digestion followed by immunoprecipitation measures linkages in K63-Ub_{2-5}

(**A**) PDB-3DVN shows the K63 linkage specific antibody (gray) interacts with the linkage (magenta) on K63-Ub_2. (**B**) The K63 containing Asp fragment (blue) mainly interacts with the antibody and does not form contacts with the rest of the K63-Ub_2 molecule (green). (**C**) Showing only the K63 containing Asp fragment with one missed cleavage (blue sticks). (**D**) MALDI-TOF following immunoprecipitation of the K63-Ub_{2-5} mixture (gel insert). Open arrows $[59\text{-}76]_n$ indicates peak representing no missed cleavages. Red arrows show the pair of Δ629 peaks expected for each product.

6.3.1 Measuring the K63 Ub modification of a model substrate

After we were convinced that we could apply our newly devised method to unanchored K63 linked Ub chains, we wanted to bring the method to the next step and analyze a substrate modified with K63 linkages. One of the major problems was that ubiquitinated substrates are not easily obtainable. Furthermore, many ubiquitinated proteins from *in vivo* sources or enzymatic reactions produce a large amount of heterogeneous chains, have low yields, and were beyond our means. From the literature it was shown that autoubiquitinating E2 and E3 enzymes, such as TRAF6 for K63 linkages and Ube2S for K11 linkages have been used to create homogeneous linkages of varying chain lengths. From my studies with DUBs, I knew that Ubch5b, an E2 conjugating enzyme could autoubiquitinate, however it also promiscuously created other linkages. For our purposes of obtaining a K63 modified substrate, the autoubiquitination E2 or E3 system was ideal since it was both practical and yielded the most material. To create our substrate I reacted Ub(K63), a Ub mutant only containing K63 with all other Ks mutated to R, with E1 and Ubch5b. This yield only mono, multiple-mono, or K63 ubiquitinated Ubch5b. After Asp digestion and immunoprecipitation, the K63 fragments would be significantly enriched and all other fragments from Ub, Ubch5b, or monoubiquitinated Ubch5b would not be selected by the antibody. To ensure that the K63 chain would have a defined length, we analyzed samples from a 15% SDS-PAGE gel band corresponding to the mass of Ubch5b + 2Ubs. The results in (**Figure 6.5**) show we identified one K63 Ub linkage at K128 of Ubch5b in the 117-130 Asp fragment, in a system we would name Ub–^{63}Ub–^{128}Ubch5b.

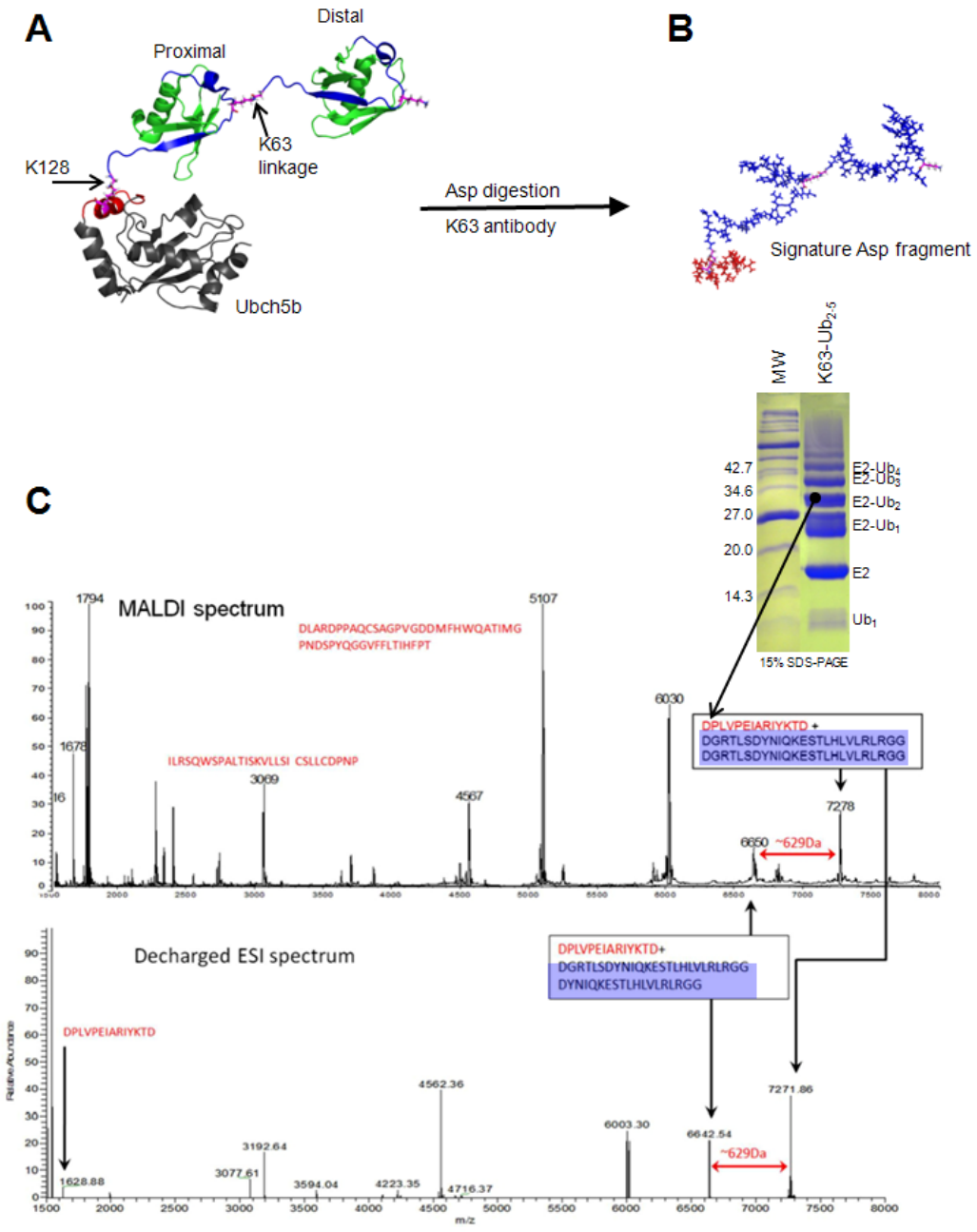


Figure 6.5 – Asp digestion of a K63 modified substrate.
(**A**) Model of Ubch5b (gray) modified with K63-Ub_2 (green) at K128. Linkages are shown as magenta sticks, the K128 Asp fragment of Ubch5b is colored red, which continues to Ub (colored blue). (**B**) Asp digest followed by immunoprecipitation with the K63 linkage specific antibody isolates an Asp fragment with a K63 Ub linkage (blue) and corresponding linkage site on Ubch5b (red). (**C**) Gel inlay shows the Ubch5b-Ub_2 band analyzed. MALDI contains the expected the signature doublet with the expected Δ629 Da split from a missed cleavage on Ub (top spectrum). The same signature fragment is observed with ESI (bottom spectrum). Sequences from Ubch5b are in red and the K63 linkage from Ub is boxed with blue.

6.3.4 Expanding to future applications

Our initial work clearly only focused on unanchored and substrate attached consecutive K63 linkages, however Asp digestion can also be used to detect other forms of ubiquitination. The possibility to isolate K33 linkages, as well as the detection of branching on any of the other Asp fragments was discussed in (**section 6.2.2**). A new application of Asp digestion is also to unambiguously reveal linkage mixing in Ub chains that contain a K63 linkage distal to another linkage. No technique to sequence mixed linkage chains has been discovered and our Asp digest method could greatly aid in the characterization of specific systems containing mixed linkages. Linkage specific antibodies and tryptic digest would detect mixed linkage chains by identifying more than one linkage type, but neither method would be able to establish the exact sequence of the linkages. Diagnostic Asp fragments for Ub–^{63}Ub–^{n}Ub that are produced by Asp cleavage of mixed linkage tri-Ubs are shown in (**Figure 6.6**). Since the K63 linkage is retained any mixed linkage with a distal K63 linkage could also be analyzed with Asp digest and immunoprecipitation as done with the homogeneously linked K63 chains. A non-K63 linkage proximal or distal to the K63 linkage will break the chain during Asp cleavage leaving only units of two Ub-Ub bonds. If there is longer sequence of K63 linkages distally attached to a non-K63 linkage, the sequence of K63 linkages will remain intact during Asp cleavage and the proximal non-K63 linkage will be detected essentially as a substrate. If however a non-K63 linkage is distally attached to a sequence of K63 linkages it will not be detectable.

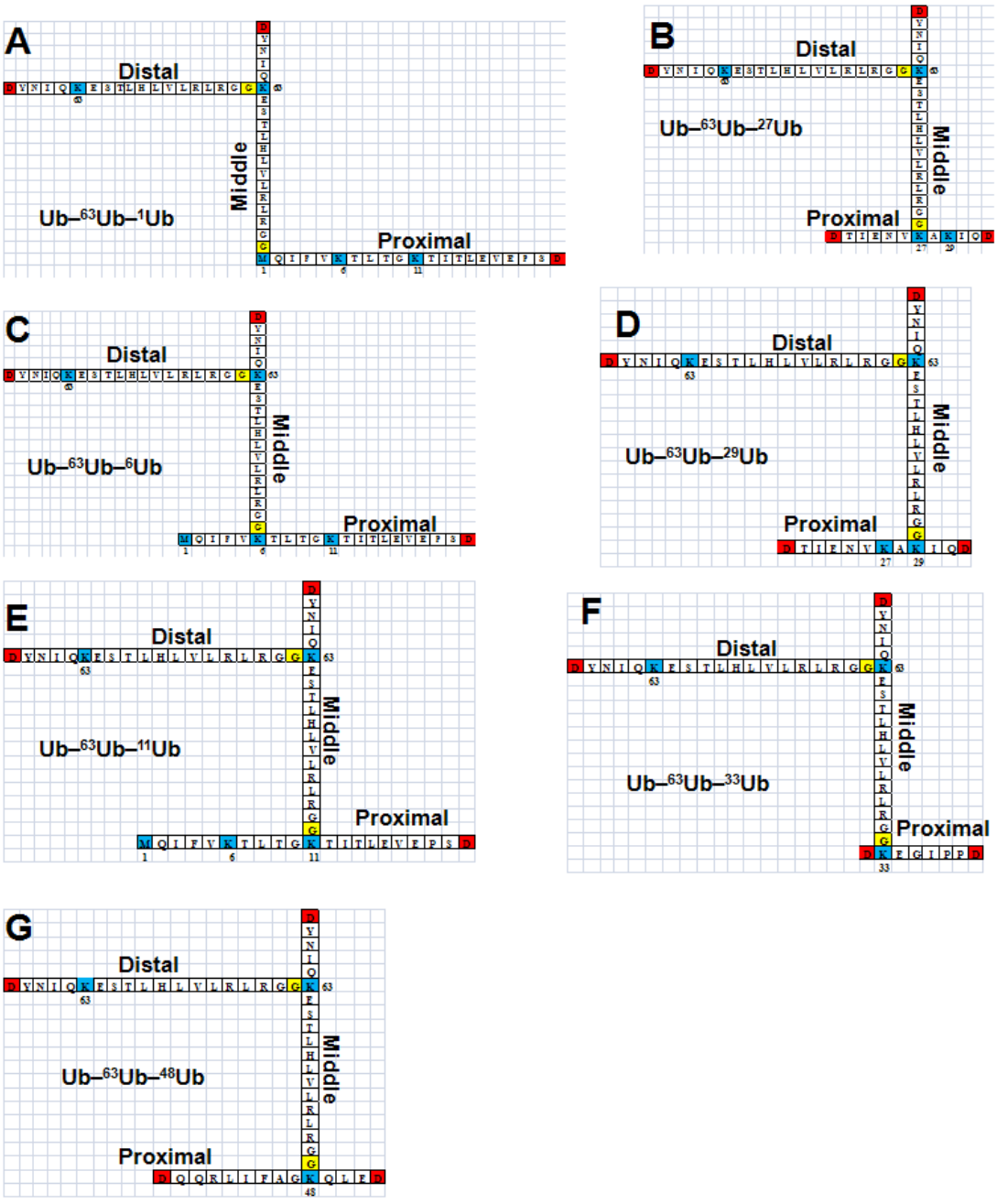


Figure 6.6 – Asp digestion of mixed linkage Ub–^{63}Ub–^{n}Ub chains.
Following Asp digestion these peptides would result from a K63 linkage distally attached to any other linkage site besides K63 on Ub. (**A**) Ub–^{63}Ub–^{1}Ub, (**B**) Ub–^{63}Ub–^{27}Ub, (**C**) Ub–^{63}Ub–^{6}Ub, (**D**) Ub–^{63}Ub–^{29}Ub, (**E**) Ub–^{63}Ub–^{11}Ub, (**F**) Ub–^{63}Ub–^{33}Ub, (**G**) Ub–^{63}Ub–^{48}Ub

With the ability of the K63 linkage specific antibody to capture the desired Asp digest fragments our methodology should allow for easy analysis of *in vivo* extracts. We have yet to attempt a bottom up proteomics approach, but we have laid the ground work to do so and are developing software for the analysis of such data. An equation to predict he mass of K63 linked Ub conjugates is a core feature of the software, show in (**Figure 6.7**). To account for missed cleavages the term "Ub" representing the mass of the Ub fragment can be 2,842.18 Da for one missed cleavage or 2,21.21 Da for no missed cleavages. Calculating the highest expected mass and subtracting 629 Da proved to be a straight forward way to search for Asp fragments. The equation can also be applied to unanchored K63 chains by setting the parameter designating the substrate fragment z equal to 0 Da.

It is unclear if polyUBL chains or heterologous Ub/UBL chains can be effectively approached with the Asp cleavage method. However, the tryptic method has been used for SUMO2/Ub chains and other cleavage methods could presumably be effective for UBL or Ub/UBL chains (*232*).

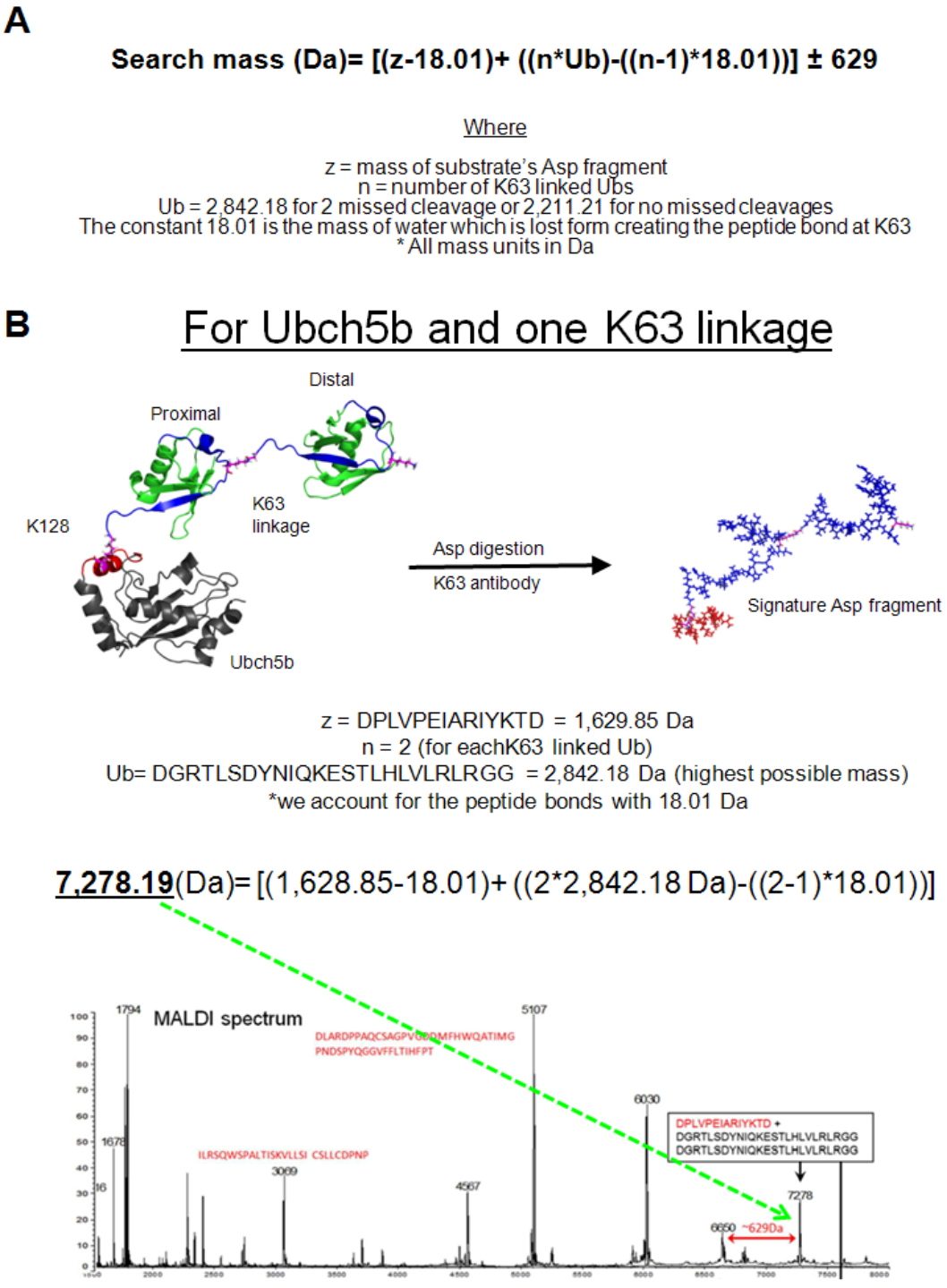


Figure 6.7 – Accurate prediction of Asp cleavage masses.
(**A**) Equation with description of parameters to predict search masses for K63 modified Asp fragments (**B**) Demonstration of the equation with Ubch5b and one K63 Ub linkage. The predicted mass of 7,278.19 Da and corresponding peak 629 Da less in mass show up in the MALDI spectrum exactly where expected.

6.4 Discussion and summary

Our method which utilizes Asp specific cleavage followed by immunoprecipitation clearly can measure the number of consecutive K63 Ub linkages in unanchored or substrate attached Ub chains. The alternate cleavage methods of Asp-N and MAAH work equally well, but with their distinct advantages and disadvantages (mostly experimental details). In the larger picture, this simple approach can easily be extended to characterize K63 linkages in other systems with a more traditional proteomics approach. However, expanding on our approach by employing alternate cleavage methods and specialized antibodies should allow us to gain much more information in future studies. A future aim will undoubtedly be to expand this method for use with mixed linkage polyUb chains and also to UBLs that form similar polymers.

The approach outlined in this chapter is by far the most versatile approach to date and can be used for both proteomics applications and the analysis of Ub conjugates in a controlled setting. Using customized instrumentation and ion mobility MS, David Clemmer and co-workers were able to differentiate between unanchored polyUb chains of K48 or K63 linkages due to their ability to retain distinct structural characteristics in the gas phase (*233*). Another top down method was used by Layfield and co-workers to show connectivity between a K48 linked Ub dimer from an *in vivo* sample, however the MS technique requires extensive optimization and is most likely impractical for longer Ub chains (*234*). Nevertheless, the success of these methods and other non-tryptic approaches have certainly gave the field an unprecedented view of Ub signaling.

Chapter 7: Short research projects

7.1 Background and research aims

The projects contained in this section represent experiments I performed as part of larger collaborations, which were not of my own design or techniques I developed.

7.2 Applications of the F45W ubiquitin mutant

7.2.1 Advantages of fluorescent Ub variants

The Ub(F45W) mutant was the first fluorescently labeled Ub species and originally used to prove the controversial "late two-state" folding pathway of Ub (*235-237*). Many other fluorescently labeled Ub variants have been used for a range of applications, but mainly for detecting DUB inhibition in high throughput formats (*238, 239*). Although sound, most of the current fluorescence assays for study of Ub have several draw backs including the concerns over the introduction of extrinsic fluorophores, the shelf life of reagents, and costs of materials. To improve upon existing methods I demonstrate several novel applications of the Ub(F45W) mutant and also devote much effort to studying the effects of the mutation. The substitution of a tryptophan at position 45 in Ub results in an almost native probe, which is in close proximity to the L8,I44,V70 hydrophobic patch. This makes the mutant ideal for monitoring Ub/receptor interactions. Furthermore, addition of the fluorescence capability also makes this Ub mutant useful for FRET applications as well.

7.2.2 Ub(F45W) retains a native Ub structure

Given that many studies have used the F45W Ub mutant in some form, but neglected to strictly assess if: (i) the native structure of monomeric Ub was preserved, (ii) the F45W mutation effected the conformation of polyUb chains, and (iii) the affinity of Ub(F45W) variants for receptors changed, I first attempted to determine if any structural changes arose from the mutation. I immediately had a concern about the stability of Ub(F45W) when I observed that it precipitated with the perchloric acid used during standard Ub purification. However, when using the heating method (see **section 3.1.4**) the F45W mutant was very stable even at temperatures nearing the T_m of Ub, which demonstrated that the mutant retained its characteristic thermal stability. After purification, the residue specific CSPs between wild type and Ub(F45W) were calculated from their ^{15}N-HSQC spectra. CSPs where observed where expected, i.e. residues in close proximity to the mutation (see **Figure 7.1**). Aside from these especially large CSPs near the mutation site, CSPs for other residues are extremely small in Ub(F45W), supporting that the mutation does not induce massive structural changes.

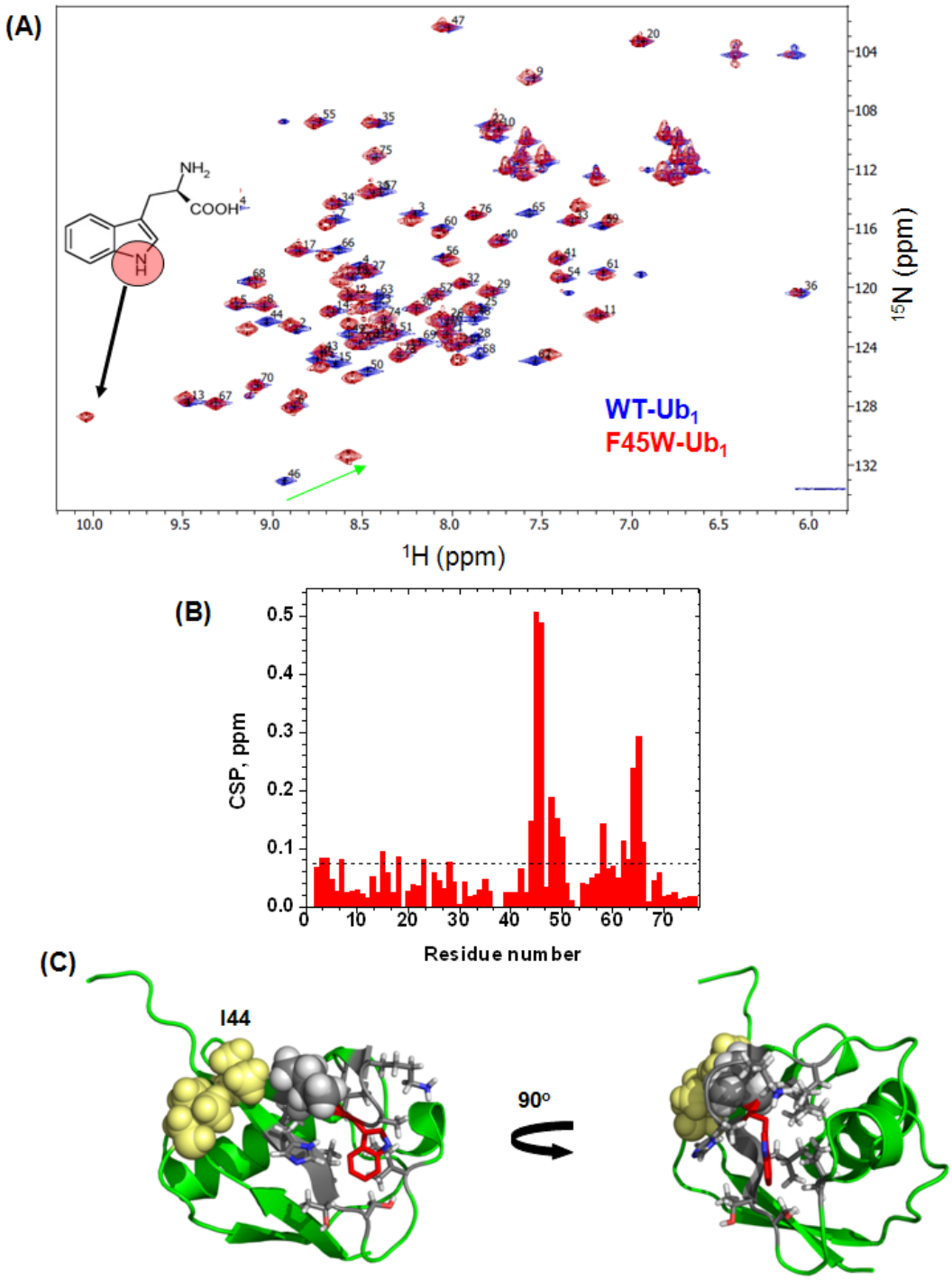


Figure 7.1 – Solution properties of monomeric Ub(F45W).
(**A**) $^{1}H,^{15}N$-HSQC overlay of wild type Ub (blue) and Ub(F45W) (red) shows an overlap for many peaks, but large changes in those around the mutation site. Note the presence of and extra speak (far left) corresponding to the side chain of tryptophan. (**B**) Residue specific CSPs are plotted between wild type and Ub(F45W), showing that the effect of the mutation is localized. (**C**) F45W modeled into (PDB-1D3Z) shows its close proximity to the L8, I44,V70 hydrophobic patch (yellow spheres), residues with in 4Å are colored in gray.

After determining that the overall structure of Ub(F45W) was essentially that of wild type, the viability of the F45W mutation was tested *in vivo*. Collaborators Dan Bolon and Ryan Hietpas used a yeast system with a rescue plasmid containing either wild type Ub or Ub(F45W) as the sole source of Ub. Ub(F45W) was shown to be just as viable as wild type (**Figure 7.2**). This results leads to several important conclusions regarding the F45W mutant. Most importantly, the F45W mutant does not interfere with Ub signaling in the cell and all pathways continue uninterrupted. Following the current dogma this implies that the structure of Ub must be well retained in the mutant. In addition, other properties such as the conjugation of Ub(F45W) into polymeric chains and their deconjugation must also be retained.

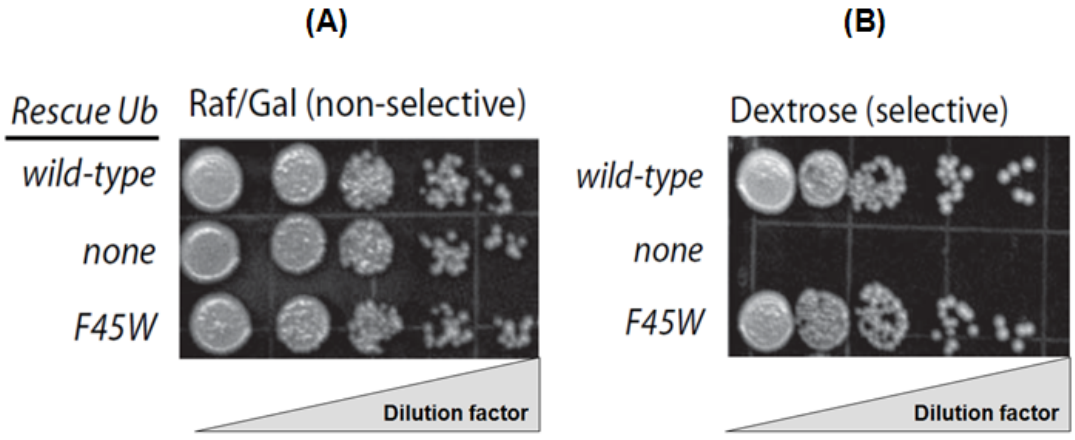


Figure 7.2 – Ub(F45W) is viable in yeast
(**A**) Ub from any source can be used in the Raf/Gal non-selective media. (**B**) In the dextrose media only Ub form the rescue plasmid is expressed. Where there is no Ub, the cells die from the lack of this essential protein. With wild type Ub and Ub(F45W), the cells flourish. This demonstrates that both forms of Ub are indistinguishable from the cell's point of view.
* Experiment carried out by Dan Bolon and Ryan Hietpas

These results were encouraging and justified treating the F45W mutant essentially as wild type Ub. Since there were no major structural changes I proceeded to the next step and formed polyUb conjugates with the F45W mutation incorporated at several different positions (distal, proximal, and higher conjugates with F45W at every position) for both K48 and K63 linkage types. Solution NMR data collected for Ub(^{15}N)–^{48}Ub(F45W) showed that signals in the distal domain in the K48 linkage did not have any significant deviation from signals in the wild type conformation and supports that the polyUb chain must also retain the same structural conformations. As a final check, crystallization was attempted with both monomeric F45W and a K48 linked dimer containing F45W in the proximal domain, Ub–^{48}Ub(F45W). In the absence of a structure at atomic resolution, this approach would provide one and in the process also reveal if there are similar crystallization properties between wild type and F45W Ub. After screening, a crystallization condition for monomeric F45W (25% PEG-3,350, 40 mM zinc acetate) very similar to PDB-1UBQ and other reported crystal structures for monomeric Ub was found (*240*). For Ub–^{48}Ub(F45W) no screening was attempted and it was simply placed in conditions identical to those used for PDB-3M3J (0.2 M lithium sulfate, 50 mM Tris pH8.0, and 25% (v/v) PEG-3,350), which resulted in the reported plate like crystals (*241*). These crystals diffracted at 2.7 Å in the same space group, again showing very similar properties between wild type Ub–^{48}Ub and Ub–^{48}Ub(F45W) (**see Figure 7.3**).

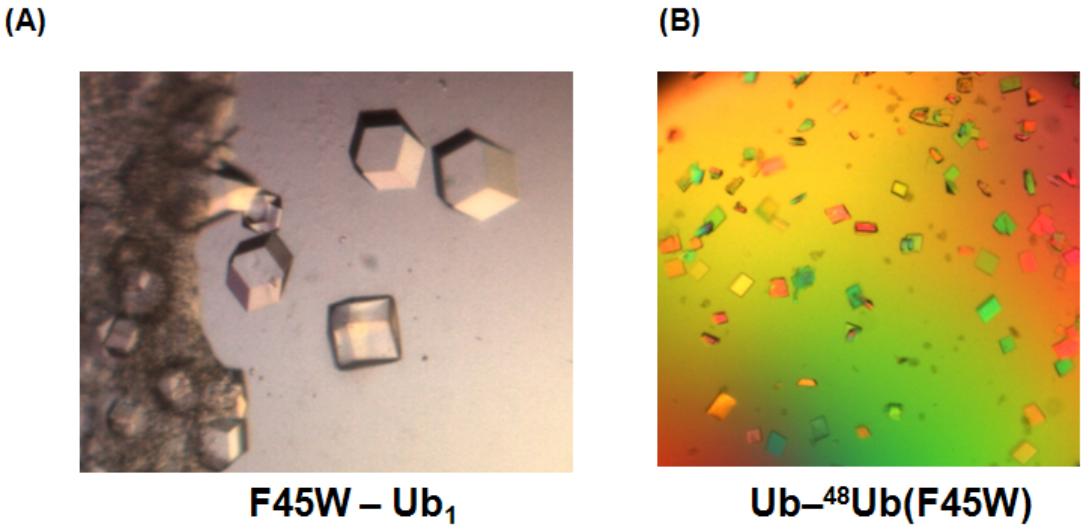


Figure 7.3 – Crystallization of F45W in monomeric and K48-Ub_2 forms. (**A**) Cubic crystals of monomeric F45W in 25% PEG-3,350, 40 mM zinc acetate after 9 days of growth at 22°C. (**B**) Plate crystals of Ub–^{48}Ub(F45W) grown in 0.2 M lithium sulfate, 50 mM Tris pH8.0, and 25% (v/v) PEG-3,350 at 22°C for 6 days.

7.2.3 Binding interactions of F45W

After establishing that the F45W mutation was essentially identical to wild type Ub structurally, I set to determine if the mutant would form the same binding interactions with receptors. The yeast data suggested F45W Ub interacted, but provided no means for quantifying isolated Ub/receptor interactions. A concern was that subtle changes around the F45W mutation could alter the binding surface resulting in different affinities for receptors, but still be viable. To determine the binding properties of F45W, titrations of the globular UBA domain from the UBL/UBA shuttle ubiquilin-1 (UBQ-1) and the α-helical UIM domain from proteasome subunit Rpn10 were performed using both NMR and tryptophan emission. Data from ^{1}H,^{15}N-HSQC titration of either ^{15}N-wild type Ub_1 or F45W Ub_1 shows similar residue specific CSPs on the L8,I44,V70 hydrophobic patch, and

also nearly identical titrations curves for select residues with UBA from UBQ-1 (**Figure 7.4**).

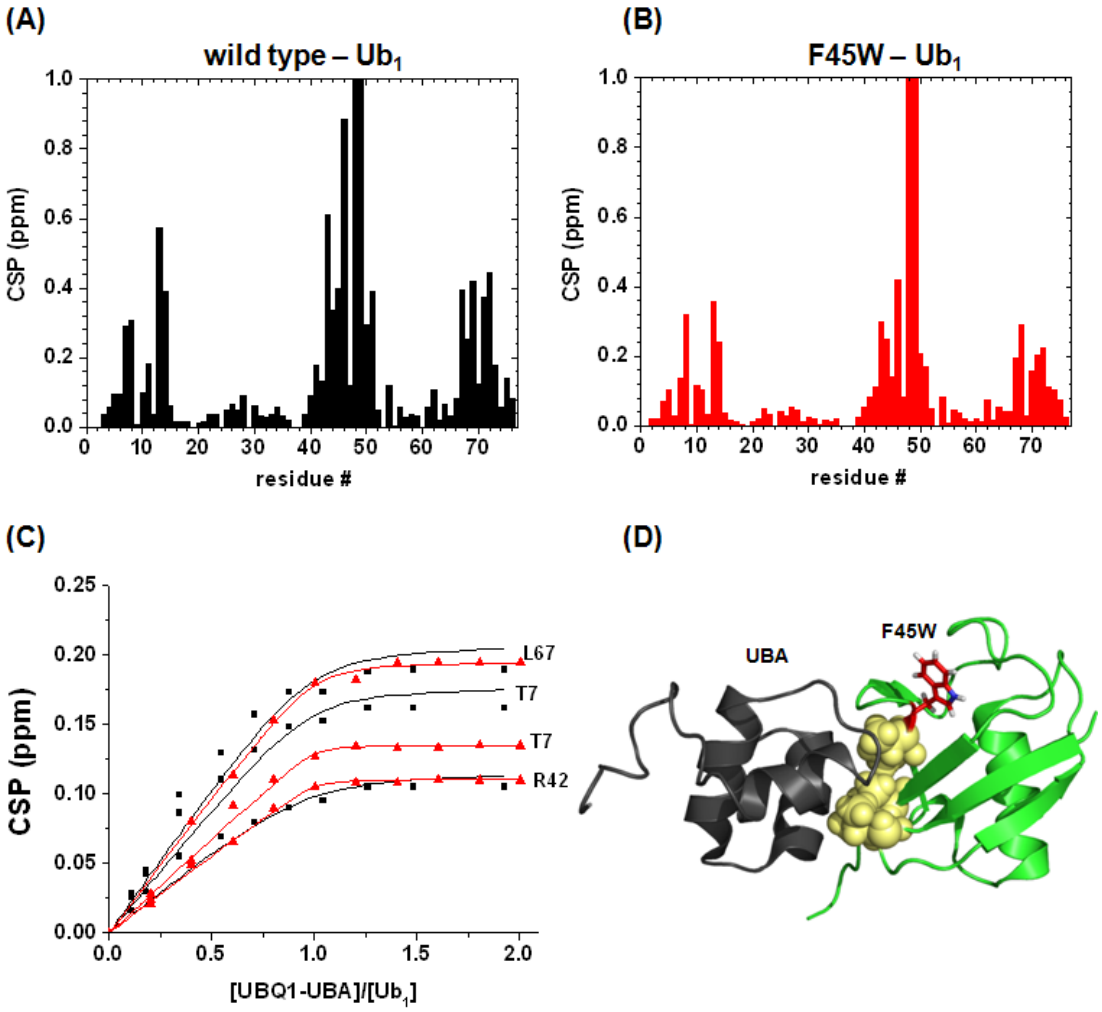


Figure 7.4 – F45W Ub_1 natively interacts with UBA from UBQ-1.
(**A**) Residues specific CSPs for wild type Ub_1 (black) and (**B**) F45W-Ub_1 (red) with the UBA domain of UBQ-1 show a similar pattern consistent with binding the L8,I44,V70 hydrophobic patch. (**C**) Titration curves for residues T7, R42 and L67 of wild type (black) and F45W (red) show a near identical affinity for each. (**D**) Solution structure of complex (PDB-2JY6) between monomeric Ub (green) and UBQ-1 UBA (gray), shows F45W (red) is very close to the interaction site, hydrophobic patch (yellow spheres).

NMR titration data with the UIM of Rpn10 also revealed a similar binding interaction (**Figure 7.5**).

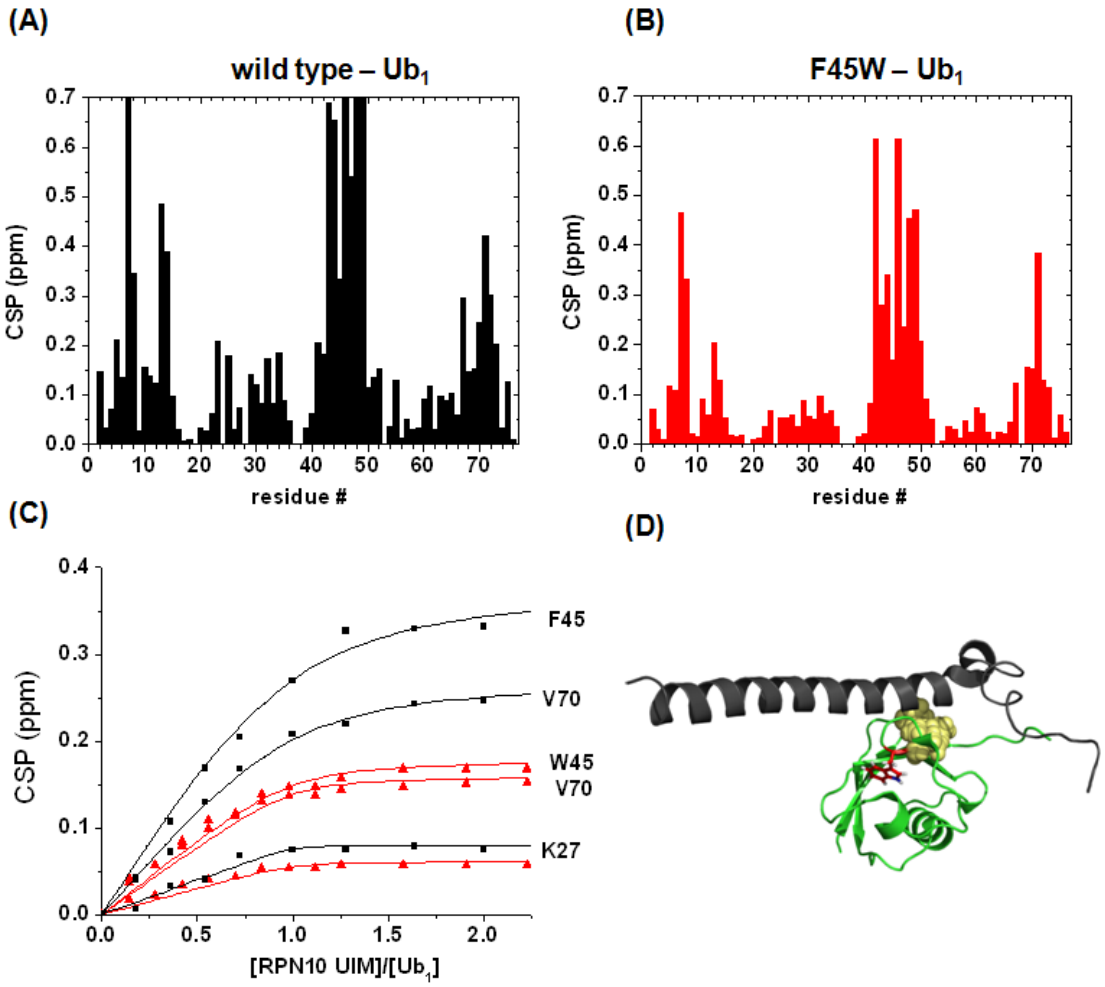


Figure 7.5 – Rpn10 UIM recognition of F45W Ub_1.
(**A**) Residues specific CSPs for wild type Ub_1 (black) and (**B**) F45W-Ub_1 (red) with the UIM domain of Rpn10 show a similar pattern in agreement with the binding mechanism. (**C**) Titration curves for residues K27, F(W)45, and V70 of wild type (black) and F45W (red) have identical profiles, but differ in magnitude. (**D**) Solution structure of complex (PDB-1YX5) between monomeric Ub (green) and Rpn10 UIUM (gray), shows the close proximity of F45W (red) to the interaction surfaces.

With NMR data in good agreement, the ability of F45W to serve as a probe for the quantification of binding was tested by tryptophan fluorescence. Generally, tryptophan emission is very sensitive to many factors including binding interactions which change the solvent exposed surface area of Ub. Data for UBQ-1 UBA are shown as an example (**Figure 7.6**). Upon receptor binding, there is a large increase in tryptophan emission for both monomeric F45W and Ub–^{48}Ub(F45W) with the same ligand. The profile of the curves also supports the binding model in that monomeric F45W saturates at exactly 1:1 and Ub–^{48}Ub(F45W), which can bind two UBAs, saturates much later. When saturation is reached there is no more increase in emission and F45W appears to function ideally as a probe.

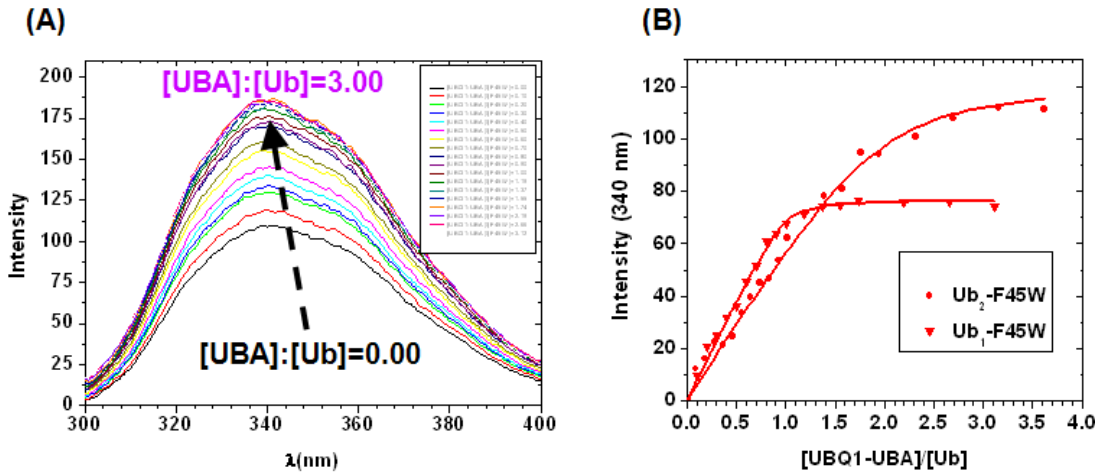


Figure 7.6 – Tryptophan emission of Ub(F45W) increases with UBA binding. (**A**) Overlay of emission spectra over the course of monomeric F45W Ub with UBA from UBQ-1. As UBA is added there is an increase in fluorescence intensity. (**B**) One wave length, 340 nm was used to calculate the K_d for monomeric and K48 linked F45W variants. Due to the difference in stoichiometry between monomeric F45W and Ub–^{48}Ub(F45W) saturation is reached at different molar ratios in the titration, which is consistent with the literature. For the titrations Ub(F45W) was 45 µM and Ub–^{48}Ub(F45W) was 43 µM. Both curves correspond to a 1:1 binding model.

After performing several titrations by both $^{1}H,^{15}N$-HSQC and tryptophan emission, the determined K_d values were evaluated between each. The NMR and fluorescence titration data were found to be in good agreement (**see Table 7.1**).

Protein	K_D [μM]			
	UBQ1-UBA NMR	RPN10 (UIM) NMR	UBQ1-UBA Fluorescence	RPN10(UIM) Fluorescence
WT–Ub_1	20 ± 5*	44.5 ± 14.8**	—	—
K48–Ub_2	4 ± 5*	13.3 ± 0.6**	—	—
F45W–Ub_1	14 ± 7	38.6 ± 6.2	0.8 ± 0.2	22.5 ± 1.1
K48–Ub_2 (F45W)	ND	ND	4.16 ± 1.08	ND

Table 7.1 – Similar recognition of F45W Ub to know receptors.
* denotes K_d values from published studies (*95, 242*)
The K_d values from NMR are very similar for both wild type and F45W Ub variants. Obtained under different conditions, the K_d values from fluorescence fall within an acceptable range of those determined by NMR. All fluorescence curves were fit to a 1:1 binding model.

Combined with the existing data, this provides quantitative evidence that the F45W mutation is appropriate for a range of *in vitro* studies. The fact that Ub–^{48}Ub(F45W) interacts essentially natively with a UBL/UBA shuttle and a polyUb receptor on the proteasome supports that the mutant will not interfere with the UPP. No other linkage selective receptors (e.g. tUIM for K63 linkages) were tested, nor were DUBs. These experiments were omitted based on the assumption that F45W would produce the same result as wild type. One interesting point is that the K_d and other parameters can vary between methods and are hard to reproduce. In this case with moderate (μM) K_d values, the data being similar between wild type Ub and Ub(F45W) was also consistent between solution NMR and fluorescence. This is interesting in that the

micro K_d values obtained from solution NMR appear to be very agreeable with a fluorescence approach that senses a macroscopic K_d.

7.2.4 Applications of the F45W-AEDANS FRET pair in polyUb

The use of fluorescently labeled and FRET capable Ub systems has increased and commercially available fluorophores have improved, however these relatively large (~1 kDa) in respect to Ub (8.65 kDa) extrinsic probes are stable for short periods of time, and also can be difficult to attach. Since F45W is the most natural fluorescent label that can be carried by Ub and is also stable for years, I set to expand its use for FRET applications. As a FRET acceptor for tryptophan, the ANS derivative IAEDANS which absorbs at 336 nm and emits at 485 nm was selected (*243*). IAEDANS is small, inexpensive, highly thiol-reactive, and has been used in many other systems aside from Ub to obtain meaningful data (*244*). With a well established precedent for the use of Ub cysteine mutants with MTSL to solve Ub/receptor complexes, (*208, 242*) IAEDANS was perfect since it could be used with our existing NMR Ub constructs. The Förster distance of AEDANS (the Iodine is a leaving group after the IAEDANS attaches to a cysteine residue) is ~22 Å which makes the tryptophan-AEDANS pair ideal for monitoring interactions in polyUb systems (*243, 245*).

An external fluorophore on Ub can be very useful for detecting different interactions (*135, 156*). To explore this with Ub $T12C^{AEDANS}$ I performed a simple gel assay with Ub $T12C^{AEDANS}$ in an enzymatic K48 chain reaction. The results (**Figure 7.7**) show AEDANS does not interfere with chain synthesis and it alone has some useful fluorescence applications, in this case detecting Ub conjugates. This

result was encouraging and implied that T12C^{AEDANS} could form conjugates with many other Ub variants aside from itself.

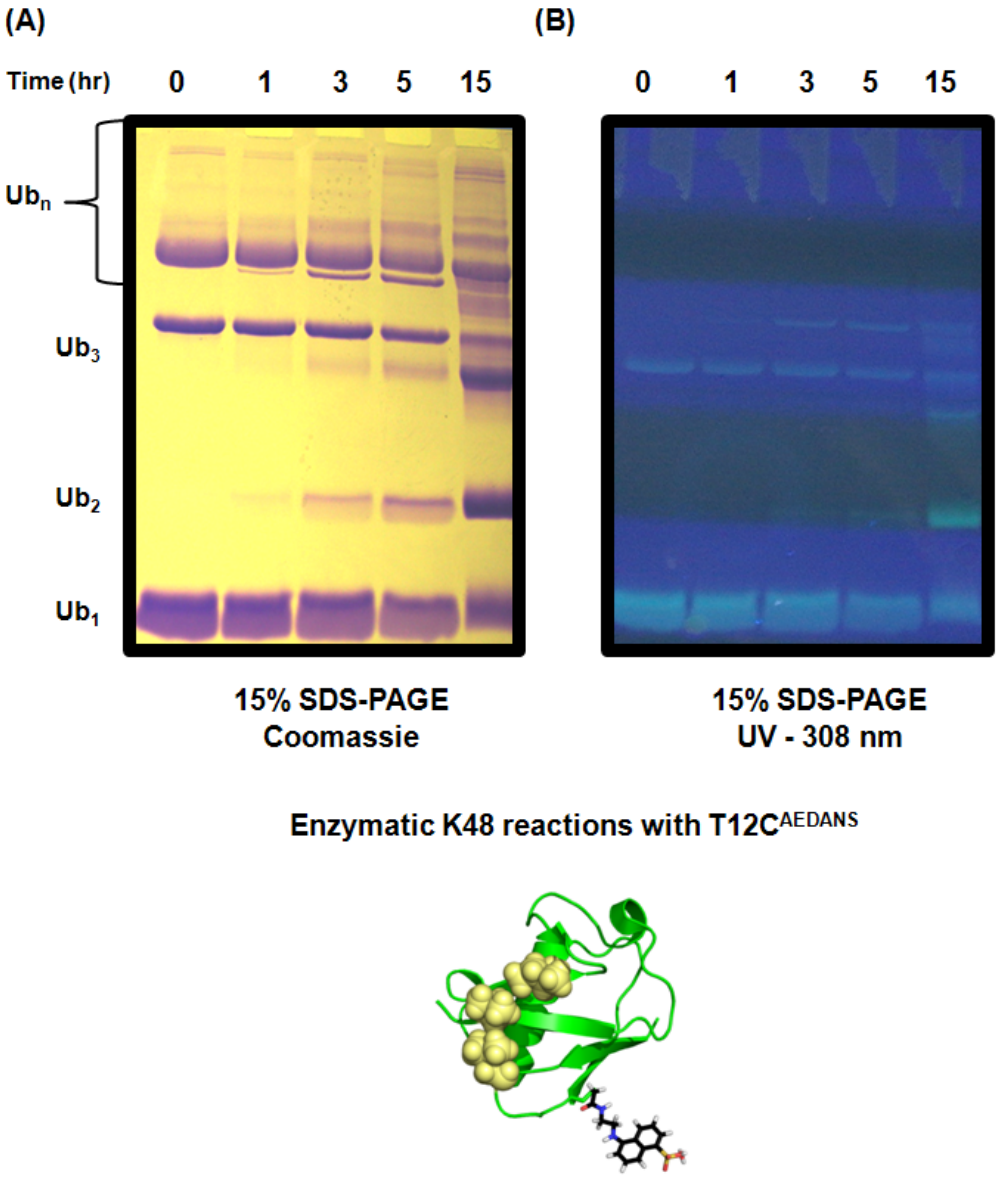


Figure 7.7 – Ub T12C^{AEDANS} as a fluorescent probe for K48-linked polyUb synthesis.
Ub T12C monomers labeled with AEDANS are reacted in a standard enzymatic K48 reaction containing E2-25K. The same exact gel demonstrating that Ub T12C^{AEDANS} K48 linked polyUb chains are formed is shown in (**A**) with Coomassie staining and before staining (**B**) on a UV transilluminator with a 308 nm lamp. All Ub species are highly fluorescent and a distribution of chains is visible. A schematic showing AEDANS attached to T12C on Ub is at on the bottom.

To create the FRET system, AEDANS labeled Ub monomers and F45W monomers were combined to form a K48 dimer with AEDANS on the distal Ub and F45W on the proximal Ub, formal name Ub(T12C^{AEDANS}) –^{48}Ub(F45W). This system provides a unique method to uncover more details from the interactions of K48 linked chains. To date most FRET applications of Ub have been used for crude methods such as gel assays or to determine DUB kinetics, but never for true distance measurements. A recent study has used single molecule FRET to determine the relative distribution of polyUb conformations (e.g. open and close in K48-Ub_2), but failed to produce any information on distances (*156*). With the F45W-AEDANS system it should be possible to accurately measure distances and obtain insightful structural information. To test this I titrated Ub(T12C^{AEDANS}) –^{48}Ub(F45W) with the K48 selective ligand UBA(2) from hHR23A. Upon binding, UBA(2) inserts between the K48 interface in the classic "sandwich mode," which should cause the distance between AEDANS and F45W to increase (*208*). Indeed the emission of AEDANS at 485 nm clearly shows this effect for titration of Ub(T12C^{AEDANS}) –^{48}Ub(F45W) with UBA(2) (**Figure 7.8**). As UBA(2) is added, the intensity of tryptophan emission at 340 nm increases, however intensity for AEDANS emission at 485 nm decreases. When Ub(T12C^{AEDANS}) –^{48}Ub(F45W) is fully saturated the decrease in AEDANS emission begins to level off. This trend is consistent with the binding mechanism and supports that the two Ubs must open to bind UBA(2), increasing their distance between each other. The assumed distance between AEDANS and F45W is 21 Å in the free form and 35 Å for the UBA(2) bound form. In the K48-Ub_2 system, these distances are ideal for detection by the FRET pair.

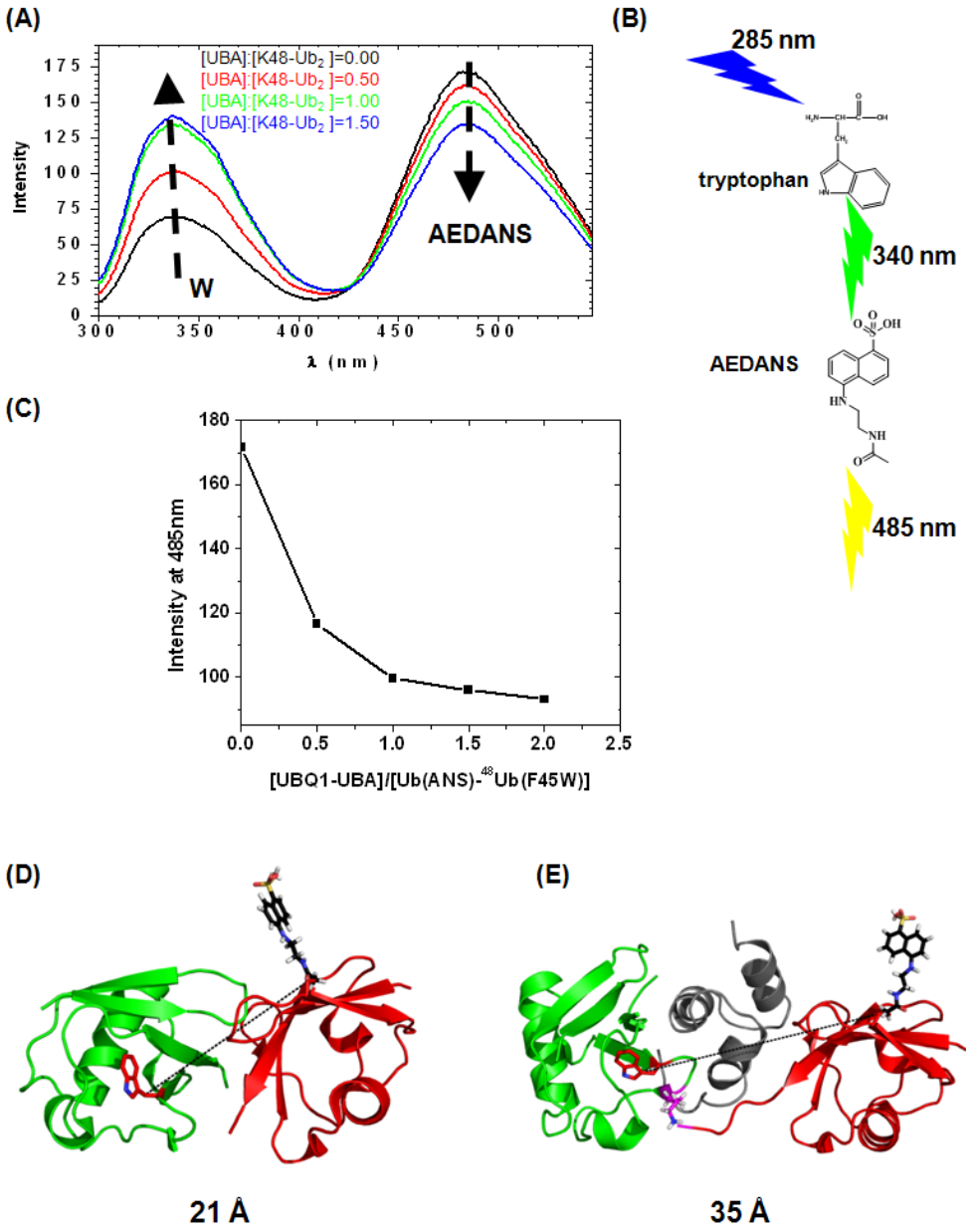


Figure 7.8 – FRET application of K48-Ub$_2$ and UBA(2).
(**A**) Emission spectra overlay for titration of Ub(T12C^{AEDANS}) –^{48}Ub(F45W) with UBA(2). With no UBA(2), the FRET emission signal from AEDANS at 485 nm is maximal, but as UBA(2) is added in the signal decreases due to UBA(2) separating the two Ubs. (**B**) Schematic showing the Förster resonance energy transfer of the excitation light through F45W and then through AEDANS. (**C**) Plot of the intensity of AEDANS emission shows a decrease that levels off as saturation is approached. (**D**) Model of Ub(T12C^{AEDANS}) –^{48}Ub(F45W) in the free state with proximal Ub (green) and distal Ub (red). The tryptophan (red sticks) is measured to be 21 Å away from the AEDANS (black sticks). (**D**) Model of the bound state with UBA(2) (gray) opening the proximal Ub (green) and distal Ub (red), increasing the distance between F45W and AEDANS to 35 Å.

7.3 Pathology of Ubb(+1) accumulation

7.3.1 Background and research aims

Abnormal ubiquitination patterns have been detected in Alzheimer's and other neurodegenerative polyglutamate diseases, Parkinson's and Huntington's (*246-249*). Given that the UPP is inhibited in these diseases, many suspect that a root cause results from the cytotoxicity of elevated levels of pathological proteins, which would normally be destroyed by the UPP (*250*). However, there is a debate as to whether these proteins are cytotoxic or merely a symptom brought on by other factors. Aside from the classic disease causing protein (e.g. alpha-synuclein, Parkin, or the Huntington's protein) there is frequently a disruption in the Ub gene itself. Specifically a frame shift mutation near the end of the Ub sequence causes a G76Y point mutation and a long 25 residue extension after G76Y of Ub, resulting in a new gene product termed Ubb(+1) (*251-253*). Recent work has uncovered that Ubb(+1) may in fact be more pathological than an aggregated protein such as alpha-synuclein (*254*). Physiologically this is a reasonable assumption given the concentration of Ub is relatively high (10μM) in the cell compared to other proteins. A great deal of *in vitro* work provides many hints as to how Ubb(+1) can be detrimental to cell function, specifically its ability to inhibit proteasomal degradation (*255, 256*). For our study we set to investigate how the Ubb(+1) gene product effected the UPP in yeast, investigate structural properties of Ubb(+1), and characterize how the mutant Ub functioned with different components of the UPP *in vitro*.

7.3.2 Synthesis and analysis of Ub–^{48}Ubb(+1)

Several variations of Ubb(+1) that differ in lengths of C-terminal extensions and also the residue at position 76 have been studied *in vitro (254, 255, 257, 258)*. Data from these studies suggests that the length of the C-terminal extension impacts how Ubb(+1) inhibits the proteasome and the residue at position 76 determines how efficiently DUBs can process Ubb(+1) conjugates. The main outcome of Ubb(+1) and other variants expressed for *in vivo* systems (mouse and drosophila) is marked by the accumulation of polyUb conjugates (*250, 252, 259*). The first solution studies of the Ubb(+1) structure by Cheryl Arrowsmith proved that the mutant Ub could bind natural partners such as E2-25K and the C-terminal extension was extremely flexible. Given that the Ubb(+1) mutant inhibited the proteasome I set to explore if the structural conformation of K48 linked Ubb(+1) conjugates deviated from those of the wild type. I used E2-25K to synthesize conjugates for study and found that Ubb(+1) was efficiently ligated (**Figure 7.9**). Intrigued by this I also explored if Ubb(+1) could be linked at other sites using Ubc13:Mms2 to make K63 linkages. Both E2s efficiently ligated K48 and K63 of Ubb(+1) suggesting that any problems caused by Ubb(+1) happen downstream of ubiquitination.

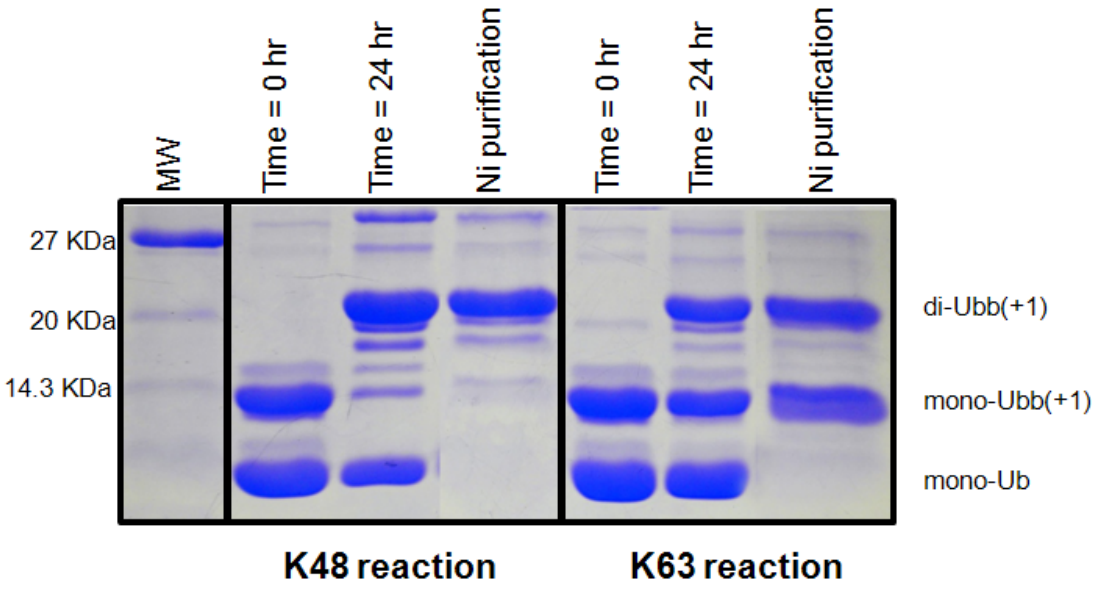


Figure 7.9 – Ubb(+1) is ligated in enzymatic K48 and K63 reactions.
The ability of Ubb(+1) to form conjugates was tested in enzymatic reactions. To assess the conjugates only K48R/K63R-Ub and an N-terminal 6xHis Ubb(+1) are used. The products are isolated with a His-Trap step. The 15% SDS-PAGE gel shows the K48 reaction with E2-25K, Ubb(+1) reacts very efficiently to form Ub–^{48}Ubb(+1) and the product is easily isolated. The right three lanes show Ubc13:Mms2 forms Ub–^{63}Ubb(+1) with some residual Ubb(+1) monomer and both Ubb(+1) species are isolated in the His-Trap step.

The specific K48 linked Ubb(+1) conjugate I chose to study by NMR was Ub(^{15}N)–^{48}Ubb(+1). By default, Ubb(+1) must be the proximal domain and the distal domain was ^{15}N enriched since its ^{1}H,^{15}N-HSQC spectrum would not be complicated by signals from residues in the C-terminal extension. In addition, this approach makes it much more feasible to test other Ubb(+1) variants with the same experiment. Since the structural properties of Ubb(+1) conjugates have not been studied in detail it was reasonable to start with the dimeric chain. Overlay of the ^{1}H,^{15}N-HSQC spectra for Ub(^{15}N)–^{48}Ubb(+1) and the monomeric Ub show many CSPs. The pattern of these CSPs suggests the distal Ub in Ub(^{15}N)–^{48}Ubb(+1) does indeed form the

classic hydrophobic interface (**Figure 7.10**). There are some subtle changes compared to wild type Ub(^{15}N)–^{48}Ub, but Ubb(+1) as a proximal Ub does not dramatically shift the overall conformation of Ub(^{15}N)–^{48}Ubb(+1) from that of the wild type. To gauge the nature of the hydrophobic interface ^{15}N T_1 relaxation was measured for the distal Ub in Ub(^{15}N)–^{48}Ubb(+1). If the interface is weakened there should be a decrease in T_1 and an increase if the interface is strengthened. The T_1 does appear to drop slightly (~100ms) suggesting the conformation is slightly shifted to the open form, however the classic K48 interface is still retained to some degree (**Figure 7.10**). Determining the actual structure of Ub(^{15}N)–^{48}Ubb(+1) will require a much more extensive investigation.

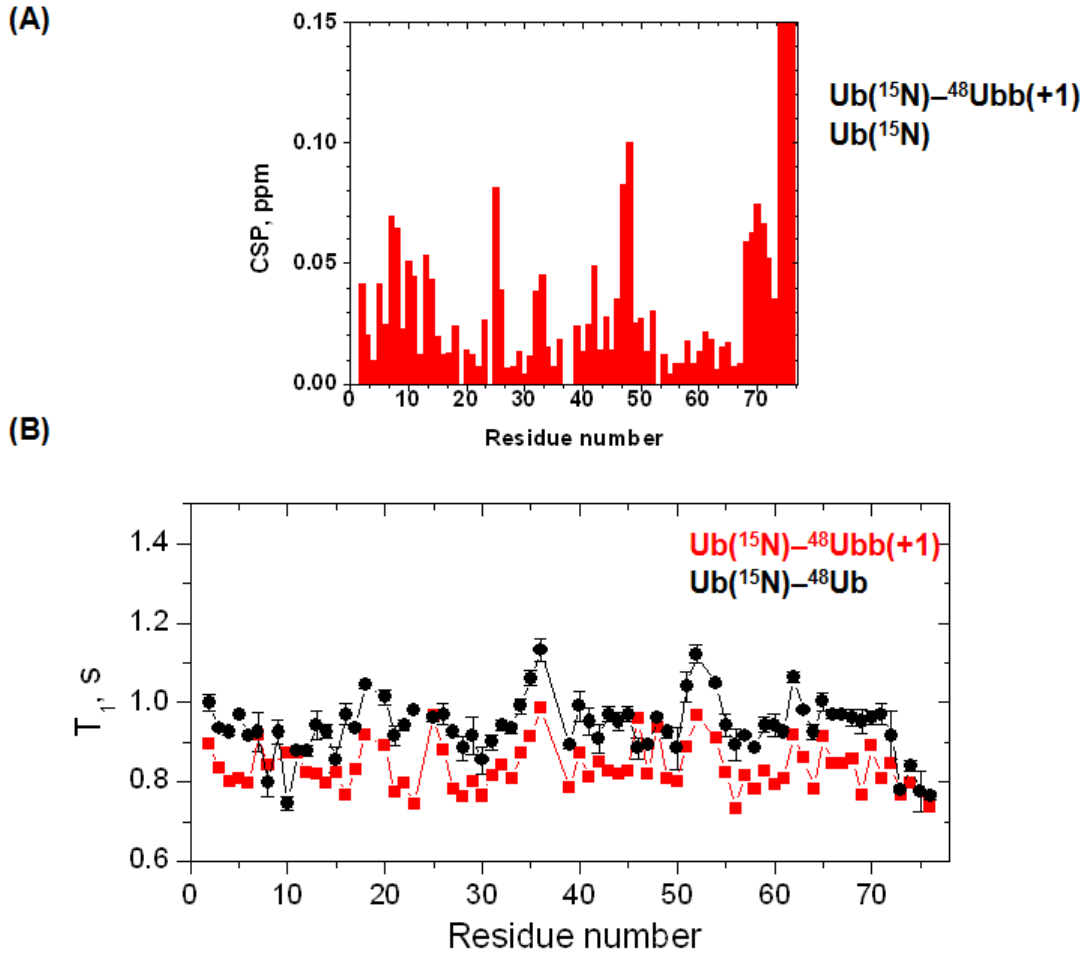


Figure 7.10 – Conformational properties of K48 linked Ubb(+1) conjugates. The ability of Ub(^{15}N)–^{48}Ubb(+1) to retain the classic K48 interface was tested with several NMR methods. (**A**) CSPs between ^{1}H,^{15}N-HSQC spectra of monomeric Ub and Ub(^{15}N)–^{48}Ubb(+1) were calculated showing distinct differences in the L8,I44,V70 residues of the hydrophobic patch. (**B**) ^{15}N T_1 between wild type Ub(^{15}N)–^{48}Ub (black circles) and Ub(^{15}N)–^{48}Ubb(+1) (red squares) show a similar profile in Ub, however Ub(^{15}N)–^{48}Ubb(+1) has a slightly shorter T_1 suggesting small conformational differences between the two forms of K48-Ub_2.

After determining that the overall conformation of K48 linked Ubb(+1) conjugates were similar to that of wild type, I then moved to investigate how DUBs would process these conjugates. Others have already demonstrated that UBDs and other binding partners could efficiently recognize Ubb(+1). Since Ubb(+1) inhibits

the proteasome I tested if the activity of Ubp6 would differ for Ub–^{48}Ubb(+1). If there is inhibition this can result for the Ubb(+1) conjugates changing the conformation of the polyUb chain, which impacts Ubp6 recognition. This effect has already been demonstrated for other Ub mutants that shift the conformational equilibrium of K48-Ub_2. Also, the tail of Ubb(+1) could be inhibitory as well. The results in (**Figure 7.11**) show that Ubp6 has absolutely no problem cleaving the K48 linkage in Ub–^{48}Ubb(+1) and it is removed essentially at the wild type Ub–^{48}Ub.

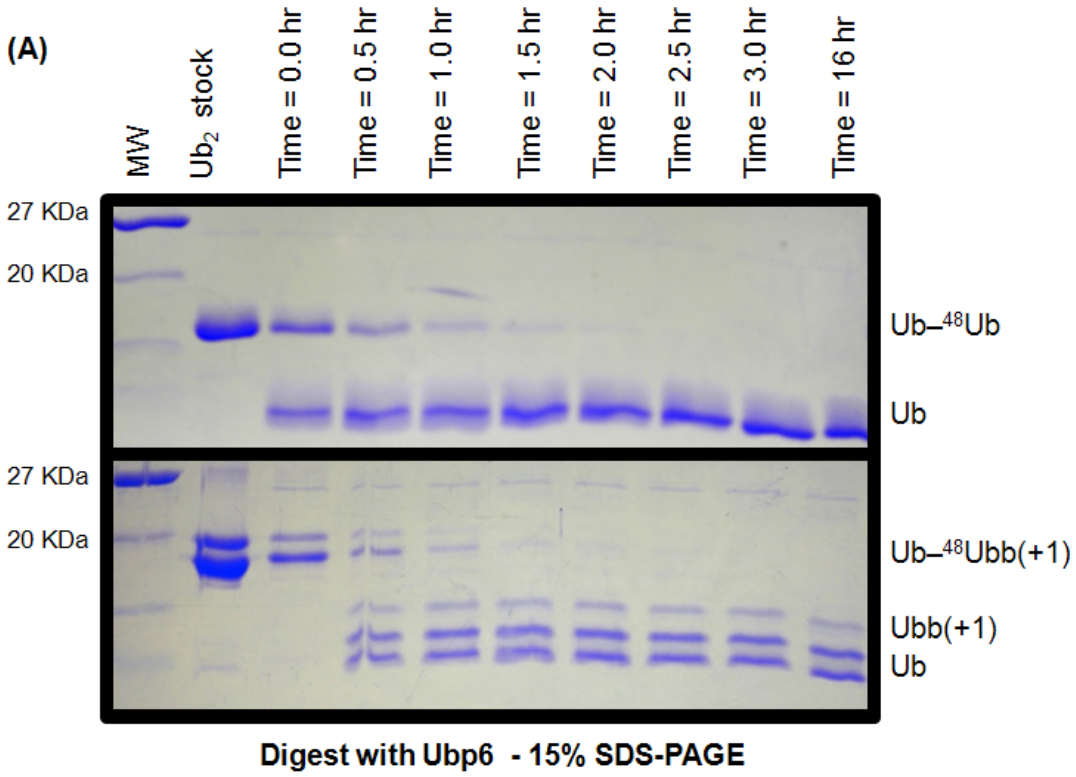


Figure 7.11 – Ub–^{48}Ubb(+1) does not interfere with Ubp6 activity.
15% SDS-PAGE gel showing cleavage of Ub–^{48}Ub and Ub–^{48}Ubb(+1) under identical conditions. After 1 hour essentially all of the dimeric Ub is converted to its component monomers.

7.3.3 Conclusions

For the first time Ubb(+1) conjugates have been isolated and analyzed structurally. My data shows that Ub–^{48}Ubb(+1) retains the classic K48 interface, but with a conformational equilibrium slightly favoring the open form. However, this does not appear to effect conjugation by E1,E2 enzymes or deconjugation by proteasomal DUBs. This does not mean that Ubb(+1) is harmless to the cell. Ubb(+1) conjugates could actually be pathological via several other mechanisms. Data from collaborators on this study also supports this. That fact that Ubb(+1) conjugates accumulate suggests other parts of cellular machinery fail to function properly. For example, the high number of Ubb(+1) conjugates may reach the proteasome and other pathways where the downstream machinery is inhibited by these atypical substrates. Ubp6 has been shown to recognize the distal Ub, however other DUBs which recognize the C-terminus of Ub have been shown to be inhibited by the atypical extension on Ubb(+1). This would certainly explain how the conjugates accumulate in the cell and it is likely that they also keep certain DUBs from functioning optimally. It is interesting that the rest of the polyUb chain attached to Ubb(+1) appears normal, thus the pathology of Ubb(+1) lies within the Ubb(+1) molecule itself and not its conjugates. The ability of Ubb(+1) to be formed into conjugates likely contributes to its pathology since once a regular polyUb signal is formed on Ubb(+1), the conjugate could be targeted to specific parts of the cell allowing Ubb(+1) to interfere with more pathways.

7.4 Monoubiquitin modification of histones

7.4.1 Background and research aims

The histone code refers to the hypothesis that many different post-translational modifications carried by histones serve to regulate gene expression and chromatin structure (*260*). Ubiquitination of histones has been well studied (*261-263*), however the structural properties of ubiquitinated histones have never been investigated. One of the most pressing issues is addressing how a post-translational modification affects both chromatin structure and transcription. In the case of Ub, a controversial hypothesis is that when attached to the tail of either H2A or H2B, the Ub modification can form specific interactions with individual histones in the nucleosome. If this is true Ub could potentially regulate specific DNA sequences in the genome. The structure of a nucleosome, containing eight histones, two of each H2A, H2B, H3, and H4 along with a 146 bp DNA fragment is show in (**Figure 7.12**). This structure is the most fundamental unit of chromatin and has a distinct arrangement of individual histones.

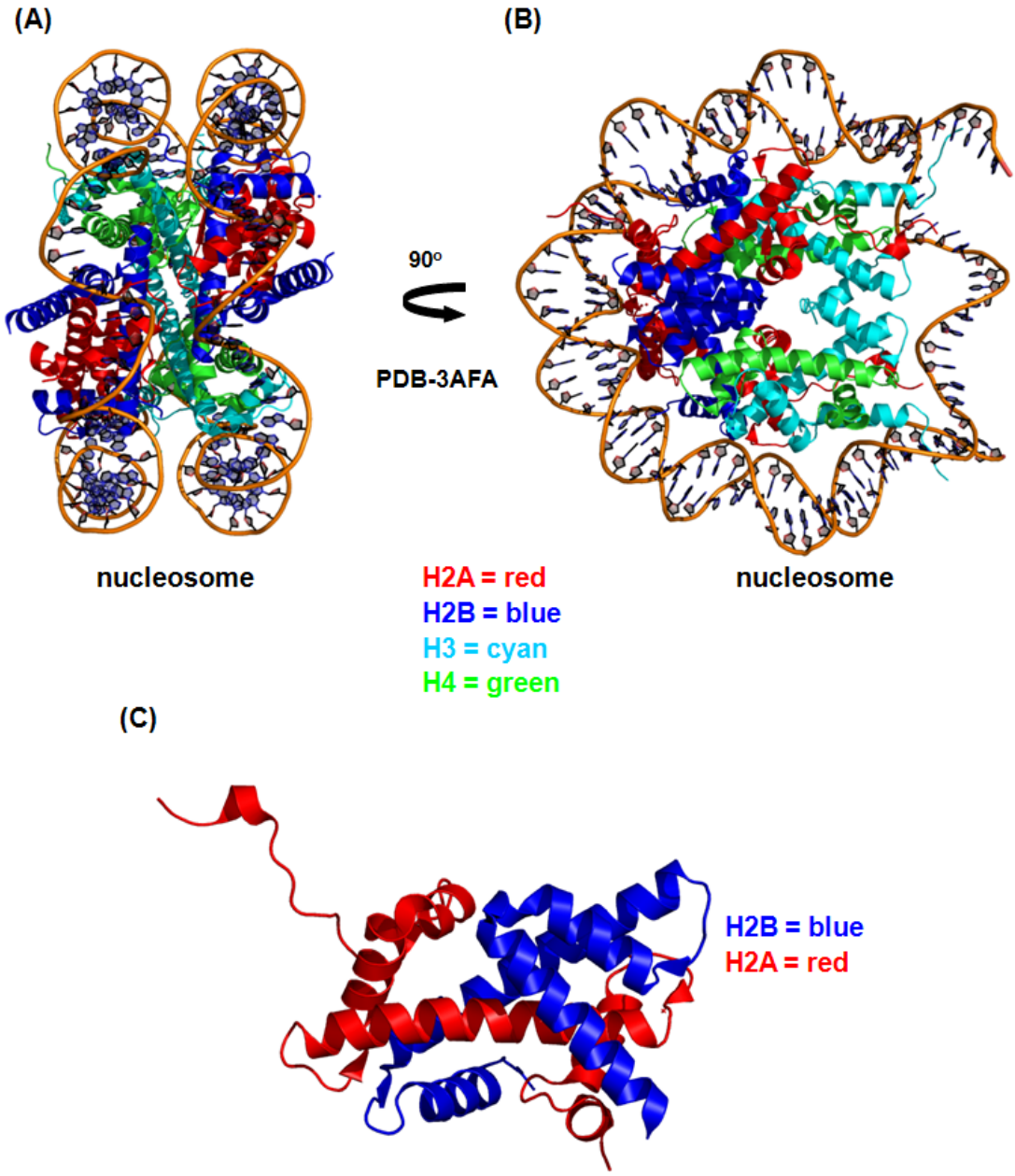


Figure 7.12 – structure of nucleosome with the H2A/H2B heterodimer.
An X-ray structure of a human nucleosome (PDB-3AFA) shows an arrangement with two tetrameric layers of histones and a single strand of DNA wrapped about both. (**A**) histones are shown as cartoons with H2A (red), H2B (blue), H3 (cyan), and H4 (green). (**B**) Another angle shows that H2A (red) and H2B (blue) are always in close proximity inside of the nucleosome. (**C**) All other components are removed and the arrangement of the H2A/H2B heterodimer is displayed.

Individual histones, H2A and H2B are the best studied and can take on a variety of post-translational modifications (*130, 131*). The acidic patch on H2A is a structural feature that has been demonstrated to participate in a variety of interactions (*264*). To test the hypothesis regarding the ubiquitination of histones, H2A and H2B would be a logical starting point. Thus the main objective of this study is to determine if Ub attached to H2A or H2B creates any detectable interactions.

7.4.2 Synthesis and analysis of monoubiquitinated H2A and H2B

The literature states that C-terminal tails of H2A and H2B harbor common sites of ubiquitination and are vital for chromatin regulation (*131, 260*). Specifically K119 in H2A and K120 in H2B, which led us to investigate the properties of mono-Ub attached to these positions. Lacking the ability to produce selectively ^{15}N enriched Ub-histone conjugates, we opted to use a chemical method relying on simple cross linking through cysteine residues. ^{15}N-Ub(G76C) with a 6xHis N-terminal tag was reacted with H2A(K119C) or H2B(K120C), then purified using His-Trap, followed by size exclusion. The opposite non-ubiquitinated histone was added and the H2A/H2B dimer formed (**Figure 7.13**). We wanted to study the Ub modification in this context since this is a more accurate representation of the nucleosome. This procedure resulted in the following two complexes for solution NMR studies, Ub(^{15}N)–^{K119C}H2A/H2B and Ub(^{15}N)–^{K120C}H2B/H2A each with a MW = 37.6 kDa.

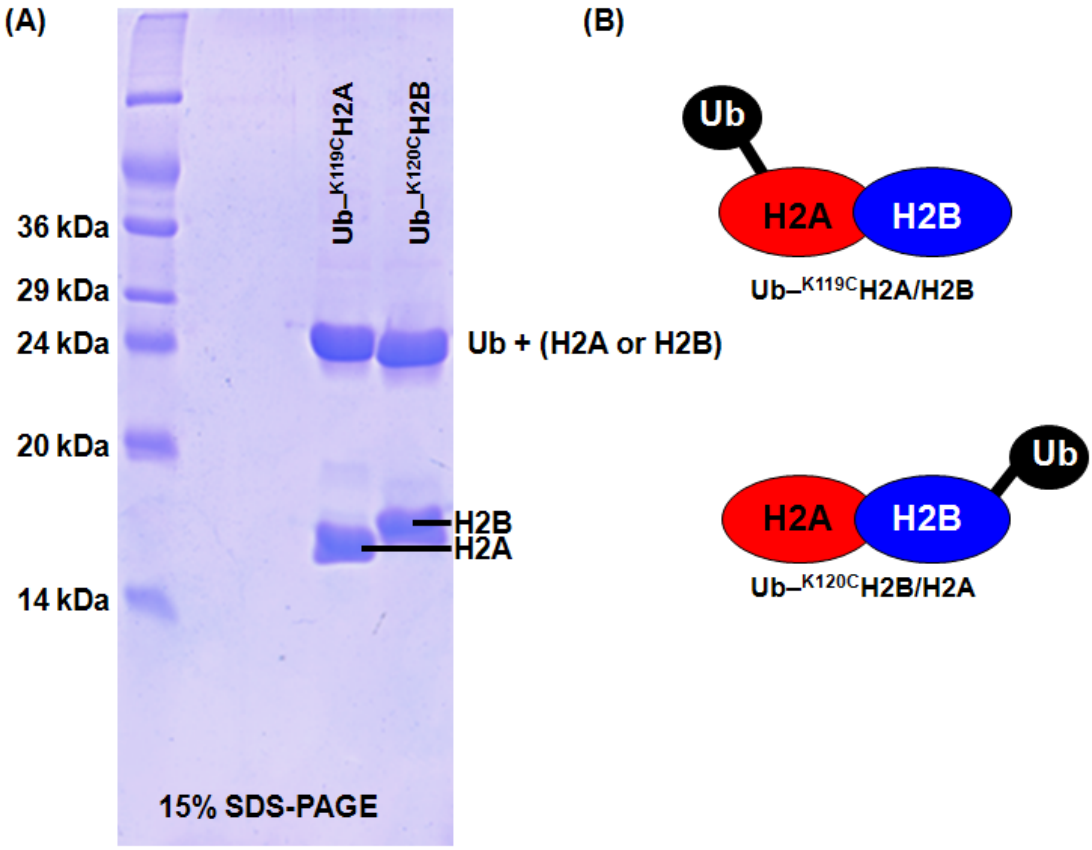


Figure 7.13 – ^{15}N-Ub is efficiently ligated to desired positions of H2A and H2B with simple cross-linking chemistry.
(**A**) A non-reducing 15% SDS-PAGE gel shows Ub(^{15}N)–^{K119C}H2A and Ub(^{15}N)–^{K120C}H2B are formed. The H2A/H2B complex is disrupted by SDS-PAGE, however the ^{15}N Ub remains attached to the respective histone. (**B**) Schematic representation of the Ub–^{K119C}H2A/H2B and Ub–^{K120C}H2B/H2A complexes with Ub attached.

To monitor any potential interactions that Ub could be created, a standard NMR approach detecting residue specific CSPs on ^{15}N-Ub in ^{1}H,^{15}N-HSQC was employed. CSPs were calculated between free monomeric 6xHis-Ub(G76C) and Ub attached to histone dimers, Ub(^{15}N)–^{K119C}H2A/H2B and Ub(^{15}N)–^{K120C}H2B/H2A. The results in (**Figure 7.14**) show that there are minimal CSPs for both, however slightly higher CSPs for Ub(^{15}N)–^{K119C}H2A. The tail region of Ub (residues 72-76) also appears to contain elevated CSPs, but this is most likely due to ligation of the Ub to either histone, not a specific interaction of Ub. These experiments were carried out at two

different salt (NaCl) concentrations 150 mM and 250 mM to ensure that we screen out any mild electrostatic interactions the acidic patch on H2A could form. However, the same result was observed in 150 mM and 250 mM NaCl.

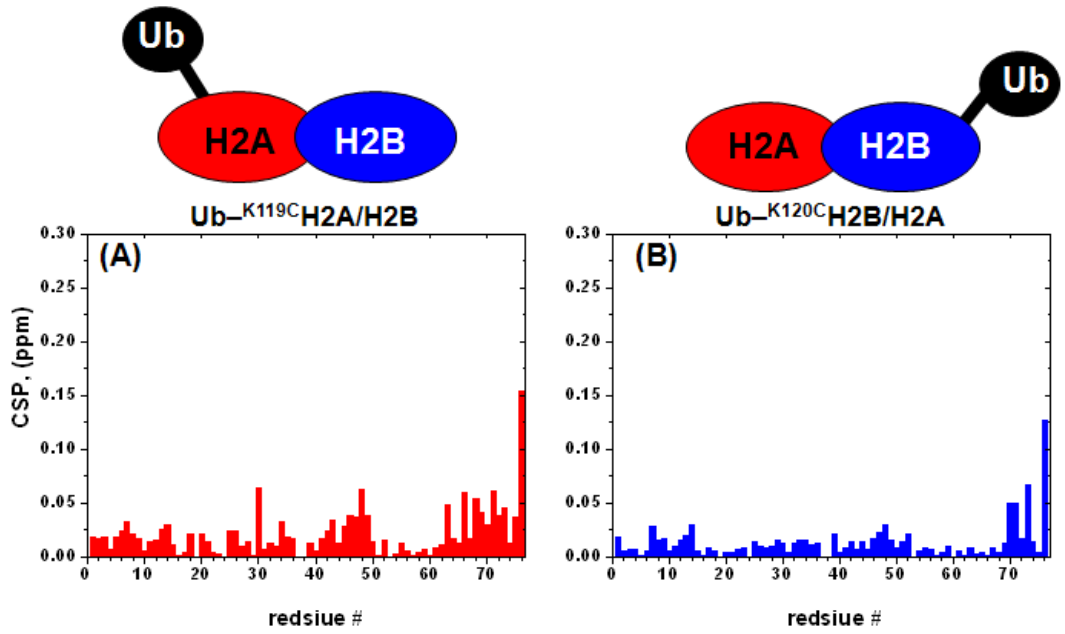


Figure 7.14 – Analysis of Ub modified H2A and H2B by $^{1}H,^{15}N$-HSQC.
If the Ub attached to either H2A or H2B is undergoing a significant interaction, there should be noticeable CSPs compared to the 6xHis-Ub(G76C) monomer, which both CSP plots use for comparison. (**A**) CSPs plot for Ub–^{K119C}H2A/H2B shows small CSPs for Ub in the C-terminal region. (**B**) Ub–^{K120C}H2B/H2A has even less CSPs and the tail region still shows the largest CSPs.

7.4.3 Conclusions

Ubiquitination of histones is an extremely important post-translational modification to understand. The exact roles of Ub are currently still unclear, but nevertheless Ub must serve some function when it is attached to histones. A few studies have shown that K63 polyUb chains form on histones at sites of DNA damage, which initiates DNA repair (*130*), however Ub appears to serve in many other roles than DNA repair for histones (*129*). This current investigation did not detect any strong interactions between the Ub on either Ub–^{K119C}H2A/H2B or Ub–^{K120C}H2B/H2A. But it is important to not rule out the possibility that Ub could still serve as a steric block to prevent other factors from interacting with histones. The approach used is sound and certainly should be extended beyond this study to address the structural properties Ub modified histones. There are many fruitful research paths to travel in this field.

Chapter 8: Summary and concluding remarks

8.1 Summary of significant results

The first structural study of mixed linkage polyUb chains presented here is the most significant biological concept of my doctoral work. The existence of mixed linkage chains goes against the traditional dogma in the field, but nevertheless is an important factor that cannot be ignored. It appears that an ever increasing number of studies are uncovering mixed linkage polyUb chains form *in vivo* systems and to understand the functions of these chains we will ultimately have to understand their structural properties. My theoretical study of branched chains explored at total of 28 out of an almost infinite number that can exist. In no way do my conclusions go as far as to say that each of these theoretical chains represents a new signal. In fact my findings, both theoretical and experimental support that even though many chains are possible, there is much degeneracy in their signaling properties. Analysis of the theoretical branched tri-Ub chains suggested that linkage mixing could lead to new polyUb signals, which I grouped into four distinct structural conformations. If linkage mixing does in fact present new signals, the structural conformations which I classified will be important for understanding how this is possible. The experimental study of K48 and K63 mixed linkage branched and unbranched chains: $[Ub]_2$–^{48,63}Ub, Ub–^{63}Ub–^{48}Ub, and Ub–^{48}Ub–^{63}Ub uncovered a common theme. Regardless of how the linkages were mixed, their individual properties were preserved. Structurally the K48 linked Ub pair always adopted the classic K48 interface characteristic of the homogeneous chain, while the lone K63 linked Ub retained the extended conformation of the homogeneous K63 linkage. After establishing that the native

structures were preserved in each mixed linkage chain, the binding properties of each linkage were also found to be preserved. Linkage selective receptors UBA(2) and tUIM each recognized the K48 and K63 linkage, respectively in $[Ub]_2$–^{48,63}Ub. Individual K48 and K63 linkages were then shown to retain the same properties for linkage selective DUBs and antibodies. After observing the same trend in each experiment, the conclusion that distinct K48 and K63 signals could be encoded in the same polyUb chain via linkage mixing was more than obvious. The physiological significance of these chains remains to be seen, but nevertheless this establishes an important principle in linkage mixing.

Determining the structure of the c112/Ub_1 complex along with the other details regarding ubistatin/Ub interactions in the process provides for a complete understanding of how ubistatins function. The concept of targeting Ub with small molecules is extremely important for therapeutic applications and the structural work presented here will help guide further development. The data show that half (c112) and full (c59) ubistatins bind the same site of Ub, but with differences in stoichiometry and affinity. Hydrophobic patch residues I44 and V70 and also nearby cationic residues R42 and R72 were shown to be essential for ubistatin/Ub interactions. This supports that ubistatins employ both a hydrophobic and an electrostatic binding mechanism. Titration data from ubistatin derivatives lacking sulfonate groups also supports these binding mechanisms. The inability of Rub1, which contains the same L8,I44,V70 hydrophobic patch as Ub to bind c112 supports that the ubistatin binding site on Ub is a unique drug target. Both half and full ubistatins are capable of binding either K48 or K63 linked polyUb chains with the

same binding mechanism, however the T_1 data suggest that c59 forms different complexes with K48-Ub_2 and K63-Ub_2. The ability of both c112 and c59 to inhibit the proteasomal DUB Ubp6 has implications for other DUBs as well and explains how ubistatins induce the accumulation of polyUb conjugates.

The new Asp cleavage method for the analysis of K63 polyUb conjugates is a real breakthrough for Ub proteomics. Aside from the novelty of measuring consecutive K63 linkages in unanchored or substrate attached polyUb chains, there are also broad implications for those studying Ub signaling with the traditional trypsin cleavage approach. The Asp method provides a clear path for determining how the K63 polyUb signal is presented in the cell, specifically where it is attached and the sequence of linkages. In addition, this method also provides a means for detection of mixed linkage polyUb chains, provided they contain a K63 linkage.

8.2 Outlook for future studies

Continuing work on mixed linkage chains is absolutely essential for the field, however new projects must be carried out logically and either address abstract concepts I failed to address or delve further into *in vivo* systems where mixed linkages have been reported. My study only addressed 2D branched chains and unbranched mixed linkage chains. It is possible that in the cell polyUb chains are highly branched, where a single Ub is simultaneously ubiquitinated at more than three positions. The greatest obstacle is deciding where to start and another major hurdle is the lack of methods available for the detection of mixed linkage chains. Given that my work represents the first structural study in the field, I am confident that there will

be many more routes to travel. It will be important to note the common trends as they are discovered, for example the observation that K48 linkages always retain their homogeneous characteristics. For the communication of future findings on this subject, the Cohen-Nakasone-Fushman nomenclature system will limit much confusion and facilitate the clear transfer of ideas. The study of mixed linkage chains should transition from an abstract idea to principles rooted in solid experimental data. The discovery of UBDs that preferentially bind mixed linkage chains would also go a long way in cementing their importance in the field. At the same time, mixed linkage chains representing the structural conformations I described should also be searched for with *in silico* methods or synthesized and assessed experimentally.

Given the importance of Ub signaling many more will attempt to target Ub with small molecules, whether they are ubistatin derivatives, gold nanoparticles, or of another design. The structure representing the c112/Ub_1 complex suggests that there is room to further derivatize c112 to better interact with the surface of Ub. Ubistatins should be approached from multiple angles for development, but the structural data cannot be ignored. With the structural knowledge, ubistatins can be modified with specialized chemistry to make them more favorable drugs without disrupting their activity. This will allow ubistatins to overcome their inherent drawbacks such as membrane permeability and solubility. Ub is the ideal drug target given its high conservation and inability to accommodate even subtle changes. The choice of Ub as a drug target is wise in that there is absolutely no mechanism for disease to acquire resistance to ubistatins. My failure to determine any structures for c59 in complex with Ub could very well be another researcher's success and is a worthwhile venture.

Furthermore, although ubistatins bind the same surface of Ub, only monomeric Ub was investigated. For chains longer than dimers it is possible that ubistatins could also form new complexes and the structures of ubistatins in complex with different polyUb linkages should be seriously evaluated.

References

1. Kaftory, A., Hershko, A., and Fry, M. (1978) Protein turnover in senescent cultured chick embryo fibroblasts, *J Cell Physiol 94*, 147-160.
2. Hadden, J. W. (1975) Thymopoietin, ubiquitin and the differentiation of lymphocytes, *Clin Bull 5*, 66-67.
3. Hershko, A., and Fry, M. (1975) Post-translational cleavage of polypeptide chains: role in assembly, *Annu Rev Biochem 44*, 775-797.
4. Hershko, A., and Ciechanover, A. (1982) Mechanisms of intracellular protein breakdown, *Annu Rev Biochem 51*, 335-364.
5. Ciechanover, A., Heller, H., Katz-Etzion, R., and Hershko, A. (1981) Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system, *Proc Natl Acad Sci U S A 78*, 761-765.
6. Hershko, A., Ciechanover, A., and Rose, I. A. (1981) Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown, *J Biol Chem 256*, 1525-1528.
7. 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.
8. 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.
9. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., and Rose, I. A. (1980) Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis, *Proc Natl Acad Sci U S A 77*, 1783-1786.
10. Hershko, A., Ciechanover, A., and Rose, I. A. (1979) Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP, *Proc Natl Acad Sci U S A 76*, 3107-3110.
11. Ciechanover, A., and Hershko, A. (1976) Early effects of serum on phospholipid metabolism in untransformed and oncogenic virus-transformed cultured fibroblasts, *Biochem Biophys Res Commun 73*, 85-91.
12. Wilkinson, K. D. (2005) The discovery of ubiquitin-dependent proteolysis, *Proc Natl Acad Sci U S A 102*, 15280-15282.
13. Gregori, L., Poosch, M. S., Cousins, G., and Chau, V. (1990) A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis, *J Biol Chem 265*, 8354-8357.
14. Komander, D. (2009) The emerging complexity of protein ubiquitination, *Biochemical Society transactions 37*, 937-953.
15. Piotrowski, J., Beal, R., Hoffman, L., Wilkinson, K. D., Cohen, R. E., and Pickart, C. M. (1997) Inhibition of the 26 S proteasome by polyubiquitin chains synthesized to have defined lengths, *J Biol Chem 272*, 23712-23721.
16. Pickart, C. M. (1997) Targeting of substrates to the 26S proteasome, *FASEB J 11*, 1055-1066.

17. Wilkinson, K. D. (1997) Regulation of ubiquitin-dependent processes by deubiquitinating enzymes, *FASEB J 11*, 1245-1256.
18. Komander, D., Clague, M. J., and Urbe, S. (2009) Breaking the chains: structure and function of the deubiquitinases, *Nat Rev Mol Cell Biol 10*, 550-563.
19. Ardley, H. C., and Robinson, P. A. (2005) E3 ubiquitin ligases, *Essays in biochemistry 41*, 15-30.
20. 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, In *J Biol Chem*, pp 17375-17386, United States.
21. Passmore, L. A., and Barford, D. (2004) Getting into position: the catalytic mechanisms of protein ubiquitylation, *Biochem J 379*, 513-525.
22. Barford, D. (2011) Structure, function and mechanism of the anaphase promoting complex (APC/C), *Q Rev Biophys 44*, 153-190.
23. Kuhlbrodt, K., Mouysset, J., and Hoppe, T. (2005) Orchestra for assembly and fate of polyubiquitin chains, *Essays in biochemistry 41*, 1-14.
24. Gonen, H., Dickman, D., Schwartz, A. L., and Ciechanover, A. (1996) Protein synthesis elongation factor EF-1 alpha is an isopeptidase essential for ubiquitin-dependent degradation of certain proteolytic substrates, *Adv Exp Med Biol 389*, 209-219.
25. Wang, X., and Jiang, X. (2012) Mdm2 and MdmX partner to regulate p53, *FEBS Lett 586*, 1390-1396.
26. Kim, H. T., and Goldberg, A. L. (2012) Formation of nondegradable forked ubiquitin conjugates by ring-finger ligases and its prevention by S5a, *Methods Mol Biol 832*, 639-652.
27. Uchiki, T., Kim, H. T., Zhai, B., Gygi, S. P., Johnston, J. A., O'Bryan, J. P., and Goldberg, A. L. (2009) The ubiquitin-interacting motif protein, S5a, is ubiquitinated by all types of ubiquitin ligases by a mechanism different from typical substrate recognition, *J Biol Chem 284*, 12622-12632.
28. Kim, H. T., Kim, K. P., Uchiki, T., Gygi, S. P., and Goldberg, A. L. (2009) S5a promotes protein degradation by blocking synthesis of nondegradable forked ubiquitin chains, *EMBO J 28*, 1867-1877.
29. 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.
30. Finley, D., Ulrich, H. D., Sommer, T., and Kaiser, P. (2012) The ubiquitin-proteasome system of Saccharomyces cerevisiae, *Genetics 192*, 319-360.
31. Hu, M., Li, P., Song, L., Jeffrey, P. D., Chenova, T. A., Wilkinson, K. D., Cohen, R. E., and Shi, Y. (2005) Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14, *EMBO J 24*, 3747-3756.

32. Ye, Y., Akutsu, M., Reyes-Turcu, F., Enchev, R. I., Wilkinson, K. D., and Komander, D. (2011) Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21, *EMBO reports 12*, 350-357.
33. Akutsu, M., Ye, Y., Virdee, S., Chin, J. W., and Komander, D. (2011) Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains, *Proc Natl Acad Sci U S A 108*, 2228-2233.
34. Davies, C. W., Paul, L. N., Kim, M. I., and Das, C. (2011) Structural and thermodynamic comparison of the catalytic domain of AMSH and AMSH-LP: nearly identical fold but different stability, *J Mol Biol 413*, 416-429.
35. Yu, H. A., Kim, S. G., Kim, E. J., Lee, W. J., Kim, D. O., Park, K., Park, Y. C., and Seo, J. H. (2007) Characterization of ubiquitin C-terminal hydrolase 1 (YUH1) from Saccharomyces cerevisiae expressed in recombinant Escherichia coli, *Protein Expr Purif 56*, 20-26.
36. Pickart, C. M., and Rose, I. A. (1985) Ubiquitin carboxyl-terminal hydrolase acts on ubiquitin carboxyl-terminal amides, *J Biol Chem 260*, 7903-7910.
37. Pickart, C. M., and Rose, I. A. (1986) Mechanism of ubiquitin carboxyl-terminal hydrolase. Borohydride and hydroxylamine inactivate in the presence of ubiquitin, *J Biol Chem 261*, 10210-10217.
38. Wilkinson, K. D., Deshpande, S., and Larsen, C. N. (1992) Comparisons of neuronal (PGP 9.5) and non-neuronal ubiquitin C-terminal hydrolases, *Biochemical Society transactions 20*, 631-637.
39. Larsen, C. N., Krantz, B. A., and Wilkinson, K. D. (1998) Substrate specificity of deubiquitinating enzymes: ubiquitin C-terminal hydrolases, *Biochemistry 37*, 3358-3368.
40. Johnston, S. C., Riddle, S. M., Cohen, R. E., and Hill, C. P. (1999) Structural basis for the specificity of ubiquitin C-terminal hydrolases, *EMBO J 18*, 3877-3887.
41. Komander, D., Lord, C. J., Scheel, H., Swift, S., Hofmann, K., Ashworth, A., and Barford, D. (2008) The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module, *Mol Cell 29*, 451-464.
42. Faesen, A. C., Luna-Vargas, M. P., and Sixma, T. K. (2012) The role of UBL domains in ubiquitin-specific proteases, *Biochemical Society transactions 40*, 539-545.
43. 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 reports 10*, 466-473.
44. Komander, D. (2010) CYLD tidies up dishevelled signaling, *Mol Cell 37*, 589-590.
45. Virdee, S., Ye, Y., Nguyen, D. P., Komander, D., and Chin, J. W. (2010) Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase, *Nature chemical biology 6*, 750-757.
46. Bremm, A., and Komander, D. (2011) Emerging roles for Lys11-linked polyubiquitin in cellular regulation, *Trends in biochemical sciences 36*, 355-363.

47. Komander, D., and Barford, D. (2008) Structure of the A20 OTU domain and mechanistic insights into deubiquitination, *Biochem J 409*, 77-85.
48. Wang, T., Yin, L., Cooper, E. M., Lai, M. Y., Dickey, S., Pickart, C. M., Fushman, D., Wilkinson, K. D., Cohen, R. E., and Wolberger, C. (2009) Evidence for bidentate substrate binding as the basis for the K48 linkage specificity of otubain 1, In *J Mol Biol*, pp 1011-1023, England.
49. Yao, T., and Cohen, R. E. (2002) A cryptic protease couples deubiquitination and degradation by the proteasome, *Nature 419*, 403-407.
50. Sato, Y., Yoshikawa, A., Yamagata, A., Mimura, H., Yamashita, M., Ookata, K., Nureki, O., Iwai, K., Komada, M., and Fukai, S. (2008) Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains, *Nature 455*, 358-362.
51. Ambroggio, X. I., Rees, D. C., and Deshaies, R. J. (2004) JAMM: a metalloprotease-like zinc site in the proteasome and signalosome, *PLoS biology 2*, E2.
52. Cooper, E. M., Cutcliffe, C., Kristiansen, T. Z., Pandey, A., Pickart, C. M., and Cohen, R. E. (2009) K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1, In *EMBO J*, pp 621-631, England.
53. Wei, N., and Deng, X. W. (2003) The COP9 signalosome, *Annual review of cell and developmental biology 19*, 261-286.
54. Lee, M. J., Lee, B. H., Hanna, J., King, R. W., and Finley, D. (2011) Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes, In *Molecular & cellular proteomics : MCP*, p R110 003871, United States.
55. Wrigley, J. D., Eckersley, K., Hardern, I. M., Millard, L., Walters, M., Peters, S. W., Mott, R., Nowak, T., Ward, R. A., Simpson, P. B., and Hudson, K. (2011) Enzymatic characterisation of USP7 deubiquitinating activity and inhibition, *Cell Biochem Biophys 60*, 99-111.
56. Hussain, S., Zhang, Y., and Galardy, P. J. (2009) DUBs and cancer: the role of deubiquitinating enzymes as oncogenes, non-oncogenes and tumor suppressors, *Cell cycle (Georgetown, Tex.) 8*, 1688-1697.
57. Singhal, S., Taylor, M. C., and Baker, R. T. (2008) Deubiquitylating enzymes and disease, *BMC biochemistry 9 Suppl 1*, S3.
58. Cummins, J. M., Rago, C., Kohli, M., Kinzler, K. W., Lengauer, C., and Vogelstein, B. (2004) Tumour suppression: disruption of HAUSP gene stabilizes p53, *Nature 428*, 1 p following 486.
59. Nicholson, B., and Suresh Kumar, K. G. (2011) The multifaceted roles of USP7: new therapeutic opportunities, *Cell Biochem Biophys 60*, 61-68.
60. Cheon, K. W., and Baek, K. H. (2006) HAUSP as a therapeutic target for hematopoietic tumors (review), *International journal of oncology 28*, 1209-1215.
61. Ikeda, F., and Dikic, I. (2006) CYLD in ubiquitin signaling and tumor pathogenesis, *Cell 125*, 643-645.
62. Melly, L., Lawton, G., and Rajan, N. (2012) Basal cell carcinoma arising in association with trichoepithelioma in a case of Brooke-Spiegler syndrome

with a novel genetic mutation in CYLD, *Journal of cutaneous pathology 39*, 977-978.
63. Massoumi, R. (2011) CYLD: a deubiquitination enzyme with multiple roles in cancer, *Future Oncol 7*, 285-297.
64. McCullough, J., Clague, M. J., and Urbe, S. (2004) AMSH is an endosome-associated ubiquitin isopeptidase, *J Cell Biol 166*, 487-492.
65. Suzuki, S., Tamai, K., Watanabe, M., Kyuuma, M., Ono, M., Sugamura, K., and Tanaka, N. (2011) AMSH is required to degrade ubiquitinated proteins in the central nervous system, *Biochem Biophys Res Commun 408*, 582-588.
66. Sierra, M. I., Wright, M. H., and Nash, P. D. (2010) AMSH interacts with ESCRT-0 to regulate the stability and trafficking of CXCR4, *J Biol Chem 285*, 13990-14004.
67. Meijer, I. M., van Rotterdam, W., van Zoelen, E. J., and van Leeuwen, J. E. (2012) Recycling of EGFR and ErbB2 is associated with impaired Hrs tyrosine phosphorylation and decreased deubiquitination by AMSH, *Cellular signalling 24*, 1981-1988.
68. Pareja, F., Ferraro, D. A., Rubin, C., Cohen-Dvashi, H., Zhang, F., Aulmann, S., Ben-Chetrit, N., Pines, G., Navon, R., Crosetto, N., Kostler, W., Carvalho, S., Lavi, S., Schmitt, F., Dikic, I., Yakhini, Z., Sinn, P., Mills, G. B., and Yarden, Y. (2012) Deubiquitination of EGFR by Cezanne-1 contributes to cancer progression, *Oncogene 31*, 4599-4608.
69. Hurtley, S. M. (1996) Lysosomal degradation of ubiquitin-tagged receptors, *Science 271*, 617.
70. Wolf, D. H. (2004) From lysosome to proteasome: the power of yeast in the dissection of proteinase function in cellular regulation and waste disposal, *Cellular and molecular life sciences : CMLS 61*, 1601-1614.
71. Fischer, M., Hilt, W., Richter-Ruoff, B., Gonen, H., Ciechanover, A., and Wolf, D. H. (1994) The 26S proteasome of the yeast Saccharomyces cerevisiae, *FEBS Lett 355*, 69-75.
72. Peters, J. M., Cejka, Z., Harris, J. R., Kleinschmidt, J. A., and Baumeister, W. (1993) Structural features of the 26 S proteasome complex, *J Mol Biol 234*, 932-937.
73. Jung, T., and Grune, T. (2012) Structure of the proteasome, *Prog Mol Biol Transl Sci 109*, 1-39.
74. Glickman, M. H., Rubin, D. M., Fried, V. A., and Finley, D. (1998) The regulatory particle of the Saccharomyces cerevisiae proteasome, *Mol Cell Biol 18*, 3149-3162.
75. Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V. A., and Finley, D. (1998) A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3, *Cell 94*, 615-623.
76. Baldwin, A. J., Religa, T. L., Hansen, D. F., Bouvignies, G., and Kay, L. E. (2010) 13CHD2 methyl group probes of millisecond time scale exchange in proteins by 1H relaxation dispersion: an application to proteasome gating residue dynamics, *J Am Chem Soc 132*, 10992-10995.

77. Tai, H. C., Besche, H., Goldberg, A. L., and Schuman, E. M. (2010) Characterization of the Brain 26S Proteasome and its Interacting Proteins, *Frontiers in molecular neuroscience 3*.
78. Kajava, A. V. (2002) What curves alpha-solenoids? Evidence for an alpha-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome, *J Biol Chem 277*, 49791-49798.
79. Elsasser, S., Gali, R. R., Schwickart, M., Larsen, C. N., Leggett, D. S., Muller, B., Feng, M. T., Tubing, F., Dittmar, G. A., and Finley, D. (2002) Proteasome subunit Rpn1 binds ubiquitin-like protein domains, *Nat Cell Biol 4*, 725-730.
80. 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.
81. Rubin, D. M., van Nocker, S., Glickman, M., Coux, O., Wefes, I., Sadis, S., Fu, H., Goldberg, A., Vierstra, R., and Finley, D. (1997) ATPase and ubiquitin-binding proteins of the yeast proteasome, *Mol Biol Rep 24*, 17-26.
82. Leggett, D. S., Glickman, M. H., and Finley, D. (2005) Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast, *Methods Mol Biol 301*, 57-70.
83. Sprangers, R., Li, X., Mao, X., Rubinstein, J. L., Schimmer, A. D., and Kay, L. E. (2008) TROSY-based NMR evidence for a novel class of 20S proteasome inhibitors, *Biochemistry 47*, 6727-6734.
84. Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N., and Tsukihara, T. (2002) The structure of the mammalian 20S proteasome at 2.75 A resolution, *Structure 10*, 609-618.
85. Rosenzweig, R., Bronner, V., Zhang, D., Fushman, D., and Glickman, M. H. (2012) Rpn1 and Rpn2 coordinate ubiquitin processing factors at proteasome, *J Biol Chem 287*, 14659-14671.
86. Sakata, E., Bohn, S., Mihalache, O., Kiss, P., Beck, F., Nagy, I., Nickell, S., Tanaka, K., Saeki, Y., Forster, F., and Baumeister, W. (2012) Localization of the proteasomal ubiquitin receptors Rpn10 and Rpn13 by electron cryomicroscopy, *Proc Natl Acad Sci U S A 109*, 1479-1484.
87. Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J. M., Villa, E., Baumeister, W., and Forster, F. (2012) Near-atomic resolution structural model of the yeast 26S proteasome, *Proc Natl Acad Sci U S A 109*, 14870-14875.
88. Nickell, S., Mihalache, O., Beck, F., Hegerl, R., Korinek, A., and Baumeister, W. (2007) Structural analysis of the 26S proteasome by cryoelectron tomography, *Biochem Biophys Res Commun 353*, 115-120.
89. Nickell, S., Beck, F., Korinek, A., Mihalache, O., Baumeister, W., and Plitzko, J. M. (2007) Automated cryoelectron microscopy of "single particles" applied to the 26S proteasome, *FEBS Lett 581*, 2751-2756.
90. Shabek, N., Herman-Bachinsky, Y., Buchsbaum, S., Lewinson, O., Haj-Yahya, M., Hejjaoui, M., Lashuel, H. A., Sommer, T., Brik, A., and Ciechanover, A. (2012) The size of the proteasomal substrate determines whether its degradation will be mediated by mono- or polyubiquitylation, *Mol Cell 48*, 87-97.

91. Jariel-Encontre, I., Bossis, G., and Piechaczyk, M. (2008) Ubiquitin-independent degradation of proteins by the proteasome, *Biochim Biophys Acta 1786*, 153-177.
92. Jacobson, A. D., Zhang, N. Y., Xu, P., Han, K. J., Noone, S., Peng, J., and Liu, C. W. (2009) The lysine 48 and lysine 63 ubiquitin conjugates are processed differently by the 26 s proteasome, *J Biol Chem 284*, 35485-35494.
93. Peth, A., Uchiki, T., and Goldberg, A. L. (2010) ATP-dependent steps in the binding of ubiquitin conjugates to the 26S proteasome that commit to degradation, *Mol Cell 40*, 671-681.
94. Nathan, J. A., Tae Kim, H., Ting, L., Gygi, S. P., and Goldberg, A. L. (2013) Why do cellular proteins linked to K63-polyubiquitin chains not associate with proteasomes?, *EMBO J*.
95. Zhang, D., Chen, T., Ziv, I., Rosenzweig, R., Matiuhin, Y., Bronner, V., Glickman, M. H., and Fushman, D. (2009) Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor, *Mol Cell 36*, 1018-1033.
96. Elsasser, S., Chandler-Militello, D., Muller, B., Hanna, J., and Finley, D. (2004) Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome, In *J Biol Chem*, pp 26817-26822, United States.
97. Raasi, S., and Pickart, C. M. (2003) Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains, *J Biol Chem 278*, 8951-8959.
98. Chen, L., and Madura, K. (2002) Rad23 promotes the targeting of proteolytic substrates to the proteasome, *Mol Cell Biol 22*, 4902-4913.
99. Walters, K. J., and Zhang, N. (2008) Rpn10 protects the proteasome from Dsk2, *Mol Cell 32*, 459-460.
100. Hanna, J., Hathaway, N. A., Tone, Y., Crosas, B., Elsasser, S., Kirkpatrick, D. S., Leggett, D. S., Gygi, S. P., King, R. W., and Finley, D. (2006) Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation, In *Cell*, pp 99-111, United States.
101. Finley, D. (2011) Misfolded proteins driven to destruction by Hul5, *Nat Cell Biol 13*, 1290-1292.
102. Braun, B. C., Glickman, M., Kraft, R., Dahlmann, B., Kloetzel, P. M., Finley, D., and Schmidt, M. (1999) The base of the proteasome regulatory particle exhibits chaperone-like activity, *Nat Cell Biol 1*, 221-226.
103. Savol, A. J., Burger, V. M., Agarwal, P. K., Ramanathan, A., and Chennubhotla, C. S. (2011) QAARM: quasi-anharmonic autoregressive model reveals molecular recognition pathways in ubiquitin, *Bioinformatics 27*, i52-60.
104. Liu, Y., Choudhury, P., Cabral, C. M., and Sifers, R. N. (1999) Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome, *J Biol Chem 274*, 5861-5867.
105. Kisselev, A. F., Akopian, T. N., Woo, K. M., and Goldberg, A. L. (1999) The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation, *J Biol Chem 274*, 3363-3371.

106. Jager, S., Groll, M., Huber, R., Wolf, D. H., and Heinemeyer, W. (1999) Proteasome beta-type subunits: unequal roles of propeptides in core particle maturation and a hierarchy of active site function, *J Mol Biol 291*, 997-1013.
107. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules, *Cell 78*, 761-771.
108. Gelman, J. S., Sironi, J., Berezniuk, I., Dasgupta, S., Castro, L. M., Gozzo, F. C., Ferro, E. S., and Fricker, L. D. (2013) Alterations of the intracellular peptidome in response to the proteasome inhibitor bortezomib, *PLoS One 8*, e53263.
109. Lopes, U. G., Erhardt, P., Yao, R., and Cooper, G. M. (1997) p53-dependent induction of apoptosis by proteasome inhibitors, *J Biol Chem 272*, 12893-12896.
110. Moreau, P., Richardson, P. G., Cavo, M., Orlowski, R. Z., San Miguel, J. F., Palumbo, A., and Harousseau, J. L. (2012) Proteasome inhibitors in multiple myeloma: ten years later, *Blood*.
111. Matthews, W., Driscoll, J., Tanaka, K., Ichihara, A., and Goldberg, A. L. (1989) Involvement of the proteasome in various degradative processes in mammalian cells, *Proc Natl Acad Sci U S A 86*, 2597-2601.
112. Thompson, J. L. (2013) Carfilzomib: a second-generation proteasome inhibitor for the treatment of relapsed and refractory multiple myeloma, *The Annals of pharmacotherapy 47*, 56-62.
113. Steele, J. M. (2013) Carfilzomib: A new proteasome inhibitor for relapsed or refractory multiple myeloma, *Journal of oncology pharmacy practice : official publication of the International Society of Oncology Pharmacy Practitioners*.
114. Nooka, A., Gleason, C., Casbourne, D., and Lonial, S. (2013) Relapsed and refractory lymphoid neoplasms and multiple myeloma with a focus on carfilzomib, *Biologics : targets & therapy 7*, 13-32.
115. Kortuem, K. M., and Stewart, A. K. (2013) Carfilzomib, *Blood 121*, 893-897.
116. Cvek, B. (2012) Proteasome inhibitors, *Prog Mol Biol Transl Sci 109*, 161-226.
117. Ruschak, A. M., Slassi, M., Kay, L. E., and Schimmer, A. D. (2011) Novel proteasome inhibitors to overcome bortezomib resistance, *Journal of the National Cancer Institute 103*, 1007-1017.
118. Blackburn, C., Gigstad, K. M., Hales, P., Garcia, K., Jones, M., Bruzzese, F. J., Barrett, C., Liu, J. X., Soucy, T. A., Sappal, D. S., Bump, N., Olhava, E. J., Fleming, P., Dick, L. R., Tsu, C., Sintchak, M. D., and Blank, J. L. (2010) Characterization of a new series of non-covalent proteasome inhibitors with exquisite potency and selectivity for the 20S beta5-subunit, *Biochem J 430*, 461-476.
119. Hanna, J., Hathaway, N. A., Tone, Y., Crosas, B., Elsasser, S., Kirkpatrick, D. S., Leggett, D. S., Gygi, S. P., King, R. W., and Finley, D. (2006) Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation, *Cell 127*, 99-111.

120. Peth, A., Besche, H. C., and Goldberg, A. L. (2009) Ubiquitinated proteins activate the proteasome by binding to Usp14/Ubp6, which causes 20S gate opening, *Mol Cell 36*, 794-804.
121. Qin, S., Wang, Q., Ray, A., Wani, G., Zhao, Q., Bhaumik, S. R., and Wani, A. A. (2009) Sem1p and Ubp6p orchestrate telomeric silencing by modulating histone H2B ubiquitination and H3 acetylation, *Nucleic acids research 37*, 1843-1853.
122. D'Arcy, P., and Linder, S. (2012) Proteasome deubiquitinases as novel targets for cancer therapy, *Int J Biochem Cell Biol*.
123. Hoeller, D., and Dikic, I. (2009) Targeting the ubiquitin system in cancer therapy, *Nature 458*, 438-444.
124. Lee, B. H., Lee, M. J., Park, S., Oh, D. C., Elsasser, S., Chen, P. C., Gartner, C., Dimova, N., Hanna, J., Gygi, S. P., Wilson, S. M., King, R. W., and Finley, D. (2010) Enhancement of proteasome activity by a small-molecule inhibitor of USP14, *Nature 467*, 179-184.
125. Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z. J. (2000) Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain, *Cell 103*, 351-361.
126. Hofmann, R. M., and Pickart, C. M. (1999) Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair, *Cell 96*, 645-653.
127. Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response, *Science 316*, 1194-1198.
128. Walsh, M. C., Kim, G. K., Maurizio, P. L., Molnar, E. E., and Choi, Y. (2008) TRAF6 autoubiquitination-independent activation of the NFkappaB and MAPK pathways in response to IL-1 and RANKL, *PLoS One 3*, e4064.
129. Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Datta, A., Greenberg, R. A., Wolberger, C., and Matunis, M. J. (2012) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage, *Science signaling 5*, ra88.
130. Moyal, L., Lerenthal, Y., Gana-Weisz, M., Mass, G., So, S., Wang, S. Y., Eppink, B., Chung, Y. M., Shalev, G., Shema, E., Shkedy, D., Smorodinsky, N. I., van Vliet, N., Kuster, B., Mann, M., Ciechanover, A., Dahm-Daphi, J., Kanaar, R., Hu, M. C., Chen, D. J., Oren, M., and Shiloh, Y. (2011) Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks, *Mol Cell 41*, 529-542.
131. Cramer, P., and Wolberger, C. (2011) Proteins: histones and chromatin, *Current opinion in structural biology 21*, 695-697.
132. Zaaroor-Regev, D., de Bie, P., Scheffner, M., Noy, T., Shemer, R., Heled, M., Stein, I., Pikarsky, E., and Ciechanover, A. (2010) Regulation of the polycomb protein Ring1B by self-ubiquitination or by E6-AP may have implications to the pathogenesis of Angelman syndrome, *Proc Natl Acad Sci U S A 107*, 6788-6793.

133. Acconcia, F., Sigismund, S., and Polo, S. (2009) Ubiquitin in trafficking: the network at work, *Experimental cell research 315*, 1610-1618.
134. Keren-Kaplan, T., Attali, I., Motamedchaboki, K., Davis, B. A., Tanner, N., Reshef, Y., Laudon, E., Kolot, M., Levin-Kravets, O., Kleifeld, O., Glickman, M., Horazdovsky, B. F., Wolf, D. A., and Prag, G. (2012) Synthetic biology approach to reconstituting the ubiquitylation cascade in bacteria, *EMBO J 31*, 378-390.
135. Sims, J. J., and Cohen, R. E. (2009) Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding preference of rap80, *Mol Cell 33*, 775-783.
136. Rahighi, S., and Dikic, I. (2012) Selectivity of the ubiquitin-binding modules, *FEBS Lett 586*, 2705-2710.
137. Husnjak, K., and Dikic, I. (2012) Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions, *Annu Rev Biochem 81*, 291-322.
138. Grabbe, C., and Dikic, I. (2009) Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins, *Chemical reviews 109*, 1481-1494.
139. Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains - from structures to functions, *Nat Rev Mol Cell Biol 10*, 659-671.
140. Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H., and Hay, R. T. (2012) Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis, *Nature 489*, 115-120.
141. Haririnia, A., Verma, R., Purohit, N., Twarog, M. Z., Deshaies, R. J., Bolon, D., and Fushman, D. (2008) Mutations in the hydrophobic core of ubiquitin differentially affect its recognition by receptor proteins, *J Mol Biol 375*, 979-996.
142. Wilkinson, K. D., Laleli-Sahin, E., Urbauer, J., Larsen, C. N., Shih, G. H., Haas, A. L., Walsh, S. T., and Wand, A. J. (1999) The binding site for UCH-L3 on ubiquitin: mutagenesis and NMR studies on the complex between ubiquitin and UCH-L3, *J Mol Biol 291*, 1067-1077.
143. Papouli, E., Chen, S., Davies, A. A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H. D. (2005) Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p, *Mol Cell 19*, 123-133.
144. Kravtsova-Ivantsiv, Y., and Ciechanover, A. (2012) Non-canonical ubiquitin-based signals for proteasomal degradation, *J Cell Sci 125*, 539-548.
145. Hu, X., Paul, A., and Wang, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates, *J Biol Chem 287*, 25510-25519.
146. Aillet, F., Lopitz-Otsoa, F., Egana, I., Hjerpe, R., Fraser, P., Hay, R. T., Rodriguez, M. S., and Lang, V. (2012) Heterologous SUMO-2/3-Ubiquitin Chains Optimize IkappaBalpha Degradation and NF-kappaB Activity, *PLoS One 7*, e51672.
147. Singh, R. K., Zerath, S., Kleifeld, O., Scheffner, M., Glickman, M. H., and Fushman, D. (2012) Recognition and cleavage of related to ubiquitin 1 (Rub1) and Rub1-ubiquitin chains by components of the ubiquitin-proteasome system, *Molecular & cellular proteomics : MCP 11*, 1595-1611.

148. Whitby, F. G., Xia, G., Pickart, C. M., and Hill, C. P. (1998) Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes, *J Biol Chem 273*, 34983-34991.
149. Tokgoz, Z., Siepmann, T. J., Streich, F., Jr., Kumar, B., Klein, J. M., and Haas, A. L. (2012) E1-E2 interactions in ubiquitin and Nedd8 ligation pathways, *J Biol Chem 287*, 311-321.
150. 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.
151. Ryabov, Y., and Fushman, D. (2006) Interdomain mobility in di-ubiquitin revealed by NMR, *Proteins 63*, 787-796.
152. Ranjani, V., Assfalg, M., and Fushman, D. (2005) Using NMR spectroscopy to monitor ubiquitin chain conformation and interactions with ubiquitin-binding domains, *Methods Enzymol 399*, 177-192.
153. Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C., and Fushman, D. (2004) Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling, *J Biol Chem 279*, 7055-7063.
154. Pickart, C. M., and Fushman, D. (2004) Polyubiquitin chains: polymeric protein signals, *Current opinion in chemical biology 8*, 610-616.
155. Varadan, R., Walker, O., Pickart, C., and Fushman, D. (2002) Structural properties of polyubiquitin chains in solution, *J Mol Biol 324*, 637-647.
156. Ye, Y., Blaser, G., Horrocks, M. H., Ruedas-Rama, M. J., Ibrahim, S., Zhukov, A. A., Orte, A., Klenerman, D., Jackson, S. E., and Komander, D. (2012) Ubiquitin chain conformation regulates recognition and activity of interacting proteins, *Nature 492*, 266-270.
157. Datta, A. B., Hura, G. L., and Wolberger, C. (2009) The structure and conformation of Lys63-linked tetraubiquitin, *J Mol Biol 392*, 1117-1124.
158. Ryabov, Y. E., and Fushman, D. (2007) A model of interdomain mobility in a multidomain protein, *J Am Chem Soc 129*, 3315-3327.
159. 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.
160. 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.
161. Bremm, A., Freund, S. M., and Komander, D. (2010) Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne, *Nat Struct Mol Biol 17*, 939-947.
162. Wu-Baer, F., Lagrazon, K., Yuan, W., and Baer, R. (2003) The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin, In *J Biol Chem*, pp 34743-34746, United States.

163. Lin, D. Y., Diao, J., Zhou, D., and Chen, J. (2011) Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7, *J Biol Chem 286*, 441-449.
164. Ben-Saadon, R., Zaaroor, D., Ziv, T., and Ciechanover, A. (2006) The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity, *Mol Cell 24*, 701-711.
165. Ozkaynak, E., Finley, D., and Varshavsky, A. (1984) The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein, *Nature 312*, 663-666.
166. Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., Sano, S., Tokunaga, F., Tanaka, K., and Iwai, K. (2006) A ubiquitin ligase complex assembles linear polyubiquitin chains, *EMBO J 25*, 4877-4887.
167. Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., Kato, R., Kensche, T., Uejima, T., Bloor, S., Komander, D., Randow, F., Wakatsuki, S., and Dikic, I. (2009) Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation, *Cell 136*, 1098-1109.
168. Yao, T., and Cohen, R. E. (2000) Cyclization of polyubiquitin by the E2-25K ubiquitin conjugating enzyme, *J Biol Chem 275*, 36862-36868.
169. Hemantha, H. P., and Brik, A. (2013) Non-enzymatic synthesis of ubiquitin chains: Where chemistry makes a difference, *Bioorganic & medicinal chemistry*.
170. Sokratous, K., Strachan, J., Roach, L. V., Layfield, R., and Oldham, N. J. (2012) Cyclisation of Lys48-linked diubiquitin in vitro and in vivo, *FEBS Lett 586*, 4144-4147.
171. Satoh, T., Sakata, E., Yamamoto, S., Yamaguchi, Y., Sumiyoshi, A., Wakatsuki, S., and Kato, K. (2010) Crystal structure of cyclic Lys48-linked tetraubiquitin, *Biochem Biophys Res Commun 400*, 329-333.
172. Dickinson, B. C., Varadan, R., and Fushman, D. (2007) Effects of cyclization on conformational dynamics and binding properties of Lys48-linked di-ubiquitin, *Protein Sci 16*, 369-378.
173. Fushman, D., and Walker, O. (2010) Exploring the linkage dependence of polyubiquitin conformations using molecular modeling, *J Mol Biol 395*, 803-814.
174. Xu, P., and Peng, J. (2008) Characterization of polyubiquitin chain structure by middle-down mass spectrometry, *Anal Chem 80*, 3438-3444.
175. Liu, C., van Dyk, D., Xu, P., Choe, V., Pan, H., Peng, J., Andrews, B., and Rao, H. (2010) Ubiquitin chain elongation enzyme Ufd2 regulates a subset of Doa10 substrates, *J Biol Chem 285*, 10265-10272.
176. Chastagner, P., Israel, A., and Brou, C. (2006) Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains, *EMBO reports 7*, 1147-1153.
177. Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Peggie, M., and Alessi, D. R. (2008) Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains, *Biochem J 411*, 249-260.
178. Licchesi, J. D., Mieszczanek, J., Mevissen, T. E., Rutherford, T. J., Akutsu, M., Virdee, S., El Oualid, F., Chin, J. W., Ovaa, H., Bienz, M., and

Komander, D. (2012) An ankyrin-repeat ubiquitin-binding domain determines TRABID's specificity for atypical ubiquitin chains, *Nat Struct Mol Biol 19*, 62-71.
179. Gustin, J. K., Douglas, J. L., Bai, Y., and Moses, A. V. (2012) Ubiquitination of BST-2 protein by HIV-1 Vpu protein does not require lysine, serine, or threonine residues within the BST-2 cytoplasmic domain, *J Biol Chem 287*, 14837-14850.
180. Wang, X., Herr, R. A., Chua, W. J., Lybarger, L., Wiertz, E. J., and Hansen, T. H. (2007) Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3, *J Cell Biol 177*, 613-624.
181. Ravid, T., and Hochstrasser, M. (2007) Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue, *Nat Cell Biol 9*, 422-427.
182. You, J., Cohen, R. E., and Pickart, C. M. (1999) Construct for high-level expression and low misincorporation of lysine for arginine during expression of pET-encoded eukaryotic proteins in Escherichia coli, *Biotechniques 27*, 950-954.
183. Pickart, C. M., Haldeman, M. T., Kasperek, E. M., and Chen, Z. (1992) Iodination of tyrosine 59 of ubiquitin selectively blocks ubiquitin's acceptor activity in diubiquitin synthesis catalyzed by E2(25K), *J Biol Chem 267*, 14418-14423.
184. Cook, W. J., Jeffrey, L. C., Carson, M., Chen, Z., and Pickart, C. M. (1992) Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2), *J Biol Chem 267*, 16467-16471.
185. Park, K. C., Woo, S. K., Yoo, Y. J., Wyndham, A. M., Baker, R. T., and Chung, C. H. (1997) Purification and characterization of UBP6, a new ubiquitin-specific protease in Saccharomyces cerevisiae, *Arch Biochem Biophys 347*, 78-84.
186. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis, *Nature methods 9*, 671-675.
187. 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.
188. Walker, O., Varadan, R., and Fushman, D. (2004) Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from (15)N relaxation data using computer program ROTDIF, *J Magn Reson 168*, 336-345.
189. Zwahlen, C., Vincent, S. J., and Kay, L. E. (1998) Analytical description of the effect of adiabatic pulses on IS, I2S, and I3S spin systems, *J Magn Reson 130*, 169-175.
190. Fleishman, S. J., Whitehead, T. A., Strauch, E. M., Corn, J. E., Qin, S., Zhou, H. X., Mitchell, J. C., Demerdash, O. N., Takeda-Shitaka, M., Terashi, G., Moal, I. H., Li, X., Bates, P. A., Zacharias, M., Park, H., Ko, J. S., Lee, H., Seok, C., Bourquard, T., Bernauer, J., Poupon, A., Aze, J., Soner, S., Ovali, S. K., Ozbek, P., Tal, N. B., Haliloglu, T., Hwang, H., Vreven, T., Pierce, B. G., Weng, Z., Perez-Cano, L., Pons, C., Fernandez-Recio, J., Jiang, F., Yang, F.,

Gong, X., Cao, L., Xu, X., Liu, B., Wang, P., Li, C., Wang, C., Robert, C. H., Guharoy, M., Liu, S., Huang, Y., Li, L., Guo, D., Chen, Y., Xiao, Y., London, N., Itzhaki, Z., Schueler-Furman, O., Inbar, Y., Potapov, V., Cohen, M., Schreiber, G., Tsuchiya, Y., Kanamori, E., Standley, D. M., Nakamura, H., Kinoshita, K., Driggers, C. M., Hall, R. G., Morgan, J. L., Hsu, V. L., Zhan, J., Yang, Y., Zhou, Y., Kastritis, P. L., Bonvin, A. M., Zhang, W., Camacho, C. J., Kilambi, K. P., Sircar, A., Gray, J. J., Ohue, M., Uchikoga, N., Matsuzaki, Y., Ishida, T., Akiyama, Y., Khashan, R., Bush, S., Fouches, D., Tropsha, A., Esquivel-Rodriguez, J., Kihara, D., Stranges, P. B., Jacak, R., Kuhlman, B., Huang, S. Y., Zou, X., Wodak, S. J., Janin, J., and Baker, D. (2011) Community-wide assessment of protein-interface modeling suggests improvements to design methodology, *J Mol Biol 414*, 289-302.

191. 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.
192. van Dijk, M., and Bonvin, A. M. (2010) Pushing the limits of what is achievable in protein-DNA docking: benchmarking HADDOCK's performance, *Nucleic acids research 38*, 5634-5647.
193. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system, *Nature protocols 2*, 2728-2733.
194. Schieborr, U., Vogtherr, M., Elshorst, B., Betz, M., Grimme, S., Pescatore, B., Langer, T., Saxena, K., and Schwalbe, H. (2005) How much NMR data is required to determine a protein-ligand complex structure?, *Chembiochem 6*, 1891-1898.
195. de Vries, S. J., van Dijk, M., and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking, *Nature protocols 5*, 883-897.
196. Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta crystallographica. Section D, Biological crystallography 60*, 1355-1363.
197. Rodrigo-Brenni, M. C., and Morgan, D. O. (2007) Sequential E2s drive polyubiquitin chain assembly on APC targets, *Cell 130*, 127-139.
198. Kirkpatrick, D. S., Hathaway, N. A., Hanna, J., Elsasser, S., Rush, J., Finley, D., King, R. W., and Gygi, S. P. (2006) Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology, *Nat Cell Biol 8*, 700-710.
199. Crosas, B., Hanna, J., Kirkpatrick, D. S., Zhang, D. P., Tone, Y., Hathaway, N. A., Buecker, C., Leggett, D. S., Schmidt, M., King, R. W., Gygi, S. P., and Finley, D. (2006) Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities, *Cell 127*, 1401-1413.
200. Goto, E., Yamanaka, Y., Ishikawa, A., Aoki-Kawasumi, M., Mito-Yoshida, M., Ohmura-Hoshino, M., Matsuki, Y., Kajikawa, M., Hirano, H., and Ishido, S. (2010) Contribution of lysine 11-linked ubiquitination to MIR2-mediated major histocompatibility complex class I internalization, *J Biol Chem 285*, 35311-35319.

201. Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komuves, L., French, D. M., Ferrando, R. E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S. G., Kelley, R. F., and Dixit, V. M. (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies, *Cell 134*, 668-678.
202. Dammer, E. B., Na, C. H., Xu, P., Seyfried, N. T., Duong, D. M., Cheng, D., Gearing, M., Rees, H., Lah, J. J., Levey, A. I., Rush, J., and Peng, J. (2011) Polyubiquitin linkage profiles in three models of proteolytic stress suggest the etiology of Alzheimer disease, *J Biol Chem 286*, 10457-10465.
203. Ziv, I., Matiuhin, Y., Kirkpatrick, D. S., Erpapazoglou, Z., Leon, S., Pantazopoulou, M., Kim, W., Gygi, S. P., Haguenauer-Tsapis, R., Reis, N., Glickman, M. H., and Kleifeld, O. (2011) A perturbed ubiquitin landscape distinguishes between ubiquitin in trafficking and in proteolysis., *Mol Cell Proteomics 10*, M111.009753.
204. Matsumoto, M. L., Dong, K. C., Yu, C., Phu, L., Gao, X., Hannoush, R. N., Hymowitz, S. G., Kirkpatrick, D. S., Dixit, V. M., and Kelley, R. F. (2012) Engineering and structural characterization of a linear polyubiquitin-specific antibody, *J Mol Biol 418*, 134-144.
205. Seyfried, N. T., Gozal, Y. M., Dammer, E. B., Xia, Q., Duong, D. M., Cheng, D., Lah, J. J., Levey, A. I., and Peng, J. (2010) Multiplex SILAC analysis of a cellular TDP-43 proteinopathy model reveals protein inclusions associated with SUMOylation and diverse polyubiquitin chains, *Molecular & cellular proteomics : MCP 9*, 705-718.
206. Winborn, B. J., Travis, S. M., Todi, S. V., Scaglione, K. M., Xu, P., Williams, A. J., Cohen, R. E., Peng, J., and Paulson, H. L. (2008) The deubiquitinating enzyme ataxin-3, a polyglutamine disease protein, edits Lys63 linkages in mixed linkage ubiquitin chains, *J Biol Chem 283*, 26436-26443.
207. Pickart, C. M., and Raasi, S. (2005) Controlled synthesis of polyubiquitin chains, *Methods Enzymol 399*, 21-36.
208. 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.
209. Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80, *EMBO J 28*, 2461-2468.
210. Edelmann, M. J., Iphofer, A., Akutsu, M., Altun, M., di Gleria, K., Kramer, H. B., Fiebiger, E., Dhe-Paganon, S., and Kessler, B. M. (2009) Structural basis and specificity of human otubain 1-mediated deubiquitination, *Biochem J 418*, 379-390.
211. Valkevich, E. M., Guenette, R. G., Sanchez, N. A., Chen, Y. C., Ge, Y., and Strieter, E. R. (2012) Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases, *J Am Chem Soc 134*, 6916-6919.

212. Guterman, A., and Glickman, M. H. (2004) Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome, *J Biol Chem 279*, 1729-1738.
213. Lee, M. J., Lee, B. H., Hanna, J., King, R. W., and Finley, D. (2011) Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes, *Molecular & cellular proteomics : MCP 10*, R110 003871.
214. Cooper, E. M., Cutcliffe, C., Kristiansen, T. Z., Pandey, A., Pickart, C. M., and Cohen, R. E. (2009) K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1, *EMBO J 28*, 621-631.
215. Verma, R., Peters, N. R., D'Onofrio, M., Tochtrop, G. P., Sakamoto, K. M., Varadan, R., Zhang, M., Coffino, P., Fushman, D., Deshaies, R. J., and King, R. W. (2004) Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain, *Science 306*, 117-120.
216. Bellare, P., Small, E. C., Huang, X., Wohlschlegel, J. A., Staley, J. P., and Sontheimer, E. J. (2008) A role for ubiquitin in the spliceosome assembly pathway, *Nat Struct Mol Biol 15*, 444-451.
217. Brancolini, G., Kokh, D. B., Calzolai, L., Wade, R. C., and Corni, S. (2012) Docking of ubiquitin to gold nanoparticles, *ACS nano 6*, 9863-9878.
218. Calzolai, L., Franchini, F., Gilliland, D., and Rossi, F. (2010) Protein--nanoparticle interaction: identification of the ubiquitin--gold nanoparticle interaction site, *Nano letters 10*, 3101-3105.
219. Shaw, B. F., Schneider, G. F., Arthanari, H., Narovlyansky, M., Moustakas, D., Durazo, A., Wagner, G., and Whitesides, G. M. (2011) Complexes of native ubiquitin and dodecyl sulfate illustrate the nature of hydrophobic and electrostatic interactions in the binding of proteins and surfactants, *J Am Chem Soc 133*, 17681-17695.
220. Schneider, G. F., Shaw, B. F., Lee, A., Carillho, E., and Whitesides, G. M. (2008) Pathway for unfolding of ubiquitin in sodium dodecyl sulfate, studied by capillary electrophoresis, *J Am Chem Soc 130*, 17384-17393.
221. Spolnik, P., Stopa, B., Piekarska, B., Jagusiak, A., Konieczny, L., Rybarska, J., Krol, M., Roterman, I., Urbanowicz, B., and Zieba-Palus, J. (2007) The use of rigid, fibrillar Congo red nanostructures for scaffolding protein assemblies and inducing the formation of amyloid-like arrangement of molecules, *Chemical biology & drug design 70*, 491-501.
222. Liang, G., Xu, K., Li, L., Wang, L., Kuang, Y., Yang, Z., and Xu, B. (2007) Using Congo red to report intracellular hydrogelation resulted from self-assembly of small molecules, *Chem Commun (Camb)*, 4096-4098.
223. Skowronek, M., Roterman, Konieczny, L., Stopa, B., Rybarska, J., Piekarska, B., Gorecki, A., and Krol, M. (2000) The conformational characteristics of Congo red, Evans blue and Trypan blue, *Computers & chemistry 24*, 429-450.
224. Bellows, D. S., and Tyers, M. (2004) Cell biology. Chemical genetics hits, *Science 306*, 67-68.
225. Denison, C., Kirkpatrick, D. S., and Gygi, S. P. (2005) Proteomic insights into ubiquitin and ubiquitin-like proteins, *Current opinion in chemical biology 9*, 69-75.

226. Sylvestersen, K. B., Young, C., and Nielsen, M. L. (2013) Advances in characterizing ubiquitylation sites by mass spectrometry, *Current opinion in chemical biology*.
227. Cannon, J., Nakasone, M., Fushman, D., and Fenselau, C. (2012) Proteomic Identification and Analysis of K63-Linked Ubiquitin Conjugates, *Analytical Chemistry 84*, 10121-10128.
228. Cannon, J., Nakasone, M., Fushman, D., and Fenselau, C. (2012) Proteomic identification and analysis of K63-linked ubiquitin conjugates, *Anal Chem 84*, 10121-10128.
229. Ibarra-Molero, B., Loladze, V. V., Makhatadze, G. I., and Sanchez-Ruiz, J. M. (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability, *Biochemistry 38*, 8138-8149.
230. Raasi, S., and Pickart, C. M. (2005) Ubiquitin chain synthesis, *Methods Mol Biol 301*, 47-55.
231. Fenselau, C., Laine, O., and Swatkoski, S. (2011) Microwave assisted acid cleavage for denaturation and proteolysis of intact human adenovirus, *International journal of mass spectrometry 301*, 7-11.
232. Schimmel, J., Larsen, K. M., Matic, I., van Hagen, M., Cox, J., Mann, M., Andersen, J. S., and Vertegaal, A. C. (2008) The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle, *Molecular & cellular proteomics : MCP 7*, 2107-2122.
233. Jung, J. E., Pierson, N. A., Marquardt, A., Scheffner, M., Przybylski, M., and Clemmer, D. E. (2011) Differentiation of compact and extended conformations of di-ubiquitin conjugates with lysine-specific isopeptide linkages by ion mobility-mass spectrometry, *Journal of the American Society for Mass Spectrometry 22*, 1463-1471.
234. Strachan, J., Roach, L., Sokratous, K., Tooth, D., Long, J., Garner, T. P., Searle, M. S., Oldham, N. J., and Layfield, R. (2012) Insights into the molecular composition of endogenous unanchored polyubiquitin chains, *J Proteome Res 11*, 1969-1980.
235. Laub, P. B., Khorasanizadeh, S., and Roder, H. (1995) Localized solution structure refinement of an F45W variant of ubiquitin using stochastic boundary molecular dynamics and NMR distance restraints, *Protein Sci 4*, 973-982.
236. Wintrode, P. L., Makhatadze, G. I., and Privalov, P. L. (1994) Thermodynamics of ubiquitin unfolding, *Proteins 18*, 246-253.
237. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993) Folding and stability of a tryptophan-containing mutant of ubiquitin, *Biochemistry 32*, 7054-7063.
238. Kao, M. W., Yang, L. L., Lin, J. C., Lim, T. S., Fann, W., and Chen, R. P. (2008) Strategy for efficient site-specific FRET-dye labeling of ubiquitin, *Bioconjugate chemistry 19*, 1124-1126.
239. Gururaja, T. L., Pray, T. R., Lowe, R., Dong, G., Huang, J., Daniel-Issakani, S., and Payan, D. G. (2005) A homogeneous FRET assay system for

multiubiquitin chain assembly and disassembly, *Methods Enzymol 399*, 663-682.
240. Cook, W. J., Suddath, F. L., Bugg, C. E., and Goldstein, G. (1979) Crystallization and preliminary x-ray investigation of ubiquitin, a non-histone chromosomal protein, *J Mol Biol 130*, 353-355.
241. Trempe, J. F., Brown, N. R., Noble, M. E., and Endicott, J. A. (2010) A new crystal form of Lys48-linked diubiquitin, *Acta Crystallogr Sect F Struct Biol Cryst Commun 66*, 994-998.
242. 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.
243. Wu, P., and Brand, L. (1994) Resonance energy transfer: methods and applications, *Analytical biochemistry 218*, 1-13.
244. Ujfalusi, Z., Barko, S., Hild, G., and Nyitrai, M. (2010) The effects of formins on the conformation of subdomain 1 in actin filaments, *Journal of photochemistry and photobiology. B, Biology 98*, 7-11.
245. Yengo, C. M., Chrin, L. R., and Berger, C. L. (2000) Interaction of myosin LYS-553 with the C-terminus and DNase I-binding loop of actin examined by fluorescence resonance energy transfer, *Journal of structural biology 131*, 187-196.
246. Dennissen, F. J., Kholod, N., Steinbusch, H. W., and Van Leeuwen, F. W. (2010) Misframed proteins and neurodegeneration: a novel view on Alzheimer's and Parkinson's diseases, *Neuro-degenerative diseases 7*, 76-79.
247. de Pril, R., Hobo, B., van Tijn, P., Roos, R. A., van Leeuwen, F. W., and Fischer, D. F. (2010) Modest proteasomal inhibition by aberrant ubiquitin exacerbates aggregate formation in a Huntington disease mouse model, *Mol Cell Neurosci 43*, 281-286.
248. Song, S., and Jung, Y. K. (2004) Alzheimer's disease meets the ubiquitin-proteasome system, *Trends Mol Med 10*, 565-570.
249. de Pril, R., Fischer, D. F., Maat-Schieman, M. L., Hobo, B., de Vos, R. A., Brunt, E. R., Hol, E. M., Roos, R. A., and van Leeuwen, F. W. (2004) Accumulation of aberrant ubiquitin induces aggregate formation and cell death in polyglutamine diseases, *Hum Mol Genet 13*, 1803-1813.
250. Lindsten, K., Menendez-Benito, V., Masucci, M. G., and Dantuma, N. P. (2003) A transgenic mouse model of the ubiquitin/proteasome system, *Nature biotechnology 21*, 897-902.
251. Chadwick, L., Gentle, L., Strachan, J., and Layfield, R. (2011) Unchained maladie - a reassessment of the role of Ubb(+1) -capped polyubiquitin chains in Alzheimer's disease, *Neuropathol Appl Neurobiol*.
252. Olive, M., van Leeuwen, F. W., Janue, A., Moreno, D., Torrejon-Escribano, B., and Ferrer, I. (2008) Expression of mutant ubiquitin (UBB+1) and p62 in myotilinopathies and desminopathies, *Neuropathol Appl Neurobiol 34*, 76-87.
253. Lindsten, K., de Vrij, F. M., Verhoef, L. G., Fischer, D. F., van Leeuwen, F. W., Hol, E. M., Masucci, M. G., and Dantuma, N. P. (2002) Mutant ubiquitin found in neurodegenerative disorders is a ubiquitin fusion degradation substrate that blocks proteasomal degradation, *J Cell Biol 157*, 417-427.

254. Chadwick, L., Gentle, L., Strachan, J., and Layfield, R. (2012) Review: unchained maladie - a reassessment of the role of Ubb(+1) -capped polyubiquitin chains in Alzheimer's disease, *Neuropathol Appl Neurobiol 38*, 118-131.
255. Ko, S., Kang, G. B., Song, S. M., Lee, J. G., Shin, D. Y., Yun, J. H., Sheng, Y., Cheong, C., Jeon, Y. H., Jung, Y. K., Arrowsmith, C. H., Avvakumov, G. V., Dhe-Paganon, S., Yoo, Y. J., Eom, S. H., and Lee, W. (2010) Structural basis of E2-25K/UBB+1 interaction leading to proteasome inhibition and neurotoxicity, *J Biol Chem 285*, 36070-36080.
256. Besche, H. C., Haas, W., Gygi, S. P., and Goldberg, A. L. (2009) Isolation of mammalian 26S proteasomes and p97/VCP complexes using the ubiquitin-like domain from HHR23B reveals novel proteasome-associated proteins, *Biochemistry 48*, 2538-2549.
257. Tijn, P. V., Dennissen, F. J., Gentier, R. J., Hobo, B., Hermes, D., Steinbusch, H. W., Van Leeuwen, F. W., and Fischer, D. F. (2012) Mutant ubiquitin decreases amyloid beta plaque formation in a transgenic mouse model of Alzheimer's disease, *Neurochem Int*.
258. Dennissen, F. J., Kholod, N., Hermes, D. J., Kemmerling, N., Steinbusch, H. W., Dantuma, N. P., and van Leeuwen, F. W. (2011) Mutant ubiquitin (UBB+1) associated with neurodegenerative disorders is hydrolyzed by ubiquitin C-terminal hydrolase L3 (UCH-L3), *FEBS Lett 585*, 2568-2574.
259. Fratta, P., Engel, W. K., Van Leeuwen, F. W., Hol, E. M., Vattemi, G., and Askanas, V. (2004) Mutant ubiquitin UBB+1 is accumulated in sporadic inclusion-body myositis muscle fibers, *Neurology 63*, 1114-1117.
260. Rusk, N. (2012) Writing the histone code, *Nature methods 9*, 777.
261. Sinha, D., and Shogren-Knaak, M. A. (2010) Role of direct interactions between the histone H4 Tail and the H2A core in long range nucleosome contacts, *J Biol Chem 285*, 16572-16581.
262. Swerdlow, P. S., Schuster, T., and Finley, D. (1990) A conserved sequence in histone H2A which is a ubiquitination site in higher eucaryotes is not required for growth in Saccharomyces cerevisiae, *Mol Cell Biol 10*, 4905-4911.
263. Pickart, C. M., and Vella, A. T. (1988) Ubiquitin carrier protein-catalyzed ubiquitin transfer to histones. Mechanism and specificity, *J Biol Chem 263*, 15076-15082.
264. Jensen, K., Santisteban, M. S., Urekar, C., and Smith, M. M. (2011) Histone H2A.Z acid patch residues required for deposition and function, *Mol Genet Genomics 285*, 287-296.