ABSTRACT

Title of Dissertation:	LYSINE-48 LINKED UBIQUITN CHAIN BINDING CYCLIC PEPTIDES AS POTENTIAL CANCER THERAPEUTICS
	Betsegaw Ephraim Lemma, Doctor of Philosophy, 2022
Dissertation directed by:	Professor David Fushman, Department of Chemistry and Biochemistry

Ubiquitin (Ub) is a protein that serves several important functions in all eukaryotes. Its role involves post-translational modification of target proteins which results in signaling for cellular activities of various kinds. Given several malignancies and neurodegenerative diseases caused by defects that arise in ubiquitin mediated pathways, understanding all the variations in ubiquitin based signaling and the selectivity and specificity involved in the pathways is imperative in potentially targeting and treating various ailments that are caused by the disruption within said pathways. For this purpose, this study examined possible therapeutic methods that could target the ubiquitin-proteasome system at the ubiquitin chain level. Binding studies utilizing NMR in addition to various SDS PAGE gel assays and crystallography were important in understanding how select therapeutic molecules interacted with ubiquitin chains.

A secondary study involved Rub1, a protein that is very similar to ubiquitin in terms of sequence and structure. Comparative mutation-based gel assays, NMR analysis and mass spectrometry methods were explored with the goal of incorporating Rub1 within ubiquitin conjugation pathways and making polyRub1 chains. In addition, NMR, DSC, and CD based temperature studies were performed to further understand the structural differences between these closely related proteins and hopefully gain more insight into the seemingly special functional nature of ubiquitin.

LYSINE-48 LINKED UBIQUITN CHAIN BINDING CYCLIC PEPTIDES AS POTENTIAL CANCER THERAPEUTICS

by

Betsegaw Ephraim Lemma

Dissertation 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
2022

Advisory Committee:
Professor David Fushman, Chair
Professor Kwaku Dayie
Professor John Orban
Professor Lai-Xi Wang
Professor Sergei Sukharev - Dean's Representative

© Copyright by
Betsegaw Ephraim Lemma
2022

Acknowledgements

I express my sincere gratitude to my advisor Professor David Fushman for his guidance, patience, and support towards completing my Ph.D. I thank him for his valuable inputs and for being very supportive throughout the whole process.

I thank my dissertation committee members for their helpful input, support, and review of my work through the years.

I thank Dr. Rajesh Singh for being my initial mentor regarding the processes of protein expression, enzymatic chain synthesis of ubiquitin and for sharing his cloning and NMR expertise. I thank Dr. Andrew Boughton for his support and encouragement throughout our years in the Fushman lab for his various inputs during my research process. I also thank Dr. Karina Herrera-Guzman for her mentorship and her encouragement. I would like to also thank them for being wonderful lab mates and friends making this long process bearable.

I would like to thank Dr. Christina Camara for sharing her NMR and CD expertise and being a wonderful lab mate. I want to thank Dr. Daoning Zhang for his support and friendship throughout my Ph.D. as well as Dr. Apurva Chaturvedi for his inputs and expertise. I thank my current lab mates Westley Pawloski, Steven Bonn, Bryan Gregory Wentz and Benjamin Thomas Lanham as well as all of the past members of the Fushman lab.

I am most thankful to my parents Ephraim Lemma and Tadelech Matewos for being supportive and helping me in every way possible as I pursued my degree and for being patient even when I couldn’t visit as often as I should. I also want to thank my sisters

Kalkidan Ephraim and Hallelujah Ephraim for being there for me and helping out in any way they can. I appreciate the love and support even if I might not express it well.

Table of Contents

Acknowledgements ii

Table of Contents iv

List of Figures vii

List of Tables xii

Abbreviations xiii

Chapter 1: Introduction 1

1.1: Ubiquitin and its activity 1

1.2: Ubiquitylation cycle 2

1.3: Ubiquitin chain topology and functional outcomes 3

Chapter 2: Cyclic peptides specifically modify lysine-48 linked polyubiquitin chains 7

2.1: Objectives 7

2.2: Enzymatic assembly of lysine-48 linked polyubiquitin chains 9

2.3: Ub4ix cyclic peptide binding to lysine-48 linked Ub chains 11

2.3.1: Mapping of Ub4ix interaction with lysine-48 linked Ub trimer 18

2.3.2: Ub4ix binding confers protection to Ub chains against DUB cleavage 23

2.3.3: Attempt to co-crystallize Ub4 with Ub4ix peptide 26

2.4: Ub4a cyclic peptide binding to lysine-48 linked Ub chains 27

2.4.1: Mapping of Ub4a interaction with lysine-48 linked poly-Ub using NMR 29

2.4.2: C-terminal tail determines which trimer moiety is preferred during binding of Ub4a to Ub_4 37

2.4.3: Binding of Ub4a to Ub_3-Ub_{R72} and Ub_2-Ub_{R72} chains 37

2.4.4: Binding of Ub4a to Ub_3-Ub_{R74} and Ub_2-Ub_{R74} chains 41

2.5: Structural analysis of Ub4a bound Ub_3 and Ub_2-Ub_{R72} using spin labeling NMR techniques 44

2.5.1: PRE based structural analysis of Ub4a bound Ub_3 45

2.5.2: PRE based structural analysis of UB4a bound Ub_3-Ub_{R72} 46

2.6: Crystal Structure of Ub_3:Ub4a identifies mechanism of interaction 47

2.6.1: General description of the Ub_3:Ub4a structure 49

2.6.2: Interactions between different subunits within Ub_3 50

2.6.3: Stabilizing interactions engaging Ub4a peptide 52

2.7: DUB assays of peptide bound Ub chains 54
2.8: Ub4a presence in Ub monomer pool does not inhibit enzymatic assembly 60
2.9: Structural model of the interaction of Ub4a with Ub_3 63
2.10: Discussion and Summary 66
Chapter 3: Ubiquitin-like proteins 72
3.1: Introduction 72
3.2: Related to ubiquitin 1 (Rub1) 73
3.3: Ubiquitin conjugating enzyme E2-25K 76
3.3.1: Ubiquitin and Rub1 bind to similar areas in E2-25K 78
3.4: Rub1/Ub chimeras 80
3.4.1: In vitro lysine-48 Rubylation reactions with Rub1-Ub chimera 81
3.4.2: In vitro lysine-11 and lysine-63 Rubylation reactions with Rub1-Ub chimera 92
3.5: From Rub1-Ub dimers to heterotrimers 94
Chapter 4: Mass spectrometric analysis of Rub1/Ub chains 98
4.1: Mass spectrometry analysis of possible enzyme modification by Rub1 and/or Ub 98
4.2: Mass spectrometry-based Top-Down analysis of Branched Rub1 proteins 105
4.2.1: Top-Down analysis of $Rub1_{T72R}$-Ub_{D77} 106
4.2.2: Top-Down analysis of $Rub1_{E63K/T72R}$-Ub_{D77} and $Rub1_{K4F/T72R}$ homodimer 109
4.2.3: Top-Down analysis of $Rub1_{T72R}$ homodimer linked through lysine-11 110
4.2.4: Discussion and Summary 112
Chapter 5: Rub1 and Ub: structural delineation and functional outcomes 113
5.1: DSC experiment results of Rub1 and Ub 113
5.2: CD based temperature change measurements of Ub and Rub1 115
5.3: NMR based temperature change measurements of Ub and Rub1 120
5.3.1: NMR spectra based analysis of Rub1 and Ub thermostability 122
5.3.2: Temperature coefficient-based analysis of Rub1 and Ub 124
5.3.3: Temperature dependent deviation from linearity in specific residues 129

5.3.4: Amide proton temperature coefficient effects on local secondary structure stability of Rub1 and Ub 134
5.3.5: Discussion and Summary 138
Chapter 6: Exploratory side projects 141
6.1: Rub1 CEST experiments 141
6.2: Characterization of lysine-48 linked chains through small-angle and very small angle neutron scattering studies 142
6.3: Partial characterization attempt of lysine-48 linked Ub trimer through NMR 147
6.4: Ub-DHFR 150
6.5: p53-UDR 153
Chapter 7: Materials and Methods 158
7.1: Ub protein expression and purification 158
7.1.1: Deuterated ubiquitin growth 158
7.1.2: Synthesis of selectively isotope enriched or deuterated Ub chains 161
7.2: Rub1 expression and purification 163
7.2.1: Enzymatic assembly of $Rub1_{E63K/T72R}$ – Ub_{74} heterodimer 163
7.2.2: Enzymatic assembly of $Rub1_{K4F/T72R}$ –$Rub1_{K4F/T72R}$ homodimer 163
7.2.3: Enzymatic assembly of $Rub1_{T72R}$–$Rub1_{T72R}$ homodimer 164
7.3: Liquid-Chromatography–Tandem Mass Spectrometry 164
7.4: NMR studies of Ub4ix and Ub4a peptides 165
7.5: PRE-Experiments 167
7.6: Deubiquitination Assays 169
7.7: X-ray crystallography data collection and analysis of Ub_4 with Ub4ix peptide 169
7.8: Crystallization, data collection and structural refinement of Ub_3:Ub4a complex 170
7.9: Differential Scanning Calorimetry experiments 172
7.10: Variable temperature NMR 173
Chapter 8: Discussion and Future Directions 174
Appendix 179
References 187

List of Figures

Figure 1.1: Enzymatic pathway of ubiquitylation 3

Figure 1.2: Ubiquitin chains and various functions 4

Figure 1.3: Structure of Ubiquitin 6

Figure 2.1: Controlled enzymatic assembly of isotopically labeled ubiquitin chains 11

Figure 2.2: Ub4ix peptide 12

Figure 2.3: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra and normalized intensity plots of selectively labeled Ub_4 upon addition of Ub4ix 17

Figure 2.4: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra and normalized intensity plots of selectively labeled Ub_3 upon addition of Ub4ix 19

Figure 2.5: Comparison of the intensities of the unbound signals of Ub_4 and Ub_3 upon addition of Ub4ix 20

Figure 2.6: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra of selectively labeled Ub_4 and Ub_3 units upon addition of Ub4ix.....21

Figure 2.7: Comparison of chemical shifts in Ub4ix bound Ub_4 and Ub_3 22

Figure 2.8: Perturbations in the distal unit of Ub_2 upon Ub4ix binding.....23

Figure 2.9: Time course disassembly of Ub_4 and Ub_3 by distal lysine-48 linkage specific OTUB1 upon addition of Ub4ix 25

Figure 2.10: Crystals of potential Ub4ix bound Ub_4 selected for diffraction.....26

Figure 2.11: Crystal structure results of potential Ub_4 bound to Ub4ix 27

Figure 2.12: Ub4a cyclic peptide 29

Figure 2.13: ^{1}H-^{15}N SOFAST-HMQC spectra and CSP plot of selectively labeled Ub_4 and Ub_3 upon addition of Ub4a 33

Figure 2.14: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra of proximally labeled Ub_4 and Ub_3 and distally labeled Ub_4 and Ub_3 upon addition of Ub4a 34

Figure 2.15: ^{1}H-^{15}N SOFAST-HMQC spectra and CSP plot of endo labeled Ub_3 upon addition of Ub4a.....35

Figure 2.16: ^{1}H-^{15}N SOFAST-HMQC spectra and CSP plot of proximally labeled Ub_2 upon addition of Ub4a.....36

Figure 2.17: ^{1}H-^{15}N SOFAST-HMQC spectra and CSP plot of distally labeled Ub_3-Ub_{R72} upon addition of Ub4a.....38

Figure 2.18: ^{1}H-^{15}N SOFAST-HMQC spectra of distally labeled Ub_2-Ub_{R72} upon addition of Ub4a.....40
Figure 2.19: ^{1}H-^{15}N SOFAST-HMQC spectra of distally labeled Ub_3-Ub_{R74} and Ub_2-Ub_{R74} upon addition of Ub4a.....43
Figure 2.20: TEMPO attached on to modified Ub4a cyclic peptide.....44
Figure 2.21: Paramagnetic spin label effect on Ub_3:$Ub4a^{TEMPO}$ complex.....46
Figure 2.22: Paramagnetic spin label effect of Ub_3-Ub_{R72}:$Ub4a^{TEMPO}$ complex.....47
Figure 2.23: Crystal structure complex of Ub_3:Ub4a.....50
Figure 2.24: Interactions between different Ub subunits within Ub_3 crystal.....51
Figure 2.25: Stabilizing interactions between different sections of Ub4a and Ub_3.....53
Figure 2.26: Time course disassembly of Ub_4 and Ub_3 by lysine-48 linkage specific OTUB1 upon addition of Ub4a.....56
Figure 2.27: Time course disassembly of Ub_4 and Ub_3 by IsoT upon addition of Ub4a..57
Figure 2.28: Time course disassembly assay of Ub_3-Ub_{R72} by lysine-48 linkage specific OTUB1 upon addition of Ub4a.....58
Figure 2.29: Time course disassembly assay of Ub_2-Ub_{R72} by lysine-48 linkage specific OTUB1 upon addition of Ub4a.....59
Figure 2.30: Time course disassembly assay of Ub_3-Ub_{R74} and Ub_2-Ub_{R74} by lysine-48 linkage specific OTUB1 upon addition of Ub4a.....60
Figure 2.31: Time course assembly of wild type Ub monomer upon addition of Ub4a ...62
Figure 2.32: Comparison of Ub_3-Ub4a crystal complex and SLfit results with TEMPO built in on Ub4a from the solved crystal structure.....65
Figure 3.1: Sequence alignments for Ub, Nedd8 and Rub1 from Saccharomyces cerevisiae (Sc) and human (Hs) 74
Figure 3.2: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra and CSP plots of titration of ^{15}N E2-25K with ubiquitin and Rub1.....78

Figure 3.3: PDB structure of E2-25K showing perturbed residues during NMR titration experiments of ^{15}N E2-25K with ubiquitin and Rub1 80
Figure 3.4: Rub1 and ubiquitin sequence differentiated into seven "cassettes" 80
Figure 3.5: Reaction assay of Wt-Rub1 with and without addition of Ub_{K48R} 83
Figure 3.6: Lysine-48 reactions with various Rub1 mutants 84
Figure 3.7: Lysine-48 reactions with various Rub1 mutants 85
Figure 3.8: E2-25K: Ub PDB crystal structure with ubiquitin bound to the UBA domain of E2-25K 86
Figure 3.9: Silver-stained lysine-48 reactions of $Rub1_{H60N/T72R}$ and Ub_{74} 87
Figure 3.10: Lysine-48 reactions with various Rub1 mutants 89
Figure 3.11: Lysine-48 reactions with various Rub1 mutants 90
Figure 3.12: Lysine-48 reactions with various Rub1 mutants 92
Figure 3.13: Lysine-48 reactions with various Rub1 mutants 93
Figure 3.14: Comparison of lysine-48 heterologous dimerization reactions of Ub_{K48C} with $Rub1_{T72R}$-Ub_{74} and Ub_{WT}-Ub_{74} 95
Figure 3.15: Trimer reaction of $Rub1_{T72R}$-Ub_{74} dimer with $Rub1_{T72R}$ 97
Figure 4.1: Enzymatic chain reaction gel of ubiquitin and Rub1 with GST-tagged and untagged E2-25K used for MS analysis 100
Figure 4.2: Deconvoluted product ion spectrum of the branched protein $Rub1_{T72R}$-Ub_{D77} 107
Figure 4.3: Graphic fragmentation maps interpreting the tandem mass spectrum of the conjugate formed between Ub_{D77} and $Rub1_{T72R}$ 108
Figure 4.4: PyMOL generated image interpreting the tandem mass spectrum of the conjugate formed between Ub_{D77} and $Rub1_{T72R}$ 109
Figure 4.5: Graphic fragmentation maps interpreting the tandem mass spectrum of heterologous and homologous dimers 110

Figure 4.6: Graphic fragmentation maps interpreting the tandem mass spectrum of the conjugate formed between $Rub1_{T72R}$ homodimer 111
Figure 5.1: DSC results of ubiquitin and Rub1 at pH 6.0..... 115
Figure 5.2: CD spectra of Rub1; Nedd8 and ubiquitin at $25^{O}C$ and pH 6.0..... 117
Figure 5.3: Temperature dependent CD spectra of ubiquitin and Rub1 at pH 6.0..... 119
Figure 5.4: Comparison of ^{1}H-^{15}N SOFAST-HMQC spectra of ubiquitin and Rub1 upon temperature change..... 122
Figure 5.5: Shift of ^{15}N-Ub signals collected over a range of temperatures 125
Figure 5.6: Temperature dependent chemical shifts of the backbone amide protons of Ub at pH 6.0 from 280K to 367.5K..... 127
Figure 5.7: Temperature dependent chemical shifts of the backbone amide protons of Rub1 at pH 6.0 from 280K to 330K..... 128
Figure 5.8: Residues with curved amide proton temperature dependence identified in ubiquitin and Rub1..... 133
Figure 5.9: $^{1}H_{NH}$ temperature coefficients plotted for ubiquitin and Rub1 residues 135
Figure 5.10: $^{1}H_{NH}$ temperature coefficients plotted for secondary structure residues of ubiquitin and Rub1..... 138
Figure 6.1: CSP plot comparison between Ub_3 and Ub_2 150
Figure 6.2: Overlay of ^{1}H-^{15}N SOFAST-HMQC spectra of ^{15}N-UbDHFR and wild type ubiquitin 152
Figure 6.3: Model of the interaction of a 53BP1 dimer with a $H_4K_{20}me_2$- and H_2AK_{15}ub-modified mononucleosome 154
Figure 6.4: Overlay of ^{1}H-^{15}N SOFAST-HMQC spectra of unlabeled GST-TUDOR-UDR and wild type ubiquitin..... 154
Figure 6.5: Overlay of ^{1}H-^{15}N SOFAST-HMQC spectra of unlabeled UDR and wild-type ubiquitin..... 155

Figure 6.6: Overlay of ^{1}H-^{15}N SOFAST-HMQC spectra of unlabeled GST-TUDOR-UDR and distally labeled K_{63}Ub dimer.....155
Figure 6.7: Alignment of the UDR region in 53BP1 orthologues 156
Figure 6.8: Overlay of ^{1}H-^{15}N SOFAST-HMQC spectra of unlabeled UDR L1619A, R1627A and distally labeled lysine-63 linked Ub dimer.....156
Figure A1: Time course assembly of wild type Ub dimer and Ub_{72} upon the addition of Ub4a 179
Figure A2: Temperature dependent chemical shifts of the backbone amide nitrogens of Ub at pH 6.0 from 280K-367.5K 180
Figure A3: Temperature dependent chemical shifts of the backbone amide nitrogens of Rub1 at pH 6.0 from 280K-330K 181
Figure A4: SANS based GNOM refined plots.....182
Figure A5: Cassette temperature coefficients for ubiquitin and Rub1 at T_m-15°C.....185

List of Tables

Table 2.1: Statistics for Ub_3:Ub4a crystal complex. 48
Table 2.2: Stabilizing interactions between the three subunits of Ub_3 52
Table 2.3: Stabilizing interactions between Ub4a and subunits of Ub_3 54
Table 4.1: MS analysis of enzyme loaded gel bands 102
Table 4.2: MS analysis of trimer and tetramer gel bands 103
Table 4.3: MS analysis of dimer gel bands 104
Table 6.1: Selectively deuterated Ub4 chains used in SANS and vSANS studies 143
Table 6.2: GNOM and Guinier based Rg results for all SANS measurements 144
Table 6.3: Cartoon model of the potential conformational equilibrium of Ub4 in solution 146
Table 6.4: GNOM and Guinier based Rg results for all SAXS measurements 147
Table A1: Amide Temperature Coefficients for residues 186

Abbreviations

26S	26S proteasome
μM	Micromolar
ATP	Adenosine triphosphate
CD	Circular Dichroism
CSP	Chemical shift perturbation
Da	Dalton
DNA	Deoxyribonucleic acid
DUB	Deubiquitinating enzyme
E1	Ubiquitin activating enzyme
E2	Ubiquitin conjugating enzyme
E3	Ubiquitin ligase
E2-25K (Ube2K)	Ubiquitin conjugating enzyme E2K
EDTA	Ethlyenediaminetetraacetic acid
ESI-MS	Electrospray ionization-mass spectrometry
HMQC	Heteronuclear multiple quantum coherence
HSQC	Heteronuclear single quantum coherence
K	Lysine
LB	Luria broth
MDM2	Murine Double Minute 2
MESNA	Sodium 2-mercaptoethanesulfonate, $C_2H_5NaO_3S_2$
mM	Millimolar
MTSL	Methyl methanesulfonothioate, a nitroxide spin-label.
MW	Molecular weight
NF- κB	Nuclear factor- κB
NMR	Nuclear magnetic resonance
OTUB1	Otubain1, lysine-48 linkage specific deubiquitinase
Kd	Dissociation constant of binding

KDFIT	A program for determining the Kd from CSPs of NMR binding assays
P53	Protein 53
PBS	Phosphate-buffered saline
PDB	Protein data bank
ppb	Parts per billion
ppm	Parts per million
PRE	Paramagnetic relaxation enhancement
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
SOFAST HMQC	Selective optimized flip angle short transient heteronuclear multiple quantum coherence
SPHP	A high performance cation column
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TC	Temperature coefficient
TCEP	Tris(2-carboxyethyl)phosphine, reducing agent
TFA	Trifluoroacetic acid
Ub	Ubiquitin
Ub_2	Ubiquitin dimer
Ub_3	Ubiquitin trimer
Ub_4	Ubiquitin tetramer
UBA	Ubiquitin-associated domain
UBD	Ubiquitin binding domain
UBL	Ubiquitin-like proteins
YUH1	Ubiquitin carboxyl-terminal hydrolase 1

Chapter 1: Introduction

To understand various complex biological processes, it is necessary to understand the mechanisms involved. In the eukaryotic cell, one can imagine the genome acting as the architects for all cellular organizations. Within this framework, the proteome could then be considered as the engineers that build from the blueprint available. As such, possible gaps that might arise between the architects and the engineers is closed through what is known as post-translational modification of proteins.

Post-translational modifications (PTMs) alter the stability of polypeptides and expand their functionality through covalent attachment of various chemical moieties thereby effectively expanding the cellular proteome[1]. Some of the most common types of PTM's include phosphorylation, acetylation, glycosylation, as well as the main focus of discussion hereafter, ubiquitylation.

1.1 Ubiquitin and its activity

Ubiquitin (Ub), a 76-residue conserved regulatory protein in eukaryotic cells, is among the most well-studied post-translational protein modifiers within the cell. It functions as a signal for various regulatory processes such as proteasomal and lysosomal degradation, DNA damage responses, endocytosis and receptor signaling. To accomplish these tasks, tagging of substrate proteins with Ub has to take place through multiple steps of enzymatic conjugation steps. The covalent attachment of Ub to ε-amino group of lysine residues in substrate proteins occurs through an isopeptide bond[2-4] that results in either a single Ub attachment (monoubiquitylation or multiple monoubiquitylation) or through

ubiquitin chains (polyubiquitylation). The type of ubiquitin attachment determines the final fate of the substrate in the cell.

1.2 Ubiquitylation cycle

In the conjugation of substrate proteins with ubiquitin, three classes of enzymes are involved. Namely, the Ub activating enzyme (E1), Ub conjugating enzyme (E2) and Ub ligase (E3)[5]. A common pathway in the ubiquitylation system initially starts with ATP activation of ubiquitin and formation of a thioester bond between the C-terminal glycine residue (G76) of Ub and the active cysteine residue on E1. The ubiquitin will then be transferred through trans-thioesterification reaction to the catalytic cysteine residue on Ubiquitin-Conjugating (UBC) domain of E2 followed by a transfer to a lysine residue of a substrate mediated by either a RING E3 ligase that acts as a scaffolding protein for a substrate and E2 or a HECT E3 ligase that forms a covalent thioester complex with Ub prior to substrate labeling[6]. The human genome encodes for 2 E1's, close to 40 E2's and close to 1000 E3's[7]. These numbers give an idea of the complicated processes involved in the post-translational modification of proteins in the cell as well as the increase in functional specificity as we move further in the enzymatic cascade.

It is noteworthy that ubiquitylation is a reversible process. De-ubiquitylating enzymes (DUBs) cleave ubiquitin from proteins, proteasome-associated peptides, and polymeric ubiquitin chains etc., thereby maintaining ubiquitin homeostasis in cells[8, 9]. Since there are DUBs that recognize and cleave ubiquitin chains with specific linkages, they are considered to be good potential targets for drug therapy.

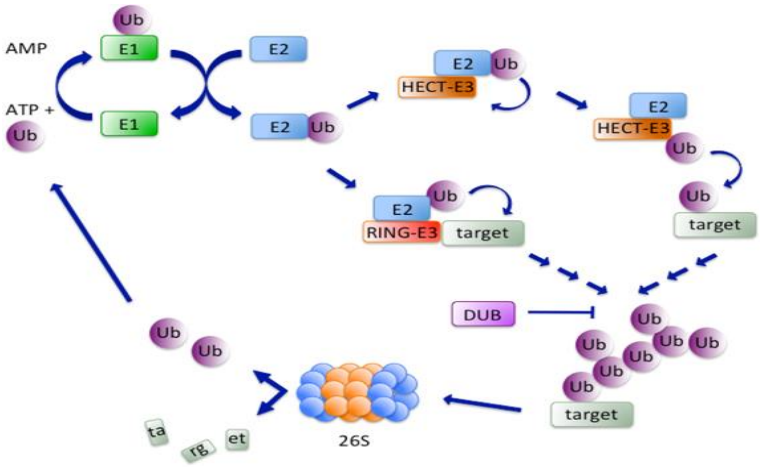


Figure 1.1: Enzymatic pathway of ubiquitylation. Ubiquitin is recycled through activity of deubiquitinases or the 26S proteasome. Image from Serena et.al., *It takes two to tango: Ubiquitin and SUMO in the DNA damage response*[10].

1.3 Ubiquitin chain topology and functional outcomes

As already mentioned, the targeting of ubiquitylated proteins for the various cellular processes depends on the type of Ub chains formed through linkages involved in the covalent attachment of one of seven lysine residues (i.e., K6, K11, K27, K29, K33, K48, K63) of a donor Ub (proximal Ub) with the glycine-76 residue in the C-terminal tail of an acceptor ubiquitin (distal Ub). It should also be noted that the N-terminus (Met) amino group can be involved in forming so-called linear chains[11]. Polyubiquitin chains that are formed can be considered homogeneous when chains are formed via elongation through the same lysine residue. On the other hand, linkages could alternate at succeeding positions within a chain leading to a mixed topology. There are also cases, both *in vivo* and *in vitro*, where chains comprised of a single ubiquitin moiety with multiple lysine modifications lead to the formation of branched chains[12]. The two canonical examples of ubiquitin function are the tagging of substrates with lysine-48 linked polyubiquitin which signals for

proteasomal degradation[13, 14] and lysine-63 linked ubiquitin chains which have been shown to have both proteolytic activity through the autophagic system[15] and non-proteolytic activity which includes DNA repair and ribosomal functioning[16-19]. Figure 1.2 illustrates the various kinds of ubiquitylation processes, and the functional consequences based on the types of modifications that substrate proteins experience. Lysine-48 linked polyubiquitin chains will be mostly discussed in this dissertation work.

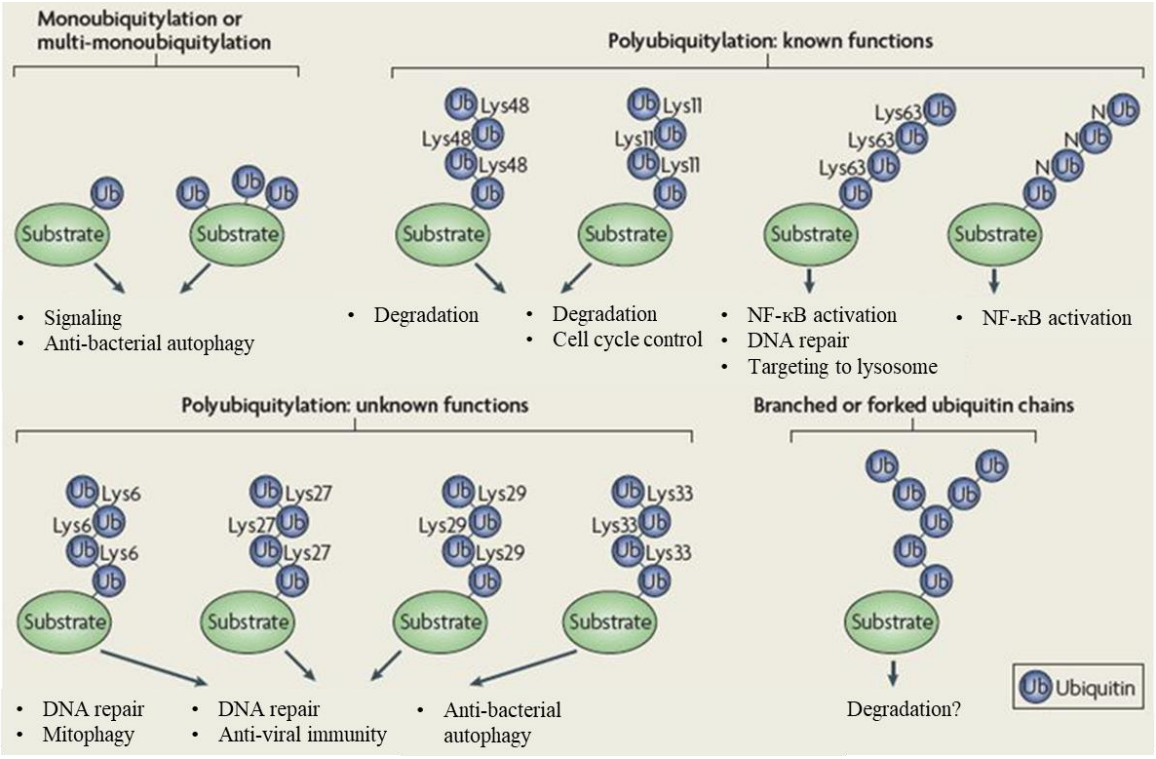


Figure 1.2: Variation in Ub chain linkages and their subsequent signaling functions. Image from Ye, Y et.al., *Building ubiquitin chains: E2 enzymes at work*[20].

In addition to the linkage-based functionality of ubiquitin, one of its more interesting features is the presence of a prominent hydrophobic patch comprising of residues Leu8, Val70, Ile44 and His68 (Fig 1.3). This hydrophobic patch along with the

flexible C-terminal tail and the electrostatic potential caused by positive charges surrounding the hydrophobic patch (Lys6, Arg42, Lys48 and Arg72) has shown to be essential for ubiquitin's interaction with multiple cellular factors which can utilize ubiquitin receptors known as Ubiquitin Binding Domains (UBDs) that interact non-covalently with ubiquitin monomers or polyubiquitin chains[4]. These cellular factors are able to interact selectively with polyubiquitin chains linked through different lysines by binding at the hydrophobic patch[20].

Among the various chains mentioned, lysine-48 linked polyubiquitin chains are one of the most studied. NMR studies of this linkage show a preference of a closed conformation due to interactions of the hydrophobic patches within the distinct Ub units in the chain[14]. This is true for both dimer (PDB: 1AAR)[21] and tetraubiquitin (PDB: 2O6V)[13] chains with the hydrophobic patch of Ub moiety-1 of the chain interacting with hydrophobic patch of moiety-2 and similar interactions being observed between moieties 3 and 4 in a tetramer leading to overall globular structure of two ubiquitin dimers. Since the main function of lysine-48 linked Ub chains is in tagging substrates for proteolytic degradation (with ≥4 Ub units as the minimum necessary length for degradative signaling)[22, 23], it is not surprising that a good portion of ubiquitin study is focused on lysine-48 linked polyubiquitin chain which is also the main focus of study in the work presented in this dissertation.

In any case, due to the variation of function based on chain type and length, the specificity and selectivity involved in formation of chains and targeting of substrates through a cascade of enzymatic reactions is worth examining. In my research, I worked with multiple E2s such as the yeast enzyme Ubc13 with its heterodimer partner Mms2

(lysine-63 linkage specific) and mammalian enzyme Ube2S (lysine-11 linkage specific) with a main focus on using the mammalian enzyme Ube2K (E2-25K) (lysine-48 linkage specific). Conveniently enough, these enzymes can have polyUb chains preassembled on the E2 enzyme itself and transferred *en-masse* to a substrate which eliminates the need for E3's in my research studies[24, 25].

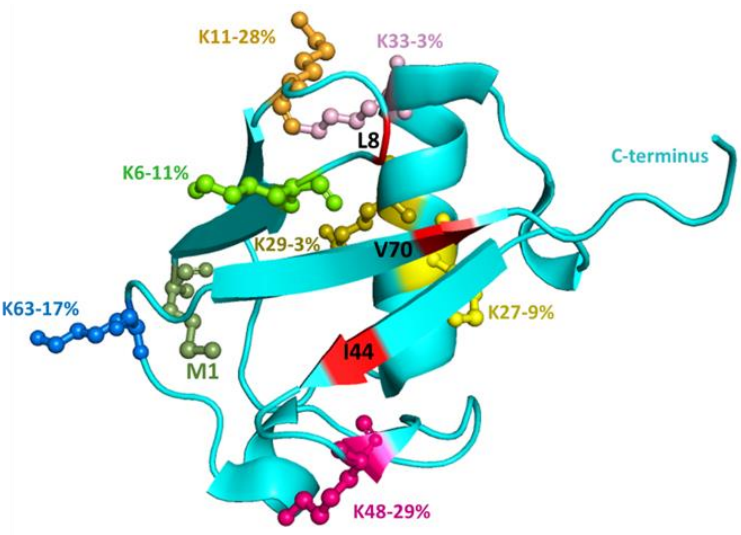


Figure 1.3: Cartoon representation of ubiquitin structure with the main residues in the canonical hydrophobic patch represented in red and all lysines and N-terminal methionine shown as sticks (PDB: 1D3Z). Figure shows abundance of the different types of polyubiquitin linkages in yeast strains as quantified by MS[26].

The following chapter discusses work done to characterize the inhibitory effects of small peptides on protein degradation. Specifically, this study focused on understanding how potential therapeutic cyclic peptides are able to target lysine-48 linked polyubiquitin chains with apparent selectivity in chain length and type. The questions raised were mainly explored through NMR, DUB reaction assays and crystallography techniques.

Chapter 2: Cyclic peptides specifically target lysine-48 linked polyubiquitin chains

2.1 Objectives

Protein homeostasis is an essential aspect of the proper functioning of biological processes. The principal mechanism by which this homeostasis is maintained is through the functioning of the ubiquitin-proteasome pathway which is involved in the degradation of proteins including those involved in cell cycle regulation (p53, p27, AMPK, mTOR)[27]. As such, since the main features of cancerous cells is uncontrolled proliferation and apoptosis evasion, the ubiquitin-proteasome system (UPS) has long been viewed as a target for anticancer drug developments. This targeting by small molecule drugs could function through the inhibition of proteasomal activity; inhibition of enzymes in Ub conjugation cascade or affecting deubiquitinases activity[28, 29]. There are several FDA approved protease inhibitor drugs such as bortezomib, carfilzomib and ixazomib used for the treatment of multiple myeloma and mantle cell lymphoma which function by binding to the catalytic site of the 26S-proteasome[30]. Nevertheless, a drawback in these types of molecules is that their therapeutic success affects various cellular networks and proteins due to their generic effects and thus are potentially toxic. Thus, there is a need for therapeutics that can have selective binding to target ubiquitin chains and directly interfere with protein-protein interactions that have been shown to be difficult to disrupt with small molecule drugs and also overcome the resistance and side effects observed from proteasome inhibitors[31, 32]. In the last decade, the use of chemical genetic screens was utilized to discover the small molecule ubistatin which was shown to bind ubiquitin (at μM levels) and thus inhibit proteasomal protein degradation[33, 34]. This was seen as a step forward in targeting Ub

chains to affect various cellular processes. More recently macrocyclic peptides have been shown to be the newest therapeutic agents by which the UPS could be targeted with tight binding and specificity[35].

Design and synthesis of peptides has shown great progress in recent times and conformational rigidity of cyclic peptides has shown to limit the entropic cost of binding observed with linear peptides. Similar to the general mechanism of action that is observed with small molecules, peptides that target the UPS pathway would be effective in attacking tumor cells that highjack this pathway leading to higher proteasomal activity and proliferation of tumor cells compared to healthy cells. This aspect of tumorigenesis would thereby limit the overall effect that peptide therapy would have on healthy cells . Furthermore, cyclizing peptides has been shown to improve permeability and reduce degradation from metabolic proteases[36, 37]. Therefore, in order to discover *de novo* macrocyclic peptides that specifically target Ub, random non-standard peptides integrated discovery (RaPID) system was utilized to generate a large number of macrocyclic peptide ligand libraries to be screened with a target of lysine-48 linked Ub chains given the significance of this chain type in degradative cellular processes[31, 38, 39]. This RaPID system used reprogrammed ribosome peptide synthesis to synthesize, and then screen extremely large ($>10^{13}$) cyclic peptide libraries for binding molecules[40]. Subsequently, several *de novo* cyclic peptides that could specifically bind lysine-48 linked chains with selectivity towards chain length were identified by our collaborators Dr. Ashraf Brik (Technion - Israel Institute of Technology) and Dr. Hiroaki Suga (University of Tokyo). The working hypothesis in this effort is that if these cyclic peptides bind to lysine-48 linked ubiquitin tetramers and prevent and/or reduce the overall-degradation of tumor suppressant proteins

thereby leading to cellular apoptosis of cancer cells, it would be possible to prevent proliferation of cancerous tumors.

The binding affinity and specificity of these cyclic peptides was tested through surface plasmon resonance (SPR) analysis and the peptides that showed the greatest binding affinity and specificity for ubiquitin chains were identified. Among the various cyclic peptides identified and tested through SPR, two cyclic peptides, named Ub4ix and Ub4a showed higher specificity for lysine-48 linked ubiquitin tetramer chains over monomer and dimer. Moreover, the binding of these peptides showed selectivity for lysine-48 linked chains of any length as compared to either lysine-11 or lysine-63 linked chains[31].

The purpose of the work discussed hereafter is to better understand the binding mode of these peptides and map their interaction with lysine-48 linked Ub chains. This would potentially improve the understanding of how these peptides act upon lysine-48 polyubiquitin chains in preventing recognition by the proteasome or DUBs. This could also provide further insights on how to improve the continuous search for other small peptides that could act in similar or better ways in targeting various other Ub chain types. As such, to identify how these peptides interacted with Ub chains, Nuclear Magnetic Resonance (NMR) spectroscopy was utilized.

2.2 Enzymatic assembly of lysine-48 linked polyubiquitin chains

Since ubiquitin chains are formed from ubiquitin units that are chemically identical, trying to distinguish between spectra of different units of a chain that are labeled simultaneously would be practically impossible. As such, the first step in conducting NMR

studies was to synthesize lysine-48 linked ubiquitin chains that were ^{15}N enriched at different units within the chain. This allows for individual assessment of the various changes that might occur on specific ubiquitin units within a chain due to a potential binding event with cyclic peptides.

For this purpose, a system that utilized a combination of the normal enzymatic synthesis of ubiquitin chains along with some extra steps to control the assembly of selectively enriched chains was undertaken[14, 23, 41, 42]. As shown in Figure 2.1, each step in the assembly was performed for all the chains used in these experiments with the only variation being the position of the isotope labeled unit within each chain. In the final product of lysine-48 linked Ub tetramer (Ub_4), the position of each Ub unit is numbered with the distal end (furthest from the C-terminus) being numbered as 4 and the proximal end (closest to the C-terminus) numbered as 1. For each chain used in this study, the distal unit has its lysine-48 mutated to an arginine (Ub_{K48R}) in order to prevent any extension of chains in the distal direction. As such, the only addition of new Ub units is done through the C-terminus in the proximal unit which allows for the control of chain length. In each step of proximal unit being added, the C-terminus has an additional aspartate at position 77 (Ub_{D77}) which prevents any uncontrolled chain extension from the proximal end. Once the distal Ub_{K48R} and the proximal Ub_{D77} are enzymatically assembled, the D77 is removed from the proximal end and another Ub_{D77} unit is added until the desired length of chain is reached. Throughout this chapter, and indeed this thesis, all Ub chains described are lysine-48 linked with the distal unit containing the aforementioned K48R mutation unless otherwise stated. Furthermore, the term distal trimer moiety or proximal trimer moiety is used when discussing Ub_4 to refer either units 2 to 4 or 1 to 3, respectively. The term “endo

unit" is used to refer to units that are in the middle of a chain. In general, this "endo" designation would mainly be relevant when discussing a trimer chain (Ub_3).

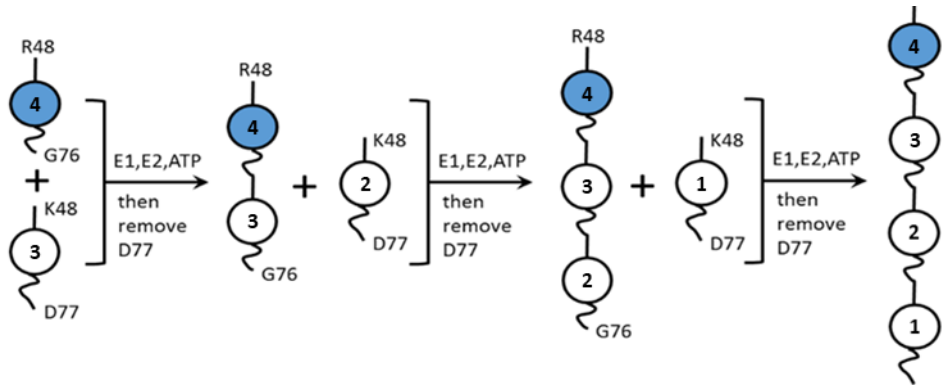


Figure 2.1: Controlled enzymatic assembly of selectively isotope-labeled lysine-48 linked di-, tri-, and tetra-Ub chains. Shown is an example of the assembly of Ub chains having the most distal Ub unit (blue) ^{15}N-isotope enriched for NMR study. Also indicated is the numbering of Ub units used in this study during the assembly of tetra-Ub chains.

2.3 Ub4ix cyclic peptide binding to lysine-48 linked poly-Ub

After acquiring both the cyclic peptides and the necessary ubiquitin chains, several NMR experiments were performed to map the binding surface on the Ub chain. Since the first peptide that was found to tightly bind lysine-48 linked ubiquitin tetramers was Ub4ix, initial rounds of NMR studies were performed on this peptide to assess the results obtained from SPR experiments performed by our aforementioned collaborators.

Ub4ix is a fifteen amino acid cyclic peptide with several hydrophobic residues. The SPR based binding experiments of Ub4ix to ubiquitin monomer and various ubiquitin chains showed a highly selective affinity for lysine-48 linked tetramers with a K_d of 6 ± 1 nM as compared to either ubiquitin dimer K_d (>1 μM) or ubiquitin monomer which

did not have binding that was measurable by SPR. It should be noted that SPR experiments did not show detectable binding to either lysine-63 or lysine-11 polyubiquitin chains[31].

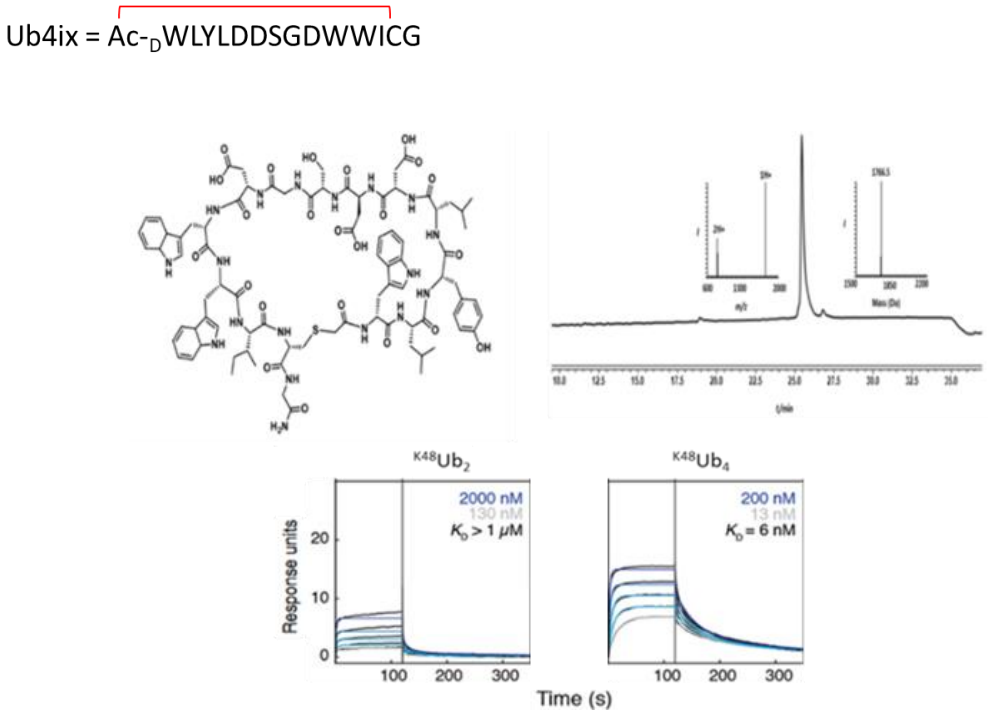


Figure 2.2: Synthesized Ub4ix peptide (top) and its binding to Ub_2 (bottom left) and Ub_4 (bottom right) detected using SPR. Highest and lowest peptide SPR concentrations are represented by blue and grey colors respectively[31].

Initial NMR experiments were conducted focusing on three units of Ub_4 namely the distal unit, 3rd unit from proximal (endo-3) and proximal units (numbering of units is as illustrated in Fig 2.1). The NMR spectra for all the titrations followed a similar pattern whereby instead of exhibiting consecutive shifts in residue signals with the progressive addition of cyclic peptide, the intensity of the unbound Ub signals gradually decreased while, at the same time, new signals appeared and increased in intensity corresponding to bound signals of ubiquitin. The signals that showed the most perturbations were mainly those corresponding to residues in the hydrophobic patch.

Generally, ligand binding causes local chemical exchange to occur in the electronic environments of the protein being observed. In cases where the chemical exchange rate is much slower than the resonance frequency difference, a slow exchange regime is observed, wherein binding leads to disappearance of NMR signals corresponding to the unbound state accompanied by the emergence of new NMR signals corresponding to the bound state[43]. This occurs because of a lack of major interconversions between bound and unbound states during the frequency detection period of the NMR experiment (~100 ms). These slow exchange regimes are typically indicative of slow off-rates (k_{off}) on the NMR experiment time scale which reflect tight binding (K_d < 0.5 μM) assuming diffusion-limited on-rates (k_{on})[44]. Thus, the titration results for Ub_4 and Ub_3, which exhibited this slow exchange regime (Fig 2.3 & 2.4), seem to indicate very tight binding in agreement with the K_d calculations based on SPR measured kinetics with NMR saturation reached after the addition of an equimolar amount of the unlabeled peptide[31, 39].

Looking at the spectra for Ub_4 titration, the distal and 3rd Ub units seem to show more interaction to the peptide upon binding (Fig 2.3; a & b) compared to the proximal unit (Fig 2.3; c). Even though, the signal attenuations in all three cases showed the participation of the hydrophobic patch in the binding process, the qualitative difference in signal attenuations observed for each unit suggested that the proximal unit was much less affected by Ub4ix binding. To further explore this observation, signal intensity comparison between each unit was performed.

Unlike fast exchange regimes, where chemical shift perturbations (CSP's) are used to assign bound peaks, the experiments performed here did not allow for exact assignment of new signals. Fortunately, the majority of the signals in the titration of Ub4ix with Ub_4

only showed signal broadening without much shift which allowed for confident signal assignment. In cases where peaks disappeared because of slow exchange, 'minimum signal shift difference' was used whereby, each free peak is linked to the signal in the bound spectrum that has moved the least from the position of the unbound[45]. As such, chemical shift changes were ascribed to most of the resonances. Even though a good number of the new peaks were assigned with some degree of confidence, to limit any error that might lead to false conclusions, the interaction between protein and peptide was assessed by measuring the normalized signal intensity for each unbound resonance instead of using chemical shift perturbations that rely on proper assignment of bound resonances. As such, upon addition of peptide during the titration, residues that are more involved in a binding event would be expected to experience a higher degree of decrease in signal intensity.

Based on the intensity plot for each titration experiment, the endo 3rd unit in the tetramer exhibited the greatest number of residues with decrease in intensity while the distal unit showed a moderate amount of relative, yet specific change. Unsurprisingly, in both cases, the main intensity dips were observed around the hydrophobic patch regions. The proximal unit on the other hand, showed an overall decrease in intensity without much indication of specific areas of interaction. This is mainly due to the fact that most signals did not attenuate upon peptide binding and thus, unlike the other units, the same signals were assigned up to the point of saturation. As such, an overall broadening of signals is observed resulting in non-specific decrease in signal intensity corroborated by the extraction of individual rows and columns from the 2D spectra with the aim of extracting 1D traces in order to compare the line width before and after peptide binding (example of S20 shown in Fig 2.3; bottom, right; unbound state line width : 31.78 Hz; bound state line

width : 53.89 Hz). Nevertheless, there are some specific signal attenuations observed in the proximal unit (residues A46, G47, H68, V70) which suggest a binding role that is played by this unit that still requires further explanation (Fig 2.3; c).

Furthermore, the NMR spectra for the endo-3rd unit was interesting in that the end point of titration showed multiple new peaks for some residues which highly resemble the bound peaks observed in both distal and endo unit of Ub_3 which is further discussed later. Since this is usually a sign of more than one conformation caused by a change in the local environment of a protein, the question of whether there was more than one peptide binding to the tetramer was raised. To test this query, ^{1}H-T_2 measurements were done in order to measure the tumbling rate of the complex upon peptide binding whereby a consistent decrease in ^{1}H-T_2 time would indicate an increase in size of the complex.

The ^{1}H-T_2 measurements from the envelope of 1D spectrum of proximally labeled Ub_4 showed a continual decrease from 13.5 ms of free Ub_4 to 10.3 ms and 9.98 ms at 1 and 2 molar equivalents of the added Ub4ix peptide, respectively. This is indicative of gradual increase in the overall size of tetramer upon Ub4ix binding but the lack of much difference in these T_2 measurements between 1:1 and 1:2 tetramer to peptide ratios suggest that a secondary binding event is not occurring. Rather, it is possible that Ub4ix binds alternatively between two sections within the tetramer which are named ‘trimer moieties’ comprising of either Ub units 4-3-2 (distal trimer moiety) or units 3-2-1 (proximal trimer moiety) with a seemingly higher preference for the distal trimer moiety. As such, it was reasonable to examine if the results could be more understood by performing similar binding studies with Ub_3 and thereby possibly eliminating the chance of more than one binding interaction occurring. It should be noted that in the initial SPR measurements, Ub_3

was not investigated as a potential binding partner of the cyclic peptides and thus no K_d measurement had been done for trimer chains. All the results regarding Ub_3 binding of Ub4ix were novel results gained through NMR work as discussed below.

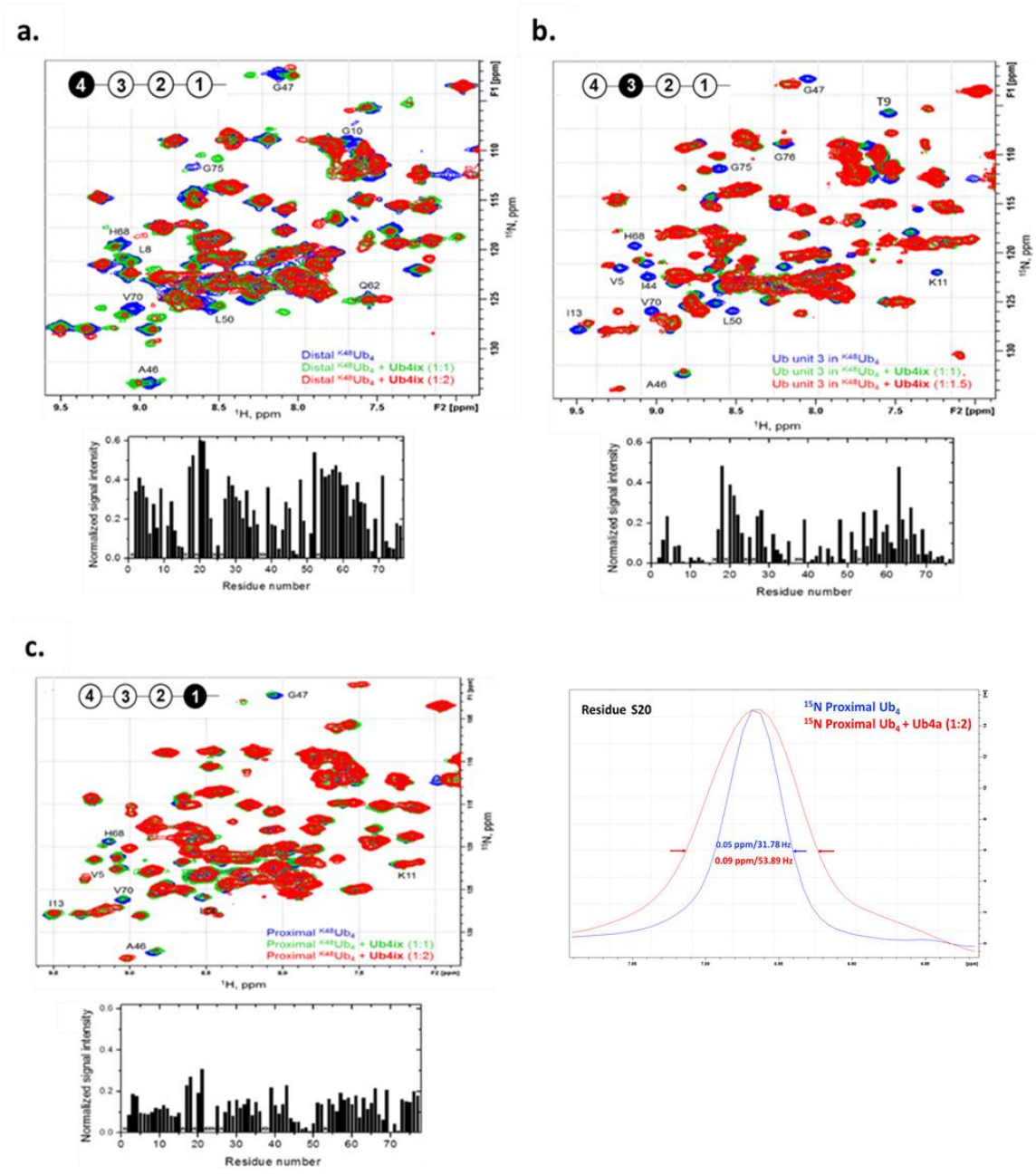


Figure 2.3: a) Perturbations in the distal Ub unit of Ub_4 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_4 (distal Ub) and in the presence of 1 (green) and 2 (red) molar equivalents of Ub4ix (top, left); **b)** Perturbations in the ^{15}N-enriched 3rd unit of Ub_4 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_4 and in the presence of 1 (green) and 1.5 (red) molar equivalents of Ub4ix (top, right); **c)** Perturbations in the proximal unit of Ub_4 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_4 and in the presence of 1 (green) and 2 (red) molar equivalents of Ub4ix (bottom, left). 1D signal Overlay of S20 in bound (blue) and unbound (red) states with measurements of line width indicating signal broadening (bottom, right). Select perturbed residues in each panel are indicated. Note several alternate (bound) peaks, for example, for residues G47 and T9. Intensities of the unbound signals at the endpoint of titration was normalized to the corresponding intensity measured in the absence of peptide (under each panel). Asterisks indicate residues that could not be detected or could not be reliably quantified due to signal overlap.

2.3.1 Mapping of Ub4ix interaction with lysine-48 linked Ub trimer

Given the results obtained for Ub_4 interaction with Ub4ix, the hypothesis for how the peptide would interact with a trimer entailed that the endo unit of the trimer would be the strongest binder while the distal and proximal units would interact the least with the peptide. As such, similar to the previous study, titration of Ub4ix into Ub_3 with individual units ^{15}N-labeled was conducted.

The NMR experiments showed that overall interaction of the cyclic peptide with Ub_3 was more pronounced than for any of the tetramer experiments. As expected, the endo unit of the trimer showed the most perturbations in almost the entire spectra indicating a very strong interaction with Ub4ix (Fig 2.4; b). For all three units, the signal intensity measurements showed the most effect in the hydrophobic patch with the proximal unit exhibiting more attenuations that were not observed in its tetramer counterpart (Fig 2.4; c).

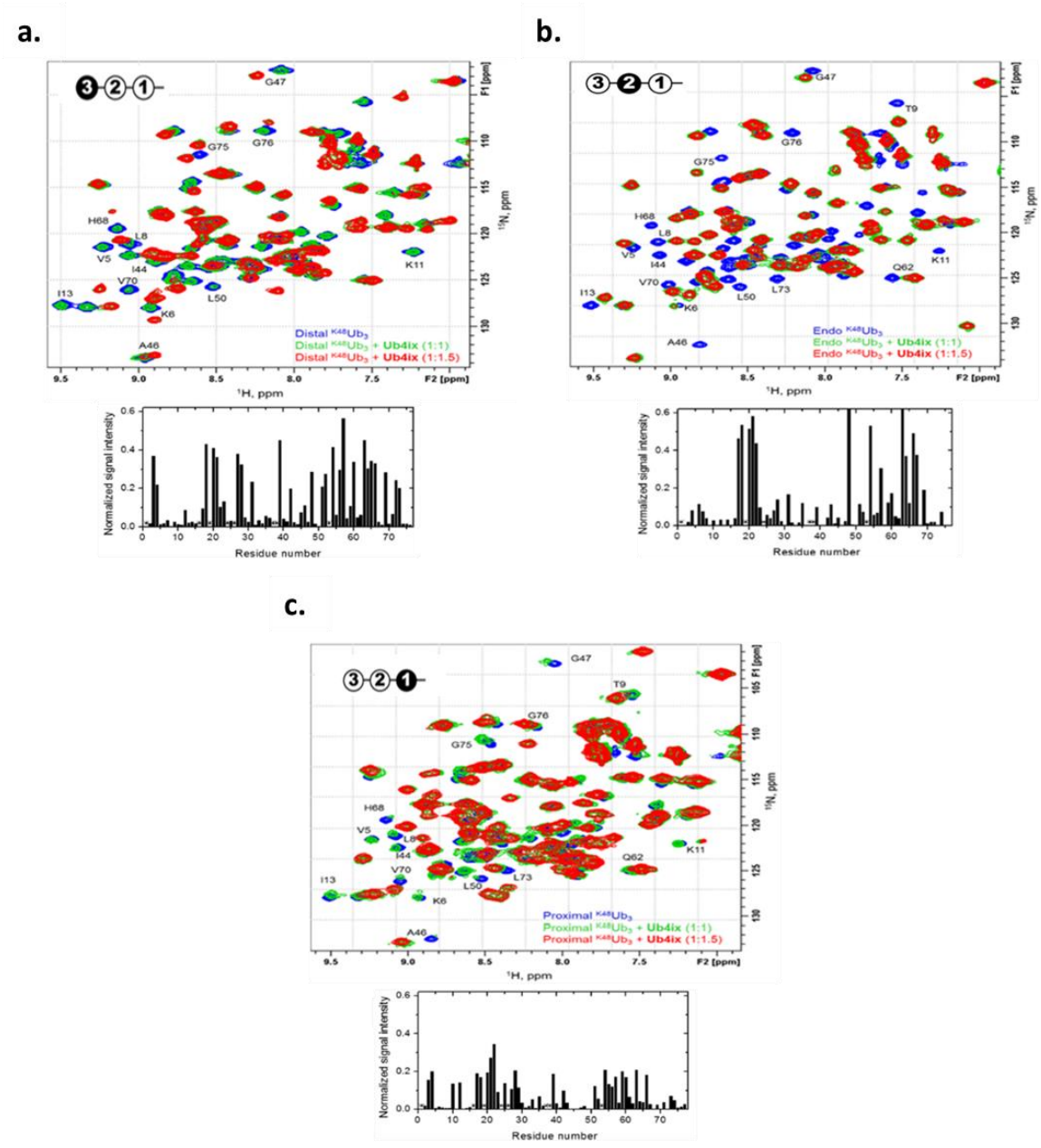


Figure 2.4: a) Perturbations in the distal Ub unit of Ub_3 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_3 (distal Ub) and in the presence of 1 (green) and 2 (red) molar equivalents of Ub4ix (top, left); **b)** Perturbations in the ^{15}N-enriched endo unit of Ub_3 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_3 and in the presence of 1 (green) and 1.5 (red) molar equivalents of Ub4ix (top, right); **c)** Perturbations in the proximal unit of Ub_3 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_3 and in the presence of 1 (green) and 2 (red) molar equivalents of Ub4ix (bottom). Select perturbed residues in each panel are indicated. Intensities of the unbound signals at the endpoint of titration was normalized to the corresponding intensity measured in the absence of peptide (under each panel). Asterisks indicate residues that could not be detected or could not be reliably quantified due to signal overlap.

Comparing the NMR spectra of Ub tetramer and Ub trimer chains indicates that Ub4ix readily binds to a trimer as much as to a tetramer. In fact, based on the signals attenuations observed, the results indicate that the peptide is more selective towards a lysine-48 linked trimer than a tetramer with the perturbations in the endo unit of Ub_3 showing full saturation at an equimolar addition of cyclic peptide.

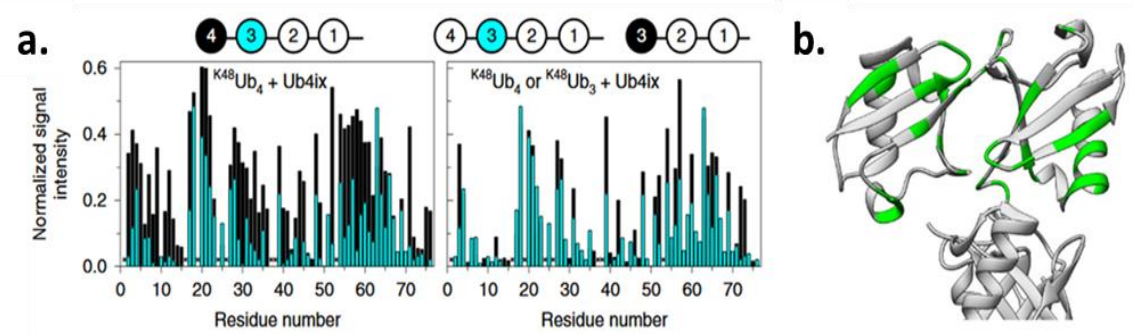


Figure 2.5: a) Comparison of the intensities of the unbound signals of the most distal (unit 4, black bars) and the next to distal (unit 3, cyan bars) Ub units in Ub_4 upon addition of Ub4ix in 2:1 and 1.5:1 molar ratio, respectively; **b)** Map of the residues showing strong signal attenuations (normalized signal intensity <0.2 colored green) in the two distal Ub units (4 and 3) of Ub_4 on the crystal structure of the closed form of Ub_4. (PDB: 2O6V)[13].

It should be noted that when comparing residue-specific signal intensities on similarly labeled units of both the trimer and tetramer upon peptide binding, the difference in strength of Ub4ix binding between the polyubiquitin chains is observable (Fig 2.5; a). The overlays of the intensity plots explicitly show that the distal unit of the trimer and the endo-3rd unit of the tetramer have very similar intensities further indicating that a trimer moiety is the main binding target of Ub4ix (Fig 2.5; a). In addition to this, spectral comparisons of both poly-Ub chains show similar patterns of change upon peptide binding with the main shifts and attenuations occurring in several of the same residues within the

units observed (Fig 2.6). This further complicates the matter since it suggests that the proximal moiety in the tetramer is preferentially selected for Ub4ix binding.

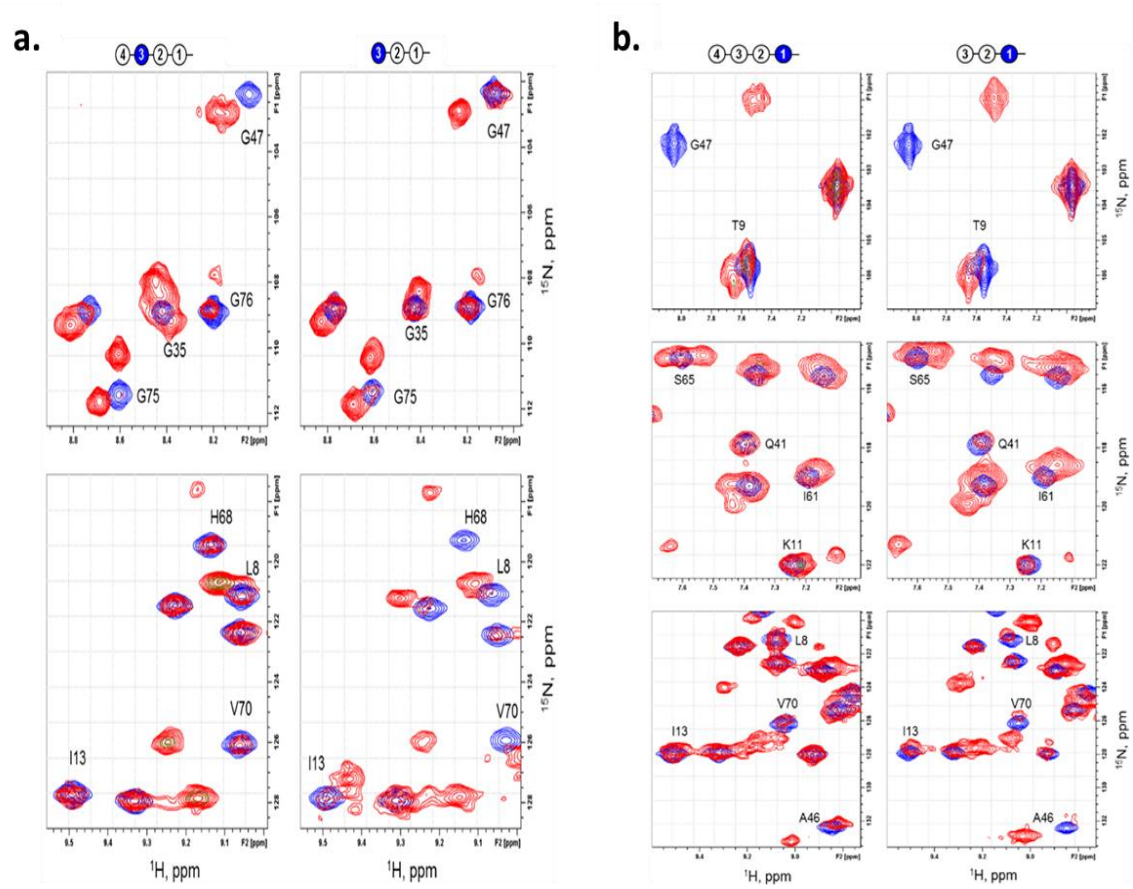


Figure 2.6: Spectra comparison of ; **a)** unit 3 in Ub_4 and the distal unit in Ub_3 and **b)** proximal Ub units in Ub_4 and Ub_3. Similar units exhibit similar patterns of residue-specific NMR signal shifts and attenuation upon peptide binding. Shown is an overlay of regions of the ^{1}H-^{15}N correlation spectra of Ub units in Ub_4 (left column) and Ub_3 (right column), in unbound free state (blue) and in the presence of 1 molar equivalent (red) of Ub4ix. Unbound signals corresponding to select residues are indicated.

In fact, the perturbations in unit 3 of Ub_4 show multiple residue peaks which are very similar to the perturbations observed in both the distal and the endo units of the trimer chain (Fig 2.7). This fully suggests that Ub4ix binding to tetramer is much more ambiguous than anticipated in terms of binding mode selectivity. Further experiments using DUB

assays were performed to try and address this ambiguity and the results are discussed in the following section.

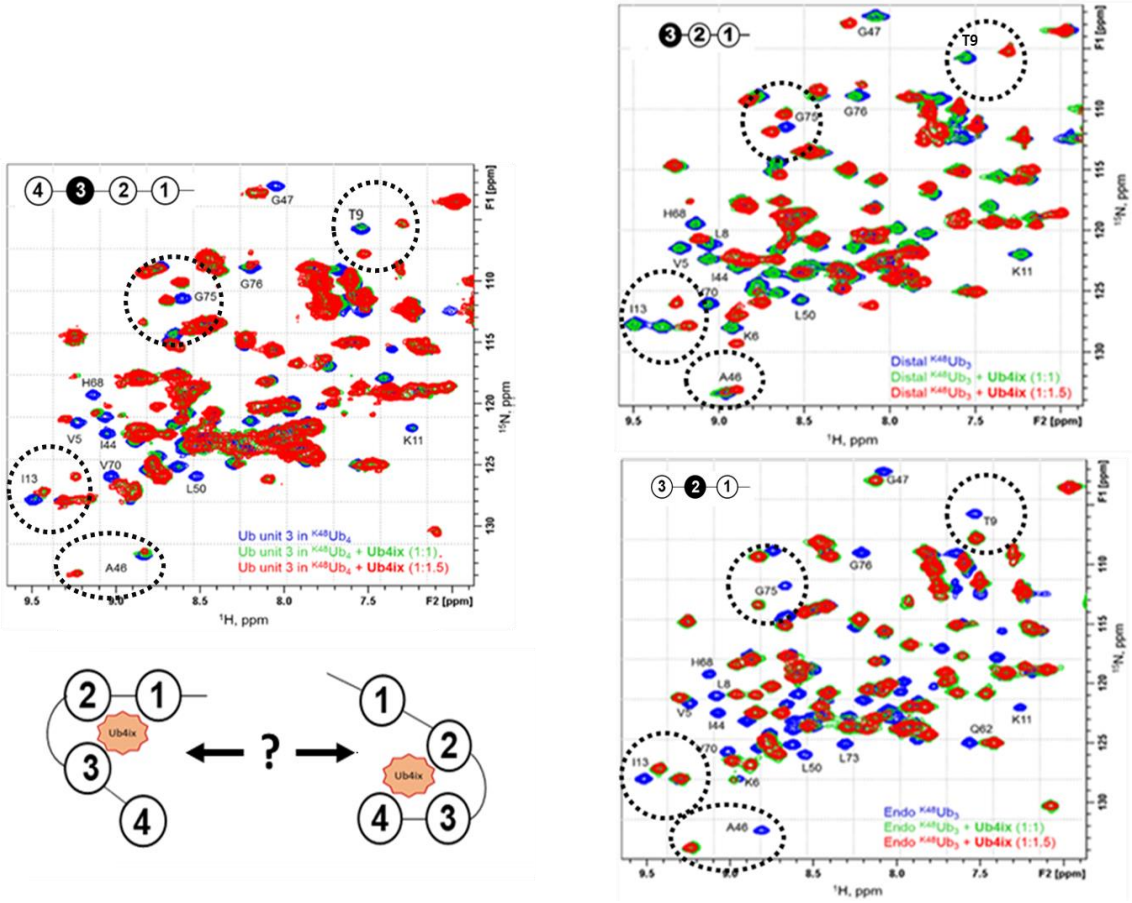


Figure 2.7: Comparison of spectra in Ub4ix bound tetramer and trimer chains. Residues in the endo 3rd unit of Ub_4 exhibiting multiple Ub4ix bound peaks are observed in spectra of distal and endo units of Ub_3 (e.g., circled signals). The potential of some form of equilibrium in preferential binding mode of Ub4ix to Ub_4 is suggested as illustrated by cartoon form.

Similar to what was done in SPR measurments, a control titration of Ub4ix into distally labeled Ub_2 was also performed. The results at the endpoint of titration (equimolar amount of peptide addition), showed a spectrum with either broadened or attenuated peaks and very few new peak appearances (Fig 2.8). The experiment suggests a case of mostly intermediate chemical exchange with most of the effect observed at the hydrophobic patch

which is already recognized as a general binding site for Ub. This result confirms the weaker interaction between dimer and the peptide observed through SPR measurements and shows that Ub4ix can be selective for a particular chain length when being considered as a potential therapeutic agent even though how that selectivity occurs requires further study.

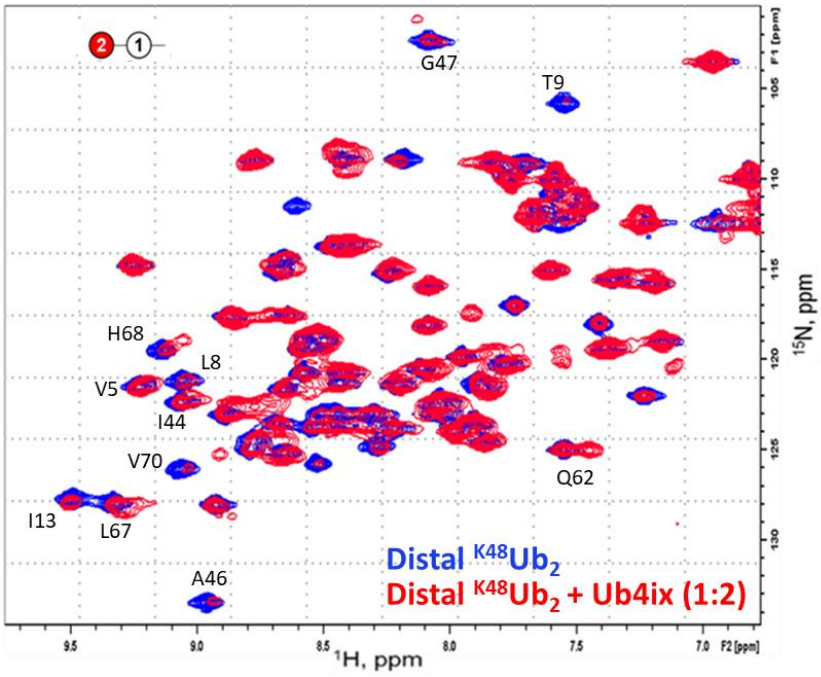


Figure 2.8: Perturbations in the distal unit of Ub_2 upon Ub4ix binding. Overlay of the ^{1}H-^{15}N correlation spectra of free (blue) Ub_2 and in the presence of an equimolar equivalent of Ub4ix (red).

2.3.2 Ub4ix binding confers protection to Ub chains against DUB cleavage

Deubiquitinases (DUBs) are a family of enzyme proteases that regulate the ubiquitination process and maintain the presence of free ubiquitins in cells[9]. In order to further examine the binding of Ub4ix to Ub chains and see if this interaction can prevent

cleavage by proteases, different DUB assays were performed. As such, initial assays with OTUB1 and Usp2 proteases had been performed by our collaborators in Dr. Brick's lab which showed that Ub4ix binding generally prevents the tetramer from being cleaved to anything less than a trimer[31].

Even though Usp2 is a general DUB which cleaves Ub chains without linkage specificity or directionality[46, 47], OTUB1 is a DUB that is specific to lysine-48 linked chains requiring a freely available distal Ub unit in order to initiate its protease activity[47, 48]. As such, the assays which resulted in cleavage to Ub_3 suggested that the distal unit in the tetramer was free to be recognized by OTUB1 indicating that Ub4ix binds across the proximal moiety. Still, an in-house repeat of the DUB assays performed on both tetramer and trimer chains show that there is very limited cleavage by OTUB1 on Ub_4 with full protection from cleavage conferred to Ub_3 (Fig 2.9; a). These results seem to conflict with the DUB assays from our collaborators suggesting that in fact the distal unit in Ub_4 interacts with Ub4ix preventing any significant OTUB1 recognition (gel results after OTUB1 assay show a small amount of cleavage for Ub_4: Fig 2.9; a). In addition, a general, linkage non-specific DUB, Ubp6 was also used to test the protection garnered by the peptide to Ub chains. The results showed Ub_4 being cleaved to a trimer chain which suggests that at least one Ub unit is cleaved from the tetramer with full protection given to trimer (Fig 2.9; b). This is confirmation that the trimer is indeed the main binding target for Ub4ix but the question of which Ub unit is being cleaved by the DUB is still unclear.

To further test directionality of binding, a lysine-48 linkage specific DUB, IsoT, which preferentially cleaves Ub chains from the proximal end was utilized. In this assay, Ub4ix bound trimer chains were virtually protected from any type of DUB activity while

some amount of Ub4ix bound tetramer chains were cleaved by IsoT indicating that the proximal end of Ub_4 is partially susceptible to IsoT cleavage (Fig 2.9; c).

Based on both NMR and DUB assays, it is quite feasible that either the distal or proximal moieties within Ub tetramer could be selected for Ub4ix binding. Based on the difference in the extent of cleavage observed through OTUB1 and IsoT, preferential binding to the distal trimer moiety is suggested. Nevertheless, it is quite evident that Ub4ix still tightly binds Ub_4 chains with a trimer as the preferential binding target.

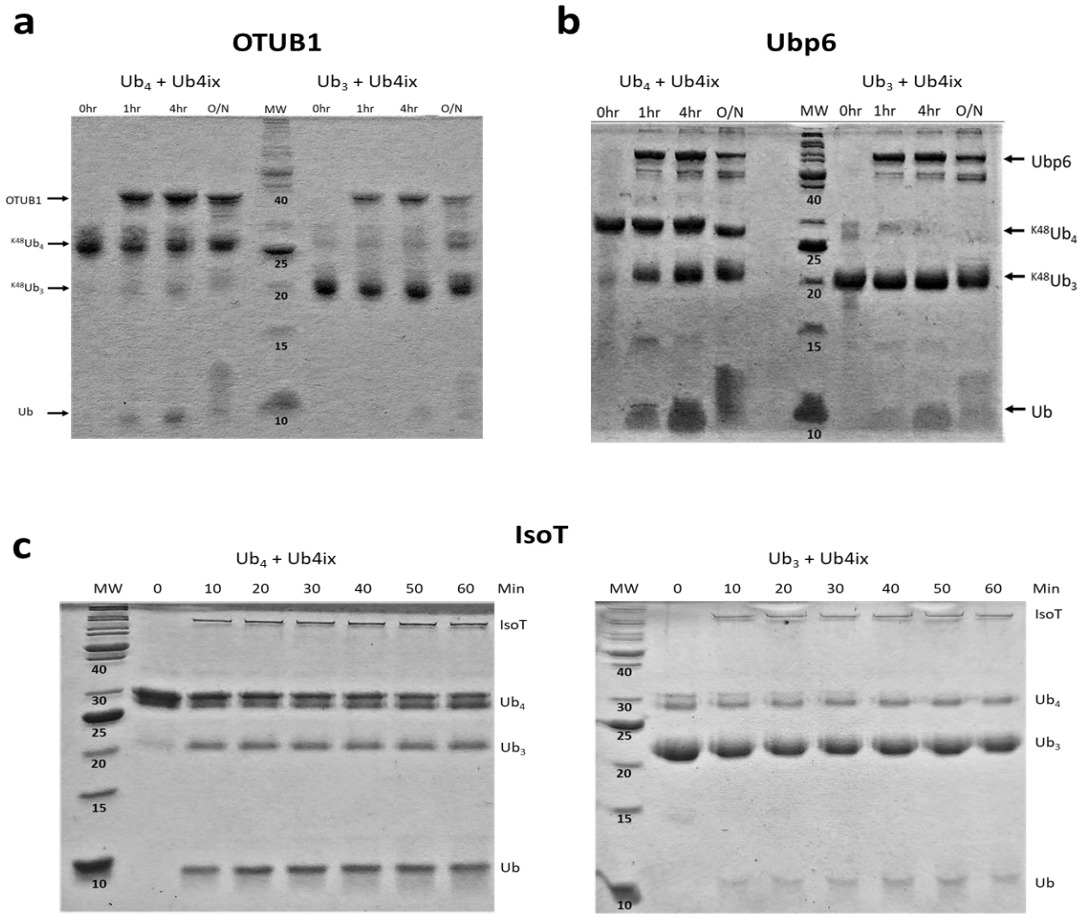


Figure 2.9: Cyclic peptide Ub4ix protects lysine-48 linked Ub-chains from disassembly by deubiquitinases. Time course disassembly of **a)** Ub_4 (left) and Ub_3 (right) by distal lysine-48 linkage specific OTUB1; **b)** Ub_4 (left) and Ub_3 (right) by linkage-nonspecific DUB Ubp6 and **c)** Ub_4 (left) and Ub_3 (right) by the proximal specific DUB IsoT in the absence and presence of Ub4ix. Corresponding molecular weight values are shown in kDa. MW = molecular markers, O/N = overnight.

2.3.3 Attempt to co-crystallize Ub_4 with Ub4ix peptide

An additional method to observe the interaction of Ub4ix with poly-Ub chains was to get a crystal structure of the bound complex. As such, a co-crystallization attempt was performed where Ub4ix-Ub_4 complex was initially screened for crystals using PEG and Index suite from Hampton research. The Phoenix© protein crystallization robot was used to set up a hanging drop vapor diffusion screen.



Figure 2.10: Crystal tray selected for diffraction with small crystal observed.

The few conditions that showed crystals in the Index screen were selected and subsequently harvested for diffraction. Only one tray had crystals that were able to be diffracted at 1.7 Å resolution (Fig 2.10) with subsequent attempts at solving the structure showed that there was no electron density observed for the peptide. The lysine-48 linked ubiquitin tetramer that was solved showed a structure that was similar to an already deposited 'closed' structure of lysine-48 linked tetraubiquitin even though slight differences are observed in the overlay due to the fact that full refinement was not concluded on the in house tetramer structure that was obtained as well as the potential difference in the resolution between the two structures (Fig 2.11)[13]. Further attempts of optimization to get diffraction-quality crystals that had the whole complex intact were not

successful leaving unanswered the question of directionality of binding within a tetramer. Given the seemingly much tighter binding of this peptide to lysine-48 linked trimer chains, it is possible that an attempt of crystallizing an Ub4ix:Ub_3 complex would result in a successful outcome.

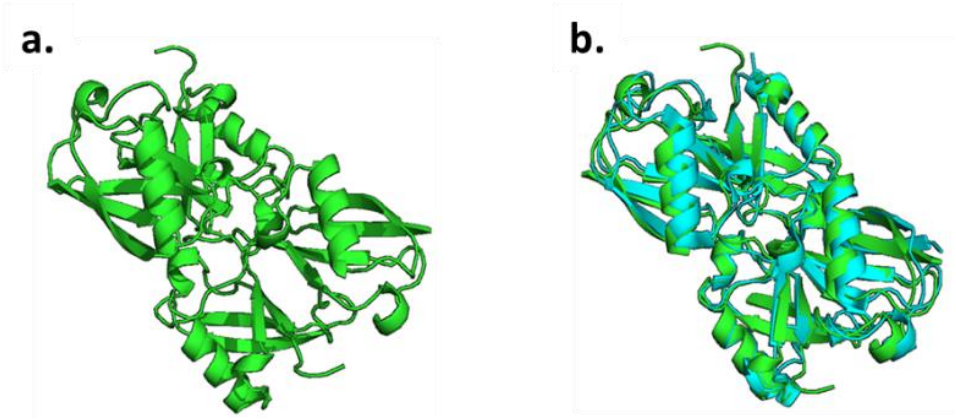


Figure 2.11: **Crystal structure results of Ub_4 bound to Ub4ix. a)** Solved tetramer structure from the electron density observed for Ub_4:Ub4ix complex. No electron density for the peptide was observed; **b)** Overlay of observed crystal structure (green; 1.7 Å) onto published lysine-48 linked tetraubiquitin (Cyan; 2.20 Å) (PDB: 2O6V)[13].

2.4 Ub4a cyclic peptide binding to lysine-48 linked Ub chains

Despite the discovery of Ub4ix peptide which was shown to specifically bind lysine-48 linked Ub chains, one of its main shortcomings was that its physicochemical properties were poor. Ub4ix is relatively large (~1.7 kDa), with proteinogenic amino acids that makes it susceptible to protease-recognition[31] as well as rich in H-bond acceptors and donors making cell permeability an issue. Thus, better cyclic peptides with improved bioactivity in targeting the UPS were necessary for research purposes as well as potential drug leads[39].

To that end, as was done in the discovery of Ub4ix, RaPID protocol was used to assemble a trillion-member plus cyclic peptide library that were ultimately incubated with Ub_4. In this process, Ub4a, an 11 amino acid macrocycle, with two N-methyl, non-canonical side chain and an initiator D-amino acid with no formally charged amino acids at pH-7.0 was discovered to be the best candidate with apparent selectivity for Ub_4. Based on SPR results, Ub4a displayed high affinity with a K_D of ~9 nM for lysine-48 linked tetramer chains while binding to the anti-target Ub_2 was considerably weaker, with >100-fold higher K_D (Fig 2.12)[39]. This peptide was comparable to the orally-available, membrane-permeable, macrocyclic peptides cyclosporin ($\leq$ 1.2 kDa) [49, 50] which similarly, has no formal charges at pH-7.0 and fewer H-bond interacting residues [51-53].

Despite the previous NMR based observations that had resulted in identifying Ub_3 as the main binding target for Ub4ix cyclic peptide, the initial SPR measurements done by collaborators for Ub4a binding still did not include trimers as a target. Nevertheless, since our lab had demonstrated that the trimer moiety was in fact the target of binding for the initial cyclic peptide discovered, subsequent NMR work studying the interaction of Ub4a with Ub chains started with the working hypothesis that trimer chains remained to be the main binding target for this new cyclic peptide.

Ub4a = Ac-^{D}FQYWOY^{Me}AT^{Me}GVCG

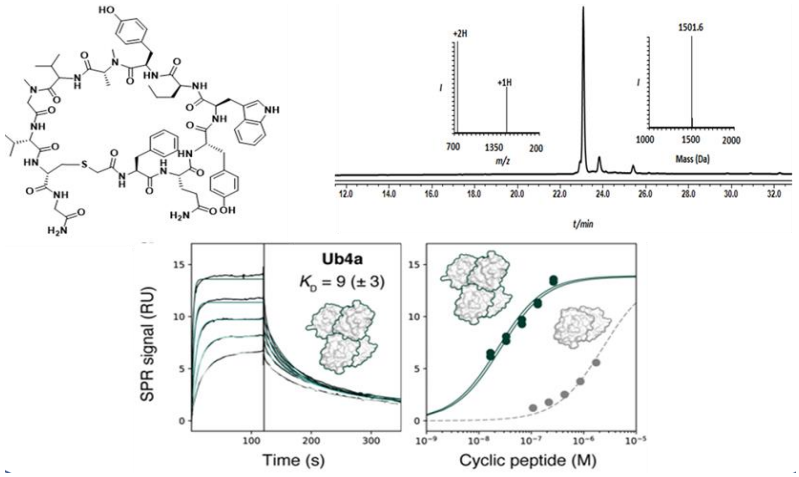


Figure 2.12: Ub4a cyclic peptide sequence along with ESI-MS analysis with observed mass of 1.5 kDa. Surface plasmon resonance (SPR) kinetics (bottom left) reveals tight (nM K_D) binding of cyclic peptide Ub4a to Ub_4. Comparison of SPR amplitudes between tetramer and dimer (bottom right) shows significantly weaker (>1 µM K_D) binding to Ub_2[39].

2.4.1 Mapping of Ub4a interaction with lysine-48 linked poly-Ub using NMR techniques

In order to map the physical interactions between Ub4a and lysine-48 linked Ub chains in addition to identifying the mode of interaction that exists between them, unlabeled Ub4a was titrated into different solution of selectively ^{15}N-labeled Ub chains, and the binding for each chain was monitored by NMR. Ub_3, Ub_4 and Ub_2 with specific Ub units uniformly ^{15}N-enriched were assembled from respective recombinant Ub monomers using controlled enzymatic chain assembly methodology as described earlier. The addition of Ub4a to Ub_4 caused strong residue-specific perturbations in the ^{1}H-^{15}N NMR spectra of the proximal Ub unit (Fig 2.13; a). Similar to Ub4ix binding, gradual decrease in intensity of the 'unbound' signals accompanied by appearance and increase of new 'bound' signals

at a different position in the spectrum was observed, consistent with the slow dissociation rates (~1 per minute) detected by SPR[39]. The perturbed residues were located primarily in and around the hydrophobic patch on Ub surface, as well as the α-helix, with most of the spectral changes saturating at about 1:1 peptide to Ub_4 molar ratio. Compared to the previous peptide, the effect of Ub4a binding on the proximal unit was much more pronounced across the entire spectra suggesting that Ub4a binding to this unit was much tighter than it was for Ub4ix. On the other hand, substantially weaker changes in the NMR spectra were observed for the distal unit of Ub_4, where the signal shifts were much smaller. Furthermore, most residues in this unit exhibited gradual signal shifts characteristic for fast exchange regime on the NMR time scale, which is indicative of weaker binding (Fig 2.13; b). Thus, unlike the higher levels of attenuation and emergence of new 'bound' peaks that was observed for the distal unit during Ub4ix titration, these results indicate that the distal Ub in the tetramer is not as directly involved in Ub4a binding and suggests that the proximal trimer moiety is the sole peptide-binding element within Ub_4.

In order to further test the hypothesis of trimer selectivity, we examined Ub4a binding to the proximal, endo, and distal Ub units of Ub_3. All three Ub units exhibited slow-exchange behavior upon Ub4a titration indicative of tight binding. The spectral changes in the proximal unit of Ub_3 were almost identical to those in the proximal unit of Ub_4 (Fig 2.14; a) pointing to strong similarity in the peptide's contacts within both types of chains. Conversely, the distal unit of Ub_3 exhibited strong NMR signal shifts (Fig 2.14; b), in stark contrast to the distal unit of Ub_4. Similar to what was observed in the Ub4ix binding to Ub_3, the endo unit in the Ub4a titration exhibited very strong attenuation and emergence of new peaks in almost the entire spectra (Fig 2.15; a). These results indicate

that all three units in Ub_3 are engaged in tight interactions with Ub4a with the middle unit exhibiting the bulk of the binding effect.

Unlike the analysis made for Ub4ix binding, in addition to 'minimum signal shift difference' used to assign signals, ^{15}N-edited TOCSY and NOESY spectra were recorded at the Ub4a titration endpoints, which allowed for more accurate assignment of a significant portion of the peptide-bound Ub chain NMR signals. Based on these assignments, the CSP's that were measured for each titration experiments (panels under each spectrum in the figures), corroborate that the hydrophobic patch in all the Ub units is involved in the peptide binding interactions. Furthermore, the extent of each signal perturbations (measured by ppm) and the number of attenuations for each unit give an indication on how strong the interaction of each respective unit is to Ub4a.

To test the SPR results of Ub4a selectivity for Ub chains longer than the dimer, Ub4a peptide was titrated into both proximally and distally ^{15}N-labeled Ub_2. Analogous to the Ub4ix experiments of a dimer, there was little or no signal shifts in the NMR spectra while strong signal broadening and disappearances for several residues located on the hydrophobic patch surface of the proximal unit were observed (Fig 2.16; a). This behavior is indicative of intermediate exchange on the NMR time scale, consistent with faster dissociation rates hence weaker binding to Ub_2 as was detected by SPR.

These NMR results corroborate tight Ub4a binding to lysine-48 linked Ub chains of length ≥ 3 and revealed differential and residue-specific physical interactions between the peptide and various Ub units in these chains. The observed spectral perturbations generally map Ub4a binding sites to the hydrophobic patch surfaces of the Ub units involved. However, there are clear differences in the strength and specificity of these

perturbations depending on the chain length and the position of the Ub unit within each chain which lends to the claim that the trimer (proximal moiety in Ub_4) is the Ub4a recognition element in lysine-48 linked chains. Unlike Ub4ix, the specificity of Ub4a for the proximal moiety in the tetramer appears is unambiguous.

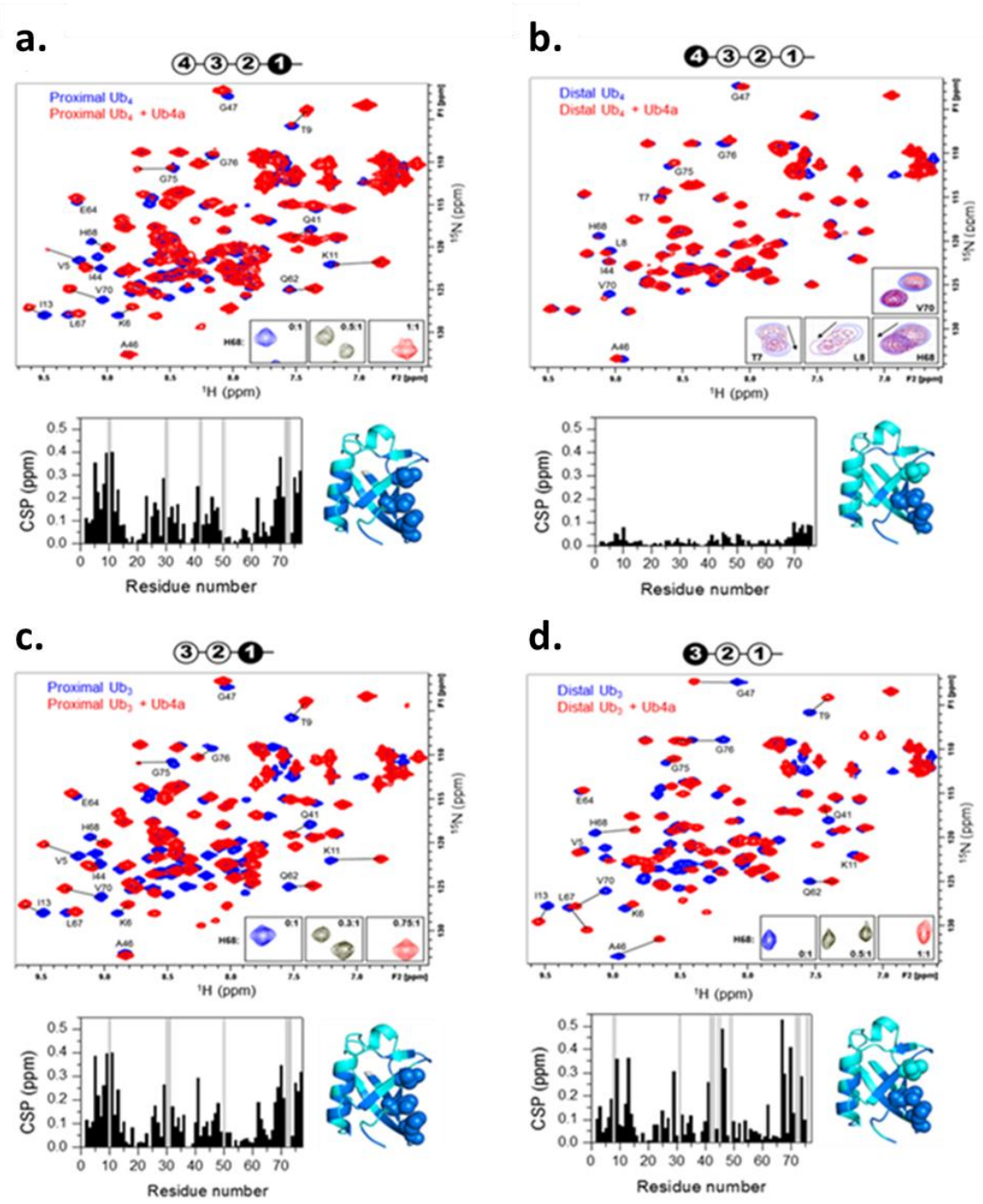


Figure 2.13: Cyclic peptide Ub4a binds to hydrophobic patch residues in lysine-48 linked tri- and tetra-Ub and engages first three Ub units in the chain: Shown are overlays of ^{1}H-^{15}N NMR spectra of free (blue) and Ub4a-bound (red) states of the proximal (**a**) and distal (**b**) Ub units of Ub_4 and proximal (**c**) and distal (**d**) Ub units of Ub_3. Select residues are indicated with free and peptide-bound states connected by lines. Insets in **a**, **c**, and **d** zoom on the signals of H68 to illustrate the slow-exchange behavior (the numbers indicate peptide/polyUb molar ratio). Insets in **b** are examples of signal behavior in the course of titration for the indicated residues: fast exchange for T7, L8, H68 and slow exchange for V70. The plots below each spectrum show residue-specific chemical shift perturbations (CSPs) for each Ub unit; the grey bars correspond to residues for which unbound signals disappeared, but the bound signals could not be unambiguously identified. Shown to the right of each plot is the structure of Ub with residues exhibiting strong CSPs mapped (colored marine); the side chains of the hydrophobic patch residues L8, I44, V70 are shown as spheres.

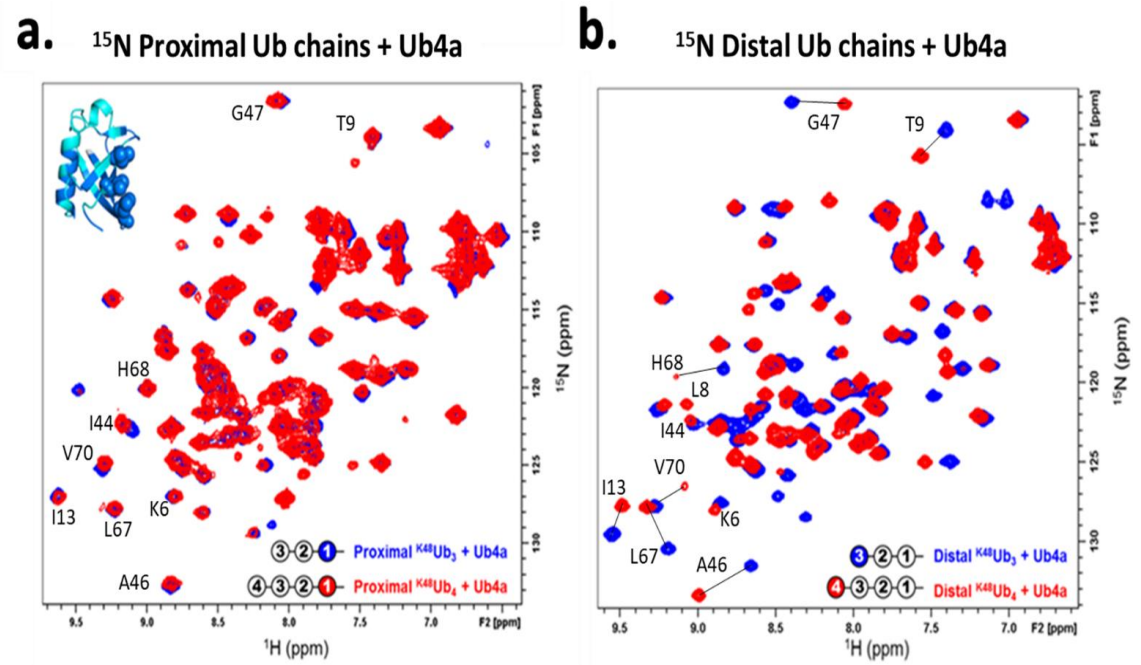


Figure 2.14: a) Overlay of ^{1}H-^{15}N NMR spectra of the proximal Ub in Ub_3 (blue) and in Ub_4 (red) at the endpoint of titration with Ub4a (peptide: polyUb molar ratio = 0.75:1 and 1:1, respectively). Note strong similarity of the two spectra. Inset in the upper right corner is the structure of Ub with residues exhibiting strong CSPs mapped (colored marine); the side chains of the hydrophobic patch residues L8, I44, V70 are shown as spheres; **b)** Overlay of ^{1}H-^{15}N NMR spectra of the distal Ub in Ub_3 (blue) and in Ub_4 (red) at the endpoint of titration with Ub4a (peptide: polyUb molar ratio = 1:1 and 2:1, respectively). Note substantial differences between the two spectra with various residue shifts shown as examples.

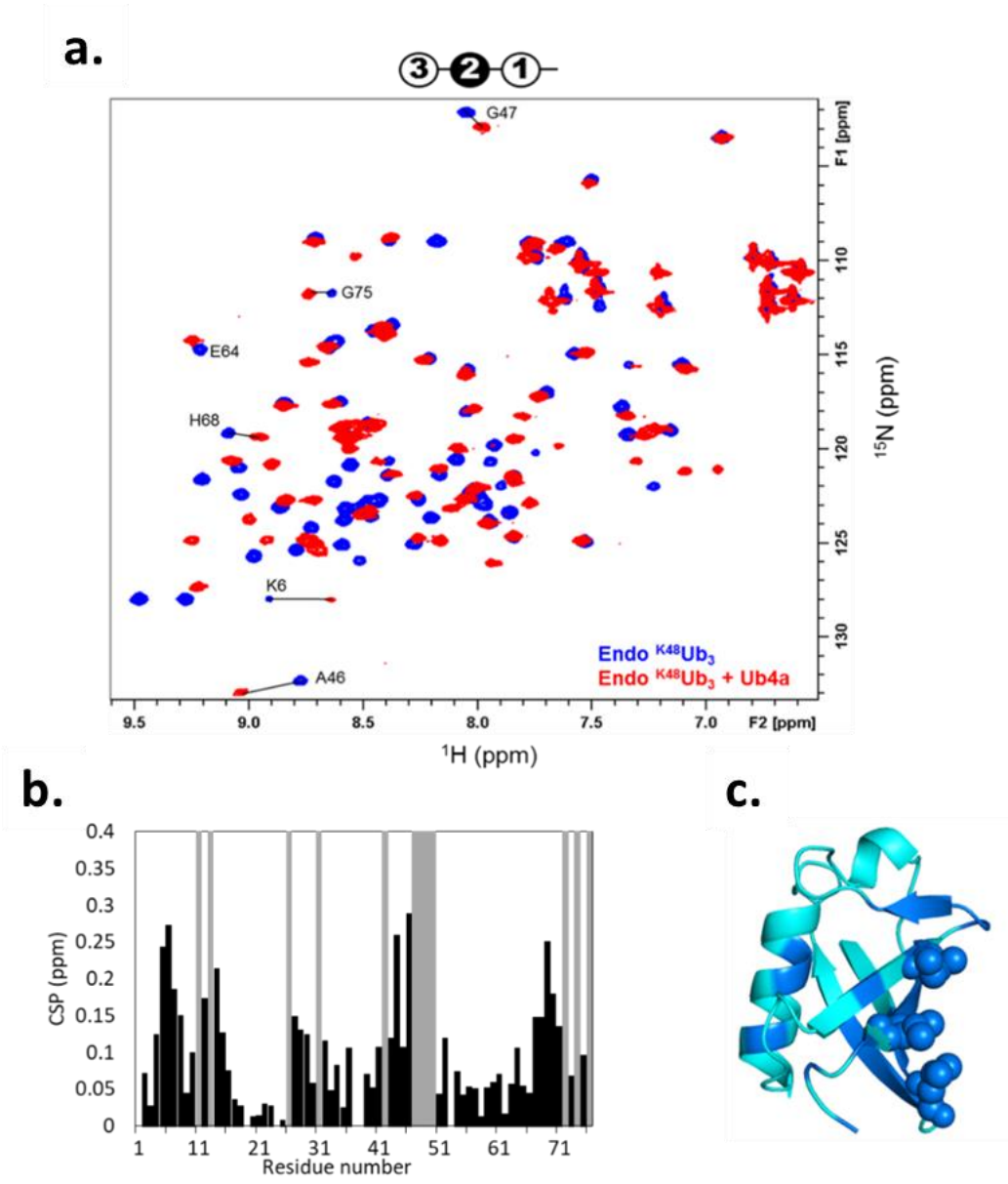


Figure 2.15: a) Overlay of ^{1}H-^{15}N NMR spectra of the middle (endo) Ub in Ub_3 free in solution (blue) and upon addition of Ub4a (red) (peptide: polyUb molar ratio = 2:1). Signals of select residues are indicated; **b**) Residue-specific CSP values (black bars). Grey bars indicate residues that exhibited signal attenuation during the titration; **c)** Structure of Ub with residues exhibiting strong CSPs mapped (colored marine); the side chains of the hydrophobic patch residues L8, I44, V70 are shown as spheres.

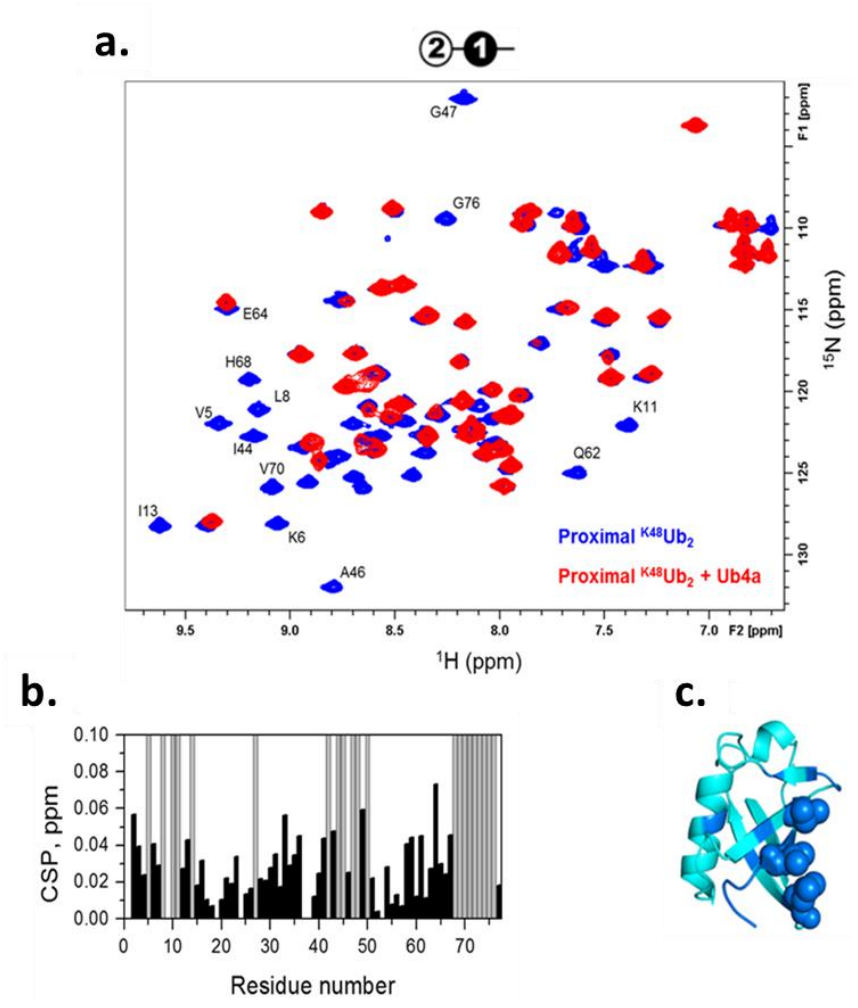


Figure 2.16: Ub4a binding to the proximal Ub of Ub_2. a) Overlay of ^{1}H-^{15}N NMR spectra of the proximal Ub of Ub_2 free in solution (blue) and upon addition of Ub4a (red) (peptide: polyUb molar ratio = 1.5:1); **b**) Residue-specific CSP values (black bars). Grey bars indicate residues that exhibited signal attenuation during the titration; **c**) Map of attenuated residues on the structure of Ub. The side chains of the hydrophobic patch residues L8, I44, V70 are shown as spheres. (PDB: 1D3Z). Notice the attenuation of many signals and the CSP magnitudes which are ~7 times less than for the proximal units in Ub_3 and Ub_4.

Given these outcomes, questions regarding the mechanism by which preferential binding to proximal trimer moiety in tetramer chains was taking place needed to be explored. The experiments conducted to address these questions are discussed in the next section.

2.4.2 C-terminal tail determines binding mode selectivity during binding of Ub4a to Ub_4

In the proximally labeled Ub-tetramer and trimer spectra (Fig 2.13; a, c), the flexible C-terminal residues seemed to exhibit higher CSP's than was expected suggesting an interaction occurring with Ub4a peptide. As such, in order to investigate if residues in the proximal unit tail of tetramer chains contributed to binding to the proximal trimer moiety of Ub_4 (explaining the weak interaction of the distal unit) as well as overall binding to a trimer moiety, ^{15}N-distally labeled Ub_3 and Ub_4 chains with truncated residues in their C-terminal tails were introduced.

For this purpose, two types of chains with partial truncation in the Ub proximal unit was prepared: **a)** truncation of the last four tail residues resulting in Ub_3-Ub_{R72} tetramer and Ub_2-Ub_{R72} trimer chains; and **b)** truncation of only the last two residues resulting in Ub_3-Ub_{R74} tetramer and Ub_2-Ub_{R74} trimer chains (Fig 2.17; a). These chains were enzymatically assembled as previously described with the NMR experiments focused on observing the effect of truncation on the binding of Ub4a to the distally ^{15}N labeled unit of each chain type. Interestingly enough, the extent of the partial cleavage of the C-terminus, had an effect on the binding of Ub4a to both tetramer and trimer Ub chains.

2.4.3 Binding of Ub4a to lysine-48 linked Ub_3-Ub_{R72} and Ub_2-Ub_{R72} chains

The NMR spectra for the distally labeled unit in Ub_3-Ub_{R72} showed an interestingly dramatic difference compared to the spectra of Ub_4 upon peptide binding suggesting a much stronger binding of Ub4a to the distal unit of this tetramer (Fig 2.17; b) as well as a

shift in the preferential binding of peptide to the distal moiety. In fact, overlaying the NMR spectra of peptide bound, distally labeled wild-type Ub_3 and Ub_3-Ub_{R72} showed an almost perfect agreement (Fig 2.17; c). This strongly suggested that the C-terminal tail residues do, in fact, have an effect in the peptide's mode of binding.

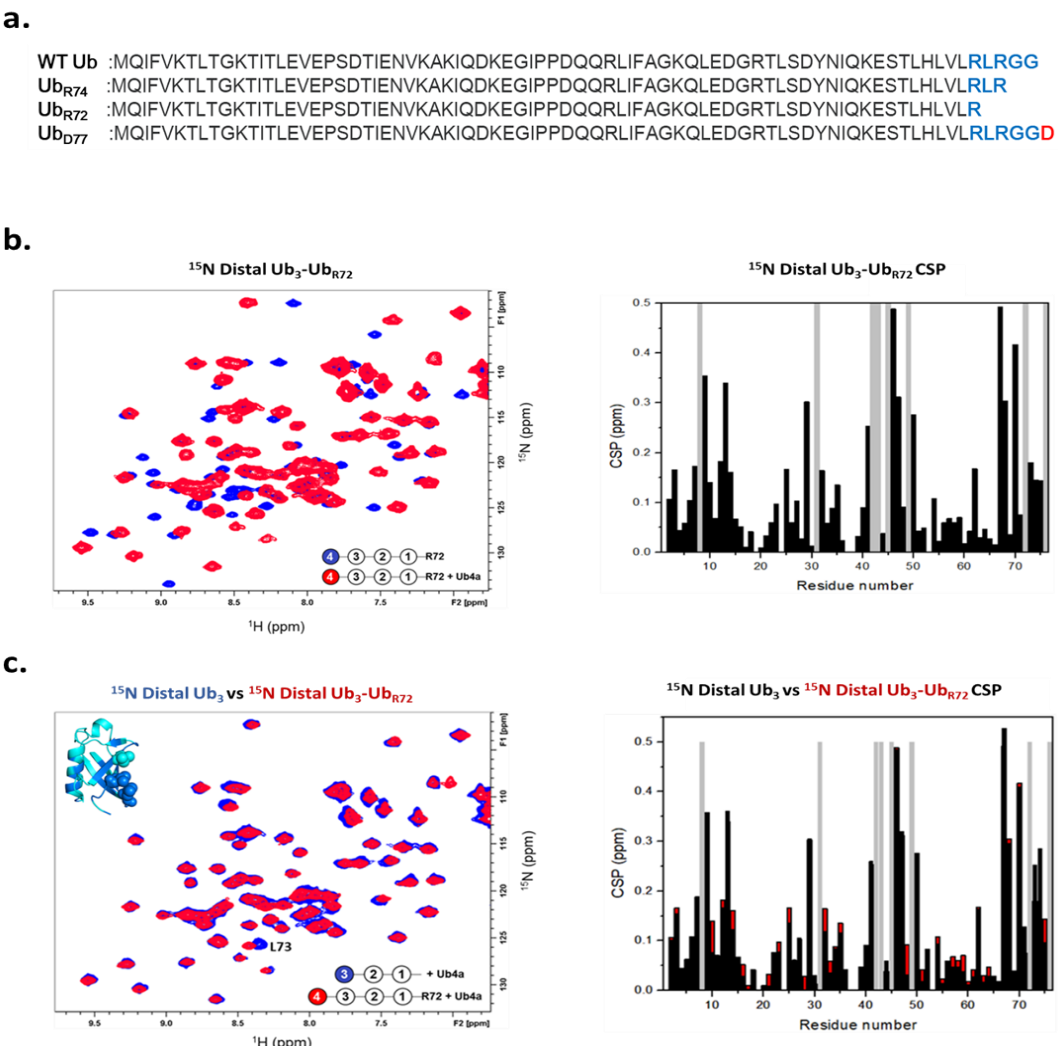


Figure 2.17: Ub4a binding to the distal Ub unit of Ub_4 with proximal Ub_{R72}. a) Amino acid sequences of wild-type ubiquitin and variants with truncated and added tail residues; **b)** Overlay of ^1H-^{15}N NMR spectra of the distal Ub of Ub_4 free in solution (blue) and upon addition of Ub4a (red) (left) with residue-specific CSP values (black bars, right); **c)** Overlay of ^1H-^{15}N NMR spectra of the distal Ub in Ub_3 (blue) and in Ub_3-Ub_{R72} (red) at the endpoint of titration with Ub4a (left) with residue-specific CSP values of wild-type Ub_3 (black bars) and Ub_3-Ub_{R72} (red bars) (right); **inset structure:** Overlapping CSPs from both chains types mapped on the structure of Ub. The hydrophobic patch residues L8, I44, V70 are shown as spheres. Grey bars indicate residues that exhibited signal attenuation during the titration. Note similarities between the two spectra and CSP. (Peptide: polyUb molar ratio = 1.5:1). (PDB: 1D3Z).

With the NMR titration of Ub4a into the trimer Ub_2-Ub_{R72}, the results showed a different binding mode to what was observed with distally labeled wild-type trimer (Fig 2.18; a). In the case of the trimer with the full tail truncation, rather than a slow exchange regime, the spectrum showed signal attenuations with increased amount of Ub4a addition. This was very similar to what was observed with Ub4a bound to dimer and the overlay of the spectrum of the end points for distally labeled wild-type Ub_2 and Ub_2-Ub_{R72} showed a striking similarity to each other which further indicated that truncation of 4 residues from the C-terminus hinders recognition of proximal trimer moiety by Ub4a. (Fig 2.18; c).

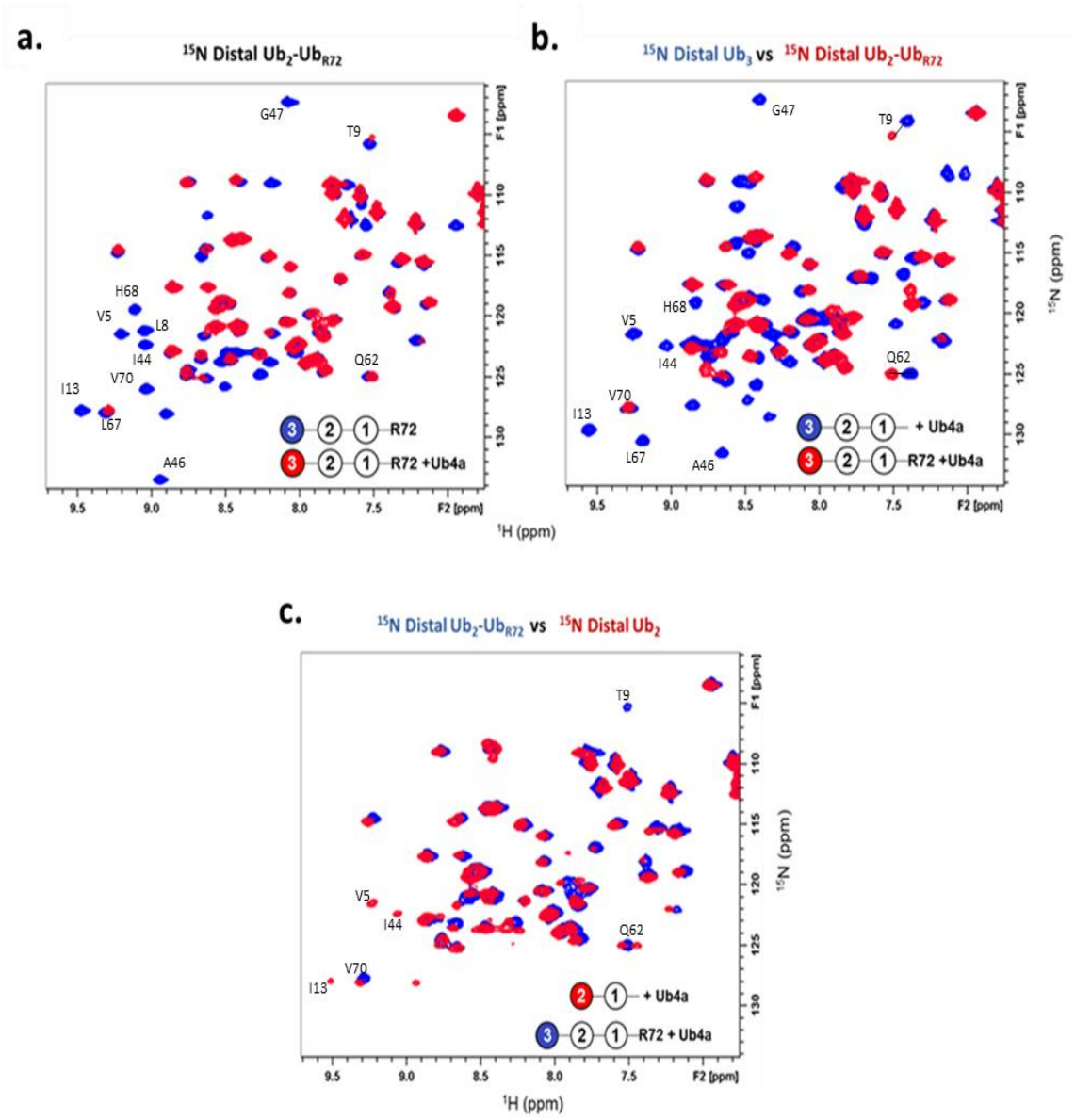


Figure 2.18: Overlay of ^{1}H-^{15}N NMR spectra of a) distal Ub in Ub_2-Ub_{R72} free in solution (blue) and upon addition of Ub4a (red) (peptide: polyUb molar ratio = 1:1); **b)** distal Ub in wild-type Ub_3 (blue) and Ub_2-Ub_{R72} (red) at the endpoint of titration with Ub4a (peptide: polyUb molar ratio = 1:1); **c)** distal Ub in Ub_2-Ub_{R72} (blue) and wild-type Ub_2 (red) at the endpoint of titration with Ub4a. Note similarities and differences in shifts and attenuations for each spectrum. (Peptide: polyUb molar ratio = 1:1)

2.4.4. Binding of Ub4a to lysine-48 linked Ub_3-Ub_{R74} and Ub_2-Ub_{R74} chains

When looking at the distal units of Ub chains in cases where titration was performed with either Ub_3-Ub_{R74} (tetramer) or Ub_2-Ub_{R74} (trimer), each spectrum showed that binding of peptide to these chains was generally similar to that of their wild-type counterparts. Ub_3-Ub_{R74} did not show a slow exchange regime while the Ub_2-Ub_{R74} trimer showed a slow exchange regime of binding. A slight difference that could be of interest arose in the spectra of Ub_3-Ub_{R74}, wherein it appears that residues within the hydrophobic patch were much more affected, as shown by their full attenuation at the end of titration (Fig 2.19; a) compared to wild-type tetramer which had exhibited virtually no change in signal intensity by the end of titration (Fig 2.19; b). On the other hand, the titration comparison between wild-type Ub_3 and Ub_2-Ub_{R74} showed much better congruence with spectra in both cases showing a typical slow exchange regime with the differences between the spectra relegated to a few residues that shifted at different places by the end of titration which might be due to allosteric changes caused by truncation of the tail (Fig 2.19; d). Although, at this junction, further structural study is needed in order to make definite claims on why there are differences between wild-type and tail truncation chains in a number of the residues during binding. But based on previous knowledge of Ub chain interactions, it is hypothesized that the truncation at the tail end has an effect on the interaction of the individual units (especially distal with proximal) within the chains and thus during peptide binding, this effect is exaggerated.

In both trimer and tetramer cases, the difference in the binding mode of Ub4a seems to arise due to the truncation of the respective tail regions. Even though this effect does not seem to reach to the point where Ub4a binding is inhibited, the titration results of chains

with truncated tail residues point to the fact that binding specificity of Ub4a within lysine-48 linked chains depends on the presence of , at least, proximal tail residues up to and including arginine-74. As such, these results demonstrate why only the proximal moieties within a longer chain are preferred during binding and also demonstrate that when the tail region of the proximal unit is not fully available (i.e. with R72 and L73 present), Ub4a will bind to a different trimer moiety that has an intact tail (i.e. distal trimer moiety) albeit in a seemingly much weaker manner suggesting the necessity of free C-terminus for maximum binding affinity in both trimer and tetramer chains. This is further explored in the section with DUB assay analysis.

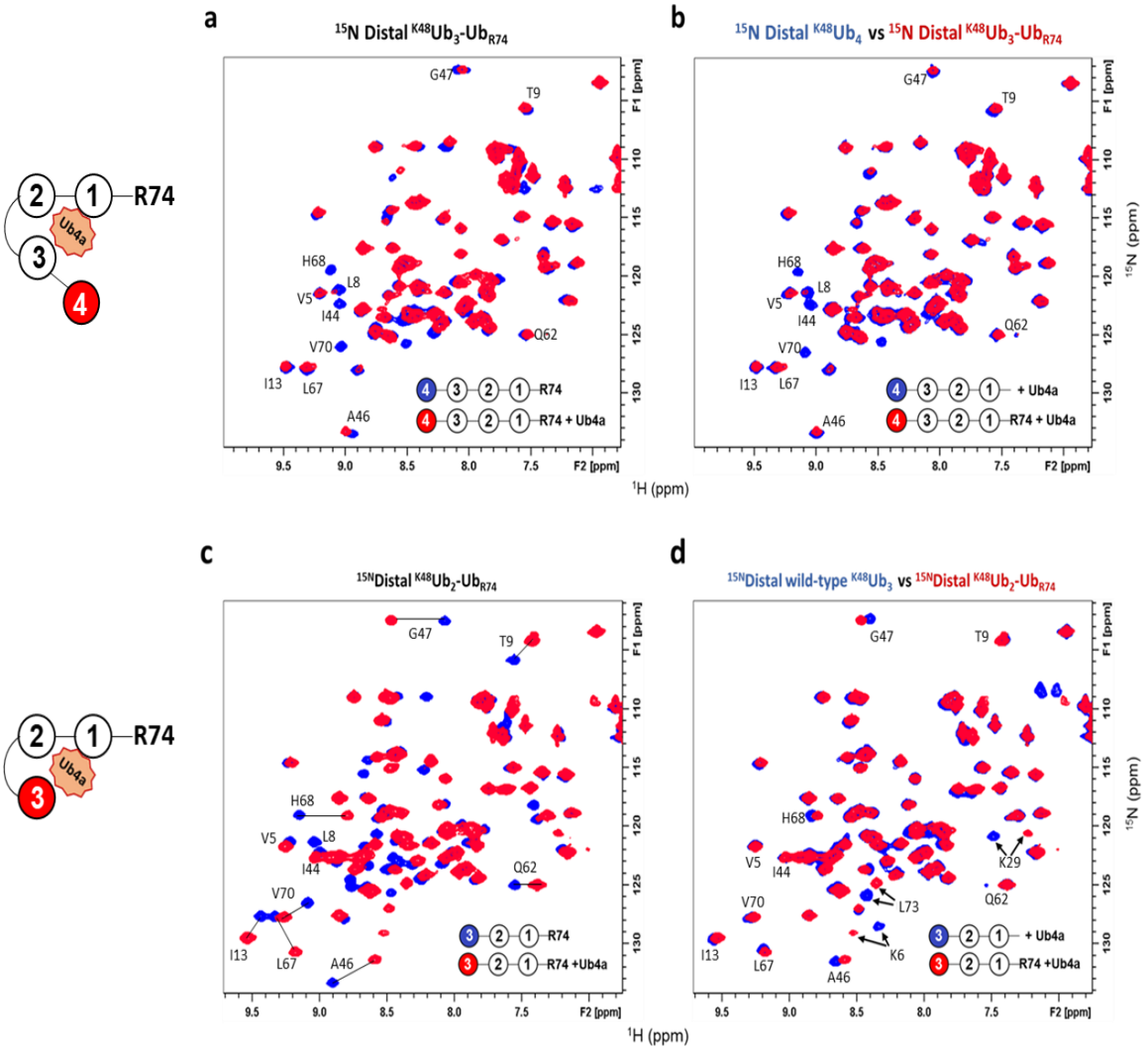


Figure 2.19: Overlay of ^{1}H-^{15}N NMR spectra of **a)** distal Ub in Ub_3-Ub_{R74} free in solution (blue) and upon addition of Ub4a (red) (peptide: polyUb molar ratio = 1:1); **b)** distal Ub in wild-type Ub_4 (blue) and Ub_3-Ub_{R74} (red) at the endpoint of titration with Ub4a (peptide: polyUb molar ratio = 1:1); **c)** distal Ub in Ub_2-Ub_{R74} free in solution (blue) and upon addition of Ub4a (red) (peptide: polyUb molar ratio = 1:1); **d)** distal Ub in wild-type Ub_3 (blue) and Ub_2-Ub_{R74} (red) at the endpoint of titration with Ub4a (peptide: polyUb molar ratio = 1:1). Attenuated signals of the patch residues in the tetramer are indicated. Signals that have shifted dramatically in the trimer comparisons are indicated by arrows.

2.5 Structural analysis of Ub4a bound Ub_3 and Ub_2-Ub_{R72} using spin labeling NMR techniques

In leu of a crystal structure of a Ub_3:Ub4a complex, spin-labeling was utilized to further explore the specific binding mode of Ub4a peptide with trimers and C-terminus truncated tetramer. To acomplish this, our collaborator, Dr. Ashraf Brik provided Ub4a cyclic peptide material which was covalently attached to the paramagnetic spin label 2, 2, 6, 6- tetramethylpiperidine 1-oxyl (TEMPO) as described in the methods section (Fig 2.20)[39]. This allowed us to measure the effect of paramagnetic relaxation rate enhancement (PRE) on the proximal, endo and distal units of Ub_3 as well as on the distally labeled Ub_2-Ub_{R72}. This helps in making specific, distance-based assessments on how the peptide binds within the trimer and try to pinpoint which residues are more highly affected based on their proximity to the unpaired electron in the paramagnetic moiety of TEMPO. Residues in the domain that comes in close proximity (~25 Å) to the paramagnetic moiety experience signal attenuation due to PRE effects. These attenuations were calculated as the ratio of signal intensities in the presence of oxidized and reduced TEMPO. The attenuations were converted to distance constraints using in-house Matlab program SLfit[54] and the position of TEMPO's unpaired electron in 3D space was determined.

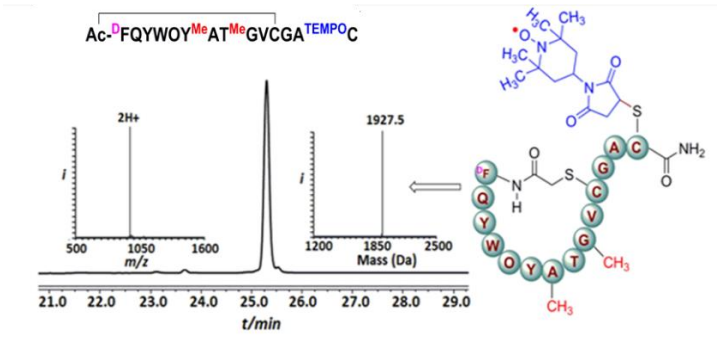


Figure 2.20: TEMPO attached on to modified Ub4a cyclic peptide[39].

2.5.1 PRE based structural analysis of Ub4a bound Ub_3

Significant attenuations in signal intensities were observed in the ^{15}N-endo unit of Ub_3 spectra, indicating the close proximity of this domain to the Cys-TEMPO on Ub4a. This effect was more pronounced in the first 3_{10} helix, the third and fourth beta-strands as well as the C-terminus of the endo unit (Fig 2.21; b). Furthermore, the results show that the proximal unit experiences some effects in the third beta-strand with minimal effect on the C-terminus while the distal unit experiences almost no effect in the main body of the domain with the observed attenuations on this unit limited to the C-terminus (Fig 2.21; a, c). The proximal and endo units of the trimer exhibited attenuations around their "canonical" hydrophobic patch residues (represented by sticks in the Ub structures). These experimentally observed profiles are in good agreement with the previous titration experiments performed and seem to confirm the domain-specific binding of Ub4a across Ub_3.

These results also give some limited insight in the orientation of the peptide as it binds to Ub_3. Given that there was no crystal structure of the peptide at this junction, we could only make speculations in terms of the binding mode of the complex. Nevertheless, by looking at the position of the spin label situated in front of the beta strands across all three units facing the hydrophobic patch, we could confidently say that the peptide was positioned within a pocket formed by all three Ub units in a trimer with the hydrophobic patch of, at least, the endo and proximal unit playing a role in stabilizing the interaction between trimer and peptide (Fig 2.21: d) which strongly suggests that Ub_3 is in a closed conformation when bound to Ub4a.

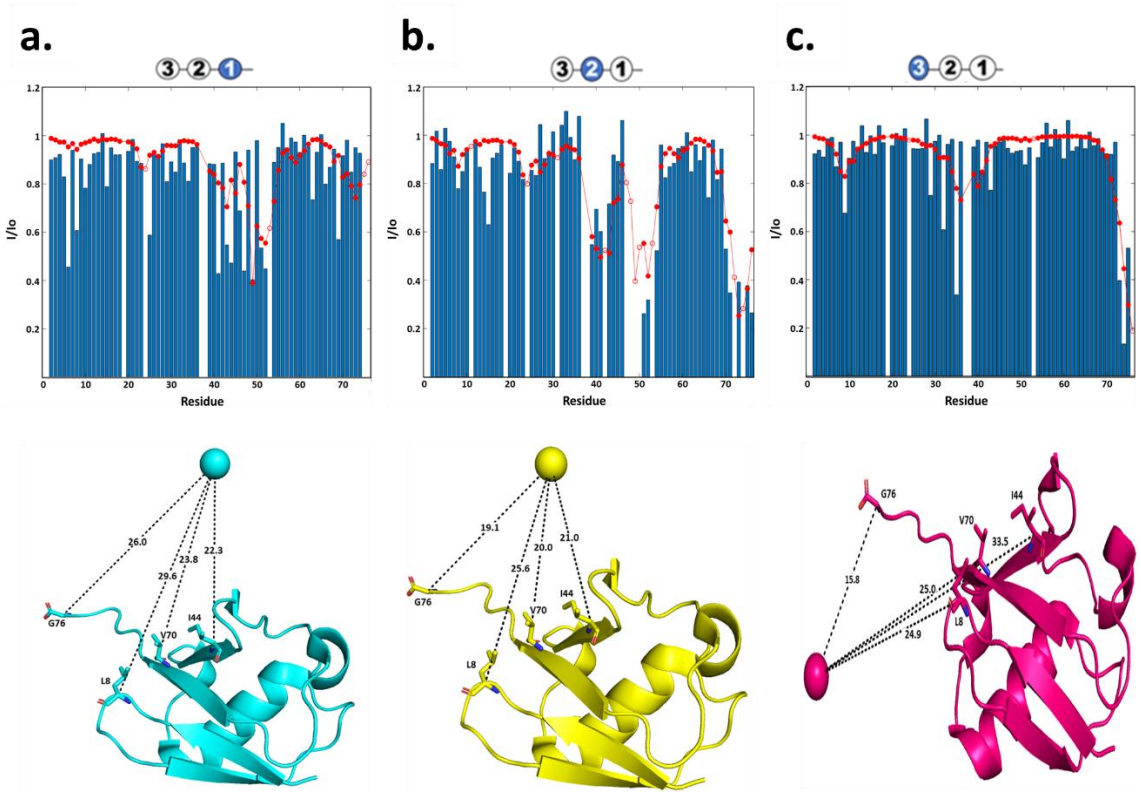


Figure 2.21: Paramagnetic spin label effect on Ub_3:Ub4a^{TEMPO} complex. Signal attenuations observed in the presence of Ub4a_TEMPO on **a)** proximal unit; **b)** endo unit and **c)** distal unit of Ub_3. Spin label for all three Ub units (cyan: proximal; yellow: endo and pink: distal) is facing the hydrophobic patch with close similarity in spin label positioning for the endo and proximal units. Distance between spin label (spheres) and hydrophobic patch residues as well as C-terminal G76 in each unit is represented (PDB: 1D3Z). Blue bars depict experimental PREs; red circles connected by lines are back-calculated PREs from the reconstructed position of the spin label. Filled circles: assigned residues used for analysis; open circles: residues not used during analysis due to lack of experimental data.

2.5.2 PRE based structural analysis of Ub4a bound Ub_3-Ub_{R72}

A similar spin label experiment performed on distally ^{15}N-labeled Ub_3-Ub_{R72} showed results that corresponded to what we would expect based on the previous NMR titration. As such, the PRE effects were observed in almost the exact areas for this unit on the tetramer as were shown for the distal unit on the trimer (Fig 2.22). This is further

corroboration that with the cleavage of most of the C-terminal residues, the binding of Ub4a shifts to the distal moiety of a tetramer thereby effectively avoiding the proximal unit in the chain. As such, it can be surmised that the C-terminus of unit 2 in the chain effectively acts as the selection mechanism used by the peptide to bind to the distal moiety instead of the proximal unit C-terminus that has the truncated residues.

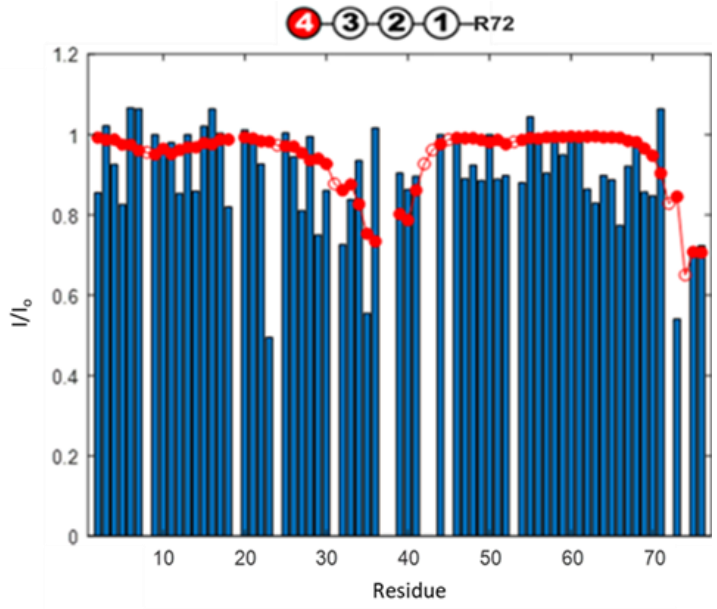


Figure 2.22: Paramagnetic spin label effect of Ub_3-Ub_{R72} bound to Ub4a^{TEMPO}. Minimal spin label effects are comparable to those of the distal trimer indicating a shift in binding mode. Blue bars depict experimental PREs for distal Ub domain in Ub_4 with partially truncated proximal tail residues; red circles connected by lines are back-calculated PREs from the reconstructed position of the spin label. Filled circles: assigned residues used for analysis; open circles: residues not used during analysis due to lack of experimental data.

2.6 Crystal Structure of Ub_3:Ub4a identifies the mechanism of interaction

To gain detailed structural insights on the binding of Ub4a to Ub trimer, the high-resolution crystal structure of Ub_3:Ub4a complex was determined at 1.85 Å resolution (Fig 2.23; a). This structure, which was solved from crystals that were grown in collaboration

with Dr. Jacek Lubkowski's lab at NCI, provided significant insights into the mechanism by which Ub trimer binds to a ligand which, up to the writing of this paper, has not been previously reported. The statistics for the data collection and refinement of the crystal complex is represented in Table 2.1.

a.

Data Collection

Parameter	Ub_3:Ub4a
Wavelength (Å)	1
Oscillation per image (˚)	0.15
Space group	*P2*
Unit cell (Å*; deg*)	83.21, 26.96, 110.47; *99.9*
Completeness (%)	93.8 (98.9)*
Redundancy	3.1 (3.2)
I/σ(I)	17.3 (2.9)
Mosaicity range (˚)	0.597 – 0.962
Unique reflections	39,731 (2070)
High resolution shell (Å)	1.85 -1.88
R-linear&	0.094 (0.414)
R-square#	0.062 (0.375)
R_{pim}@	0.061 (0.263)
$CC_{1/2}$	0.987 (0.907)

b.

Refinement

Parameter	Ub_3:Ub4a
Resolution (Å)*	1.85 – 30.0 (1.85 – 1.90)#
No. of reflections:	
refinement	37,740 (2841)
validation	1986 (133)
No. of Ub_3 in AU	6
Number of non-H atoms:	
all	4183
Ub subunits	3631
Ub4a	214
other ligands	36
water molecules	302
ADP model&	Isotropic
Average ADP (Å^2)	
all non-H atoms	20.3
Ub subunits	20.5
Ub4a	16.3
other ligands	47.7
water molecules	29.3
R-factor	0.228 (0.248)
R_{free}	0.298 (0.337)
Rmsd values:	
bond lengths (Å)	0.02
bond angles (˚)	2.11
Est.coordinate error (Å)@	0.122
Ramachandran statistics:	
preferred (No, %)	446, 99.1
allowed (No, %)	4, 0.9

Table 2.1: Statistics for Ub_3:Ub4a crystal complex. X-ray diffraction data collection (a) and structure refinement (b); values in parentheses are for the highest-resolution shell. Processing method is described in methods section.

2.6.1 General description of the Ub_3:Ub4a structure

In discussing the crystal structure results in this section, to succinctly convey the architecture of the tri-ubiquitin chain, we will follow the nomenclature in which lysine-48 linked trimer is represented as Ub_C-Ub_B-Ub_A, where Ub_C, Ub_B and Ub_A stand for the distal, middle (endo) and proximal Ub units, respectively, within the trimer chain. Specific residues within each unit are also represented by the respective subscript identifiers.

Three Ub subunits within the trimeric assembly are mutually related by ~120˚ rotation around the axis passing through the assembly's center (Fig 2.23; a). The cavity at the center of Ub_3 is occupied by Ub4a. Despite the anisotropic mosaicity of X-ray data, the final electron density maps were easily interpretable. Guided by the electron density, it was possible to model almost all residues in each Ub_3 assembly. Only two C-terminal residues (Gly76 and Asp77) in the proximal Ub unit were not modeled. The C-terminus of this subunit interacts closely with the side chain of Arg72 and a symmetry-mate of Thr12 (this is a crystal-specific interaction), leading to structural disorder. In fact, fragmented peaks of electron density suggest alternative traces for the C-terminal region of subunit Ub_A; however, they are too weak to complete the structural refinement. Importantly, all four covalent linkages between Ub subunits in two Ub_3 molecules are well defined by the electron density. Two trimeric complexes are very similar, with the r.m.s.d deviation between 1601 non-H equivalent atoms being 0.22 Å. Contacts between different Ub subunits and interactions with Ub4a are common for both complexes.

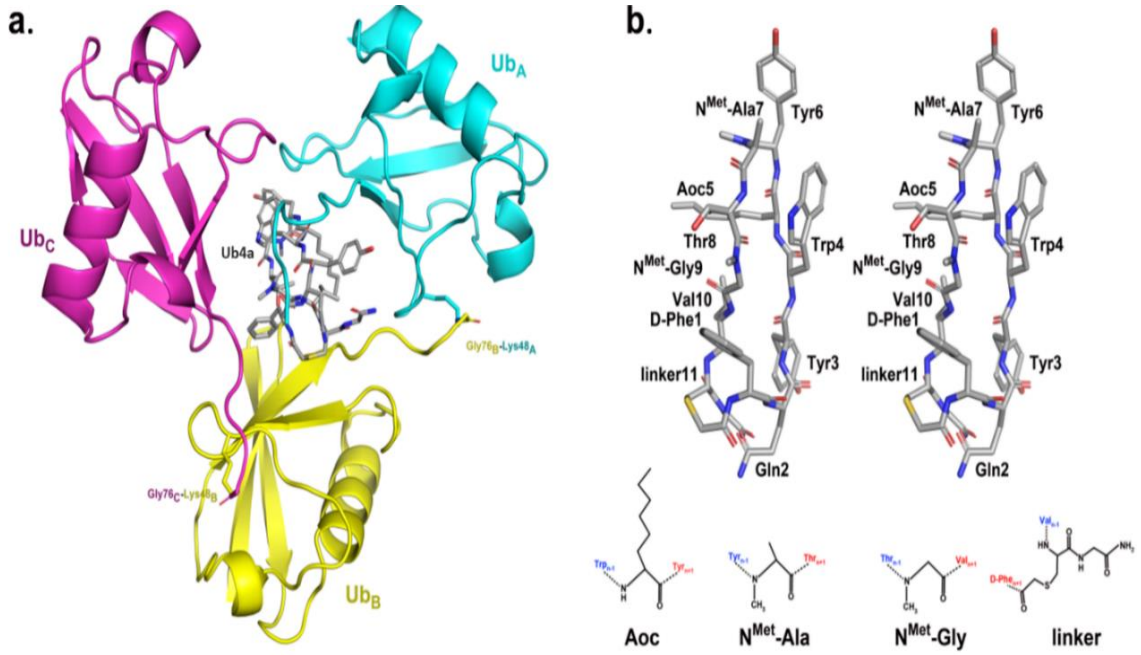


Figure 2.23: Crystal structure complex of Ub_3:Ub4a: a) Three covalently linked units of Ub (drawn in a ribbon representation, colored cyan (proximal unit-Ub_A)-magenta (distal unit-Ub_C) –yellow (endo unit-Ub_B) and annotated) are shown in complex with one molecule of Ub4a (drawn as gray sticks at the center); **b)** Stereo view of Ub4a peptide in the conformation found in the crystal structure of the complex with Ub_3. Structural depictions of non-standard blocks of the peptide are illustrated under the stereo image.

2.6.2 Interactions between different subunits within lysine-48 linked Ub_3

In addition to the isopeptide bonds that link subsequent Ub units within the trimer chain, several additional interactions within each interface define the relative positioning of adjacent subunits. These interactions are listed in detail in Table 2.2 and illustrated in Figure 2.24. It appears that contacts between different Ub subunits involve canonical hydrophobic interactions between the Ub subunits surrounding the peptide [13, 21], as well as H-bonds observed between all three Ub units.

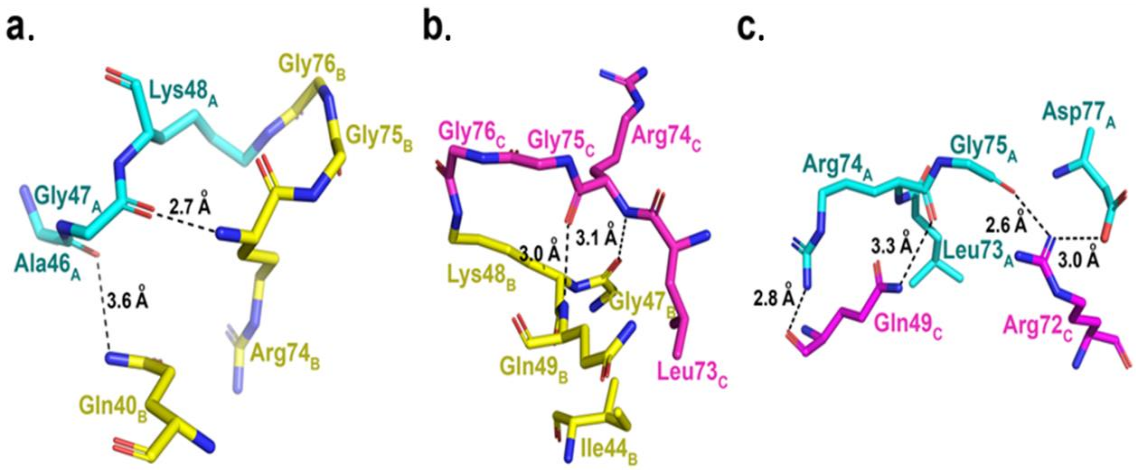


Figure 2.24: Interactions between different Ub subunits within Ub_3 crystal. In all panels, residues are shown in stick representation and Ub subunits are distinctly colored (Ub_A - cyan, Ub_B - yellow and Ub_C - magenta). **a)** Illustration of the interface between Ub_A and Ub_B; **b)** shows the Ub_B and Ub_C interface; **c)** shows the interactions between Ub_A and Ub_C. All residues illustrated in this figure are labeled. H-bonds are indicated by black dashed lines and the corresponding distances are shown.

Most extensive interactions are observed for the interface between Ub_A and Ub_C that are not covalently linked. In this case, four H-bonds stabilize relative positions of both subunits (Fig 2.24; c). It should be noted that $Asp77_A$ is not a native residue to wild-type Ub and thus, the H-bond formed between this residue and $Arg72_C$ would not be present *in vivo*. NMR experiments and DUB assays using Ub chains with $Asp77_A$ removed do not indicate a major difference in the interaction of Ub units with each other or with the peptide. Nevertheless, the existence of $Asp77_A$ in the C-terminal tail of the proximal Ub unit does in fact have an interesting interaction with the distal unit which seemingly gives the structure an additional anchor. It is also worth noting that the relative spatial arrangement of the subunits is almost certainly affected by their interactions with the cyclic peptide which is further discussed.

	Stabilizing interactions		
Interface	**Isopeptide bond**	**H-bonds**	**Hydrophobic contacts**
Ub_A---Ub_B	$Lys48_{(NZ)A}$-$Gly76_{(C)B}$	$Gly47_{(O)A}$-$Arg74_{(N)B}$ [2.7 Å]	$Ile44_{(CD1)A}$-$Leu73_{(CD2)B}$
		$Ala46_{(O)A}$-$Gln40_{(NE2)B}$ [3.6Å]	$Ile44_{(CD1)A}$-$Leu71_{(CD2)B}$
Ub_B---Ub_C	$Lys48_{(NZ)B}$-$Gly76_{(C)C}$	$Gln49_{(N)B}$-$Arg74_{(O)C}$ [3.0 Å]	$Ile44_{(CD1)B}$-$Leu73_{(CD1)C}$
		$Gly47_{(O)B}$-$Arg74_{(N)C}$ [3.1 Å]	$Leu8_{(CD2)C}$-$Ile44_{(CD2)B}$
Ub_C---Ub_A		$Arg72_{(NH1)C}$-$Asp77_{(OD2)A}$ [3.0 Å]	
		$Arg72_{(NH1)C}$-$Gly75_{(O)A}$ [2.6 Å]	
		$Gln49_{(NE2)C}$-$Arg74_{(O)A}$ [3.3 Å]	
		$Gln49_{(O)C}$-$Arg74_{(NH2)A}$ [2.8 Å]	

Table 2.2: Stabilizing interactions between the three subunits of Ub_3. Parenthesis indicates atoms that are involved in the interaction.

2.6.3 Stabilizing interactions engaging Ub4a peptide

These interactions can logically be partitioned into two categories. The first category includes stabilizing contacts within the Ub4a molecule itself. While they are naturally described here in the context of the complex, it is possible that some or many of them exist (stably or transiently) even in free, un-complexed, molecules of Ub4a. The cyclic peptide Ub4a does not appear circular in shape but rather elongated and stabilized by five intramolecular H-bonds. These include $Aoc5_{(O)}$-$_{(HN)}$Thr8, $Aoc5_{(NH)}$-$_{(O)}$Thr8, $Tyr3_{(O)}$-$_{(HN)}$Val10, $Tyr3_{(NH)}$-$_{(O)}$Val10, and $Gln2_{(OD1)}$-$_{(HN)}$(linker). Except for D-Phe1, the side chain of which is accommodating the sulfur atom of the linker, no stabilizing hydrophobic contacts within Ub4a are observed.

The second category encompasses complementary interactions between different sections of Ub4a and Ub_3 which likely provides the major stabilization of the complex.

Despite the small size of the cyclic peptide, its interactions with Ub_3 are quite extensive (surface of the interface between two entities, calculated with the PISA Server[55], is ~800 Å^2), resulting from the fact that Ub4a occupies the central cavity in Ub_3 and has comparable interactions with all three Ub subunits. The interface area can approximately be partitioned between three Ub subunits into 43% for Ub_A, 25% for Ub_B and 32% for Ub_C. The structure also shows Ub_3 with the hydrophobic patch of each unit buried within the complex forming a closed conformational interface around Ub4a. Since most Ub interactions occur through this patch[20, 56], it is not surprising to observe Ub4a bound within the pocket. Even though, stabilizing contacts consist of a variety of H-bonds and hydrophobic interactions, it appears that the latter group is defining the strength of binding (affinity). A list of specific interactions is shown in Table 2.3 with selected examples illustrated in Figure 2.25.

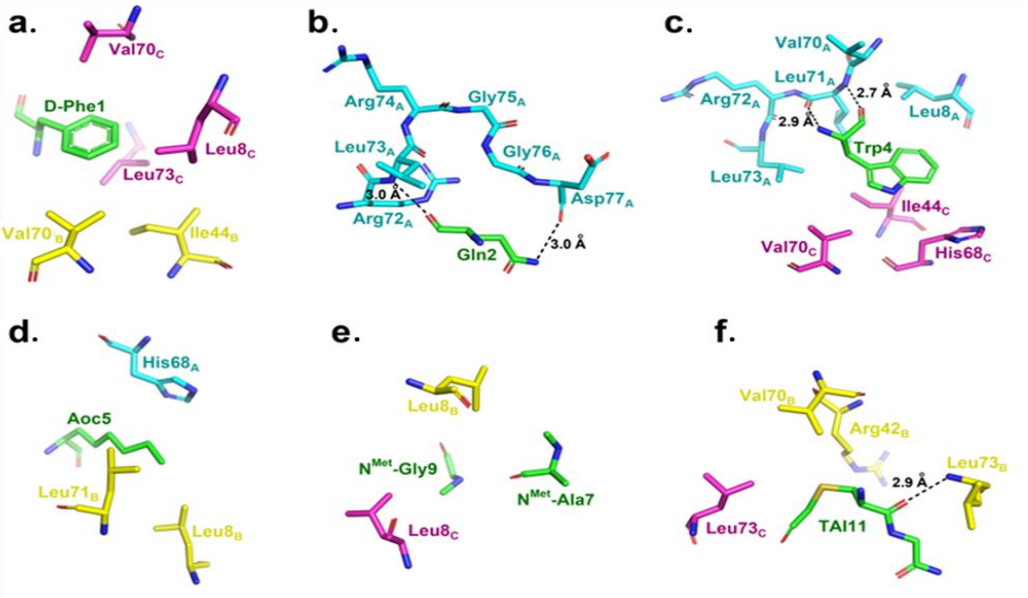


Figure 2.25: Stabilizing interactions between different sections of Ub4a and subunits of Ub_3. Residues are shown in stick representation and labeled. In all examples fragments of Ub4a are colored green and Ub_3 subunits are colored according to common scheme. H-bonds between Ub4a and Ub_3 are indicated by black dashed lines and associated distances are also shown. More extensive description of contacts is listed in Table 2.3 below.

Residue in Ub4a	Interacting residues from Ub_3
D-Phe1	O-atom in two H-bonds with guanidine group of $Arg42_C$ Phenyl ring occupying cavity formed by side chains of $Leu8_C$, $Val70_C$, $Leu73_C$, $Ile44_B$ and $Val70_B$
Gln2	O-atom forms H-bond with $Leu73_{(N)A}$ (NE2)-atom H-bonded to $Asp77_{(O)A}$
Tyr3	OH-function H-bonded with $Arg42(NH_2)_A$ and $Gln49(NE2)_A$
Trp4	N-atom H-bonded with $Leu71_{(O)A}$ O-atom H-bonded with $Leu71_{(N)A}$ Side chain in the cavity formed by $Leu8_A$, $Leu71_A$ and $Leu73_A$, $Ile44_C$, $His68_C$, $Val70_C$
Aoc5	Side chain stabilized by $His68_A$, $Leu8_B$ and $Leu71_B$
Tyr6	No obvious stabilizing interactions - protrudes from Ub_3 central cavity
N^{Met}-Ala7	N^{Met}-group interacts with $Leu8_B$
Thr8	Side chain wedged between side chains of $Leu8_B$ and $Val70_B$

Table 2.3: Stabilizing interactions between Ub4a and subunits of Ub_3. Parenthesis indicates atoms that are involved in the interaction.

2.7 DUB assays of Ub4a bound Ub chains

In order to further examine the protection Ub4a provides to Ub chains from degradation as well as test the role of the tail in the preferential binding of Ub4a to the proximal trimer moiety of a tetramer, various types of Ub4a bound tetramer and trimer complexes were incubated with different DUBs that recognize Ub-chains and cleave the connecting isopeptide bond. The DUB, OTUB1 has been introduced earlier as a lysine-48 linkage specific protease with a preference for cleaving from distal ends of longer Ub chains. On the other hand, IsoT, Ubp6 and Usp2 have been discussed as linkage non-specific DUBs with IsoT showing some preference for cleavage from the proximal end of a chain. These assays were performed on peptide bound complexes of Ub_3, Ub_2-Ub_{R72}, Ub_2-Ub_{R74} as well as Ub_4, Ub_3-Ub_{R72} and Ub_3-Ub_{R74} chains which enabled us to determine, both qualitatively and quantitatively, the extent of DUB protection offered by Ub4a binding to lysine-48 linked chains.

The NMR finding that only the first three Ub units in Ub_4 bind Ub4a tightly, was independently verified by comparing the ability of the peptide to protect Ub_4 and Ub_3 against disassembly by both linkage-specific and non-specific DUBs. The results of polyUb incubation with OTUB1 showed that Ub4a effectively shields Ub_3 from disassembly (Fig 2.26; a). This is consistent with the NMR data showing tight peptide binding to all Ub units in the trimer. Interestingly, in the presence of Ub4a, OTUB1 cleaves only one Ub molecule from Ub_4 resulting in visible and sustained accumulation of a trimer, thus supporting the conclusion that the distal Ub unit in the tetramer binds Ub4a loosely. A slightly different trend was observed for Ub4a protection against proteasome-associated DUB Ubp6, where some amount of Ub_4 was reduced to a trimer which shows that Ubp6 cleavage activity can occur from the distal end but a significant portion of Ub_4 remained unaffected, likely reflecting Ubp6's reduced activity when not bound to the proteasome[57]. By contrast, Ub4a effectively blocked Ub_4 disassembly by a general linkage-nonspecific DUB, Usp2 (Fig 2.26; b, c). A secondary possible reason for the reduced activity observed for Ubp6 (as well as Usp2) could be related to the fact that this DUB appears to show some preference to endo-isopeptide bonds and thus binding to two Ub units becomes important for efficient cleavage[58]. As such, the binding of Ub4a could be preventing these DUBs from fully accessing a secondary Ub unit and thus reducing their overall efficiency.

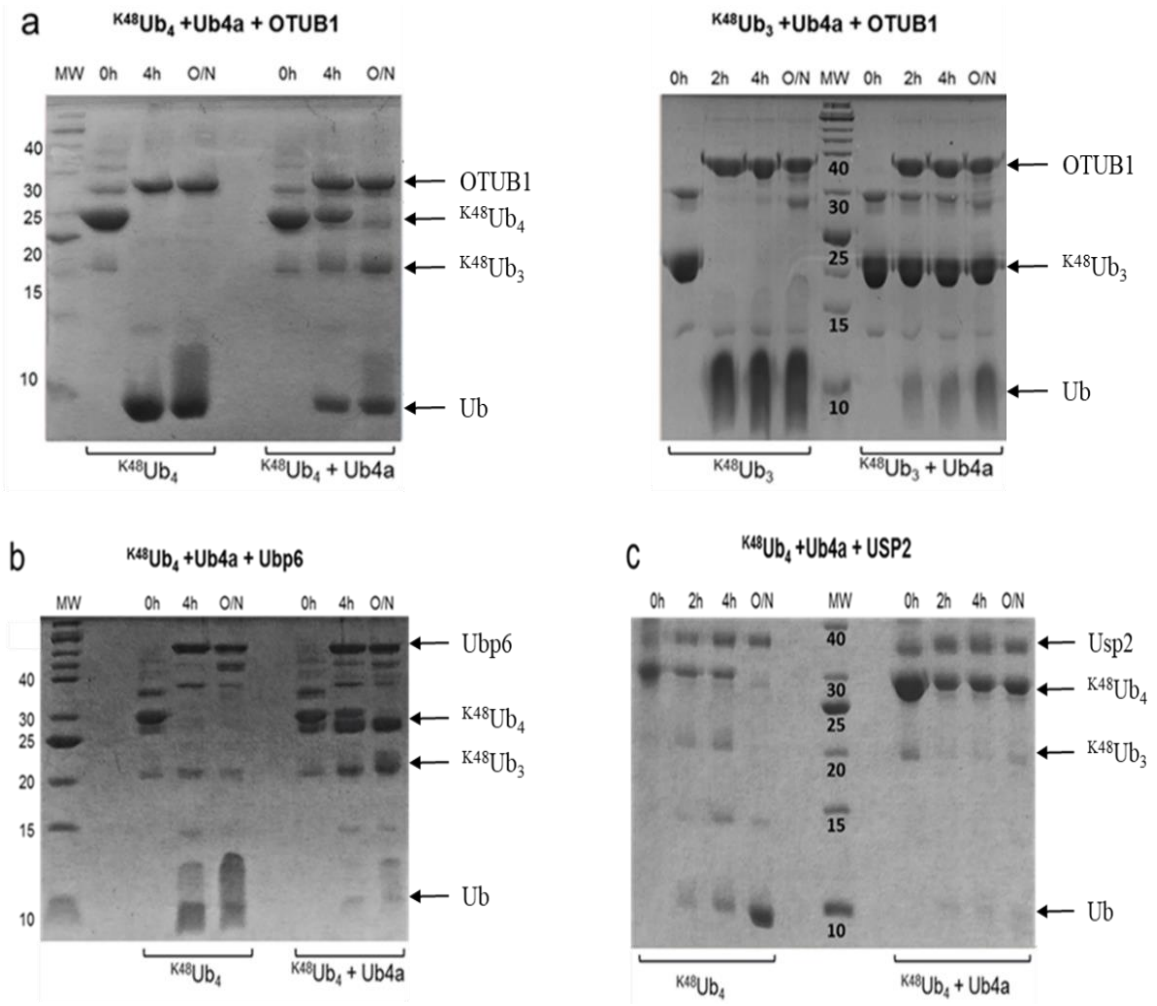


Figure 2.26: Cyclic peptide Ub4a protects lysine-48 linked Ub-chains from disassembly by deubiquitinases. Time course disassembly of **a)** Ub_4 (left) and Ub_3 (right) by lysine-48 linkage specific OTUB1; **b**) Ub_4 by the proteasome-associated DUB Ubp6 and **c)** Ub_4 by linkage-nonspecific DUB Usp2 in the absence and presence of Ub4a. Corresponding molecular weight values are shown in kDa. 0hr in panel c shows Usp2 band since gel samples at initial point were taken immediately after addition of Usp2. MW = molecular markers, O/N = overnight.

On the other hand, IsoT assays showed that trimer chains that bind Ub4a were fully protected from proximal cleavage which, as expected, was also true for Ub4a bound tetramer chains (Fig 2.27). This seems to corroborate that Ub4a binding to the proximal moiety of lysine-48 linked chains prevents access of IsoT to the chain from the proximal end. In fact, given the initial hypothesis of Ub4a preferring the proximal trimer moiety and

interaction with the tail residues, that fact that ubiquitin pentamer chains are protected from IsoT cleavage in the presence of Ub4a strongly corroborates that the tail residues and proximal Ub unit within the chain are prevented from being recognized by IsoT DUB (Fig 2.27; a).

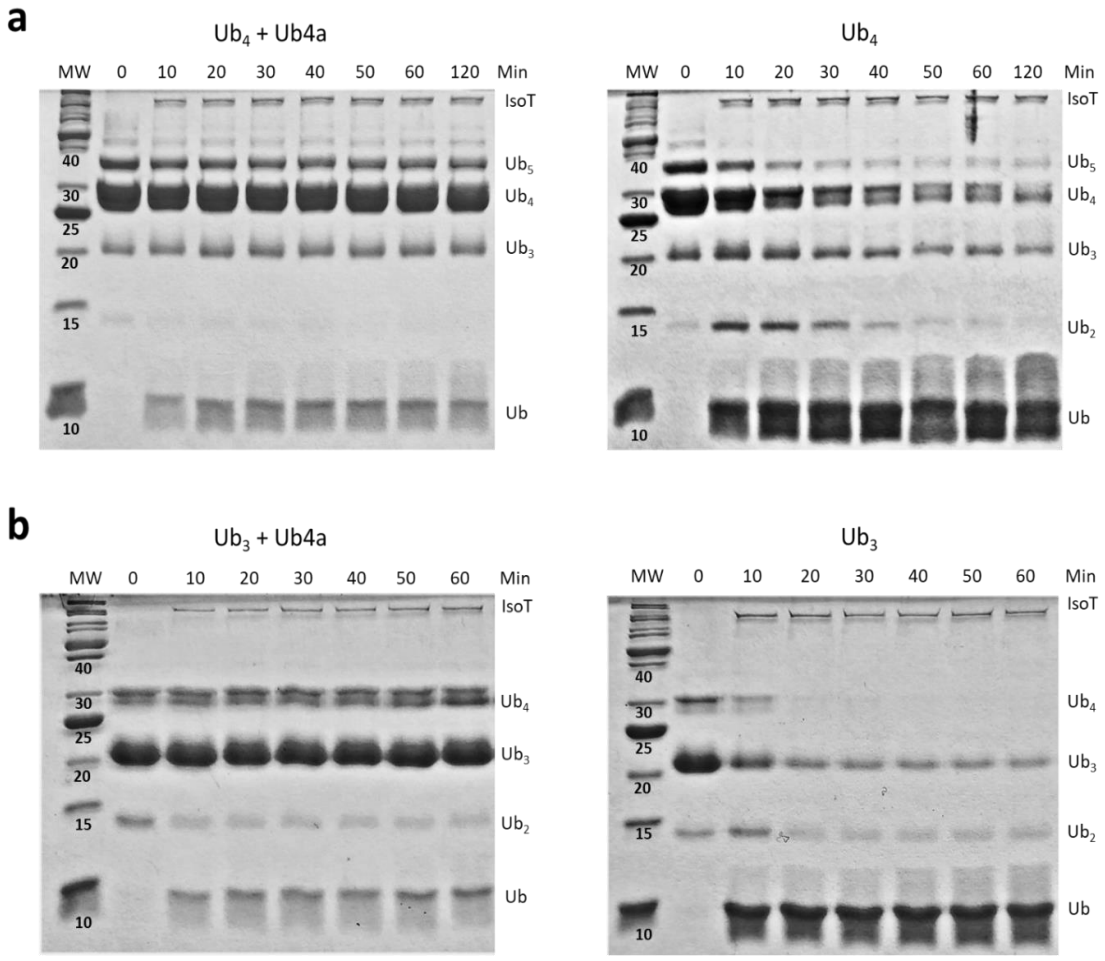


Figure 2.27: Cyclic peptide Ub4a protects lysine-48 linked Ub-chains from disassembly by IsoT. Time course disassembly of **a)** Ub_4 + Ub4a (left) and Ub_4 (right) by IsoT; **b**) Ub_3 + Ub4a (left) and Ub_3 IsoT. Comparison between peptide bound and unbound assays clearly shows the difference in the deubiquitinase activity of IsoT. Corresponding molecular weight values are shown in kDa. MW = molecular markers.

Furthermore, deubiquitinase assay was conducted on Ub_3-Ub_{R72} chains in order to investigate whether the NMR results observed for this tetramer would translate to a situation where it would be more protected from DUB assays because of the apparent preferential binding of the peptide to the distal moieties (units 2-4). As such, given that the cleavage activity of OTUB1 occurs mainly from the distal end of Ub chains, it was chosen for this investigation. Surprisingly, the results indicated that binding of Ub4a to this tetramer with truncated C-terminal tail residues did not prevent OTUB1 from fully acting on the chain. In fact, repeated experiments showed that within 2 hrs, almost all the Ub tetramer was consumed (Fig 2.28) which was starkly different from how OTUB1 acted on wild-type tetramer and trimer chains that were bound to Ub4a (Fig 2.26; a).

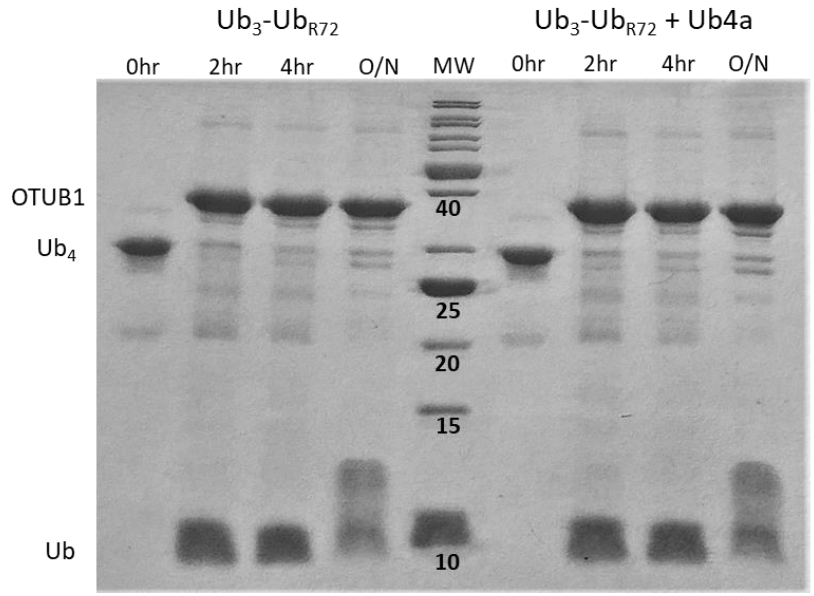


Figure 2.28: Cyclic peptide Ub4a does not protect Ub_3-Ub_{R72}-chains from deubiquitinase disassembly. Gel assays of time course disassembly assay of Ub_3-Ub_{R72} (left) and Ub_3-Ub_{R72} with Ub4a peptide (right) by OTUB1.

These seemingly contradicting results between NMR and DUB assays raised more questions on why the Ub_3-Ub_{R72} tetramer was not protected from DUB activity despite being apparently bound to the cyclic peptide in a similar manner to that of a trimer moiety even if the positions of binding might have shifted to the distal trimer end. Nevertheless, in order to assess how OTUB1 would act upon a trimer with similar C-terminal cleavage, a reaction assay with Ub_2-Ub_{R72} was conducted. The result from this assay was consistent with the respective NMR results in that there was no protection provided by Ub4a against OTUB1 cleavage (Fig 2.29). This suggested that despite the NMR results that definitely show binding was shifted to the distal moiety of the tetramer, the strength of binding is apparently not strong enough to provide competitive advantage against the activities of OTUB1 even in the truncated tetramer chain.

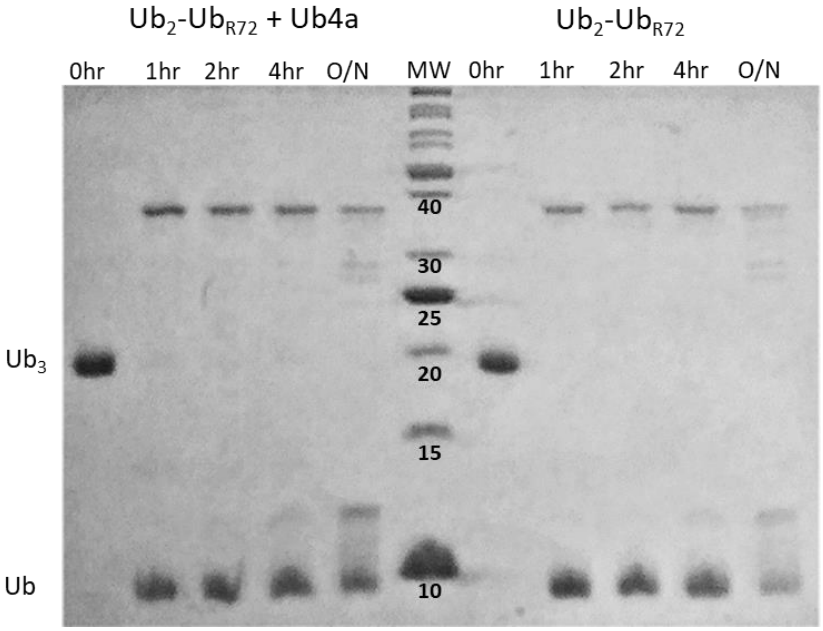


Figure 2.29: Cyclic peptide Ub4a does not protect Ub_2-Ub_{R72}-chains from deubiquitinase disassembly. Gel assays of time course disassembly assay of Ub_2-Ub_{R72} with Ub4a peptide (left) and Ub_2-Ub_{72} (right) by OTUB1.

To confirm the results from NMR on the chains with Ub_{R74} truncation, similar DUB assays were conducted on both Ub_3-Ub_{R74} and Ub_2-Ub_{R74}. The ensuing assays where further validation on the hypothesis that only the truncation of the full C-terminal tail in the proximal unit prevents the strong binding of the peptide to the Ub chains (Fig 2.30)

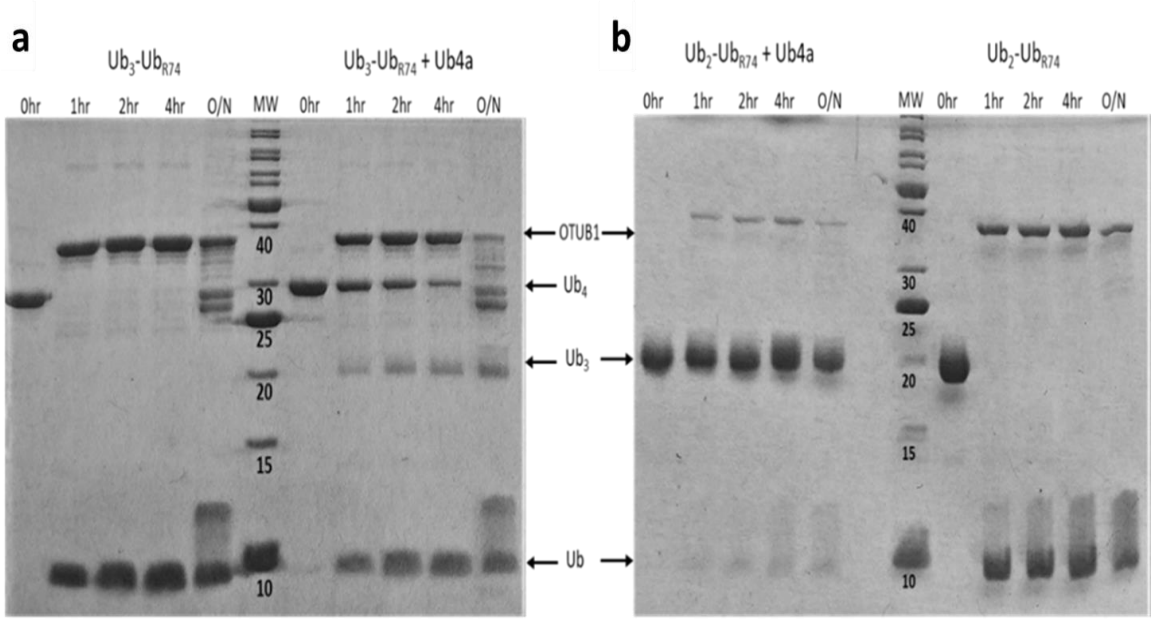


Figure 2.30: Extent of C-terminal tail cleavage has an effect on peptide protection of chains from DUB activity. Assay of time course disassembly of **a)** Ub_3-Ub_{R74} (left) and Ub_3-Ub_{R74} with Ub4a peptide (right); **b)** Ub_2-Ub_{R74} with Ub4a peptide (left) and Ub_2-Ub_{R74} (right) by OTUB1.

2.8 Ub4a presence in Ub monomer pool does not inhibit enzymatic assembly

The end goal of peptide binding to Ub chains is in its potential to inhibit proteasomal activity. Unlike *in vitro* assays where binding to chains is artificially directed, Ub4a interacts with varying lengths of Ub during chain elongation process *in vivo*. As such, the question arises whether Ub4a binding to different stages during Ub chain elongation

process i.e., Ub monomer or lysine-48 linked dimer, would inhibit said elongation process resulting in the reduction of longer chains that are needed for efficient recognition by the proteasome. In order to test this, *in vitro* enzymatic chain elongation assays in the presence of Ub4a were performed.

To determine the effect of Ub4a binding to Ub monomer at the beginning of polyubiquitin chain assembly, wild-type ubiquitin was incubated with Ub4a followed by an enzymatic conjugation reaction with time point samples taken for gel visualization followed by quantification of resulting bands with Image J[59]. Compared to control reaction, the results of this assay showed that conjugation is not inhibited due to peptide binding up until a trimer length was reached after which elongation to tetramer chain was sufficiently slowed down (Fig 2.31; a). This result was additional corroboration on the chain length selectivity of Ub4a and further indication that Ub_3 is the main binding partner of the cyclic peptide within a tetramer chain. A similar experiment was performed with the starting material being a lysine-48 linked dimer incubated with Ub4a. To this dimer, Ub_{R72} was enzymatically added in order to assess if a trimer formation is inhibited. As expected, the result showed that Ub4a binding to dimer does not, in fact, inhibit elongation of the chain (Fig A1)**.**

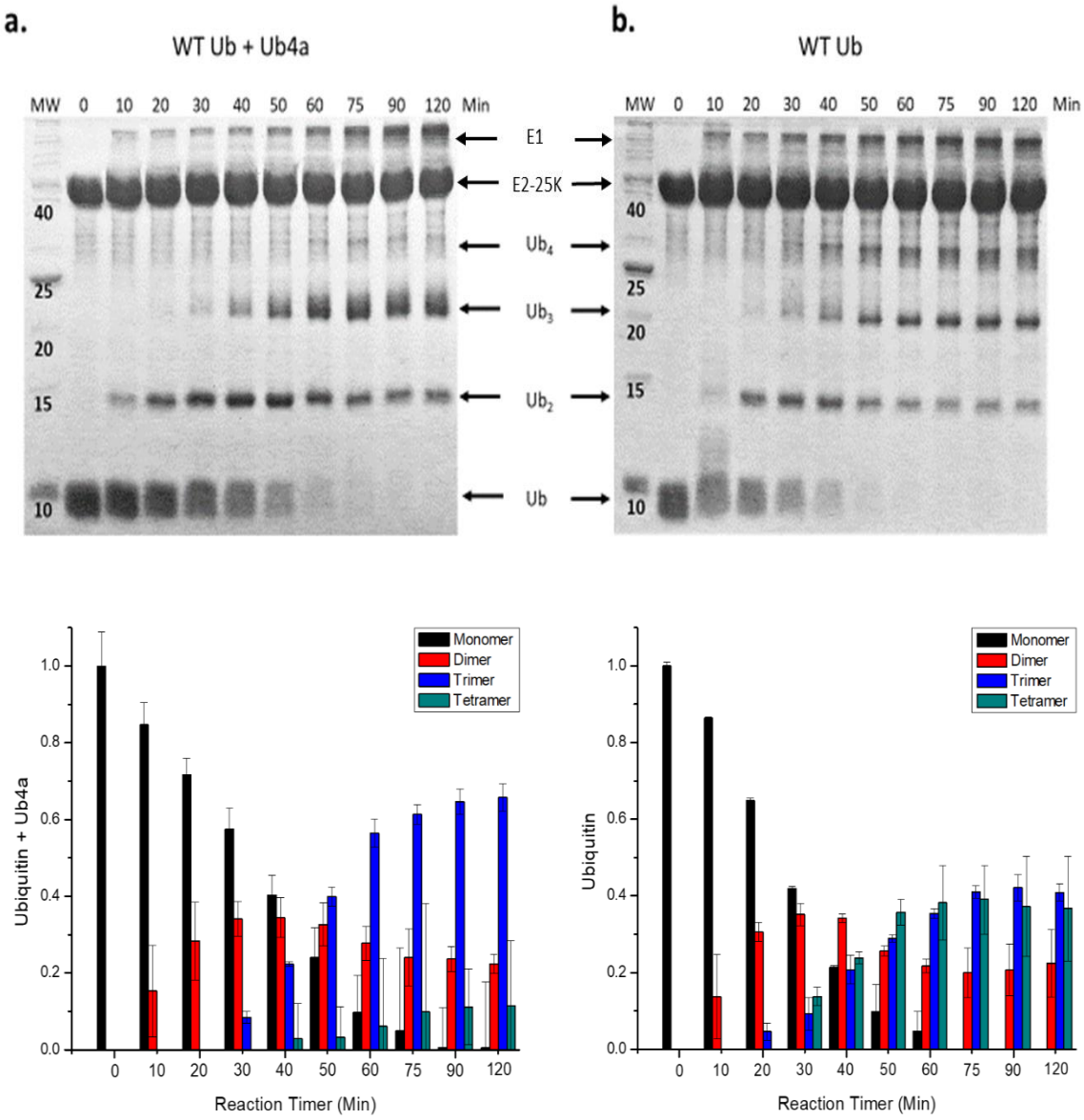


Figure 2.31: Ub4a binding of lower order chains does not inhibit initial chain elongation reaction. Assay of time course assembly of wild type Ub monomer with Ub4a peptide (left) and control without Ub4a peptide (right). Compare the trimer band for both assays. Bottom graphs show the quantified results for the gel bands using Image J software. The difference in formation of tetramer chains between the two assays is clearly observed.

2.9 Comparison of PRE based results and crystal structure of Ub_3:Ub4a

Once the crystal structure of the complex was solved, comparison with the previous PRE data that was obtained was used in order to further characterize the structure of Ub_3:Ub4a complex in solution. As such, in addition to distance information obtained between TEMPO bound peptide and individual Ub (1D3Z) units, further PRE-measurements were conducted with the newly solved trimer crystal structure utilized as a model structure.

Once SLfit was used to identify the position of the spin label, in order to map out and compare the results with the crystal structure, TEMPO was manually built into the crystal structure of Ub4a from the solved complex available in PyMOL, and the subsequent results from the SLfit measurements were superimposed onto the crystal structure obtained. Comparison of TEMPO positions and subsequent PRE effects were in good agreement with the crystal structure allowing for some small variation due to the allowed flexibility of the tag in solution. This result indicated that the crystal structure faithfully reflected the Ub_3:Ub4a structure in solution (Fig 2.32).

Interestingly, by overlaying the SLfit results from 1D3Z used as model structure onto the crystal structure and comparing this to the overlay results of the trimer SLfit results showed a very similar positioning of TEMPO demonstrating that the initial PRE results were quite useful in giving some information on the arrangement of the peptide when bound to Ub_3 (Fig 2.32, a; b). The positioning of spin label for the distal unit showed a greater variation in distance from the position of TEMPO which was expected given the minimal PRE effects experienced by this unit and thus the potential for errors in the SLfit measurements for this domain. Nevertheless, the position of spin label from the perspective

of the distal unit when using either the Ub_3-Ub4a complex or ubiquitin monomer as the model structure was consistent (Fig 2.32). In general, this reflects the limitations of using PRE based measurements to determine position of spin label or obtained information about structural conformation when the effective distances between the spin label and the affected residues is too large. In terms of the slight differences observed between the position of the spin label from the perspective of both the proximal and endo domains, this can be explained by taking into consideration that TEMPO attached to Ub4a has conformational flexibility which is reflected by the consistency observed in the positioning of the spin labels when using either the Ub_3-Ub4a crystal complex or ubiquitin monomer as the model structure for SLfit analysis. Overall, these results suggest that the crystal structure faithfully reflects the Ub_3:Ub4a interactions in solution.

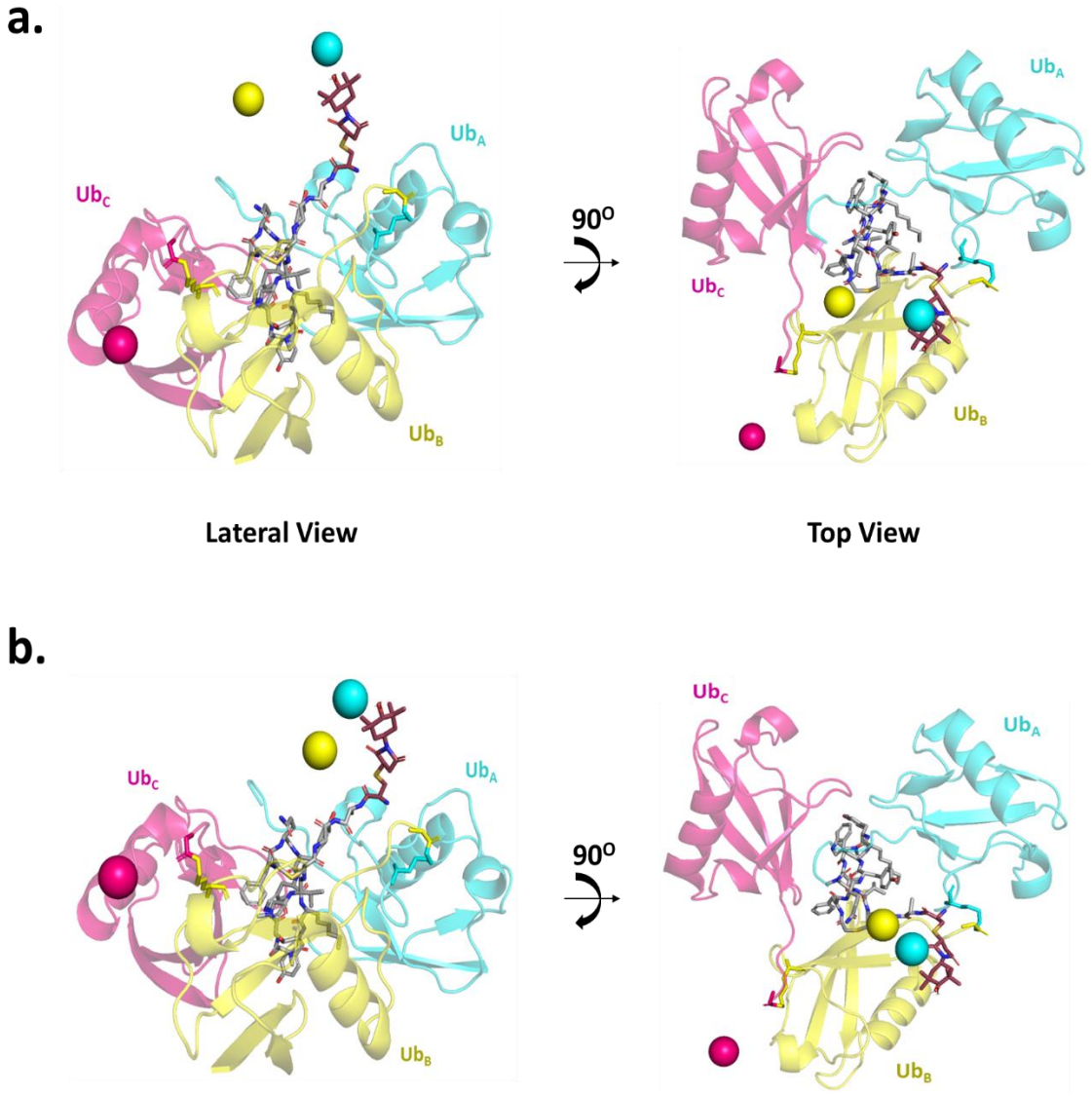


Figure 2.32: Crystal structure results compared to SLfit based positioning of paramagnetic spin label TEMPO (spheres) attached to Ub4a. a) Position of spin label from SLfit results using Ub_3-Ub4a complex aligned with Ub_3-Ub4a crystal complex with TEMPO built onto Ub4a; **b)** Position of spin label from SLfit results using individual Ub monomers (1D3Z) aligned with Ub_3-Ub4a crystal complex with TEMPO built onto Ub4a. Spheres: spin label position from SLfit (Pink: distal Ub; Yellow: endo Ub; Cyan: proximal Ub). TEMPO was manually built in the crystal structure in PyMOL.

2.10 Discussion and Summary

Due to the important role played by Ub chains as signaling molecules for proteasomal degradation, lysine-48 linked Ub chains are potential targets for therapeutic intervention in diseases that hijack the ubiquitin-proteasomal system in their progression. Previous efforts had described the discovery of proteinogenic and non-proteinogenic cyclic peptides as agents that specifically bind to longer (≥3) lysine-48 linked Ub chains involved in proteasomal protein degradation[31, 39]. The studies presented in this chapter showed that these peptides were shown to mainly interact with trimer units within a longer polyUb chain even though they were originally selected to preferentially bind to tetramer chains over dimer chains. The mechanism by which the cyclic peptide Ub4ix and more specifically, the subsequent non-proteinogenic cyclic peptide Ub4a, selects for and binds to trimer moieties within Ub tetramer chains and the extent by which the C-terminal tail determines the specific binding to proximal trimer moiety with regards to Ub4a peptide was reported. Further studies on the protection from DUB activity that arises from said binding and the overall strength of the interaction was also discussed.

For both peptide studies here, trimer chains seem to be the favored binding partners. Interestingly, the NMR results for Ub4ix binding to Ub tetramers indicated that both the distal trimer moiety and the proximal trimer moiety could be selected for peptide binding. This was further corroborated by DUB assays which show Ub4ix bound tetramer being protected from OTUB1 cleavage and partially being protected from IsoT cleavage. As indicated earlier, since OTUB1 and IsoT prefer to cleave from the distal and proximal ends of Ub chains, respectively, the results suggest that Ub4ix can bind to both the distal and proximal trimer moieties interchangeably. The general DUB UBp6, on the other hand, was

able to cleave Ub4ix bound tetramer chains much more than either OTUB1 or IsoT. This suggests that since Ubp6 does not have a preference on which end it starts to cleave from, there is a cumulative effect of cleavage activity on the tetramer from either end which might explain the pronounced cleavage that occurs with this DUB. In either case, a unit of Ub is cleaved with a trimer remaining as the main binding moiety for Ub4ix. Still, the NMR results seem to indicate that the proximal unit has the least amount of signal perturbations suggesting that the distal trimer moiety has the larger share of Ub4ix binding.

In the case of Ub4a peptide, in addition to NMR results that showed the specific preference of Ub4a binding to the proximal trimer moiety, DUB assays further demonstrated that Ub4a binding protects the proximal trimer moiety in lysine-48 linked tetramer chains with the distal Ub unit being cleaved off when subjected to distal end specific DUB's. This was also corroborated with the activity of IsoT which, even though it has a much lower K_d (K_d < 2 nM)[60] than Ub4a for Ub chain binding, was not able to cleave either Ub tetramer or trimer chains in the presence of the peptide due to its inability to access the proximal tail residues within the aforementioned chains. Further assays with the general DUB Usp2 also did not exhibit any cleavage activity on peptide bound chains.

On the other hand, DUB assays on Ub chains with truncation of C-terminal residues showed a significant difference compared to the wild-type chains. In all cases where the tail residues ended at R72, Ub4a binding did not offer any protection from cleavage but chains with C-terminus ending at residue R74 had some protection from DUB activity for tetramers and full protection for trimers giving important insight on the role played by the C-terminal residues in Ub4a binding. Our NMR binding studies with Ub_3-Ub_{R72} revealed a shift in binding mode from proximal trimer moiety to distal trimer moiety while similar

binding studies with Ub_3-Ub_{R74} showed spectra that was similar to wild-type Ub tetramer. Even though the CSP results in Ub_3-Ub_{R72} binding to peptide suggest a comparable binding to that of wild-type trimer, it is still not fully understood why OTUB1 DUB cleavage is not inhibited, as was expected. It is possible that the shifting of Ub4a binding to distal trimer moiety and absence of free tail residues reduces overall peptide binding affinity enough to allow competitive binding with DUBs. In any case, these observations signify which tail residues are important in the binding of Ub4a to lysine-48 linked chains with length ≥3 units and suggest that glycine-75 and glycine-76 might not play as much of a major role in binding selectivity. This seems to be corroborated by the crystal structure results that do not show a role played by these two residues in directly interacting with Ub4a apart from the somewhat artificial interaction between residue D77 and Q2 of Ub4a (since wild-type Ub sequence ends at G76). Nevertheless, the closeness of G76 in the proximal unit to Q2 in Ub4a (as observed in the crystal structure) indicates a possibility of some interaction represented by the change in chemical environment of the signal in the NMR spectra. In addition, these tail residues are important in trimer conformation and stability through interactions with the distal unit where $G75_A$ makes two H-bond with $R72_C$. These interactions might explain the slight difference in spectra between wild-type chains and R74 chains as well as the results of the DUB assays suggesting that the trimer pocket where Ub4a sits in is not as stable if G75 and G76 in the proximal units are truncated.

Fortunately, the perturbations observed in the C-terminal region of Ub_A during NMR experiments seem to be explained by several interactions that are illustrated in the crystal structure. The polar interactions that exist between $G75_A$ and $R72_C$ as well as those between $R74_A$ and $Q49_C$ appear to close the “loop” and hold the cyclic peptide within the

trimer pocket. Further interactions between the peptide and C-terminal tail of Ub_A show how the tail region is important for the binding process. This can, to a certain degree, explain the specificity observed in trimer binding since a flexible tail seems to be required to make the necessary polar interactions with the peptide.

Further extrapolation can be made on how the partial truncation of the tail region can affect peptide binding. Given that the H-bond interactions of peptide with the C-terminal tail of Ub_A ends with L73 making an H-bond with Q2 of Ub4a (Fig 2.25; b), it therefore follows that truncation of residues that interrupt this interaction will have a bigger impact on the binding affinity that exists. This is shown by the results where truncations that included $L73_A$ and $R74_A$ made a difference in the preference of binding for both trimer and tetramer chains. Furthermore, even if $R74_A$ does not make contacts with the peptide, the fact that it makes two H-bonds with $Q49_C$ could play a role in the anchoring of the peptide within the trimer pocket and thus acting as a contributing factor in the binding selection for the proximal trimer moiety. As such, upon truncation of $L73_A$ and $R74_A$, it appears that $L73_B$ acts as $_a$ substitute binding partner for Q2 in Ub4a explaining the shift in binding mode in both tetramer and trimer chains. Nevertheless, the DUB assay results indicate that this shift does not create the same tight binding that was observed either for the wild-type chains or those with partial C-terminal tail residue truncations.

Crystal structure results of the Ub_3:Ub4a complex shows how binding of Ub4a to lysine-48 linked chains occurs. Spatial arrangement of the Ub units observed in the structure as well as the interaction of the peptide to the C-terminal tail of the proximal Ub unit within the trimer shows how the proximal trimer moiety in a lysine-48 tetramer chain is preferentially selected for binding with individual Ub units interacting amongst each

other to incorporate the peptide within a central pocket. This process involves the canonical hydrophobic patch of each Ub unit which further magnifies the importance of this patch in the interactions Ub has with various substrates. These crystal structure results appear to be corroborated by SLfit results which shows that the positioning of TEMPO bound peptide in solution is very similar to what is depicted in the crystal structure. Furthermore, the combination of hydrophobic and polar interactions that the peptide has with the trimer gives insight into future designs of cyclic-peptides that strike a balance in terms of incorporating non-proteinogenic amino acids along with H-bond forming residues to maximize target binding efficiency.

Given the various results that were discussed regarding the cyclic peptides, an investigation on using Ub4a to induce apoptosis in cancer cell lines was undertaken by our collaborators. As such, several cancer-derived cell lines, namely U87 (brain cancer), SH-SY5Y cells (neuroblastoma), MDA-MB-231 cells (epithelial human breast cancer) and HeLa cells (cervical cancer) were incubated with Ub4a with the results showing a clear increase in apoptosis after 24 hr treatment with the cyclic peptide. These results were similar to the effect that was observed when direct proteasome inhibitors: bortezomib, ixazomib, and carfilzomib were utilized[39, 61]. Furthermore, comparison of the results from Ub4a treatment of these cancer cells with similar experiments that were performed using Ub4ix clearly showed that Ub4a was able to induce apoptosis in a greater proportion than Ub4ix[39].

Our collaborators followed the above experiments with further *in vivo* studies that involved the treatment of mice that were injected with human CAG myeloma cells with Ub4a in order to investigate its efficacy in inhibiting the growth of tumors. These

experiments showed a significant reduction in tumor progression similar to what was observed when mice were treated with the proteasomal inhibitor bortezomib indicating that Ub4a does interfere in the proteasomal degradation process within the cell[39].

Here, our *in vitro* chain elongation assay has shown that enzymatic extension of lysine-48 linked Ub's is hampered after reaching a trimer length. Given the physiological processes that follow similar routes, it appears that a possible method by which Ub4a affects proteasomal degradation of oncogenic proteins is by binding to lysine-48 linked trimer chains thereby diminishing the availability of lysine-48 linked tetramer chains necessary for degradative signaling. In cases where tetramer is in fact synthesized, Ub4a can potentially bind to the proximal trimer moiety of this chain with enough binding affinity to prevent binding to other substrates as well as reduce the activity of DUBs which leads to the decrease in the pool of available free Ub monomers. As such, we contend that Ub4a binding provides a cumulative signal inhibition effect in its interaction with lysine-48 linked Ub chains making it a promising avenue for cancer therapy.

Chapter 3: Ubiquitin-like proteins

3.1 Introduction

Ubiquitin is not the only post-translational modifier protein in eukaryotic cells. Since its discovery, a class of other proteins that share similarities at the structural and residue level have been identified. Most of these proteins known as ubiquitin-like proteins (UBLs) also utilize a similar enzymatic cascade (E1-E2-E3) resulting in post-translational modification of substrate proteins[38]. Some of the well-known UBLs include proteins such as Rub1 (Related to ubiquitin 1, Nedd8 in mammals), the SUMO (Small Ubiquitin-like Modifier) family of UBLs[62], APG8 and APG12 (Autophagy related 8 and 12)[63]. Despite the similarities between all these UBLs, the uniqueness of ubiquitin lies in the role it plays in regulating so many cellular processes as well as its ability to make chains through its various lysines in order accomplish these varying tasks. On the other hand, most of the other UBL attachments are transient and seem to have evolved from prokaryotic systems[64]. Currently, along with the aforementioned Ubls, including ubiquitin, there are about ten distinct UBLs that are able to covalently modify other macromolecules.[64]

One of the interesting aspects of the presence of these many proteins that have similar types of roles with similar enzymatic cascades is the ability of each cascade to select for the exact type of UBLs necessary for specific modifications. As far as we know, the only overlap in the recognition of UBLs by their respective E1 and E2 enzymes is relegated to a single E1 being shared by Atg8 and Atg12 which are both autophagy related proteins[65]. As such, trying to understand the molecular mechanisms by which ubiquitin is distinctively recognized by its conjugation mechanism could potentially lead to a better understanding of the selectivity that exists in various UBL related processes. This understanding of UBL

selectivity could shed more light in the discovery of therapeutics that can be designed for specific purposes with various UBL systems as was demonstrated in the previous chapter.

In order to explore this potential, a model UBL was chosen as protein of interest for study. This model protein named Rub1 shares so many similarities with ubiquitin that it was determined to be the best protein to test its interaction with the ubiquitin conjugation system. In this chapter, the research conducted in assessing Rub1's potential interaction with ubiquitin and its conjugation system is discussed.

3.2 Related to ubiquitin 1 (Rub1)

Rub1 is a yeast protein with a mammalian orthologue Nedd8 (neural-precursor-cell-expressed developmentally down-regulated 8) which, among the known UBLs, is the closest relative to ubiquitin. It has 76 amino acids that have about 53% identity and close to 77% similarity to ubiquitin with the three-dimensional structure comparison of these proteins showing a high-level of similarity (Fig 3.1)[66-68]. In this study, even though Nedd8 could be considered as the natural UBL choice to compare with ubiquitin within the context of a mammalian system, the complications in the practical aspects of Nedd8 growth and purification (especially the mutant versions) made utilizing Rub1 a reasonable compromise. This is further cemented by the fact that ubiquitin is found in all eukaryotic cells with only three differences in the amino acid sequences between mammalian and yeast ubiquitin.

Most of the amino acid differences between Rub1 and ubiquitin map to the surface of the structure while the hydrophobic patch on Rub1 remains unchanged from that of ubiquitin. Given the similarities that exist between these two proteins, it was initially

thought that Rub1 could be involved in regulation and/or degradation of some cellular proteins as was discovered for ubiquitin. Nevertheless, the apparent similarities did not translate into Rub1's post-translational modifying activities[69, 70]. It should be noted that much of the work discussed in this chapter almost exclusively involved Rub1 instead of its mammalian counterpart, Nedd8. As such, unless stated otherwise, the experiments performed, and results obtained in this chapter deal solely with Rub1.

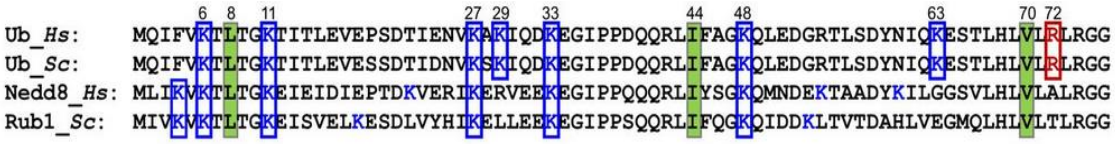


Figure 3.1: Sequence alignments for Ub, Nedd8 and Rub1from *Saccharomyces cerevisiae* (Sc) and human (Hs). Blue boxes show lysine residues conserved in all three proteins with the shared hydrophobic patch residues (L8, I44, and V70) shown in green shaded boxes. The red box indicates the positively charged arginine residue that is found only in Ub[71].

Like Ub, the function of Rub1 involves being attached to substrates in order to perform mainly regulatory tasks[72]. Rub1 is conjugated as a monomer to its most well-known post translational modification substrate CDC53/cullin protein which is a subunit of the SCF (Skp, Cullin, F-box containing complex) E3 ubiquitin ligase. This complex catalyzes the ubiquitination of proteins destined for 26S proteasomal degradation[73] and along with the anaphase-promoting complex[74] SCF has an important role in the ubiquitination of proteins involved in the cell cycle. Studies have shown that Rub1 conjugation is necessary for the functional assembly of this ligase and thereafter the continuation of the ubiquitination process necessary for various regulatory functions[69, 70, 75, 76]. Relatively recent developments have shown that Rub1(Nedd8) can

modify other various substrates such as p53[77], Mdm2[78, 79], p73[80],epidermal growth factor receptor (EGFR)[81] as well as parkin which is an E3 ubiquitin ligase[82].

Similar to ubiquitin, conjugation of Rub1 to substrates is mediated by the consecutive action of an E1 activating enzyme, UBA3-ULA1 heterodimer in *S. cerevisiae* (NEDD8 activating enzyme; NAE for Nedd8), and its own E2 conjugating enzyme (Ubc12) followed by ligase activity by E3 enzymes which have been shown to be shared by both Ub and Rub1[67, 69]. Furthermore, it has been well documented that during cellular stress, levels of Rubylation seem to increase and the occurrence of Rubylation and ubiquitination on the same target has been reported[83, 84]. Given the low level of substrate modification by Rub1 in these conditions, the actual mode and effect of the conjugation is still unknown. Recently, *in vitro* studies, including those discussed in this dissertation have demonstrated cases were Rub1 can be attached to ubiquitin chains and act as a cap inhibiting any further extension of the polyUb chain thereby limiting the amount of ubiquitin that could be extended. This can be an explanation for why, "atypical Rubylation/Neddylation" of non-cullin targets upon either overexpression of Rub1/NEDD8 or during stress related depletion of free Ub has been known to occur[68, 83, 85].

The respective activating enzymes for Ub and Rub1 have repeatedly been shown to distinguish between the two proteins despite their structural similarities. UBA3-ULA1 has an arginine side chain (R190) that sterically clashes with an arginine residue (R72) at the carboxyl tail of Ub (Fig 3.1, red box) which hinders any possible activation of Ub by Rub1/Nedd8 E1 enzyme[86, 87]. Although the ubiquitin activating enzyme, UAE (Uba1 or Ube1) shows a preference for Ub over Rub1[66], as stated earlier, some excess levels of Rub1

can be activated by UAE even with the presence of uncharged threonine at position 72 of Rub1. Experiments where R72 of Ub had been mutated to the same amino acid as Rub1 has also shown a comparable level of activation of Ub by either NAE or UBA3-ULA1[71]. Interestingly, there have also been other *in vitro* studies that show Nedd8 forming chains on its substrates through its various lysine residues as well as being able to make these chains on the active site of its E2 enzyme as is usually observed on Ub chain formation[88]. How this polyNeddylation event affects protein modification and the resulting effect on the substrates is an avenue worth exploring.

Given the stated scenarios, the question remains on how there is selectivity in the conjugation pathways of the various UBLs whether *in vivo* or *in vitro*. Even though previous work had attempted to address the question of how the proteasomal and non-proteasomal Ub receptors differentiate among the same target proteins modified by either Ub and/or Rub1[68], the bulk of the research discussed here focused on trying to understand the differentiation in the conjugation cascade between ubiquitin and Rub1. More specifically, why can Ub make chains *in vivo* through its various lysines while Rub1 is unable to do so despite having many of the same residues available as well as having an extra lysine that can potentially be used for making chains?

3.3 Ubiquitin conjugating enzyme E2-25K

Unlike the studies on the E1 enzymes, studies on how the E2 enzymes differentiate between UBLs have so far not given a clear answer. As has been discussed previously, there are a number of E2 enzymes that can be utilized in ubiquitin assembly through

different lysine residues. Since it is not practical to study the interaction of Rub1 with all the Ub E2 enzymes, it was necessary to work with a model E2. As such, given that lysine-48 linked chains in the UPS are the most abundant, a decision was made to focus on E2-25K (Ube2K) as said model E2 enzyme.

This 25-kDa mammalian E2 conjugating enzyme has proven to be a largely non-promiscuous enzyme that catalyzes polyubiquitylation through lysine-48 Ub residues[89]. E2-25K is part of a class II Ub conjugating enzyme family that contains a catalytic ubiquitin conjugating domain (UBC) and unique among other E2 enzymes, a C-terminal ubiquitin associated domain (UBA). The UBA domains of these E2 families contain a conserved MGF (M^{172}, G^{173} and F^{174}) loop which plays an integral part in Ub binding and propagation[75]. This domain, like most other UBA domains, binds non-covalently with the hydrophobic-patch region of Ub. Furthermore, even though E2-25K can carry out E2-ubiquitin thioester formation synthesis without the UBA domain, polyubiquitylation chain assembly appears to be much more enhanced with it[89]. It has been suggested that the UBA domain is able to interact with a growing ubiquitin chain on a substrate without actually interfering with the catalytic domain's ability to perform its Ub additions in a reaction[90]. Given the number of different E2's that make distinct ubiquitin chains through the varying lysine residues available, it appears that the conjugating enzyme is quite important in chain type and UBL specificity and selectivity within the conjugation cascade.

The aforementioned 'atypical Rubylation/Neddylation' cases within the Ub pathway have been observed with E2-25K[68]. The results obtained through *in vitro* experiments using E2-25K and other E2 enzymes in this dissertation work do not imply any *in vivo* parallels. Rather, this study tries to understand any residue specific or structural

based reasons behind the differences observed in the post-translational modification roles played by the different UBLs.

3.3.1 Ubiquitin and Rub1 bind to similar areas in E2-25K

To demonstrate that E2-25K can recognize Rub1, NMR titrations were performed with both proteins by looking at ^{15}N labeled E2-25K through SOFAST-HMQC during addition of increasing amounts of unlabeled Rub1 or ubiquitin. By following the chemical shift perturbations in the residues of E2-25K due to binding of either ligand to the enzyme, the main binding sites of both proteins was mapped.

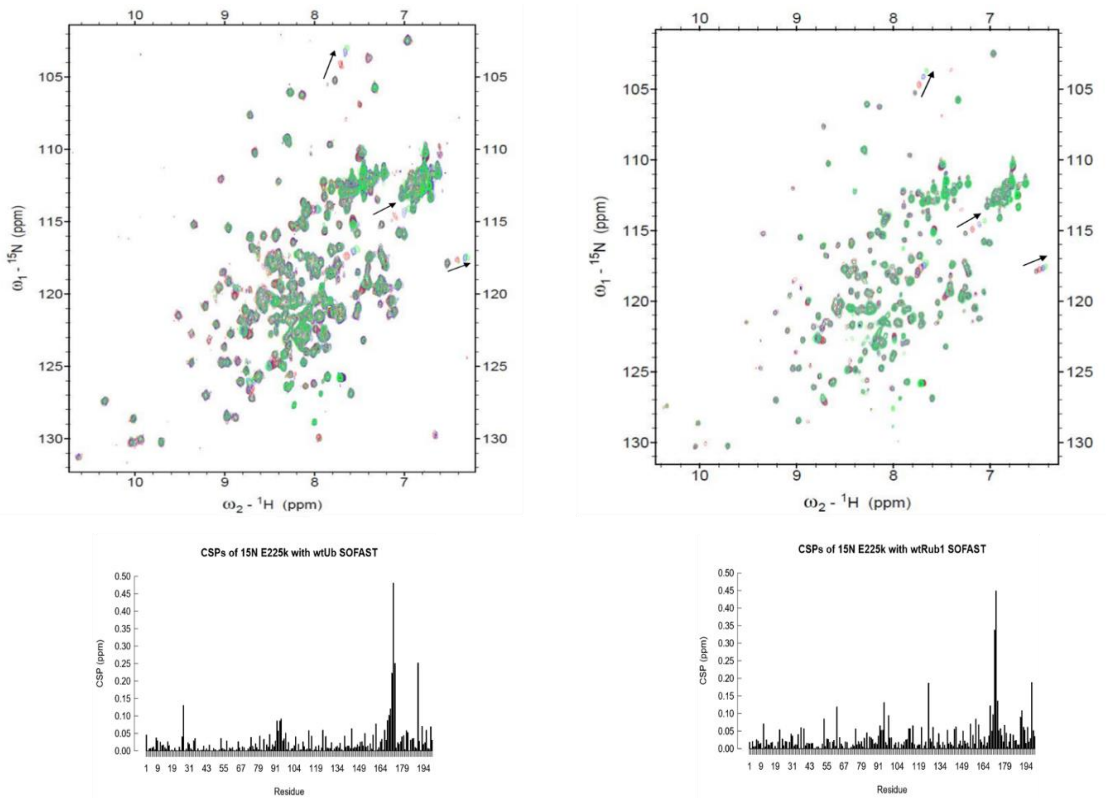


Figure 3.2: ^{1}H-^{15}N SOFAST-HMQC spectra of titration of ^{15}N E2-25K with ubiquitin (top, left). ^{1}H-^{15}N SOFAST-HMQC spectra of titration of ^{15}N E2-25K with Rub1 (top, right). Respective chemical shift perturbations (CSPs) are shown (bottom). Arrows indicate the MGF loop residues of the UBA domain (residues 160 to 200).

In both titration experiments, the main binding site for both Rub1 and Ub was on the UBA domain's MGF loop (Fig 3.2). This was expected for ubiquitin but previously unknown for Rub1. The K_d of binding of Ub to either full length E2-25K or the UBA domain of E2-25K was similar at ~1mM[91]. More recent results from NMR titrations with ^{15N}Ub and E2-25K done in our lab as well other labs has shown that ubiquitin uses its hydrophobic patch to bind to the UBA domain of E2-25K with a K_d of ~400±100 μM[92, 93]. With regards to Rub1, in-house experiments with ^{15}N labeled Rub1 titration on unlabeled, full length E2-25K had shown similar perturbations in the hydrophobic patch of Rub1 with similar K_d of binding (~310±70 μM) (in-house unpublished work: Adithya Sundar). These results suggest that E2-25K interacts with both proteins in a similar manner. Nevertheless, the method by which it differentiates between them during its enzymatic process is still unknown.

The mapping of the perturbed residues on the structure of E2-25K shows the UBA domain involved in the non-covalent interactions that are observed (Fig 3.3). The MGF loop (red) is 34 Å away from the active site cysteine in the UBC domain (pink) and thus, there does not seem to be any steric hindrance caused by non-covalent interaction this domain could have with Ub or Rub1. Yet, this interaction still does not shed light into the exact role the UBA domain plays in the chain making process.

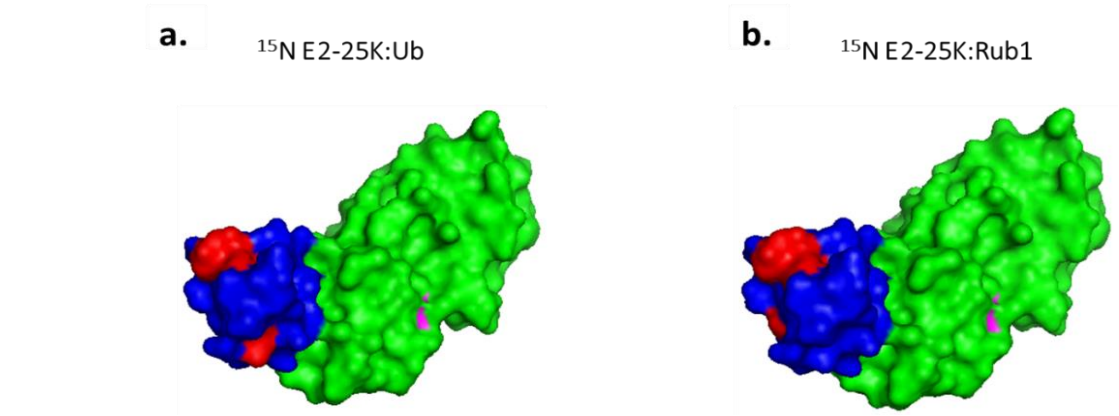


Figure 3.3: PDB structure of E2-25K showing perturbed residues (red) during NMR titration experiments of ^{15}N-E2-25K with Ub (left) and Rub1 (right). E2-25K UBA domains are represented in blue and active cysteine of UBC domain in pink. (PDB ID: 3E46)[94]

3.4 Rub1/Ub chimeras

The first step in trying to use the ubiquitin conjugation enzymes to make Rub1 chains was to identify sequences within each protein that were different electrostatically and/or showed variation in their R-group that could cause sterical limitations during the conjugation reaction process. Once these amino acids termed as "cassettes" were identified (Fig 3.4), the next step was to do mutation-based studies by interchanging residues within these cassettes and testing for conjugation.



Figure 3.4: Rub1 and ubiquitin sequence differentiated into seven "cassettes". Green shaded boxes indicate the main hydrophobic patch residues conserved in both proteins.

Ub/Rub1 chimeras were previously explored in our lab whereby certain residue "cassettes" in ubiquitin were mutated to look like Rub1 residues. The reaction assays that were undertaken after regions I to V of this chimera were individually mutated resulted in similar polyubiquitin chain assembly as the wild-type Ub. The only main difference observed was when dealing with ubiquitin mutants that involved cassettes VI and VII. Chain assembly assays with these mutants (individual or tandem cassette changes) showed slower and less efficient polyubiquitin chain assembly compared to wild type. This suggested that the residues within these cassettes might be important in the selectivity and reactivity of ubiquitin by E2-25K. Therefore, focusing on these two cassettes for making Rub1/Ub chimeras was further explored.

Subsequently, Rub1/Ub chimeras with cassettes VI and VII mutated to look like ubiquitin were made and the resulting assays showed an inefficient yet novel covalent Rub1-Rub1 dimer formation. The low efficiency of Rub1 dimers formed confirmed that even though E2-25K recognizes Rub1, it still has a hard time making chains. These results were also interesting in that initially, the threonine in position 72 of Rub1/Ub chimera was still not changed to an arginine which might have made it hard for Uba1 to activate these chimeras efficiently. As such, all the assays that were performed hereafter and discussed in this dissertation have the T72R mutation incorporated in them unless otherwise stated.

3.4.1 *In vitro* lysine-48 Rubylation reactions with Rub1-Ub chimera

As stated already, the Rub1/Ub chimeras were able to make a limited amount of dimer chain. Still, higher-level chains beyond a dimer size were not formed. As such, even

though cassettes VI and VII showed some promise, other cassettes needed to be tested to check if chain assembly could be improved. Before doing these experiments, control assays with WT-Rub1 being incubated with ubiquitin Uba1 and E2-25K_GST was undertaken (Fig 3.5). The GST tag attached through the N-terminus of the E2 enzyme is used for purification purposes and subsequent assays with Ub had shown that it does not hinder polyubiquitylation. Throughout this work, E2-25K with or without GST has been interchangeably used and any effect that this tag might have in reactions is discussed accordingly. E2-25K enzyme with GST tag is referred as E2-25K_GST throughout this thesis.

Even though Rub1 is known to make heterologous dimers with Ub where it acts as a distal unit acting as a chain cap, a reaction to test if Rub1 could be recognized and incorporated as a proximal unit in a heterologous chain along with Ub was undertaken. To this end, WT-Rub1 was again incubated with ubiquitin enzymes along with Ub_{K48R} (ubiquitin with lysine 48 mutated to arginine to prevent chain elongation from the distal end) (Fig 3.5). As expected, the assays showed that E2-25K is unable to recognize Rub1's lysine-48 either in a homologous WT-Rub1 reaction or a heterologous Ub_{K48R}-Rub1 reaction. Nevertheless, there was an important observation that was made here as well as in subsequent assays whereby very high-level bands above 50kDa were appearing after an overnight reaction suggesting that there is some type of modification occurring on the E2 enzyme. When Ub_{K48R} is in the mix, this phenomenon appears to be much stronger. The possible reason for this occurrence is discussed later.

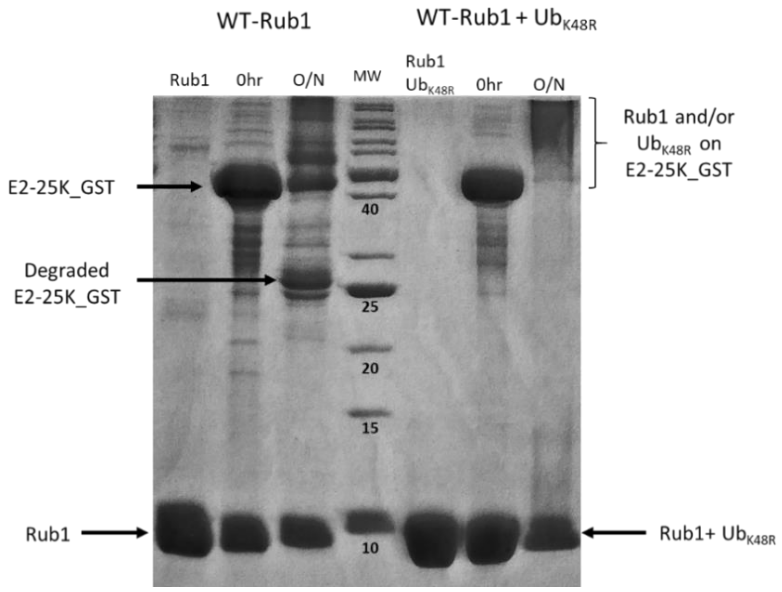


Figure 3.5: Reaction assay of Wt-Rub1 with and without addition of Ub_{K48R}. In overnight (O/N) reactions for Rub1 alone, no chains are formed, and the enzyme degrades even if there is accumulation of Rub1 on enzyme. When Ub_{K48R} is added, still no chain formation but it appears there is full loading of the enzyme (high molecular weight bands).

Following this result, subsequent cassette changes were done in combination to make different Rub1-Ub chimeras. To this end, since an E63K mutation from cassette VII had proven to be enough in producing dimers in Rub1 chimeras (result not shown), this single mutation from the last cassette was included in tandem with the other cassette changes. As such, $Rub1_{I/VI/E63K/T72R}$, $Rub1_{I//II/VI/E63K/T72R}$ and $Rub1_{I/V/VI/E63K/T72R}$ proteins were made, and reactions performed with the hope of finding out any cassettes that could bolster the dimer formation from previous results. As Figure 3.6 shows, the results did not present any observable improvements in terms of chain elongation. Reaction of the chimeras with either Ub_{K48R} or Ub_{D77} (Ub with an aspartate added to the C-terminus to prevent elongation from the proximal end) showed similar results as that of wild type Rub1.

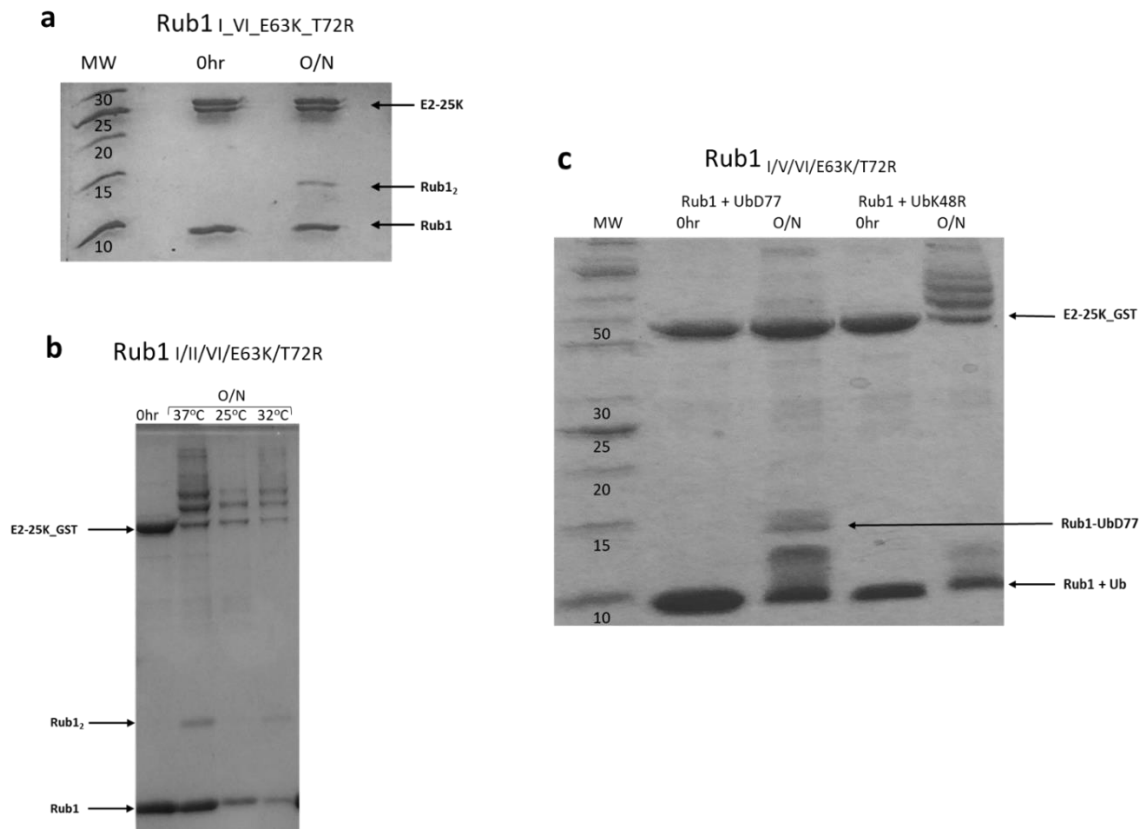


Figure 3.6: Lysine-48 reactions with various Rub1 mutants. **a)** $Rub1_{I/VI/E63K/T72R}$ and **b)** $Rub1_{I//II/VI/E63K/T72R}$. Assay shows that dimerization was dependent on temperature of reaction. **c)** Rub1 $_{I/V/VI/E63K/T72R}$ reacted with Ub_{D77} or Ub_{K48R} indicates that Rub1 can be readily available as a distal unit (left) in a reaction but not as a proximal unit (right). Unless indicated otherwise, all reactions were done at 37ºC.

Given these results, it became prudent to step back and begin mutational analysis by identifying specific residues that are different between Rub1 and Ub within the cassettes and target them instead of changing whole cassettes at once. This was especially necessary given that the cassettes mentioned earlier included both surface and internal residues. Since mutating internal residues could potentially alter the 3D fold of Rub1 and make the study more complicated, focus was shifted to residues that were more surface exposed in addition to having significant charge and/or size difference between Rub1 and Ub.

The easiest starting point was to make Rub_{T72R} and add on to this mutant any relevant mutations and observe any changes that might occur. Since this single mutation was enough to make a Rub1 dimer (Fig 3.7; a), it was not surprising to see that the heterogeneous reactions with Ub_{D77} resulted in a heterologous dimer band as was observed with wild-type Rub1 reactions[68]. In order to prevent any leakage of wild-type ubiquitin that was observed during a control reaction of Ub_{D77}, a $Rub1_{T72R}$ and Ub_{74} heterologous reaction was performed (Fig 3.7; b & c). This showed a similar result to the Ub_{D77} reactions albeit with no ambiguity on the type of dimer formed from this reaction. These results were further validation that it was not necessary to mutate a large number of residues within cassettes in order to form Rub1 chains.

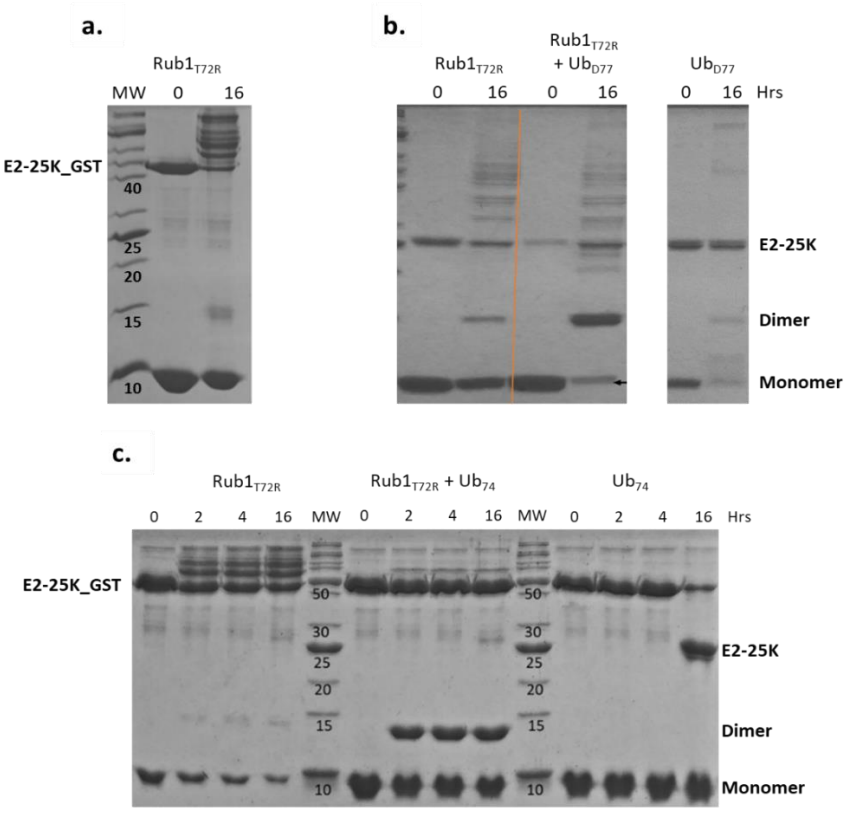


Figure 3.7: Lysine-48 reactions of **a)** $Rub1_{T72R}$; **b)** $Rub1_{T72R}$ with Ub_{D77} and **c)** $Rub1_{T72R}$ with Ub_{74}. $Rub1_{T72R}$ control with E2-25K_GST and E2-25K as well as Ub_{D77} and Ub_{74} control reactions are shown. Robust dimerization for heterologous reactions is observed.

As such, a new criterion for choosing residues to change was utilized based on the NMR results shown earlier on the binding of the two proteins to the UBA domain of E2-25K. The crystal structure of the complex E2-25K:Ub where E2-25K is bound to Ub (PDB: 3K9P), was used to align Rub1 structure with that of Ub, in order to visualize which residues were in close contact with the MGF loop of E2-25K and focus on the ones that were different between Ub and Rub1 (Fig 3.8). As mentioned before, care was taken to only focus on residues that were solvent exposed.

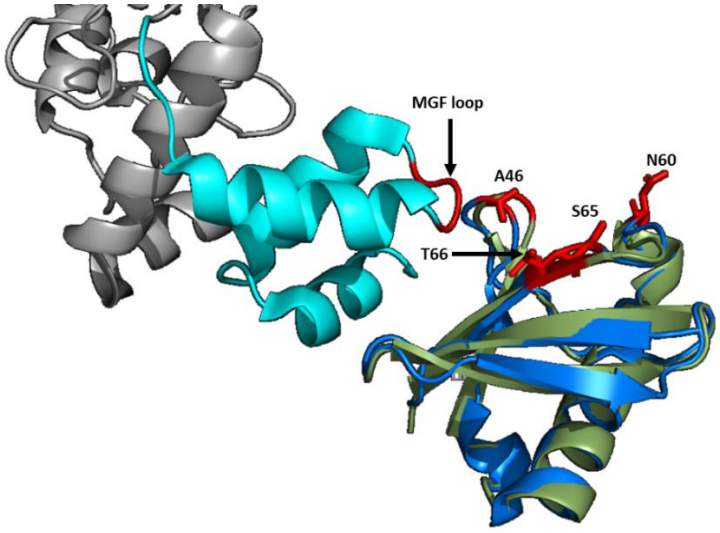


Figure 3.8: E2-25K: Ub crystal structure with Ub (blue) bound to the UBA domain of E2-25K (cyan). The MGF loop on the UBA domain and residues nearby that were mutated in Rub1 are represented in red. Rub1 structure (green) is aligned with Ub structure (blue). (Rub1 PDB ID: 1BT0 and E2-25K: Ub PDB ID: 3K9P)[67, 95]

With this in mind, the next residue chosen for mutation was residue 60 which is a histidine in Rub1 and an asparagine in Ub. Given the proximity of asparagine-60 to lysine-48 in Ub, a previous study looking at a K60N mutation in Nedd8 had reported some increase in $Nedd8_{K60N}$'s ubiquitin accepting activity[96], during a lysine-48, heterologous enzymatic reaction. Thus, in a similar manner, a H60N mutation on Rub1 was worth

exploring with the hypothesis that if residue 60 had a role to play in formation of lysine-48 linked chains, then the addition of the H60N mutation on $Rub1_{T72R}$ would enhance chain formation of Rub1. Disappointingly, the assays with this mutant showed no appreciable difference in chain elongation but, $Rub1_{H60N/T72R}$ in a reaction with Ub_{74} showed dimer bands indicating a relatively robust dimerization occurring which appears to be a combination of heterologous Rub1_Ub dimers as well as homologous Rub1_Rub1 dimers (Fig 3.9).

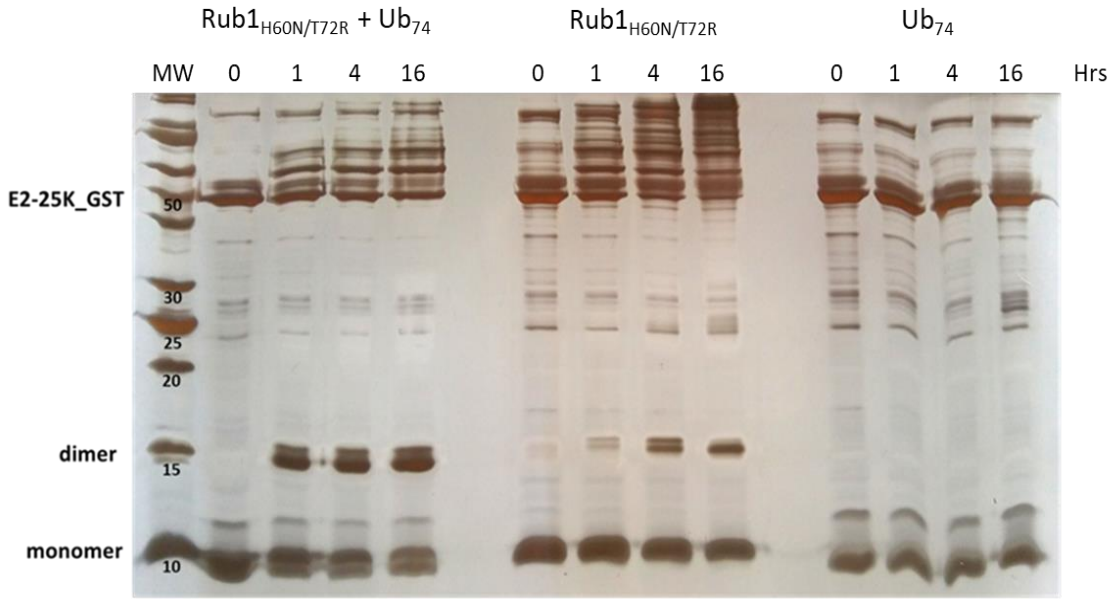


Figure 3.9: Silver-stained lysine-48 reactions of $Rub1_{H60N/T72R}$ and Ub_{74} with $Rub1_{H60N/T72R}$ and Ub_{74} control reactions shown. Heterologous dimers are more readily made without further chain elongation.

This result was repeated when assessing further mutants such as $Rub1_{E63K/T72R}$ and $Rub1_{K4F/T72R}$ (Fig 3.10). In the former case, the apparent difference in the charges at position 63 between Rub1 and Ub was a hopeful target in terms of affecting the charges nearby. In the latter case, not only was there a charge difference but also a potential steric difference between the lysine resides on Rub1 and a phenylalanine residue on Ub. In these

reactions, E2-25K with and without and GST tag was utilized. Compared with previous assays, it was apparent that the amount of Rub1 modification on the enzyme dramatically declined when there is no GST attached. Based on this, it was surmised that it was possible that even if E2-25K is unable to transfer Rub1 to lysine residues in order to make higher order chains, it could potentially still transfer Rub1 to other substrates that have available lysine residues (GST in this case). Based on these results, one cannot say if the modifications in either the E2 or the GST tag are multiple monoRubylation events or formation of actual poly-Rub1 chains which would in fact shed more light on the capacity of E2-25K to make Rub1 chains. Mass spectrometry experiments were used to shed light on this question which will be discussed in the next chapter.

In either case, when comparing the dimer bands formed with the use of E2-25K_GST and E2-25K alone, a more efficient formation of dimer is observed with GST tagged enzyme (Fig 3.10, b). This suggests that the tag might be playing an additional role of somehow increasing the efficiency of Rub1 chain formation. Based on the same figure, it would be tempting to say that the band seen on the $Rub1_{K4F/T72R}$ reaction gel at ~25kDa corresponds to a trimer but as shown on the next lane indicated with a star, the Rub1 protein fully precipitated during the reaction period and thus could not be purified for further analysis. Subsequent attempts to repeat this reaction at lower temperatures resulted in either the protein degrading or no formation of bands at the trimer level. Even though this was the most promising Rub1 dimer chain elongation that has so far been observed through assays, the fact that previous work on mutating the entire first cassette (where lysine-4 is located) did not give the same results compounded on the mystery of this whole project.

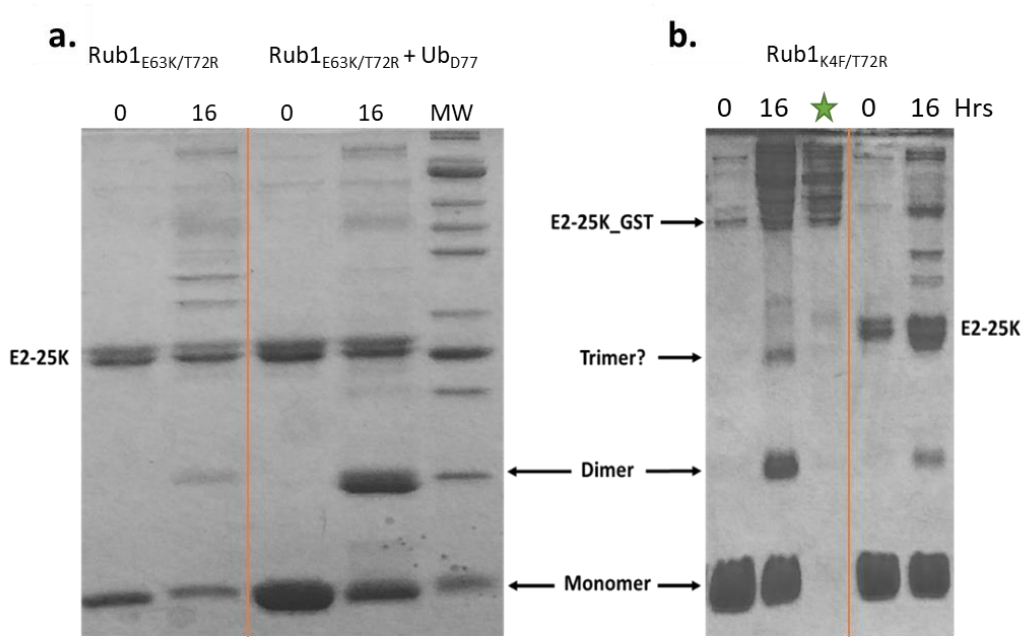


Figure 3.10: Lysine-48 reaction assay of **a)** $Rub1_{E63K/T72R}$ and Ub_{D77} and **b)** $Rub1_{K4F/T72R}$ with E2-25K_GST (left) and E2-25K (right). Lane with green star indicates 16 hr reaction after centrifuging out precipitate formed. Heterologous dimerization is robust in E63K/T72R mutant while E2-25K_GST enzyme allows for greater dimerization and enzyme (GST) modification.

The next attempt focused on the differences observed in previous NMR titrations of Ub and Rub1 on labeled E2-25K. More specifically, glutamine-46 (Q46) in Rub1 showed very little shift when binding to E2-25K while alanine-46 (A46) in Ub experienced a much greater environmental change. Therefore, separate reactions with Ub_{A46Q} and Rub1 variants incorporating Q46A mutations were used for an assay in order to see if these changes could have any effect in the efficiency of lysine-48 chain formation on both proteins.

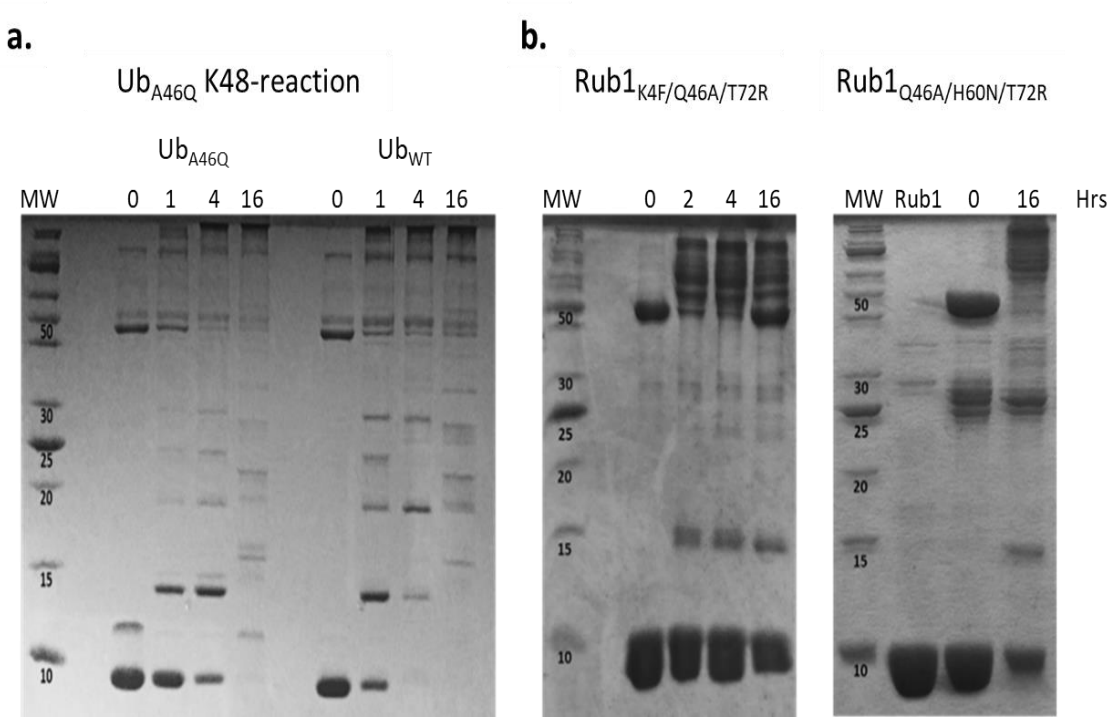


Figure 3.11: Lysine-48 time course reactions of **a)** Ub_{A46Q} and WT-Ub with an observable difference by the 4th hour that the mutant Ub is much slower in making higher level chains and **b)** Rub1 variants with and addition of a Q46A mutation. Both reaction gels showed a maximum of dimer chains with a high level of loading on E2.

Interestingly, the assay showed that despite chains being ultimately formed, the Ub_{A46Q} mutant, lysine-48 linkage reaction was much slower as compared to wild type ubiquitin reaction (Fig 3.11; a). Given this observation, the reverse mutation was made on Rub1, and subsequent reactions made with the hope of getting a promising result. This reverse mutation was combined with other mutations that seemingly showed the most promise in potentially elongating chains higher than a dimer. The subsequent assays indicated no visible enhancement in chain formation (Fig 3.11; b) but in the process of purifying, storing, and making reactions with these mutants, it was observed that the overall stability of Rub1 was improved.

The results of the reaction assays conducted so far where discouraging in that as new mutated residues were added on Rub1, either the efficiency of chain formation did not change or in some cases, was drastically reduced from earlier results. This case of dimer disappearance specifically occurred in Rub1 which had an M65S mutation incorporated. This mutant was interesting in that there was a pre-reaction band that appeared corresponding to a dimer size, but this band could not be purified and/or be observed by mass spectrometry. An additional matter of interest was that reactions with this mutation showed a higher than normal amount of accumulation of Rub1 on the E2 enzyme which raises the question of whether or not this mutant is highly susceptible to making chains on the E2 enzyme itself (Fig 3.12; a & b). A reverse mutation on Ub was conducted where a UbS65M mutant was subjected to a lysine-48 reaction with the expectation of observing a reduced efficiency in poly-ubiquitination. Alas, the reaction did not show anything out of the ordinary in that the ubiquitin chain reaction went to completion within the given timeframe (Fig 3.12; c).

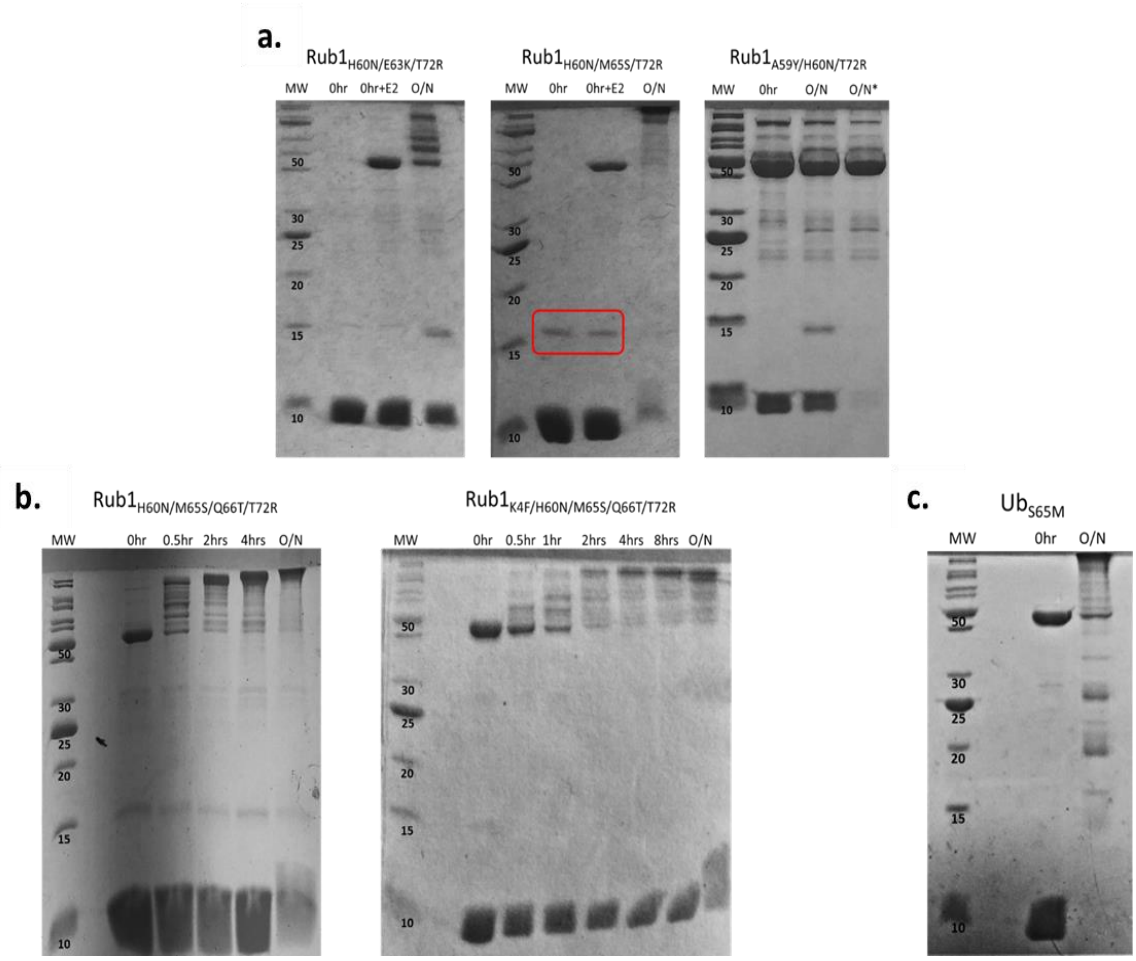


Figure 3.12: Lysine-48 reactions of a) $Rub1_{H60N/T72R}$ mutant with an addition of E63K (left); M65S (middle) and A59Y (right). Red box indicates pre-reaction dimer band observed with M65S mutant; **b)** $Rub1_{H60N/M65S/T72R}$ with an addition of Q66T (left) and K4F/Q66T (right). Notice the lack of dimer formation with Rub1 that has the M65S mutation along with the extensive modification on the enzyme with monomer band drastically reduced. **c)** Control reaction with Ub_{S65M} mutant where poly-Ub formation was not hampered.

3.4.2 *In vitro* lysine-11 and lysine-63 Rubylation reactions with Rub1/Ub chimera

Given the partial results with E2-25K, it was considered useful to test whether or not the lack of chain elongation was relegated to lysine-48 based reactions. As such, other E2 enzymes were used to assess the possibility of higher order chains forming through other Rub1 lysines. To that end, lysine-11 chains were explored using the ubiquitin E2,

Ube2S. Additionally, the ubiquitin E2 heterodimeric enzyme Ubc13/Mms2 was used to specifically observe possible lysine-63 chain formation even though Rub1 does not have a native lysine at position 63. As such, this specific reaction was conducted only on the mutant $Rub1_{E63K/T72R}$. As was with the lysine-48 reactions, only dimers were formed throughout these experiments (Fig 3.13) and the enzyme modifications for these reactions were significantly less compared to that of E2-25K.

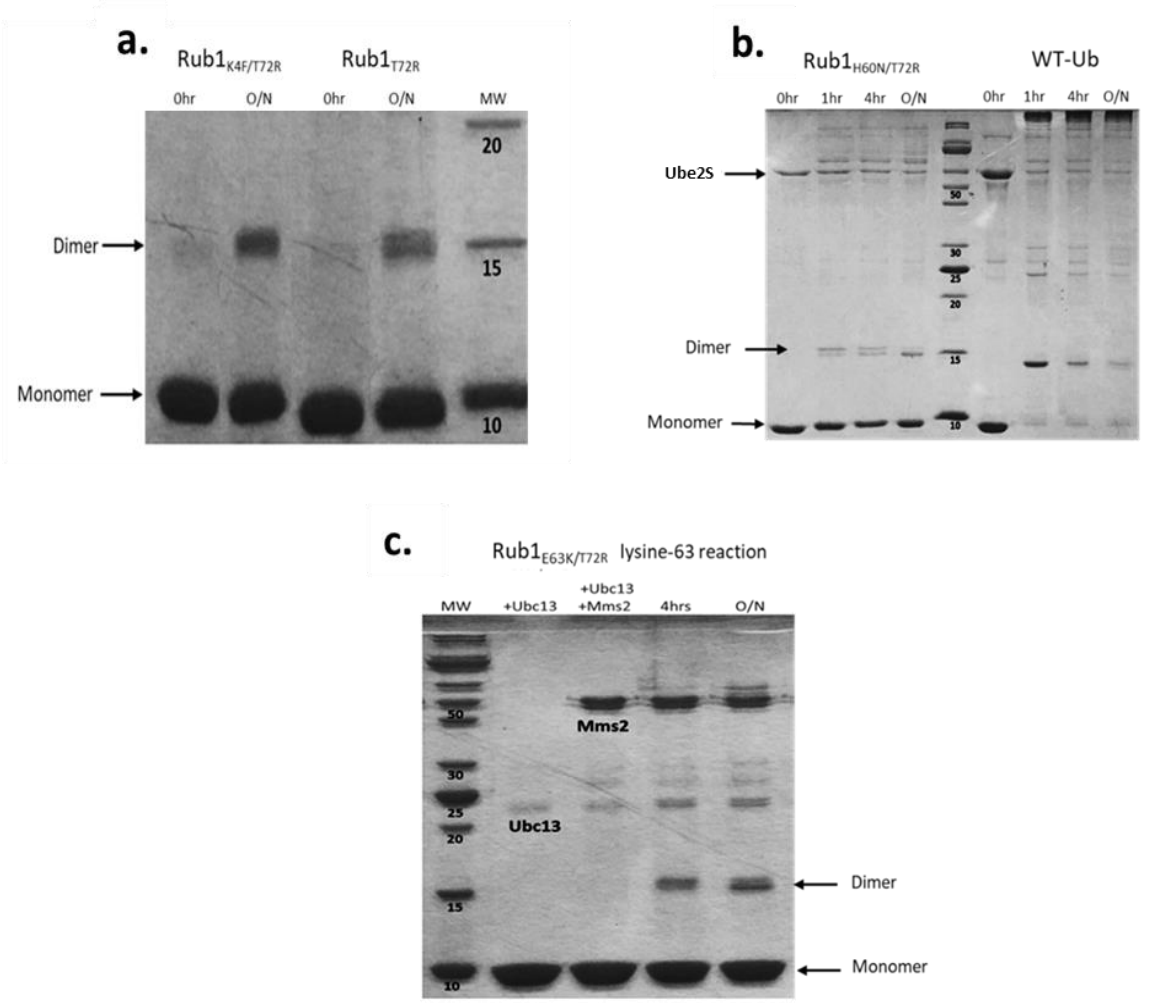


Figure 3.13: Lysine-11 reactions of **a)** $Rub1_{K4F/T72R}$; $Rub1_{T72R}$ and **b)** $Rub1_{H60N/T72R}$. **c)** Lysine-63 reaction of $Rub1_{E63K/T72R}$. Only dimer chains were observed in all reaction assay gels. Bands around 70kDa in panel-b indicate a potential loading of $Rub1_{H60N/T72R}$ or WT-Ub onto the Ube2S enzyme.

All the results discussed so far indicate that Rub1 can covalently attach to one or several lysines of E2-25K_GST during *in vitro* reactions. Thus, it can be safely assumed from these assays that Rub1-E2 thioester intermediate is formed and yet the following step of unanchored Rub1 chain formation is generally hampered. The most puzzling part of these reactions lies in the fact that there is some nucleophilic attack of the thioester by the lysine residue of an acceptor Rub1 in order to form dimer bands. Nevertheless, for reasons that are still unresolved, the repetition of this phenomenon to enable elongation of the chain is still prevented. Several of the mutations that were discussed above did not change this effect which points to some unfavorable conformational state which the acceptor Rub1 takes upon binding to any of the E2 conjugating enzymes that were utilized for this study. In any case, a slightly different approach needed to be utilized in order to specifically test the efficacy of E2-25K reactivity on Rub1.

3.5 From Rub1-Ub dimers to heterotrimers

Given the proven ability of Rub1 to be incorporated in a heterologous ubiquitin chain and act as a chain terminator, it was hypothesized that using Rub1 mutant that was able to independently make Rub1 dimers (in this case $Rub1_{T72R}$) might assist in furthering the chain of this heterologous dimer and at least add one more Rub1 or Ub to make a lysine-48 linked trimer. For this purpose, initial tests were done with the addition of Ub to a $Rub1_{T72R}$-Ub_{74} dimer. Since this dimer cannot incorporate an additional unit on the proximal end, a Ub_{K48C} mutant was used as the proximal Ub to be added. This K48C mutation prevents ubiquitin from making chains on its own and thus any reaction can only occur through an addition on the free lysine-48 of the Rub1 unit of the heterologous dimer

present. A control reaction was conducted in parallel where Ub_{K48C} was added onto a Ub_{WT}-Ub_{74} homologous dimer which should form a trimer during a reaction.

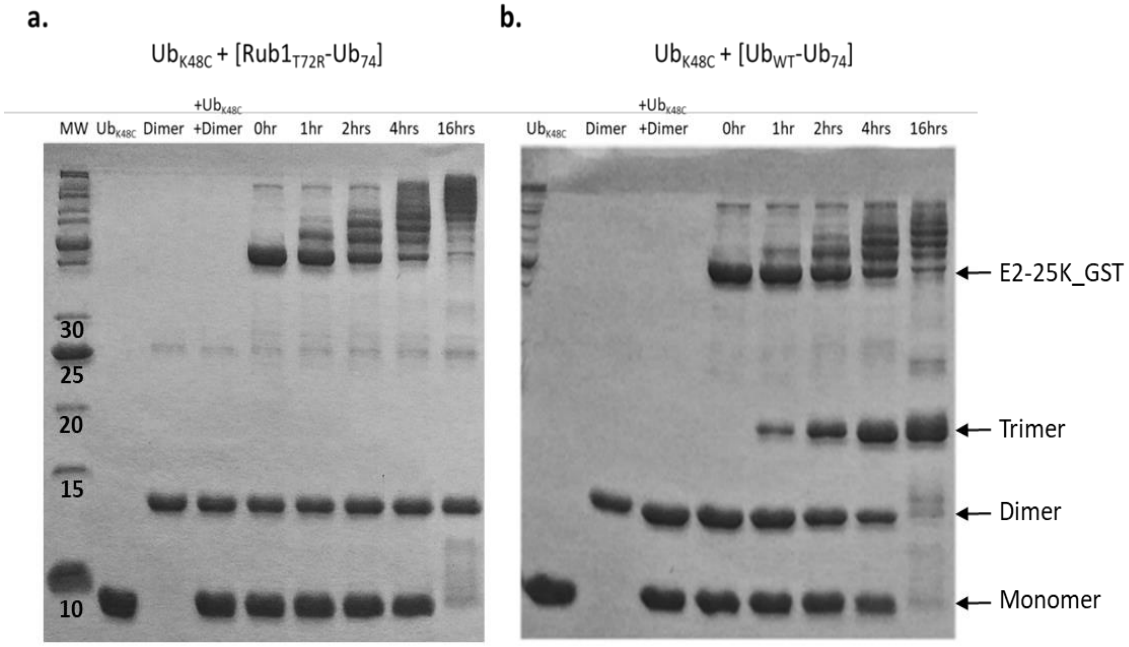


Figure 3.14: Lysine-48 trimer reactions of a) $Rub1_{T72R}$-Ub_{74} heterologous dimer and **b)** Ub_{WT}-Ub_{74} homologous dimer with the addition of Ub_{K48C} as the proximal unit. Rub1 still acts as a chain terminator.

It is quite clear from the gel results that the ubiquitin conjugation system does not have any issues in extending a Ub_{K48C} unit on a ubiquitin dimer, but a $Rub1_{T72R}$-Ub hybrid dimer does not appear to be extended further. It is interesting to note that in the $Rub1_{T72R}$-Ub_{74} dimer assay, despite lack of chain elongation, there was an observable accumulation of Ub_{K48C} on the E2 enzyme (Fig 3.14; a). This seems to indicate that E2-25K does not recognize the availability of free lysine-48 residues on the distal $Rub1_{T72R}$ unit of the dimer and instead prefers to transfer Ub_{K48C} onto free lysines found on the enzyme itself. These reactions were conducted in the presence of TCEP and thus formation

of disulfide bonds with the enzyme was not expected. Furthermore, similar reactions were performed with Ub_{K48R} used instead of Ub_{K48C} with reproducible results.

In order to test the reverse reaction where a Rub1 unit (rather than Ub) was used to extend the heterologous dimer, a reaction assay where $Rub1_{T72R}$ was added to a reaction mixture in an attempt to form chains where this $Rub1_{T72R}$ would act as a proximal unit of a $Rub1_{T72R}$-Ub_{74} dimer was performed. Observing the result through SDS gel methods did not indicate a conclusive positive outcome. When the gel was subjected to silver staining there was some indication of a trimer being formed in this reaction (Fig 3.15). Nevertheless, as the gel shows, the reaction conditions necessary to give this result cause the very small amount of trimer formed to crash out of solution making any further analysis difficult. Still, this gives some hope that E2-25K might be able to transfer a mutated Rub1 onto an already formed dimer. Further attempts in trying to maximize the outcome shown here need to be undertaken. Of course, in the larger picture, this is still not enough of a positive result to make any conclusions regarding why and how Ub is selected by specific enzymes to accomplish its tasks. Still, the research done so far leads us a few steps further in trying to understand this complicated system. Furthermore, given the results described, a mass spectrometric approach of trying to analyze the various bands formed during the assays was conducted. The results from these attempts are discussed below.

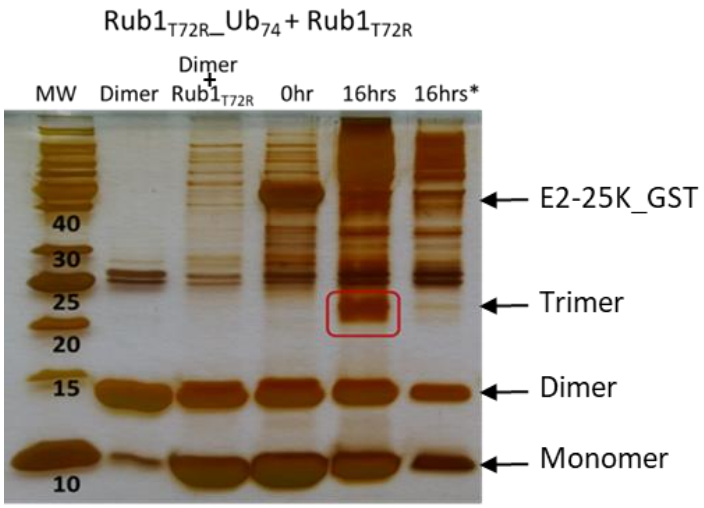


Figure 3.15: Silver stained gel of trimer reaction of Rub1$_{T72R}$-Ub$_{74}$ dimer with Rub1$_{T72R}$. The red box indicates a formation of possible trimer. Last lane shows sample after centrifuging the precipitate fromed during the reaction. Bands at ~25kDa in lanes 2 and 3 are dimer impurites while similar bands in lanes 4,5 and 6 indicate E2-25K that has dissasmebled from GST tag during the course of the reaction.

Chapter 4. Mass spectrometric analysis of Rub1/Ub chains

Even though purification of the poly-Rub1 formed through the reaction assays proved to be unsuccessful, there was still a desire to decipher the relationship between protein modifier structures and their potential functions. As such, in order to characterize the components and lysine attachment sites of any of the branched protein products observed through assays as well as confirm the presence of any expected mutations, certain proteomic strategies needed to be utilized.

Most of what is known about the cellular occurrence of proteins modified by ubiquitin (or Rub1) has been acquired by analysis of the tryptic products of ubiquitinated (Rubylated) proteins. Both ubiquitin and Rub1 are truncated at R74 to leave -GG tags on the modified lysines in peptides from the substrate proteins[97-99]. Thus, using this approach, a number of reaction assays involving Rub1 and Ub were explored.

4.1 Mass spectrometry analysis of possible enzyme modification by Rub1 and/or Ub

As mentioned earlier, some of the reactions of $Rub1_{T72R}$ and Ub_{K48R} showed an accumulation of bands higher than 25kDa on SDS gels. Based on these observations, it was hypothesized that Rub1 could be modifying the E2-25K (Ube2K) and potentially, E1 (Uba1) enzymes by either through several mono-Rubylation instances or potentially making longer chains on the enzymes. The E2-25K enzyme has 14 lysine residues that could conceivably make a covalent attachment to either Rub1 or Ub. In utilizing E2-25K enzyme with a GST-tag (GST26_SCHJA) attached, the number of lysine residues balloons to 21. This is further complicated by the fact that the E1 used in these reactions has an

additional 53 residues that could possibly be modified as well. With this in mind, several enzymatic reactions were undertaken with the purpose of making heterologous chains between Ub and Rub1 as well as homologous chains of Rub1 and try to utilize tryptic digestion-based mass spectrometry analysis to try and distinguish the type of chains being formed as well as the modification occurring on the enzymes. As such, Ub_{K48R} and Rub1 were incubated with both E2-25K_GST and E2-25K conjugating enzymes.

In analyzing gel extracted proteins through either ionization mass spectrometry (MALDI-MS) or electrospray mass spectrometry (ESI-MS) techniques, trypsin is the most commonly used enzyme for digestion of proteins. This serine protease specifically cleaves protein at the C-terminal side of arginine or lysine producing tryptic peptides[100]. As already stated, since both Ub and Rub1 end in a C-terminal tail sequence of -LRLRGG, tryptic digestion would cleave both enzymes at the last arginine amino acid resulting in a -GG tag for both Ub and Rub1 making differentiation of the modifier (distal) unit from the receptor (proximal) unit of a chain impossible. For this reason, a ^{15}N-labeleing scheme was undertaken that would allow to make detailed and unambiguous analysis based on the higher mass that a labeled protein would exhibit. As such, after the relevant labeling and reactions were conducted in-house, tryptic digestion of the resulting gels bands of interest was done and mass spectrometry analysis performed by Dr. Marta Mendes from Prof. Gunnar Dittmar's lab at the Luxembourg Institute of Health.

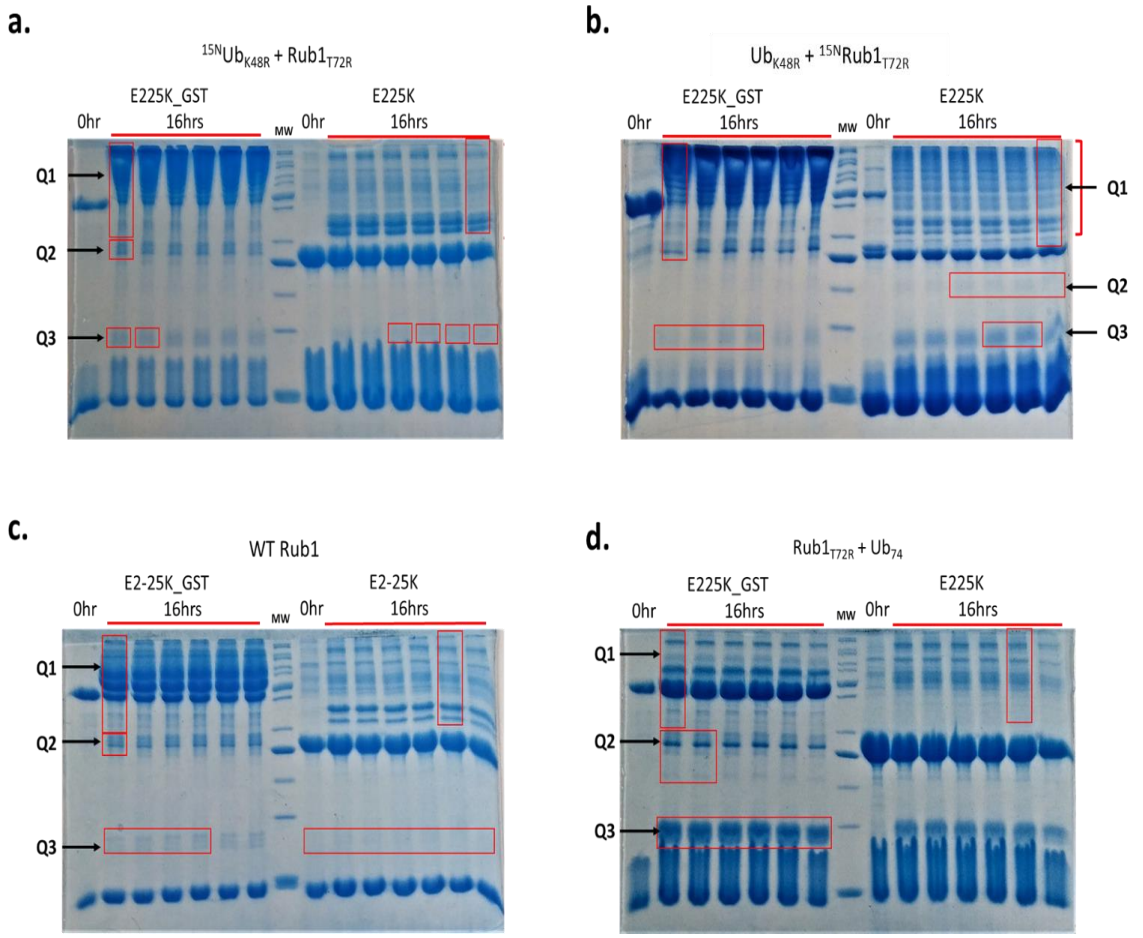


Figure 4.1: Enzymatic chain reaction gel of Ub and Rub1 with GST tagged and untagged E2-25K used for MS analysis: Heterologous Ub/Rub1 reactions with ^{15}N-labeling of **a)** Ub_{K48R} and **b)** $Rub1_{T72R}$; **c)** homologous reaction of wild-type Rub1; **d)** heterologous reaction of $Rub1_{T72R}$ with Ub_{74} which forces the Ub to be a proximal unit. Three main sections of bands are indicated for MS analysis.

As Fig 4.1 shows, each gel was sent with specific bands identified for analysis. These bands were generally selected to ask three main questions:

Q1) Are the high molecular weight bands Ub or Rub1 loaded onto E1 and/or E2 enzymes?

Q2) Do the bands between MW 20-30 KDa represent trimer or tetramer chains? If so what kind?

Q3) What kind of dimer do the bands around 15 KDa represent?

The answers for these questions gave interesting results in terms of the ability of Rub1 to make both heterologous and homologous chains. Furthermore, both Ub and Rub1 showed a relatively robust modification ability on both the activating and conjugating enzymes. As expected, there was some variation in modification of enzymes depending on the presence or absence of GST tag on E2-25K. It is not surprising that when GST is available, both proteins will readily modify the GST lysines. Interestingly, in cases where both Rub1 and Ub are present in a reaction, Ub takes the lion's share in terms of higher order modification of GST and both enzymes (Table 4.1; a, b). This shows the high reactivity of Ub, and the potential selectivity shown by the enzymes for Ub.

Furthermore, despite the fact that E2-25K is a lysine-48 specific enzyme, the chains identified by MS analysis show a high degree of promiscuity in the chain linkage types. One could expect this in the presence of Ub_{K48R} mutants which cannot be extended on the distal end and yet, the availability of lysine-48 in Rub1 did not appear to push the reaction to form a high amount of heterologous Ub-Rub1 chains through this lysine linkage. In fact, lysine-11 modifications of Rub1 were consistently present in almost all types of the reactions analyzed and the presence of K53 linkages on Rub1 was quite interesting given that this lysine is unique to Rub1 in these reactions (Tables 4.1-4.3).

a. $^{15N}Ub_{K48R}$ + $Rub1_{T72R}$

E2-25K

Modifier(distal)	Receptor (proximal)	GG: 15N-adduct (from Ub)	GG-adduct (from Rub1)	N° of peptides
15N-Ub_{K48R}	Rub1	K11, K18		10
Rub1	15N-Ub_{K48R}		K11, K63	9
15N-Ub_{K48R}	15N-Ub_{K48R}	K6		9
15N-Ub_{K48R}	GST	K27		17
15N-Ub_{K48R}	UBE2K	K18, K24, K72, K97, K165, K186, K28, K135		31
15N-Ub_{K48R}	UBA1	K806, K296, K89, K97, K304, K322, K385, K411, K593, K604, K528, K470, K923, K746, K671, K984, K465		85

E2-25K_GST

Modifier(distal)	Receptor (proximal)	GG: 15N-adduct (from Ub)	GG-adduct (from Rub1)	N° of peptides
15N-Ub_{K48R}	Rub1	K11, K18, K27, K33, K53, K48		10
Rub1	Rub1		(K11)?	
Rub1	15N-Ub_{K48R}		K6, K11, K27, K29, K33, K63	14
15N-Ub_{K48R}	15N-Ub_{K48R}	K6, K11, K33, K63		14
15N-Ub_{K48R}	GST	K40, K119, K125, K64, K87, K113, K191/K194, K11, K45, K181, K27, K131, K9, K44		41
Rub1	GST	K27		
15N-Ub_{K48R}	UBE2K	K14, K18, K24, K28, K72, K97, K164/K165, K186, K135, K61		34
15N-Ub_{K48R}	UBA1	K296, K89, K304, K322, K385, K671, K465		79

b. Ub_{K48R} + $^{15N}Rub1_{T72R}$

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Rub1)	GG-adduct (from Ub)	N° of peptides
Ub_{K48R}	15N-Rub1		K48	4
15N-Rub1	Ub_{K48R}	K63		18
Ub_{K48R}	Ub_{K48R}	K6,K11,K33, K63		18
Ub_{K48R}	GST	K113, K125, K191, K45, K27		25
Ub_{K48R}	UBE2K	K14, K18, K24, K28, K72, K164, K165, K186, K135		28
Ub_{K48R}	UBA1	K806, K838, K89, K304, K322, K385, K604, K671, K465		63

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Rub1)	GG-adduct (from Ub)	N° of peptides
Ub_{K48R}	15N-Rub1		K18, K53	7
15N-Rub1	15N-Rub1	K11		7
15N-Rub1	Ub_{K48R}	K11,K63		19
Ub_{K48R}	Ub_{K48R}	K6,K11,K63		19
15N-Rub1	GST	K45, K64, K27		37
Ub_{K48R}	GST	K119, K125, K64, K113, K191, K194, K197, K11, K45, K181, K27, K131, K9, K44		25
15N-Rub1	UBE2K	K72		30?
Ub_{K48R}	UBE2K	K14, K18, K24, K28, K72, K97, K164, K165, K135, K61		30?
Ub_{K48R}	UBA1	K89, K304, K322, K385, K671, K465		64

c. Rub1-WT

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	N° of peptides
Rub1	Rub1	K11, K18,K48	14
Rub1	GST	K27	15
Rub1	UBE2K	K14, K18, K24, K72, K97, K165, K28	35
Rub1	UBA1	K89, K322, K671	70

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	N° of peptides
Rub1	Rub1	K11, K18, K27, K33, K48, K53	16
Rub1	GST	K119, K125, K113, K191, K194, K197, K45, K181, K27, K40, K131, K44	40
Rub1	UBE2K	K14, K18, K72, K97, K164, K165, K28, K135	34
Rub1	UBA1	K89, K322, K671, K604, K806	70

d. $Rub1_{T72R}$ + Ub_{74}

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	N° of peptides
Rub1	Rub1	K11, K27	9
Rub1	Ub	K48	12
Rub1	UBA1	K322, K746, K806	88

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	N° of peptides
Rub1	Ub	K48, K63	14
Rub1	GST	K113, K191, K45, K27, K44	40
Rub1	UBE2K	K14, K164	33

Table 4.1: MS analysis of enzyme loaded gel bands identified by Q1 in fig 4.1 for reactions with E2-25K (left) and E2-25K_GST (right): Almost all modifications of enzymes and GST is through Ub_{K48R} in the heterologous reactions (a) and (b). Rub1 mostly modifies GST when availbale (c) with limited modifcaton observed in the presence of Ub_{74} as proximal Ub unit(d). All GG-adducts are from either a distal Ub or distal Rub1 unit.

Despite the limited instances observed, there were in fact, a number of cases were Rub1 was modified through its lysine-48 at different chain levels in both homologous and heterologous reactions (Table 4.2 and 4.3) with or without GST tagged E2-25K enzyme. Along with the fact that Rub1 has consistently been shown to act as a chain terminator by modifying Ub on the distal end, these *in vivo* results show that the Ub enzymatic system can, to a certain degree, recognize Rub1 as a proximal unit to form lysine-48 linkages both in the presence and absence of Ub.

a. $^{15N}Ub_{K48R}$ + $Rub1_{T72R}$

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Ub)	GG-adduct (from Rub1)	Nº of peptides
15N-Ub_{K48R}	Rub1	K33, K48, K53		7
Rub1	Rub1		K11, K27	-
15N-Ub_{K48R}	15N-Ub_{K48R}	K6		1
15N-Ub_{K48R}	GST	K27, K45		28

b. Ub_{K48R} + ^{15N}Rub1-WT

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Ub)	GG-adduct (from Rub1)	Nº of peptides
15N-Ub_{K48R}	Rub1	K11, K48		7
Ub_{K48R}	Ub_{K48R}	K6,K11,K63		8
Ub_{K48R}	GST	K191		12
Ub_{K48R}	UBE2K	K61, K72		9

c. Rub1-WT

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	Nº of peptides
Rub1	GST	K113	21

d. $Rub1_{T72R}$ + Ub_{74}

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	Nº of peptides
Rub1	Ub	K48 (Tetramer band)	8
Rub1	Rub1	K11 (Trimer band)	8
Rub1	Ub	K48, K63 (Trimer band)	12

Table 4.2: MS analysis of trimer and tetramer gel bands identified by Q2 in fig 4.1: All reactions except (c) show bands that contain either a heterologous or homologous chain formed. Only in reaction (d) does Rub1 act as a modifier unit through K48 both in a trimer and dimer chain.

Even though MS analysis using tryptic digestion is not able to indicate the length of any of the chains formed, based on the visualization of the reaction gels, it is quite possible that Rub1 could potentially make longer chains once it is loaded on E2 enzyme. This is somewhat ambiguous when looking at the homologous wild-type Rub1 reaction gel

results (Fig 4.1; c) which show an increase in higher order bands upon reaction. Since the increase in bands in this gel could be because of mono-Rubylation of enzymes or Rub1 chain assembly (Table 4.1; c), further exploration into the mechanism of lysine-48 based Rub1 conjugation by E2-25K needs to be made. Modifications that were observed through MS analysis are listed in Tables 4.1, 4.2 and 4.3. Specific analysis of heterologous dimers formed through Rub1-Ub linkages as well as homologous Rub1-Rub1 linkages were further investigated through a top-down mass spectrometry analysis method and discussed in the following section.

a. $^{15N}Ub_{K48R}$ + $Rub1_{T72R}$

E2-25K

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Ub)	GG-adduct (from Rub1)	Nº of peptides
15N-Ub_{K48R}	Rub1	K11, K48		7

E2-25K_GST

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Ub)	GG-adduct (from Rub1)	Nº of peptides
15N-Ub_{K48R}	Rub1	K33, K48		9
Rub1	Rub1		K11	9

b. Ub_{K48R} + $^{15N}Rub1_{T72R}$

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Rub1)	GG-adduct (from Ub)	Nº of peptides
Ub_{K48R}	Ub_{K48R}		K33, K6, K11, K63	12

Modifier(distal)	Receptor (proximal)	GG15N-adduct (from Rub1)	GG-adduct (from Ub)	Nº of peptides
Ub_{K48R}	15N-Rub1		K11,K48	2
Ub_{K48R}	Ub_{K48R}		K6, K11, K63	12

c. Rub1-WT

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	Nº of peptides
Rub1	Rub1	K11, K27,K48	10

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	Nº of peptides
Rub1	Rub1	K18, K27, K33, K48, K53, K11	15

d. $Rub1_{T72R}$ + Ub_{74}

Modifier(distal)	Receptor (proximal)	GG-adduct (from Rub1)	Nº of peptides
Rub1	Rub1	K11,K27,K48	13
Rub1	Ub	K6, K48	15

Table 4.3: MS analysis of dimer gel bands identified by Q3 in fig 4.1 for reactions with E2-25K (left) and E2-25K_GST (right): All reactions except (c) show bands that contain either a heterologous or homologous chain formed. Only in reaction (d) does Rub1 act as a modifier unit through K48 both in a trimer and dimer chain. Empty cells represent no adducts detected.

4.2 Mass spectrometry based Top-Down analysis of Branched Rub1 proteins

Parallel to the enzymatic experiments described in chapter 3 and independent from the previous MS based analysis shown, a separate Top-down mass spectrometry approach was utilized to gain insight into the dimeric architecture of Ub-Rub1 and Rub1-Rub1 chains. As previously described, bottom-up proteomic methods relies on the fact that -GG-tagged peptide be located within a complex mixture of tryptic peptide products which has the drawback of not allowing distinction of ubiquitination from Rubylation since both proteins have the same C-terminal sequence and does not provide information on the length and connectivity of the modifications formed[101, 102]. Despite the labeling method used to circumvent this difficulty, a separate approach of top-down analysis of branched proteins was expected to allow for characterization of complete structures and investigate the relationship between said structure and function without the need of isotopic labeling of proteins.

Even though, top-down MS/MS spectra of homopolymers of ubiquitin[103], isomeric ubiquitin trimers[104], and isomeric ubiquitin tetramers[105] have been recently reported, examples of top-down analyses of branched proteins or proteins modified by ubiquitin or Rub1/Nedd8 had not been reported. This is in large part because of the problem of purifying and isolating these modified proteins from cells in sub-microgram amounts. Still, given the ability of our lab to make relatively stable, *in vivo* enzymatic conjugations of ubiquitin and Rub1, samples of Rub1-Ub heterodimers were used for localization and identification of sites of attachment.

Tandem instrumentation has allowed for the analysis of branched proteins at high resolution, high mass accuracy, and extensive activation. Still, manual interpretation of

complex spectra that arise from this analysis is not easy. As such, a collaboration with Dr. Catherine Fenselau's mass spectrometry lab at UMDCP was undertaken in order to further investigate enzymatically synthesized Rub1-Ub branched proteins as well as novel Rub1-Rub1 proteins with Dr. Fabio Gomes performing the specific mass spectrometry experiments and analysis.

4.2.1 Top-Down analysis of $Rub1_{T72R}$-Ub_{D77}

A dimer of $Rub1_{T72R}$ and Ub_{D77} was enzymatically synthesized through Ube2K (E2-25K) as already described with Rub1 being recognized as either a possible modifier protein or anchor protein while Ub_{D77} could only be rendered as a potential anchor but not a modifier. After synthesis and purification, ProSight Lite was used to evaluate which protein was the anchor and which the modifier. Only a series of b/c ions can be mapped onto the modifier-protein whose C-terminal is conjugated, while series of both b/c and y/z ions are usually recognized in the anchor protein. In the case of $Rub1_{T72R}$-Ub_{D77} dimer, b/c ions were mapped for Rub1 with a few y/z ions detected as well while Ub showed both b/c and y/z ions (Fig 4.2). The small amount of y/z ions mapped to Rub1 are indications of a certain level of Rub1-Rub1 dimerization occurring during reaction which is not surprising given previously described assay results.

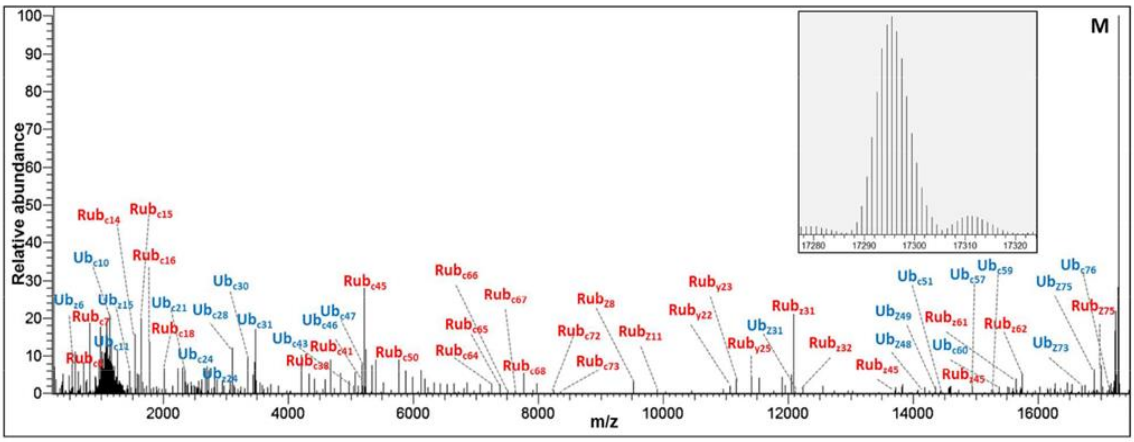


Figure 4.2: Deconvoluted product ion spectrum of the branched protein Rub1$_{T72R}$-Ub$_{D77}$. Fragment ions were assigned using ProSight Lite. The inset shows the deconvoluted isotope cluster of the molecular ions. A number of y/z ions are detected for Rub1 indicating Rub1dimerization[106].

The fragment ions in the complete spectrum were then mapped to the sequences of both unmodified proteins. During this process, fragment ions in the MS/MS spectrum were mapped onto dimer structures in which attachment is tested at each lysine residue in the anchor protein. The correct linkage site was assigned based on quasi-manual interpretation and consideration of the number of peptide bond cleavages characterized as a per cent of the total available. Any mutations were confirmed by molecular masses, and their positions located by dense fragmentation or localized within each branch by incomplete fragmentation. After all this, the optimal fragmentation map established the site of attachment (Fig 4.3).

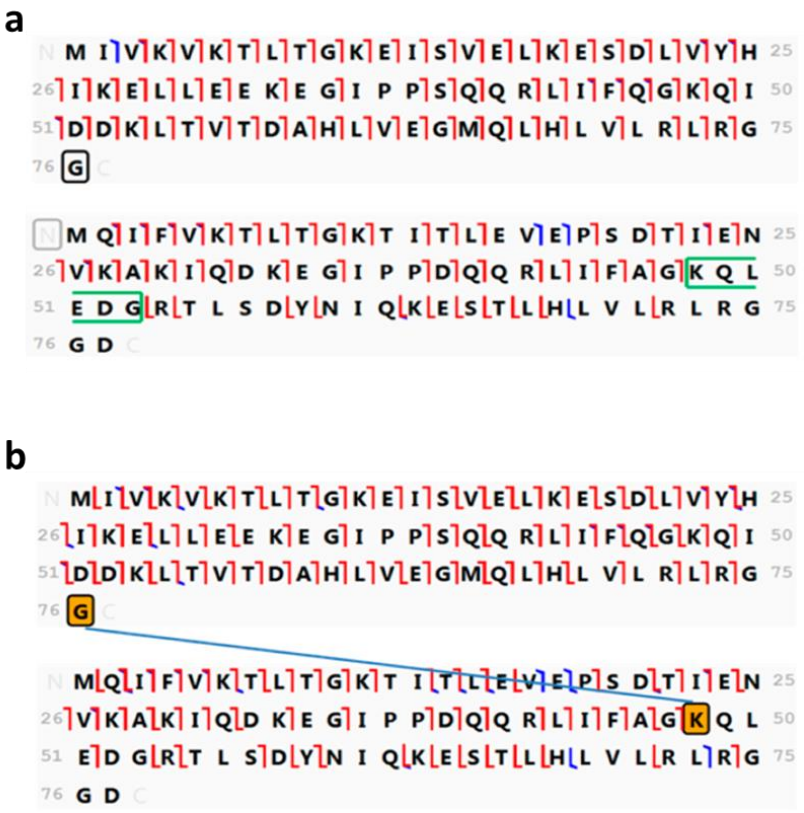


Figure 4.3: Graphic fragmentation maps interpreting the tandem mass spectrum of the conjugate formed between Ub_{D77} and $Rub1_{T72R}$. a) Foundation maps of the two proteins that constitute the branched protein (upper: $Rub1_{T72R,}$ lower: Ub_{D77}); **b)** Final fragment-ion map and structure assigned as the branched protein $Rub1_{T72R}$-Ub_{D77} linked through lysine-48[106].

Given the pattern observed, the assay-based conclusion of Rub1 being a chain terminator was supported in that the heterologous dimer was a rubylated ubiquitin and not a ubiquitinated Rub1. In the map shown, a gap between the b/c and y/z series when they are mapped onto the sequence indicates the site of the modification between the two proteins (lysine-48 in this case outlined in green). PyMOL was utilized to generate a visualization of the lysine-48 linked heterodimer structure of $Rub1_{T72R}$-Ub_{D77} interpreting the results obtained from the tandem mass spectrometry results (Fig 4.4). Of course, further work

needs to be done in utilizing crystallographic or NMR based analysis in order to make any definitive conclusions on the structural arrangement of this dimer.

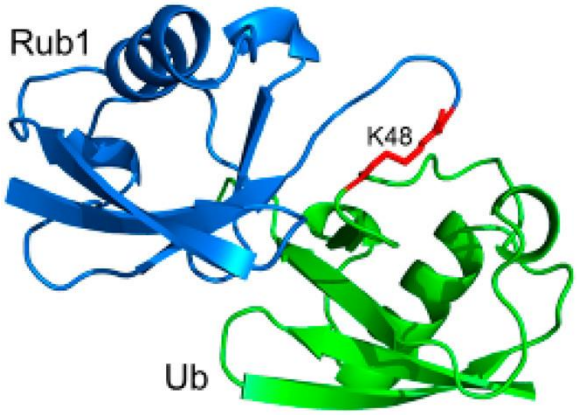


Figure 4.4: PyMOL generated image interpreting the tandem mass spectrum of the conjugate formed between Ub_{D77} and $Rub1_{T72R}$. This does not represent actual structure of the heterologous dimer[106].

4.2.2 Top-Down analysis of $Rub1_{E63K/T72R}$-Ub_{D77} and $Rub1_{K4F/T72R}$ homodimer

To further utilize the strategy of top-down analysis of branched proteins, a $Rub1_{E63K/T72R}$-Ub_{D77} heterodimer and the lysine-48 linked Rub1 homodimer of $Rub1_{K4F/T72R}$ was examined. A similar strategy was applied in synthesis and mass spec analysis of these dimers as described earlier.

As expected, in both cases, a dimer with linkage at lysine-48 was observed. These results explicitly showed that neither of the mutations prevented activation by ubiquitin E1 nor E2 conjugation on either Ub or Rub1. The fact that E2-25K recognized $Rub1_{K4F/T72R}$ as both the modifier and the anchor protein was a welcomed confirmation of the assay-based research done so far.



Figure 4.5: Graphic fragmentation maps interpreting the tandem mass spectrum of heterologous and homologous dimers. a) Fragment-ion map and structure assigned as the branched protein $Rub1_{E63K/T72R}$-Ub_{D77} linked through lysine-48; **b)** Fragment-ion map and structure assigned as the homodimer $Rub1_{K4F/T72R}$ linked through lysine-48[106].

4.2.3 Top-Down analysis of $Rub1_{T72R}$ homodimer linked through lysine-11

As mentioned in a previous chapter, a Rub1 homodimer was found to be formed in small amounts when the protein was incubated with Ube1 (E1) and the E2 enzyme Ube2S which preferentially makes ubiquitin chains linked via lysine-11, with only a small percentage of lysine-63 linkages and some small possibility of making lysine-48 linkages[107]. As such, it was prudent to subject this dimer to the same top-down exploration that was shown to be useful and see if the dimers observed in reaction assays were indeed linked through the expected lysine pathways.

The foundation map (Fig 4.6, a) showed a gap at residues 52–57 between the b/c series and a sparse y/z series. This sequence contains one lysine residue at position 53 which is not present in ubiquitin. Given that lysine residues are the canonical sites of enzymatic attachment, the results indicate a formation of $Rub1_{T72R}$ homodimer linked

through lysine-53. As in previous examples, the result here demonstrates that the ubiquitin E2 (Ube2S) still recognizes Rub1$_{T72R}$ as both a modifier and the anchor (substrate) protein. However, as demonstrated, the resulting conjugation site is different than was expected for this particular ubiquitin homodimer. As already mentioned, wild type Rub1 lacks a lysine amino acid at position 63, and conjugation at either lysine-11 or lysine-48 was not detected in the Rub1$_{T72R}$ homodimer formed by this enzyme. Given that ubiquitin does not have a lysine at position 53, this linkage site appears to be unique for Rub1 conjugation by Ube2S. This suggests a novel function for Ube2S and opens the door for further exploration of the actions of this specific enzyme with different UBLs.

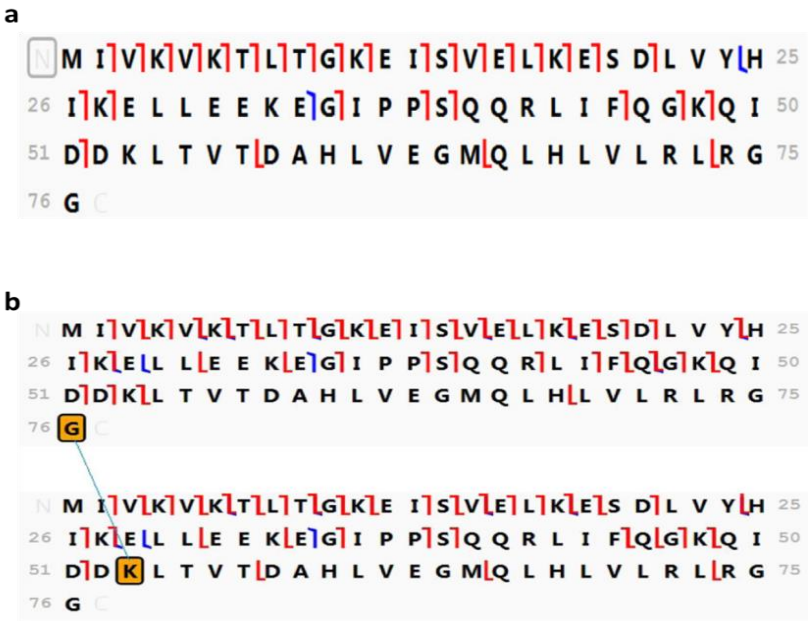


Figure 4.6: Graphic fragmentation maps interpreting the tandem mass spectrum of the conjugate formed between Rub1$_{T72R}$ homodimer. a) Foundation map of Rub1$_{T72R}$ homodimer. Gap at [52-57] between the b/c and y/z series indicates site of modification; **b)** Final fragment-ion map and structure assigned as the homodimer protein Rub1$_{T72R}$ linked through lysine-53[106].

4.2.4 Discussion and Summary

The various homologous and heterologous enzymatic assays performed with Rub1 and Ub have resulted in verifiable chain making potential using the ubiquitin conjugation system. Based on all the work done, the inclusion of T72R mutation in Rub1 increases the activation of Rub1 by ubiquitin E1. It appears that once Rub1 is activated, the bottleneck arises in the recognition of Rub1 by E2's as a potential proximal partner that could be added onto either a Rub1 or Ub unit that has already made a successful thioester bond with E2's active cysteine. Various mutation based experiments have not been able to pinpoint why this bottleneck exists requiring further exploration in this area. Nevertheless, based on reaction assay and mass spectrometry data, several heterologous and homologous chains (dimers) have been shown to be constructed with a number of novel Rub1 homodimers being made through linkages that were previously not shown. Further exploration into the extent by which E2s recognize Rub1 and the mechanisms involved in the limited chain making capability needs to take place to fully understand why Ub is selectively preferred for various signaling functions by its conjugation cascade.

Since the bottleneck appears to occur in E2 recognition stage, there exists an avenue of research that focuses on mutation based study of the E2 enzyme and how that affects interactions with both Ub and Rub1. Some of the common points of enzymatic interaction that exists for both Ub and Rub1 have been mentioned earlier but more work needs to be done to pinpoint areas that have not been yet discovered that could play a much more important role in the selective mechanism of E2 enzymes.

Chapter 5: Rub1 and Ub: structural delineation and functional outcomes

In all cases of Rub1 purification, the protein was consistently kept at pH 6.0 because of its sensitivity to pH changes while ubiquitin could withstand a much larger range of pH changes. Furthermore, during most of the reaction assays, it was noticed that Rub1 and many of its mutants tended to precipitate at 'normal' enzymatic reaction conditions (37°C, pH 8.0, ~16 hr reaction time). As such, in order to investigate if the stability difference between these two proteins could have an effect on reactivity and thus the resulting chain formation ability, temperature-based stability comparison of these two proteins was undertaken through the utilization of DSC, CD and NMR. The results of this examination are discussed hereafter.

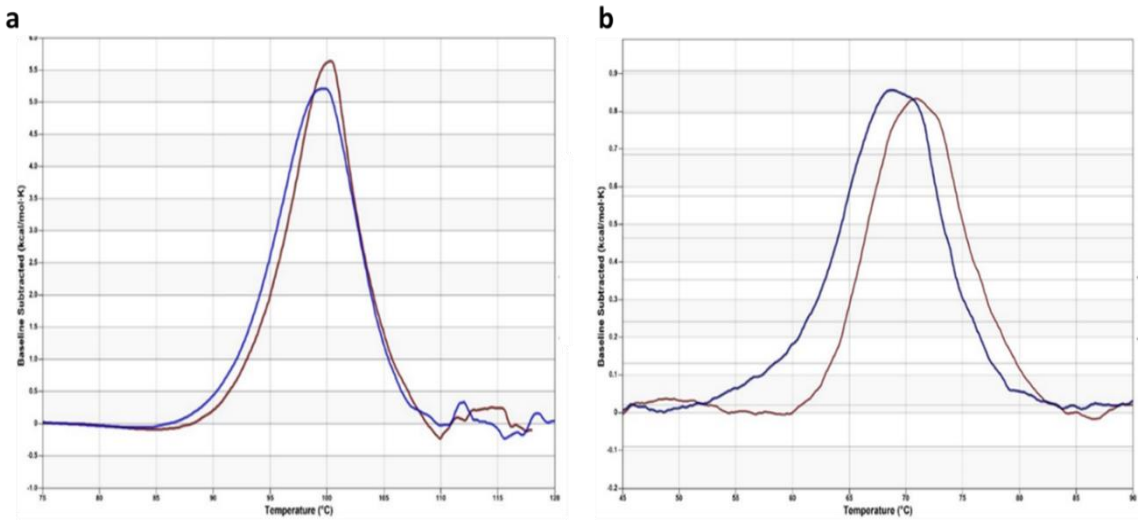
5.1 DSC experiment results of Rub1 and Ub

Among the various techniques used to measure thermal properties of proteins is differential scanning calorimetry (DSC). This thermodynamic technique measures heat capacity as a function of temperature and allows for determination of enthalpy and establish a link between temperature and physical properties of proteins. As such, by monitoring the change in phase transitions between Ub and Rub1, the enthalpy and thermal transition midpoints (T_m) for each transition can be obtained and a comparison of conformational change of these proteins caused by temperature variation can be made[108].

For this specific experiment, a collaborative effort was undertaken with Dr. Robert G. Brinson at IBBR where DSC for wild-type Ub and wild-type Rub1 was run. The

resulting data curves report a measurable difference in the temperature based structural integrity of the two proteins (Fig 5.1). Comparison at two different salt concentrations was recorded for both proteins at pH 6.0 which was found to be the pH were Rub1 was the most stable at. The difference in salt concertation did not show a significant effect on ubiquitin T_m but the results indicate that thermal stability of Rub1 is slightly more affected by the difference in salt concentration. Despite the fact that the pH during these experiments was done at close to the isoelectric point of both Rub1 (6.2) and Ub (6.7)[109], it appears that some electrostatic interactions within Rub1 are disrupted at higher salt concentrations indicating the relative sensitivity of Rub1 to different variables. In any case, the reported T_m difference between the proteins shows a significant thermal stability for ubiquitin at pH 6.0. The change in enthalpy also points out the relatively weaker van der Waals and hydrogen bonds with Rub1 as indicated by the much lower ΔH upon unfolding (Fig 5.1).

Aggregation caused during thermal denaturation prevented the calculation of precise thermodynamic parameters for both proteins. Even though it is beyond the scope of this dissertation work, further careful measurements through different techniques could be done to calculate said parameters. Nonetheless, in this particular case, the results show that despite the seemingly high similarity between these two proteins, Rub1 is evidently much more thermodynamically unstable than ubiquitin.



	Melting Temperature (°C)	ΔH (kcal/mol)
High Salt	99.33 (± 0.19)	137.6
Low Salt	99.80 (± 0.09)	137.60 (± 5.85)

	Melting Temperature (°C)	ΔH (kcal/mol)
High Salt	68.88 (± 0.05)	85.49 (±2.73)
Low Salt	71.18 (± 0.14)	91.60 (±2.21)

Figure 5.1: DSC results of (a) Ub and (b) Rub1 at pH 6.0. Data indicates more of a change with salt concentration on the stability of Rub1 as opposed to ubiquitin. High salt buffer included 130 mM NaCl (blue trace) while low salt buffer had no NaCl (red trace).

5.2 CD based temperature change measurements of Ub and Rub1

Another valuable spectroscopic technique that is useful in determining the unfolding of proteins as a function of temperature is circular dichroism (CD). CD is a measure of the difference between the absorption of left-handed and right-handed circularly polarized light reported in either the difference in absorbance of the aforementioned polarized light ($\Delta\varepsilon = \varepsilon_R - \varepsilon_L$) which gives the molar circular dichroism or in degrees ellipticity which is just circular dichroism that has been corrected for concentration. Molar ellipticity, [θ] is represented in deg·cm^2/dmol[110].

In comparing the secondary structure of Ub, Rub1 and Nedd8, the initial hypothesis was that all three proteins would show generally similar CD spectra that indicated protein

structures that were characteristically helical in nature. Surprisingly, Rub1 and Need8 exhibited a much-diminished negative dip at 222 nm which suggests a much lower percentage of helical structure (Fig 5.2). As would be expected, the two UBLs had the most similarity in their CD spectra but both of them were starkly different from that of Ub. These results were in contradiction with the structural data that is available indicating a similar number of helical amino acids for all three proteins. As such, the unexpected initial CD measurements added to the question of structural differences between Ub and the other UBLs that might play a role in the selective reactivity of Ub as a prolific tagging protein. To independently verify the result from DSC, CD based temperature dependence experiments were performed for Rub1 and Ub.

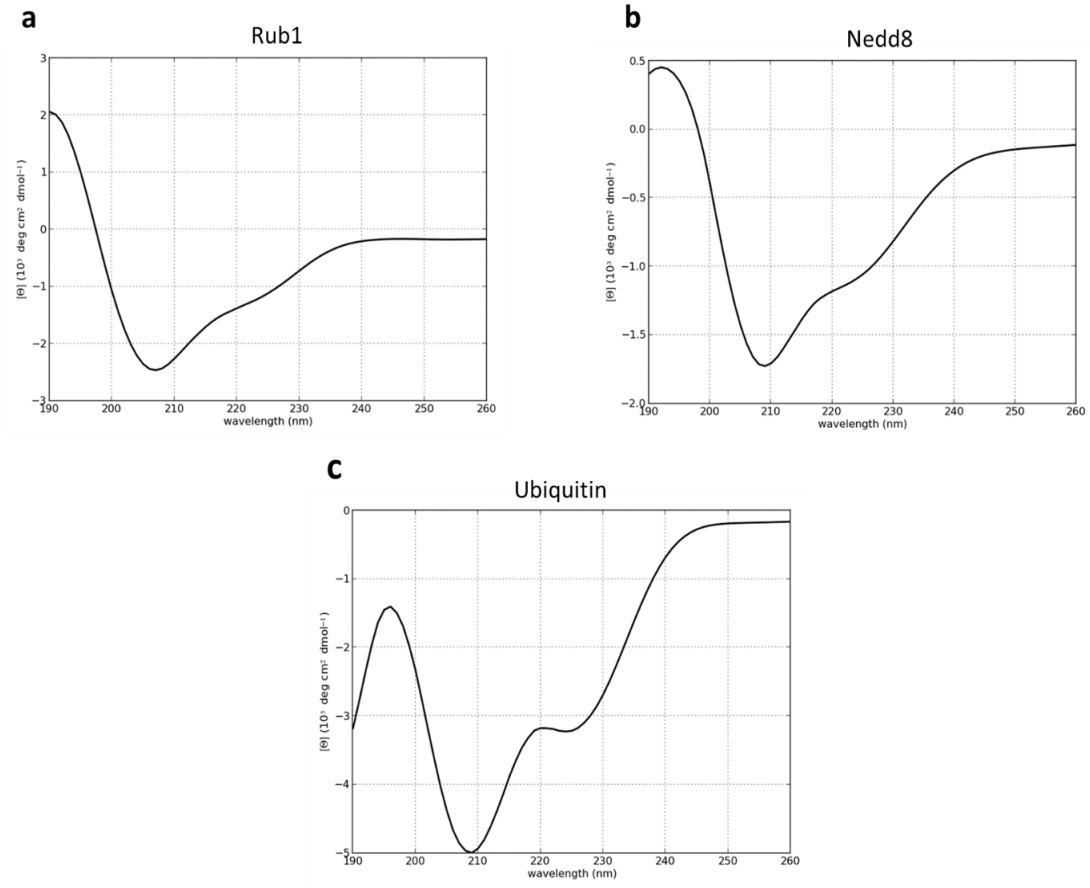


Figure 5.2: CD spectra of (a) Rub1; (b) Nedd8 and c) Ub at 25 ^{O}C and pH 6.0. The significantly diminished dip at 222 nm is an indication of a structure that is less α-helical in nature. These spectra show that there is some extent of structural difference between all three proteins with the two UBLs showing the most similarity.

These CD measurements at different temperatures were conducted from 20°C to 90°C in 20 mM sodium phosphate buffer, pH 6.0. As expected, the CD curve for Ub did not show much of a difference at both characteristic wavelengths of 222 nm and 208 nm. On the other hand, Rub1 started to show a change at 208nm by 70°C. By 77.5^{o}C, the dips at both 208 nM and 222 nm had shifted dramatically corroborating the results from DSC (Fig 5.3; b). By the end of the measurements, precipitation of Rub1 was observed but a re-

measurement after cooling the sample showed a spectrum that was relatively similar to that of the folded protein at the beginning of the experiment. This suggests that Rub1 can re-fold even after denaturation due to temperature increase. By plotting the changes in the mean residue ellipticity of Rub1 at 222 nm as a function of temperature, we were able to obtain an apparent melting temperature (T_m) of ~77.8^oC which is higher than the T_m measured through DSC. The T_m of Ub was not reached within the temperature range of the experiments.

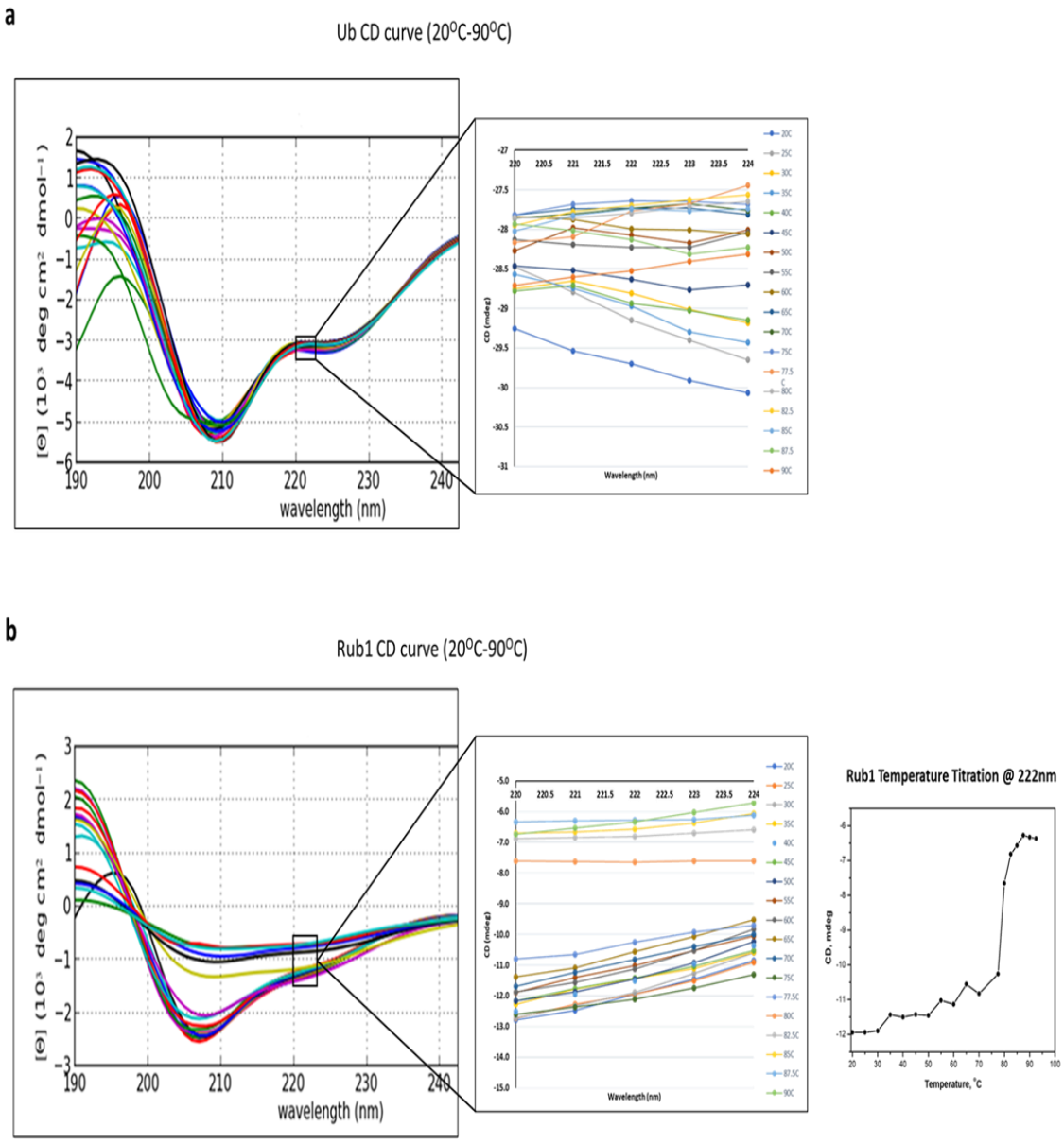


Figure 5.3: Temperature dependent CD spectra of (a) Ub and (b) Rub1 at pH 6.0. There was no significant change in the Ub spectra after reaching end point. Spectra of Rub1 shows a significant change by the end of CD measurements. Inset shows the changes at 222nm during the changes in temperature with the melting profile of Rub1 shown.

5.3 NMR based temperature change measurements of Ub and Rub1

Another approach available for the assessment of temperature-based stability of proteins is variable-temperature NMR analysis. This technique gives insight into the dynamics and structure of proteins, chemical shifts, and attenuation of backbone amide protons due to temperature changes. More specifically, the linear dependence of these amide proton chemical shifts on temperature can give useful information about the intramolecular hydrogen bonds in the protein of interest[111]. This correlation of temperature coefficients (TC) (i.e., slopes from fitting chemical shift vs. temperature data to a linear model) of amide protons to hydrogen bond strength is an established method to probe hydrogen bond status of individual amides[112, 113]. In some cases where chemical shifts exhibit deviations from a linear model of chemical shifts due to temperature (curving), the result could be an indication of an exchange between distinct conformations of the protein[114, 115]. Generally, most shifts exhibit linearity until the highest temperatures are reached. It has been demonstrated that the temperature at which curvature might start to occur is around 15° below melting temperature (T_m) of the proteins[113]. It is commonly accepted that temperature coefficient values that are more positive than −4.5 ppb/K are generally good indicators of hydrogen bonds, while those more negative than −4.5 ppb/K tend to be associated with non-hydrogen bonded or solvent exposed amides[116, 117]. Furthermore, signal attenuations and observable protein precipitation at different temperatures can be interpreted as a measure of temperature-dependent loss of structure and in this particular case, give corroboration to the apparent T_m results from DSC and CD.

In previous works, temperature coefficient measurements were utilized to assess the thermostability differences of a protein either when a ligand is introduced or when

mutation is incorporated into the protein of interests. The difference in temperature coefficients can give insight into specific areas within the protein that are affected by the introduction of said ligand or mutation[118]. For the purpose of this project, since Ub and Rub1 are very similar proteins, temperature coefficient measurements were used as one of the methods to try and understand the various differences in the biophysical characteristics of Rub1 and Ub. As such, temperature variation NMR experiments were performed, and the amide proton and nitrogen chemical shifts were measured through a series of $^{1}H/^{15}N$ heteronuclear multiple quantum coherence (SOFAST, HMQC) spectra collected at increasing temperatures. The comparisons between Ub and Rub1 were based on the chemical shifts of these two proteins which are affected by local magnetic environments of nuclei that are directly related to temperature changes.

In the process of collecting data from temperature-based NMR measurements, the magnetic field needs to be homogeneous and stable. As such, in solution NMR, "locking" adjusts the strength of magnetic field to preserve a constant resonant frequency which compensates for the temperature dependent shift of D_2O nuclei, which is used as a lock signal. Since all non-water residue peaks shift by an amount equal to the temperature dependence of D_2O, there needs to be a way to find out the real position of each peak during temperature changes. As such, to overcome this "deuterium lock" temperature dependence, spectra were referenced to 4,4-dimethyl-4-silapentane-1-sulfonicacid (DSS) which is a standard used since it has negligible temperature change dependence[119].

5.3.1 NMR spectra based analysis of Rub1 and Ub thermostability

Based on the attenuation of the 2D spectra of each protein during temperature increases, the NMR measurements indicated that wild-type ubiquitin is more thermostable compared to Rub1. For Ub, temperature measurements were taken from 280K (6.85^{o}C) to a maximum of 370K (96.85^{o}C) with the spectra of ubiquitin showing global attenuation by the endpoint (Fig 5.4; a). Further increases in temperature were not applied given the lack of any signals that were visible. Upon lowering the temperature back to room temperature of 296K (22.85^{o}C), ubiquitin signals retained their original positions showing that the temperature based structural changes in Ub are reversible.

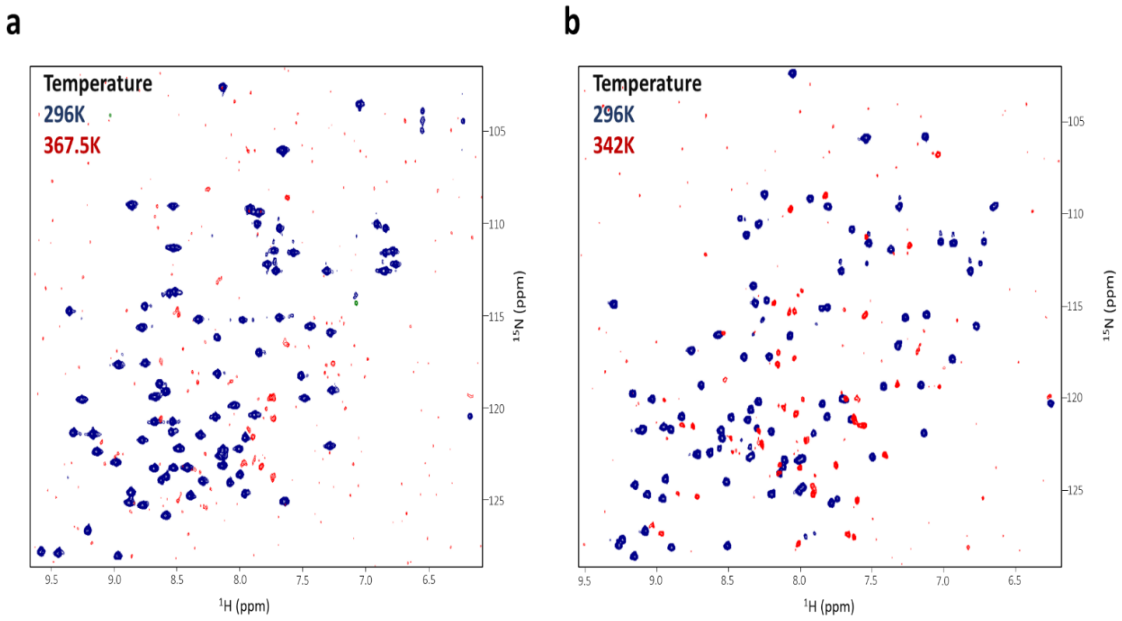


Figure 5.4: ^{1}H-^{15}N SOFAST-HMQC spectra of (a) Ub at 296K (blue) and 367.5K (red) and (b) Rub1 at 296K (blue) and 342K (red). Notice the signals of both Ub and Rub1 have either diminished greatly or have fully attenuated at their respective final temperatures. Rub1 precipitation began at 340K (66.85^{o}C) while that of Ub began ~362.5K (89.35^{o}C). Both measurements were done in 20mM NaP, pH 6.0 buffer.

On the other hand, the thermostability of Rub1 was predictably lower than ubiquitin. In a similar temperature dependent experiment performed, Rub1 showed beginnings of precipitation when the temperature reached 340K (66.85^{o}C) and the measurement was stopped at 342K (68.85^{o}C) because of further increase in the precipitation of this protein and the loss of signal intensity (Fig 5.4; b). It should be mentioned that despite the low thermostability of Rub1 and attenuation of several peaks, spectral spread of any remaining unprecipitated protein was preserved suggesting that a higher temperature could be reached until full denaturation of Rub1 occurs. Additionally, folding of this protein was also shown to be somewhat reversible given the return of signals to their original position when temperature was reduced to room temperature. Nevertheless, the peaks that returned were much lower in intensity given the extent of precipitation caused by unfolding of the protein at 342K. In fact, for both proteins, the spectral spread of signals was maintained at higher temperatures which suggests that the overall structure of the proteins was still intact. Furthermore, the reduction in signal intensity could be a result of both precipitation (reducing the total amount of structured protein) and fast exchange of protons with bulk water.

The results from both NMR studies gave an indication of the melting temperature for each protein which seemed to agree with the results from DSC. Of course, more rigorous quantitative experimental studies need to be done to get a full picture of the difference between the proteins but for the limited purposes of this study, the similar results from NMR, DSC and CD are a good indication of the thermodynamic differences between Ub and Rub1.

5.3.2 Temperature coefficient-based analysis of Rub1 and Ub

As mentioned already, measurement of temperature coefficients (ppb/K) for the chemical shift of the amide hydrogens (^{1}H) of each residue can be used to identify temperature dependent loss of structural integrity within secondary structures as internal H-bonding is disturbed. In this study, as temperature was increased, cross peaks for both Ub and Rub1 tended to move downfield in both dimensions in a linear fashion prior to DSS referencing (Fig 5.5; a). The true temperature dependence of the amide proton chemical shifts is displayed after referencing to DSS (Fig 5.5; b). Intramolecular hydrogen bonds extend when temperature is increased which weakens amide proton electron polarization causing less de-shielding, thus causing up-field movements of chemical shifts which is shown for both Ub and Rub1. As such, stronger intramolecular hydrogen bonds cause chemical shifts to be less affected by temperature compared to solvent-exposed or non-hydrogen-bonded amides and measuring linear temperature coefficients of the amide protons of both proteins would allow for assessing the difference in hydrogen bond strengths between Rub1 and Ub. As indicated earlier, temperature coefficients for each protein were measured with the acknowledgement that several signals have the propensity to either attenuate or show a semblance of curvature that is too subtle to be considered 'real' within 15°C of T_m. As such, much of the analysis was done at T_m-15°C. The relevance of this analysis is to try and gain more understanding on how the two proteins differ from each other at and around physiological temperatures.

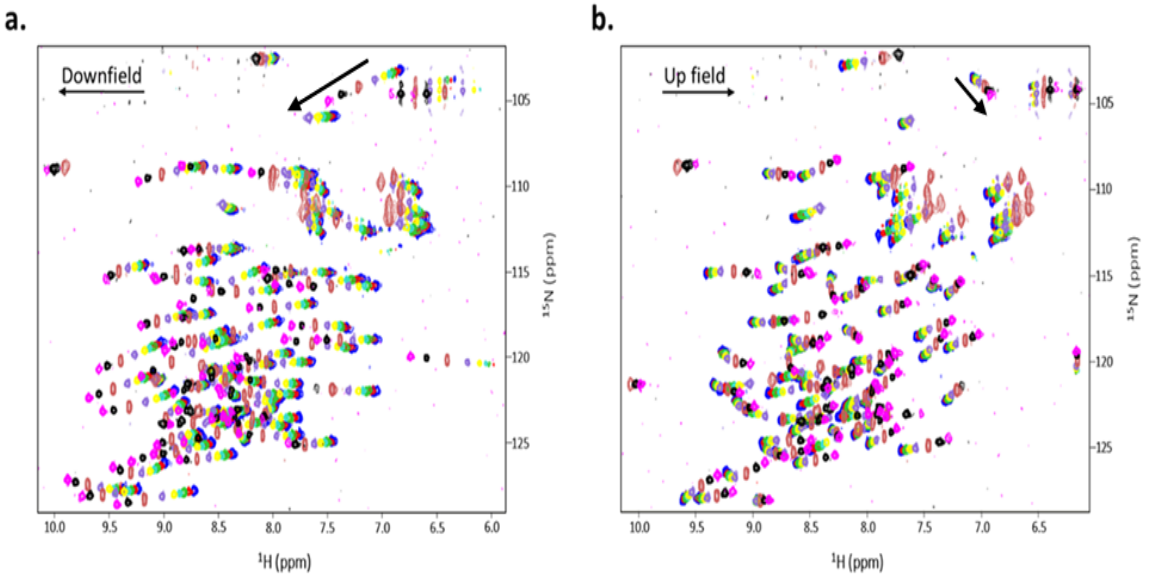


Figure 5.5: Shift of ^{15}N-Ub signals collected over a range of temperatures; a) Un-referenced Ub HMQC spectra. The deuterium lock artefact is evident with signals systematically shifting downfield in the proton dimension; **b)** HMQC spectra referenced to DSS, with the up field shifting of signals in the proton dimension showing true temperature dependences of amide chemical shifts. DSS referencing was done using SPARKY.

The linearly fitted experimental data for each residue in Ub and Rub1 show that the majority of all residues have linear temperature dependence. Still, there were residues in both proteins that showed systematic deviations from linearity (Fig 5.6 & Fig 5.7 respectively). This is expected for residues that are solvent exposed and thus experience some form of fast conformational exchange due to temperature change. It is possible that a few of these types of residues in both proteins may not be involved in a stable secondary or tertiary interaction based on where they are situated in the protein structure (i.e., loops, end of secondary structures). On the other hand, there are other residues with observable curvature that are found within beta strands and α-helices that suggest temperature dependent disturbance in polar contacts lending to possible conformational dynamics. Interestingly, there are more minimal deviations from linearity observed in Ub compared to Rub1 where fewer residues with deviations are observed with a number of them showing

pronounced curvature in their plots. This could be an indication that Ub experiences more temperature based local perturbations while Rub1 experiences much more pronounced regional dynamics in the areas that are being affected by temperature change.

In these same experiments, amide nitrogen temperature coefficients for both proteins were also calculated and although a similar mechanism of interpreting these TC measurements to structural dynamics of proteins has yet to be successfully demonstrated, it is interesting to compare the amide nitrogen TC plots that were obtained for Ub and Rub1with their corresponding amide proton plots. In both proteins, the nitrogen TC's showed several residues with positive slopes and well pronounced deviations from linearity that was not observed in the proton TC plots. More work needs to be done to determine the relationship (if any exists) between amide nitrogen temperature coefficients and structural integrity of proteins which is beyond the scope of this specific study. Nevertheless, the data obtained from these experiments can be used for future such endeavors. Following is a slightly more detailed discussion focused on the residues exhibiting curvatures in Ub and Rub1 based solely on their amide proton temperature coefficients.

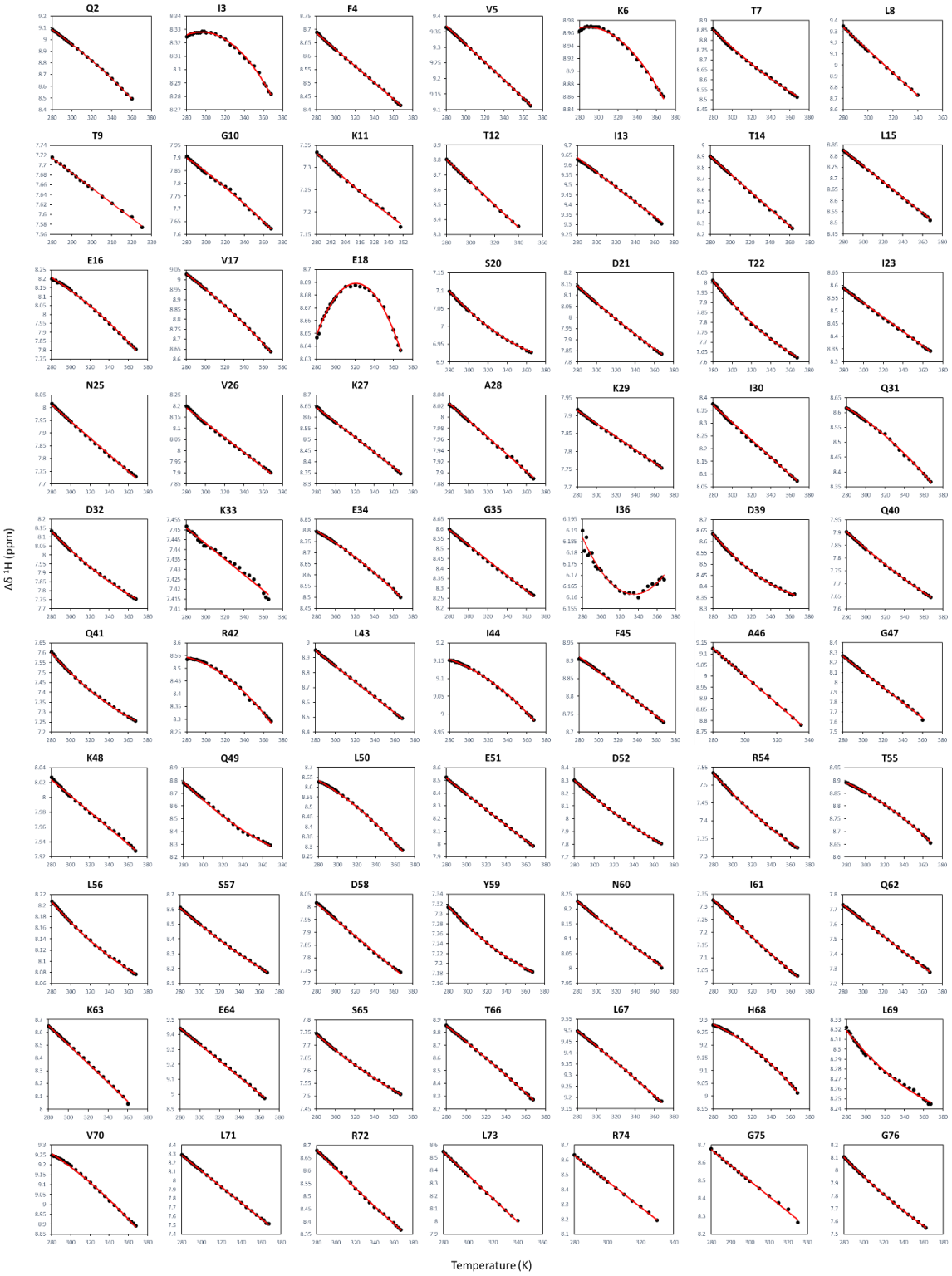


Figure 5.6: Temperature dependent chemical shifts of the backbone amide protons of Ub at pH 6.0 from 280K-367.5K. Notice the curvature in a few of the residues. Several signals attenuated at high temperatures and are not used in hydrogen bond strength analysis. Corresponding chemical shift behaviors of amide nitrogens is shown in Figure A2.

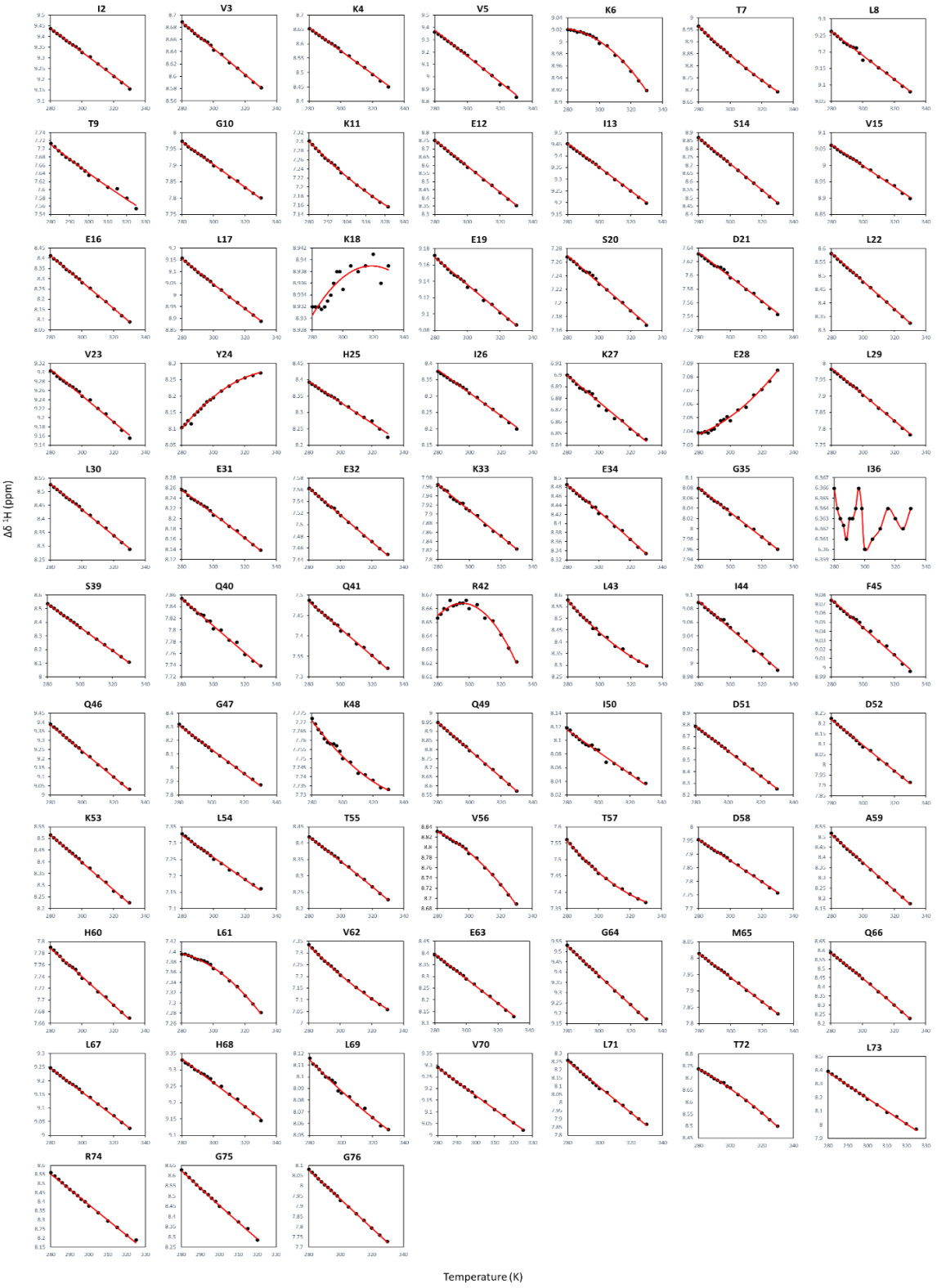


Figure 5.7: Temperature dependent chemical shifts of the backbone amide protons of Rub1 at pH 6.0 from 280K to 330K. Notice the positive coefficient in a few of the residues. Several signals attenuated at high temperatures and are not used in hydrogen bond strength analysis. Corresponding chemical shift behaviors of amide nitrogens is shown in Figure A3.

5.3.3 Temperature dependent deviation from linearity in specific residues

As already stated, deviations from linear temperature dependence at the highest temperature measurements could be an indication of protein unfolding[114]. For Rub1, the NMR data appears to support this claim based on the results whereby a number of Rub1 signals at high temperatures fully attenuated in tandem with an increase in the level of precipitation of the protein. Nevertheless, a gradual curvature in temperature coefficient plots for many of the residues in both proteins before T_m is reached indicate some type of fast exchange between at least two species in said regions over the temperature range studied. This is mostly observed in Ub with curvature more pronounced when the temperature is below ambient temperature (<300K). Overall, it appears more residues in Ub have visible deviation from linearity compared to Rub1.

Figure 5.6 and 5.7 show variation in the extent of curvature exhibited by several amino acids in both Ub and Rub1 respectively. Interestingly, out of all residues in both proteins, Y24 and E28 in Rub1 have a distinctive positive slope indicating a strong H-bonding interaction. In fact, these two residues are in the main α-helix structure with a backbone H-bond between them in addition to H-bonding with other residues within the helix (Fig 5.8; a). In the data analysis for Rub1, K18 also seems to show a small positive slope but the changes in the shift of the residue peaks in ppm is so small that not much information can be gleaned from this data point.

Interestingly, in ubiquitin, E24 and A28 are found in the same helix region with similar H-bond contacts with residues in their vicinity and yet, there is no positive temperature coefficient nor any observable curvature for A28 to suggest any non-local conformational dynamics experience by Ub in this region over the temperature range

studied. Since E24 is in an intermediate conformational exchange at room temperature, its signal is too broadened and does not appear in Ub spectra until high temperature is reached. As such, no direct information can be gleaned to make any assessment of its behavior. But, in an interesting manner, E18 which is also found in a loop region of Ub shows minimal overall shifts, but the profile of these shifts shows a negative parabola suggesting temperature based conformational variability that is potentially different from Rub1.

Another main curvature that is observed for both Ub and Rub1 is in residue R42 in the β3 region. In ubiquitin, R42 forms a backbone H-bond with V70 in the adjacent β5 strand. This is similar to contacts observed in Rub1 with the exception that R42 in Rub1 makes additional H-bonds with S39 and Q41 which might lend to the slightly more exaggerated curvature observed in R42 of Rub1 as more H-bonds are potentially disturbed during temperature change. Interestingly, none of the contact partners of R42 in Rub1 exhibit any deviation from linearity while V70 in Ub shows some change over the temperature range. Furthermore, in a crystal structure of Ub (PDB:1D3Z), R42 in Ub can potentially make H-bonds to R72 which could potentially have some effect in how Ub is selected (especially by E1 enzyme). Nevertheless, the TC measurements do not show deviation from linearity for R72 while conversely, Rub1 shows a slightly curved slope for its T72 amino acid which also could be explained by the two H-bonds it makes with Q40 that are susceptible to temperature change effects. These data seemingly show that the number of H-bonds that are made might have an additive effect in the extent of curvatures observed in the chemical shifts of the proteins during temperature variation.

Interestingly, in both Ub and Rub1, K6 forms backbone H-bonds with both L67 and L69. But the TC curvature results show that in Ub, K6 and L69 show curvatures while

in Rub1, only K6 shows curvature in its profile. A similar feature is repeated where both proteins form H-bond interactions between I44 and H68 and yet only the residues in Ub show curved profiles. The difference observed here between the two proteins is harder to explain. Given that several residues within the β-strands are involved in intra-strand H-bonds, it is possible that several of the curvatures observed in one protein over the other is the product of a general H-bond disturbance in the respective regions of interest.

Furthermore, both proteins share the same hydrophobic patch residues with V70 located between L8 and I44 in the structure and the general changes in curvature observed for residues in β1 and β2 are indicative of the overall intra-strand H-bonds that are affected during temperature change. Nevertheless, the differences observed in the dynamics of the hydrophobic region can give some insights into the difference in how these two proteins interact with ligands. Exactly how that could be is still an unanswered question. This is further complicated by the fact that comparing the two proteins shows an observable difference in the slope of the signals that showed curvatures. Some signals (e.g., K6, R42) in both proteins show negative slopes that curve inward while some signals (e.g., Y59 in Ub; L43 in Rub1) curve outward. Why this phenomenon occurs and how it pertains to the question of temperature based structural dynamics seems to be an addition to the multitude of queries that this course of study tries to address.

Figure 5.8 shows Rub1 and Ub structures with residues that exhibit temperature-dependent curvature. If temperature driven, major conformational changes were to occur, they would be expected in the structured regions of each protein. Of the residues listed, the ones that are located in unstructured regions of each protein may not be as involved in stable secondary or tertiary interactions. These residues are most likely involved in

H-bonds with solvent and thus are not as informative in terms of temperature dependent structural change. Figure 5.8 shows an obvious difference in the extent by which either protein responds to temperature changes. A good portion of the β-sheet face in Ub shows potential structural flexibility which could potentially indicate why Ub is more stable at higher temperatures. It is possible that Ub has a much more predilection to 'bend' rather than 'break' upon temperature changes while Rub1 is more conformationally limited and thus unable to withstand much higher temperature changes. This could give some important insight into the difference in structural integrity between the two proteins. As such, any future studies with these two proteins could potentially focus on mutational work on the residues that are located in the main structured sections of the proteins that have been identified to respond differently to temperature changes.

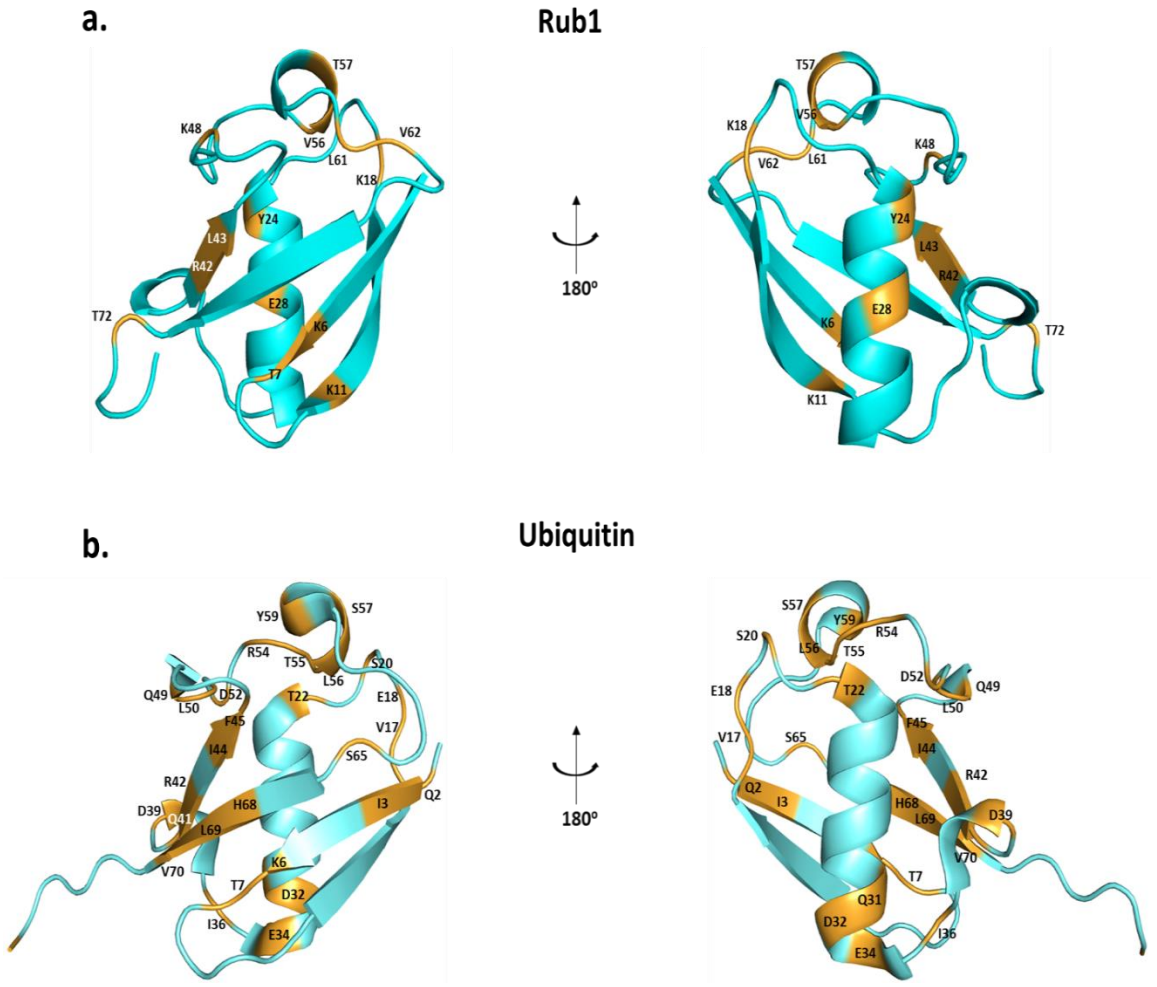


Figure 5.8: Residues with curved amide proton temperature dependence identified in **a)** Rub1 (NMR based in-house structure) and **b)** ubiquitin (PDB: 1D3Z)[120]. All residues that show curvature are labeled and shown in orange. Notice the extensive number of curved residues in Ub with minimal curvature in the main helical region for both proteins.

It should be noted that temperature-dependent conformational changes that do not result in curvature of NMR signals are possible. Care must be taken so as to not extrapolate any curvature where it does not exist as well as in the explanation of curvatures that are observed. Additionally, it has been reported that temperature coefficients are not necessarily good indicators of the strength of hydrogen bonds to amides but rather a result of local melting of structures that occurs due to thermal expansion of their intermolecular

hydrogen bond lengths[121]. As such, to test this, an attempt was made to relate the temperature coefficient results to the potential thermostability of the secondary structures of Rub1 and Ub without relying on observed curvature for any extant explanations. The results are discussed in the next section.

5.3.4 Amide proton temperature coefficient effects on local secondary structure stability of Rub1 and Ub

The magnitudes of Rub1 temperature coefficients for specific residues were frequently greater than those of corresponding Ub residues indicating that the magnetic environments of the amide protons change more with temperature (Fig 5.9). Since the analysis performed was restricted to changes that occur up to T_m-15°C, ensuring that each protein did not lose its tertiary structure, the spectra of Rub1 and Ub remained well resolved with no attenuations within the 280K-330K and 280K-360K temperature ranges, respectively. This indirect comparison gives insight into the difference in thermostability as each protein reaches its corresponding T_m. Thus, the total average TC calculated for Rub1 at 330K and ubiquitin at 360K (−4.25±0.33 ppb/K and −4.00±0.28 ppb/K respectively), which correlate to the global stability of the proteins. Overall, the global TC's for the proteins were very comparable suggesting a similar response to temperature changes with Ub showing a slight edge in stability when looking at specific structured regions as described hereafter. Generally, both proteins show a similar, relatively large TC in their C-terminal tail (residues 73-76) suggesting that the amide protons in these highly flexible regions are essentially involved in hydrogen bonds with solvent.

Ubiquitin and Rub1 share a similar beta-grasp fold with five β-strands, one α-helix and 2 short 3_{10} helices. Since most of the inter-residue hydrogen bonds occur within these secondary structure elements, the changes in local stability upon temperature changes could shed light on the difference between the two proteins in terms of the integrity of their respective secondary structures upon external pressures. Therefore, temperature coefficients for each protein were plotted against those specific amides in the secondary structures which are involved in the main intramolecular hydrogen bonding (Fig 5.10; a).

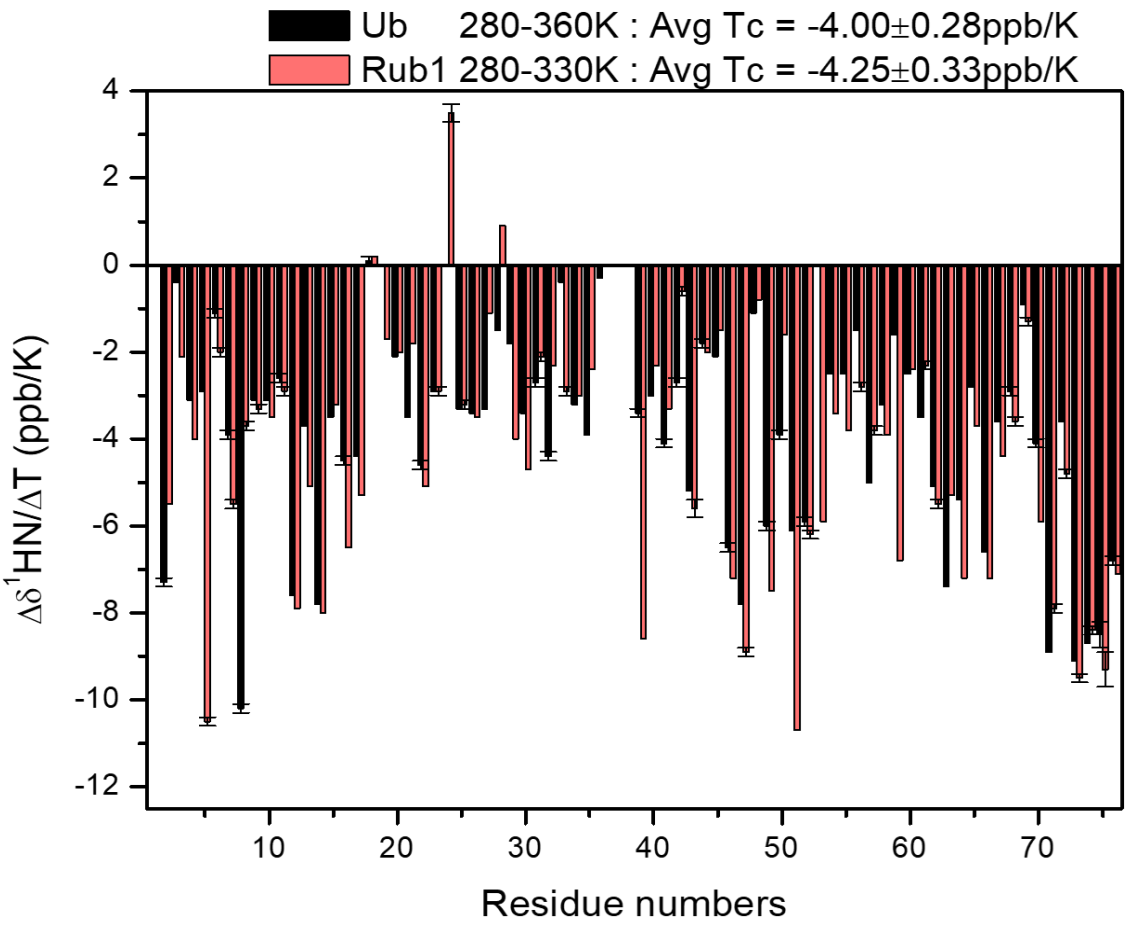


Figure 5.9: $^{1}H_{NH}$ temperature coefficients plotted for ubiquitin residues (black bars) and Rub1 residues (salmon bars) with temperature points ending at their respective T_m-15°C. Y24 and E28 in Rub1 exhibit a positive TC indicating more H-bond contacts than in Ub's E24 and A28 residues.

This assessment illustrates the differences in residues that are involved in H-bonding and gives potential targets for further mutation-based studies. Despite some amides in Ub having more negative TC's, further evaluation of the local stability of secondary structures through the averaging of the coefficients for each section shows areas in the secondary structures which give a better insight in the TC of structured regions within each protein (Fig 5.10; b). It appears that in regions where there is a big difference in TC, Ub shows a much more positive TC and thus suggests a sturdier structural integrity. One of the main differences observed in TC between the two proteins is in the first 3_{10} helix where there is a larger negative TC observed for Rub1. Nevertheless, it should be noted that this region shows a 3_{10} helix in the in-house NMR structure of yeast Rub1 used for this study but is not found in the crystal structure of Rub1 from Arabidopsis thaliana. As such, given the lack of crystal structure for yeast Rub1 and thus the possibility that this reflects a TC in Rub1 for an unstructured loop region, the data point is not considered when making preliminary conclusions after analyzing the difference in coefficients between Ub and Rub1 (Fig 5.10; c*).

To better show the regions where there is a potential increase in local stability for Rub1, the difference in secondary structure temperature coefficients ($\Delta TC = Rub1_{TC} - Ub_{TC}$) between the proteins was calculated, whereby the more positive the value, the more stable that region is in Rub1 protein. It is interesting to note that the regions of α-helix, β_3 and β_4 in the structure of Rub1 seem to be more stable than Ub (Fig 5.10; c). Given the positive slope of Y24 and E28 in Rub1, it is likely that the relatively stronger H-bonding among these residues correlates to the stronger stability in the helical region (Fig 5.8; a). It is harder to see why β_3 shows much of a difference given both proteins

show a level of conformational dynamics in that region while the difference in β_4 could potentially be attributed to the fact that this small region is folded in Ub while Rub1 has an unfolded section which might mean there is more H-bonding with water for this region and thus less perturbation in the main body of Rub1.

By looking at the regions in each structure with residues that show curvature and comparing them to secondary structure regions in each protein that show stability, there seems to be a correlation between temperatures based local dynamics and overall structural integrity in that more conformational dynamics allows for more flexibility in the respective structures. Specific studies on how this correlation plays out in terms of overall protein stability and function remain to be explored.

Still, in an attempt to further see if the TC based analysis could be used in conjunction with previously mentioned cassette based mutational studies, regions that were selected based on cassettes between the two proteins where subjected to similar TC assessments as mentioned earlier. The results show an overall stability for Ub with only two cassette regions showing Rub1 as more stable (Fig S2). By overlapping these two TC results (secondary structure vs cassettes) it might be possible to identify specific residues to perform mutational studies upon and further understand the difference between structure and function of Rub1 and Ub. Individual coefficient data for both proteins is provided in Table A3 in the appendix.

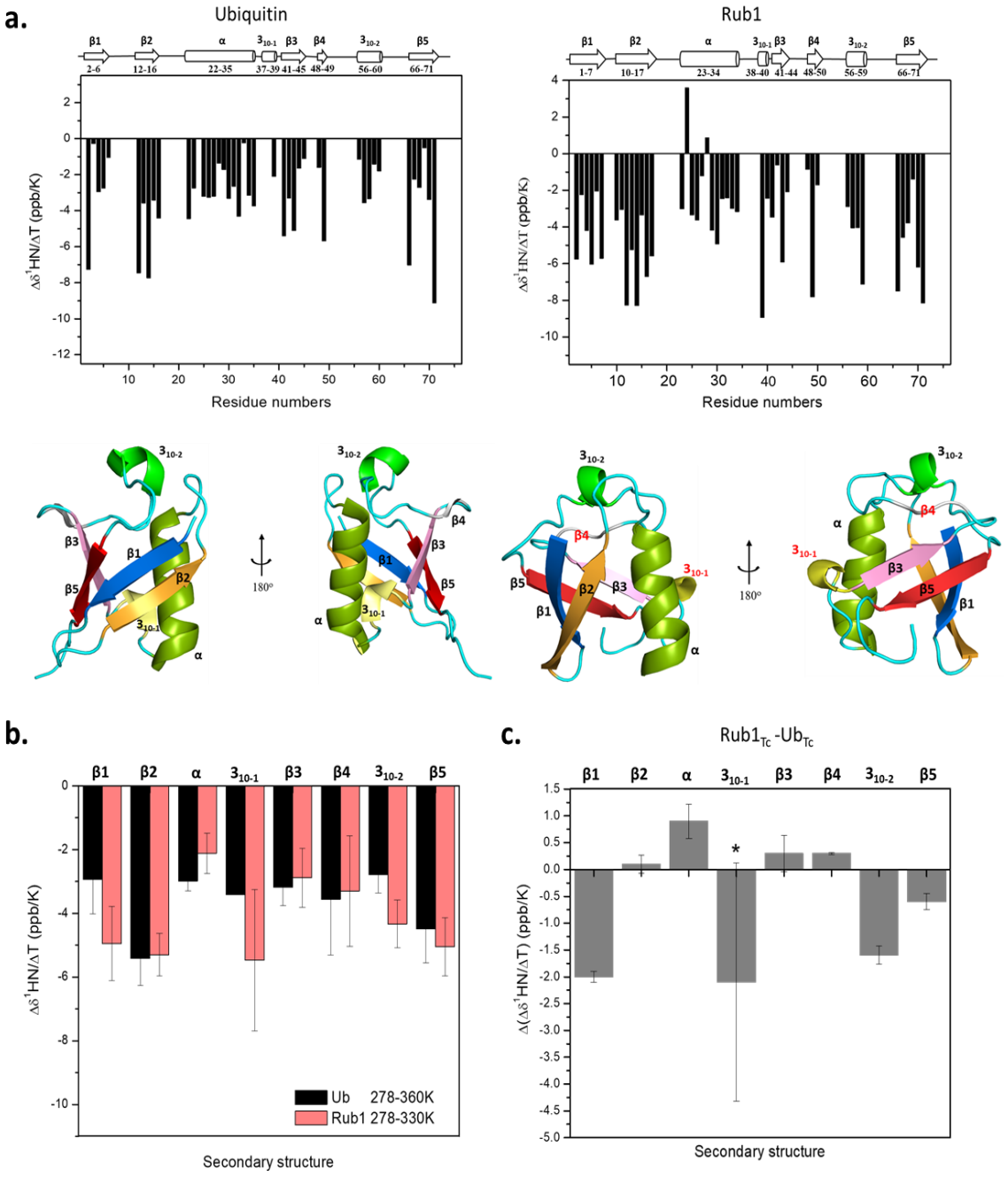


Figure 5.10: $^{1}H_{NH}$ temperature coefficients plotted for **a)** secondary structure residues of ubiquitin (left) and Rub1 (right). Secondary structures for both proteins are represented in different colors (bottom); **b)** Averaged secondary structure temperature coefficients for Ub and Rub1 from 278K to respective T_m-15°C temperatures; **c)** Difference in amide temperature coefficients between Ub and Rub1 averaged over the secondary structures. The two 3_{10} helices are represented as 3_{10-1} and 3_{10-2} for simplicity. In Rub1 structure, 3_{10-1} and β4 are represented with red to indicate the variation observed between available crystal and NMR structure of Rub1 protein. (Ub PDB: 1D3Z)[120] (Rub1: in-house NMR structure).

5.3.5 Discussion and Summary

In trying to understand the differences between Rub1 and Ub and to address the question of how Ub is selected to function as a robust signaling molecule, several experiments were attempted utilizing various techniques of investigation that are available. Many of the results have shown the extent in which the ubiquitin conjugation system can and does recognize Rub1 in its processes *in vitro* given mutational manipulations that are introduced in the system. Some of the mass spectrometry results indicated that even wild type Rub1 can be recognized by the Ub enzymes in a very limited manner without any specific selectivity for the type of linkages available. None of the experimental results have so far, provided a clear-cut demonstration for why the selection of ubiquitin over Rub1 both *in vitro* and *in vivo* is highly favored.

The $^{1}H_{NH}$ temperature coefficient measurements performed here depicted the effect of temperature change in stability difference between Rub1 and Ub. Despite the similarities between these two proteins, the results indicated a relatively more pronounced stability in Ub. Whether this method can be used to show how Rub1 differs from Ub upon conjugation enzyme binding remains to be seen. Nevertheless, from the results, individual residues or a series of residues can be targeted for mutational experiments in order to assess their contribution to the variation in stability between the two proteins. In fact, one of the methods by which temperature coefficient measurements can be useful is in comparing the difference of protein structure stability (via H-bond strength) in a ligand bound and unbound state. By extrapolating from this process, it is possible to compare two proteins i.e., Ub and Rub1 by measuring their TC in the apo state as well as after adding a ligand (E2, UBA's etc.) and observing which regions in the structure show a significant change.

This could help in pinpointing areas that are necessary for ubiquitin selection upon ligand binding and allow for focused mutation based experiments on Rub1 protein. As such, measurement of temperature coefficients can be a good technique to use in addition to other dynamics studies to gain more insight into how structural changes play a role in functionality.

Overall, the work that is presented in this chapter demonstrates the complexity involved in understanding the ubiquitin-proteasome system and the uniqueness of Ub as a signaling molecule. It is evident that much more work needs to be done to fully grasp the reasons behind the difference that exists between Ub and other UBLs with very similar structural and sequential organizations. It is my hope that the work done so far and presented here pushes the inquiry a step closer to finding the explanations that answer many of the questions that still remain.

Chapter 6: Exploratory experiments and side projects

6.1 Rub1 CEST experiments

Given the possibility of conformational change for Rub1 at higher temperatures based on curvature of shifts as well as the somewhat uncommon results observed on CD measurements for Rub1 it was necessary to investigate a possible presence of "invisible" conformational states that were not observable by the experiments that were undertaken so far. To address this issue, an NMR method of Chemical Exchange Saturation Transfer (CEST)[122] was utilized .

CEST is an experiment that relies on saturation transfer to provide information on invisible states in chemical exchange with visible state by exploiting a difference in resonance frequency between the two states[123]. This method applies a saturation field with an offset from the visible state frequency if the invisible state is saturated, and the saturation will be transferred by chemical exchange to the visible state, leading to a decrease in the intensity of the latter. Plotting the intensity of the visible state against the offset of the saturation field will produce a saturation profile that can be analyzed. A direct way of observing more than one state would be to detect more than one intensity dip for the residue that is being observed. CEST can detect minor states that are in slow exchange (k_{ex} between 10 and 500s^{-1}) provided that they are populated to at least 1%. Faster exchange (k_{ex} >500s^{-1}) approaches the intermediate exchange regime in which the two states do not have different chemical shifts. For Rub1, CEST experiments were taken at 280K, 296K and 310K. In all these cases, there was no detection of more than one dip for any signals

suggesting that Rub1 does not show conformational dynamics in solution that could be detected by NMR (results not shown).

6.2 Characterization of lysine-48 linked chains through Small-angle and very small angle neutron scattering studies

In order to determine the shape of lysine-48 linked ubiquitin trimers and tetramers in solution, a collaborative effort with Dr. Susan Krueger at the NIST Center for Neutron Research (NCNR) was undertaken by utilizing the capabilities of very small and small angle neutron scattering (vSANS and SANS) probes. These measurements were used for determination of radii of gyration of the scattering entities[124], their molecular weights from extrapolation of the scattering to its intensity at zero angle[125] and their shapes in solution on length scales between approximately 10 Å to about 1000 Å.

In addition to the crystal structure and NMR based structural analysis of ubiquitin tetramer, a SANS based study would enhance the knowledge of how ubiquitin chains in a tetramer are arranged in solution. Because neutrons interact with atomic nuclei and are sensitive to different isotopes of an element, such as hydrogen and deuterium, the technique of contrast matching was utilized whereby the ubiquitin tetramers and trimers used in this experiment were selectively labeled by deuteration in addition to the use of deuterated solvent. This would give information on the structural organization of the chains in solution and selective deuteration would mask the scattering from labeled units and thus allow for scattering only from units chosen for observation (non-deuterated units). This would give

sets of data that would add on to the already acquired understanding of ubiquitin structure and give a potentially complete picture on the dynamics involved in ubiquitin chains.

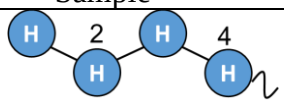
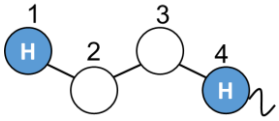
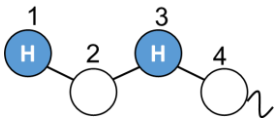
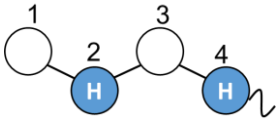
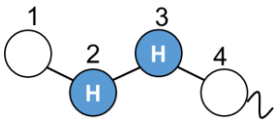
Sample		pD 6.8
1		2.00 mg/ml
2	 2 & 3 deuterated	3.55 mg/ml
3	 2 & 4 deuterated	2.5 mg/ml
4	 1 & 3 deuterated	4.0 mg/ml
5	 1 & 4 deuterated	4.0 mg/ml

Table 6.1: Selectively deuterated Ub_4 chains used in SANS and vSANS studies. The ubiquitin binding protein UBA2 was added to sample 4 at 1:1 and 1:2 molar ratios to observe any possible conformational changes.

The various chains were enzymatically assembled in the same manner as previously described. Once a variation of deuterium labeled chains were obtained, SANS and vSANS measurements were performed on samples that were mainly in pD 6.8. Even though differences in radius of gyration (Rg) measurements from SANS was anticipated given the

varied deuterium labeling of the units within the tetramers, there was still an expectation that samples 3 and 4, which are similar constructs in terms of the interval of labeling, would give similar results. Yet, as shown in Table 6.2, it appears that samples 1, 2 and 4 were actually more similar in terms of their radius of gyration. This was an interesting finding which raised questions on how the tetramer was arranged. It is quite possible that given the closed conformation of Ub tetramer, that sample 3 indicates a close interaction between units 1 and 3 which is similar to what is observed for units that are arranged next to each other (sample 5) while the non-deuterated units in samples 2 and 4 are more similar in their distance between each other.

SANS Results

Sample (all D_2O)	Concentration (mg/ml)	GNOM					Guinier			
		Rg(Å)	Error	I(0)	Error	Dmax	Rg(Å)	Error	I(0)	Error
1	2.0	24.8	0.2	0.0814	0.0004	80	24.8	0.3	0.0818	0.0005
2	3.6	25	1	0.0264	0.0007	85	22.1	0.7	0.0255	0.0005
3	2.5	22	2	0.0158	0.0006	80	18.5	0.7	0.0152	0.0003
4	4.0	26	1	0.036	0.001	90	23.9	0.7	0.0342	0.0006
5	4.0	17.8	0.3	0.0432	0.0004	60	17.6	0.5	0.0434	0.0006

Table 6.2: GNOM and Guinier based Rg results for all SANS measurements. Samples 1, 2 and 4 have similar Rg values as compared to samples 3 and 5.

Based on the atom pair distributions, P(r), for all the samples (Figure A4), it is evident that the tetramer exists in several conformations. This can be seen from the well-defined secondary P(r) maximums observed for all the samples between ~40 Å and 60 Å. Given the flexibility of ubiquitin linkages and polyubiquitin dynamics, the potential for several conformations is not surprising. Nevertheless, it is interesting to note that, samples 2 and 3 show a third maximum around 75 Å either indicating an even more extended conformation for Ub_4 or a possible artifact that presented itself during the conversion of I(q) to P(r).

By referencing the crystal structure of closed tetramer (PDB ID: 2O6V) as well as the results obtained from SANS, hypothetical solution states could be envisioned in which Ub_4 could be arranged in (Table 6.3). As such, if we follow the results from the SANS experiment, it makes sense that sample 5 would give the smallest R_g given that in all possible states, the non-deuterated units (2 and 3) are next to each other. As the distance between the non-deuterated Ub units increases, the arrangement of said units will have an effect on the resulting Rg. Therefore, given that sample 3 has the second lowest Rg, we can eliminate all the states that would potentially give samples 2 and 4 a lower Rg than sample 3. This would mean State A and E could not be the best representation for the observed SANS results. By the same rationale, State C has to be eliminated since that arrangement would make it so that sample 2 would have its non-deuterated units much closer to each other than sample 3. This leaves two potential states that would correspond to the SANS results observed. Still, given the information available, it is not possible to pinpoint the best options for the arrangement of the tetramer which means that further detailed work needs to be done to elucidate on the conformational population states that would best explain the SANS results obtained.

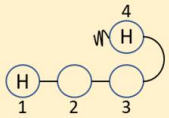
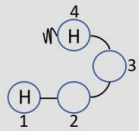
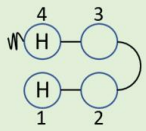
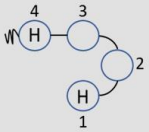
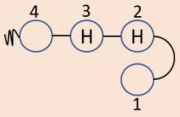
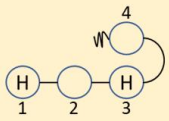
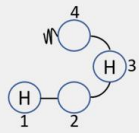
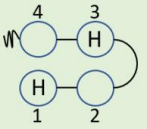
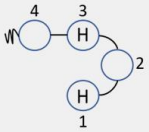
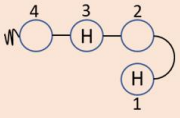
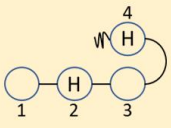
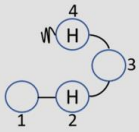
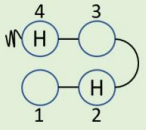
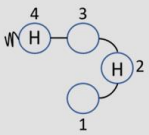
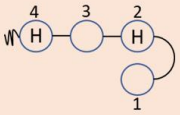
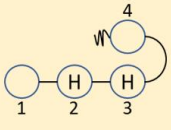
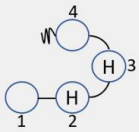
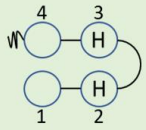
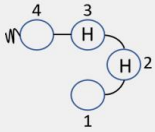
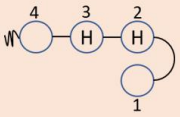
Sample	State A	State B	State C	State D	State E
2					
3					
4					
5					

Table 6.3: Cartoon model of the potential conformational equilibrium of Ub_4 in solution. Unit 1 indicates the distal unit while unit 4 indicates the proximal unit within the chain. States A, C and E are the least populated states given their potential contradictory explanation to SANS data.

Along with SANS work, SAXS experiments were also conducted by Dr. Alex Grishaev at IBBR to assess whether or not our samples were adversely affected either by D_2O buffer or deuteration of subunits. Surprisingly, the SAXS data Rg for all samples show high similarity with sample 5 indicating the most deviation when compared to SANS data (Table 6.4). Nevertheless, the SAXS curves for deuterated samples all look very similar when superimposed to that of non-deuterated samples which leads to the conclusion that deuteration did not adversely affected the results in the SANS experiments. Still, SAXS results give Rg's that are quite larger than that of SANS (especially for sample 5) which further complicates the results obtained so far. As already stated, this indicates that much

more rigorous work needs to be done to fully understand what the experimental results mean in terms of the solution based structural arrangement of lysie-48 linked Ub tetramer. All the relevant GNOM refined plots are represented in supplemental information (Figure A4).

SAXS Results

Sample	Concentration	GNOM					Guinier			
	(mg/ml)	Rg (Å)	Error	I(0)	Error	Dmax	Rg (Å)	Error	I(0)	Error
1 H_2O	1.1	28.2	0.3	0.0351	0.0003	95	26.8	0.5	0.0344	0.0004
1 D_2O	1.6	27.4	0.2	0.0416	0.0003	95	27.2	0.3	0.0418	0.0003
2 H_2O	1.5	26.7	0.2	0.0488	0.0003	95	26.2	0.2	0.0486	0.0003
2 H_2O	3.0	27.2	0.2	0.0986	0.0004	95	26.4	0.1	0.0971	0.0004
2 D_2O	1.8	27.5	0.2	0.0568	0.0003	95	26.6	0.2	0.0562	0.0003
3 H_2O	1.5	27.6	0.3	0.0522	0.0004	95	26.5	0.2	0.0513	0.0004
3 D_2O	1.25	27.0	0.3	0.0391	0.0003	95	27.1	0.3	0.0397	0.0003
3 D_2O	2.5	27.1	0.1	0.0865	0.0003	95	26.3	0.2	0.0856	0.0004
4 H_2O	1.4	28.9	0.2	0.0507	0.0003	100	29.4	0.4	0.0522	0.0005
4 H_2O	2.8	27.6	0.2	0.0987	0.0005	95	27.1	0.2	0.0985	0.0004
4 D_2O	1.5	27.4	0.2	0.0498	0.0003	95	26.9	0.3	0.0497	0.0003
4 D_2O	3.0	27.0	0.1	0.1080	0.0004	95	26.6	0.2	0.1080	0.0004
5 H_2O	1.6	28.6	0.3	0.0452	0.0004	95	28.5	0.3	0.0455	0.0004
5 D_2O	1.8	29.0	0.5	0.0564	0.0005	110	28.1	0.3	0.0562	0.0004

Table 6.4: GNOM and Guinier based Rg results for all SAXS measurements. All results between deuterated and non-deuterated samples are comparable to each other. SAXS results have high Rg which is especially evident for sample 5.

6.3 Attempt at partial characterization of lysine-48 linked Ub trimer through NMR

During the process of collecting data for peptide binding of Ub trimers, it became prudent to gain some structural insights of free Ub trimers on a residue-by-residue basis by means of the 2D heteronuclear ^{1}H, ^{15}N HSQC, SOFAST NMR spectroscopy that was being applied. Given the scarce information on the structural and dynamic information of lysine-48 linked Ub trimers, there was interest to determine whether these types of chains adopt

distinct conformations in solution in comparison with ubiquitin monomer and lysine-48 linked dimer.

As already stated, Ub_3 was generated enzymatically with the incorporation of selective ^{15}N-isotope labeled ubiquitin unit through a stepwise manner. Chemical shift mapping using NMR was used to identify the interactions between each unit through comparisons with CSPs of dimer and Ub monomer. To this end, CSPs for the amide backbone residues were compared between the corresponding ^{15}N-enriched distal and proximal units in the trimer and the corresponding units in ^{15}N labeled dimers under identical NMR experimental conditions (20 mM sodium phosphate, 130 mM NaCl, and pH 6.8 at 296 K). In the case of endo (middle) labeled trimer, since it has the property of both a proximal and distal labeled Ub, it was compared to both distal and proximal labeled Ub dimers.

The results show an interesting overlap in the way Ub_3 and Ub_2 are arranged in 3D space. Based on the published structures of lysine-48 linked Ub_2, we know that dimers have a closed conformation in both crystal form and in solution[21, 126]. As such, the small CSP's observed between Ub_3 and Ub_2 structures (<0.05 ppm average CSP) strongly suggests a form of a closed conformation in Ub_3 as well. Any real difference is seen in the distal unit of the trimer, which shows a slightly higher average CSP compared to the other units in the trimer (Fig 6.1; a). This is also observed when separately comparing the CSP's of dimers and trimers compared to unanchored wild-type Ub monomers. All the perturbations observed are relegated to the hydrophobic patch for both chain types which strongly suggests that the closed conformation in both chains similarly involves the interactions of this patch between the various Ub units within each chain.

Interestingly, newly published independent studies have suggested that Ub_3 is in a conformational equilibrium between linear and some version of a closed conformation with higher probability given to the population where the proximal and endo units form a compact interaction[127]. This appears to support our NMR results indicating a closed conformation with some variation observed in the distal unit. Of course, further detailed structural studies need to be undertaken to confirm this data with solution NMR that incorporates paramagnetic effects (using MTSL or lanthanides bonded to various units within the trimer) as well as SANS data in order to gain better insights in how the various units within Ub_3 interact with each other.

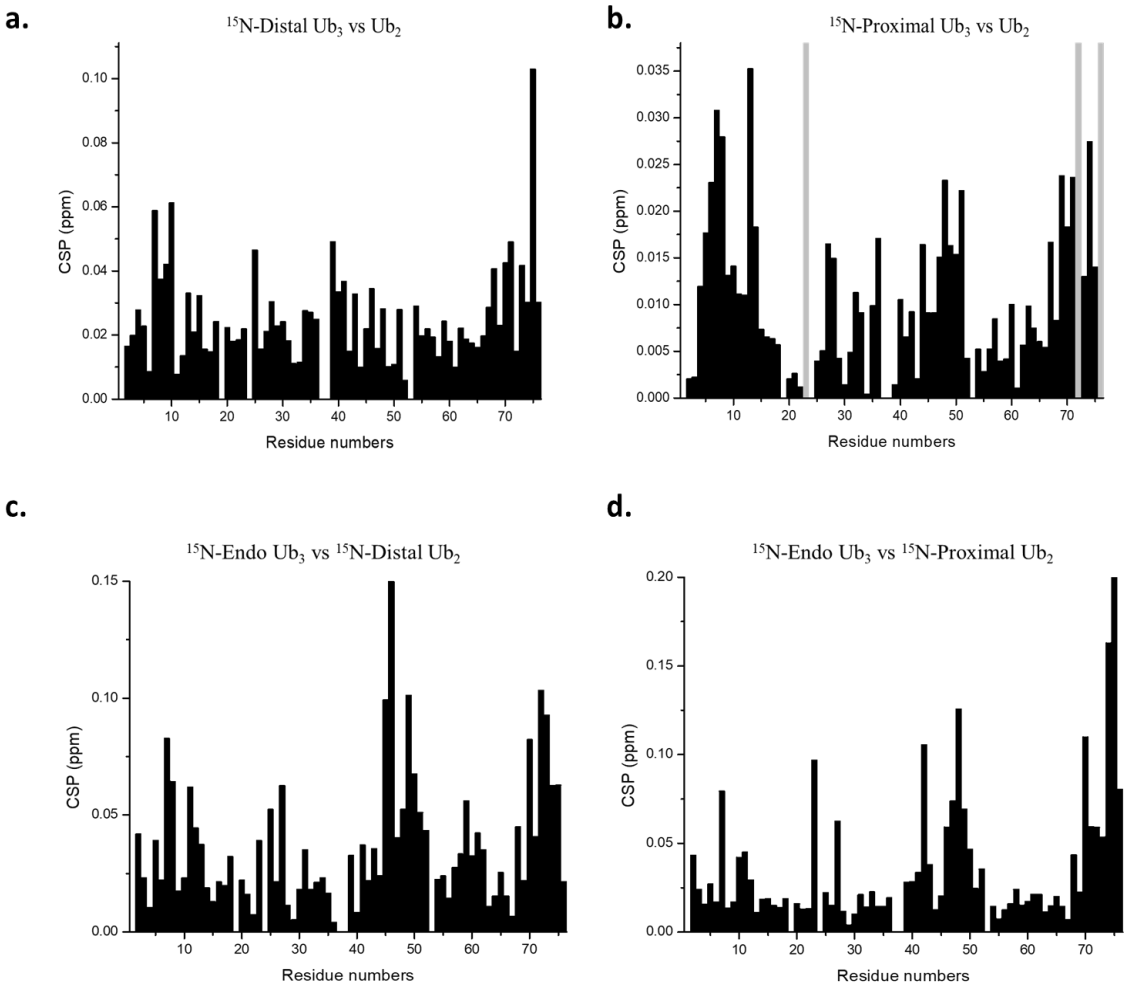


Figure 6.1: CSP plot comparison between Ub_3 and Ub_2. a) Distal unit shows an overall large CSP with less specific areas of perturbation. b) Proximal unit comparisons exhibit slight differences in the hydrophobic patch regions. c) and d) show how the endo unit compares to both units in a dimer. CSP's indicate areas where there is difference in interaction between dimer units and corresponding trimer units.

6.4 Ub-DHFR

While trying to understand the mechanism of ubiquitylation by E2-25K, a parallel interest asking the question of how ubiquitylation affects substrate structure and activity was explored. For this purpose, the enzyme dihydrofolate reductase (DHFR) was used as a model substrate due to its well folded and homogeneous nature[128]. Just like Ub, DHFR is a protein that is quite ubiquitous. In fact, it is found in all organisms to catalyze the

reduction of dihydrofolate to tetrahydrofolate through the cofactor NADPH acting as the electron donor. As an enzyme, it is essential in maintaining the amount of tetrahydrofolate in the cell which is used for de-novo synthesis of nucleic acids. Being important in the role of DNA precursor synthesis, DHFR is an attractive therapeutic target in the treatment of cancer by inhibiting its function and thereby limiting the growth of cancerous cells[90]. A number of proteasome related studies have used human and mouse DHFR as a model substrate but in this project, *E.coli* DHFR (17.8 KDa) was used for experiments related to ubiquitylation. More specifically, a head to tail Ub fused, C-terminal His-tagged DHFR (Ub-DHFR-IR-His6; 32.7 KDa) was used as a potential substrate for the ubiquitylation experiments that were explored.

As previously mentioned, studies have already shown that Ub-DHFR can be ubiquitylated without the need of an E3 enzyme and is readily recognized and degraded by the 26S proteasome[38] making this substrate ideal for working in our lab. Even though lysine-48 linked tetra ubiquitylation of Ub-DHFR through the use of E2-25K had been used to show the minimum number of lysine-48 linked Ub chains required for substrate degradation, the difference in the Ub-DHFR constructs used in previous studies compared for this project and the preliminary results thereof suggest a relationship between the type of Ub fusion to DHFR and a consequent change in structure which has an effect in ubiquitylation process. Therefore, the goal of this preliminary inquiry was to look at the structure of Ub-DHFR as a model system to elucidate on how ubiquitylation might change the structural integrity and activity of substrates during ubiquitylation prior to degradation. Apart from NMR studies, an attempt was also made to try and get a crystal structure of Ub_4-(Ub-DHFR) complex.

For this study, Ub fused DHFR (Ub-DHFR) was provided by Dr. Markus R. Eisele's lab at the Max Planck Institute of Biochemistry. Even though NMR results showed the difference in Ub signals caused by the interaction with DHFR, the spectra for Ub-DHFR showed a high degree of signals congregating in the middle of the spectra suggesting a potentially unfolded protein. Furthermore, none of the crystallization attempts of Ub-DHFR: Ub complex were successful. Unfortunately, based on these initial results, this line of examination was indefinitely discontinued.

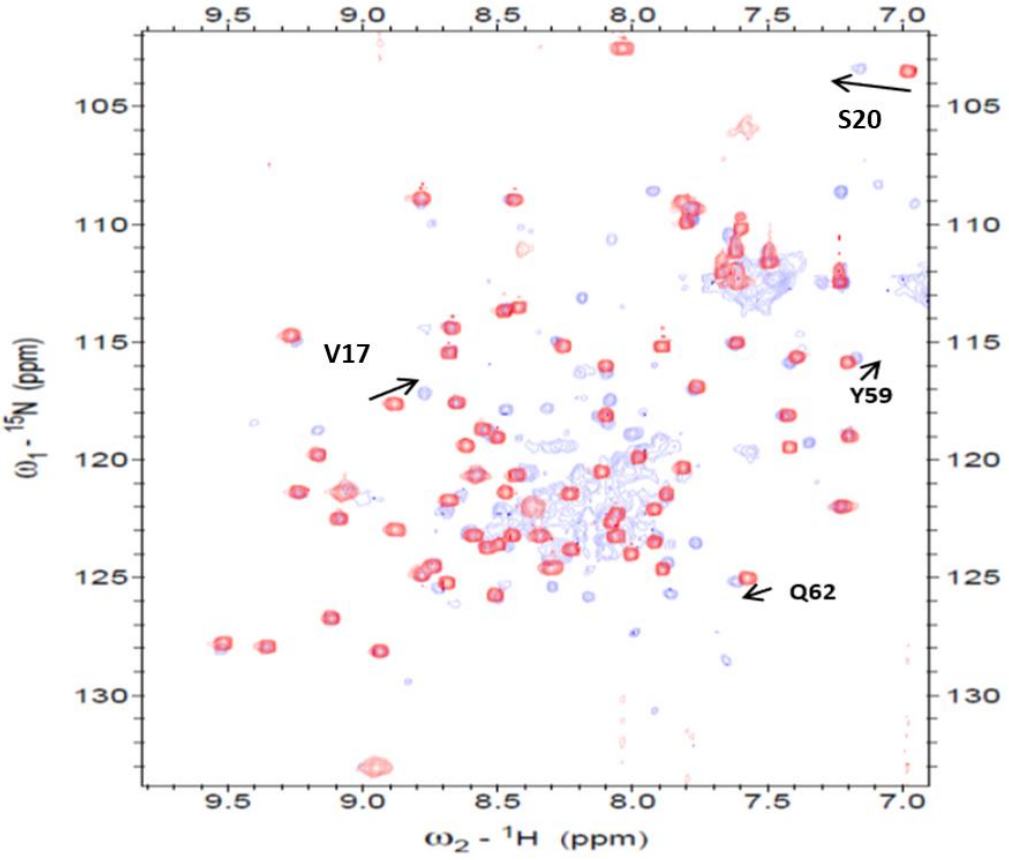


Figure 6.2: Overlay of ^{1}H-^{15}N-HMQC Sofast spectra of ^{15}N-UbDHFR with wild type Ub. Red: wild type ^{15}N-Ub. Blue: ^{15}N Ub-DHFR. Arrows show shifted residues of Ub.

6.5 p53-UDR

Another side project focused on the protein p53-Binding protein 1 (53BP1) which is a protein that plays a role in cellular response to DNA damage and has been proposed as being a transcriptional coactivator of the p53 tumor suppressor protein[129]. More specifically, the study focused on whether or not the ubiquitination-dependent recruitment (UDR) motif of 53BP1 recognized and interacted with lysine-27 and lysine-63 linked ubiquitin dimers. The genesis of this project was based on previous work that indicated UDR specifically recognized ubiquitin chains that are linked via lysines 27 or 63[130]. As such further exploration was performed with the utilization of NMR binding studies on this protein with ubiquitin monomer and lysine-63 linked dimer.

53BP1 responds to DNA damage by acting as a checkpoint protein that accumulates on the chromatin surrounding DNA double-strand break (DSBs) to accomplish its repairing functions. 53BP1 accumulation at DSB sites requires the recognition of histone methylation, in particular H4K20me2[131], by its tandem Tudor domain[132] (Fig 6.3). However, the formation of 53BP1 foci also requires the UDR motif which is a region that endows 53BP1 with the ability to accumulate at DSB sites with the help of the RNF168 E3 ligase[133,134]. This raised the question of how a ubiquitin ligase promotes the accumulation of a methylated-histone binding protein at sites of DNA damage.

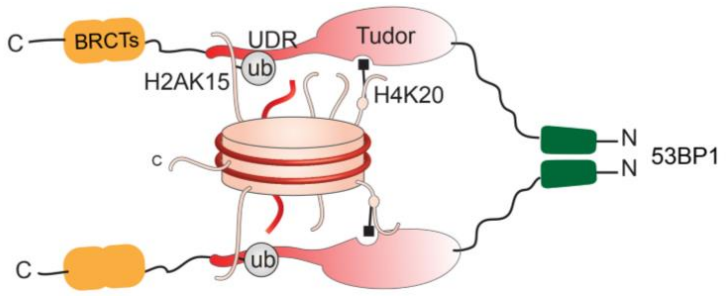


Figure 6.3: Model of the interaction of a 53BP1 dimer with a H4K20me2- and H2AK15ub-modified mononucleosome[132].

To that extent, the UDR and TUDOR-UDR plasmids of 53BP1 were received from the lab of Dr. Daniel Durocher at the University of Toronto and the proteins were grown. Subsequent NMR titration studies of these proteins onto ^{15}N labeled mono-ubiquitin and ^{15}N labeled lysine-63-linked di-ubiquitin were performed. The NMR results showed negligible to no binding of any kind while the same type of study with lysine-27 linked diubiquitin has yet to be done.

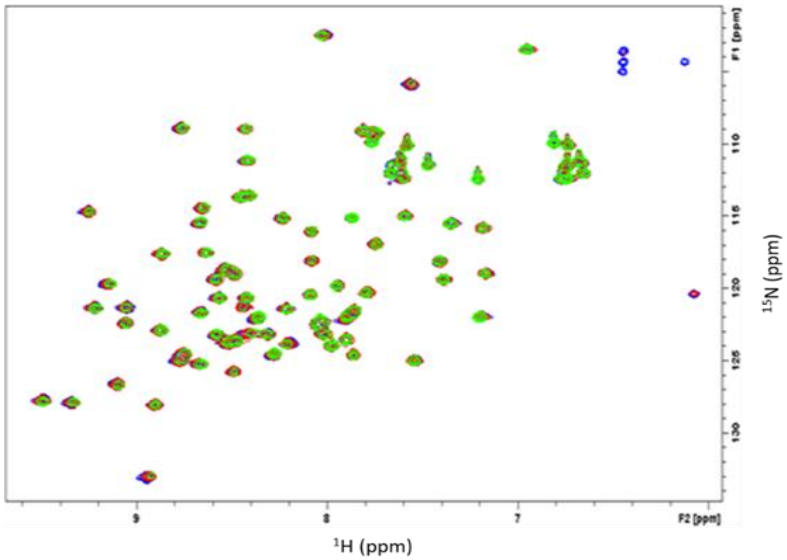


Figure 6.4: Overlay of ^{1}H-^{15}N-HMQC Sofast spectra of ^{15}N-Ub into which unlabeled GST-TUDOR-UDR is titrated. ^{15}N-Ub (blue spectra): ^{15}N-Ub: GST-TUDOR-UDR (1:1) (red spectra): ^{15}N-Ub: GST-TUDOR-UDR (1:5) (green spectra). No shifts were observed.

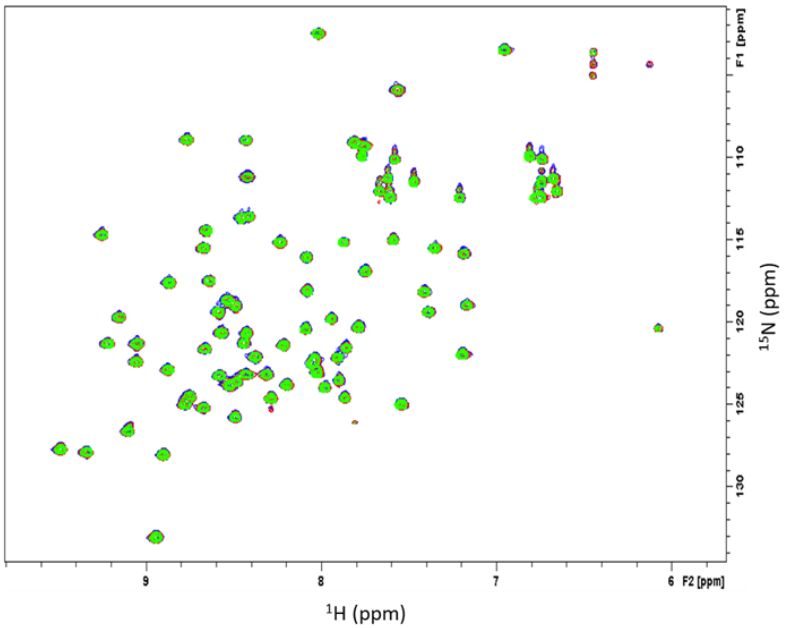


Figure 6.5: Overlay of ^{1}H-^{15}N-HMQC Sofast spectra of ^{15}N-Ub into which unlabeled UDR is titrated. ^{15}N-Ub (blue spectra): ^{15}N-Ub: UDR (1:1) (red spectra): ^{15}N-Ub: UDR (1:3) (green spectra).

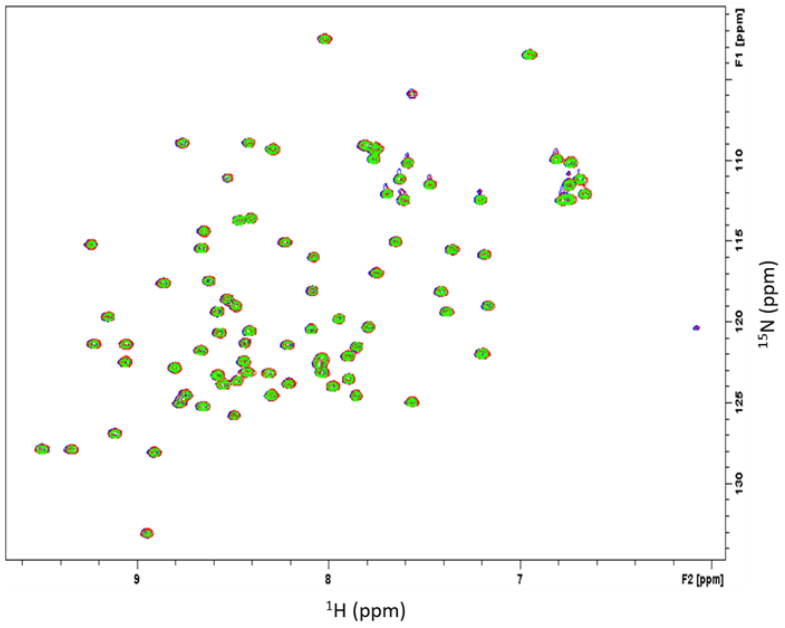


Figure 6.6: Overlay of ^{1}H-^{15}N-HMQC Sofast spectra of distally labeled ^{15}N lysine-63 linked Ub dimer into which unlabeled UDR is titrated. Distal ^{15}N lysine-63 linked Ub dimer (blue spectra): ^{15}N lysine-63 linked Ub dimer: UDR (1:1) (red spectra): ^{15}N lysine-63 linked Ub dimer: UDR (1:3) (green spectra).

In order to identify residues that participate in the recruitment of 53BP1 to DNA damage sites, a number of UDR mutants had been identified that disrupt 53BP1 recruitment[130] (Fig 6.7). Among these, a titration trial with UDR that contained two of the mutations that had the strongest impact, namely L1619A and R1627A was undertaken. Comparing this titration to that of wild type UDR showed no difference whatsoever.

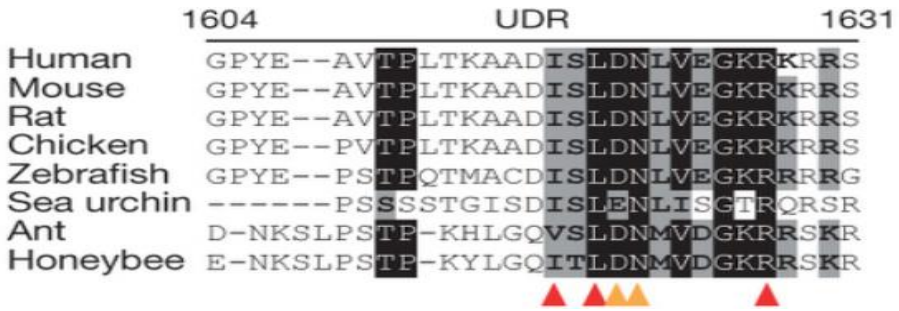


Figure 6.7: Alignment of the UDR region in 53BP1 orthologues. Arrowheads highlight key UDR residues[130].

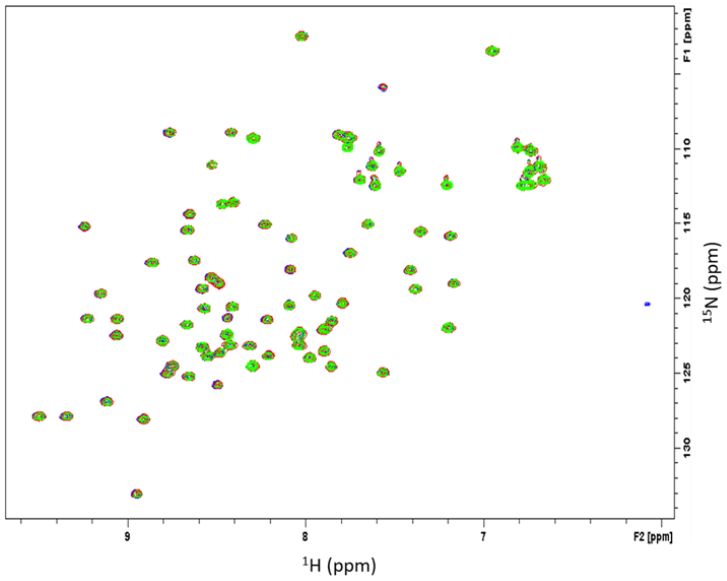


Figure 6.8: Overlay of ^{1}H-^{15}N-HMQC Sofast spectra of distally labeled ^{15}N lysine-63 linked Ub dimer into which unlabeled UDR L1619A, R1627A is titrated. Distal ^{15}N lysine-63 linked Ub dimer (blue spectra): ^{15}N lysine-63 linked Ub dimer: UDR mutant (1:1) (red spectra): ^{15}N lysine-63 linked Ub dimer: UDR mutant (1:3) (green spectra).

In fact, all the NMR titration results of TUDOR-UDR and UDR with either mono ubiquitin or lysine-63 linked di-ubiquitin showed little to no change in chemical shifts or any visible signal attenuation. This suggests that there is very little interaction between these two proteins. Based on the data presented in previous studies, it could well be that interaction between ubiquitin and the UDR domain of 53BP1 is dependent on the presence of RNF168 E3 ligase ubiquitylated histone H2A[130]. The NMR results align more to the report that indicated failure in detecting interaction between UDR and free ubiquitin through NMR methods[130]. These experiments further seem to indicate that even TUDOR-UDR does not interact with free ubiquitin. Even though there is not enough data to make any definitive claims about the general relationship between TUDOR-UDR and ubiquitin, these preliminary experiments show that further study of TUDOR-UDR interaction with different ubiquitin chains could be performed to observe the mechanism by which histone ubiquitylation allows for 53BP1 interaction.

Chapter 7: Materials and Methods

7.1 Ub protein expression and purification

All the constructs used in this work encode human and yeast proteins in commercially available DNA plasmids for expression in *E.Coli* cells. The plasmids were propagated in DH5α cells and mainly expressed in BL21 (DE3) variants unless mentioned otherwise. Any additional cloning was carried out through site directed mutagenesis.

All the ubiquitin constructs were expressed in *E.coli* pJY2 cells. The ubiquitin proteins used for Rub1 conjugation or self-assembly were expressed and purified as previously described[135].

Most of the GST-tagged proteins were expressed in *E.Coli* BL21 (DE3) cells using LB media. Proteins were purified using GST columns (GE LifeSciences) and eluted using 50 mM Tris, 10 mM Glutathione, and pH 7.6. The purified proteins were concentrated using 10 kDa MWCO Amicon concentrators and stored in 50 mM Tris, pH 8, in small aliquots at -80 °C. If the GST-tag needed to be cleaved, it was done by using thrombin protease at 4 °C for 4hrs and further purified using size-exclusion chromatography. Protein concentration was measured using the NanoDrop 2000c (Thermo Scientific)

7.1.1 Deuterated ubiquitin growth

As mentioned already, the incorporation of D_2O labeled ubiquitin in a chain required the growth of deuterated Ub_{D77} and Ub_{K48R} ubiquitin proteins. For this purpose, Dr. Zvi Kelman's BL2 lab at IBBR was helpful in making their facilities and equipment available

whereby with the assistance of Dr. James Todd Hoopes, a 5 L fermenter was used in order to optimize and maximize the yield of D_2O labeled ubiquitin growth.

Initially, concentrated stocks of essential nutrients were used for both starter media and media used as feedstock during the main 5L growth. Salts were provided as 20x concentrated nitrogen phosphate stock (NPS) containing 1 M sodium phosphate dibasic, 1 M potassium phosphate monobasic, 0.1 M sodium sulfate, and 0.93 M ammonium chloride all dissolved in D_2O. Vitamins were provided by a 1000x concentrated vitamin stock (VS) containing 0.2mM vitamin B12, 0.2mM p-aminobenzoic acid, 0.2mM nicotinic acid, 0.2mM thiamine-HCl, 0.2mM pyridoxine-HCl (vitamin B6), 5μM folic acid and 5μM riboflavin (vitamin B2). Trace metals were provided by a 10,000x concentrated trace metals stock (TMS) containing 50mM Iron(III) chloride, 20mM calcium chloride, 10mM manganese(II) chloride, 10mM zinc chloride, 2mM cobalt(II) chloride, 2mM copper(II) chloride, 2mM nickel(II) chloride, 2mM sodium molybdate, 2mM sodium selenite, and 2mM boric acid.

High-density cultures were achieved using Winpact Parallel Fermentation system equipped with a 5-liter fermentation vessel. The starting media for the vessel was composed of 125 mL 20x NPS, 2.5 mL 1M magnesium sulfate, and 2.3L D_2O. The vessel was autoclaved for 50 minutes and allowed to cool. Cooling can be accelerated by connecting the water jacket of the fermenter vessel and the condenser to their water sources and turning temperature control to 37 °C. Once at 37 °C, the media in the jar was supplemented by 2.5 mL of 1000x VS, 250 μL of 10000x TMS, and the appropriate antibiotics (ampicillin and chloramphenicol) to make a total of 2.5 L of a Modified Taylor Media that was used

throughout the growing process and will be referred henceforth as MTM media. The successful growth of ubiquitin was assured through use of 10% inoculum.

To begin growing deuterated ubiquitin, the mutants had to be acclimated to D_2O. As such, BL21 (DE3) LB agar plates of ubiquitin with ampicillin and chloramphenicol resistance were used to inoculate 20 ml starter of H_2O/D_2O (50/50) LB media. The 20 ml culture was allowed to grow for 18 hours at 37°C and 220 rpm. This culture was then centrifuged at 2000 g for 10 minutes, and the pellets transferred to a warm, 250 ml MTM media made up of H_2O/D_2O (30/70). This culture was again grown at 37°C and 220 rpm for 18 hours. Finally, this culture was pelleted down as already stated and re-grown in 250 ml, 100% D_2O MTM media. Throughout this process, the OD of the cultures was monitored to ensure that they were at least above 6 OD's. Once the cultures were fully acclimated, they were spun down and re-suspended in approximately 80 ml's of 20% glycerol/ H_2O, divided into 20 ml inoculates and placed in -80°C after being flash frozen.

Once all the media and inoculate were prepared, one 20 ml inoculate was added to the fermentation jar containing the MTM media. The OD of the fermentation vessel was monitored by extracting a sample through the sampling arm, diluted as necessary and measuring A600 with a spectrophotometer.

All settings involved in maintaining the appropriate environment for successful growth were regulated using values set before start of fermentation. Temperature was maintained at 37°C for the duration of the culture and the pH was held at 7 using 5 M sodium hydroxide. Agitation, provided by a removable agitation motor with an impeller, was set at a constant 350 rpm. An autoclavable oxygen probe was used to monitor and control dissolved oxygen levels. The DO probe was calibrated such that un-inoculated media with

air flow and 350rpm agitation was 100% oxygen saturation. Oxygen levels were controlled using the oxygen enrichment/D.O. Cascade system, in which oxygen is piped into the gas flow to maintain a minimum D.O. at 36%.

Nutrients for growth were provided both in a constant feed. Initial 1liter feed was composed of 100 mL 20x NPS, 150 g glycerol, 6 mL of 1,000x VS, and 600 μL of 10,000x TMS and 30 g ammonium chloride dissolved in D_2O. The slow initial growth of culture was done overnight until OD measurement reached 5 by which time the initial feed was started. The rate of feed was regulated to dispense its contents over a certain amount of time.

After the contents of the first feed were dispensed and the OD reaches at least above 15, induction was done with 1mM IPTG and a second feed containing 15% glycerol in D_2O was started. From inoculation to harvest, a normal growth lasted anywhere from 32 hours to 49 hours. This variation in growth time depended on the feed rates applied and the total inoculation time.

Given the media and nutrients above, a single growth is expected to reach an OD close to 20 with an expected final volume of 5 L. Cells were harvested when the change in cell growth decreased with highly reduced demand for supplemental oxygen. In this process of growing Ub, a little over 800 mgs of non-deuterated Ub and 400 mgs of deuterated Ub was grown.

7.1.2 Synthesis of selectively isotope enriched or deuterated Ub chains

Recombinant Ub monomers carrying chain-terminating mutations (either Ub_{K48R} or Ub_{K63R} mutations for distal termination or C-terminal D77 extension (Ub_{D77}) for proximal

termination) were expressed and purified as previously detailed. Isotopic enrichment (^{15}N) of Ub_{WT} and the abovementioned Ub variants was achieved as detailed[136]. The assembly of selectively isotope-enriched lysine-48 linked Ub trimers and tetramers was achieved using the following steps, as illustrated in Fig 2.1. First, the Ub_{K48R}-Ub_{D77} dimers with desired starting amount and isotope-labeling scheme were made in a 2 mL reaction mixture in the presence of 500 nM E1, 50 μM E2-25K (E2), protein breakdown mix, 2 mM ATP and 3 mM TCEP, for 12 hrs at 37°C. . The reaction was quenched with glacial acetic acid to adjust the pH to 4.5. The efficiency of the reaction was checked on a 15% SDS-PAGE gel. These di-ubiquitin chains were separated using cation chromatography on 5 mL Hi-Trap SPHP columns (GE LifeSciences) using a gradient of 50 mM Ammonium acetate, 1 M NaCl, pH 4.5. The chain-terminating D77 at the C-terminus of the proximal Ub was removed enzymatically using yeast Ub C-terminal hydrolase YUH1 and buffer exchanged into 50 mM Tris, pH 8.0 making the dimer available for further conjugation. As the next step, this dimer was conjugated to Ub_{D77} to form a trimer. Subsequently, the C-terminal D77 of the trimer was removed to enable assembly of the lysine-48 linked tetramer. Final Ub chains were buffer exchanged into 20 mM sodium phosphate, pH 6.8 for NMR studies unless otherwise described.

Synthesis of selectively deuterated chains followed the exact same protocol as described with difference occurring only in the utilization of a deuterated Ub domain instead of an isotope-enriched one.

7.2 Rub1 expression and purification

Rub1 (Resembles Ubiquitin Protein 1), from Saccharomyces cerevisiae was used along with variants that contained the following mutations: E63K & T72R ($Rub1_{E63K/T72R}$), K4F & T72R ($Rub1_{K4F/T72R}$), or T72R ($Rub1_{T72R}$). All Rub1 proteins were expressed in *E. coli* BL21 (DE3) cells as a fusion with an intein tag containing a chitin-binding domain. The cells were lysed using sonication and then centrifuged at 6000 rpm. The supernatant was fractionated on a chitin-bead column as detailed elsewhere (NEB: Impact system). The purity of Rub1 fractions was determined using SDS-PAGE and mass spectrometry.

7.2.1 Enzymatic assembly of $Rub1_{E63K/T72R}$ - Ub_{74} heterodimer

$Rub1_{E63K/T72R}$ and Ub_{74} as well as $Rub1_{T72R}$ and Ub_{74} (all 10 mgs each) were incubated with 20 μM ubiquitin-conjugating E2 enzyme (Ube2K, aka E2-25K), 500 nM ubiquitin-activating E1 enzyme (Uba1), 10 mM creatine phosphate, 5 mM $MgCl_2$, 5 mM ATP, and creatine phosphokinase in 50 mM Tris-HCl buffer, pH 8.0 at room temperature for approximately 16 hrs. The solution was centrifuged at 13000 rpm in order to remove the precipitated enzymes (E1 and E2) and further fractionated using FPLC (GE Healthcare Life Sciences)[106].

7.2.2 Enzymatic assembly of $Rub1_{K4F/T72R}$ –$Rub1_{K4F/T72R}$ homodimer

$Rub1_{K4F/T72R}$ (20 mg) was incubated with 20μM ubiquitin conjugating E2 enzyme Ube2K, 500 nM Ube1, 10 mM creatine phosphate, 5 mM $MgCl_2$, and creatine phosphokinase in 50 mM Tris-HCl buffer, pH 8.0 and 5 mM ATP for approximately 16 hrs. The solution was centrifuged at 13000 rpm in order to remove the precipitated enzymes (E1 and E2) and

then fractionated using FPLC (GE Healthcare Life Sciences)[106]. The reaction was monitored by SDS PAGE.

7.2.3 Enzymatic assembly of $Rub1_{T72R}$–$Rub1_{T72R}$ homodimer

$Rub1_{T72R}$ (10 mg) was incubated with 20 μM ubiquitin conjugating E2 enzyme Ube2K, 500 nM Ube1, 10 mM creatine phosphate, 5 mM $MgCl_2$, and creatine phosphokinase in 50 mM Tris-HCl buffer, pH 8.0 and 5 mM ATP for approximately 16 hrs. The solution was centrifuged at 13000 rpm in order to remove the precipitated enzymes (E1 and E2). The reaction was monitored by SDS PAGE. The supernatant was then injected slowly onto a 5 mL cation-exchange column at 1 mL/min using FPLC (GE Healthcare Life Sciences)[106]. The Rub1–Rub1 containing species were eluted with cation buffer B (50 mM ammonium acetate containing 1 M NaCl, pH 4.5) and the purity was checked using SDS-PAGE.

7.3 Liquid-Chromatography–Tandem Mass Spectrometry

After synthesis and initial purifications of either Rub1-Rub1 and/or Rub1-Ub dimers, the following subsequent steps were undertaken by Dr. Fabio Gomes from Dr. Catherine Fenselau's lab. Intact proteins were separated using an Ultimate 3000 ultrahigh-performance liquid chromatograph (Thermo Fisher). Approximately 100 ng of rubylated ubiquitin was desalted and concentrated in a PepSwift Monolithic trap (200 μm × 5 mm) and subsequently separated on a ProSwift RP-4H column (100 μm × 25 cm, Thermo Fisher) at a flow rate of 1.5 μL/min using a linear gradient of 5–55% solvent B (75% acetonitrile, 25% water, and 0.1% formic acid) for 20 min. Solvent A was 97.5% water, 2.5% acetonitrile, and 0.1% formic acid. The column oven and auto sampler temperatures

were set to 35 and 4 °C, respectively. Mass spectra were acquired on an orbitrap Fusion Lumos mass spectrometer (Thermo Fisher) in intact-protein mode with an ion-routing multipole pressure of 3 mTorr. The potential for in-source fragmentation was maintained at 10 V. A resolving power of 120 000 was used to acquire both the precursor and fragment ions. The automatic-gain-control (AGC) target was defined as 10^6 ions during both the precursor- and fragmentation-ion acquisitions. All of the MS/MS spectra were produced in the data-dependent mode using electron-transfer dissociations (ETD) supported by high-energy-collision- induced dissociations (EThcD) with a 6 ms ETD reaction time and a supplemental activation with 10% normalized high-energy collisions[104, 106].

7.4 NMR studies of Ub4ix and Ub4a peptides

All NMR studies were on Bruker Avance III 600 MHz or 800 MHz spectrometer equipped with TXI or TCI cryoprobe. The sample temperature was set to 296K (23°C) for Ub4ix and Ub4a studies, unless indicated otherwise. Selectively ^{15}N-labelled Ub_4, Ub_3 and Ub_3 samples were prepared in 20 mM sodium phosphate buffer, pH 6.8 containing 5-10% D_2O and 0.02% (w/v) NaN_3. Both Ub4ix and Ub4a were synthesized as previously described[31, 39]. Peptide preparation for NMR based titration studies is described below.

Ub4ix and Ub4a peptide were dissolved in 40% D_6-DMSO, 30% D_2O and 30% 20 mM sodium phosphate pH 6.8 buffer and titrated in μL amounts into lysine-48 linked Ub_2, Ub_3, and Ub_4 chains that were ^{15}N-enriched at different unit positions, with a starting Ub chain concentration of 200 μM. The endpoint chain/peptide molar ratio varied from 1:1 (Ub_2) to 1:1.5 or 1:2 for Ub_3 and Ub_4. Several of these titrations were performed twice and at

differing protein concentrations, to verify the reproducibility of the observed spectral changes. Separately, a similar amount of deuterated and non-deuterated DMSO was added to ^{15}N-labeled Ub chains to verify that the presence of DMSO in the peptide-containing stock solution had negligible effect on Ub spectra.

1D ^{1}H spectra and 2D ^{1}H-^{15}N SOFAST-HMQC spectra were acquired at each titration point. Assignment of the peptide-bound NMR signals of Ub_3 and Ub_4 were complicated by signal broadening (due to slow tumbling of tetra-Ub), signal overlap and by the slow-exchange regime which precluded tracing gradual shifts in the NMR signals upon titration. The spectra of free protein and at the titration endpoint were also recorded at 37°C in order to reduce line broadening (for signal assignment purposes) and to check the reproducibility of the spectral changes at a physiologically relevant temperature. The residues affected by peptide binding were identified based on strong attenuation/disappearance of their unbound signals combined with the absence of new, bound signals in close proximity as seen in the spectra. Spectra were processed and analyzed using TopSpin 3.6.3 and 4.1.1 (Bruker Inc.) and Sparky[137, 138].

NMR signal shifts were quantified as chemical shift perturbations using the following equation:

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

where δ_{HA} ; δ_{HB} and δ_{NA} ; δ_{NB} correspond to chemical shifts of ^{1}H and ^{15}N resonances for the unbound and bound species, respectively.

Assignment of peptide-bound NMR signals of Ub_3 and Ub_4 was complicated by signal overlap, signal broadening (due to slow tumbling of tri- and tetra-Ub) as well as) by

the slow-exchange regime which precluded tracing gradual shifts of the NMR signals upon titration. Generally, in the cases of slow or intermediate exchange, the residues affected by peptide binding were identified based on strong attenuation/disappearance of their unbound signals in the spectra as already discussed. In addition, for the experiments performed for Ub4a peptide, ^{15}N-edited TOCSY and NOESY spectra were recorded at the titration endpoint for assignment purposes, which allowed for assignment of significant portion of the peptide-bound NMR signals.

To quantify the reduction in intensity of the unbound NMR signals of Ub chains upon peptide binding, intensities in the ^{15}N-^{1}H spectra recorded at each titration point were measured at the positions corresponding to peaks of the peptide-free species, and then normalized for each residue by dividing by the intensity of the corresponding signal in the spectra recorded before the addition of the peptide. This allowed for the identification and monitoring of residues involved in slow exchange wherein the unbound signals were observed to gradually decrease in intensity upon addition of cyclic peptide.

7.5 PRE-Experiments

Ub4a labeling with TEMPO was performed in Dr. Ashraf Brik's lab at the Technion Institute of Technology by dissolving Ub4a in 6 M Gn·HCl/200 mM phosphate buffer (pH 7.3, 2386 µL, 1 mM) and cooled to 0°C for 5 min. A solution of TEMPO (4-maleimido-2, 2, 6,6-tetramethyl-1-piperidinyl oxy) free radical (2.70 mg, 4.5 equivalent) dissolved in 100 µl of DMF at 0°C under dark conditions was added to the peptide solution and the reaction was kept at 0°C under dark. Under dark conditions, the HPLC analysis was carried

out on a C4 analytical column with a gradient of 0-60% B over 30 min and found the reaction was completed in 1 h. Under dark conditions, the crude peptide was purified using a C4 semi-preparative column with the gradient flow of 0-60% B over 30 min to give TEMPO labeled peptide in 70% (3.23 mg) in yield[39].

^{1}H-^{15}N HSQC spectra were obtained with TEMPO in the oxidized (paramagnetic) and reduced (diamagnetic) states. TEMPO was reduced by incubating the sample with 6-10 molar excess ascorbate for an hour. In order to remove TEMPO completely sample was incubated with 6-10 molar excess TCEP for an hour. It is important to mention that the NMR spectra of samples reduced with ascorbate agreed well with equivalent spectra that were collected after removing TEMPO, indicating that addition of a spin label did not adversely affect the binding of Ub4a to various Ub chains.

The paramagnetic relaxation enhancement effects were expressed for each amino acid as the ratio of signal intensities (I/Io) in the ^{1}H-^{15}N HSQC spectra recorded with spin label in the oxidized (I) and reduced (Io) state. The reconstruction of paramagnetic spin label position and distance between spin label and particular amino acid position was determined with the in-house Matlab program SLfit[54] according to the following equations:

$$\frac{I}{Io} = \frac{R_{2,\mathrm{dia}}}{R_{2,\mathrm{dia}} + \Delta R_{2,\mathrm{para}}} \exp\left(-t\Delta R_{2,\mathrm{para}}\right)$$

where t is the total experimental time when amide ^{1}H in the magnetization is in the transverse plane during the pulse sequence; $R_{2,\mathrm{dia}}$ is the intrinsic transverse relaxation rate measured for the amide ^{1}H in the diamagnetic state; $\Delta R_{2,\mathrm{para}}$ is the amide ^{1}H transverse relaxation rate enhancement induced by spin label in paramagnetic state.

Value of the $\Delta R_{2,para}$ depends on the distance (r) between the amide ^1H and the unpaired electron of the spin label.

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

where K is $1.23*10^{-44}$ m^6s^{-2}, τ_c is the rotational correlation time of the molecule and ω_H is the ^1H resonance frequency.

7.6 Deubiquitination Assays

Deubiquitination assays were performed as described[139] using 35 μM of each Ub chains with a 1:1 ratio of either Ub4ix or Ub4a peptide and the appropriate enzyme concentration in Tris, pH 8.0 at 32°C. Enzyme concentrations were: 3 μM OTUB1, 3 μM IsoT/USP5, 3 μM Usp2, and 3 μM UBP6. Reaction samples were taken at various timepoints and quenched with SDS-PAGE loading dye, then resolved on a 15% polyacrylamide gel by SDS-PAGE and visualized by Coomassie staining.

7.7 X-ray crystallography data collection and analysis of Ub_4 with Ub4ix peptide

The initial Ub4_Ub4ix concentration was 5.0 mg/mL in 50 mM HEPES, pH 7.4. Screening of conditions was performed in sitting drop vapor diffusion 96-well plates at 22°C. Crystallization was observed after one week in 100 mM Bis-Tris, 200 mM NaCl, 25% polyethylene glycol (PEG) 3350, pH 6.5 and were left for another week to form better growth. Crystals were briefly soaked in 25% PEG 3350 and submerged in liquid nitrogen

prior to data collection. Diffraction data were collected at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA) on beam line NE-CAT 24-ID-E with a wavelength of 0.97921 Å and rotation angle of 0.8°.

Diffraction images were indexed and integrated with iMOSFLM[140] and scaled with AIMLESS[141]. Molecular replacement with the crystal structure of Ub (PDB: 1UBQ) as the search model was carried out in Phaser[142]. Six copies of Ub were present in the asymmetric unit, corresponding to two trimers. Refinement was performed with Coot[143, 144], CCP4i[145], and Phenix[146]. Electron density was not visible for residues 72-77 in the proximal Ub. Nor was any electron density observed for Ub4ix in the crystal. Furthermore, the Ub_4 structure obtained from this attempt was essentially the same to already published structure of Ub tetramer with closed structure as shown in Figure 2.11.

7.8 Crystallization, data collection and structural refinement of Ub_3:Ub4a complex

The following steps were performed by Dr. Jacek Lubkowski and his lab at NCI. Crystals of the Ub_3:Ub4a complex were grown from the hanging droplets composed of equal volumes of Ub_3:Ub4a (8 mg/ml) and the crystallization solution containing 0.15 M NaCl, 23% (w/v) PEG 3350 and 0.1 M Hepes (pH 7.5). Incubated at 20˚C, crystals appeared within 3-5 days and continued to grow for additional 20-30 days. For X-ray data collection, crystals were transiently transferred to the solution containing the abovementioned components with part of water replaced by glycerol (to the final concentration of 25% v/v) and Ub4a (to a final concentration of 5 μM). Subsequently, crystals were frozen in liquid nitrogen, followed by collection of X-ray data.

The X-ray data were collected at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA), on beamline 22-ID. All tested crystals displayed significant anisotropic mosaicity, a problem that was circumvented by acquisition of narrow-oscillation (0.15˚) individual images. Images were processed and scaled with HKL3000[147].

Structure of the complex was solved by molecular replacement (MR) with the program Phaser[142] using the monomer A extracted from the structure deposited in the PDB (1UBQ)[148]. The MR showed unambiguous solution for six Ub units, arranged as two trimers. The MR solution was first subjected to rigid-body refinement at the resolution 2.5 Å with the program Refmac5[149], followed by several cycles of a structural refinement of positions and isotropic atomic displacement parameters (B's) for non-H atoms. At this stage, traces of covalent linkages between the Ub subunits within each of the two trimers became quite apparent. Also, the difference electron density peaks (F_o-F_c) clearly indicated the presence of unaccounted molecular content. Guided by this electron density, it was possible to model one molecule of Ub4a within each Ub_3 complex. The restraint libraries for non-standard linkages between Ub subunits, as well as unnatural fragments of Ub4a, needed for subsequent structural refinement, were created with the aid of the program JLigand[150]. The refinement was carried out to the extent of experimental data (i.e., the resolution of 1.85 Å). Emerging model was intermittently inspected with the aid of the program Coot[143] and appropriate corrections were introduced. At the final stages of the refinement, we could identify and model two molecules of the Hepes buffer (one per trimer) and one molecule of glycerol (both were components of the crystallization or cryo-protectant solution). Additionally, 302 water molecules could be modeled within the

asymmetric unit, based on the difference electron density peaks. Final models were evaluated by the MolProbity server[151]. The crystal structure of the complex was solved by Dr. Jacek Lubkowski.

A number of processing and refinement parameters were obtained through the following mathematical equations:

$$R_{linear} = \frac{\sum(|I - \langle I \rangle|)}{\sum(I)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

$$R_{square} = \frac{\sum((I - \langle I \rangle)^2)}{\sum I^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

$$R_{p.i.m} = \frac{\sum \sqrt{\frac{1}{n_h - 1}} \sum |\langle I_h \rangle - I_{h,i}|}{\sum_h \sum_i I_{h,i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

where R_{linear} is Intensity R-factor, R_{square} is weighted R-factor, $R_{p.i.m}$ is precision of merged data, I is observed intensity; <I> is predicted intensity from calculated model, n_h is number of reflections with index h where h and i represent indices h and i.

7.9 Differential Scanning Calorimetry experiments

The following steps were conducted by Dr. Rob Brinson at IBBR. The proteins were thermally denatured using a nano differential calorimeter (TA Instruments, Inc.). The scan rate was 1 °C/min from 25 °C to 100 °C or 120 ° C for WT Rub1 and WT Ub, respectively. All samples were recorded in 20 mM sodium phosphate pH 6.0 and with 0 mM sodium chloride or 130 mM sodium chloride. Protein concentration ranged from 85 μM to 130 μM. Before each protein scan, three buffer scans were performed to establish the instrument

thermal history. In all cases, the proteins aggregated upon thermal denaturation, and reheating produced negligible signal. Processing was performed in Nano-Analyze (TA Instruments). All DSC traces were baseline subtracted and fit to a 2-state model. Due to the aggregation, precise thermodynamic parameters could not be calculated.

7.10 Variable temperature NMR

Temperature variable 1D ^{1}H spectra and 2D ^{15}N-^{1}H SOFAST-HMQC spectra were manually acquired using Bruker Avance III 800 MHz spectrometer equipped with a triple resonance TXI probe. Amide proton chemical shifts used for temperature coefficient calculations and curvature analysis were directly referenced to DSS (by shifting by the same correction required to move the 1D ^{1}H DSS peak in the spectrum to 0 ppm), while amide nitrogen chemical shifts are indirectly referenced to DSS using a $^{15}N/^{1}H$ Ξ ratio of 0.101329118 (i.e., $^{15}N_{(DSS\ correction)} = {}^{1}H_{(DSS\ correction)} \times 0.101329118$) to give results in Hz. The referenced shift in Hz is multiplied by 10 to convert the ^{15}N chemical shifts to ppm.

Chapter 8: Discussion and Future Directions

Among the various projects presented in this thesis, the first section dealt with characterizing the binding mechanism of two macrocyclic peptides, Ub4ix and Ub4a, with lysine-48 linked ubiquitin chains. NMR studies and enzymatic assays were utilized in trying to reveal the mechanism by which the interactions occurred with novel observations indicating that these peptides were actually partnering with the trimer moieties within a tetramer chain. In fact, our studies indicate that these peptides tightly interact with lysine-48 linked trimers given the NMR results showing similar slow exchange regime of binding as was observed in the tetramer binding experiments. The work done with Ub4ix further indicated that, within the context of a tetramer, this peptide can interact with either the distal or proximal trimer moieties with some preference for the distal end while the results from Ub4a studies, which were complemented by crystallography results, clearly show that the proximal trimer moiety within a tetramer is the preferable binding partner.

Within the context of Ub4a binding to ubiquitin chains, our studies showed how the C-terminal tail residues of the proximal Ub domain affect both the affinity of binding and the selectivity of binding partners within the chains. Furthermore, both NMR and crystallography results showed the role played by the hydrophobic patches of all the ubiquitin domains of a trimer which contribute to the binding of Ub4a. This further illustrates how cyclic peptides interact with target proteins and thereby perform their inhibitory tasks by interfering with the protein-protein interactions that are prominent within the proteome.

Moreover, these findings provide insights into the compositional characteristics of cyclic peptides which could serve as potential therapeutic agents for various disorders. The differences observed in terms of how Ub4ix and Ub4a interacted with ubiquitin chains further illustrate how incorporation of non-proteogenic amino acids can be useful in garnering better selectivity and affinity as well as improving membrane permeability. Both the *in vitro* and *in vivo* studies clearly show that incorporation of such non-proteogenic amino acids within Ub4a did not adversely affect the efficacy of the peptide and in fact, showed improvement due to its potential increase in resistance to hydrolysis by deubiquitinases or the proteasome.

One of the interesting aspects of the Ub4a peptide study was the importance of the free tail residues of the proximal Ub domain in peptide binding since in the cell, the C-terminal tail of a Ub chain is involved in a covalent linkage with a target protein. As such, unlike our *in vitro* experiments, chains within the cell will not have a free tail that could potentially affect the binding of Ub4a. In fact, our studies that looked at Ub4a binding with truncated tail residues mostly resemble what would be present within the cell with the proximal domain of such chains acting as a substrate protein that was tagged by a Ub trimer. If that holds true, one can imagine that Ub4a binding to a Ub_4 chain that is attached to a target protein would potentially interact with the distal trimer moiety of the tetramer. On the other hand, if Ub_4 association with a target protein occurs mainly through the very ends of the proximal tail, the Ub_4:Ub4a binding *in vivo* could potentially resemble our binding study that involved tetramers with partially truncated tail regions. As such, in order to get a clearer understanding on how the mechanism of action takes place *in vivo*, peptide binding studies with Ub chains that are covalently bound to substrate proteins through their

C-terminus should be examined. This would give a clearer picture into how Ub4a and other similar peptides interact with and affect the functions of ubiquitin chains.

Furthermore, in addition to using ubiquitin trimers as targets of peptide screening, future research should take into consideration the various chain types that are present within various ubiquitin signaling pathways in trying to find peptide therapeutics that are targeted for specific purposes. As such, it would be beneficial to look at peptides that specifically interact with chains linked through different lysines (K63, K11 etc.) and compare the mechanisms of action among various peptides in order to further understand how the chain selection mechanism works as well as the role played by the various ubiquitin chain conformations in the binding process.

Finally, one of the main advantages of using peptides is their efficacy in interfering with protein-protein interactions as demonstrated by their binding mechanism to Ub chains. Still, in the process of incorporating hydrophobic residues during peptide synthesis, the issue of solubility needs to be considered. This was reflected when dealing with both Ub4ix and Ub4a in the progression of this study whereby in addition to problems with solubility of the peptides in the initial stages, precipitation was observed after a certain excess of peptide was added during the NMR titration experiments (>1:2 Ub: peptide ratio). As such, future peptide synthesis and screenings need to have a balance between overall binding efficacy and solubility.

With regards to the Rub1/Ub studies, much is still unknown or unclear on what the differences are between these two very similar proteins which result in their non-overlapping functions. The studies presented here have clearly shown that Rub1 can be partially incorporated into the ubiquitin conjugation system given a number of specific

mutational changes. Nevertheless, despite these various mutations that were introduced with the goal of making Rub1 resemble ubiquitin, the results have not been able to conclusively demonstrate that Rub1 is able to make homogeneous Rub1 chains longer than a dimer. Furthermore, even with the structural and sequential similarity between Rub1 and ubiquitin, various stability experiments clearly indicate that Ub is much more stable than Rub1. How this fact relates to the functional differences observed between these two proteins is still unclear. As such, it is evident that there is a lot of work that still needs to be done in order to understand the differences between these two proteins.

Several methods of exploration that were utilized in the Rub1/Ub study were intended to pinpoint the potential differences between their sequences that could explain their functional differentiation. In fact, the experiments that measured temperature coefficients were intended to identify areas within Rub1 that show significant difference from Ub and thereby potentially lead to mutational studies that focus on said areas. As such, future studies can be informed from the TC results obtained here in identifying residues within Rub1 that show significant differences compared to Ub. In addition, given that TC analysis is an indication of the state of H-bonding within a structure, a complementary NMR experiment of hydrogen-deuterium exchange can lead to a better analysis in the stability difference between Rub1 and Ub. Since it is understood that areas with stronger hydrogen bonding will have a slower hydrogen-deuterium exchange than solvent exposed or structurally unstable areas, in addition to stability related analysis, this type of study, could reveal areas that should be focused on to make further mutation based studies.

Lastly, given that the genesis of the Rub1/Ub study was intended to understand why, despite its similarity to Ub, Rub1 is not being recognized by the ubiquitin conjugation system, it would be prudent to look at the enzymatic components of the conjugation in the future. More specifically, since the activation of Rub1 by the Ub E1 enzyme has already been shown to be successful, a more concerted effort could be applied to understand the mechanism by which E2 interacts with Rub1. This would especially be interesting given the results presented which showed that Rub1 can make some limited chains through different types of E2 enzymes. Future studies that explore this question from the perspective of E2 enzymes could improve our overall understanding of the selective mechanism employed by the ubiquitin conjugation system.

Appendix

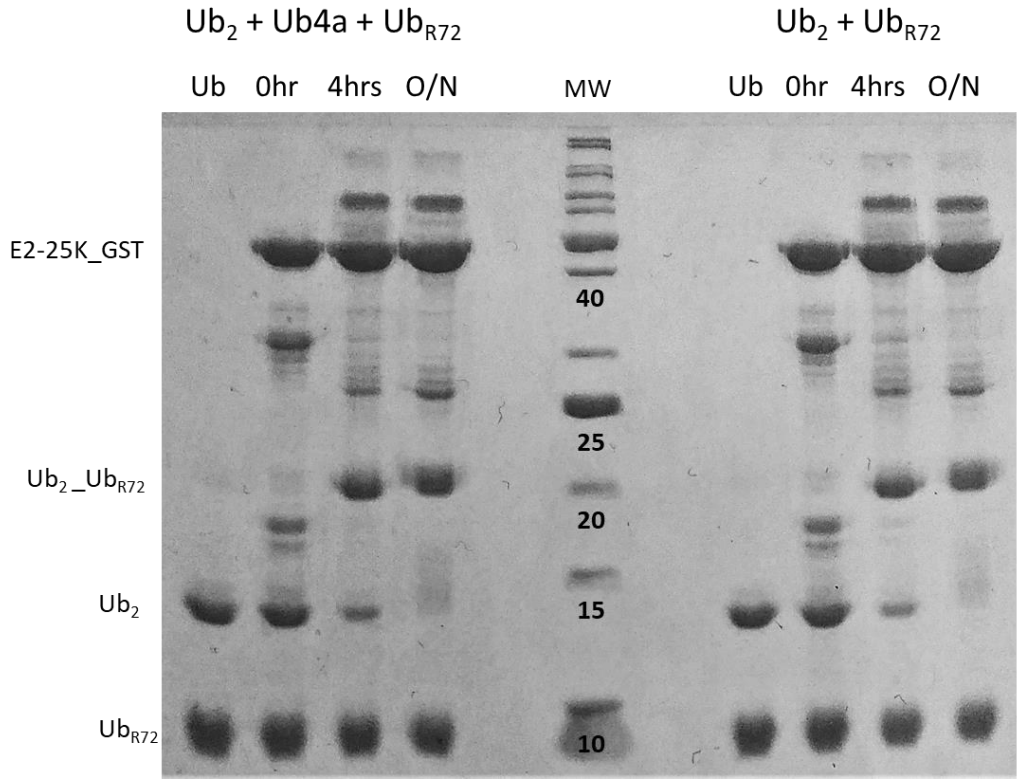


Figure A1: Ub4a binding of dimer chains does not inhibit chain elongation reaction. Assay of time course assembly of wild type Ub dimer and Ub_{72} with Ub4a peptide (left) and control without Ub4a peptide (right). Both assays show uninhibited elongation to a trimer chain.

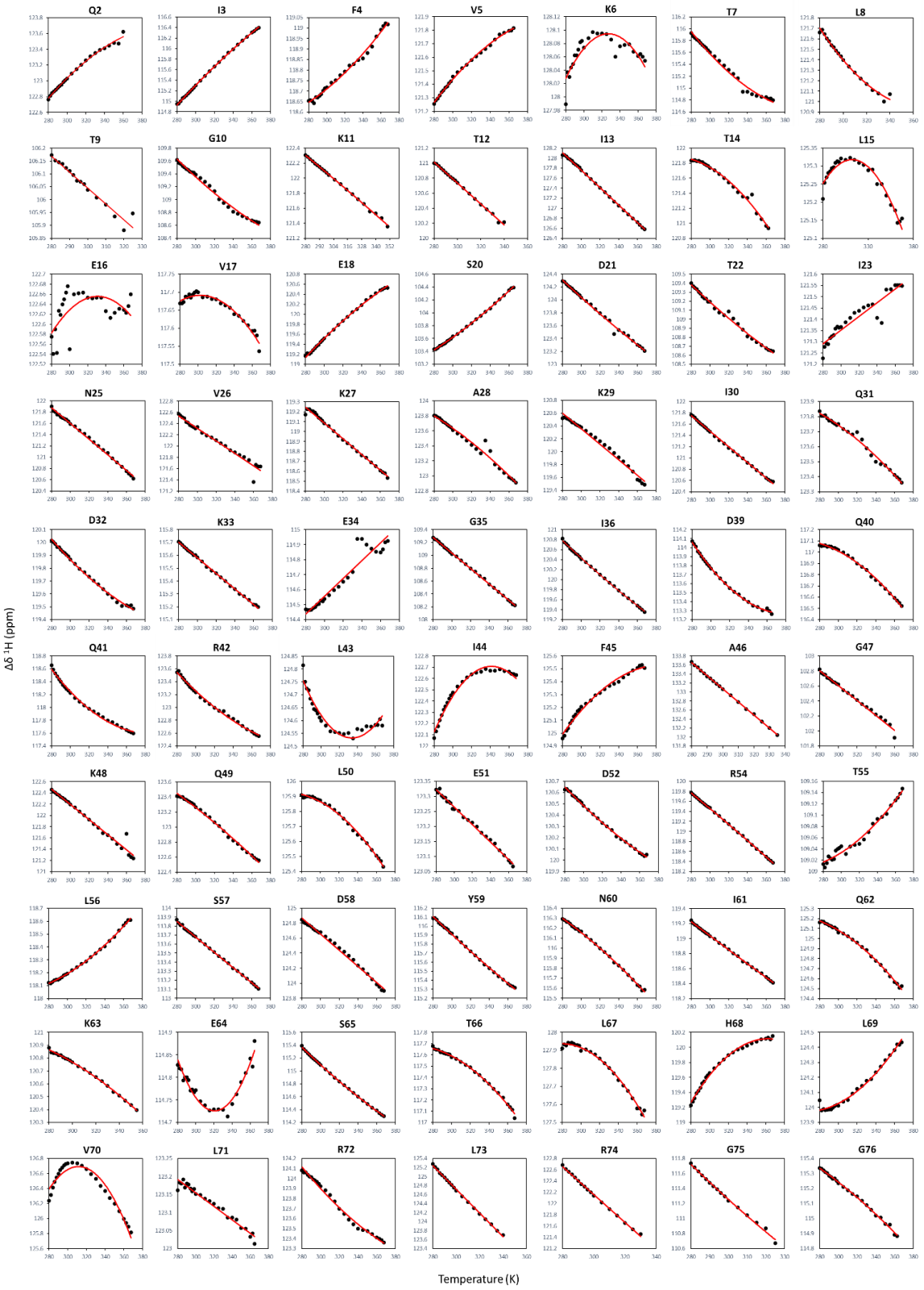


Figure A2: Temperature dependent chemical shifts of the backbone amide nitrogens of Ub at pH 6.0 from 280K-367.5K. Notice the positive slopes in many of the residues. More signals show curvatures than their proton TC plot counterparts (see Fig 5.6).

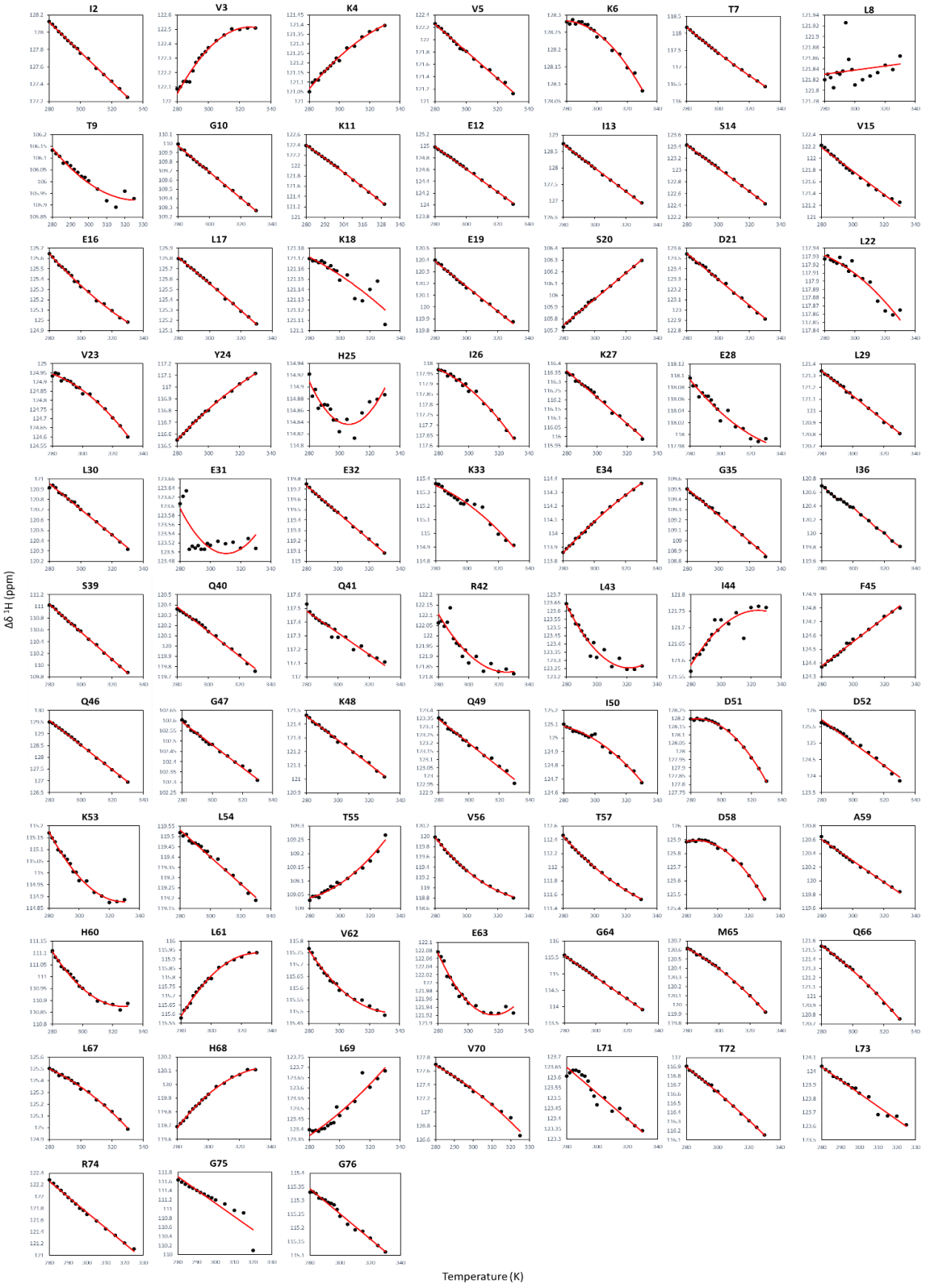
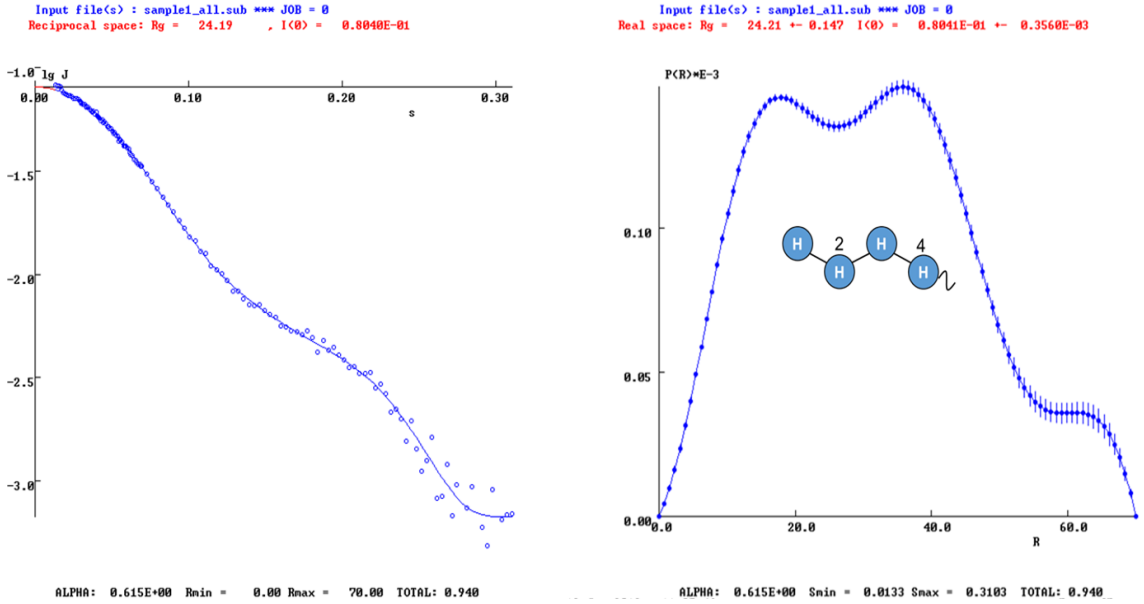
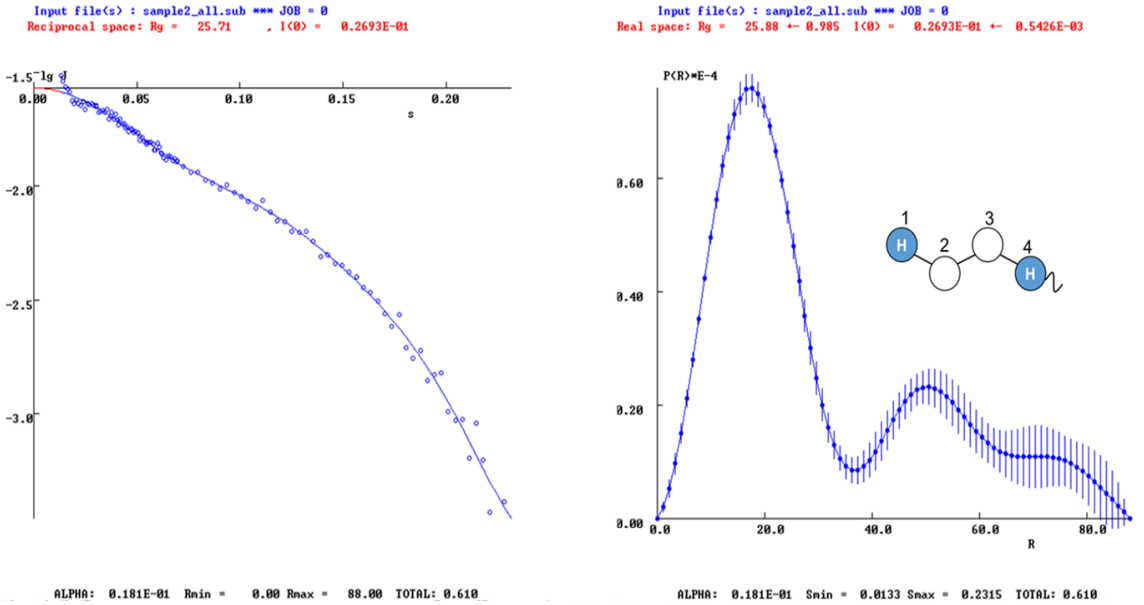


Figure A3: Temperature dependent chemical shifts of the backbone amide nitrogens of Rub1 at pH 6.0 from 280K-330K. Notice the positive slopes in many of the residues. More signals show curvatures than their proton TC plot counterparts (see Fig 5.7).

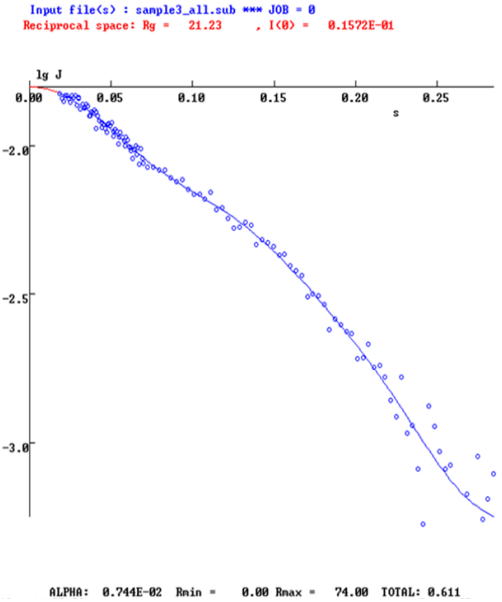
1.

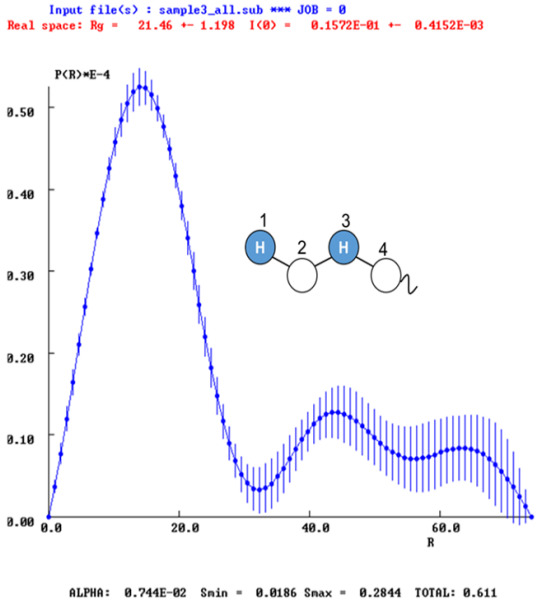

Input file(s) : sample1_all.sub *** JOB = 0
Reciprocal space: Rg = 24.19 , I(0) = 0.8040E-01
lg J
s
ALPHA: 0.615E+00 Rmin = 0.00 Rmax = 70.00 TOTAL: 0.940
Input file(s) : sample1_all.sub *** JOB = 0
Real space: Rg = 24.21 +- 0.147 I(0) = 0.8041E-01 +- 0.3560E-03
P(R)*E-3
R
H
2
H
H
4
ALPHA: 0.615E+00 Smin = 0.0133 Smax = 0.3103 TOTAL: 0.940

2.

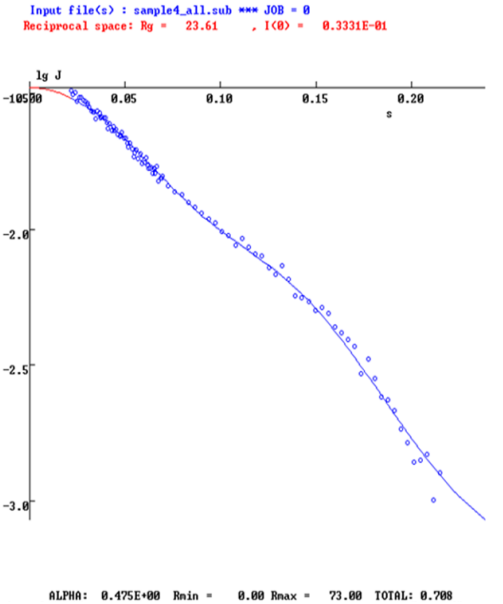

Input file(s) : sample2_all.sub *** JOB = 0
Reciprocal space: Rg = 25.71 , I(0) = 0.2693E-01
lg J
s
ALPHA: 0.181E-01 Rmin = 0.00 Rmax = 88.00 TOTAL: 0.610
Input file(s) : sample2_all.sub *** JOB = 0
Real space: Rg = 25.88 +- 0.985 I(0) = 0.2693E-01 +- 0.5426E-03
P(R)*E-4
R
1
H
2
3
4
H
ALPHA: 0.181E-01 Smin = 0.0133 Smax = 0.2315 TOTAL: 0.610

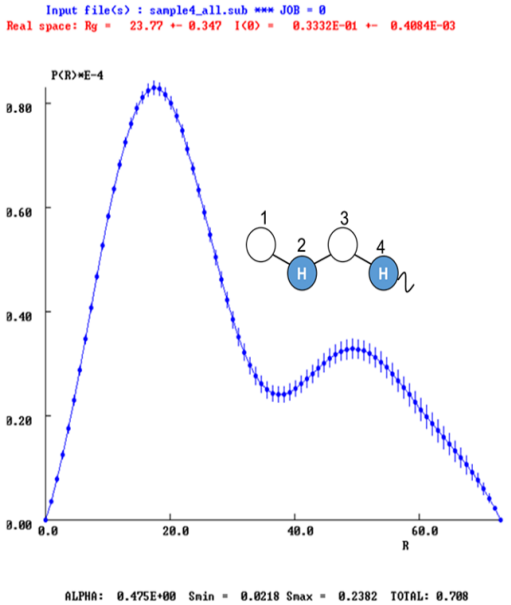
3.


Input file(s) : sample3_all.sub *** JOB = 0
Reciprocal space: Rg = 21.23 , I(0) = 0.1572E-01
lg J
s
ALPHA: 0.744E-02 Rmin = 0.00 Rmax = 74.00 TOTAL: 0.611


Input file(s) : sample3_all.sub *** JOB = 0
Real space: Rg = 21.46 +- 1.198 I(0) = 0.1572E-01 +- 0.4152E-03
P(R)*E-4
R
ALPHA: 0.744E-02 Smin = 0.0186 Smax = 0.2844 TOTAL: 0.611

4.


Input file(s) : sample4_all.sub *** JOB = 0
Reciprocal space: Rg = 23.61 , I(0) = 0.3331E-01
lg J
s
ALPHA: 0.475E+00 Rmin = 0.00 Rmax = 73.00 TOTAL: 0.708


Input file(s) : sample4_all.sub *** JOB = 0
Real space: Rg = 23.77 +- 0.347 I(0) = 0.3332E-01 +- 0.4084E-03
P(R)*E-4
R
ALPHA: 0.475E+00 Smin = 0.0218 Smax = 0.2382 TOTAL: 0.708

5.

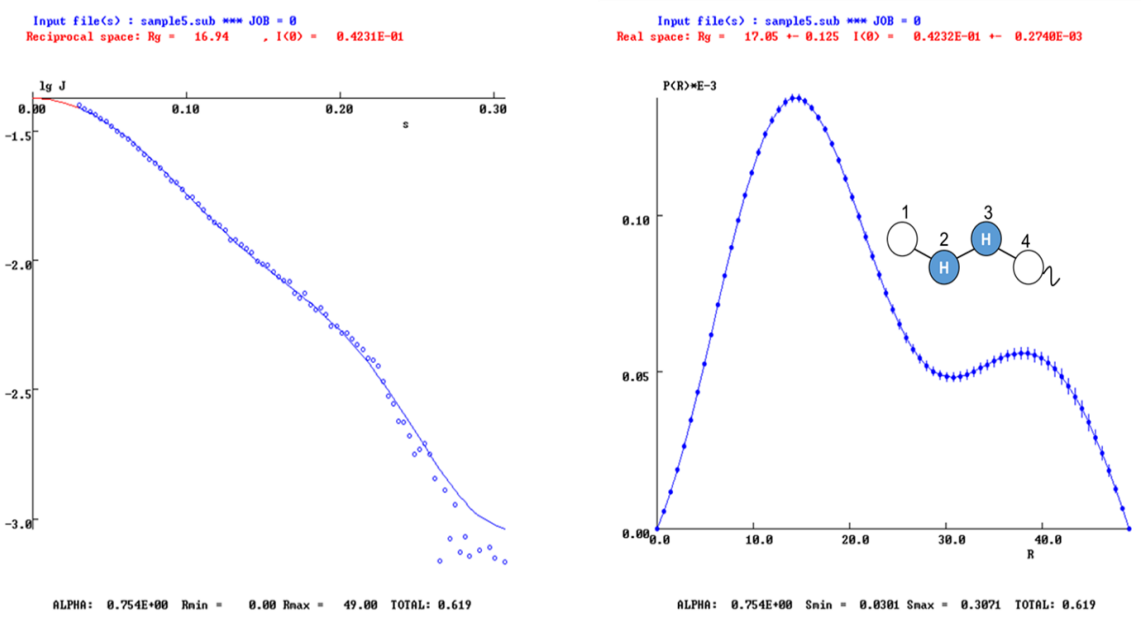


Figure A4: SANS based GNOM refined plots for samples 1-5. **Left:** experimental SANS profile representation of regularized intensity extrapolated for all I(s) and **Right:** experimental P(r) profiles for each sample. s: scattering length in $Å^{-1}$; R: chord length in Å

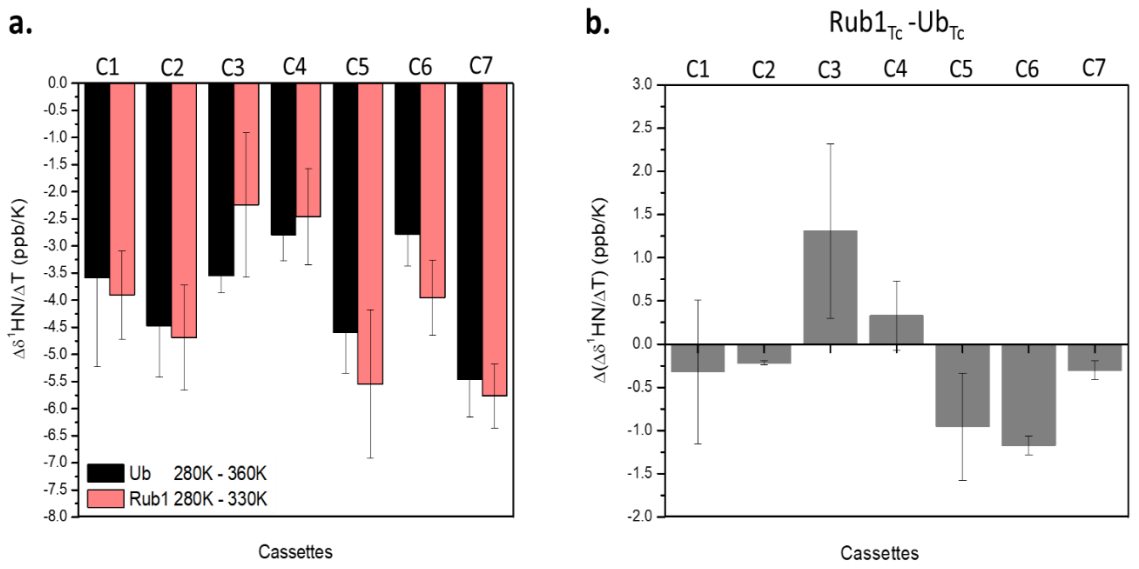


Figure A5: a) Averaged cassette temperature coefficients for Ub and Rub1 at T_m-15ºC; **b)** Difference in amide temperature coefficients between Ub and Rub1 averaged over selected cassette regions. (Ub PDB: 1D3Z) (Rub1: in-house NMR structure).

a.

Residue	Ub TC (ppb/K)
Q2N-H	-7.28
I3N-H	-0.40
F4N-H	-3.06
V5N-H	-2.87
K6N-H	-1.10
T7N-H	-3.93
L8N-H	-10.18
T9N-H	-3.05
G10N-H	-3.13
K11N-H	-2.27
T12N-H	-7.56
I13N-H	-3.66
T14N-H	-7.84
L15N-H	-3.53
E16N-H	-4.51
V17N-H	-4.36
E18N-H	0.15
S20N-H	-2.09
D21N-H	-3.54
T22N-H	-4.61
I23N-H	-2.86
N25N-H	-3.33
V26N-H	-3.38
K27N-H	-3.32
A28N-H	-1.51
K29N-H	-1.83
I30N-H	-3.43
Q31N-H	-2.75
D32N-H	-4.43
K33N-H	-0.38
E34N-H	-3.24
G35N-H	-3.87
I36N-H	-0.28
D39N-H	-3.41
Q40N-H	-3.00

Residue	Ub TC (ppb/K)
Q41N-H	-4.08
R42N-H	-2.71
L43N-H	-5.23
I44N-H	-1.82
F45N-H	-2.06
A46N-H	-6.53
G47N-H	-7.81
K48N-H	-1.08
Q49N-H	-6.04
L50N-H	-3.92
E51N-H	-6.14
D52N-H	-5.86
R54N-H	-2.46
T55N-H	-2.52
L56N-H	-1.54
S57N-H	-5.05
D58N-H	-3.21
Y59N-H	-1.59
N60N-H	-2.53
I61N-H	-3.53
Q62N-H	-5.15
K63N-H	-7.41
E64N-H	-5.43
S65N-H	-2.76
T66N-H	-6.56
L67N-H	-3.61
H68N-H	-2.86
L69N-H	-0.85
V70N-H	-4.13
L71N-H	-8.93
R72N-H	-3.63
L73N-H	-9.14
R74N-H	-8.68
G75N-H	-8.74
G76N-H	-6.76

b.

Residue	Rub1 TC (ppb/K)
I2N-H	-5.55
V3N-H	-2.12
K4N-H	-4.04
V5N-H	-10.53
K6N-H	-2.00
T7N-H	-5.45
L8N-H	-3.65
T9N-H	-3.32
G10N-H	-3.47
K11N-H	-2.90
E12N-H	-7.95
I13N-H	-5.06
S14N-H	-7.97
V15N-H	-3.21
E16N-H	-6.47
L17N-H	-5.33
K18N-H	0.17
E19N-H	-1.67
S20N-H	-2.01
D21N-H	-1.77
L22N-H	-5.12
V23N-H	-2.89
Y24N-H	3.52
H25N-H	-3.20
I26N-H	-3.49
K27N-H	-1.13
E28N-H	0.92
L29N-H	-4.01
L30N-H	-4.75
E31N-H	-2.15
E32N-H	-2.30
K33N-H	-2.88
E34N-H	-3.02
G35N-H	-2.40
I36N-H	-0.02
S39N-H	-8.61
Q40N-H	-2.32

Residue	Rub1 TC (ppb/K)
Q41N-H	-3.31
R42N-H	-0.64
L43N-H	-5.62
I44N-H	-2.00
F45N-H	-1.54
Q46N-H	-7.19
G47N-H	-8.86
K48N-H	-0.78
Q49N-H	-7.53
I50N-H	-1.61
D51N-H	-10.66
D52N-H	-6.23
K53N-H	-5.87
L54N-H	-3.38
T55N-H	-3.84
V56N-H	-2.81
T57N-H	-3.81
D58N-H	-3.88
A59N-H	-6.82
H60N-H	-2.43
L61N-H	-2.29
V62N-H	-5.50
E63N-H	-5.27
G64N-H	-7.16
M65N-H	-3.68
Q66N-H	-7.22
L67N-H	-4.38
H68N-H	-3.63
L69N-H	-1.30
V70N-H	-5.90
L71N-H	-7.86
T72N-H	-4.77
L73N-H	-9.47
R74N-H	-8.41
G75N-H	-9.30
G76N-H	-7.09

Table A1: Amide Temperature Coefficients for residues of **a)** Ub and **b)** Rub1.

References

1. Kumar, A.; Narayanan, V.; Sekhar, A., Characterizing Post-Translational Modifications and Their Effects on Protein Conformation Using NMR Spectroscopy. *Biochemistry* **2020,** *59* (1), 57-73.

2. Hochstrasser, M., Lingering mysteries of ubiquitin-chain assembly. *Cell* **2006,** *124* (1), 27-34.

3. Kerscher, O.; Felberbaum, R.; Hochstrasser, M., Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual Review of Cell and Developmental Biology* **2006,** *22*, 159-80.

4. Pickart, C. M., Back to the future with ubiquitin. *Cell* **2004,** *116* (2), 181-90.

5. Scheffner, M.; Nuber, U.; Huibregtse, J. M., Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **1995,** *373* (6509), 81-3.

6. Pickart, C. M., Mechanisms underlying ubiquitination. *Annual Review of Biochemistry* **2001,** *70*, 503-33.

7. Stewart, M. D.; Ritterhoff, T.; Klevit, R. E.; Brzovic, P. S., E2 enzymes: more than just middle men. *Cell Research* **2016,** *26* (4), 423-40.

8. Weissman, A. M., Themes and variations on ubiquitylation. *Nature Reviews Molecular Cell Biology* **2001,** *2* (3), 169-78.

9. Komander, D.; Clague, M. J.; Urbé, S., Breaking the chains: structure and function of the deubiquitinases. *Nature Reviews Molecular Cell Biology* **2009,** *10* (8), 550-63.

10. Bologna, S.; Ferrari, S., It takes two to tango: Ubiquitin and SUMO in the DNA damage response. *Frontiers in Genetics* **2013,** *4*, 106.

11. Dittmar, G.; Winklhofer, K. F., Linear Ubiquitin Chains: Cellular Functions and Strategies for Detection and Quantification. *Frontiers in Chemistry* **2019,** *7*, 915.

12. Boughton, A. J.; Krueger, S.; Fushman, D., Branching via K11 and K48 Bestows Ubiquitin Chains with a Unique Interdomain Interface and Enhanced Affinity for Proteasomal Subunit Rpn1. *Structure* **2020,** *28* (1), 29-43.e6.

13. Eddins, M. J.; Varadan, R.; Fushman, D.; Pickart, C. M.; Wolberger, C., Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH. *Journal of Molecular Biology* **2007,** *367* (1), 204-11.

14. Varadan, R.; Walker, O.; Pickart, C.; Fushman, D., Structural properties of polyubiquitin chains in solution. *Journal of Molecular Biology* **2002,** *324* (4), 637-47.

15. Tan, J. M.; Wong, E. S.; Kirkpatrick, D. S.; Pletnikova, O.; Ko, H. S.; Tay, S. P.; Ho, M. W.; Troncoso, J.; Gygi, S. P.; Lee, M. K.; Dawson, V. L.; Dawson, T. M.; Lim, K. L., Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases. *Human Molecular Genetics* **2008,** *17* (3), 431-9.

16. Beau, I.; Esclatine, A.; Codogno, P., Lost to translation: when autophagy targets mature ribosomes. *Trends in Cell Biology* **2008,** *18* (7), 311-4.

17. Lundgren, K.; Montes de Oca Luna, R.; McNeill, Y. B.; Emerick, E. P.; Spencer, B.; Barfield, C. R.; Lozano, G.; Rosenberg, M. P.; Finlay, C. A., Targeted expression of MDM2 uncouples S phase from mitosis and inhibits mammary gland development independent of p53. *Genes & Development* **1997,** *11* (6), 714-25.

18. Hofmann, R. M.; Pickart, C. M., Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **1999,** *96* (5), 645-53.

19. Ulrich, H. D., Conservation of DNA damage tolerance pathways from yeast to humans. *Biochemical Society Transactions* **2007,** *35* (Pt 5), 1334-7.

20. Ye, Y.; Rape, M., Building ubiquitin chains: E2 enzymes at work. *Nature Reviews Molecular Cell Biology* **2009,** *10* (11), 755-64.

21. Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M., Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *Journal of Biological Chemistry* **1992,** *267* (23), 16467-16471.

22. Chau, V.; Tobias, J. W.; Bachmair, A.; Marriott, D.; Ecker, D. J.; Gonda, D. K.; Varshavsky, A., A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **1989,** *243* (4898), 1576-1583.

23. Piotrowski, J.; Beal, R.; Hoffman, L.; Wilkinson, K. D.; Cohen, R. E.; Pickart, C. M., Inhibition of the 26 S Proteasome by Polyubiquitin Chains Synthesized to Have Defined Lengths. *Journal of Biological Chemistry* **1997,** *272* (38), 23712-23721.

24. Li, W.; Tu, D.; Li, L.; Wollert, T.; Ghirlando, R.; Brunger, A. T.; Ye, Y., Mechanistic insights into active site-associated polyubiquitination by the ubiquitin-conjugating enzyme Ube2g2. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (10), 3722-7.

25. Song, S.; Kim, S. Y.; Hong, Y. M.; Jo, D. G.; Lee, J. Y.; Shim, S. M.; Chung, C. W.; Seo, S. J.; Yoo, Y. J.; Koh, J. Y.; Lee, M. C.; Yates, A. J.; Ichijo, H.; Jung, Y. K., Essential role of E2-25K/Hip-2 in mediating amyloid-beta neurotoxicity. *Molecular Cell* **2003,** *12* (3), 553-63.

26. Xu, P.; Duong, D. M.; Seyfried, N. T.; Cheng, D.; Xie, Y.; Robert, J.; Rush, J.; Hochstrasser, M.; Finley, D.; Peng, J., Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* **2009,** *137* (1), 133-145.

27. Adams, J., The development of proteasome inhibitors as anticancer drugs. *Cancer Cell* **2004,** *5* (5), 417-21.

28. Nguyen, T.; Ho, M.; Ghosh, A.; Kim, T.; Yun, S. I.; Lee, S. S.; Kim, K. K., An ubiquitin-binding molecule can work as an inhibitor of ubiquitin processing enzymes and ubiquitin receptors. *Biochemical and Biophysical Research Communications* **2016,** *479* (1), 33-9.

29. Nguyen, T.; Ho, M.; Kim, K.; Yun, S. I.; Mizar, P.; Easton, J. W.; Lee, S. S.; Kim, K. K., Suppression of the Ubiquitin Pathway by Small Molecule Binding to Ubiquitin Enhances Doxorubicin Sensitivity of the Cancer Cells. *Molecules* **2019,** *24* (6).

30. Cusack, J. C., Rationale for the treatment of solid tumors with the proteasome inhibitor bortezomib. *Cancer Treat Rev* **2003,** *29 Suppl 1*, 21-31.

31. Nawatha, M.; Rogers, J. M.; Bonn, S. M.; Livneh, I.; Lemma, B.; Mali, S. M.; Vamisetti, G. B.; Sun, H.; Bercovich, B.; Huang, Y.; Ciechanover, A.; Fushman, D.; Suga, H.; Brik, A., De novo macrocyclic peptides that specifically modulate Lys48-linked ubiquitin chains. *Nature Chemistry* **2019,** *11* (7), 644-652.

32. Kaplan, G. S.; Torcun, C. C.; Grune, T.; Ozer, N. K.; Karademir, B., Proteasome inhibitors in cancer therapy: Treatment regimen and peripheral neuropathy as a side effect. *Free Radical Biology and Medicine* **2017,** *103*, 1-13.

33. Verma, R., Ubistatins Inhibit Proteasome-Dependent Degradation by Binding the Ubiquitin Chain. *Science* **2004,** *306* (5693), 117-120.

34. Nakasone, M. A.; Lewis, T. A.; Walker, O.; Thakur, A.; Mansour, W.; Castañeda, C. A.; Goeckeler-Fried, J. L.; Parlati, F.; Chou, T. F.; Hayat, O.; Zhang, D.; Camara, C. M.; Bonn, S. M.; Nowicka, U. K.; Krueger, S.; Glickman, M. H.; Brodsky, J. L.; Deshaies, R. J.; Fushman, D., Structural Basis for the Inhibitory Effects of Ubistatins in the Ubiquitin-Proteasome Pathway. *Structure* **2017,** *25* (12), 1839-1855.e11.

35. Jongkees, S. A. K.; Hipolito, C. J.; Rogers, J. M.; Suga, H., Model foldamers: applications and structures of stable macrocyclic peptides identified using in vitro selection. *New Journal of Chemistry* **2015,** *39* (5), 3197-3207.

36. Hruby, V. J., Peptide Science: Exploring the Use of Chemical Principles and Interdisciplinary Collaboration for Understanding Life Processes†. *Journal of Medicinal Chemistry* **2003,** *46* (20), 4215-4231.

37. Purkayastha, A.; Kang, T. J., Stabilization of Proteins by Covalent Cyclization. *Biotechnology and Bioprocess Engineering* **2019,** *24* (5), 702-712.

38. Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M., Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal* **2000,** *19* (1), 94-102.

39. Rogers, J. M.; Nawatha, M.; Lemma, B.; Vamisetti, G. B.; Livneh, I.; Barash, U.; Vlodavsky, I.; Ciechanover, A.; Fushman, D.; Suga, H.; Brik, A., In vivo modulation of ubiquitin chains by N-methylated non-proteinogenic cyclic peptides. *RSC Chemical Biology* **2021**.

40. Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H., Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem Biol* **2011,** *18* (12), 1562-70.

41. Varadan, R.; Assfalg, M.; Raasi, S.; Pickart, C.; Fushman, D., Structural Determinants for Selective Recognition of a Lys48-Linked Polyubiquitin Chain by a UBA Domain. *Molecular Cell* **2005,** *18* (6), 687-698.

42. Pickart, C. M.; Raasi, S., Controlled Synthesis of Polyubiquitin Chains. In *Methods in Enzymology*, Elsevier: 2005; pp 21-36.

43. Williamson, M. P., Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013,** *73*, 1-16.

44. Latham, M. P.; Zimmermann, G. R.; Pardi, A., NMR Chemical Exchange as a Probe for Ligand-Binding Kinetics in a Theophylline-Binding RNA Aptamer. *Journal of the American Chemical Society* **2009,** *131* (14), 5052-5053.

45. Farmer, B. T.; Constantine, K. L.; Goldfarb, V.; Friedrichs, M. S.; Wittekind, M.; Yanchunas, J.; Robertson, J. G.; Mueller, L., Localizing the NADP+ binding site on the MurB enzyme by NMR. *Nature Structural Biology* **1996,** *3* (12), 995-997.

46. Renatus, M.; Parrado, S. G.; D'Arcy, A.; Eidhoff, U.; Gerhartz, B.; Hassiepen, U.; Pierrat, B.; Riedl, R.; Vinzenz, D.; Worpenberg, S.; Kroemer, M., Structural Basis of Ubiquitin Recognition by the Deubiquitinating Protease USP2. *Structure* **2006,** *14* (8), 1293-1302.

47. Faesen, C., Alex; Luna-Vargas, P. A., Mark; Geurink, P., Paul; Clerici, M.; Merkx, R.; Dijk, V., J., Willem; Hameed, S., Dharjath; Oualid, E., Farid; Ovaa, H.; Sixma, K., Titia, The Differential Modulation of USP Activity by Internal Regulatory Domains, Interactors and Eight Ubiquitin Chain Types. *Chemistry & Biology* **2011,** *18* (12), 1550-1561.

48. Mevissen, E. T., Tycho; Hospenthal, K., Manuela; Geurink, P., Paul; Elliott, R., Paul; Akutsu, M.; Arnaudo, N.; Ekkebus, R.; Kulathu, Y.; Wauer, T.; Oualid, E., Farid; Freund, M. V., Stefan; Ovaa, H.; Komander, D., OTU Deubiquitinases Reveal Mechanisms of Linkage Specificity and Enable Ubiquitin Chain Restriction Analysis. *Cell* **2013,** *154* (1), 169-184.

49. Pye, C. R.; Hewitt, W. M.; Schwochert, J.; Haddad, T. D.; Townsend, C. E.; Etienne, L.; Lao, Y.; Limberakis, C.; Furukawa, A.; Mathiowetz, A. M.; Price, D. A.; Liras, S.; Lokey, R. S., Nonclassical Size Dependence of Permeation Defines Bounds for Passive Adsorption of Large Drug Molecules. *Journal of Medicinal Chemistry* **2017,** *60* (5), 1665-1672.

50. Matsson, P.; Kihlberg, J., How Big Is Too Big for Cell Permeability? *Journal of Medicinal Chemistry* **2017,** *60* (5), 1662-1664.

51. Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J., Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chemistry & Biology* **2014,** *21* (9), 1115-42.

52. Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D. t.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W.; Duvall, J. R.; Kihlberg, J., Structural and conformational determinants of macrocycle cell permeability. *Nature Chemical Biology* **2016,** *12* (12), 1065-1074.

53. Naylor, M. R.; Bockus, A. T.; Blanco, M. J.; Lokey, R. S., Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Current Opinion in Chemical Biology* **2017,** *38*, 141-147.

54. Ryabov, Y.; Fushman, D., Interdomain mobility in di-ubiquitin revealed by NMR. *Proteins: Structure, Function, and Bioinformatics* **2006,** *63* (4), 787-796.

55. Krissinel, E.; Henrick, K., Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology* **2007,** *372* (3), 774-797.

56. Hurley, H., James; Lee, S.; Prag, G., Ubiquitin-binding domains. *Biochemical Journal* **2006,** *399* (3), 361-372.

57. Hu, M.; Li, P.; Song, L.; Jeffrey, P. D.; Chenova, T. A.; Wilkinson, K. D.; Cohen, R. E.; Shi, Y., Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *The EMBO Journal* **2005,** *24* (21), 3747-3756.

58. Mansour, W.; Nakasone, M. A.; Von Delbrück, M.; Yu, Z.; Krutauz, D.; Reis, N.; Kleifeld, O.; Sommer, T.; Fushman, D.; Glickman, M. H., Disassembly of Lys11 and Mixed Linkage Polyubiquitin Conjugates Provides Insights into Function of Proteasomal Deubiquitinases Rpn11 and Ubp6. *Journal of Biological Chemistry* **2015,** *290* (8), 4688-4704.

59. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **2012,** *9* (7), 671-675.

60. Reyes-Turcu, F. E.; Shanks, J. R.; Komander, D.; Wilkinson, K. D., Recognition of Polyubiquitin Isoforms by the Multiple Ubiquitin Binding Modules of Isopeptidase T. *Journal of Biological Chemistry* **2008,** *283* (28), 19581-19592.

61. Deshaies, R. J., Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. *BMC Biology* **2014,** *12* (1).

62. Gareau, J. R.; Lima, C. D., The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews Molecular Cell Biology* **2010,** *11* (12), 861-871.

63. Schreiber, A.; Peter, M., Substrate recognition in selective autophagy and the ubiquitin-proteasome system. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2014,** *1843* (1), 163-81.

64. Hochstrasser, M., Origin and function of ubiquitin-like proteins. *Nature* **2009,** *458* (7237), 422-429.

65. Ichimura, Y.; Kirisako, T.; Takao, T.; Satomi, Y.; Shimonishi, Y.; Ishihara, N.; Mizushima, N.; Tanida, I.; Kominami, E.; Ohsumi, M.; Noda, T.; Ohsumi, Y., A ubiquitin-like system mediates protein lipidation. *Nature* **2000,** *408* (6811), 488-492.

66. Whitby, F. G.; Xia, G.; Pickart, C. M.; Hill, C. P., Crystal Structure of the Human Ubiquitin-like Protein NEDD8 and Interactions with Ubiquitin Pathway Enzymes. *Journal of Biological Chemistry* **1998,** *273* (52), 34983-34991.

67. Rao-Naik, C.; Delacruz, W.; Laplaza, J. M.; Tan, S.; Callis, J.; Fisher, A. J., The Rub Family of Ubiquitin-like Proteins. *Journal of Biological Chemistry* **1998,** *273* (52), 34976-34982.

68. Singh, R. K.; Zerath, S.; Kleifeld, O.; Scheffner, M.; Glickman, M. H.; Fushman, D., Recognition and cleavage of related to ubiquitin 1 (Rub1) and Rub1-ubiquitin chains by components of the ubiquitin-proteasome system. *Molecular & Cellular Proteomics* **2012,** *11* (12), 1595-1611, 17 pp.

69. Hochstrasser, M., Biochemistry. All in the ubiquitin family. *Science* **2000,** *289* (5479), 563-4.

70. Liakopoulos, D., A novel protein modification pathway related to the ubiquitin system. *The EMBO Journal* **1998,** *17* (8), 2208-2214.

71. Gurevich, S. Z.; Sinha, A.; Longworth, J.; Singh, R. K.; Lemma, B. E.; Thakur, A.; Popp, O.; Kornitzer, D.; Reis, N.; Scheffner, M.; Dittmar, G.; Pick, E.; Fushman, D.; Glickman, M. H., Rub1/NEDD8, a ubiquitin-like modifier, is also a ubiquitin modifier. Cold Spring Harbor Laboratory: 2020.

72. Hori, T.; Osaka, F.; Chiba, T.; Miyamoto, C.; Okabayashi, K.; Shimbara, N.; Kato, S.; Tanaka, K., Covalent modification of all members of human cullin family proteins by NEDD8. *Oncogene* **1999,** *18* (48), 6829-6834.

73. Ou, Y.; Rattner, J. B., The centrosome in higher organisms: structure, composition, and duplication. *International Review of Cytology* **2004,** *238*, 119-82.

74. Fischer, M.; Dang, C. V.; DeCaprio, J.; A., "Control of Cell Division", Hematology. Seventh Edition ed.; Elsevier: 2018; pp 176-185.

75. Rabut, G.; Peter, M., Function and regulation of protein neddylation. *EMBO reports* **2008,** *9* (10), 969-976.

76. Merlet, J.; Burger, J.; Gomes, J.-E.; Pintard, L., Regulation of cullin-RING E3 ubiquitin-ligases by neddylation and dimerization. *Cellular and Molecular Life Sciences* **2009,** *66* (11-12), 1924-1938.

77. Abida, W. M.; Nikolaev, A.; Zhao, W.; Zhang, W.; Gu, W., FBXO11 Promotes the Neddylation of p53 and Inhibits Its Transcriptional Activity. *Journal of Biological Chemistry* **2007,** *282* (3), 1797-1804.

78. Harper, J. W., Neddylating the Guardian. *Cell* **2004,** *118* (1), 2-4.

79. Xirodimas, D. P.; Saville, M. K.; Bourdon, J.-C.; Hay, R. T.; Lane, D. P., Mdm2-Mediated NEDD8 Conjugation of p53 Inhibits Its Transcriptional Activity. *Cell* **2004,** *118* (1), 83-97.

80. Watson, I. R.; Blanch, A.; Lin, D. C. C.; Ohh, M.; Irwin, M. S., Mdm2-mediated NEDD8 Modification of TAp73 Regulates Its Transactivation Function*. *Journal of Biological Chemistry* **2006,** *281* (45), 34096-34103.

81. Oved, S.; Mosesson, Y.; Zwang, Y.; Santonico, E.; Shtiegman, K.; Marmor, M. D.; Kochupurakkal, B. S.; Katz, M.; Lavi, S.; Cesareni, G.; Yarden, Y., Conjugation to Nedd8 Instigates Ubiquitylation and Down-regulation of Activated Receptor Tyrosine Kinases. *Journal of Biological Chemistry* **2006,** *281* (31), 21640-21651.

82. Shimura, H.; Hattori, N.; Kubo, S.; Mizuno, Y.; Asakawa, S.; Minoshima, S.; Shimizu, N.; Iwai, K.; Chiba, T.; Tanaka, K.; Suzuki, T., Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nature Genetics* **2000,** *25* (3), 302-5.

83. Leidecker, O.; Matic, I.; Mahata, B.; Pion, E.; Xirodimas, D. P., The ubiquitin E1 enzyme Ube1 mediates NEDD8 activation under diverse stress conditions. *Cell Cycle* **2012,** *11* (6), 1142-1150.

84. Lane, D. P., Stress, specificity and the NEDD8 proteome. *Cell Cycle* **2012,** *11* (8), 1488-1489.

85. Hjerpe, R.; Thomas, Y.; Chen, J.; Zemla, A.; Curran, S.; Shpiro, N.; Dick, L. R.; Kurz, T., Changes in the ratio of free NEDD8 to ubiquitin triggers NEDDylation by ubiquitin enzymes. *Biochemical Journal* **2012,** *441* (3), 927-36.

86. Bohnsack, R. N.; Haas, A. L., Conservation in the Mechanism of Nedd8 Activation by the Human AppBp1-Uba3 Heterodimer. *Journal of Biological Chemistry* **2003,** *278* (29), 26823-26830.

87. Walden, H.; Podgorski, M. S.; Huang, D. T.; Miller, D. W.; Howard, R. J.; Minor, D. L.; Holton, J. M.; Schulman, B. A., The Structure of the APPBP1-UBA3-NEDD8-ATP Complex Reveals the Basis for Selective Ubiquitin-like Protein Activation by an E1. *Molecular Cell* **2003,** *12* (6), 1427-1437.

88. Ohki, Y.; Funatsu, N.; Konishi, N.; Chiba, T., The mechanism of poly-NEDD8 chain formation in vitro. *Biochemical and Biophysical Research Communications* **2009,** *381* (3), 443-7.

89. Cook, B. W.; Barber, K. R.; Shilton, B. H.; Shaw, G. S., The HIP2~ubiquitin conjugate forms a non-compact monomeric thioester during di-ubiquitin synthesis. *PLoS One* **2015,** *10* (3), e0120318.

90. Schnell, J. R.; Dyson, H. J.; Wright, P. E., Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annual review of biophysics and biomolecular structure* **2004,** *33*, 119-40.

91. Wilson, R. C.; Edmondson, S. P.; Flatt, J. W.; Helms, K.; Twigg, P. D., The E2-25K ubiquitin-associated (UBA) domain aids in polyubiquitin chain synthesis and linkage specificity. *Biochemical and Biophysical Research Communications* **2011,** *405* (4), 662-666.

92. Mueller, T. D.; Kamionka, M.; Feigon, J., Specificity of the Interaction between Ubiquitin-associated Domains and Ubiquitin. *Journal of Biological Chemistry* **2004,** *279* (12), 11926-11936.

93. Ryu, K.-S.; Lee, K.-J.; Bae, S.-H.; Kim, B.-K.; Kim, K.-A.; Choi, B.-S., Binding Surface Mapping of Intra- and Interdomain Interactions among hHR23B, Ubiquitin, and Polyubiquitin Binding Site 2 of S5a. *Journal of Biological Chemistry* **2003,** *278* (38), 36621-36627.

94. Wilson, R. C.; Hughes, R. C.; Flatt, J. W.; Meehan, E. J.; Ng, J. D.; Twigg, P. D., Structure of full-length ubiquitin-conjugating enzyme E2-25K (huntingtin-interacting protein 2). *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **2009,** *65* (5), 440-444.

95. 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.; Lee, W., Structural Basis of E2–25K/UBB+1 Interaction Leading to Proteasome Inhibition and Neurotoxicity. *Journal of Biological Chemistry* **2010,** *285* (46), 36070-36080.

96. Choi, Y.-S.; Jeon, Y. H.; Ryu, K.-S.; Cheong, C., 60th residues of ubiquitin and Nedd8 are located out of E2-binding surfaces, but are important for K48 ubiquitin-linkage. *FEBS Letters* **2009,** *583* (20), 3323-3328.

97. Peng, J.; Schwartz, D.; Elias, J. E.; Thoreen, C. C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S. P., A proteomics approach to understanding protein ubiquitination. *Nature Biotechnology* **2003,** *21* (8), 921-926.

98. Kim, W.; Bennett, J., Eric; Huttlin, L., Edward; Guo, A.; Li, J.; Possemato, A.; Sowa, E., Mathew; Rad, R.; Rush, J.; Comb, J., Michael; Harper, W., J.; Gygi, P., Steven, Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. *Molecular Cell* **2011,** *44* (2), 325-340.

99. Udeshi, N. D.; Mertins, P.; Svinkina, T.; Carr, S. A., Large-scale identification of ubiquitination sites by mass spectrometry. *Nature Protocols* **2013,** *8* (10), 1950-1960.

100. Huynh, M.-L.; Russell, P.; Walsh, B., Tryptic Digestion of In-Gel Proteins for Mass Spectrometry Analysis. In *Methods in Molecular Biology*, Humana Press: 2009; pp 507-513.

101. Lee, A. E.; Geis-Asteggiante, L.; Dixon, E. K.; Kim, Y.; Kashyap, T. R.; Wang, Y.; Fushman, D.; Fenselau, C., Preparing to read the ubiquitin code: characterization of ubiquitin trimers by top-down mass spectrometry. *Journal of Mass Spectrometry* **2016,** *51* (4), 315-321.

102. Cannon, J. R.; Martinez-Fonts, K.; Robotham, S. A.; Matouschek, A.; Brodbelt, J. S., Top-Down 193-nm Ultraviolet Photodissociation Mass Spectrometry for Simultaneous Determination of Polyubiquitin Chain Length and Topology. *Analytical Chemistry* **2015,** *87* (3), 1812-1820.

103. Ansong, C.; Wu, S.; Meng, D.; Liu, X.; Brewer, H. M.; Deatherage Kaiser, B. L.; Nakayasu, E. S.; Cort, J. R.; Pevzner, P.; Smith, R. D.; Heffron, F.; Adkins, J. N.; Pasa-Tolic, L., Top-down proteomics reveals a unique protein S-thiolation switch in Salmonella Typhimurium in response to infection-like conditions. *Proceedings of the National Academy of Sciences* **2013,** *110* (25), 10153-10158.

104. Chen, D.; Gomes, F.; Abeykoon, D.; Lemma, B.; Wang, Y.; Fushman, D.; Fenselau, C., Top-Down Analysis of Branched Proteins Using Mass Spectrometry. *Analytical Chemistry* **2018,** *90* (6), 4032-4038.

105. Lee, A. E.; Geis-Asteggiante, L.; Dixon, E. K.; Miller, M.; Wang, Y.; Fushman, D.; Fenselau, C., Preparing to read the ubiquitin code: top-down analysis of unanchored ubiquitin tetramers. *Journal of Mass Spectrometry* **2016,** *51* (8), 629-637.

106. Gomes, F.; Lemma, B.; Abeykoon, D.; Chen, D.; Wang, Y.; Fushman, D.; Fenselau, C., Top-down analysis of novel synthetic branched proteins. *Journal of Mass Spectrometry* **2019,** *54* (1), 19-25.

107. Bremm, A.; Freund, S. M. V.; Komander, D., Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nature Structural & Molecular Biology* **2010,** *17* (8), 939-947.

108. Gill, P.; Moghadam, T. T.; Ranjbar, B., Differential scanning calorimetry techniques: applications in biology and nanoscience. *Journal of Biomolecular Techniques* **2010,** *21* (4), 167-93.

109. Nir, I.; Huttner, D.; Meller, A., Direct Sensing and Discrimination among Ubiquitin and Ubiquitin Chains Using Solid-State Nanopores. *Biophysical Journal* **2015,** *108* (9), 2340-2349.

110. Greenfield, N. J., Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **2006,** *1* (6), 2876-2890.

111. Ohnishi, M.; Urry, D. W., Temperature dependence of amide proton chemical shifts: The secondary structures of gramicidin S and valinomycin. *Biochemical and Biophysical Research Communications* **1969,** *36* (2), 194-202.

112. Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z.; Tong, H., Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging. *Journal of the American Chemical Society* **1997,** *119* (36), 8547-8561.

113. Baxter, N. J.; Williamson, M. P., *Journal of Biomolecular NMR* **1997,** *9* (4), 359-369.

114. Baxter, N. J.; Hosszu, L. L. P.; Waltho, J. P.; Williamson, M. P., Characterisation of low free-energy excited states of folded proteins 1 1Edited by P. E. Wright. *Journal of Molecular Biology* **1998,** *284* (5), 1625-1639.

115. Doyle, C. M.; Rumfeldt, J. A.; Broom, H. R.; Sekhar, A.; Kay, L. E.; Meiering, E. M., Concurrent Increases and Decreases in Local Stability and Conformational Heterogeneity in Cu, Zn Superoxide Dismutase Variants Revealed by Temperature-Dependence of Amide Chemical Shifts. *Biochemistry* **2016,** *55* (9), 1346-1361.

116. Cierpicki, T.; Otlewski, J., *Journal of Biomolecular NMR* **2001,** *21* (3), 249-261.

117. Cierpicki, T.; Zhukov, I.; Byrd, R. A.; Otlewski, J., Hydrogen Bonds in Human Ubiquitin Reflected in Temperature Coefficients of Amide Protons. *Journal of Magnetic Resonance* **2002,** *157* (2), 178-180.

118. Kumar, A.; Srivastava, S.; Hosur, R. V., NMR Characterization of the Energy Landscape of SUMO-1 in the Native-state Ensemble. *Journal of Molecular Biology* **2007,** *367* (5), 1480-1493.

119. Hoffman, R. E., Standardization of chemical shifts of TMS and solvent signals in NMR solvents. *Magnetic Resonance in Chemistry* **2006,** *44* (6), 606-616.

120. Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A., Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *Journal of the American Chemical Society* **1998,** *120* (27), 6836-6837.

121. Tomlinson, J. H.; Williamson, M. P., Amide temperature coefficients in the protein G B1 domain. *Journal of Biomolecular NMR* **2012,** *52* (1), 57-64.

122. Vallurupalli, P.; Bouvignies, G.; Kay, L. E., Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *Journal of the American Chemical Society* **2012,** *134* (19), 8148-8161.

123. Forsén, S.; Hoffman, R. A., Study of Moderately Rapid Chemical Exchange Reactions by Means of Nuclear Magnetic Double Resonance. *The Journal of Chemical Physics* **1963,** *39* (11), 2892-2901.

124. Guinier, A., La diffraction des rayons X aux très petits angles : application à l'étude de phénomènes ultramicroscopiques. *Annales de Physique* **1939,** *11*, 161.

125. Jacrot, B.; Zaccai, G., Determination of molecular weight by neutron scattering. *Biopolymers* **1981,** *20* (11), 2413-2426.

126. Cook, W. J.; Jeffrey, L. C.; Kasperek, E.; Pickart, C. M., Structure of tetraubiquitin shows how multiubiquitin chains can be formed. *Journal of Molecular Biology* **1994,** *236* (2), 601-9.

127. Hiranyakorn, M.; Yanaka, S.; Satoh, T.; Wilasri, T.; Jityuti, B.; Yagi-Utsumi, M.; Kato, K., NMR Characterization of Conformational Interconversions of Lys48-Linked Ubiquitin Chains. *International Journal of Molecular Sciences* **2020,** *21* (15), 5351.

128. Johnson, E. S.; Bartel, B.; Seufert, W.; Varshavsky, A., Ubiquitin as a degradation signal. *The EMBO Journal* **1992,** *11* (2), 497-505.

129. Schultz, L. B.; Chehab, N. H.; Malikzay, A.; Halazonetis, T. D., P53 Binding Protein 1 (53bp1) Is an Early Participant in the Cellular Response to DNA Double-Strand Breaks. *Journal of Cell Biology* **2000,** *151* (7), 1381-1390.

130. Fradet-Turcotte, A.; Canny, M. D.; Escribano-Díaz, C.; Orthwein, A.; Leung, C. C. Y.; Huang, H.; Landry, M.-C.; Kitevski-Leblanc, J.; Noordermeer, S. M.; Sicheri, F.; Durocher, D., 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **2013,** *499* (7456), 50-54.

131. Botuyan, M. V.; Lee, J.; Ward, I. M.; Kim, J.-E.; Thompson, J. R.; Chen, J.; Mer, G., Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell* **2006,** *127* (7), 1361-1373.

132. Huyen, Y.; Zgheib, O.; Ditullio Jr, R. A.; Gorgoulis, V. G.; Zacharatos, P.; Petty, T. J.; Sheston, E. A.; Mellert, H. S.; Stavridi, E. S.; Halazonetis, T. D., Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **2004,** *432* (7015), 406-411.

133. Lukas, J.; Lukas, C.; Bartek, J., More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. *Nature Cell Biology* **2011,** *13* (10), 1161-1169.

134. Jackson, P., Stephen; Durocher, D., Regulation of DNA Damage Responses by Ubiquitin and SUMO. *Molecular Cell* **2013,** *49* (5), 795-807.

135. Varadan, R.; Assfalg, M.; Haririnia, A.; Raasi, S.; Pickart, C.; Fushman, D., Solution Conformation of Lys63-linked Di-ubiquitin Chain Provides Clues to Functional Diversity of Polyubiquitin Signaling. *Journal of Biological Chemistry* **2004,** *279* (8), 7055-7063.

136. Varadan, R.; Walker, O.; Pickart, C.; Fushman, D., Structural Properties of Polyubiquitin Chains in Solution. *Journal of Molecular Biology* **2002,** *324* (4), 637-647.

137. Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015,** *31* (8), 1325-1327.

138. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004,** *25* (13), 1605-1612.

139. Castañeda, A., Carlos; Dixon, K., Emma; Walker, O.; Chaturvedi, A.; Nakasone, A., Mark; Curtis, E., Joseph; Reed, R., Megan; Krueger, S.; Cropp, A., T.; Fushman, D., Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. *Structure* **2016,** *24* (3), 423-436.

140. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W., iMOSFLM: a new graphical interface for diffraction-image processing withMOSFLM. *Acta Crystallographica Section D Biological Crystallography* **2011,** *67* (4), 271-281.

141. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography* **2013,** *69* (7), 1204-1214.

142. Mccoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phasercrystallographic software. *Journal of Applied Crystallography* **2007,** *40* (4), 658-674.

143. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* **2004,** *60* (12), 2126-2132.

144. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development ofCoot. *Acta Crystallographica Section D Biological Crystallography* **2010,** *66* (4), 486-501.

145. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; Mccoy, A.; Mcnicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of theCCP4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography* **2011,** *67* (4), 235-242.

146. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Echols, N.; Headd, J. J.; Hung, L.-W.; Jain, S.; Kapral, G. J.; Grosse Kunstleve, R. W.; Mccoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., The Phenix software for automated determination of macromolecular structures. *Methods* **2011,** *55* (1), 94-106.

147. Minor, W.; Cymborowski, M.; Otwinowski, Z.; Chruszcz, M., HKL-3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Crystallographica Section D Biological Crystallography* **2006,** *62* (8), 859-866.

148. Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J., Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology* **1987,** *194* (3), 531-544.

149. Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography* **2011,** *67* (4), 355-367.

150. Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N., JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallographica Section D Biological Crystallography* **2012,** *68* (4), 431-440.

151. Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* **2010,** *66* (1), 12-21.