Analytical Methods to Study Ubiquitin Signaling

By

Ellen M. Valkevich

A dissertation submitted in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

(Chemistry)

at the

University of Wisconsin-Madison

2015

Date of final oral examination: January 15, 2015

This dissertation is approved by the following members of the Final Oral Committee:
Eric R. Strieter, Assistant Professor, Department of Chemistry
Helen E. Blackwell, Professor, Departments of Chemistry and Biochemistry
Samuel H. Gellman, Professor, Department of Chemistry
Ying Ge, Assistant Professor, Departments of Chemistry and Cell and Regenerative Biology
Tehshik P. Yoon, Professor, Department of Chemistry

In honor and memory of my dad, Bob Valkevich

Acknowledgments

I have many people I would like to thank. First, my advisor, Eric Strieter, has helped me develop as a scientist throughout my years of grad school, and encouraged me to keep going, even through the hard parts. I especially appreciate his passion for science, which is often contagious. He settles for nothing but the best, and instills this attitude in the lab.

I would also like to thank Professor Ying Ge. Ying has been a wonderful collaborator for the past 4 years, and is always available to share her mass spec knowledge with me. The members of the Ge lab have also been very helpful over the years, always there to lend a helping hand when the instrument is acting fussy, or when I have questions about my experiments.

I am so appreciative of many helpful people in the chemistry department over the years, especially older grad students from other labs, who provided guidance (as well as reagents and instruments) to our young lab, when we were starting out.

I am glad to have had the pleasure of working with the world's best lab mates. I send HUGE thanks to all Strieter lab members, past and present. The group has provided me a little Madison family through all of the ups and downs that come with studying science. Thanks for providing me with many buffers and baked goods. I will especially miss our fun discussions, many being scientific, but many filling out the "philosophy" portion of the PhD. (Julia is an honorary lab member. Although she had a different advisor, she did spend a lot of time using our sonicator, and keeping me company at Graze.) Nick, I hope you read my thesis someday! Thanks for always climbing trees with me, and for making buffers, and filling pipette boxes. I loved teaching you all about ubiquitin, and getting to learn everything else from all of your classes in exchange. HP, thanks for being

the best desk neighbor. You have spoiled me. Sean. I'm glad you like T Swift. Ambar, thanks for being awesome.

And finally, I would not have made it through grad school without the wholehearted support from my friends (near and far) and family. Alice, Caitlin, and Sarah, you have been there with me even through crazy thesis writing, always checking in, and putting a smile on my face. I know you all understand the challenges and successes of academia, and it has been amazing having you all here, every step of the way. Thanks to my family, for putting up with my absence for the past 5 years, and for welcoming me back to CA with open arms. Mom and Jeannie, you have been my biggest cheerleaders, thanks for making the trek to WI (in January!) to be here for me. I'll be back soon, and Grant and Owen will have the coolest 5th grade science fair project this year.

Table of Contents

List of Figures

List of Abbreviations

AMC	7-amino-4-methyl coumarin
BRCA1	breast cancer associated protein
CAD	collision-activated dissociation
CoSMoS	colocalization single molecule spectroscopy
DDR	DNA damage response
DUB	deubiquitinase
ECD	electron capture dissociation
FT-ICR	fourier-transform ion cyclotron
IPTG	β-D-1-thiogalactopyranoside
IsoT	isopeptidase T
JAMM/MPN+	Josephins and JAB1/ MPN/MOV34 metalloenzyme
LAP	lithium acyl phosphinate
LB	luria broth
MALDI-TOF	matrix assisted laser desorption ionization-time of flight
MS	mass spectrometric
NCL	native chemical ligation
NEM	N-ethylmaleimide
OTU	ovarian tumor protease
PEG	poly(ethylene) glycol
PMSF	phenylmethanesulfonylfluoride
SEC	Size exclusion chromatography
SMF	single molecule fluorescence
SPPS	solid phase peptide synthesis
SUMO	small ubiquitin modifier
TAMRA	tetramethylrhodamine
TCEP	tris-(2-carboxyethyl)phosphine
TEC	thiol–ene coupling
TIRF	total internal reflection fluorescence
Ub	ubiquitin
UBD	ubiquitin binding domain
UCH	C-terminal hydrolase
UIM	ubiquitin interacting motif
UPS	ubiquitin proteasome system
USP	ubiquitin-specific protease
Yuh1	yeast ubiquitin C-terminal hydrolase

Chapter 1: Introduction to Ubiquitin Signaling

1.1 Introduction

Intracellular communication is vital to life. When DNA is damaged, warning signals call for an enzymatic cascade to come repair it. When proteins are no longer needed in the cell, they are flagged to be recycled. Deviation from this highly regulated process can cause chaos and malfunction. An important mechanism of regulation is post-translational modification (PTM) of proteins. PTMs function by changing the surface of the modified protein and fine-tuning the flux through different biochemical pathways. One important PTM is the small protein ubiquitin (Ub), which has been shown to regulate DNA damage response, protein degradation, and a host of other cellular functions[1,2]. My graduate work focuses on providing new chemical tools to improve understanding of Ub signaling. In this chapter, we describe what is known about the language of Ub and specifically highlight what we have added to better understand Ub biology.

Ubiquitin

Ubiquitin (Ub) is a small and highly stable 76 amino acid protein. The C-terminus of Ub can be conjugated via an isopeptide bond to the ε-amino group on a substrate protein (termed mono-ubiquitination). The attachment of Ub to a substrate protein creates an additional recognition site for binding partners and can subsequently change the biological fate of a substrate. Ub can also act as a substrate, leading to the formation of Ub chains linked at one of the seven lysine residues (K6,

K11, K27, K29, K33, K48, K63) or the N-terminus (M1) (Figure 1.1)[3,4,5]. Myriad linkage combinations lead to a diversity of Ub chain types.

The substrate scope for Ub is vast. In 2003, Gygi and coworkers identified 1075 potential ubiquitinated proteins from *Saccharomyces cerevisiae* through a proteomics study[6]. Gygi later used an antibody to enrich for ubiquitinated proteins and identified 19,000 modified lysine residues in 5,000 proteins[7]. This prevalence of Ub, underscores the importance of understanding the causes and the outcomes of ubiquitination in different situations.

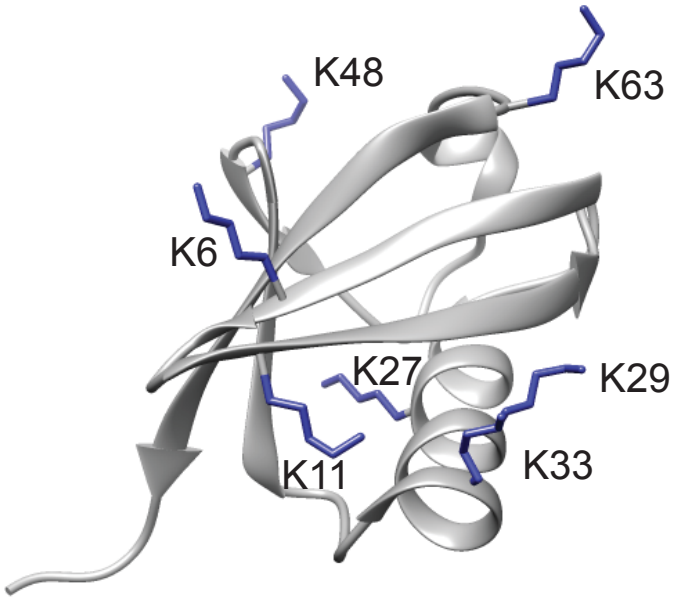


Figure 1.1. The crystal structure of Ub (PDB: 1ubq)

Mechanism of Ub attachment

The covalent attachment of Ub to a substrate protein is facilitated by three enzymes: E1 Ub-activating enzyme, E2 Ub-conjugating enzyme, and E3 Ub ligating enzyme (Figure 1.2)[8,9]. E1 activates the C-terminus of Ub by forming an adenylate

and subsequently forms a thioester bond to Ub via the E1 active site cysteine. Ub is then transferred to an active site cysteine of an E2 enzyme. The ubiquitinated E2 and a substrate for ubiquitination are brought together by the E3, and Ub is then transferred to the ε-amino group of a lysine residue on the substrate. Deubiquitinating enzymes (DUBs) cleave Ub chains to reverse the modification. There are eight E1s[10], 35 E2s[11], ~600 E3s[9], and ~100 DUBs[12] in humans. Overall, these enzymes control substrate specificity and Ub chain topology.

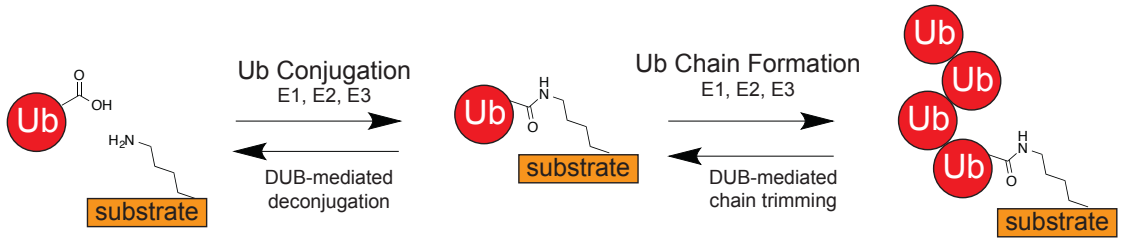


Figure 1.2. Mechanism of Ub modification

Ub Chain Topology

Ub has seven lysine residues and can polymerize in a variety of combinations, leading to Ub chains with different topologies. Homotypic Ub chains extend by attachment to the same lysine of each monomer. Heterotypic linear chains are comprised of isopeptide linkages from different lysine residues of each Ub subunit[13,14,15]. Branched Ub chains contain Ub monomers that are modified at two or more lysine residues[15, 16,17]. (Figure 1.3)

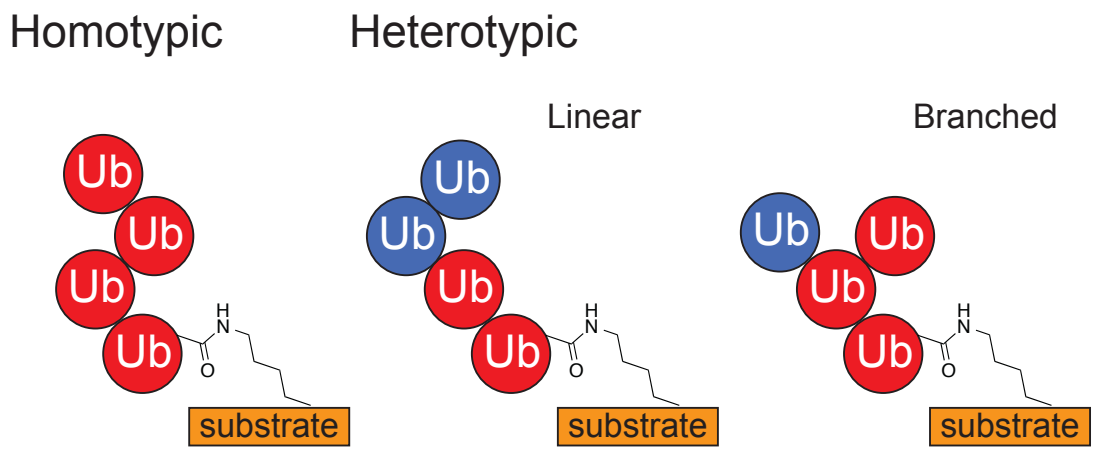


Figure 1.3. Ubiquitin chains linkage can vary, leading to different possible chain topologies.

Ub Chain Function

Different chain topologies have been linked to distinct cellular functions. Lys48 linked chains escort substrates to the 26S proteasome and lead to degradation of the ubiquitinated protein[18,19]. Lys63 linked chains are involved in the immune response[20] and DNA damage response[21,22]. Lys11 linked chains are proteolytic signals as well, specifically during mitosis[23,24].

The diversity in function arises from different chain topologies having different three-dimensional structures, thus different binding partners. A hydrophobic patch on Ub (Ile44, Leu8, Val70, His68) dictates many of its binding interactions[25,26,27] (Figure 1.4). The hydrophobic patch is displayed in different orientations depending on the chain linkage[28,29,30,31].

In order to study the functions of different Ub chain types, it is helpful to have access to chains of defined length and linkage for biochemical assays. It is also

helpful to have analytical techniques to observe chain branching. The projects presented here aim to advance the Ub field through both of these strategies. We have developed a method to chemically synthesize well-defined, biomimetic Ub chains that employs a radical thiol-ene coupling (TEC) reaction. We have also developed a middle-down mass spectrometry (MS) technique to characterize Ub linkage type. This method has been expanded to characterize linear and branched Ub chains. Prior to discussing the advances we have made, it is important to put them into context.

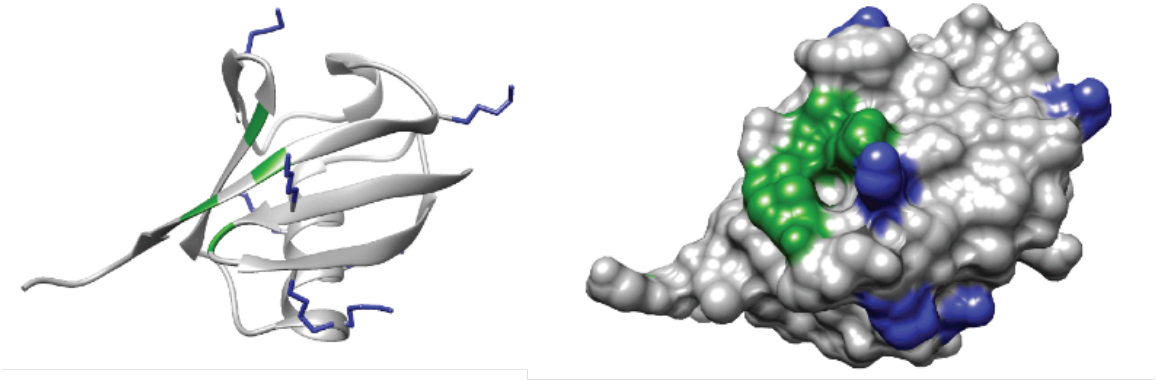


Figure 1.4. Ubiquitin crystal structure. Lysine residues are in blue and the hydrophobic patch (L8, I44, H68, V70) is in green.

1.2 Synthesis of Ub chains

Enzymatic Methods

One of the most effective ways to synthesize Ub chains is by exploiting the capabilities of enzymes that have inherent linkage selectivity. Pickart and co-workers were the first to demonstrate the synthetic utility of Ub conjugating enzymes[32]. Since then, four enzymes have been used to build free (unattached to

substrate protein) Ub chains: K48-specific UBE2K (UbcH1, E2-25K)[33], K48-specific Cdc34[34], K63 specific Ubc13Mms2[35,36], and K11-specific UBE2S[37,38].

Linkage specificity is dictated by the conjugating enzyme by restricting orientation of the approaching of the Ub monomers. Length specificity was originally achieved by combining linkage specific conjugating enzymes with a series or protection and deprotection steps, both enzymatic and chemical, to synthesize Ub oligomers[32]. With current purification techniques, chains of varied length can be synthesized together, and then separated by size exclusion chromatography, yielding milligram quantities of chains[39]. Access to these substrates opened up the field to study enzymatic activity on defined chains[40,41,30]. However, there are many Ub chain types that are not accessible through enzymatic synthesis, and this limitation has led to an expansion of chemical and semi-chemical chain synthesis strategies to generate Ub chains.

Chemical Methods

Chemical approaches towards the synthesis of defined Ub chains overcome many of the limitations associated with enzymatic methods. For example, chemical methods are not limited to specific isopeptide linkages. To achieve regioselectivity with chemical methods, the lysine residue at the desired linkage site must be differentiated from the other six lysine residues on Ub. This has been accomplished by incorporating artificial amino acids at the lysine of interest, and also by incorporating differential protecting groups on the lysine residues[42,43]. There are two main methods to install artificial amino acids into Ub: the entire protein can be

synthesized by peptide synthesis methods[44], or Ub can be expressed and purified from cells containing an orthogonal tRNA/tRNA synthetase pair that is optimized for artificial amino acid incorporation[45]. Brik and coworkers employ the former to afford a native isopeptide bond[46,47], while Chin and co-workers have focused on efforts to efficiently incorporate an unnatural amino acid into recombinant Ub[43]. Chemically synthesized Ub oligomers have played an instrumental role in defining the selectivity of DUBs[43,48]; however, the chemical synthesis of Ub chains is very low yielding as a result of many inefficient steps in the process. Thus this approach is impractical for generating large quantities of Ub chains that are necessary for biochemical and structural analyses.

1.3 Characterization of Atypical Ub Chain Topologies

The biological presence and function of branched Ub chains are poorly understood because methods for generating and characterizing branching are underdeveloped. Branching at adjacent lysine residues on Ub has been observed by standard Ub proteomic techniques[49,50]. Trypsin will not cleave after a lysine residue that is engaged in an isopeptide bond. Thus if modifications occur on adjacent lysine residues, both of the characteristic GG Ub motifs can be observed stemming from one peptide (Figure 1.5). If the branching lysine residues contain a tryptic digestion point between them, then the GG modifications are separated in the digestion process, and the resulting peptides are indistinguishable from those derived from mixed linear Ub chains. New analytical methods that are amenable to the characterization of branched chains are important because Rape and co-workers

recently showed that these Ub conjugates enhance proteasomal degradation during mitosis[16].

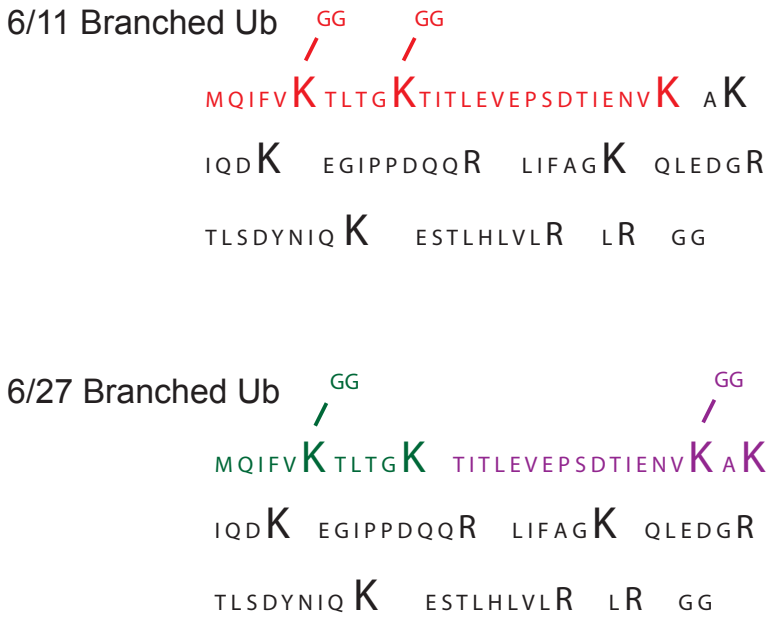


Figure 1.5. Peptides resulting from full tryptic digest. Branching on adjacent lysine residues (6/11 branched Ub) results in one peptide with two GG modifications (red). Branching on non-adjacent lysine residues (6/27 branched Ub) leads to a division of GG modifications onto two separate peptides (green and purple).

1.4 Ub chain characterization methods

Hampering our understanding of Ub chain branching is a lack of characterization techniques. Analytical techniques that have provided the most insight to connect Ub linkage to function include, linkage specific antibodies, the use of K-to-R Ub variants, and mass spectrometry (MS). Each technique has limitations in the context of characterizing branched chains.

Linkage specific antibodies are useful for gaining knowledge about the co-localization of specific Ub chains and other cellular proteins, but antibodies are only available for M1[51], K11[52], K48, and K63[53] linked chains and therefore report on the presence of a linkage between two Ub subunits. The use of Ub K-to-R variants blocks chain extension through specific sites, thus preventing downstream events associated with a particular polyUb chain[54,55]. By knocking out specific points of chain attachment, phenotypic observations can be tracked to the absence of a specific Ub chain linkage. However, the dynamics of chain formation and recognition can be affected when native residues are altered. Furthermore, *in cellular* studies with KxR Ub is prone to contamination by residual wild type Ub. Mass spectrometry (MS) is arguably the most powerful approach to distinguish which linkages are present in a given sample of Ub chains. MS techniques, as described below, have afforded a wealth of information regarding the identity of ubiquitinated protein substrates.

Bottom-up mass spectrometry

Bottom-up MS is a highly used technique to probe biological systems[56,57]. Advancements in instrumentation and reference database analysis have opened up the field of proteomics to high throughput analysis that can detect even low levels of proteins in any system.

A typical sample preparation and analysis involves: obtaining a protein mixture of interest, digesting the proteins to small peptide fragments, injecting the peptide mixture into a chromatography column (reverse phase chromatography is

widely used), ionization of the sample, detection of the mass of the peptides, and in situ fragmentation of abundant peptides to discover their amino acid sequence. Detected peptides are searched through a database that matches predicted peptides to full proteins. This type of analysis is helpful in identifying protein binding partners or complexes, or looking at the presence/absence of a particular protein in a given system. Developments in quantification techniques also provide insight into the relative and absolute levels of proteins under different biological conditions[58,59], e.g. when comparing diseased and healthy samples.

Bottom-up MS has contributed to the Ub field by identifying chain linkages and ubiquitinated substrates[60]. Access to this information is possible because of the distinct structure of the Ub isopeptide linkage. The ε-amino group of the lysine residue is attached to the C-terminal glycine residue of Ub. When Ub chains are digested with trypsin, peptide bonds after basic residues are hydrolyzed, releasing Arg74, and leaving any ubiquitinated lysines still tethered to G75G76 via an isopeptide bond. This canonical GG marker has been tracked extensively to provide much of the current insight into Ub chain linkage. Since bottom up MS involves full digestion of proteins to peptides, information regarding multiple modifications on one protein, e.g. chain branching, cannot be unambiguously identified.

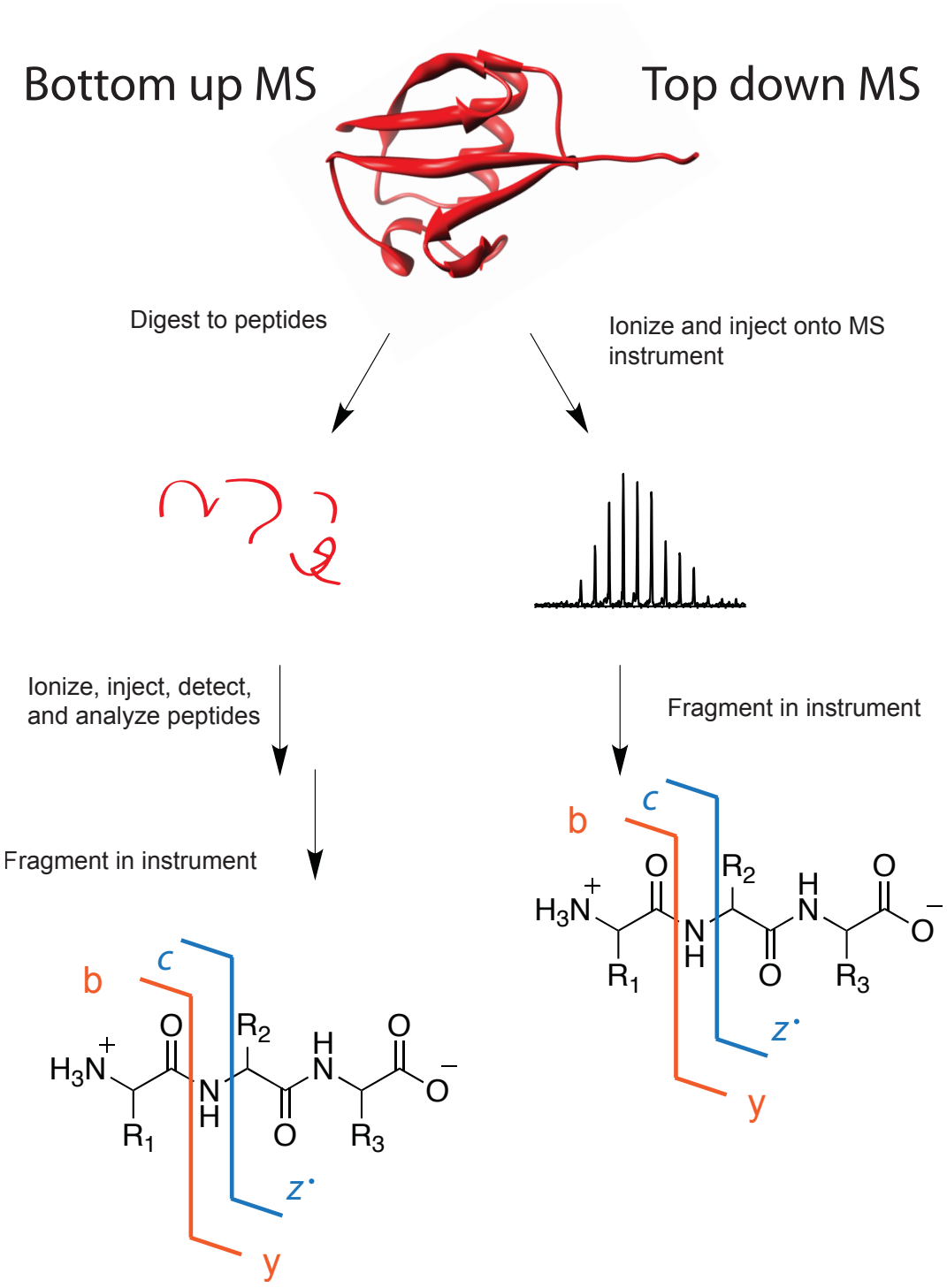


Figure 1.6. Bottom-up versus top-down MS techniques

Top-down mass spectrometry

Top-down MS techniques have been developed to analyze multiple modifications on one protein, but this approach has not been applied to the analysis of Ub chain branching. Advancements in protein ionization and high-resolution detection methods have expanded the utility of top-down MS[61,62,63].

The typical sample preparation and analysis involves: obtaining a protein mixture of interest, ionizing the protein, gathering spectra for the full mass of the intact protein, isolation of a desired charge state, fragmentation by electron capture dissociation (ECD) or collision induced dissociation (CID), analysis of fragments, extracting sequence data from observed fragments by comparing to predicted fragments.

This technique has advantages for analyzing proteins with multiple PTMs. Mixtures of singly and doubly modified proteins are isolated within the instrument, and then fragmented and analyzed separately. Thus, distinctions between modified residues in different isoforms are observable. Also, ECD fragmentation cleaves a variety of sites, without bias of bond strength[64]. The utility of ECD fragmentation has been demonstrated extensively for phosphorylated proteins[65]. Weaker bonds tethering phospho groups remain intact, while the peptide backbone is cleaved in different places, giving rise to more complete coverage of all possible fragmentation sites. High coverage of all possible fragments leads to greater accuracy when identifying residues that contain modifications.

Bottom-up and top-down MS can be combined to create a new method for Ub chain analysis - termed middle-down MS[66]. Due to the inherent stability of Ub,

minimal trypsin digest conditions can be modified to cleave only after R74, which resides in the flexible linker near the isopeptide bond. The GG isopeptide modification remains on any lysine residue that served as a chain linkage site. The partial digestion conditions only cleave Ub chains to monomers, leaving the Ub core intact. Thus monomers with two GG modifications are distinguishable from those with one GG modification, providing an opportunity to observe and characterize Ub chain branching.

1.5 Conclusion

Ubiquitin plays an important role in regulating many cellular functions. Access to synthetic Ub substrates and development of analytical techniques have advanced mechanistic understanding, but there is still more to uncover. In particular, very little is known about atypical Ub chains with alternating chain linkages or chain branching. With new techniques, we can discover more about the biological functions of branched Ub chains. I have developed two techniques to aid in this endeavor: thiol-ene coupling chemistry provides access to branched Ub substrates and middle-down MS to unambiguously detects branched chains.

1.6 Forward to thesis

In this thesis, I will highlight the work I have done to: synthesize Ub substrates for biochemical investigations; and advance middle-down MS techniques to characterize Ub chains. Chapter 2 describes the development of thiol-ene chemistry to form specific Ub dimers linked at any lysine residue. Chapter 3 expands

on this technique to form longer chains and analyzes their stability in the presence of deubiquitinating enzymes. Chapter 4 addresses the advancement of middle-down MS as an analytical technique to characterize chain branching. Finally, chapter 5 shows preliminary data for expansion of the middle-down MS technique to identify chain branching from cells.

1.7 References

[1] Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479.

[2] Grabbe, C., Husnjak, K., and Dikic, I. (2011) The spatial and temporal organization of ubiquitin networks. *Nat Rev Mol Cell Biol* 12, 295–307.

[3] Pickart, C. M., and Fushman, D. (2004) Polyubiquitin chains: polymeric protein signals. *Current Opinion in Chemical Biology* 8, 610–616.

[4] Ikeda, F., and Dikic, I. (2008) Atypical ubiquitin chains: new molecular signals. "Protein Modifications: Beyond the Usual Suspects" review series. *EMBO* Rep 9, 536–542.

[5] Komander, D., and Rape, M. (2012) The Ubiquitin Code 1–27.

[6] Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* 21, 921–926.

[7] Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W., and Gygi, S. P. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* 44, 325–340.

[8] Pickart, C. M. (2001) Mechanisms underlying ubiquitination. *Annu. Rev. Biochem*. 70, 503–533.

[9] Deshaies, R. J., and Joazeiro, C. A. (2009) RING Domain E3 Ubiquitin Ligases. *Annu. Rev. Biochem*. 78, 399–434.

[10] Schulman, B. A., and Wade Harper, J. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* 10,

319–331.

[11] van Wijk, S. J. L., and Timmers, H. T. M. (2010) The family of ubiquitin conjugating enzymes (E2s): deciding between life and death of proteins. *The FASEB Journal* 24, 981–993.

[12] Popovic, D., Vucic, D., and Dikic, I. (2014) Ubiquitination in disease pathogenesis and treatment. *Nat Med* 20, 1242–1253.

[13] Nakasone, M. A., Livnat-Levanon, N., Glickman, M. H., Cohen, R. E., and Fushman, D. (2013) Mixed-Linkage Ubiquitin Chains Send Mixed Messages. *Structure/Folding and Design* 1–14.

[14] Ben-Saadon, R., Zaaroor, D., Ziv, T., and Ciechanover, A. (2006) The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro 36 histone H2A ligase activity. *Mol Cell* 24, 701–711.

[15] Boname, J. M., Thomas, M., Stagg, H. R., Xu, P., Peng, J., and Lehner, P. J. (2010) Efficient Internalization of MHC I Requires Lysine-11 and Lysine-63 Mixed Linkage Polyubiquitin Chains. *Traffic* 11, 210–220.

[16] Meyer, H.-J., and Rape, M. (2014) Enhanced Protein Degradation by Branched Ubiquitin Chains. *CELL* 157, 910–921.

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

[18] Finley D. 2009. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem*. 78:477–513.

[19] Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, et al. 1989. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 243:1576–83.

[20] Deng L, Wang C, Spencer E, Yang L, Braun A, et al. 2000. Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* 103:351–61.

[21] Al-Hakim A, Escribano-Diaz C, Landry MC, O'Donnell L, Panier S, et al. 2010. The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair* 9:1229–40.

[22] Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, et al. 2007. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 316:1198–202.

[23] Jin L, Williamson A, Banerjee S, Philipp I, Rape M. 2008. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* 133:653–65.

[24] Matsumoto ML, Wickliffe KE, Dong KC, Yu C, Bosanac I, et al. 2010. K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol. Cell* 39:477–84

[25] Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M., and Hicke, L. (2001) Distinct Functional Surface Regions on Ubiquitin. *Journal of Biological Chemistry* 276, 30483–30489.

[26] Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains — from structures to functions 1–13.

[27] Hospenthal, M. K., Freund, S. M. V., and Komander, D. (2013) Assembly, analysis and architecture of atypical ubiquitin chains. *Nat Struct Mol Biol*

[28] Cook, W. J., Jeffrey, L. C., Carson, M., Chen, Z., and Pickart, C. M. (1992) Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J Biol Chem* 267, 16467–16471.

[29] Eddins, M. J., Varadan, R., Fushman, D., Pickart, C. M., and Wolberger, C. (2007) Crystal Structure and Solution NMR Studies of Lys48-linked Tetraubiquitin at Neutral pH. *Journal of Molecular Biology* 367, 204–211.

[30] Castañeda, C. A., Kashyap, T. R., Nakasone, M. A., Krueger, S., and Fushman, D. (2013) Unique Structural, Dynamical, and Functional Properties of K11-Linked Polyubiquitin Chains. *Structure/Folding and Design* 21, 1168–1181.

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

[32] Pickart, C. M., and Raasi, S. (2005) Controlled Synthesis of Polyubiquitin Chains. *Meth Enzymol* 399, 21–36.

[33] Chen, Z., and Pickart, C. M. (1990) A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. *J Biol Chem* 265, 21835–21842.

[34] Petroski, M. D., and Deshaies, R. J. (2005) Mechanism of Lysine 48-Linked Ubiquitin-Chain Synthesis by the Cullin-RING Ubiquitin-Ligase Complex SCF-Cdc34.

CELL 123, 1107–1120.

[35] Hofmann, R. M., and Pickart, C. M. (1999) Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *CELL* 96, 645–653.

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

[37] Wu, T., Merbl, Y., Huo, Y., Gallop, J. L., Tzur, A., and Kirschner, M. W. (2010) UBE2S drives elongation of K11-linked ubiquitin chains by the Anaphase-Promoting Complex. *Proceedings of the National Academy of Sciences* 107, 1355–1360.

[38] Bremm, A., Freund, S. M. V., and Komander, D. (2010) Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat Struct Mol Biol* 17, 939–947.

[39] Dong, K. C., Helgason, E., Yu, C., Phu, L., Arnott, D. P., Bosanac, I., Compaan, D. M., Huang, O. W., Fedorova, A. V., Kirkpatrick, D. S., Hymowitz, S. G., and Dueber, E. C. (2011) Preparation of Distinct Ubiquitin Chain Reagents of High Purity and Yield. *Structure/Folding and Design* 19, 1053–1063.

[40] Thrower, J. S., Hoffman, L., Rechsteiner, M., and Pickart, C. M. (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J* 19, 94–102.

[41] Jacobson, A. D., Zhang, N. Y., Xu, P., Han, K. J., Noone, S., Peng, J., and Liu, C. W. (2009) The Lysine 48 and Lysine 63 Ubiquitin Conjugates Are Processed Differently by the 26 S Proteasome. *Journal of Biological Chemistry* 284, 35485–35494.

[42] Castañeda, C., Liu, J., Chaturvedi, A., Nowicka, U., Cropp, T. A., and Fushman, D. (2011) Nonenzymatic Assembly of Natural Polyubiquitin Chains of Any Linkage 48 Composition and Isotopic Labeling Scheme*. J Am Chem Soc* 133, 17855–17868.

[43] Virdee, S., Ye, Y., Nguyen, D. P., Komander, D., and Chin, J. W. (2010) Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* 6, 750–757.

[44] Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J Am Chem Soc* 85, 2149–2154.

[45] Virdee, S., Kapadnis, P. B., Elliott, T., Lang, K., Madrzak, J., Nguyen, D. P., Riechmann, L., and Chin, J. W. (2011) Traceless and Site-Specific Ubiquitination of Recombinant Proteins. *J Am Chem Soc* 133, 10708–10711.

[46] Kumar, K. S. A., Spasser, L., Erlich, L. A., Bavikar, S. N., and Brik, A. (2010) Total Chemical Synthesis of Di-ubiquitin Chains. *Angew. Chem. Int. Ed*. 49, 9126–9131.

[47] Kumar, K. S. A., Bavikar, S. N., Spasser, L., Moyal, T., Ohayon, S., and Brik, A. (2011) Total chemical synthesis of a 304 amino acid K48-linked tetraubiquitin protein. *Angew Chem Int Ed Engl* 50, 6137–6141.

[48] Licchesi, J. D. F., Mieszczanek, J., Mevissen, T. E. T., Rutherford, T. J., Akutsu, M., Virdee, S., Oualid, El, F., Chin, J. W., Ovaa, H., Bienz, M., and Komander, D. (2011) An ankyrin-repeat ubiquitin-binding domain determines TRABID's specificity for atypical ubiquitin chains. *Nat Struct Mol Biol* 19, 62–71.

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

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

[51] Matsumoto, M. L., Dong, K. C., Yu, C., Phu, L., Gao, X., Hannoush, R. N., Hymowitz, S. G., Kirkpatrick, D. S., Dixit, V. M., and Kelley, R. F. (2012) Engineering and structural characterization of a linear polyubiquitin-specific antibody, *J. Mol. Biol*. 418, 134-144.

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

[53] Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komuves, L., French, D. M., Ferrando, R. E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S. G., Kelley, R. F., and Dixit, V. M. (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies, *Cell* 134, 668-678.

[54] Finley, D., Sadis, S., Monia, B. P., Boucher, P., Ecker, D. J., Crooke, S. T., and Chau, V. (1994) Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant, *Mol. Cell. Biol*. 14, 5501-5509.

[55] Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination, *Mol. Cell. Biol*. 15, 1265-1273.

[56] Aebersold, R.; Mann, M. (2003). Mass Spectrometry-Based Proteomics. *Nature*, *422*, 198–207.

[57] Henzel, W. J.; Watanabe, C.; Stults, J. T. (2003). Protein Identification: The Origins of Peptide Mass Fingerprinting. *J. Am. Soc. Mass Spectrom. 14*, 931–942.

[58] Kaiser, S. E., Riley, B. E., Shaler, T. A., Trevino, R. S., Becker, C. H., Schulman, H., & Kopito, R. R. (2011). Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nature Methods*, *8*, 691–696

[59] Bennett, E. J., Shaler, T. A., Woodman, B., Ryu, K.-Y., Zaitseva, T. S., Becker, C. H., et al. (2007). Global changes to the ubiquitin system in Huntington's disease. *Nature*, *448*, 704–708.

[60] Xu, P., and Peng, J. (2006) Dissecting the ubiquitin pathway by mass spectrometry, Biochim. *Biophys. Acta* 1764, 1940-1947.

[61] Catherman, A. D., Skinner, O. S., & Kelleher, N. L. (2014). Top Down proteomics: facts and perspectives. *Biochemical and Biophysical Research Communications*, *445*, 683–693.

[62] Compton, P. D., Zamdborg, L., Thomas, P. M., & Kelleher, N. L. (2011). On the scalability and requirements of whole protein mass spectrometry. *Analytical Chemistry*, *83*, 6868–6874.

[63] Breuker, K., Jin, M., Han, X., Jiang, H., & McLafferty, F. W. (2008). Top-down identification and characterization of biomolecules by mass spectrometry. *Journal of the American Society for Mass Spectrometry*, *19*, 1045–1053.

[64] Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. (1998). Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Pro- cess. *J. Am. Chem. Soc. 120*, 3265–3266.

[65] Ayaz-Guner, S., Zhang, J., Li, L., Walker, J. W., & Ge, Y. (2009). In vivo phosphorylation site mapping in mouse cardiac troponin I by high resolution top-down electron capture dissociation mass spectrometry: Ser22/23 are the only sites basally phosphorylated. *Biochemistry*, *48*, 8161–8170.

[66] Xu, P., and Peng, J. (2008) Characterization of polyubiquitin chain structure by middle-down mass spectrometry, *Anal. Chem*. 80, 3438-3444.

Chapter 2: Forging Isopeptide Bonds Using Thiol-Ene Chemistry: Site-Specific Coupling of Ubiquitin Molecules for Studying the Activity of Isopeptidases

Portions of this work have been published in:

Valkevich, E. M., Guenette, R. G., Sanchez, N. A., Chen, Y.-C., Ge, Y., and Strieter, E. R. (2012) Forging Isopeptide Bonds Using Thiol-Ene Chemistry: Site-Specific Coupling of Ubiquitin Molecules for Studying the Activity of Isopeptidases. *J Am Chem Soc* 134, 6916-6919.

Contributions:

Mass Spectrometry help was provided by Y.-C. Chen

Isopeptidase assays were performed by R. G. Guenette

2.1 ABSTRACT

Chemical methods for modifying proteins can enable studies aimed at uncovering biochemical function. Herein, we describe the use of thiol-ene coupling (TEC) chemistry to report on the function of branched (also referred to as forked) trimeric ubiquitin (tri-Ub) polymers. First, we show how site-specific isopeptide (*N*ε-Gly-L-homothiaLys) bonds are forged between two molecules of Ub, demonstrating the power of TEC in protein conjugation. Second, we demonstrate that the *N*ε-Gly-L-homothiaLys isopeptide bond is processed to a similar extent by deubiquitinases (DUBs) as that of a native *N*ε-Gly-L-Lys isopeptide bond, thereby establishing the utility of TEC in the generation of Ub-Ub linkages. Lastly, TEC is applied to the synthesis of branched tri-Ub derivatives, where the position of branching Ub units is varied. Interrogation of these branched polymers with DUBs reveals that the relative orientation of the two Ub units has a dramatic impact on how they are hydrolyzed. In particular, cleavage of K48C-linkages is suppressed when the central Ub unit is also conjugated through K6C, whereas cleavage proceeds normally when the central unit is conjugated through either K11C or K63C. The results of this work presage a role for branched polymeric Ub chains in regulating linkage-selective interactions.

2.2 Introduction

Addition of thiyl radicals to alkenes, termed thiol-ene coupling (TEC), has the potential to serve as a powerful method for chemically modifying proteins.[1] The reason is manifold: 1) bimolecular rate constants on the order of 10^6 $M^{-1}s^{-1}$ have been measured for the addition of thiyl radicals to alkenes,[2] which is ideal for carrying out reactions with proteins at mM concentration;[3] 2) the ability to employ standard recombinant proteins;[4] and 3) the potential to forge stable thioether linkages that closely mimic amino acid side chains. Despite these advantages, TEC has seen limited use in the direct modification of cysteines residues in proteins, presumably due to the number of side reactions a thiyl (CysS•) radical can undergo.[5] Nevertheless, there are examples of TEC with peptides and proteins, suggesting in the presence of an alkene side reactions based on CysS• are not competitive.[4,6] Inspired by these studies, we reasoned TEC could be exploited in the construction of isopeptide bonds, an abundant linkage established during the posttranslational modification of proteins with information-rich acyl groups such as ubiquitin (Ub) and Ub-like proteins.[7]

A number of reports have recently emerged describing chemical approaches to the site-specific conjugation of Ub molecules through native *N*ε-Gly-L-Lys isopeptide linkages as well as various nonnative linkages.[8] Indeed, some of the methods have elucidated key structural distinctions between Ub dimers linked through different lysines[9] (there are seven available: K6, K27, K29, K33, K48, and K63), and enabled studies that uncovered how the structure and function of target proteins is altered upon Ub modification.[10] However, many of the chemical methods designed to recapitulate the *N*ε-Gly-L-Lys linkage suffer from drawbacks such as (i) instability in the case of disulfide bonds, (ii) the number of

synthetic manipulations required, and (iii) the use of specialized recombinant DNA technologies for incorporating unnatural amino acids. Development of additional methods is therefore necessary to gain rapid access to a diverse range of Ub modified targets.

We hypothesized TEC would provide an alternative to known chemical approaches for forming isopeptide linkages, as standard recombinant proteins can be employed with minimal synthetic effort. For instance, the conjugation of Ub to a target protein requires a protein harboring a cysteine residue in lieu of a lysine –– a mutation introduced using site-directed mutagenesis –– and Ub bearing small alkene such as allylamine (AA) appended to the C-terminus. Upon TEC, an *N*ε-Gly-L-homothiaLys isopeptide bond would be furnished, which is only one atom longer (~2.4 Å) than the native linkage: an alteration not expected to perturb function.[11] Herein, we describe the successful application of TEC in the concatenation of Ub molecules that exhibit similar behavior as those constructed enzymatically. Moreover, we show how TEC can be used to reveal biochemical details associated with a unique set of Ub polymers referred to as branched chains, i.e., where a single Ub unit is attached through multiple lysines to other Ub molecules.

2.3 Results and Discussion

Our studies commenced with the installation of AA at the C-terminus of Ub. To accomplish this goal, we turned to a class of proteases referred to as Ub C-terminal hydrolases (UCHs). UCHs promote hydrolysis of Ub variants carrying C-terminal extensions through the formation of a $Ub_{1\text{-}76}$-*S*-UCH acyl-enzyme intermediate.[12] Based on this enzymatic logic along with the reversible nature of proteases, AA was added in large excess relative to a Ub variant harboring a C-terminal aspartate cap (UbD77) in the presence of the yeast

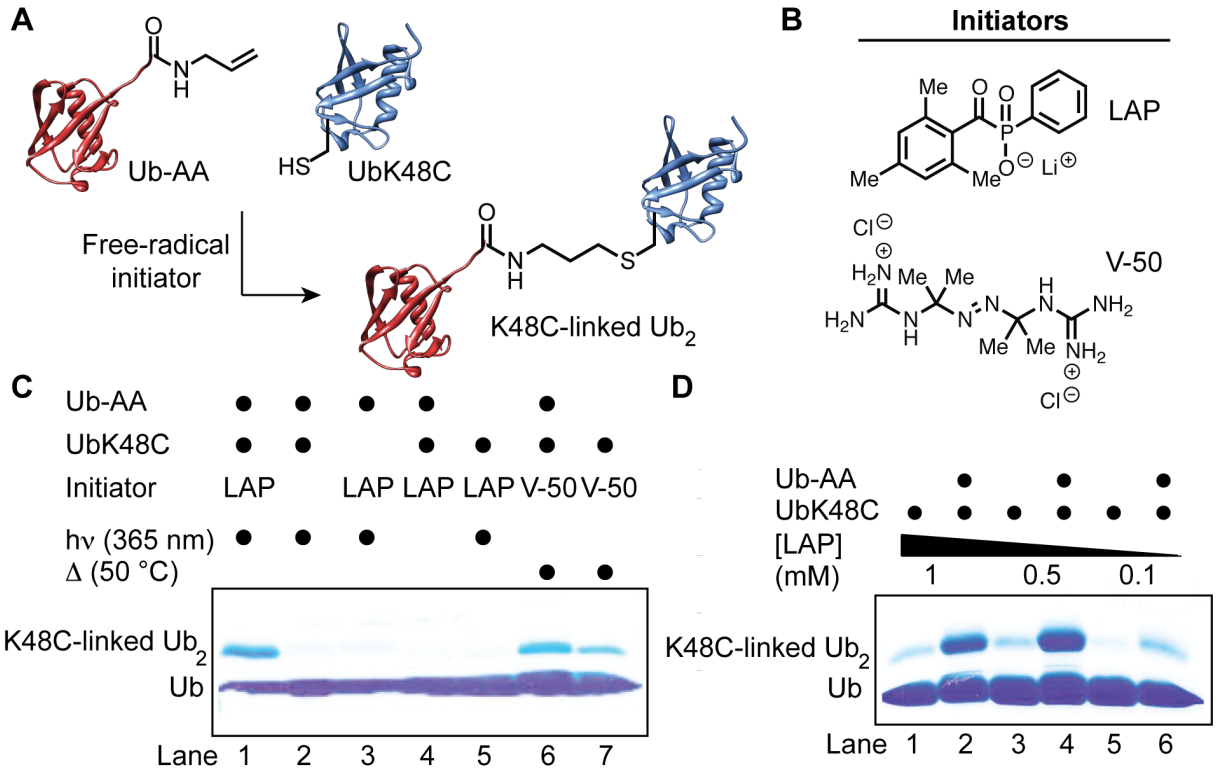


Figure 2.1. Construction of K48C-linked Ub_2 using TEC. (A) Reaction scheme for the TEC of Ub-AA and UbK48C (PDB code for Ub structure shown: 1UBQ). (B) Structures of the free-radical initiators used in this study. (C) Coomassie-stained SDS-PAGE analysis of TEC reactions carried out with different initiators. Each lane represents a reaction conducted with Ub-AA (1 mM) and UbK48C (1 mM), and free-radical initiator LAP (0.1 mM) or V-50 (100 mM) at pH 5.0. In the case of LAP, the reactions were irradiated with long wavelength light at 365 nm. Black dot indicates presence of specified reaction component. (D) SDS-PAGE analysis of TEC reactions with varying concentrations of the LAP photoinitiator.

UCH YUH1. Under these conditions, we found the AA adduct of Ub_{1-76} (herein termed Ub-AA) could readily be obtained in ~ 30% yield (see Materials and Methods). This procedure enables production of Ub-AA on milligram scale.

With a method to generate Ub-AA, we examined conditions to carry out TEC with Ub K48C obtained through site-directed mutagenesis (Figure 2.1A). After screening a series of water-soluble free-radical initiators, the two that proved most effective in terms of the amount of Ub dimer formed were the lithium acyl phosphinate (LAP)[13] photoinitiator and

heat-activated AIBN derivative Vazo-50, also called V-50 (Figure 2.1B). Control experiments, however, revealed clear distinctions between these two initiators. Namely, when Ub-AA was omitted from the reaction with V-50 and replaced with wild type Ub, dimer formation was not abolished. By contrast, in similar control experiments with LAP, a band for the dimer was not observed (Figure 2.1C, compare lanes 5 and 7). These results suggest that at the temperature required to activate V-50 (50 °C), the Ub structure becomes more dynamic possibly promoting reaction pathways other than TEC. Hence, our attention shifted to LAP as the photoinitiator in the coupling of UbK48C to Ub-AA, as it was evident that formation of dimer was dependent on the presence of each reaction component (Figure 2.1C, lane 1 vs. lanes 2-5). Subsequent optimization studies revealed a marked improvement in the amount of dimer formed with higher concentrations of LAP (Figure 2.1D). Accordingly, the use of 0.5 mM LAP furnished milligram quantities of the desired product after ion-exchange chromatography.

Extensive characterization of the K48C-linked dimer was conducted using high-resolution tandem mass spectrometry (MS/MS) on a Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR). Crude TEC reactions were monitored by MS analysis of intact proteins. Peaks corresponding to two distinct ionization states of the dimer (z = 17 and 18) are readily detected when TEC reactions are carried out with all components (Figure 2.2A). In addition, a peak 16 daltons (Da) larger than the dimer

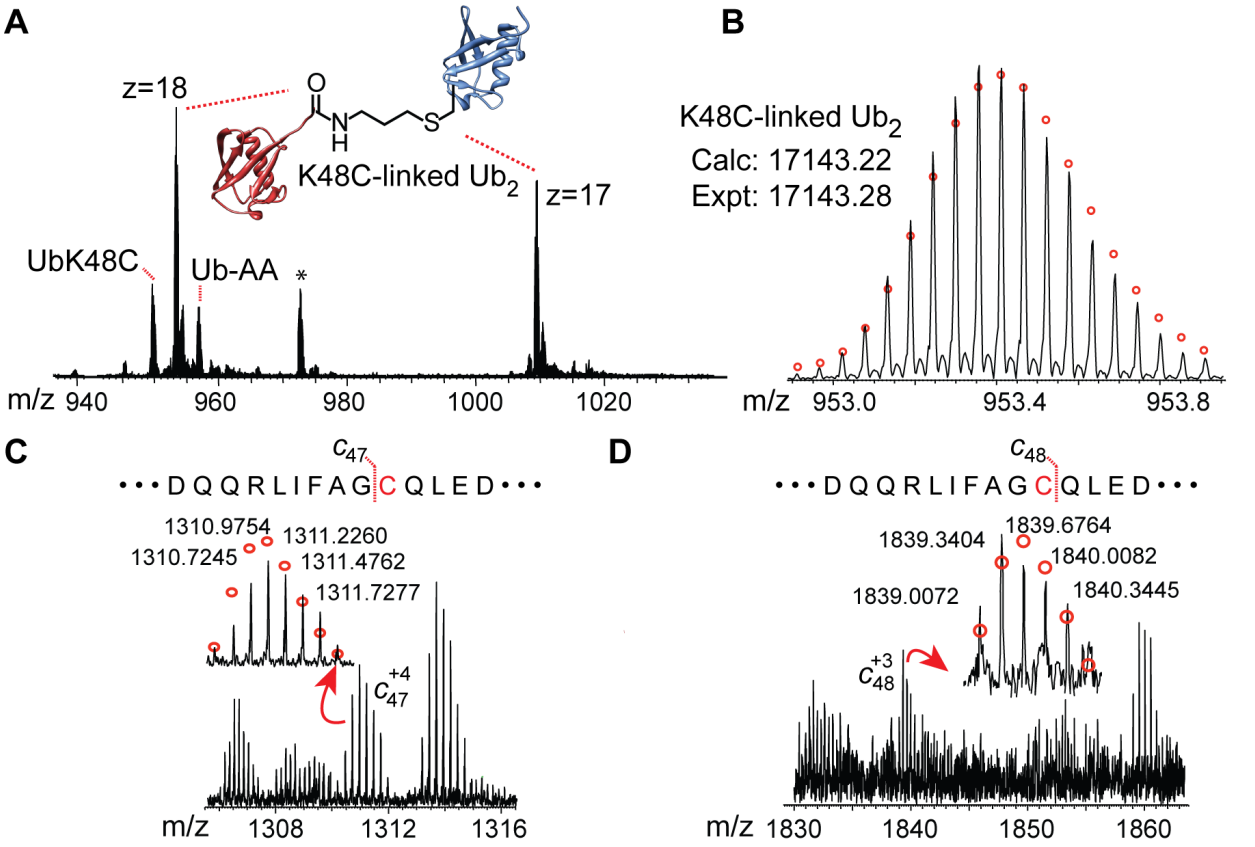


Figure 2.2. Representative mass spectrometric (MS) analysis of K48C-linked Ub dimer. (A) FT-ICR MS analysis of crude reaction mixture shown in Figure 2C. * corresponds to the mass of Ub-AA plus the phosphinate portion of LAP (Scheme S3). (B) An individual charge state (z = 18 or M^{18+}) of purified K48C-linked Ub dimer. The isotopic distribution shown here represents the intact mass of full-length dimer. The red circles correspond to the theoretical distribution of isotopic abundance. (C and D) Representative ECD spectra of a minimal trypsin digest of K48C-linked Ub dimer indicating the installation of a Gly-Gly-AA motif at Ub residue 48. The *c* ions are shown for a fragment containing the first 47 residues (C) and a fragment containing the first 48 residues (D).

was always observed, which might be attributed to the oxidation of the thioether linkage since it is not observed in control experiments. The highly accurate mass measurement of the purified K48C-linked dimer (3.5 ppm between the experimental and theoretical molecular weights) provides strong evidence for the formation of the desired product (Figure 2.2B). To then verify modification of position-48, electron capture dissociation (ECD)[14], a non-ergodic MS/MS technique, was performed with a minimal trypsin digest of

K48C-linked dimer. Minimal digest removes the C-terminal diGly motif leaving a Ub_{1-74} unit with a Gly-Gly-AA (171 amu) appendage (Figure 2.12); this greatly simplifies ECD analysis. ECD spectra report on the extensive fragmentation of N-terminal ***c*** ions and C-terminal ***z***• ions (Figure 2.15), and the fragmentation pattern surrounding position-48 unambiguously verified incorporation of the desired modification at this position (Figures 2.2C and D). Taken together, SDS-PAGE and MS data argue that TEC is a robust method for site-specifically coupling two proteins and generating K48C-linked Ub dimers.

Encouraged by these results, we investigated the application of TEC in the synthesis of topoisomers, e.g., dimers linked through K6C, K11C, K27C, K29C, K33C, and K63C. Using similar conditions to those optimal for the K48C-linked dimer, K6C-, K11C-, and K63C-linked dimers were obtained in yields comparable to K48C (Figure 2.7). However, synthesis of K29C-, K33C-, and in particular the K27C-linked dimer proved more challenging. Specifically, high-resolution FT-ICR MS analysis of crude reaction mixtures containing 1:1 ratios of UbK29C/Ub-AA and UbK33C/Ub-AA showed dimers in low abundance relative to the monomeric substrates; in similar experiments with UbK27C, dimer formation was not observed. Based on these results, we reasoned the addition of CysS• to Ub-AA could be suppressed by the steric bulk surrounding the radical species. To test this hypothesis, different concentrations of

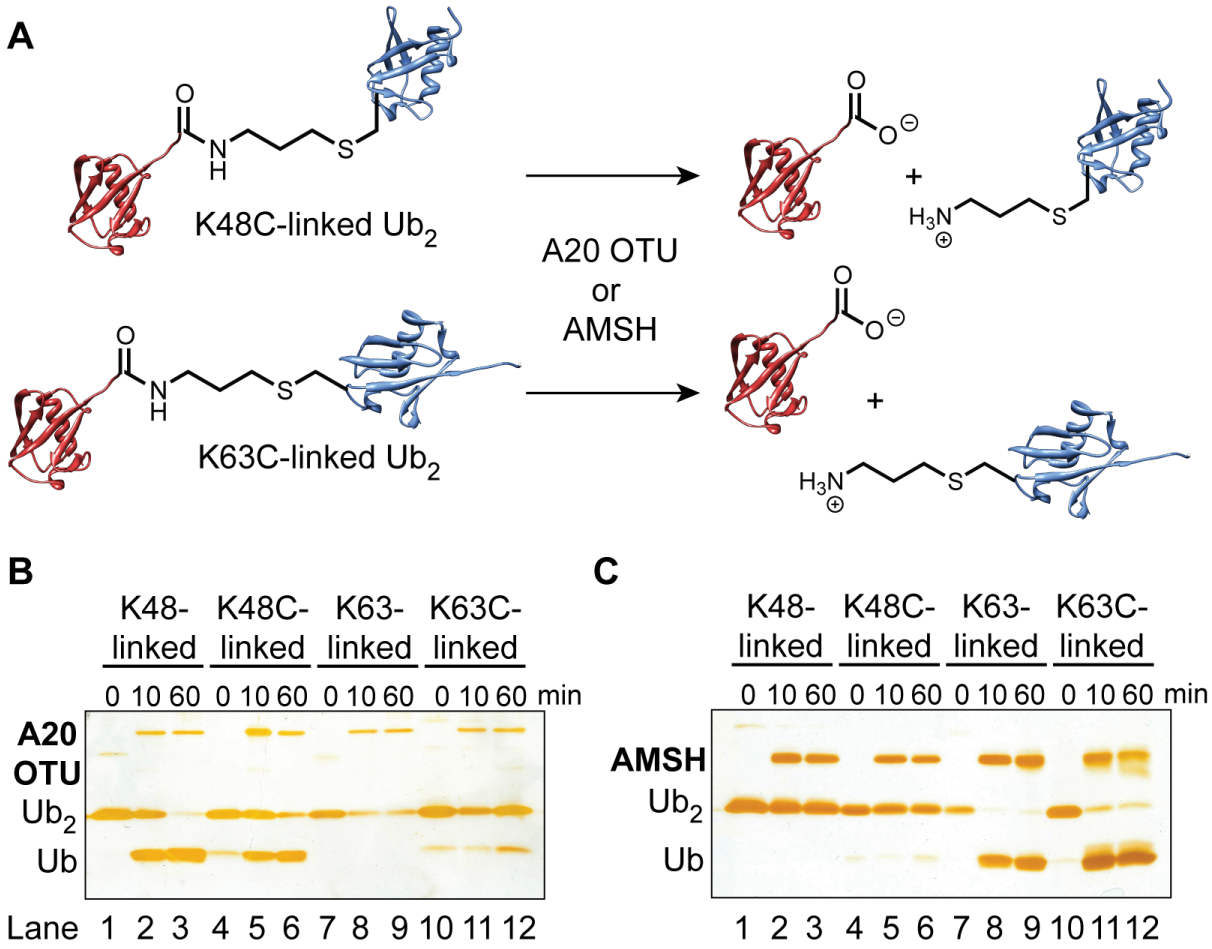


Figure 2.3. DUB-catalyzed hydrolysis of dimers forged through TEC. (A) The general scheme for the DUB-catalyzed hydrolysis of K48C- and K63C-linked Ub dimers. (B) Silver-stained SDS-PAGE analysis of cleavage reactions with the K48-linkage selective A20-OTU domain. At t = 0 min, the DUB is absent from the reaction mixture. (C) Silver-stained SDS-PAGE analysis of cleavage reactions with K63-linkage selective AMSH.

Ub-AA were employed. For UbK29C and UbK33C, MS analysis pointed to a clear trend in the relative amount of dimer to monomeric substrate, i.e., higher concentrations of Ub-AA led to an increase in the abundance of the dimer peak (Figures 2.23 and 2.24). Conversely, UbK27C remained refractory toward coupling even at UbK27C/Ub-AA ratios of 1:4. This result, however, was not surprising considering residue-27 is the least accessible of those tested.

With the reaction scope realized in the synthesis of Ub dimers, we tested the function of these molecules. To accomplish this goal, the hydrolytic cleavage of *N*ε-Gly-L-homothiaLys isopeptide linkages was investigated using isopeptidases (also referred to as deubiquitinases or DUBs).[15] Several members of the DUB family of enzymes preferentially cleave specific linkages within a polyUb chain. For instance, the ovarian tumor (OTU) domain-containing protein referred to as A20 prefers K48-linkages, whereas the DUB AMSH (*a*ssociated *m*olecule with the *SH*3 domain of STAM) cleaves K63-linkages.[16] In many cases, and in particular with A20 and AMSH, linkage specificity arises from (i) the unique sequence context of each Ub lysine residue and (ii) direct contact with all atoms of the lysine side chain.[15] Due to the discriminating features of linkage-selective DUBs, we surmised A20-OTU- and AMSH-catalyzed hydrolysis of K48C- and K63C-linked dimers would provide a stringent test for the ability of *Nε-Gly*-L-homothiaLys to mimic the native linkage (Figure 2.3A). For direct comparison, native K48- and K63-linked dimers were hydrolyzed alongside the dimers synthesized by TEC. Analysis of A20-OTU-catalyzed cleavage of *Nε-Gly*-L-homothiaLys indicated that while the K63C-linked dimer is not hydrolyzed (Figure 2.3B, lanes 10-12), the K48C-linked native dimer is almost completely converted to the respective monomeric units within 60 minutes (Figure 2.3B, lanes 4-6). These results are congruent with those obtained for dimers linked through the native isopeptide bond (Figure 2.3B, lanes 1-3 and 7-9). In the case of K63-linkage specific DUB AMSH, the K63C-linked dimers are rapidly hydrolyzed, whereas the K48C-linked dimers are not processed (Figure 2.3C, lanes 4-6 vs. 10-12). It is important to note that although there is a small amount of each thioether dimer remaining after an hour, the half-life of *Nε-Gly*-L-homothiaLys is nearly identical to that of *Nε-Gly*-L-Lys. These results indicate TEC

affords fully functional Ub dimers that can readily be obtained in two straightforward steps from standard recombinant proteins.

Given the ease with which functional Ub dimers can be forged, we surmised TEC could be applied towards the synthesis of branched polyUb chains. This class of polymers has been observed as products of specific pairs of E2 Ub-conjugating and E3 Ub-ligating enzymes, but their abundance and function *in vivo* remains unclear, in part, due to the inability to identify branched linkages from tryptic digests.[17] What is known about branched chains is that they display a low affinity for 26S proteasomes, and certain Ub-binding chaperones can prevent their formation thereby promoting protein degradation.[18] To gain more insight into the function of this class of polymers, our established TEC protocol was used to synthesize three branched tri-Ub topoisomers starting from the Ub lysine-to-cysteine double mutants: K6C, K48C; K11C, K48C; and K48C, K63C (Figure 2.4A). This particular set was chosen in order to systematically investigate the influence of an additional Ub unit on the hydrolysis of the K48C-linkage. Similar to the dimers, the trimers were purified using ion-exchange chromatography and characterized by ECD analysis of intact proteins minimally digested with trypsin (Figures 2.16-2.21). Western blot analysis with a Ub antibody shows two bands for all trimers along with a faint band corresponding to a dimer (Figure 2.4B). The trimer bands can be ascribed to the presence of reduced and oxidized forms of the *Nε-Gly*-L-homothiaLys linkage as both are observed by MS analysis. A shift in electrophoretic mobility between different forms of Ub polymers is common as the polymer increases in size, which explains why distinct bands for the reduced and oxidized forms of the thioether linkage are not observed with dimers.[19]

Functional studies with the branched trimers revealed clear differences in the DUB-catalyzed hydrolysis of the K48C-linkage (Figure 2.4B). Cleavage of each trimer was examined with three different DUBs: IsoT, A20-OTU, and AMSH. IsoT hydrolyzes free polyUb chains, i.e., those not conjugated to a target protein, with little selectivity over linkage type.[20] Given the presence of a free C-terminus in each trimer we anticipated that IsoT-catalyzed hydrolysis would rapidly furnish dimeric and monomeric products. Indeed, IsoT efficiently processed all three trimers as evidenced by western blot analysis (Figure 2.4B). The most striking result, however, came while studying A20-OTU-catalyzed cleavage. That is, western blot analysis indicated A20-OTU cleaved the K48C-linkage in K11C, K48C- and K48C, K63C-linked trimers, whereas the same linkage remained intact in the K6C, K48C-linked trimer (Figure 2.4B). Since other non-selective DUBs such as those in the USP (Ub-specific protease) family,[21] in particular USP7, trim K6C, K48C-linked tri-Ub down to the monomer (see Materials and Methods), the results with A20-OTU suggest the additional Ub unit appended to position-6 abrogates hydrolytic cleavage by K48-linkage selective DUBs. In the context of other linkage selective DUBs such as AMSH, the presence of a Ub appendage at position 48 does not influence cleavage of the K63C-linkage as indicated by the formation of di-Ub and Ub upon hydrolysis of the K48C, K63C-linked

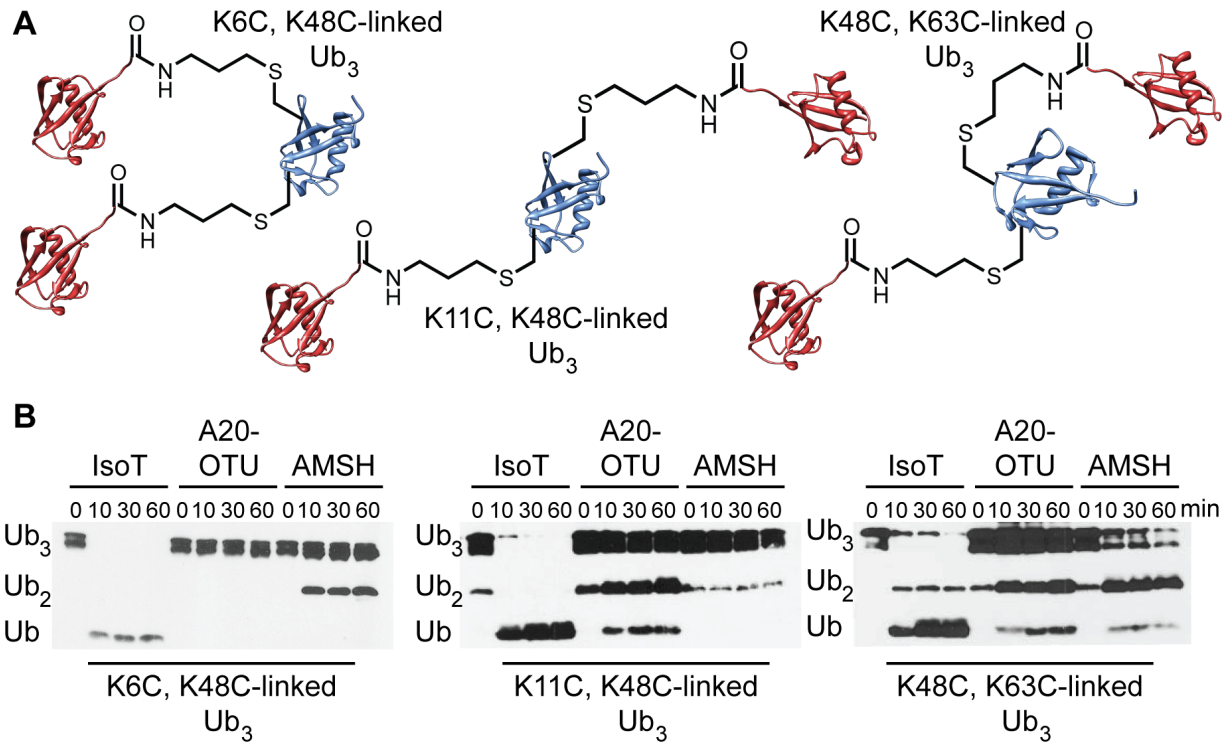


Figure 2.4. Structure and function of branched tri-Ub derivatives. (A) The three branched tri-Ub (Ub_3) derivatives synthesized in this study: K6C, K48C-; K11C, K48C-; and K48C, K63C-linked. The central Ub unit conjugated to two Ub molecules is shown in blue. (B) Western blots developed with a Ub antibody (P4D1) showing the extent to which different DUBs (IsoT, A20-OTU, and AMSH) hydrolyze the three branched tri-Ub derivatives. Hydrolysis is indicated by the formation of dimers (Ub_2) and monomers.

trimer (Figure 2.4B). Additional work is necessary to determine whether other branch points, e.g., K6C and K11C, affect hydrolysis of K63C-linkages.

Our systematic examination of branched trimer topologies suggests that branch points in a polyUb chain furnish a regulatory mechanism for linkage-selective interactions. Consistent with this analysis, K6-linkages are proposed to suppress degradation of target proteins by 26S proteasomes.[22] In principle, this could lead to the accumulation, and possibly aggregation, of the target protein, which in turn would set the stage for clearance by the lysosomal pathway.[23] If the latter is either unable or slow to process the aggregated proteins bearing polyUb chains, then toxic levels may begin to accrue in the cell: this is a

hallmark of many neurodegenerative diseases. Interestingly, mixed K6-, K11-, and K48-linked polyUb chains have been observed in Tau aggregates isolated from brain tissue of individuals with Alzheimer's disease.[24]

In summary, the work described herein showcases the utility of TEC in the construction as well as the biochemical analysis of dimeric and trimeric Ub conjugates. With these new tools, future work will focus on understanding the abundance and function of branched polyUb chains *in vivo*.

2.4 Materials and Methods

2.4.1 Ubiquitin (Ub) cloning and expression

Cloning. Ubiquitin (1-76) (herein referred to as Ub_{1-76}) was purchased from Addgene and cloned into pET-22b using the forward primer ggcggt**CATATG**CAGATCTTCGTCAAG and reverse primer ggcggt**GCGGCCGC**TCAACCACCTCTTAGTCT containing NdeI and NotI restrictions sites. Lysine-to-cysteine mutations (KxC; where x is the position within Ub primary sequence) were introduced using the mutagenesis technique of splice overlap extension (SOE).[25] Primers containing the TGC mutation were used to replace the respective codon for lysine. Aspartate 77 was encoded in the reverse primer to afford the clone for UbD77.

Expression. All Ub variants were expressed and purified from Rosetta™ 2(DE3)pLysS cells (Novagen) following a procedure adapted from Pickart 2005.[26] A starter culture was inoculated (10 mL LB media, 100 μg/mL Ampicillin), grown to OD_{600}=0.5, and kept at 4 °C overnight. The starter culture was added to 1L 2x YT media (16 g Peptone, 5 g NaCl, 10 g

Yeast extract, 100 μg/ml Ampicilin) and grown at 37 °C. Cultures were then induced with 0.4 mM IPTG at OD_{600}= 0.6 and incubated for an additional 4 hours at 37 °C. Cells were pelleted and resuspended in 150 mL lysis buffer (50 mM Tris pH 7.5, 0.5 mM EDTA, 1 mM EgTA, 0.02% Igepal, 1 mM PMSF, 1 mM DTT). After sonication, the lysate was clarified by centrifugation at 8,000 rpm for 30 min. Perchloric acid (70%, 0.19 mL) was added dropwise to the soluble layer and stirred for 20 min to precipitate impurities. After centrifuging at 8,000 rpm for 30 min, the supernatant was exchanged into FPLC Buffer A (50 mM NH_4OAc pH 4.4, 1 mM EDTA, 1 mM DTT) with 2 rounds of dialysis (3.5 kD molecular weight cutoff snake-skin tubing). Ub variants were further purified by cation exchange chromatography with a gradient of 0% to 60% Buffer B (50 mM NH_4Oac pH 4.4, 1 mM EDTA, 1 mM DTT, 1 M NaCl) over 35 column volumes. Fractions containing Ub (monitored by SDS-PAGE) were combined, concentrated, exchanged into H_2O and lyophilized: the purpose of which is to determine yields and minimize variation in the concentration of stock solutions.

Yeast Ub Hydrolase 1 (YUH1) expression and purification

YUH1 in pET-3a was purchased from Addgene and expressed from Rosetta™ 2(DE3)pLysS cells (Novagen). A starter culture was inoculated (10 mL LB media, 100 μg/mL ampicillin), grown at 37 °C for 6 h, and stored at 4 °C overnight. 1 L of LB was inoculated with starter culture and grown to OD_{600}=0.8. A culture was induced with 1mM IPTG and grown 13 h at 18 °C. Cells were harvested by centrifugation (8,000xg at 4 °C, 30 min) and pellet was resuspended in lysis buffer (20 mM $NaPO_4$, 0.5 M NaCl, pH 7.4). Cells were lysed by sonication, and clarified by centrifugation (30,000xg at 4°C, 30 min). YUH1 was purified by

ammonium sulfate precipitation. The 40% and 60% ammonium sulfate fraction was dialyzed into 25 mM NaCl, 50 mM HEPES pH 6.8, 1 mM DTT, and further purified by anion exchange (Buffer A: 25 mM Tris pH 8; Buffer B: 25 mM Tris pH 8, 1 M NaCl; 0-100% B, 30 column volumes).

2.4.2 Synthesis of Ub allyl amine adduct (Ub-AA)

UbD77 (185.6 mg, 21.7 μmols) was dissolved in a buffer containing 50 mM Hepes pH 8, 1 mM EDTA, 30% DMSO, and 250 mM allylamine to a total reaction volume of 25 mL. To this mixture was added YUH1 (25 nM). After two hours of shaking at room temperature, the reaction mixture was quenched with TFA to a pH of 1-2, exchanged into Buffer A (50 mM NH_4Oac pH 4.4, 1 mM DTT, 0.5 mM EDTA) and purified by cation exchange chromatography using the same method as in the Ub purification described above. Fractions containing Ub-AA were verified by MALDI mass spectrometry, combined, concentrated, exchanged into water, and lyophilized for use in thiol-ene reactions. Final characterization was done by high resolution Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry using the methods explained below (**Figure 2.5**).

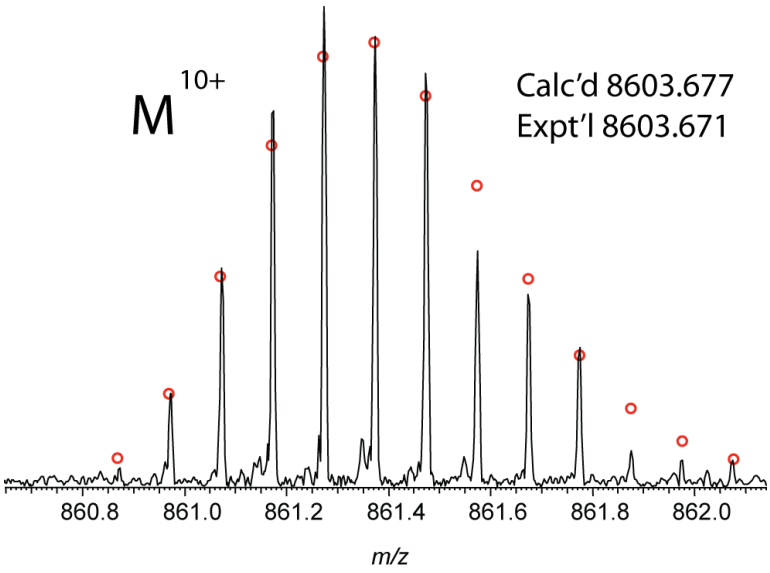


Figure 2.5. High resolution FT-ICR MS analysis of intact full-length Ub-AA (M^{10+} charge state is shown). Red circles represent the theoretical isotopic distribution. Calc'd and Expt'l refer to the calculated and experimental molecular weights of full length Ub-AA, respectively.

2.4.3 Synthesis of the lithium acyl phosphinate (LAP) free-radical photoinitiator

Lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) was synthesized according to the procedures described by Majima[27] and Fairbanks[28]without modifications (**Figure 2.6**).

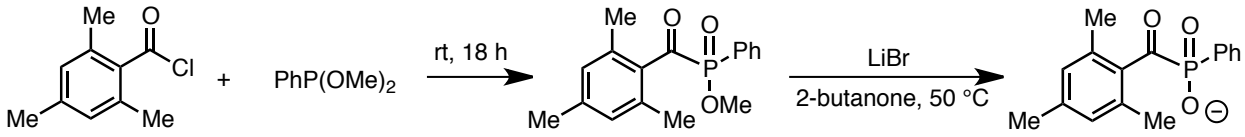


Figure 2.6. Synthetic scheme for LAP free-radical photoinitiator.

2.4.4 Thiol-ene coupling (TEC) reactions

Reaction procedure. Typical TEC reactions contained UbKxC (1 mM), Ub-AA (1 mM), LAP (0.5 mM) in 250mM NaOAc buffer pH 5.1 (50 mL reaction volume). Samples were placed on ice and irradiated from above with 365 nM light for 30 min using an OmniCure series 1500 light source placed 15 cm away from the sample. Control reactions contained wild

type Ub (1 mM) instead of Ub-AA. TEC reactions were performed on all seven single UbKxC mutants (x = 6, 11, 27, 29, 33, 48, 63) and analyzed by Coomassie-stained SDS-PAGE (**Figure 2.7**). The same procedure was performed with the K6C, K48C; K11C, K48C; and K48C, K63C Ub double mutants, but in this case Ub-AA (2 mM) was used to furnish the respective branched trimers.

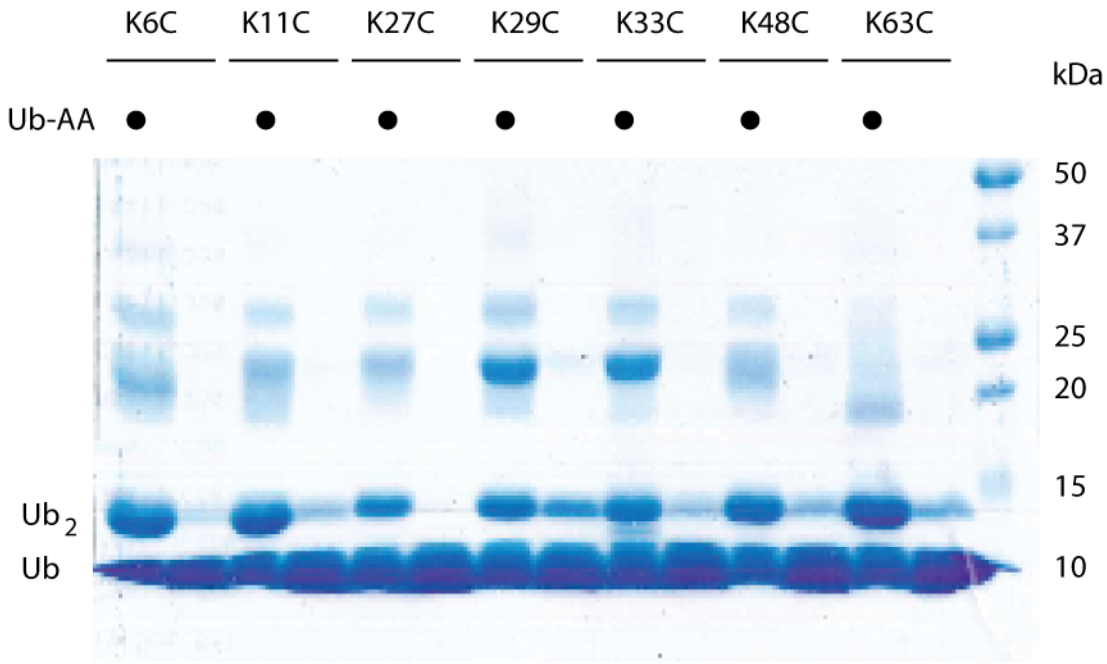


Figure 2.7. Coomassie-stained SDS-PAGE analysis of TEC reactions with all seven UbKxC mutants. Dimers are observed in all reactions containing Ub-AA. The higher MW bands observed in the reactions conducted with Ub-AA are present in differing amounts depending on the UbKxC mutant.

2.4.5 Purification procedure for Ub dimers and trimers. Multiple TEC reactions for each dimer (K6C-, K48C-, and K63C-linked) and branched trimer (K6C, K48C-; K11C, K48C-; and K48C, K63C-linked) were combined (15x50 mL) and purified by cation exchange chromatography with a gradient of 0-60% B over 35 column volumes (**Figure 2.8**).

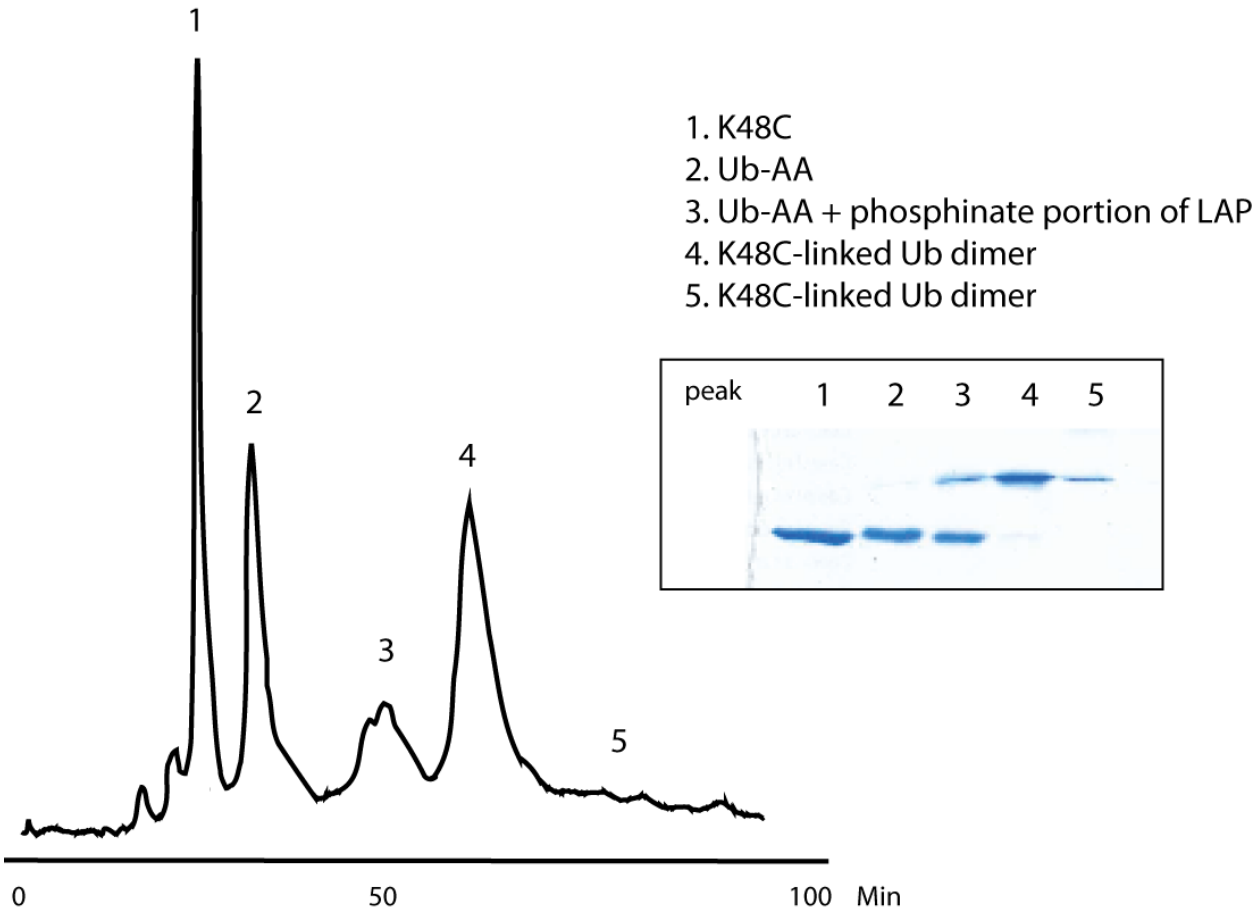


Figure 2.8. Representative purification of TEC products: FPLC chromatogram for the K48C-linked Ub dimer. The inset shows Coomassie-stained SDS-PAGE analysis for each peak observed in the chromatogram: peak 4 contains the purified K48C-linked Ub dimer. MS analysis of peak 3 corresponds to the mass of Ub-AA plus the phosphinate portion of the LAP photoinitiator (for more details, see below).

2.4.6 MS analysis of intact full-length Ub dimers and trimers

General procedure. Crude TEC reactions for all KxC-linked Ub dimers were reduced with dithiothreitol (DTT) then desalted using Amicon 3.5 kD MW cutoff filters. Samples were dissolved in a solution of water/MeCN/acetic acid (45:45:10) and injected into a 7T linear ion trap/Fourier transform ion cyclotron resonance (LTQ/FT-ICR) hybrid mass

spectrometer (Thermo Scientific Inc., Bremen, Germany) equipped with an automated chip-based nanoESI source (Triversa NanoMate, Advion BioSciences, Ithaca, NY) as described previously.[29] The resolving power of the FT-ICR mass analyzer was set at 200,000. All FT-ICR spectra were processed with in-house developed Software (MadTHRASH version 1.0) using a signal to noise threshold of 3 and fit factor of 60%, and then validated manually.

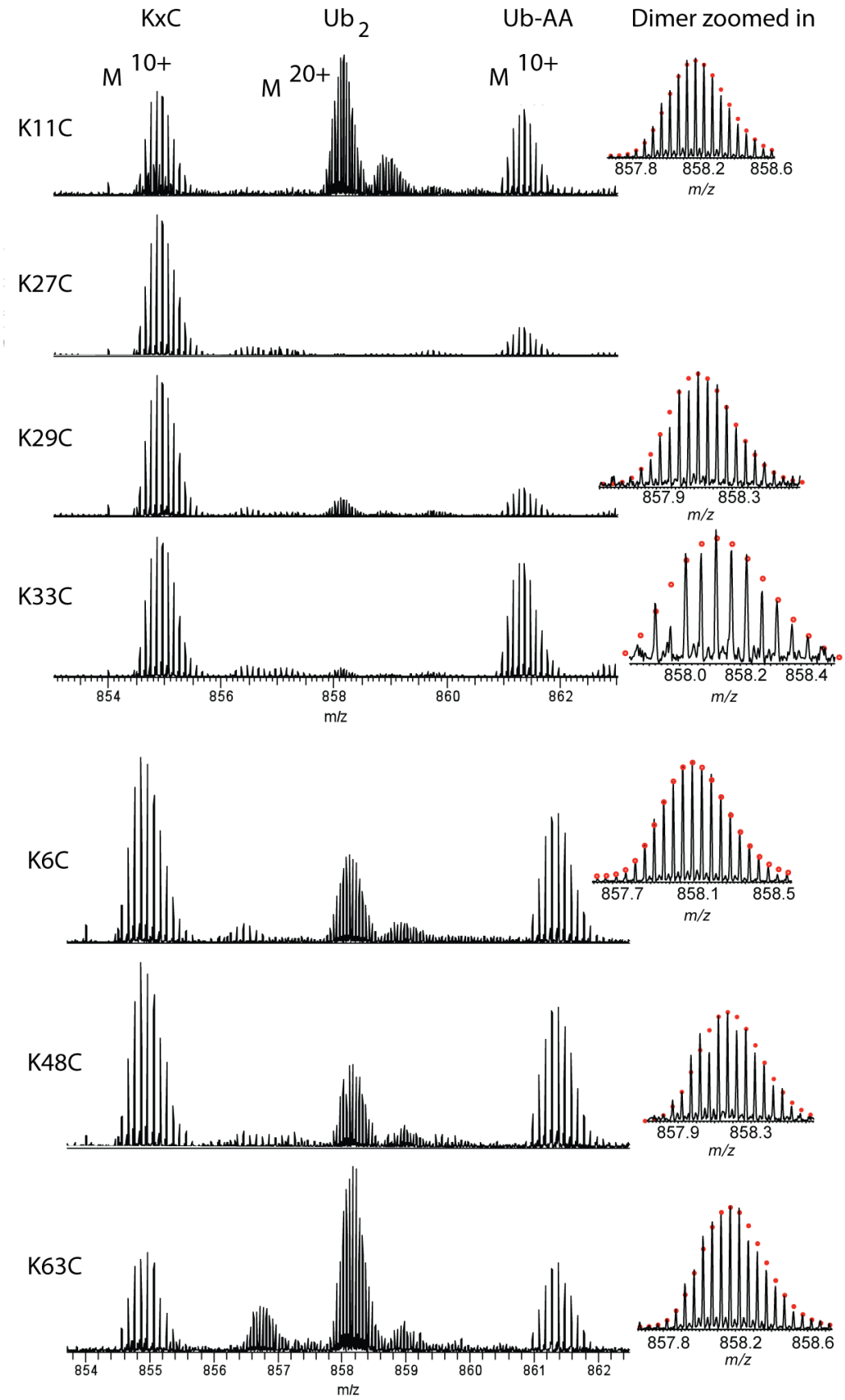

KxC
Ub 2
Ub-AA
Dimer zoomed in
M 10+
M 20+
M 10+
K11C
K27C
K29C
K33C
K6C
K48C
K63C
857.8 858.2 858.6
m/z
857.9 858.3
m/z
858.0 858.2 858.4
m/z
857.7 858.1 858.5
m/z
857.9 858.3
m/z
857.8 858.2 858.6
m/z
854 856 858 860 862
m/z
854 855 856 857 858 859 860 861 862
m/z

Figure 2.9. High resolution FT-ICR MS analysis of crude TEC reactions using intact full-length proteins. Wide view shows abundance of Ub dimers in comparison to the starting materials UbKxC and Ub-AA (M^{10+} charge state for starting materials, M^{20+} charge state for dimer is shown). Zoom in shows each dimer compared to the theoretical isotopic distribution (red dots).

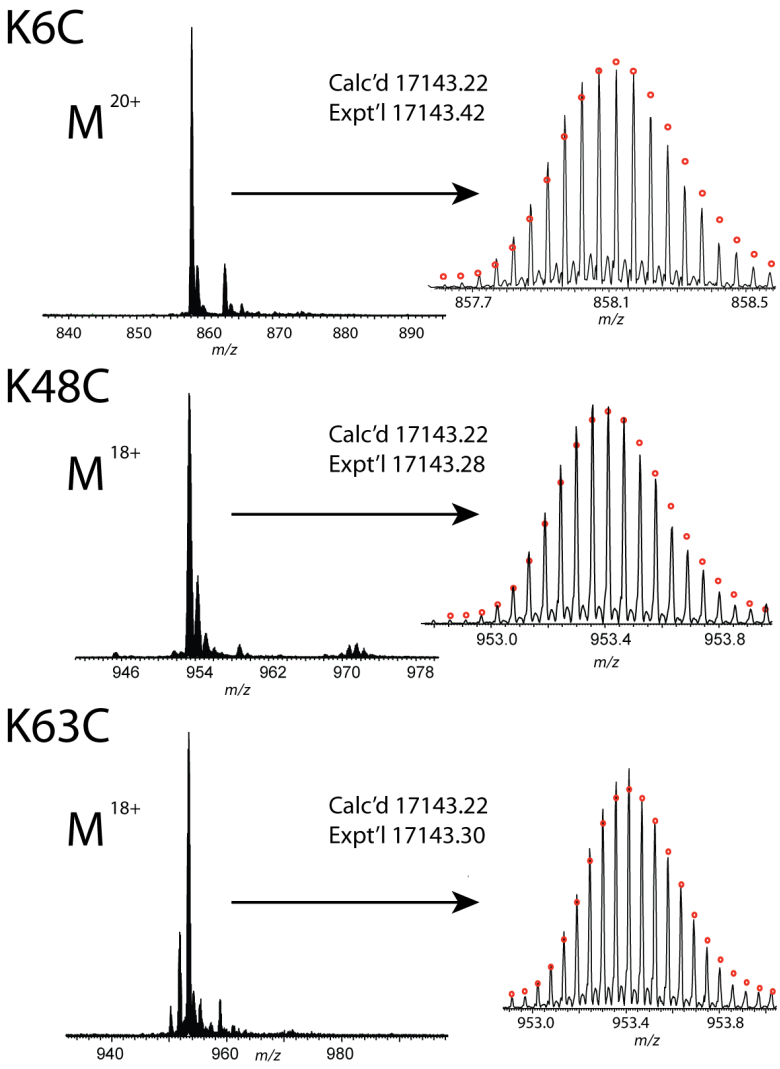


Figure 2.10. High resolution FT-ICR MS analysis of each purified dimer. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

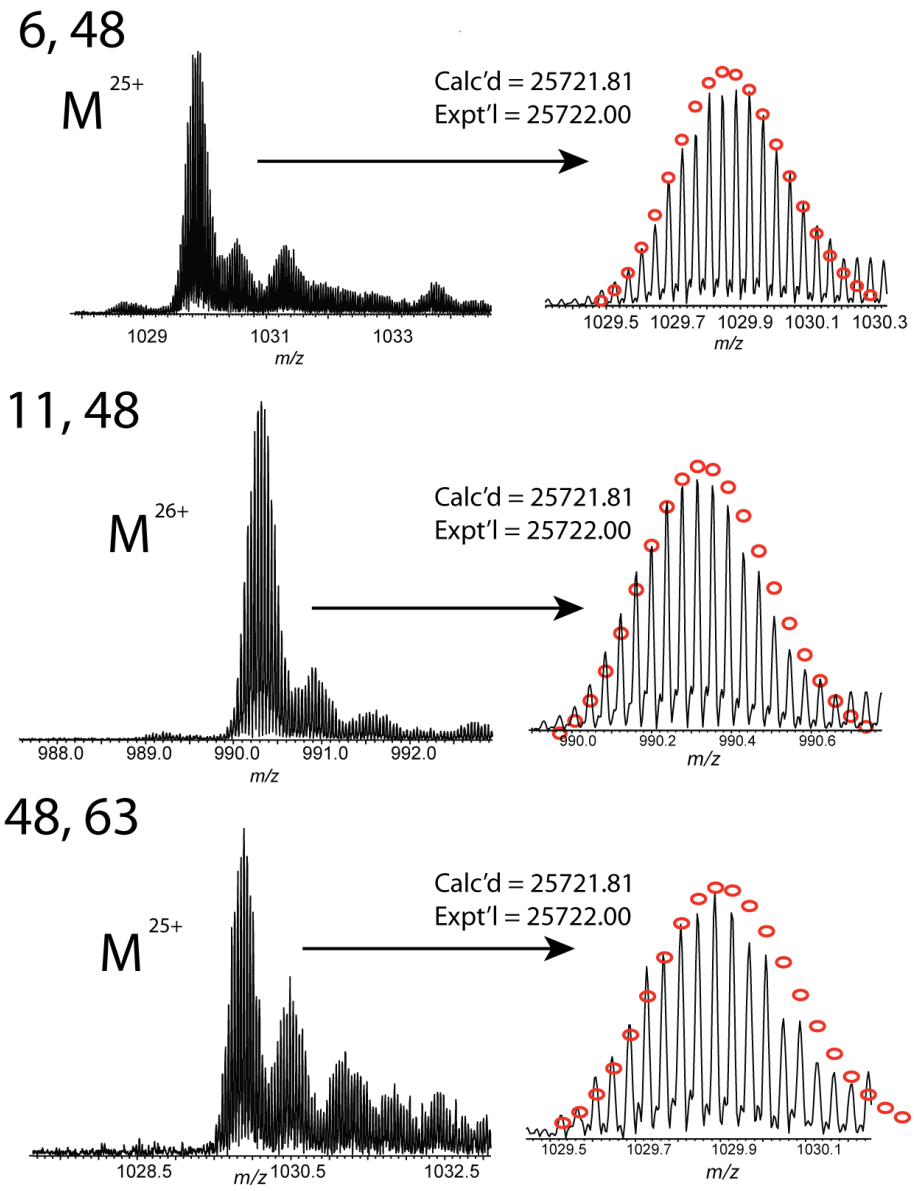


Figure 2.11. FT-ICR analysis of each purified branched trimer. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

2.4.7 Electron capture dissociation (ECD) analysis of the *Nε-Gly*-L-homothiaLys linkage

General procedure for ECD. For tandem mass spectrometry (MS/MS) experiments using ECD, individual charge states of protein molecular ions were first isolated. Then, the ions were dissociated by ECD using 3% "electron energy" and a 70 ms duration with no delay. All FT-ICR spectra were processed with in-house developed Software (MadTHRASH version 1.0) using a signal to noise threshold of 3 and fit factor of 60%, and then validated manually. The resulting mass lists were further assigned based on the protein sequence of Ub with or without the modification (GlyGly-AA) at each Cys using a 10 and 20 ppm tolerance for precursor and fragment ions, respectively. All reported M_r values are most abundant molecular weights.

Sample preparation for ECD analysis. Purified K6C-, K48C-, and K63C-linked Ub dimers and branched Ub trimers (K6C, K48C-; K11C, K48C-; and K48C/K63C-linked) were minimally digested with trypsin (3 h digestion, 1:100 trypsin:Ub ratio) for detailed analysis of the *Nε-Gly*-L-homothiaLys thioether linkage. Since trypsin cleaves at position-74 of Ub, the products in this case are single Ub units missing the C-terminal GlyGly motif, but attached to GlyGly-AA at the respective cysteine residue (the product of this minimal trypsinolysis is referred to as Ub_{1-74}GlyGly-AA): this greatly simplifies analysis by ECD fragmentation (**Figure 2.12**).

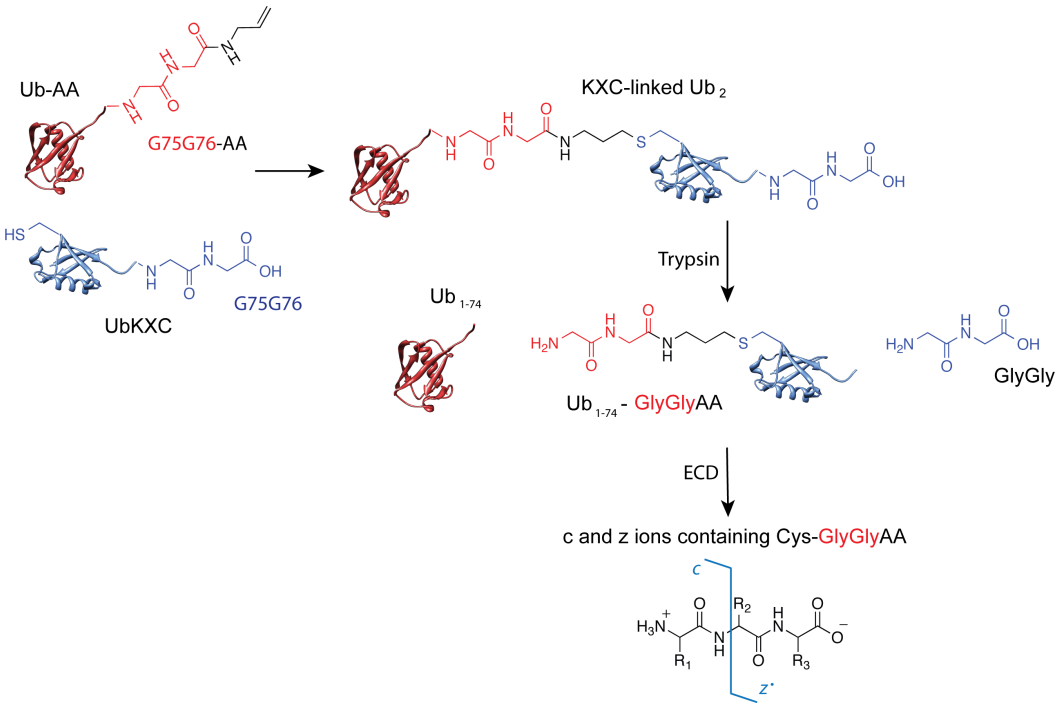


Figure 2.12. ECD analysis of the *Nε-Gly*-L-homothiaLys linkage.

Detailed procedure for ECD analysis. Ub_{1-74}GlyGly-AA was fragmented by ECD, and the resulting fragments were analyzed to verify the linkage at the desired cysteine. Analysis was performed using in-house data analysis software (described above). The Ub_{1-74} sequence was used to predict fragment molecular weights of all possible c ions (N-terminal ions, numbered from amino acid 1) and $\boldsymbol{z}^{\cdot}$ ions (C-terminal ions numbered starting from 1 from the C-terminal arginine and counting in reverse of the conventional amino acid numbering scheme). Raw data were analyzed to find molecular weights of each observed fragments. Observed fragments and predicted fragments are compared to assign ion type to the observed peaks. Ion assignments were then verified and analyzed with and without inclusion of a cysteine modification in the theoretical peak predictions. When the thiol-ene modification was not taken into account in the analysis, ***c*** ions after the cysteine and $\boldsymbol{z}^{\cdot}$ ions

before the cysteine were lacking. Upon addition of 171 amu, which represents the GlyGly-AA motif, to the molecular weight of cysteine the ion types in observed data were reassigned. In this case, the ***c*** and ***z·*** ions were both present throughout the entire sequence. This data supports the modification at the desired cysteine residue and holds true for all mutants analyzed.

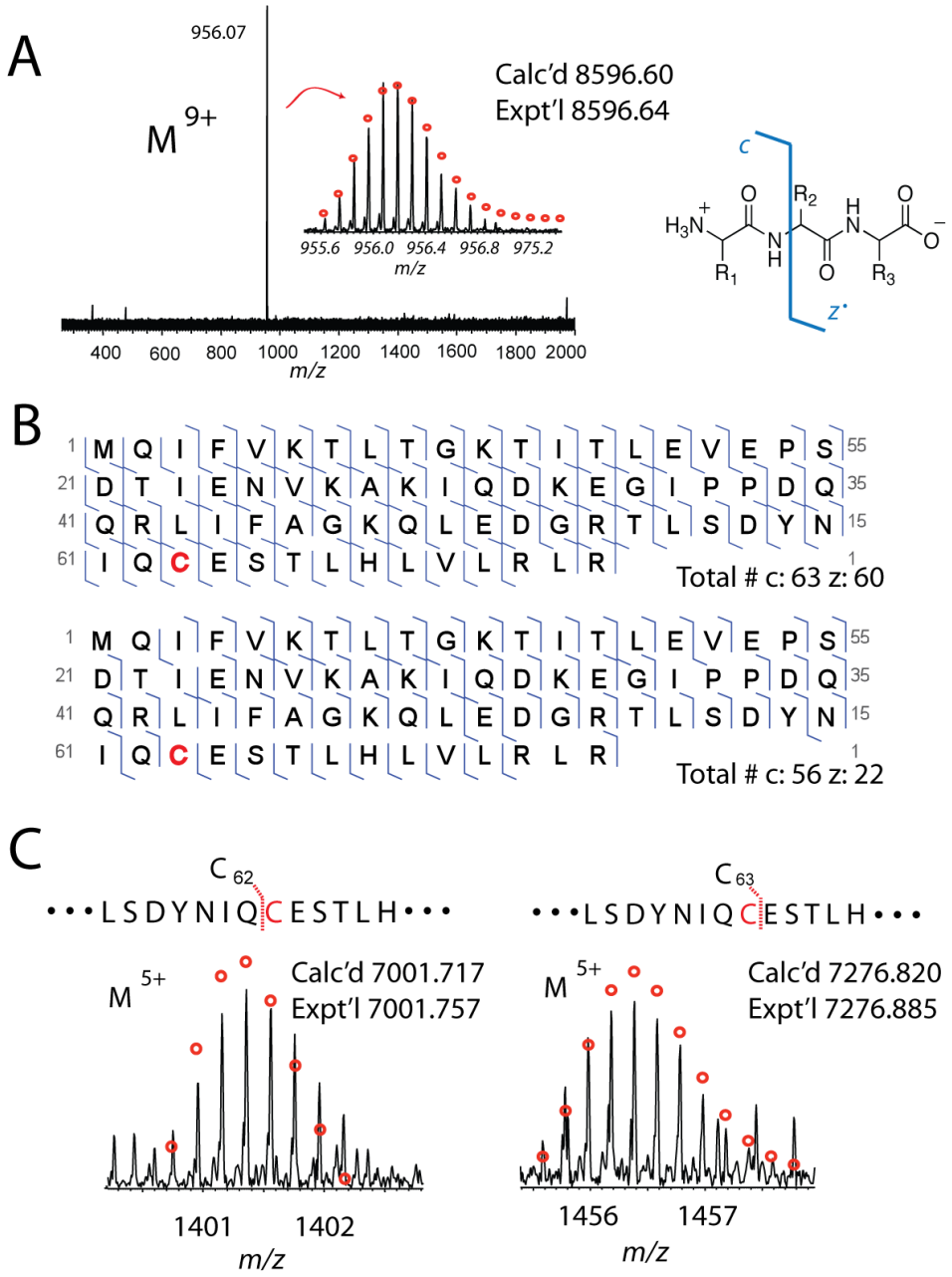


Figure 2.13. ECD analysis of K63C-linked dimer. (A) K63C-linked $Ub_{1\text{-}74}$ GlyGly-AA parent ion isolation (M^{9+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top includes *Nε-Gly*-L-homothiaLys thioether linker modification at cysteine-63 (red) in ***c*** and ***z·*** ion predictions. Bottom map does not include thioether linker modification in theoretical analysis. (C) Key ECD fragment ions for

mapping thioether linkage site on UbK63C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

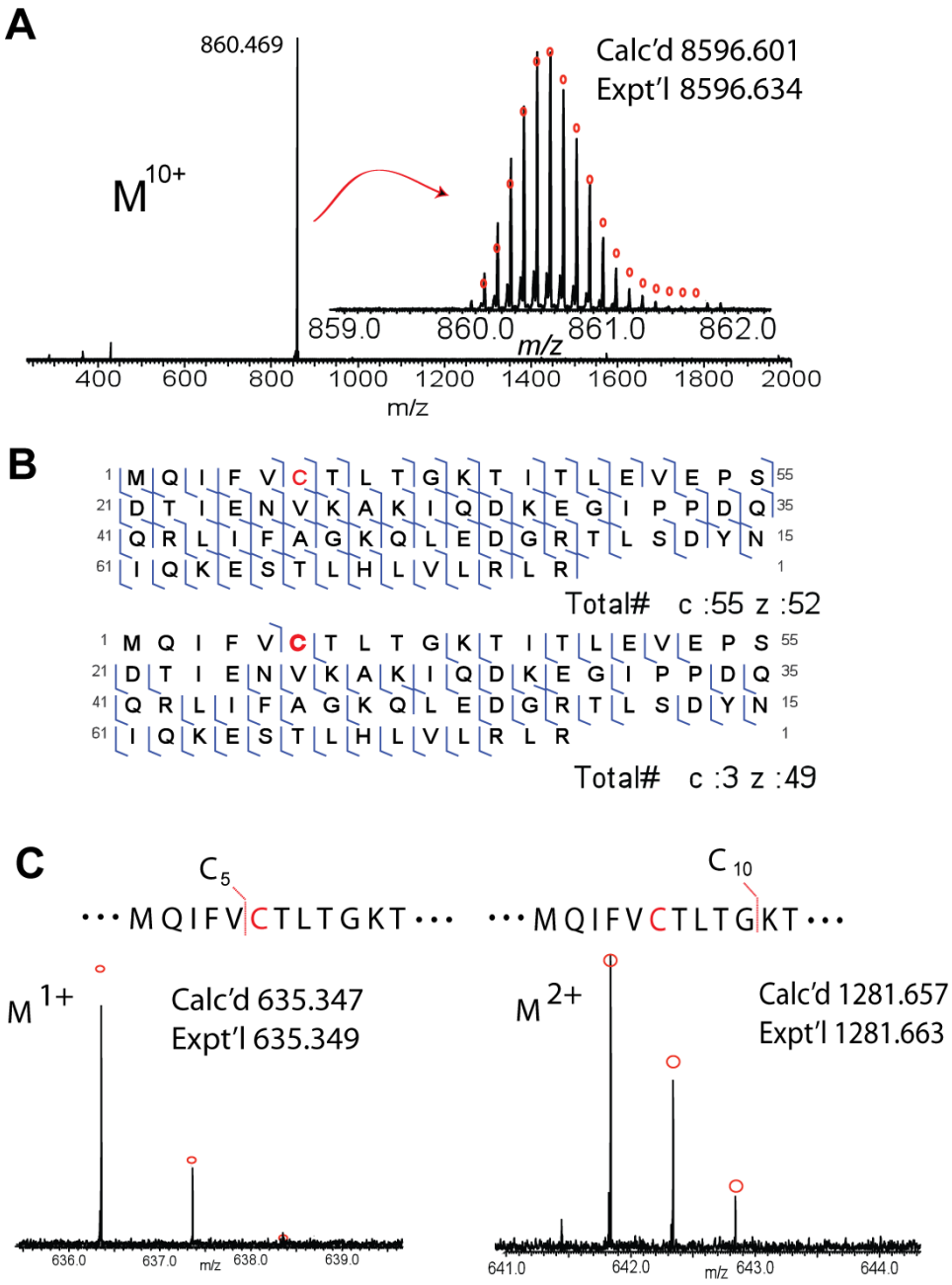


Figure 2.14. ECD analysis of K6C-linked dimer. (A) K6C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top includes *Nε-Gly*-L-homothiaLys thioether linker

modification at cysteine-6 (red) in ***c*** and ***z***· ion predictions. Bottom map does not include thioether linker modification in theoretical analysis. (C) Key ECD fragment ions for mapping thioether linkage site on UbK6C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant mass. Expt'l: experimental most abundant mass.

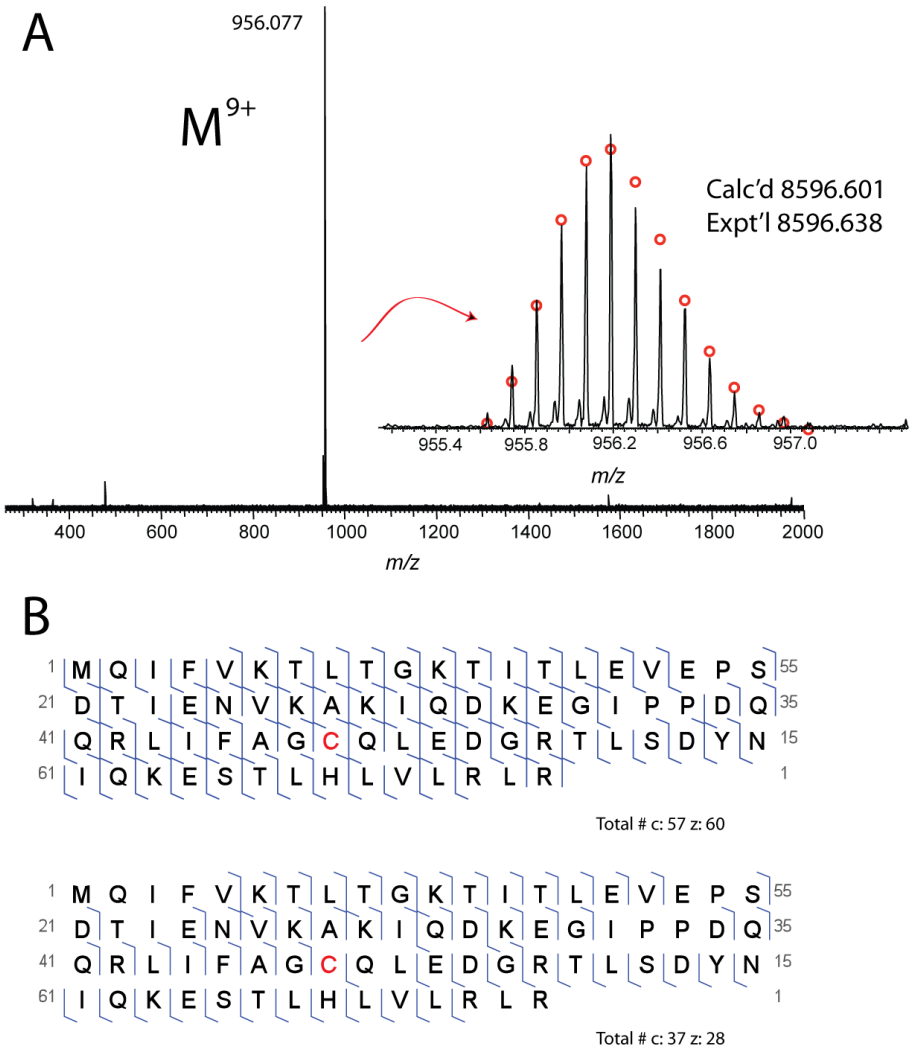


Figure 2.15. ECD analysis of K48C-linked Ub dimer. (A) K48C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{9+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top includes *Nε-Gly*-L-homothiaLys thioether linker modification at cysteine-48 (red) in ***c*** and ***z***· ion predictions. Bottom map does not include thioether linker modification in theoretical analysis.

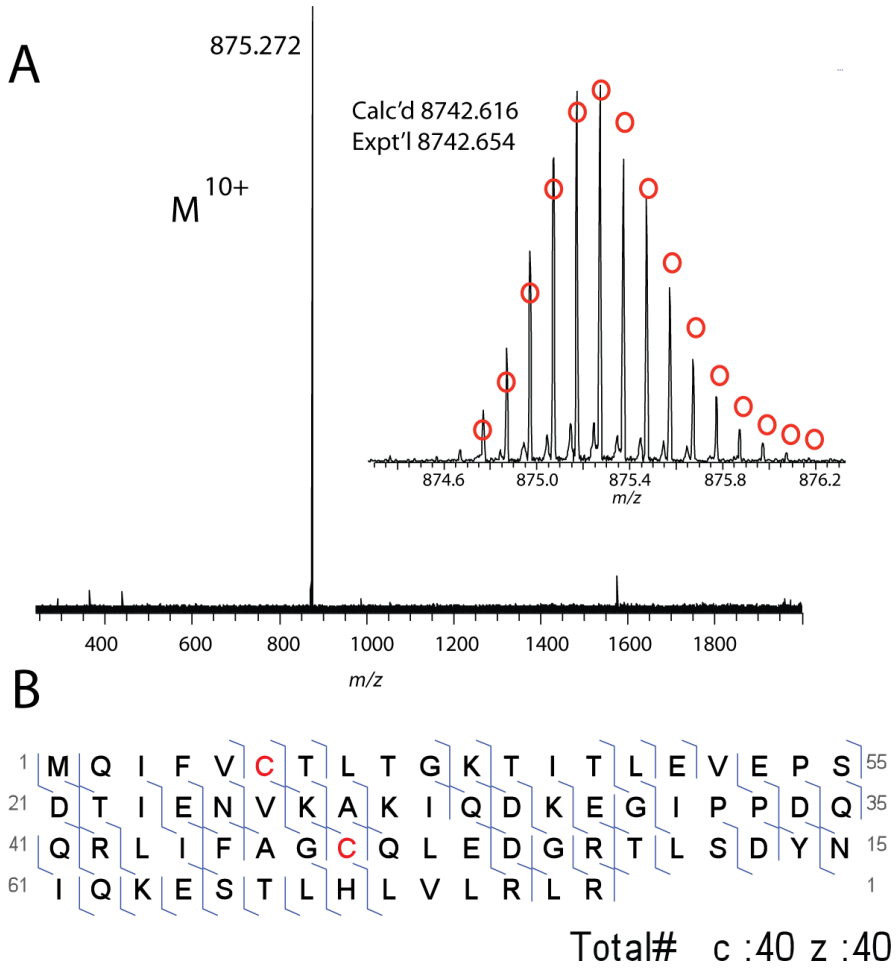


Figure 2.16. ECD analysis of K6C, K48C-linked branched Ub trimer. (A) K6C, K48C Ub_{1-74} $GlyGlyAA_2$ parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis includes *Nε-Gly*-L-homothiaLys thioether linker modification at cysteine-6 and cysteine-48 (red) in ***c*** and ***z·*** ion predictions.

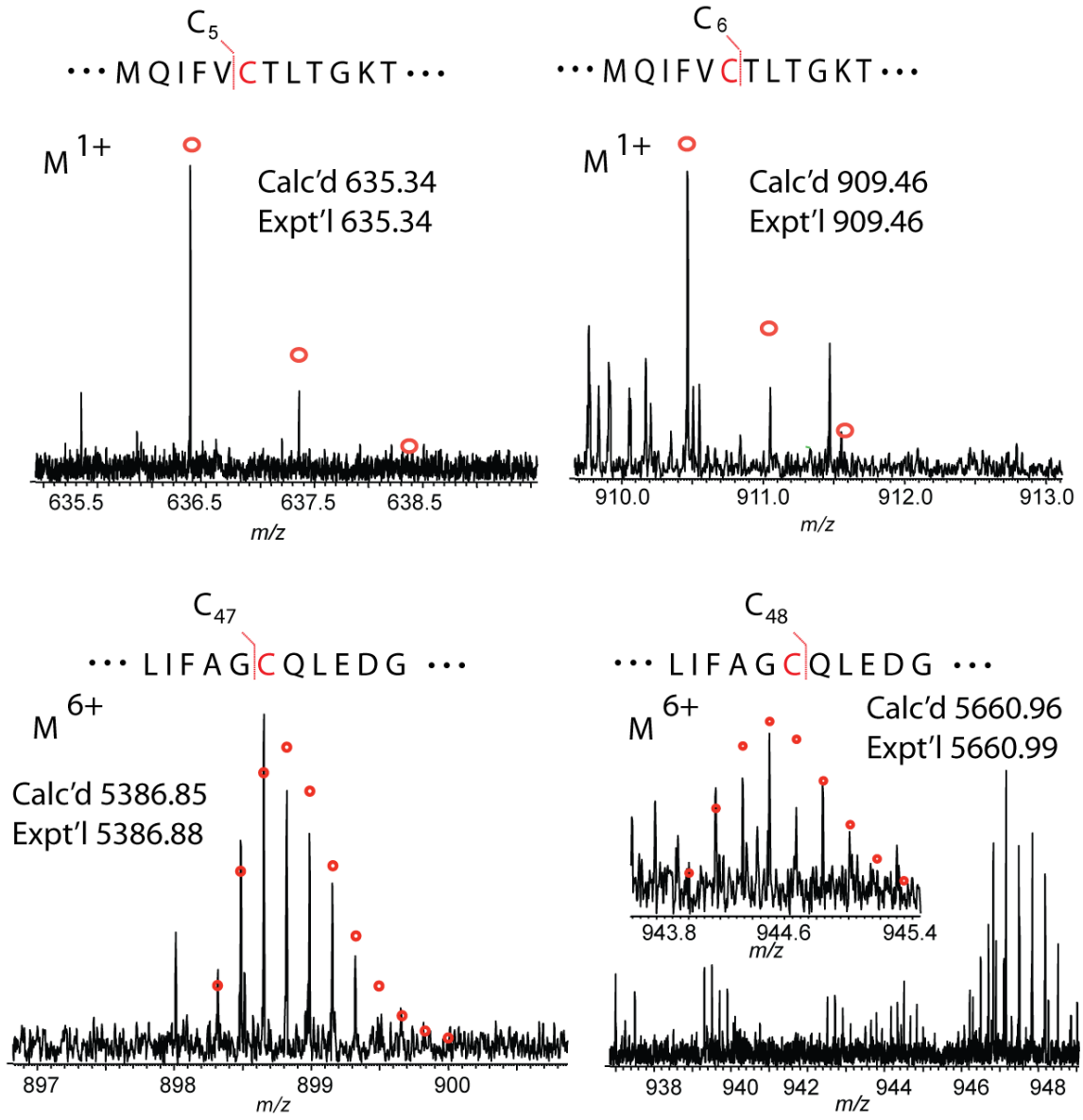


Figure 2.17. Key ECD fragment ions for K6C, K48C-linked trimer. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

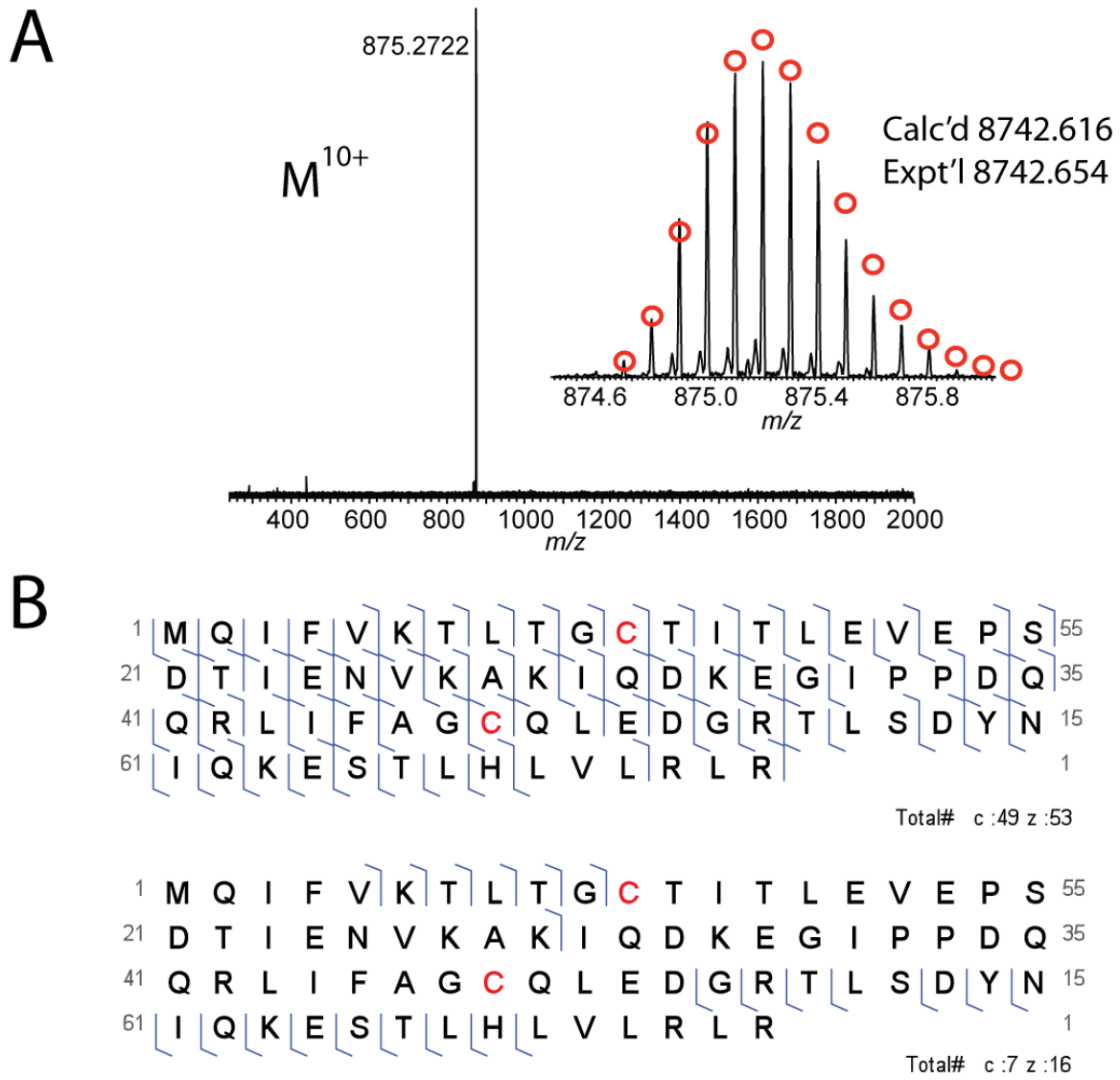


Figure 2.18. ECD analysis of K11C, K48C-linked trimer. (A) K11C, K48C-linked Ub_{1-74} $GlyGlyAA_2$ parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on the top includes the *Nε-Gly*-L-homothiaLys thioether linker modification at cysteine-11 and cysteine-48 (red) in ***c*** and ***z·*** ion predictions. Bottom map does not include thioether linker modifications in the sequence.

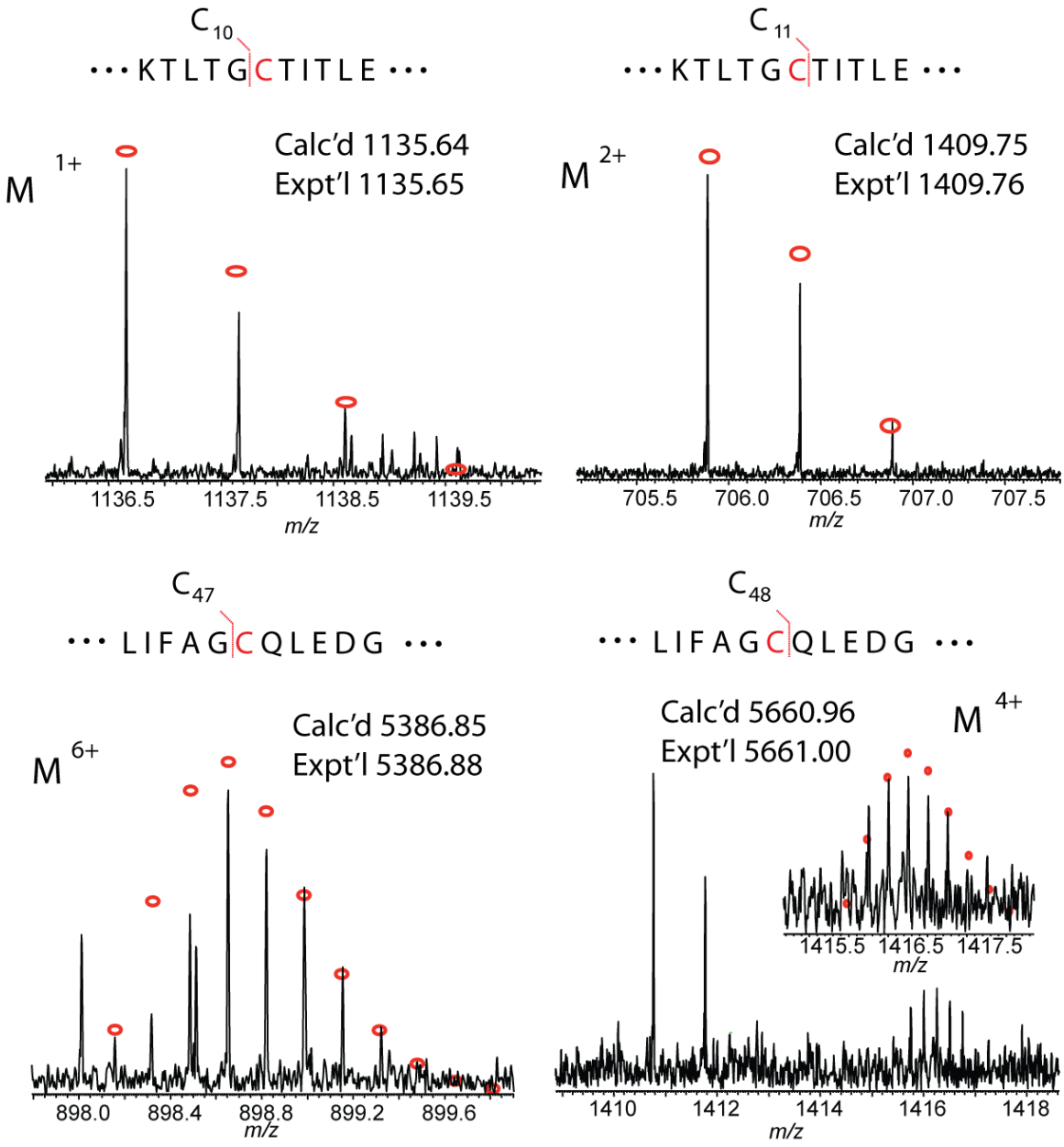


Figure 2.19. Key ECD fragment ions for K11C, K48C-linked trimer. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

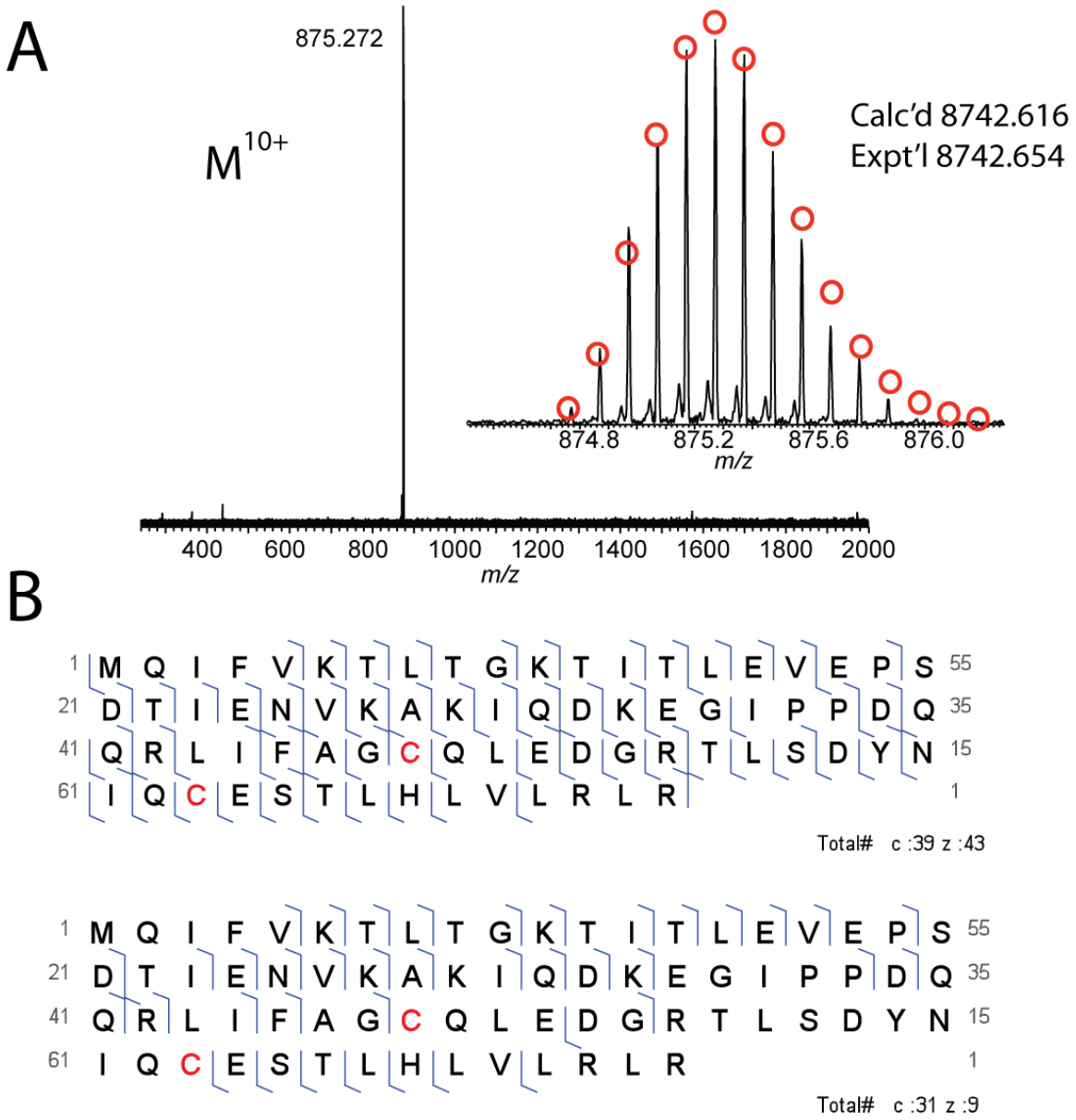


Figure 2.20. ECD analysis of K48C, K63C-linked trimer. (A) K48C, K63C-linked Ub_{1-74} GlyGlyAA_2 parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top includes the *Nε-Gly*-L-homothiaLys thioether linker modification at cysteine-48 and cysteine-63 (red) in ***c*** and ***z·*** ion predictions. Bottom map does not include thioether linker modifications in theoretical analysis.

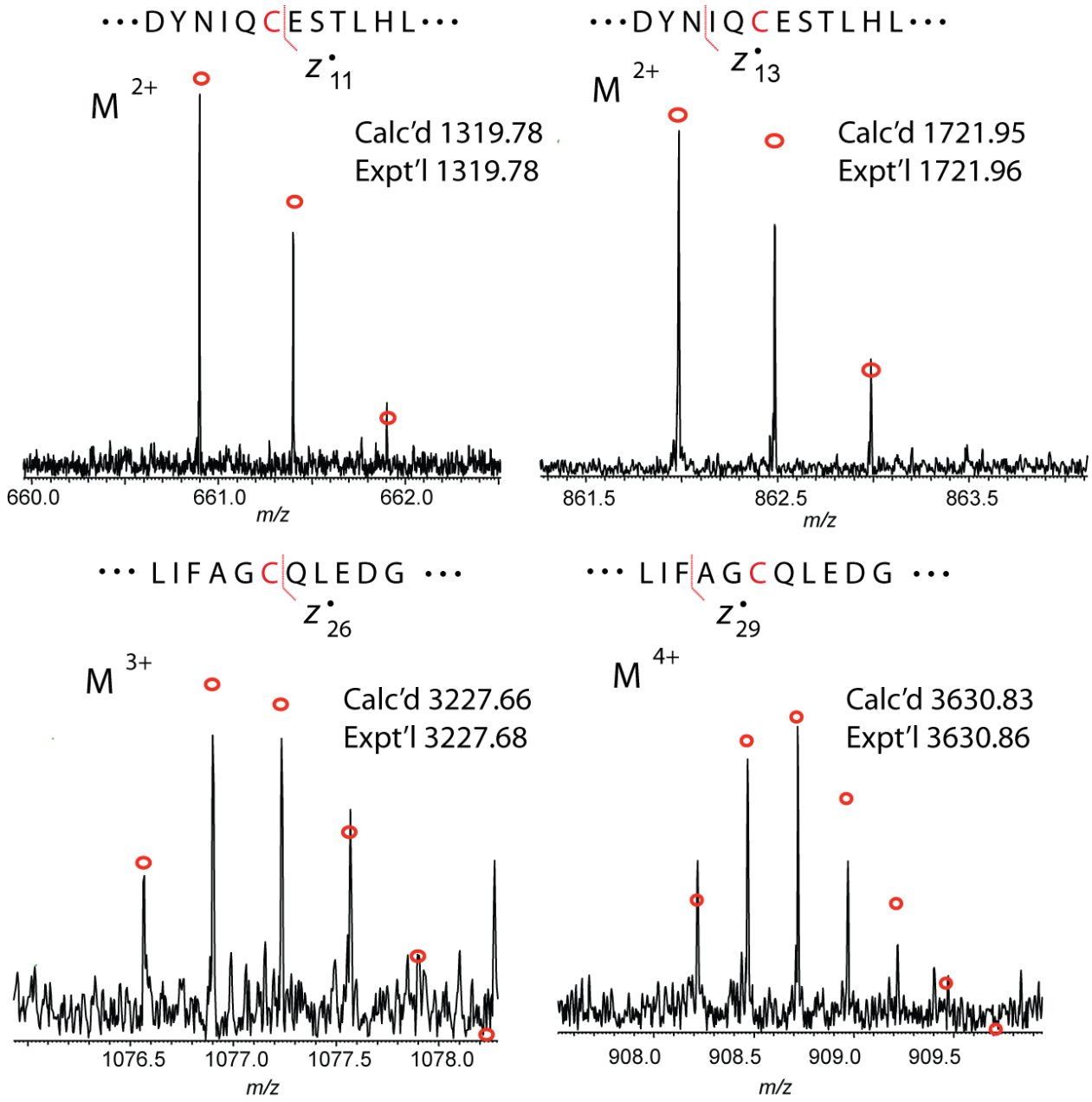


Figure 2.21. Key ECD fragment ions for K48C, K63C-linked trimer. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

VIII. Optimization of difficult linkages (K27C, K29C, K33C)

TEC reactions were performed with varying amounts of Ub-AA and analyzed by high resolution FT-ICR MS. Relative amounts of product can be compared between each spectrum because the amount of UbKxC was the same in each reaction and can therefore be used as an internal standard.

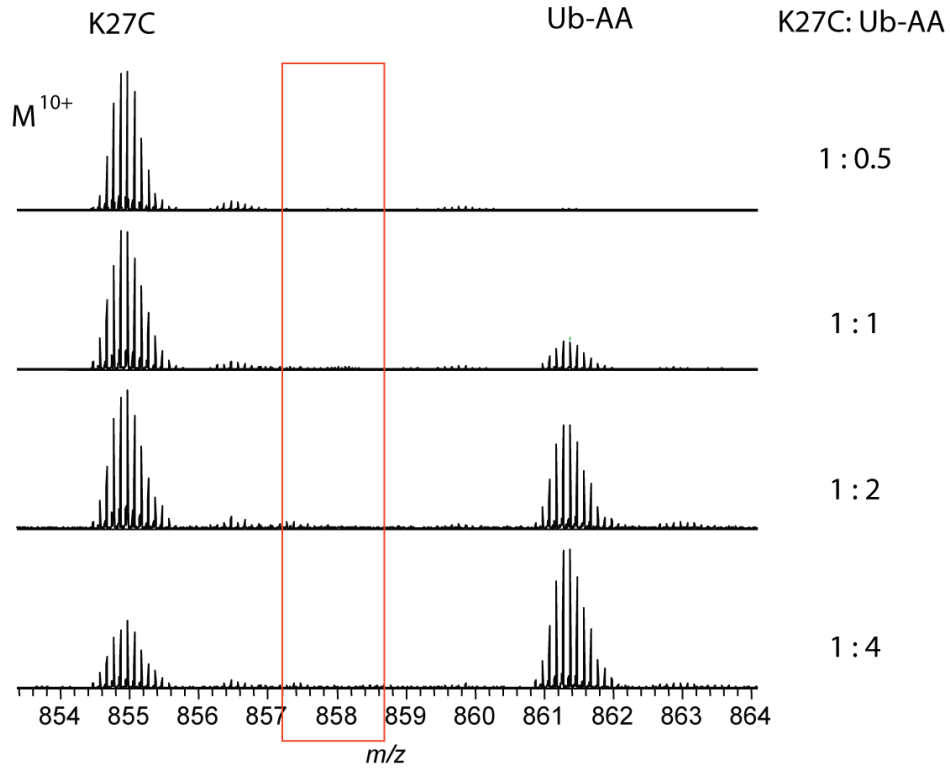


Figure 2.22. High resolution FT-ICR MS analysis of K27C-linked dimer with varying amounts of Ub-AA. Relative amounts of Ub-AA to UbK27C are shown on the right. Red box shows where M^{20+} dimer should appear.

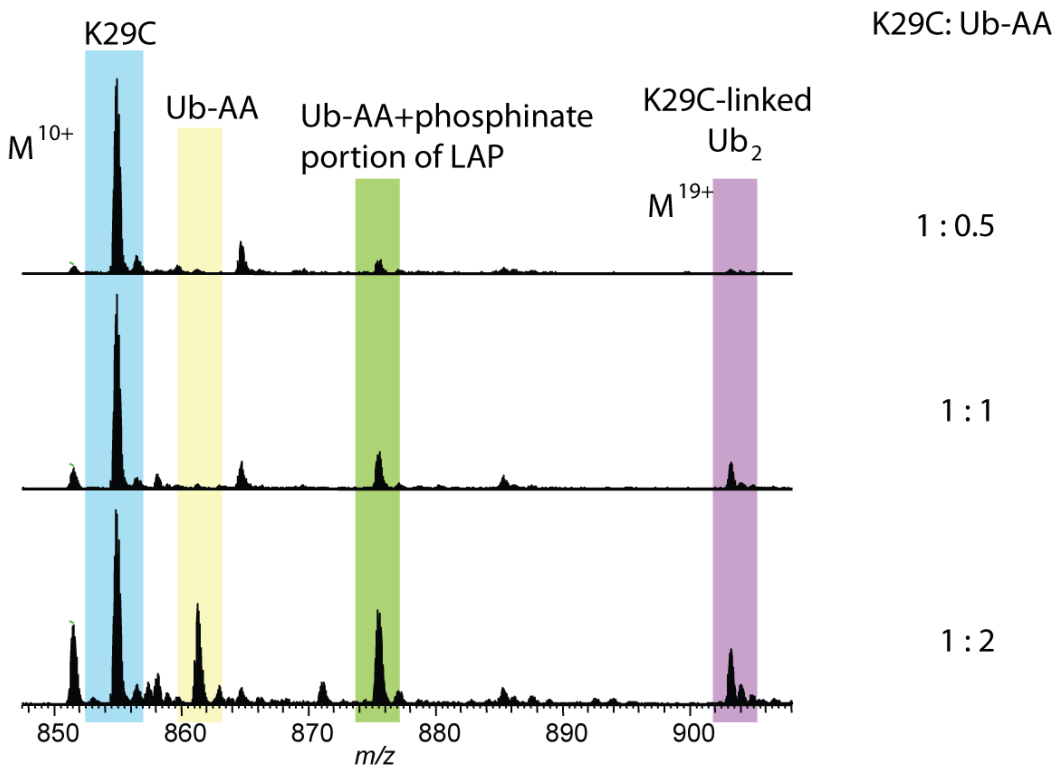


Figure 2.23. FT-ICR analysis of K29C-linked dimer with varying amounts of Ub-AA. The purple box highlights the formation of the desired dimer. Relative concentrations of Ub-AA to UbK29C are shown on the right.

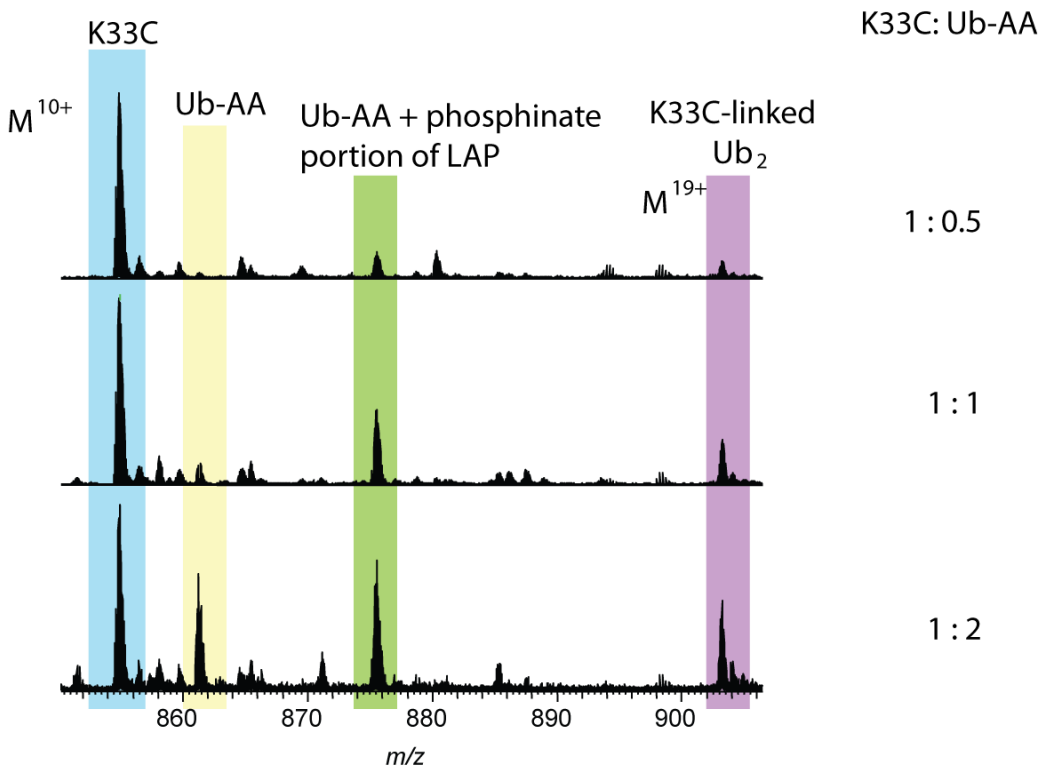


Figure 2.24. High resolution FT-ICR MS analysis of K33C-linked dimer with varying amounts of Ub-AA. The purple box highlights the formation of the desired dimer. Relative concentrations of Ub-AA to UbK33C are shown on the right.

2.4.8 Addition of phosphinate portion of LAP to Ub-AA. Mass spectra of crude reaction mixtures shows a peak that corresponds to the mass of Ub-AA plus the phosphinate portion of the LAP photoinitiator (in Figures 2.23 and 2.24 this peak is highlighted in green). This observation has precedent from the work of Jockusch and Turro,[30] which describes the rapid addition ($k \sim 10^7$ $M^{-1}s^{-1}$) of phosphinoyl radicals to acrylates. Based on this work we propose the process shown in **Figure 2.25** occurs.

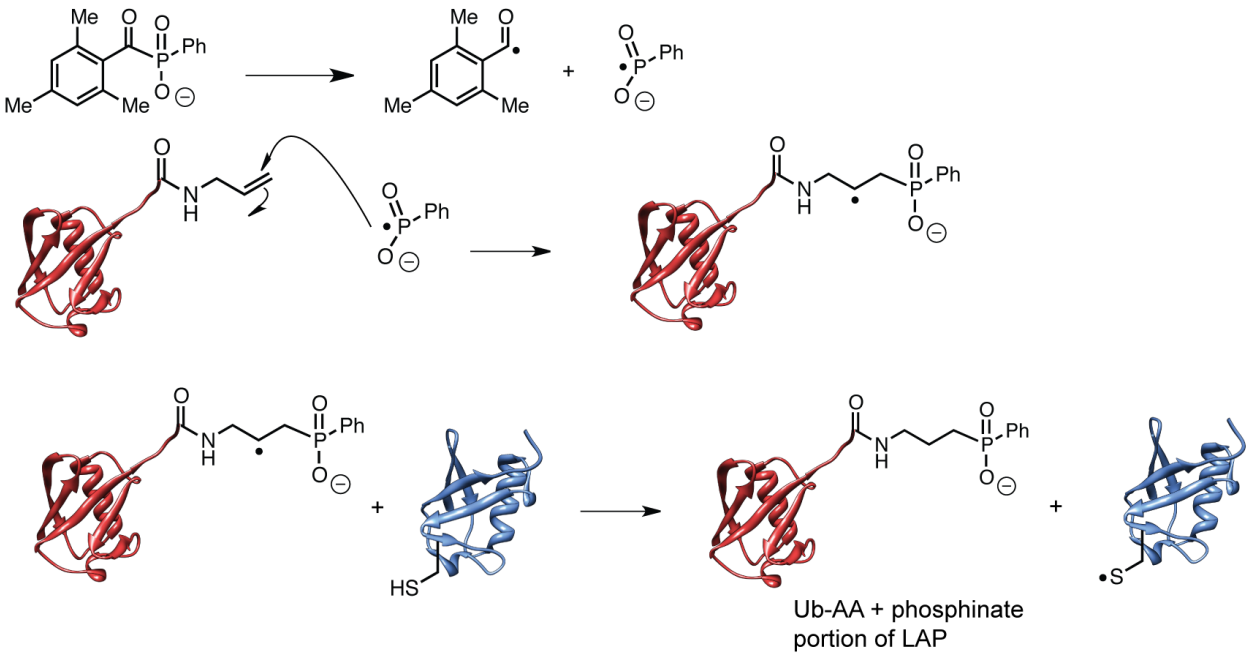


Figure 2.25. Proposed mechanism for the generation of Ub-AA/phosphinate adduct observed in crude reaction mixtures for the TEC reactions.

2.4.9 DUB-catalyzed hydrolysis of Ub dimers and trimers

General procedure. IsoT and A20 were purchased from Boston Biochem, while AMSH was purchased from LifeSensors. For the DUB-catalyzed hydrolysis of Ub dimers, reactions contained a particular KxC-linked Ub dimer (5 mM) and the DUB (5 μM AMSH or 500 nM A20-OTU) in the DUB reaction buffer (50 mM Tris pH 7.6, 25 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) at 37 °C. The DUB was added to the reaction last to initiate hydrolysis. At the time points indicated, 10 μL aliquots were taken and mixed with 3 μL of 6X Laemmli sample loading buffer. Samples were subjected to SDS-PAGE analysis and visualized using silver stain. For the DUB-catalyzed hydrolysis of Ub trimers, reactions contained Ub trimer (20 μM) and the DUB (5 μM AMSH, 500 nM A20-OTU, or 1 mM IsoT) in the DUB reaction buffer at 37 °C. At the time points indicated, 10μL aliquots were taken and mixed with 3 μL of 6X

Laemmli sample loading buffer. Samples were subjected to SDS-PAGE analysis and detected by western blot using anti-ubiquitin antibody (P4D1) from Cell Signaling Technologies.

To provide additional support for the hydrolytic cleavage of K6C, K48C-linked branched tri-Ub with DUBs lacking linkage selectivity, we investigated the activity of USP7. USP7 is a member of the ubiquitin-specific protease (USP) family, and recently Sixma and co-workers reported that the majority of isopeptidases in this family display little linkage selectivity.[31] The reason for specifically investigating the activity of USP7 towards the K6C, K48C-linked trimer is that Ciechanover and co-workers demonstrated the regulation of RING1B by USP7.[32] Autoubiquitylation of RING1B generates a putative branched polyUb chain linked through K6, K27, and K48.[33] We surmised that if a branched chain containing K6- and K48-linkages is indeed attached to RING1B and USP7 is responsible for removing this chain, then USP7 should process K6C, K48C-linked tri-Ub. As shown in **Figure 2.26**, dimeric and monomeric Ub products are immediately produced upon treatment of the branched chain with USP7.

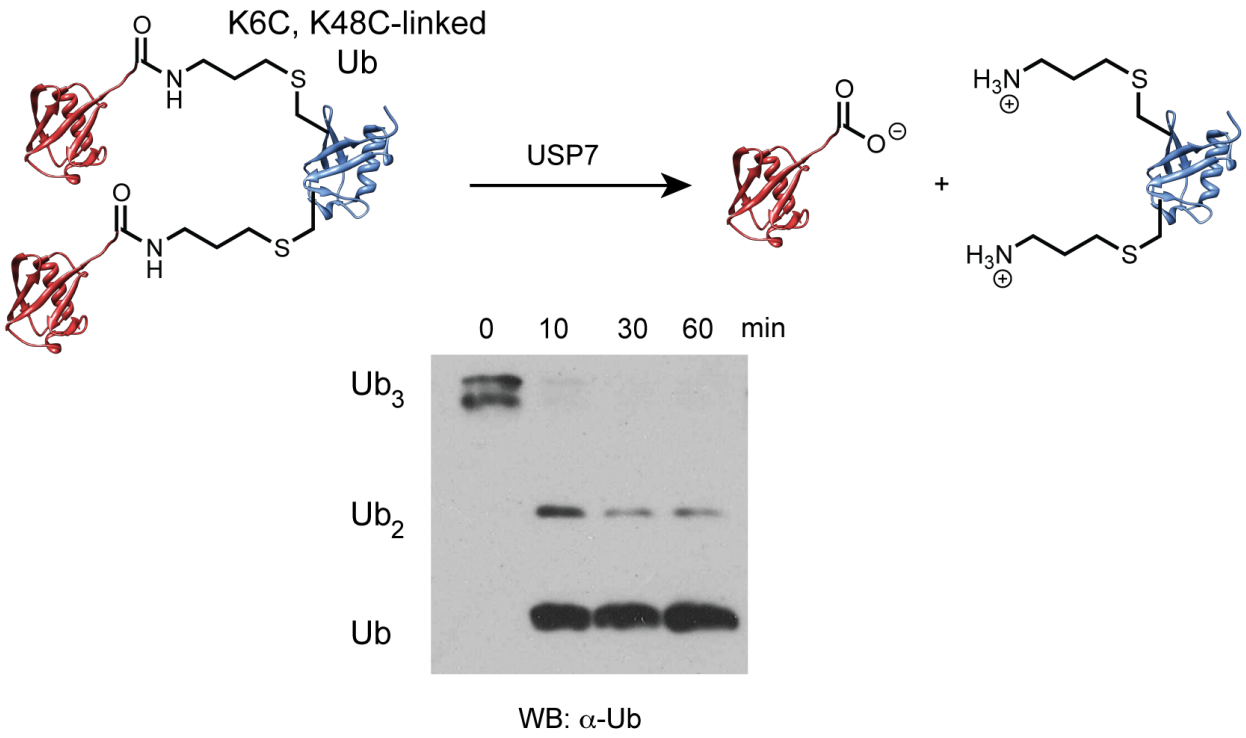


Figure 2.26. Hydrolytic cleavage of K6C, K48C-linked Ub trimer with USP7.

2.5 References

(1) (a) Dondoni, A. *Angew. Chem. Int. Ed. Engl* **2008**, *47*, 8995. (b) Hoyle, C. E.; Bowman, C. N. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 1540.

(2) Reddy, S. K.; Cramer, N. B.; Bowman, C. N. *Macromolecules* **2006**, *39*, 3673.

(3) Kalia, J.; Raines, R. T. *Curr. Org. Chem.* **2010**, *14*, 138.

(4) (a) Chalker, J. M.; Bernardes, G. J.; Davis, B. G. *Acc.Chem. Res.* **2011**, *44*, 730. (b) Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7798.

(5) (a) Nauser, T.; Schoneich, C. *J. Am. Chem. Soc.* **2003**, *125*, 2042. (b) Schoneich, C. *Chem. Res. Toxicol.* **2008**, *21*, 1175.

(6) (a) Dondoni, A.; Massi, A.; Nanni, P.; Roda, A. *Chem-Eur J.* **2009**, *15*, 11444. (b) Lo Conte, M.; Staderini, S.; Marra, A.; Sanchez-Navarro, M.; Davis, B. G.; Dondoni, A. *Chem. Commun.* **2011**, *47*, 11086. (c) Wittrock, S.; Becker, T.; Kunz, H. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 5226. (d) Weinrich, D.; Lin, P. C.; Jonkheijm, P.; Nguyen, U. T.; Schroder, H.; Niemeyer, C. M.; Alexandrov, K.; Goody, R.; Waldmann, H. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 1252. (e) Aimetti, A. A.; Feaver, K. R.; Anseth, K. S. *Chem. Commun.* **2010**, *46*, 5781.

(7) (a) Pickart, C. M. *Annu. Rev. Biochem.* **2001**, *70*, 503. (b) Grabbe, C.; Husnjak, K.; Dikic, I. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 295.

(8) (a) Strieter, E.; Korasick, D. *ACS chemical biology* **2011**. (b) Fekner, T.; Li, X.; Chan, M. K. *Chembiochem* **2011**, *12*, 21.

(9) (a) Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W. *Nat. Chem. Biol.* **2010**, *6*, 750. (b) Castaneda, C. A.; Spasser, L.; Bavikar, S. N.; Brik, A.; Fushman, D. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 11210.

(10) (a) McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W. *Nature* **2008**, *453*, 812. (b) Fierz, B.; Chatterjee, C.; McGinty, R. K.; Bar-Dagan, M.; Raleigh, D. P.; Muir, T. W. *Nat. Chem. Biol.* **2011**, *7*, 113. (c) Chen, J.; Ai, Y.; Wang, J.; Haracska, L.; Zhuang, Z. *Nat. Chem. Biol.* **2010**, *6*, 270. (d) Hejjaoui, M.; Haj-Yahya, M.; Kumar, K. S.; Brik, A.; Lashuel, H. A. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 405.

(11) Chatterjee, C.; McGinty, R. K.; Fierz, B.; Muir, T. W. *Nat. Chem. Biol.* **2010**, *6*, 267.

(12) Johnston, S. C.; Riddle, S. M.; Cohen, R. E.; Hill, C. P. *EMBO J.* **1999**, *18*, 3877.

(13) Fairbanks, B. D.; Schwartz, M. P.; Bowman, C. N.; Anseth, K. S. *Biomaterials* **2009**, *30*, 6702.

(14) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265.

(15) Komander, D.; Clague, M. J.; Urbe, S. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 550.

(16) (a) Komander, D.; Barford, D. *Biochem. J.* **2008**, *409*, 77. (b) Sato, Y.; Yoshikawa, A.; Yamagata, A.; Mimura, H.; Yamashita, M.; Ookata, K.; Nureki, O.; Iwai, K.; Komada, M.; Fukai, S. *Nature* **2008**, *455*, 358.

(17) Kim, H. T.; Kim, K. P.; Lledias, F.; Kisselev, A. F.; Scaglione, K. M.; Skowyra, D.; Gygi, S. P.; Goldberg, A. L. *J. Biol. Chem.* **2007**, *282*, 17375.

(18) Kim, H. T.; Kim, K. P.; Uchiki, T.; Gygi, S. P.; Goldberg, A. L. *EMBO J.* **2009**, *28*, 1867.

(19) Dong, K. C.; Helgason, E.; Yu, C.; Phu, L.; Arnott, D. P.; Bosanac, I.; Compaan, D. M.; Huang, O. W.; Fedorova, A. V.; Kirkpatrick, D. S.; Hymowitz, S. G.; Dueber, E. C. *Structure* **2011**, *19*, 1053.

(20) Reyes-Turcu, F. E.; Shanks, J. R.; Komander, D.; Wilkinson, K. D. *J. Biol. Chem.* **2008**, *283*, 19581.

(21) Faesen, A. C.; Luna-Vargas, M. P.; Geurink, P. P.; Clerici, M.; Merkx, R.; van Dijk, W. J.; Hameed, D. S.; El Oualid, F.; Ovaa, H.; Sixma, T. K. *Chem. Biol.* **2011**, *18*, 1550.

(22) Ben-Saadon, R.; Zaaroor, D.; Ziv, T.; Ciechanover, A. *Mol. Cell* **2006**, *24*, 701.

(23) Pandey, U. B.; Nie, Z.; Batlevi, Y.; McCray, B. A.; Ritson, G. P.; Nedelsky, N. B.; Schwartz, S. L.; DiProspero, N. A.; Knight, M. A.; Schuldiner, O.; Padmanabhan, R.; Hild, M.; Berry, D. L.; Garza, D.; Hubbert, C. C.; Yao, T. P.; Baehrecke, E. H.; Taylor, J. P. *Nature* **2007**, *447*, 859.

(24) Cripps, D.; Thomas, S. N.; Jeng, Y.; Yang, F.; Davies, P.; Yang, A. J. *J. Biol. Chem.* **2006**, *281*, 10825.

(25) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 51.

(26) Pickart, C. M.; Raasi, S. *Meth. Enzymol.* **2005**, *399*, 21.

(27) Majima, T.; Schnabel, W.; Weber, W. *Makromol. Chem.* **1991**, *192*, 2307.

(28) Fairbanks, B. D.; Schwartz, M. P.; Bowman, C. N.; Anseth, K. S. *Biomaterials* **2009**, *30*, 6702.

(29) (a) Ayaz-Guner, S.; Zhang, J.; Li, L.; Walker, J. W.; Ge, Y. *Biochemistry* **2009**, *48*, 8161. (b) Ge, Y.; Rybakova, I. N.; Xu, Q.; Moss, R. L. *Proc. Natl. Acad. Sci. U S A* **2009**, *106*, 12658. (c) Zhang, J.; Guy, M. J.; Norman, H. S.; Chen, Y. C.; Xu, Q. G.; Dong, X. T.; Guner, H.; Wang, S. J.; Kohmoto, T.; Young, K. H.; Moss, R. L.; Ge, Y. *J. Proteome Res.* **2011**, *10*, 4054.

(30) Jockusch, S.; Turro, N. J. *J. Am. Chem. Soc.* **1998**, *120*, 11773.

(31) Faesen, A. C.; Luna-Vargas, M. P.; Geurink, P. P.; Clerici, M.; Merkx, R.; van Dijk, W. J.; Hameed, D. S.; El Oualid, F.; Ovaa, H.; Sixma, T. K. *Chem. Biol.* **2011**, *18*, 1550.

(32) de Bie, P.; Zaaroor-Regev, D.; Ciechanover, A. *Biochem. Biophys. Res. Commun.* **2010**, *400*, 389.

(33) Ben-Saadon, R.; Zaaroor, D.; Ziv, T.; Ciechanover, A. *Mol. Cell* **2006**, *24*, 701.

Chapter 3: Nonenzymatic Polymerization of Ubiquitin: Single-Step Synthesis and Isolation of Discrete Ubiquitin Oligomers

Portions of this work have been published in:

Trang, V. H., Valkevich, E. M., Minami, S., Chen, Y. C., Ge, Y., and Strieter, E. R. (2012) Nonenzymatic polymerization of ubiquitin: single-step synthesis and isolation of discrete ubiquitin oligomers, Angew. Chem. 51, 13085-13088

Contribution:

Chain synthesis and analysis help was provided by V. H. Trang, and Y.-C. Chen

Ubiquitin purification help was performed by S. Minami

3.1 Introduction

In eukaryotes, covalent attachment of ubiquitin (Ub) to target proteins is involved in regulating nearly all biological processes.[1] The most well known function of Ub is in targeting proteins for degradation through the Ub proteasome system (UPS).[2] For UPS processing, the prevailing view has been that a target protein must be modified with a polyUb chain consisting of a minimum of four Ub subunits linked between the C-terminus of one unit to lysine (Lys)-48 of the preceding unit (the *N*ε-Gly-L-Lys linkage is commonly referred to as an isopeptide bond).[3,4] Recent studies, however, suggest the polyUb signal can be much more diverse. That is, chains bearing linkages originating from any of the seven Ub lysines (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63) or the N-terminal methionine can promote protein turnover *in vivo*.[5-9]

We have been interested in understanding the function of unconventional Lys6-linked polyUb chains, as these particular signals are a product of the E3 Ub ligase activity exhibited by the breast cancer associated protein (BRCA1).[10,11] BRCA1 is a major player in the DNA damage response (DDR) pathway and hereditary mutations predispose women to breast and ovarian cancers.[12,13] However, the role of Lys6-linked chains in DDR remains unclear. Some studies implicate these chains in proteasomal degradation,[7,8,14,15] while others argue for a non-proteolytic role. [16,17] Even the importance of BRCA1 E3 Ub ligase activity in DDR and tumor suppression has been called into question.[18] Many of these issues could be resolved by deciphering how Lys6-linked chains are recognized and processed by the proteasome and DDR machinery.

To this end, we sought to develop a straightforward synthesis of long, proteolytically-competent versions of Lys6-linked chains (with ≥ 4 units). Unlike for other chain linkages

(e.g., Lys11, Lys48, and Lys63), [19-23] exploiting the enzymatic activity of BRCA1 to generate Lys6-linked oligomers is not feasible due to the difficulty of expressing and purifying full-length BRCA1. Moreover, current methods to chemically synthesize natively-linked chains with four or more subunits require multiple steps, protecting group manipulations, and often denaturing conditions hence the need for additional approaches. [24-28]

We have previously described the use of thiol-ene coupling (TEC) for the site-specific conjugation of Ub dimers through non-native *N*ε-Gly-L-Lys linkages.[29] These linkages are tolerated by a variety of deubiquitinase enzymes (DUBs), which make TEC an acceptable alternative for the synthesis of Ub chains. Herein we report on a straightforward strategy using free radical TEC to produce discrete Ub oligomers with well-defined linkages. Moreover, we show that these oligomers can be used to investigate substrate preferences of proteasomal components.

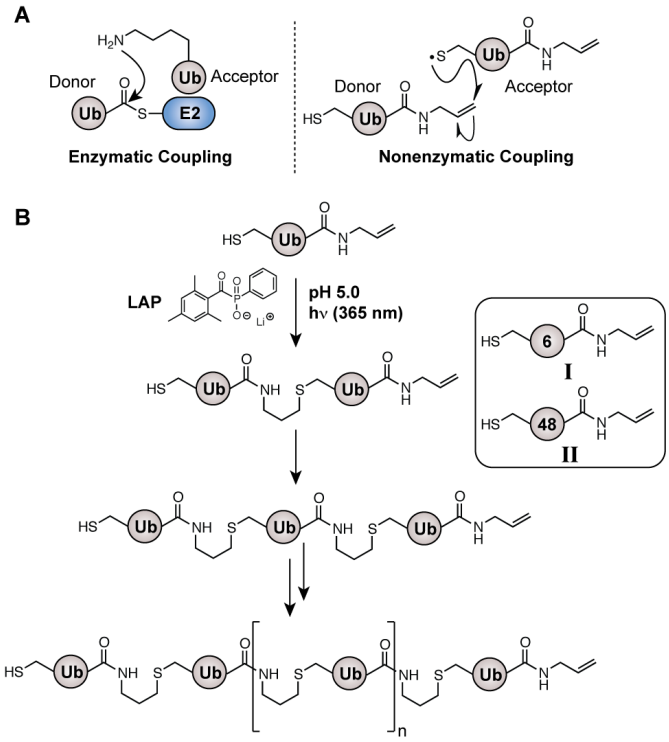


Figure 3.1. (A) Comparison between enzymatic and nonenzymatic coupling. Ub charged E2 thioester (E2-S-Ub) interacts with an acceptor Ub to catalyze isopeptide (N ε--Gly-L-Lys) bond formation. For the nonenzymatic approach a free-radical TEC strategy is shown with a dually functionalized Ub monomer harboring a C-terminal allyl amine adduct and a lysine-to-cysteine mutation. (B) Scheme depicting nonenzymatic polymerization initiated using lithium acyl phosphinate (LAP) and 365 nm light. Discrete oligomers are linked through an N ε-Gly-L-homothiaLys isopeptide-like bond. The two dually functionalized Ub monomers (I and II) used in this study are also shown, wherein the numbers denote the lysine residue mutated to cysteine.

3.2 Results and Discussion

Our plan was to emulate the enzymatic logic of polyUb chain formation. Three enzymes–E1 Ub-activating, E2 Ub-conjugating, and E3 Ub-ligating–catalyze Ub polymerization.[30,31] In some cases, a few E2 enzymes are capable of catalyzing the formation of linkage-specific polyUb chains *in vitro* in the absence of an E3.[20,23,32] E2 enzymes achieve linkage specificity by orienting a particular lysine of an acceptor Ub for nucleophilic attack on the donor ubiquityl thioester (Figure 3.1A).[33-35] Single-linkage (homotypic) oligomers of different lengths are then generated after successive rounds of conjugation through a step-growth polymerization process. With this mechanism as a model, we envisioned expanding the repertoire of available homotypic chains beyond Lys11-, Lys48-, and Lys63-linkages by using a dually functionalized Ub monomer (Figure 3.1A). In this Ub variant, the C-terminal allyl amine appendage acts as the activated E2-*S*-Ub intermediate and the thiol moiety of cysteine serves as the lysine surrogate providing linkage-specificity. Upon photopolymerization via free-radical thiol-ene chemistry[29,36] under non-denaturing conditions, discrete homotypic oligomers can therefore be obtained (Figure 3.1B).

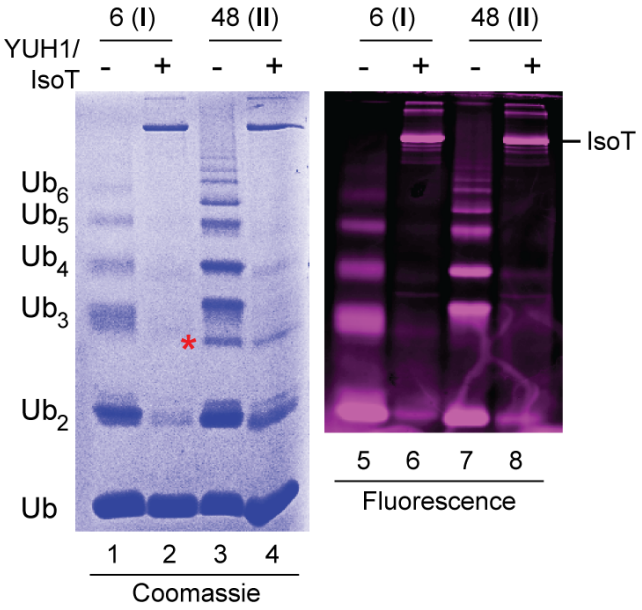


Figure 3.2. Thiol-ene polymerization reactions using monomers **I** and **II**. Reaction conditions for polymerization: **I** or **II** (2 mM), LAP (500 μM) irradiated at 365 nm for 30 minutes. Polymerization reactions were visualized by Coomassie staining and fluorescence, the latter requiring addition of tetramethylrhodamine-5-maleimide (TAMRA) post polymerization. For the removal of polymers via the action of IsoT, polymerization reaction mixtures were incubated with IsoT (2 μM) for 2 hours at 37 °C after a 16 h incubation with YUH1 (750 nM) to remove allyl amine appendages. The cyclic trimer is marked with an asterisk (*).

We initially studied the polymerization of monomers **I** and **II** with the goal of constructing (i) a set of oligomers harboring linkages currently unattainable by enzymatic methods (6-linked chains), and (ii) a control series with enzymatically available 48-linkages.[19,21,23] Exposing **I** and **II** to thiol-ene conditions revealed rapid formation of a series of discrete oligomers (Figure 3.2; lanes 1 and 3). Polymerization of **II** occurred to a much greater extent than that for **I** based on the formation of higher molecular weight species with 48-linkages. This difference is likely due to the relative steric hindrance of the

two positions; position-6 is part of a β-sheet making it less exposed relative to position-48, which is located in a loop region. Another feature of these polymerization reactions is the amount of monomer that remains. This could be a consequence of the fast termination kinetics (10^8 $M^{-1}s^{-1}$) typically observed for thiol-ene polymerizations.[37-39] Consistent with this notion, we found that the conversion of monomers into higher molecular weight species came to a halt almost immediately after irradiation (Figure 3.6).

Encouraged by these results, it was important for us to verify that the chain ends were not altered during polymerization given modifications such as cyclization may affect biochemical function.[40] Accordingly, reactions were monitored by fluorescence to determine whether each of the species observed by Coomassie contained a reactive thiol moiety. We reasoned any species lacking a free thiol moiety could represent a cyclic oligomer in which the distal and proximal Ub units are tethered through an *N*ε-Gly-L-homothiaLys bond. The results obtained by labeling with TAMRA (tetramethylrhodamine-5-maleimide) closely mirrored Coomassie staining with one key exception. That is, a band appearing as a lower molecular weight 48-linked trimer was not observed by fluorescence (Figure 3.2; compare lanes 3 and 7). Since isopeptidase T (IsoT) is a DUB sensitive to cyclization[40], we sought to further characterize these products with this enzyme. Prior to IsoT-catalyzed hydrolysis, C-terminal allyl amine modifications were removed by the yeast Ub hydrolase YUH1. This step was necessary to maximize hydrolytic cleavage of all the oligomers because IsoT prefers a C-terminal carboxy motif.[41,42] Introduction of IsoT led to disassembly of all 6-linked oligomers and most of the 48-linked oligomers except for the faster migrating trimer as evidenced by Coomassie staining (Figure 3.2; lane 4).

Fluorescence analysis of the IsoT reactions showed an absence of a band corresponding to the putative cyclic trimer, consistent with the loss of a free thiol moiety (Figure 3.2; lane 8).

To characterize the remaining acyclic products, we isolated individual oligomers and assessed the *N*ε-Gly-L-homothiaLys linkages. Using a size-exclusion protocol, milligram quantities of the entire series from tetramers to heptamers could be obtained for both 6- and 48-linked oligomers (Figure 3.3A). SDS-PAGE analysis showed that 48-linked oligomers produced via thiol-ene chemistry migrate through polyacrylamide in the same manner as those produced enzymatically.[23] Moreover, a commercially available rabbit monoclonal antibody (D9D5) raised against polyUb chains linked through Lys48 exhibited binding to oligomers with *N*ε-Gly-L-homothiaLys$_{48}$ linkages (Figure 3.3B). To confirm the linkage between each subunit, we used Fourier-transform ion cyclotron (FT-ICR) MS analysis. Exploiting the stability of Ub in the presence of trypsin, oligomers were minimally digested to generate two species: Ub$_{1\text{-}74}$ (**IV**) and Ub$_{1\text{-}74}$ with a 171.2 amu addition due to the Gly-Gly-allyl amine appendage (**V**) (Figure 3.3C).[29,43,44] MS analysis of these digests demonstrated the presence of both **IV** and **V** (Figure 3.3D). Note that if undesired C-S or C-C radical recombination products formed during polymerization, our minimal digest approach would identify these products. However, no other Ub variants carrying cysteine modifications besides **V** were observed. All position-specific modifications were verified by electron capture dissociation (ECD) and/or collision-activated dissociation (CAD) MS/MS analysis (see materials and methods). Together, these data indicate thiol-ene chemistry is a powerful method for producing long homotypic Ub oligomers.

To study the function of 6-linked chains in the context of proteasomal components, we carried out hydrolysis assays with proteasome-associated DUBs. For comparison, 48-linked

tetramers were also analyzed, as these are the canonical signals for proteasomal turnover.[3] Prior to evaluating the stability of tetramers in the presence of proteasome-associated DUBs we wanted to verify that these oligomers could serve as substrates for DUBs other than IsoT. To this end, we used a DUB with

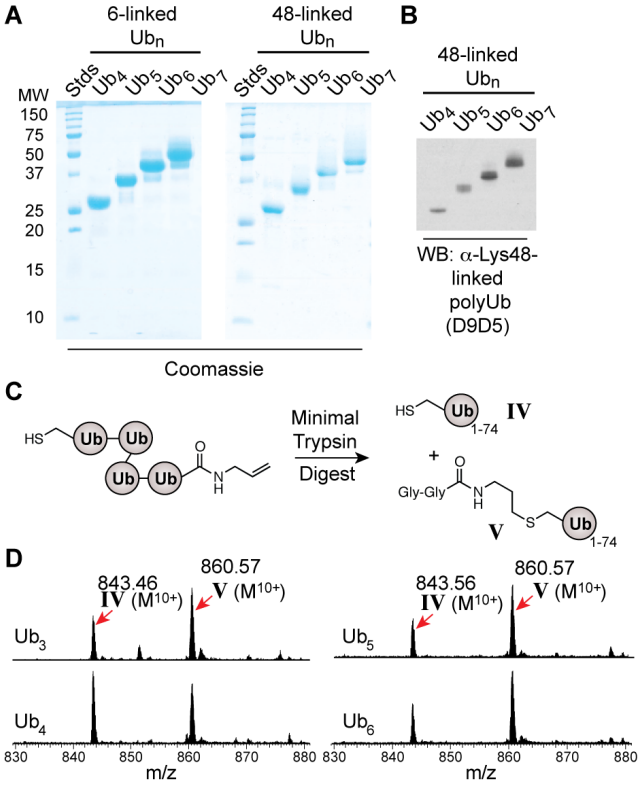


Figure 3.3. Isolation and characterization of discrete 6- and 48-linked Ub oligomers. (A) SDS-PAGE analysis of purified oligomers. (B) Western blot (WB) analysis of 48-linked oligomers is shown on the right. For this experiment, the D9D5 monoclonal antibody (Cell Signaling Technology) was used as it recognizes the *N*ε-Gly-L-Lys$_{48}$ isopeptide linkage. (C) Minimal trypsin digest of oligomers results in the formation of two distinct Ub monomers **IV** and **V**. (D) Representative FT-ICR MS spectra of minimal trypsin digests of various 6-linked oligomers. All peaks are shown in the M^{10+} ionization state.

selectivity for Lys48-linked chains (OTUB1; otubain 1)[45] and one without a preference for linkage type (USP7; Ub specific protease 7).[46] Consistent with the reported activity of OTUB1 we observed disassembly of 48-linked tetramers into smaller fragments within 90

minutes, whereas the 6-linked tetramer remained intact (Figure 3.4A). USP7 on the other hand efficiently degraded both 6- and 48-linked tetramers (Figure 3.4A). With these observations along with the antibody-binding assay (Figure 3.3B) we conclude that tetramers forged through thiol-ene polymerizations are good models for native oligomers.

For experiments with proteasomal DUBs, we examined 6- and 48-linked tetramers in the presence of UCH37 and the 19S regulatory particle (RP). The 19S RP, which acts as a receptor for polyubiquitinated target proteins bound to the 26S proteasome, is composed of three DUBs (UCH37, USP14, and RPN11/POH1) but only two hydrolyze bonds between subunits within a chain (USP14 and UCH37).[47] With UCH37 alone, we observed a preference for 48-linked tetramers over 6-linked substrates (Figure 3.4B). However, in the context of 19S both linkage types were hydrolyzed to a similar extent (Figure 3.4B). These data suggest Ub-binding proteins present in the 19S RP could override intrinsic substrate preferences of proteasome-associated DUBs.

These findings raise another important point related to what constitutes a proteolytically-competent Ub signal. Lys48- and Lys63-linked polyUb chains have long been thought to govern distinct biological pathways: the former is a proteolytic signal whereas the latter is non-proteolytic. The underlying biochemical details of these discrete functions are perplexing considering Lys63-linked chains promote proteasomal degradation *in vitro*.[48] However, a recent study by Liu and coworkers shed light on this paradox.[15] They found Lys63-linked chains are sequestered away from the proteasome by linkage-specific Ub-binding domains, but also proteasome-associated DUBs hydrolyze these linkages faster than Lys48-linked chains. These latter results advocate for a kinetic proofreading mechanism to ensure the right substrate is degraded. If a chain is particularly

susceptible to hydrolysis by proteasomal DUBs, e.g., Lys63-linked chains, the rate of substrate unfolding and proteolytic cleavage cannot compete with dissociation. This scenario prevents substrate degradation. By contrast, chains less vulnerable to the action of proteasomal DUBs, such as Lys48-linked chains, allow for unfolding and proteolysis to compete with substrate dissociation. Given the similar stabilities of 6- and 48-linked oligomers in the presence of the 19S RP, it is tempting to speculate Lys6-linked chains can support proteasomal degradation. Additional kinetic and substrate degradation studies are necessary to validate this proposal.

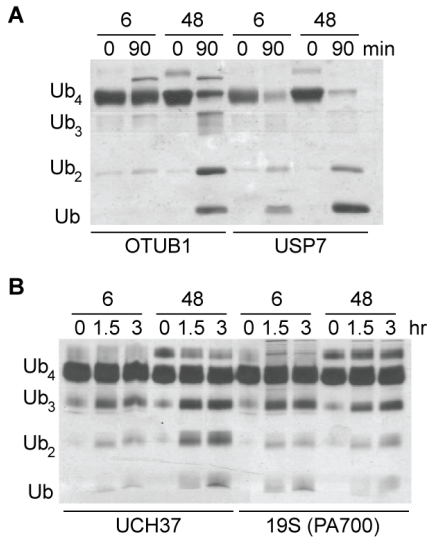


Figure 3.4. DUB-catalyzed cleavage of 6- and 48-linked Ub tetramers. (A) Silver-stained SDS-PAGE analysis of disassembly of 6- and 48-linked tetramers (10 μM) using the 48-linkage selective DUB OTUB1 (1 μM) and the promiscuous DUB USP7 (1 μM). (B) Silver-stained SDS-PAGE analysis of tetramer (10 μM) hydrolysis using proteasomal components UCH37 (1 μM) and the 19S (PA700) regulatory particle (100 nM), which harbors three deubiquitinating enzymes USP14, UCH37, and POH1/RPN11.

In summary, this work highlights how free-radical thiol-ene polymerizations can be applied to the nonenzymatic synthesis of polyUb chains with well-defined linkages. Using

this approach we have shown for the first time that homotypic polyUb chains with up to seven subunits can be obtained in a single step. With rapid access to long chains we will be able to gain more insight into the role of 6-linked chains in the DNA damage response pathway as well as the UPS.

3.3 Materials and Methods

3.3.1 Ubiquitin (Ub) cloning and expression

Cloning Ubiquitin (Ub) KxCD77 (x denotes the position within Ub amino acid sequence) was constructed by introducing cysteine mutations at specific sites using splice overlap extension.[49] Primers containing the TGC mutation were inserted at the desired lysine codon. D77 was encoded in the reverse primer to afford KxCD77.

Expression and Purification. The Ub variants were expressed and purified as described previously[50,51] with one key exception related to purification. That is, upon dialysis into Buffer A (50 mM NH_4OAc pH 4.4, 1 mM EDTA, 1 mM DTT), the lysate was batch bound to 10 mL of SP Sepharose Fast Flow resin (GE life sciences) for 1 hr. The resin was then washed with 50mL Buffer A, 50mL 5% Buffer B (50 mM NH_4OAc pH 4.4, 1 mM EDTA, 1 mM DTT, 1M NaCl), and 50mL 10% Buffer B. The Ub variants were then eluted with 25% Buffer B. Fractions containing Ub (monitored by SDS PAGE) were exchanged into H_2O and lyophilized.

Synthesis of Ub allyl amine adduct (UbKxC-AA)

UbKxCD77 (185.6 mg, 21.7 μM) was dissolved in a buffer containing 50 mM Hepes pH 8, 1 mM EDTA, 30% DMSO, and 250 mM allylamine to a total reaction volume of 25 mL. YUH1[51]

(0.5 μM) was added to initiate the reaction and the mixture was allowed to incubate at room temperature for two hours. The reaction mixture was quenched with TFA (pH~ 2), exchanged into Buffer A (50 mM NH_4OAc pH 4.4, 1 mM DTT, 0.5 mM EDTA) and purified by cation exchange chromatography using previously described methods.[51] The molecular weight (MW) of the UbKxC-AA conjugate was verified using a matrix assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometer Bruker REFLEXII. Spectra were processed using mMass[52,53].

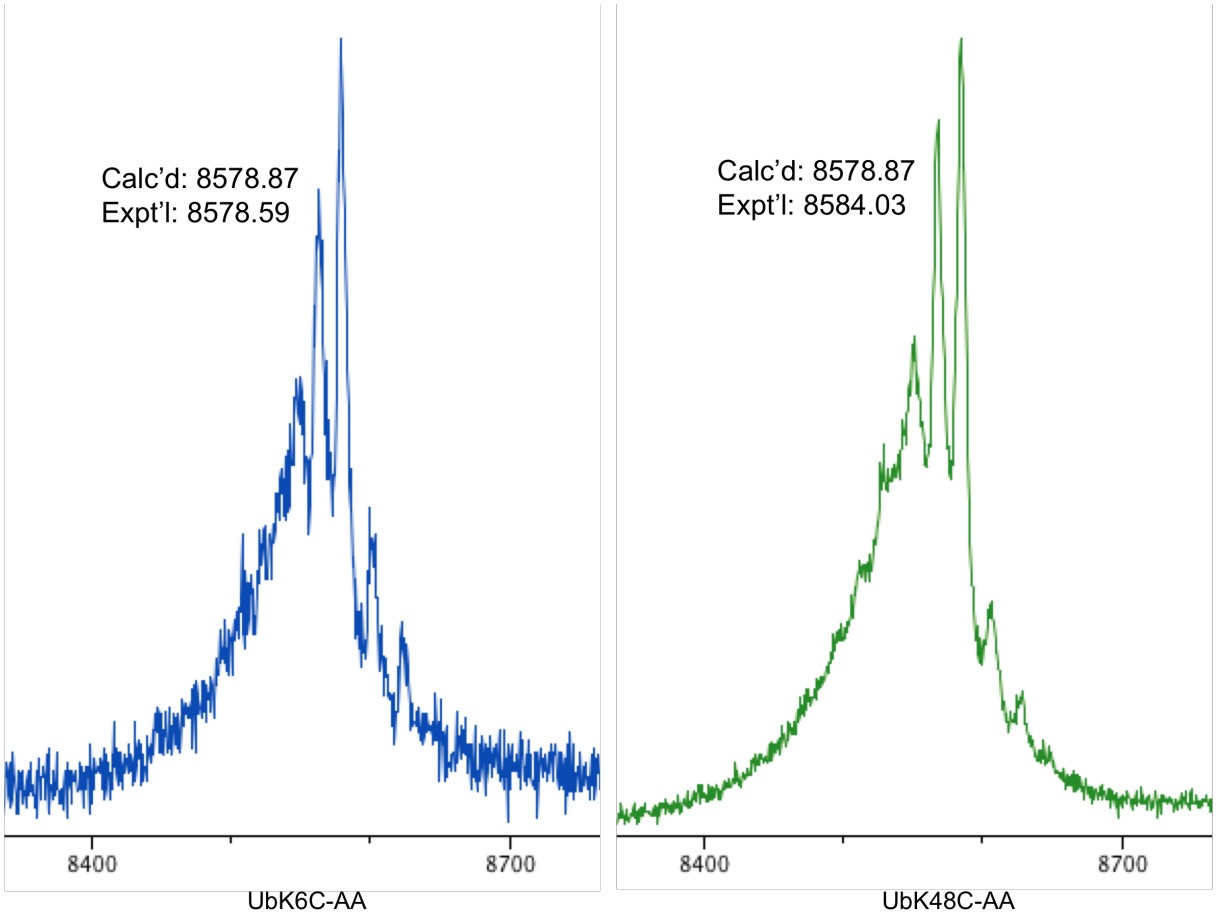


Figure 3.5: MALDI/TOF MS spectrum of UbK6C-AA and UbK48C-AA, calibrated to wildtype Ub. Calc'd: calculated average MW. Expt'l: experimental calculated average MW.

3.3.2 Ub Polymerization via Thiol-ene coupling (TEC)

Ub polymerization reactions contained UbKxC-AA (2 mM), lithium acyl phosphinate[51, 54, 55] (LAP) (0.5 mM) in 250 mM NaOAc buffer pH 5 in 100 mL reaction volume. Samples were placed on ice and irradiated with 365 nM light for 30 minutes using an OmniCure series 1500 light source placed 15 cm above the sample. As a negative control, reactions were carried out in the absence of the LAP photoinitiator. The dually functionalized monomers used in this study were: UbK6C-AA and UbK48C-AA.

A time course was performed to highlight the rate at which discrete oligomers are formed. As shown in **Figure 3.6** discrete oligomers along with high molecular weight products are produced almost instantaneously. This experiment also shows the reaction stops immediately.

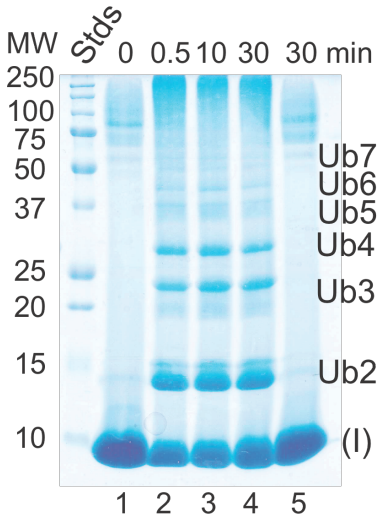


Figure 3.6: Time course of K48C-AA. Reactions shown in lanes 1 and 5 are missing LAP, but are irradiated for the indicated time.

Purification of Ub polymers: 20 TEC reactions (2 mL total) were combined for each polymerization and purified using a Superdex 75 HiLoad size exclusion column at a flow rate of 0.2 mL/min, collecting 2.5 mL fractions over 0.7 column volumes. This allowed for

separation of monomers up to heptamers. Further purification of polymers via SEC was performed when necessary. Anti-Lys48-linked polyUb (D9D5) from Cell Signaling Technology was used to verify the K48 linkage.

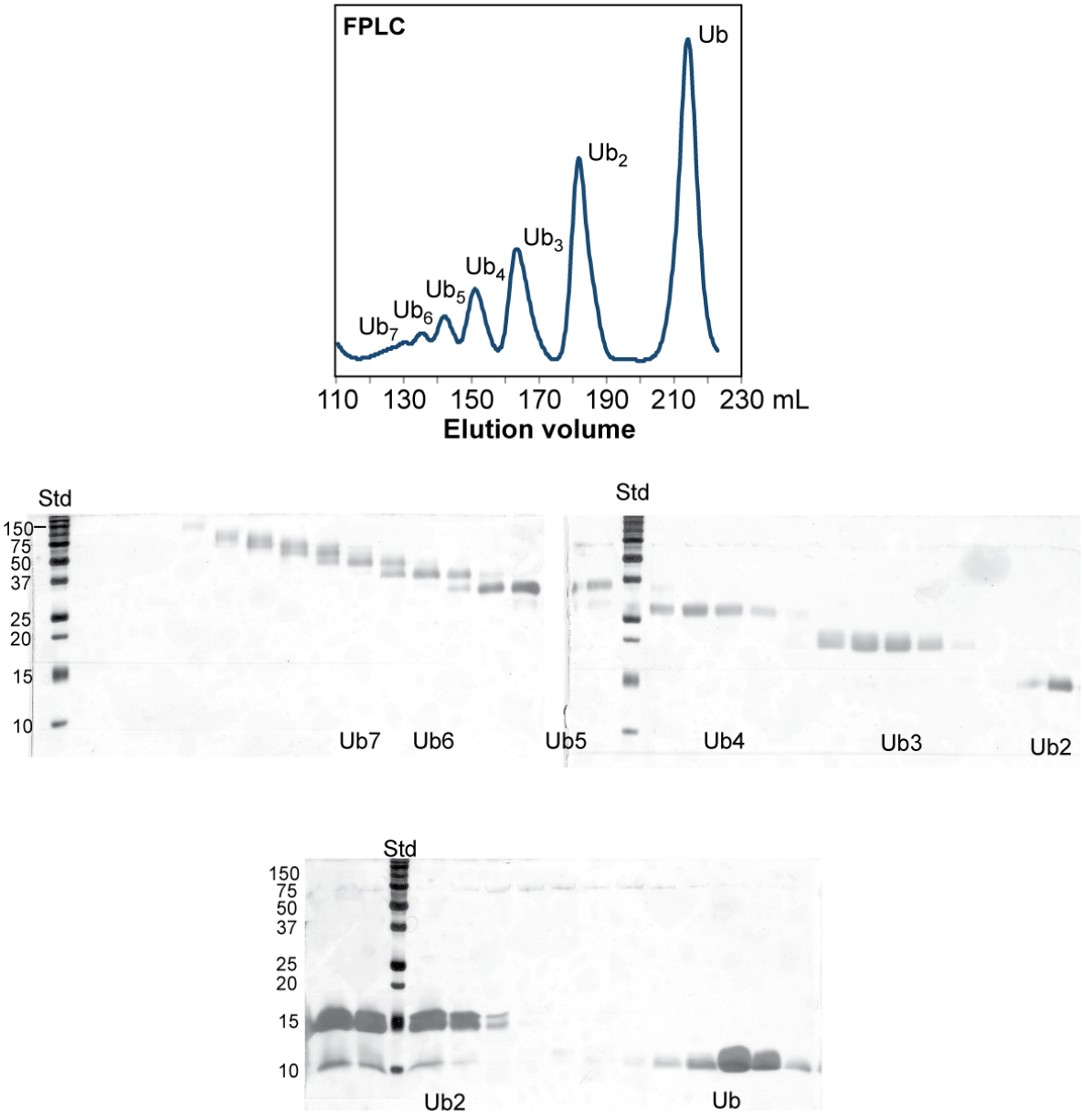


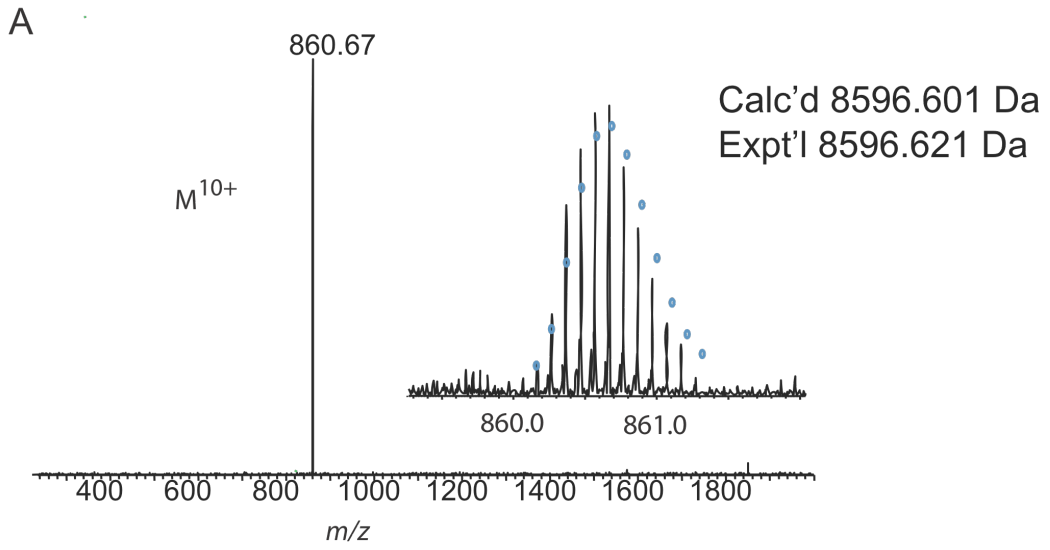
Figure 3.7: Representative purification of thiol-ene polymerization. FPLC chromatogram for 6-linked polymers shows separation of distinct polymers from ubiquitin (Ub) monomer to ubiquitin heptamer (Ub7). Fractions were analyzed by SDS-PAGE and visualized using silver stain.

3.3.3 Characterization of 6- and 48-linked polymers

Sample preparation using the minimal digest method: To verify the K6C and K48C linkages, polymers were minimally digested[51] with trypsin (6-12h, 1:25-1:50 trypsin:Ub ratio) in ammonium bicarbonate (160mM) buffer. Digests were quenched with 10% acetic acid and diluted 1:4 (acetonitrile: sample) prior to injection into Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer.

Mass Spectrometry General Procedure: Samples were injected into a 7T linear ion trap LTQ/FT-ICR hybrid mass spectrometer (Thermo Scientific Inc., Bremen, Germany) equipped with an automated chip-based nanoESI source (Triversa NanoMate, Advion BioSciences, Ithaca, NY).[56] The resolving power of the FT-ICR mass analyzer was set to either 100,000 or 200,000. All FT-ICR spectra were processed with in-house developed Software (MASH SUITE version 1.0) using a signal to noise threshold of 3 and a fit factor of 60%. Each measurement was then validated manually.

Electron capture dissociation (ECD) and Collisionally Activated Dissociation (CAD) analysis: ECD (using a 3-5% "electron energy" and a 70-75 ms duration with no delay) and CAD (using a 19-20% normalized collision energy) were used to map out the location of modification. The resulting mass lists were further assigned based on the protein sequence of Ub with or without the modification (GlyGly-AA, 171.2 amu) at each Cys using a 10 and 20 ppm tolerance for precursor and fragment ions, respectively. All reported M_r values are the most abundant molecular weights.



B

1 M Q I F V C T L T G K T I T L E V E P S D T I E N 50
26 V K A K I Q D K E G I P P D Q Q R L I F A G K Q L 25
51 E D G R T L S D Y N I Q K E S T L H L V L R L R 1

Total# c :3 z :36

1 M Q I F V C T L T G K T I T L E V E P S D T I E N 50
26 V K A K I Q D K E G I P P D Q Q R L I F A G K Q L 25
51 E D G R T L S D Y N I Q K E S T L H L V L R L R 1

Total#c :42 z : 36
Modified cysteine at C-6 with Mw of 171.2 Da

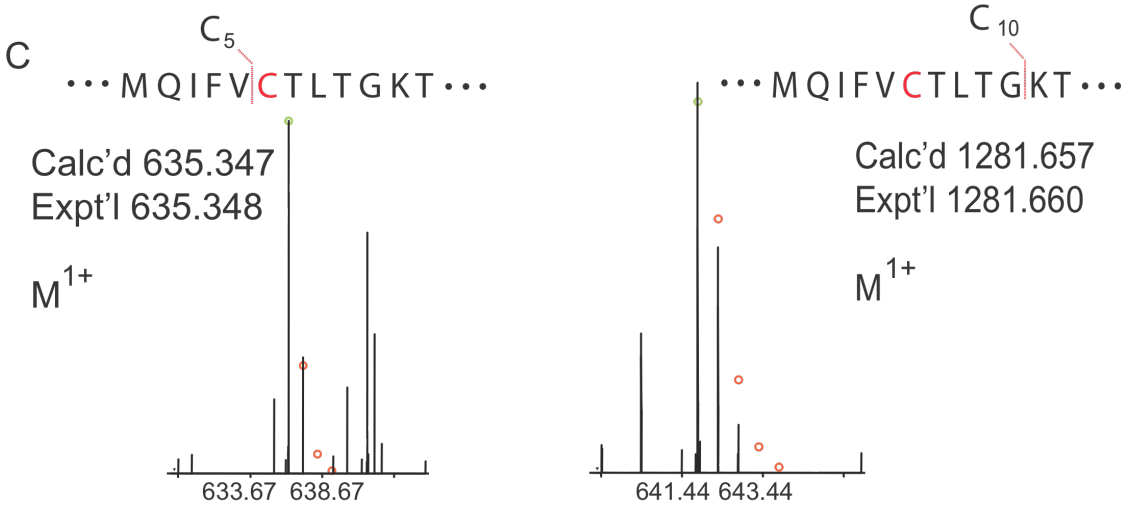
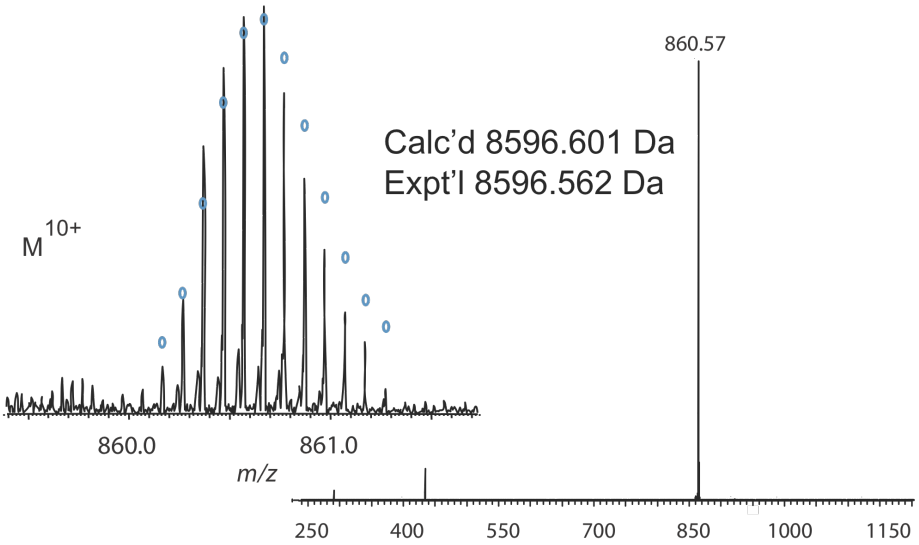


Figure 3.8: ECD analysis of K6C-linked tetramer. (A) K6C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*ε-Gly-L-homothiaLys thioether linker modification at C-6 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK6C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

A



B

1 M Q I F V C T L T G K T I T L E V E P S D T I E N 50
26 V K A K I Q D K E G I P P D Q Q R L I F A G K Q L 25
51 E D G R T L S D Y N I Q K E S T L H L V L R L R 1

Total# c :8 z :38

1 M Q I F V C T L T G K T I T L E V E P S D T I E N 50
26 V K A K I Q D K E G I P P D Q Q R L I F A G K Q L 25
51 E D G R T L S D Y N I Q K E S T L H L V L R L R 1

Total# c :32 z :39
Modified cysteine at C-6 with Mw of 171.2 Da

C

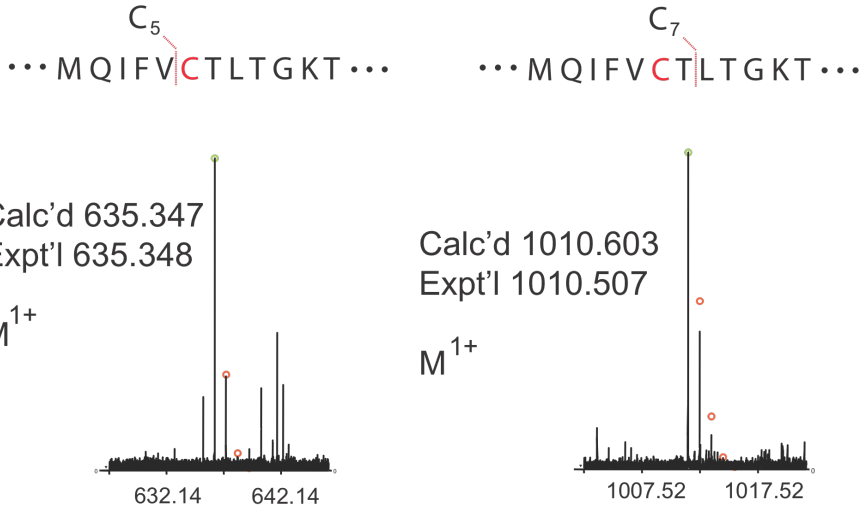
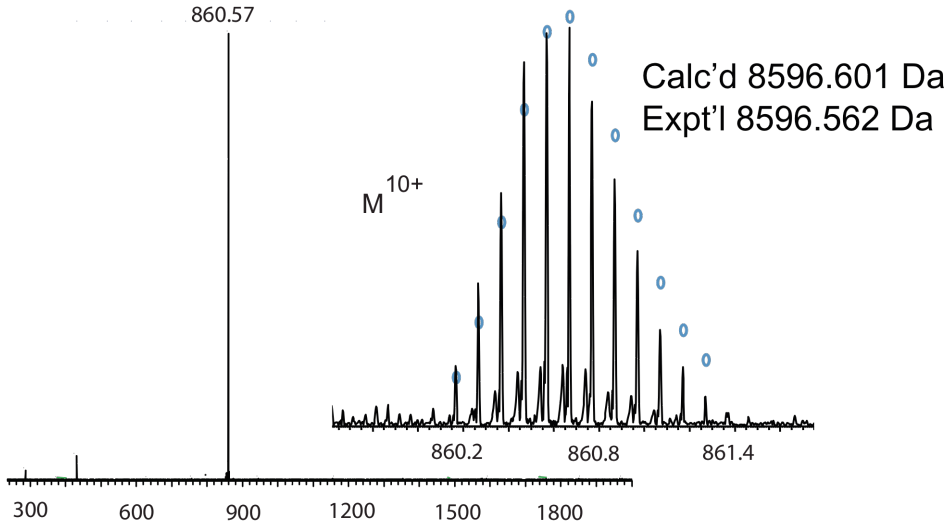
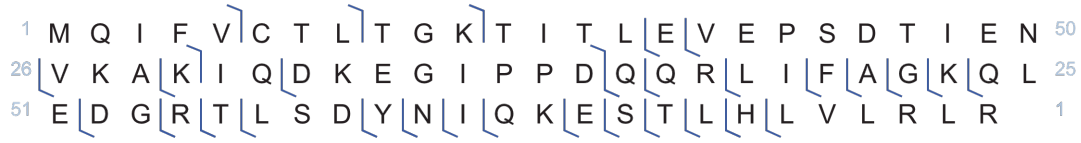


Figure 3.9: ECD analysis of K6C-linked pentamer. (A) K6C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*ε-Gly-L-homothiaLys thioether linker modification at C-6 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK6C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

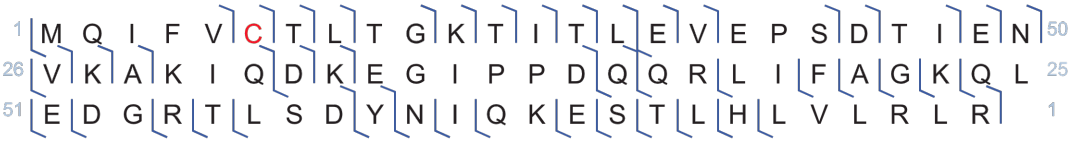
A



B



Total# c :6 z :28



Total# c :28 z :27

Modified cysteine at C-6 with Mw of 171.2 Da

C

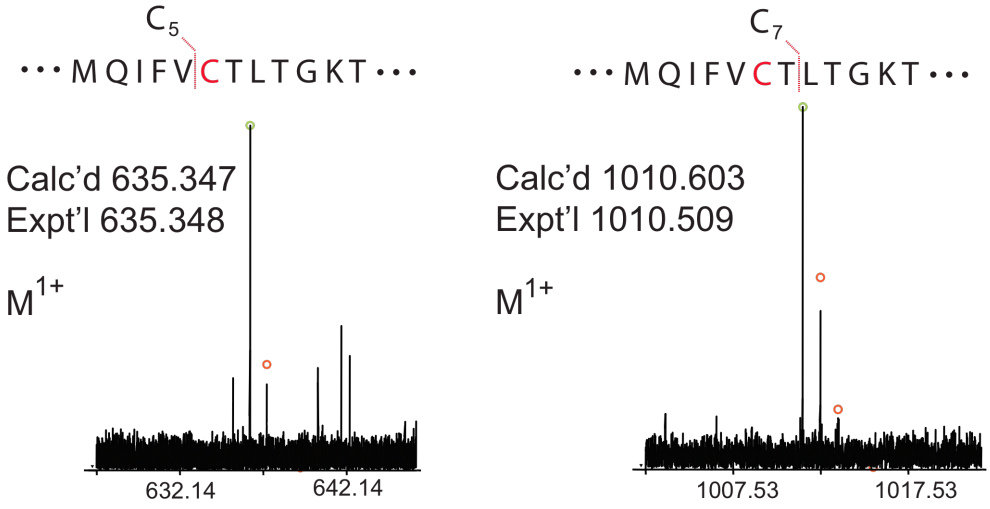
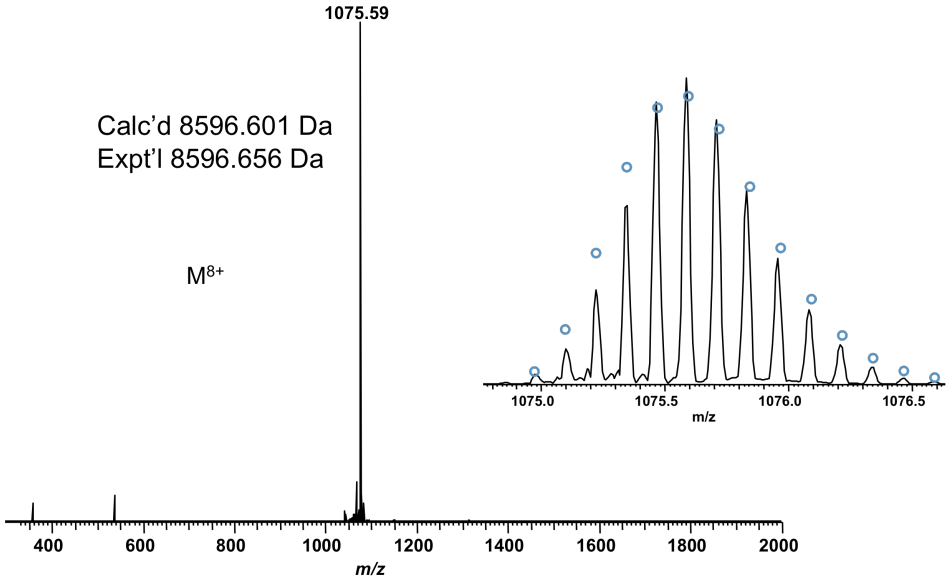
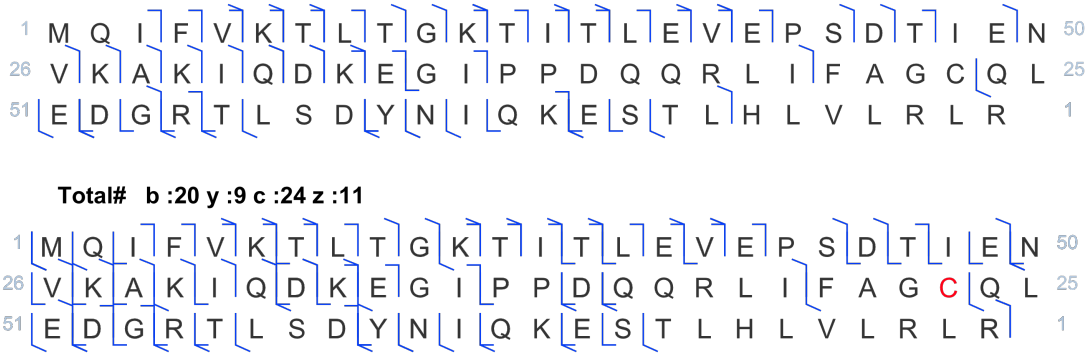


Figure 3.10: ECD analysis of K6C-linked hexamer. (A) K6C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{10+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*ε-Gly-L-homothiaLys thioether linker modification at C-6 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK6C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

A



B



Total# b :20 y :9 c :24 z :11

Total# b :21 y :27 c :25 z :22
Modified cysteine at C-48 with Mw of 171.2 Da

C

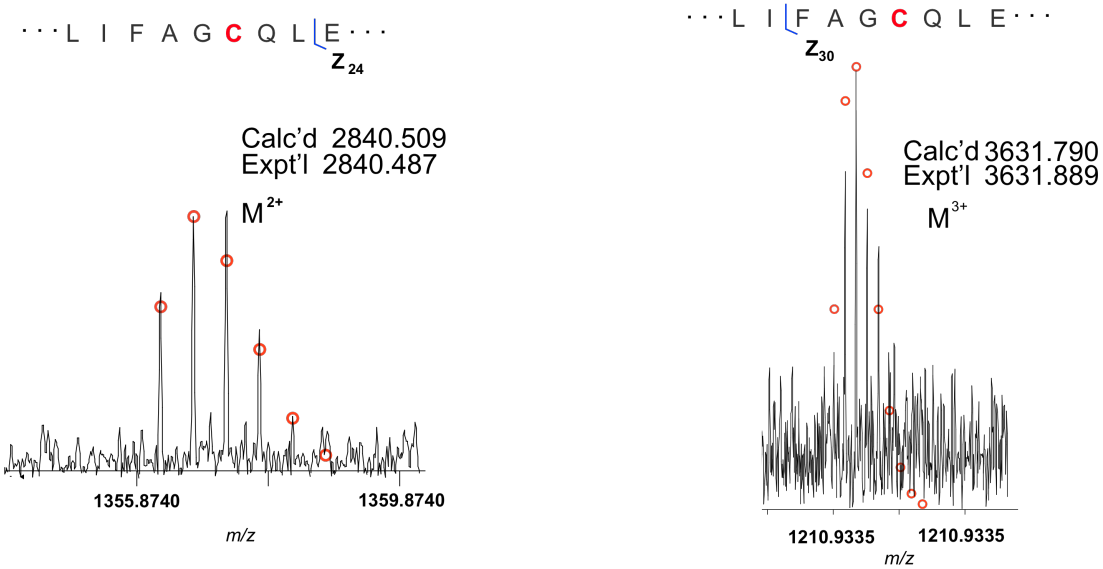
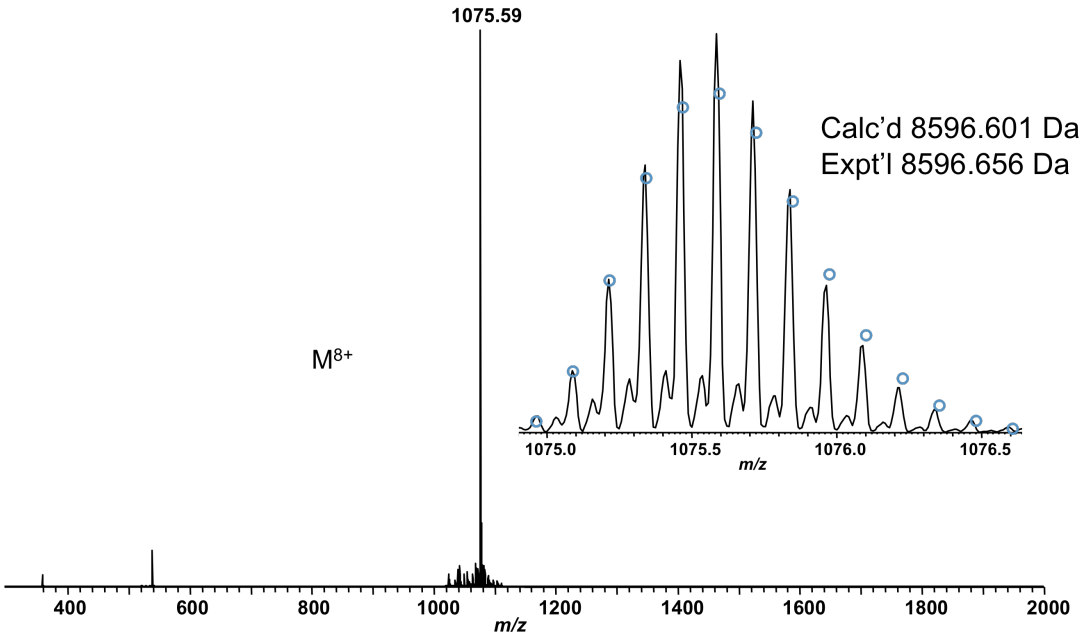
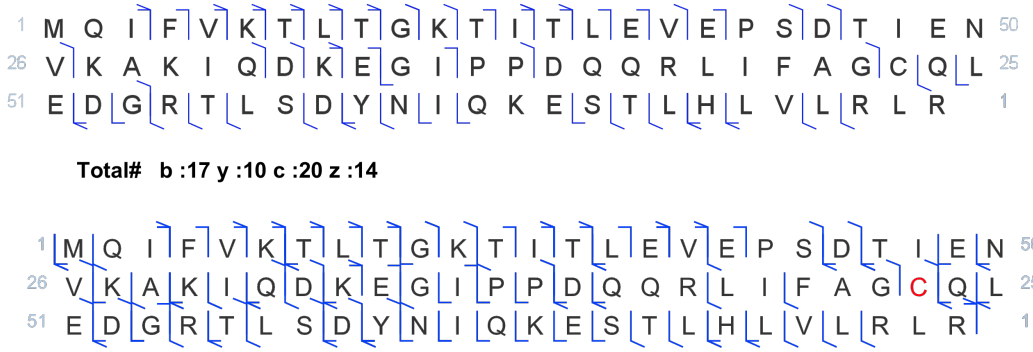


Figure 3.11: ECD and CAD analysis of K48C-linked tetramer. (A) K48C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{8+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*e-Gly-L-homothiaLys thioether linker modification at C-48 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK48C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

A



B



Total# b :17 y :10 c :20 z :14

Total# b :19 y :32 c :28 z :35
Modified cysteine at C-48 with Mw of 171.2 Da

C

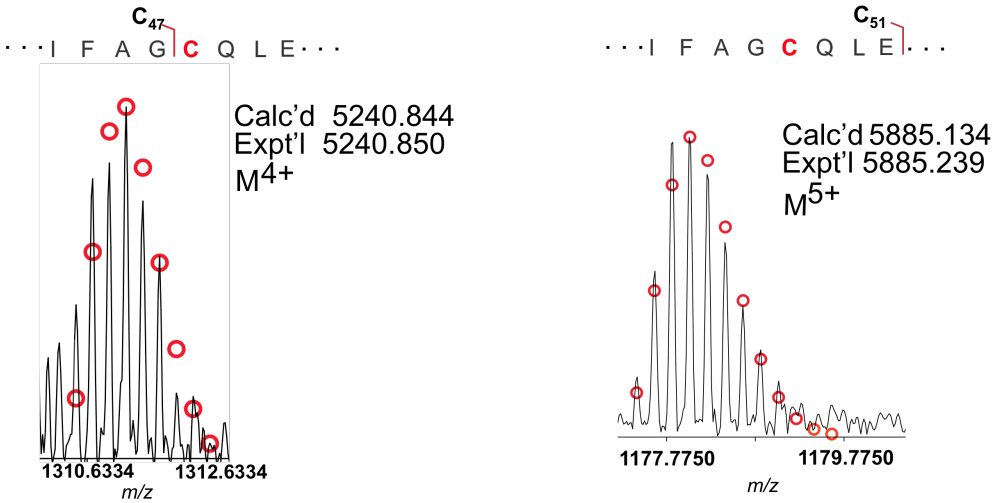


Figure 3.12: ECD and CAD analysis of K48C-linked pentamer. (A) K48C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{8+} charge state) with insert of isotopomers. (B) Map of observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*e-Gly-L-homothiaLys thioether linker modification at C-48 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK48C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

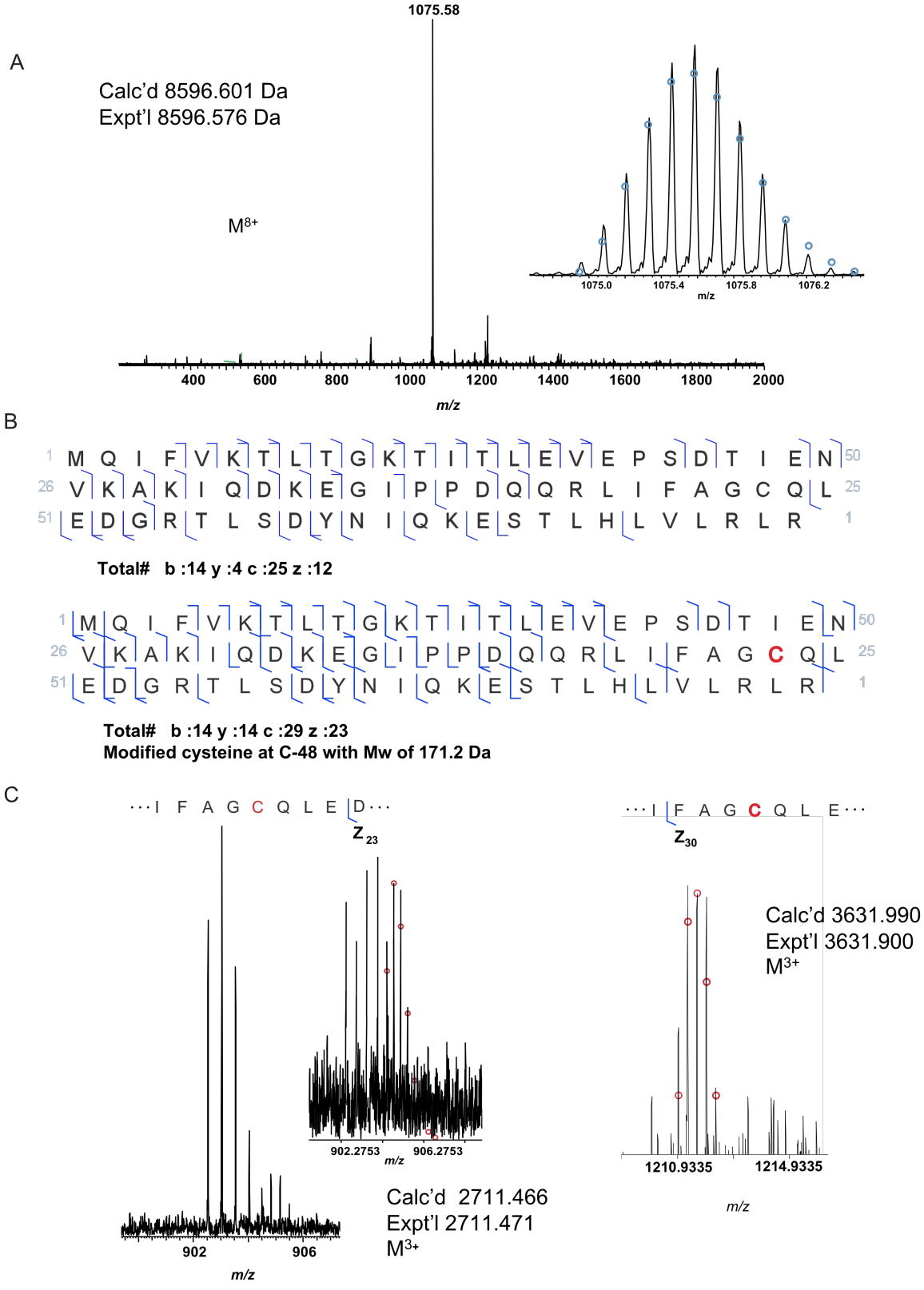


Figure 3.13: ECD and CAD analysis of K48C-linked hexamer. (A) K48C-linked Ub_{1-74} GlyGly-AA parent ion isolation (M^{8+} charge state) with insert of isotopomers. (B) Map of

observed fragments. Data analysis for the map on top does not include thioether linker. Bottom map includes *N*ε-Gly-L-homothiaLys thioether linker modification at C-48 (red) with a modified mass value of 171.2 Da. (C) Key ECD fragment ions for mapping thioether linkage site on UbK48C. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant MW. Expt'l: experimental most abundant MW.

3.4 References

[1] C. Grabbe, K. Husnjak, I. Dikic, *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 295–307.

[2] A. Hershko, A. Ciechanover, *Annu. Rev. Biochem.* **1998**, *67*, 425–479.

[3] J. S. Thrower, L. Hoffman, M. Rechsteiner, C. M. Pickart, *EMBO J.* **2000**, *19*, 94–102.

[4] V. Chau, J. W. Tobias, A. Bachmair, D. Marriott, D. J. Ecker, D. K. Gonda, A. Varshavsky, *Science* **1989**, *243*, 1576–1583.

[5] T. Kirisako, K. Kamei, S. Murata, M. Kato, H. Fukumoto, M. Kanie, S. Sano, F. Tokunaga, K. Tanaka, K. Iwai, *EMBO J.* **2006**, *25*, 4877–4887.

[6] D. S. Kirkpatrick, N. A. Hathaway, J. Hanna, S. Elsasser, J. Rush, D. Finley, R. W. King, S. P. Gygi, *Nat. Cell Biol.* **2006**, *8*, 700–710.

[7] P. Xu, D. M. Duong, N. T. Seyfried, D. Cheng, Y. Xie, J. Robert, J. Rush, M. Hochstrasser, D. Finley, J. Peng, *Cell* **2009**, *137*, 133–145.

[8] E. B. Dammer, C. H. Na, P. Xu, N. T. Seyfried, D. M. Duong, D. Cheng, M. Gearing, H. Rees, J. J. Lah, A. I. Levey, et al., *J. Biol. Chem.* **2011**, *286*, 10457–10465.

[9] Y. Saeki, T. Kudo, T. Sone, Y. Kikuchi, H. Yokosawa, A. Toh-e, K. Tanaka, *EMBO J.* **2009**, *28*, 359–371.

[10] F. Wu-Baer, *J. of Biol. Chem.* **2003**, *278*, 34743–34746.

[11] H. Nishikawa, S. Ooka, K. Sato, K. Arima, J. Okamoto, R. E. Klevit, M. Fukuda, T. Ohta, *J. Biol. Chem.* **2004**, *279*, 3916–3924.

[12] M. S. Y. Huen, S. M.-H. Sy, J. Chen, *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 138–148.

[13] M. L. Li, R. A. Greenberg, *Trends Biochem. Sci.* **2012**.

[14] L. Bedford, R. Layfield, R. J. Mayer, J. Peng, P. Xu, *Neurosci. Lett.* **2011**, *491*, 44–47.

[15] A. D. Jacobson, N. Y. Zhang, P. Xu, K. J. Han, S. Noone, J. Peng, C. W. Liu, *J. of Biol. Chem.* **2009**, *284*, 35485–35494.

[16] K. Sato, R. Hayami, W. Wu, T. Nishikawa, H. Nishikawa, Y. Okuda, H. Ogata, M. Fukuda, T. Ohta, *J. Biol. Chem.* **2004**, *279*, 30919–30922.

[17] X. Yu, S. Fu, M. Lai, R. Baer, J. Chen, *Genes Dev.* **2006**, *20*, 1721–1726.

[18] R. Shakya, L. J. Reid, C. R. Reczek, F. Cole, D. Egli, C.-S. Lin, D. G. deRooij, S. Hirsch, K. Ravi, J. B. Hicks, et al., *Science* **2011**, *334*, 525–528.

[19] J. Piotrowski, R. Beal, L. Hoffman, K. D. Wilkinson, R. E. Cohen, C. M. Pickart, *J. Biol. Chem.* **1997**, *272*, 23712–23721.

[20] R. M. Hofmann, C. M. Pickart, *Cell* **1999**, *96*, 645–653.

[21] C. M. Pickart, S. Raasi, *Methods Enzymol.* **2005**, *399*, 21–36.

[22] A. Bremm, S. M. V. Freund, D. Komander, *Nat. Struct. Mol. Biol.* **2010**, *17*, 939–947.

[23] K. C. Dong, E. Helgason, C. Yu, L. Phu, D. P. Arnott, I. Bosanac, D. M. Compaan, O. W. Huang, A. V. Fedorova, D. S. Kirkpatrick, et al., *Structure* **2011**, *19*, 1053–1063.

[24] E. R. Strieter, D. A. Korasick, *ACS Chem. Biol.* **2012**, *7*, 52–63.

[25] L. Spasser, A. Brik, *Angew. Chem., Int. Ed.* **2012**, *51*, 6840–6862.

[26] K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 6137–6141.

[27] S. N. Bavikar, L. Spasser, M. Haj-Yahya, S. V. Karthikeyan, T. Moyal, K. S. Ajish Kumar, A. Brik, *Angew. Chem., Int. Ed.* **2011**, *51*, 758–763.

[28] C. Castañeda, J. Liu, A. Chaturvedi, U. Nowicka, T. A. Cropp, D. Fushman, *J. Am. Chem. Soc.* **2011**, *133*, 17855–17868.

[29] E. M. Valkevich, R. G. Guenette, N. A. Sanchez, Y.-C. Chen, Y. Ge, E. R. Strieter, *J. Am. Chem. Soc.* **2012**, *134*, 6916–6919.

[30] C. M. Pickart, *Annu. Rev. Biochem.* **2001**, *70*, 503–533.

[31] R. J. Deshaies, C. A. Joazeiro, *Annu. Rev. Biochem.* **2009**, *78*, 399–434.

[32] Z. Chen, C. M. Pickart, *J Biol Chem* **1990**, *265*, 21835–21842.

[33] M. C. Rodrigo-Brenni, S. A. Foster, D. O. Morgan, *Mol. Cell* **2010**, *39*, 548–559.

[34] K. E. Wickliffe, S. Lorenz, D. E. Wemmer, J. Kuriyan, M. Rape, *Cell* **2011**, *144*, 769–781.

[35] A. Saha, S. Lewis, G. Kleiger, B. Kuhlman, R. J. Deshaies, *Mol. Cell* **2011**, *42*, 75–83.

[36] C. E. Hoyle, C. N. Bowman, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 1540–1573.

[37] T. M. Lovestead, K. A. Berchtold, C. N. Bowman, *Macromolecules* **2005**, *38*, 6374–6381.

[38] S. K. Reddy, N. B. Cramer, C. N. Bowman, *Macromolecules* **2006**, *39*, 3673–3680.

[39] C. N. Bowman, C. J. Kloxin, *AIChE J.* **2008**, *54*, 2775–2795.

[40] T. Yao, R. E. Cohen, *J. Biol. Chem.* **2000**, *275*, 36862–36868.

[41] K. D. Wilkinson, V. L. Tashayev, L. B. O'Connor, C. N. Larsen, E. Kasperek, C. M. Pickart, *Biochemistry* **1995**, *34*, 14535–14546.

[42] F. E. Reyes-Turcu, J. R. Horton, J. E. Mullally, A. Heroux, X. Cheng, K. D. Wilkinson, *Cell* **2006**, *124*, 1197–1208.

[43] K. D. Wilkinson, M. J. Cox, A. N. Mayer, T. Frey, *Biochemistry* **1986**, *25*, 6644–6649.

[44] P. Xu, J. Peng, *Anal. Chem.* **2008**, *80*, 3438–3444.

[45] M. J. Edelmann, A. Iphöfer, M. Akutsu, M. Altun, K. di Gleria, H. B. Kramer, E. Fiebiger, S. Dhe-Paganon, B. M. Kessler, *Biochem. J.* **2009**, *418*, 379–390.

[46] A. C. Faesen, M. P. A. Luna-Vargas, P. P. Geurink, M. Clerici, R. Merkx, W. J. van Dijk, D. S. Hameed, F. El Oualid, H. Ovaa, T. K. Sixma, *Chem. Biol.* **2011**, *18*, 1550–1561.

[47] M. J. Lee, B.-H. Lee, J. Hanna, R. W. King, D. Finley, *Mol. Cell. Proteomics* **2011**, *10*, R110.003871.

[48] R. M. Hofmann, C. M. Pickart, *J. Biol. Chem.* **2001**, *276*, 27936–27943.

[49] Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 51.

[50] Pickart, C. M.; Raasi, S. *Meth. Enzymol.* **2005**, *399*, 21.

[51] Valkevich, E.M.; Guenette, R. G.; Sanchez, N.A.; Chen, Y.; Ge, Y.; Strieter, E.R. *J. Am. Chem. Soc.*, **2012**, 134, 6916- 6919

[52] Strolhalm, M.; Kavan, D.; Novák, P.; Halíček, V. *Anal Chem*, **2010**, *11*, 4648-4651.

[53] Strolhalm, M.; Hassman, M.; Košata, B.; Kodíček, M. **2008**, *6*, 905-908.

[54] Majima, T.; Schnabel, W.; Weber, W. *Makromol. Chem.* **1991**, *192*, 2307.

[55] Fairbanks, B. D.; Schwartz, M. P.; Bowman, C. N.; Anseth, K. S. *Biomaterials* **2009**, *30*, 6702

[56] (a) Ayaz-Guner, S.; Zhang, J.; Li, L.; Walker, J. W.; Ge, Y. *Biochemistry* **2009**, *48*, 8161. (b) Ge, Y.; Rybakova, I. N.; Xu, Q.; Moss, R. L. *Proc. Natl. Acad. Sci. U S A* **2009**, *106*, 12658. (c) Zhang, J.; Guy, M. J.; Norman, H. S.; Chen, Y. C.; Xu, Q. G.; Dong, X. T.; Guner, H.; Wang, S. J.; Kohmoto, T.; Young, K. H.; Moss, R. L.; Ge, Y. *J. Proteome Res.* **2011**, *10*, 4054.

Chapter 4: Middle-Down Mass Spectrometry Enables Characterization of Branched Ubiquitin Chains

Portions of this work have been published in:

Valkevich, E. M., Sanchez, N. A., Ge, Y., and Strieter, E. R., (2014) Middle-Down Mass Spectrometry Enables Characterization of Branched Ubiquitin Chains, *Biochemistry.* 53, 4979-4989

4.1 ABSTRACT

Protein ubiquitylation, one of the most prevalent post-translational modifications in eukaryotes, is involved in regulating nearly every cellular signaling pathway. The vast functional range of ubiquitylation has largely been attributed to the formation of a diverse array of polymeric ubiquitin (polyUb) chains. Methods that enable characterization of these diverse chains are necessary to fully understand how differences in structure relate to function. Here, we describe a method for the detection of enzymatically-derived branched polyUb conjugates in which a single Ub subunit is modified by two Ub molecules at distinct lysine residues. Using a middle-down mass spectrometry approach in which restricted trypsin-mediated digestion is coupled with mass spectrometric analysis, we characterize the polyUb chains produced by bacterial effector E3 ligases NleL (non-Lee-encoded effector ligase from enterohemorrhagic *E. coli* O157:H7) and IpaH9.8 (from *Shigella flexneri*). Since Ub is largely intact after minimal trypsinolysis, multiple modifications on a single Ub moiety can be detected. Analysis of NleL- and IpaH9.8-derived polyUb chains reveals branch points are present in approximately 10 % of the overall chain population. When unanchored, well-defined polyUb chains are added to reactions containing NleL, longer chains are more likely to be modified internally, forming branch points rather than extending from the end of the chain. These results suggest that middle-down mass spectrometry can be used to assess the extent to which branched polyUb chains are formed by various enzymatic systems and potentially evaluate the presence of these atypical conjugates in cell and tissue extracts.

4.2 INTRODUCTION

Through covalent attachment to intracellular proteins – a process termed protein ubiquitylation – the small protein ubiquitin (Ub) regulates a wide range of biological processes.[1-6] Ub is tethered to proteins via the action of three proteins: E1 Ub activating enzymes, E2 Ub conjugating enzymes, and E3 Ub ligating enzymes (Figure 4.1A).[7, 8] This enzymatic cascade results in the formation of an isopeptide linkage between the Ub C-terminus and an ε-amino group of a substrate lysine residue. Analogous to protein glycosylation, once Ub is anchored to a substrate protein additional rounds of conjugation afford polymeric Ub (polyUb) chains. A variety of chains can be assembled due to the presence of seven Ub lysine residues (K6, K11, K27, K29, K33, K48, and K63) and an amino terminus (M1) (Figure 4.1B). The prevailing view is that distinct polyUb chains govern specific biological pathways.[5] This supposition is based, in large part, on the disparate activities of single-linkage (homotypic) K48- and K63-linked polyUb chains. K48-linked Ub chains with a minimum of four subunits endow cells with the ability to remove intracellular proteins by acting as the principal signal for proteasomal degradation.[9-11] By contrast, K63-linked polyUb chains serve as non-degradative signals during the DNA damage response and cytokine signaling.[12, 13] These functional discrepancies have lead to efforts to uncover the roles of other polyUb chains, particularly since proteomics studies have revealed all eight Ub-Ub linkages are present in eukaryotic cells.[14] While it has become evident non-K48 and K63 linkages increase in abundance under certain cellular conditions (e.g., M1, K6, and K11),[15-17] the precise function of these conjugates is unclear. Moreover, the exact topology of chains (e.g., whether a chain is linear or branched) that govern a particular cellular pathway has been a mystery.

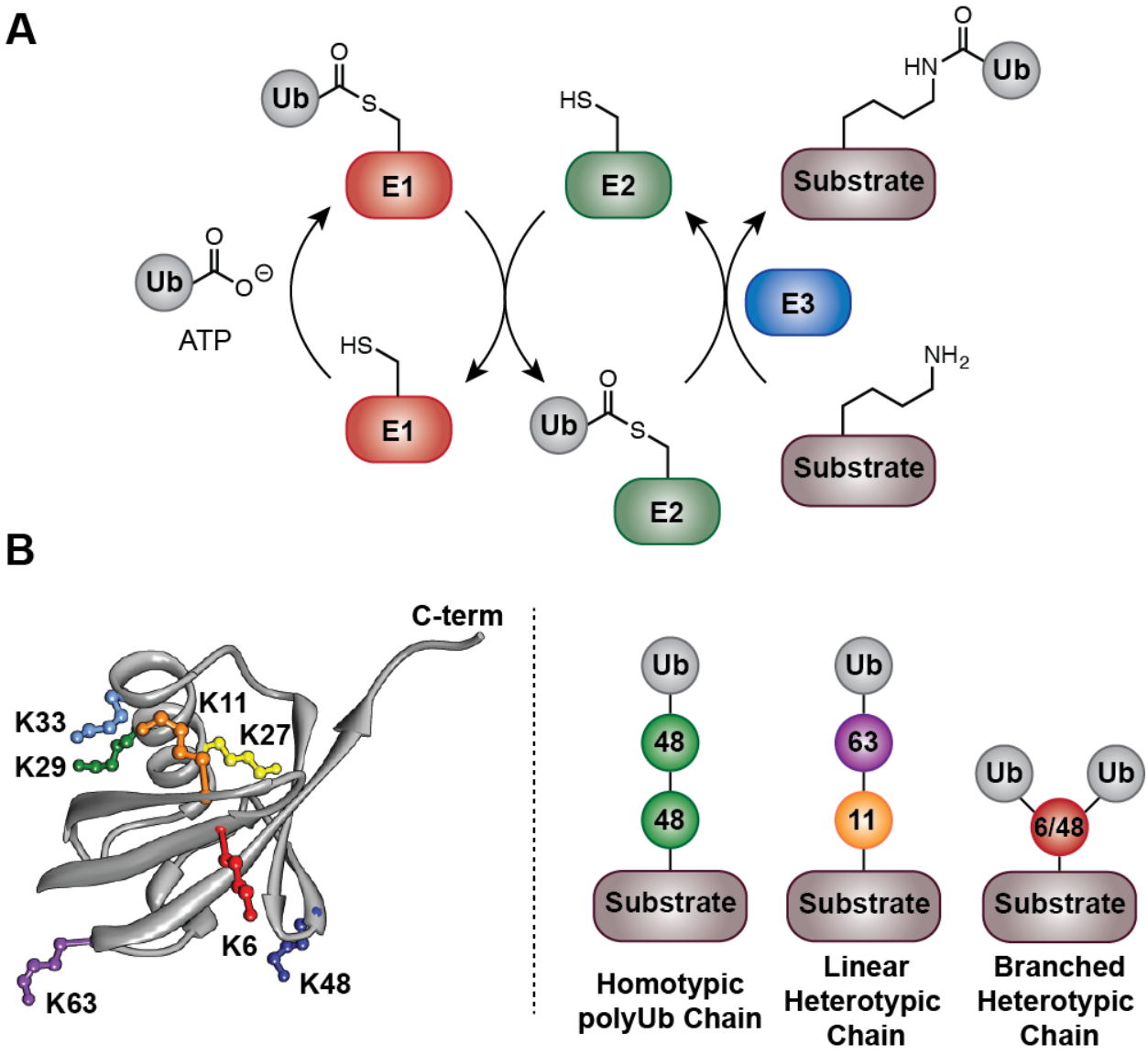


Figure 4.1. Protein ubiquitylation. **A.** Cascade of E1, E2, and E3 enzymes catalyze the formation of an isopeptide bond between a substrate protein and ubiquitin (Ub). **B.** Structure of Ub (1ubq.pdb) showing the seven lysines (K6, K11, K27, K29, K33, K48, and K63) and the types of polymeric Ub (polyUb) chains that form due to the presence of these residues. The lines between Ub subunits in the chains denote an isopeptide linkage and the numbers indicate the lysine used to link subunits together. In the case of a branched heterotypic chain, a single subunit is modified with two or more Ub molecules via two or more isopeptide linkages.

Connecting chain topology to a particular biological function presents a significant analytical challenge. The three most common methods for analyzing polyUb chains include

expression of Ub lysine-to-arginine (K-to-R) variants in different cell lines, linkage-specific antibodies, and mass spectrometry (MS).[18] Use of Ub K-to-R variants blocks chain extension through specific sites and therefore prevents downstream events associated with a particular polyUb chain. The implementation of this approach has led to seminal discoveries related to K48 and K63 linkages.[10, 12] However, Ub K-to-R variants are overexpressed in living cells and the presence of wild type Ub often complicates analysis. Furthermore, definitive conclusions cannot be drawn with regards to the formation of linear or branched heterotypic chains, and particular K-to-R substitutions may affect the surface of Ub and bias polyUb assembly. Although linkage-specific antibodies (four of which are available: M1, K11, K48, and K63)[15, 17, 19] provide an opportunity to explore the impact of chain linkage on the fate of endogenous polyUb-modified proteins, it is difficult to differentiate between homotypic and heterotypic chains.[19, 20]

MS approaches have also been instrumental in characterizing polyUb chains. The most common strategy is to use a bottom-up approach in which Ub conjugates are first digested extensively by trypsin then analyzed by MS.[21] Modified lysine residues of a particular protein, including Ub itself, are then identified by the presence of a diglycine (GG) motif appended to the ε-amino group (trypsin cleaves between R74 and G75 of Ub leaving behind the GG motif). While bottom-up MS enables characterization of the linkages between two Ub molecules, it is difficult to assess chain length and topology because connectivity of Ub fragments harboring branch points is destroyed during digestion.[22] To address this, there has been interest in using other approaches that examine intact proteins such as middle-down MS.[23-25] Middle-down MS combines the benefits of both bottom-up and top-down approaches by exploiting minimal protease digestion of protein samples and

the ability to detect multiple post-translational modifications on a single polypeptide chain.[26-28] However, this approach has not been applied to the analysis of Ub chain topology.

Here we demonstrate the utility of middle-down MS in the characterization of branched polyUb chains produced by bacterial E3 ligases. Several pathogenic bacteria deliver proteins, termed effectors, into host cells to undermine the defense response.[29] Since the ubiquitylation network plays a major role in the immune response, it is one of the primary targets of these effector proteins. Many effectors functionally mimic eukaryotic E3 ligases and catalyze the assembly of polyUb chains on a distinct subset of host proteins.[30] One particular example is the E3 ligase NleL (non-Lee-encoded effector ligase) from enterohemorrhagic *E. coli* (EHEC) O157:H7.[31, 32] NleL is important for modulating the actin cytoskeleton of the host cell, and has recently been shown to build heterotypic polyUb chains bearing K6 and K48 linkages *in vitro*. [31, 33] Because of the challenges associated with the assessment of chain topology, however, the extent to which NleL constructs polyUb chains bearing branch points is unknown. Therefore, we sought to employ middle-down MS in the characterization of polyUb chains produced by NleL as well as another bacterial E3 ligase from *Shigella flexneri*, IpaH9.8.[34] The results of our investigation suggest middle-down MS can be used to evaluate polyUb chain branching and dissect the factors that contribute to the formation of this underexplored chain topology.

4.3 MATERIALS AND METHODS

4.3.1 Protein Expression and Purification. Wild type ubiquitin (Ub) and lysine-to-cysteine (K-to-C) Ub variants were expressed in *E. coli* RosettaTM 2(DE3)pLysS cells

(Novagen) and purified by perchloric acid precipitation, following a procedure adapted from Pickart 2005.[35] DNA encoding the human E1 Ub-activating enzyme was amplified from the HeLa cell cDNA library and cloned into pET24a(+). The UbcH5c(Ube2D3)-pET14a DNA construct was purchased from Addgene. The catalytic domain of IpaH9.8$_{254\text{-}545}$ was cloned into pET28a(+). Human E1, UBE2D3, and IpaH9.8 (*Shigella flexneri*) with 6x-histidine (His) tags were expressed and purified as previously described.[36] A bacterial expression vector encoding glutathione S-transferase (GST)-tagged NleL$_{170–782}$ was a gift from D.Y.W. Lin and J. Chen. Briefly, GST-tagged NleL was expressed in *E. coli* BL-21 cells grown in LB medium (OD_{600} of 0.6) at 37 ^{0}C, induced with IPTG (0.1 mM) and grown at 16 ^{0}C (16 h). GST-NleL was then purified by glutathione sepharose affinity chromatography. The GST tag was cleaved from the eluted protein with TEV protease (4 °C, 16 hours) and further purified by gel filtration (Superdex 200, GE Healthcare). The gene construct for UbcH7 (UBE2L3) was purchased from DNASU Plasmid Repository and cloned into pGEX-4-T2 bacterial expression vector with an N-terminal GST tag (BamHI and XhoI restriction sites). Cells were grown in LB medium at 37 ^{0}C (OD_{600} of 0.6), induced using IPTG (0.4 mM), and grown for 4 h. As with GST-tagged NleL, GST-tagged UBE2L3 was purified by glutathione sepharose affinity chromatography. The GST tag was again cleaved from the eluted protein with thrombin protease (4 °C, 16 h, minimal N-terminal perturbation is imperative for chain synthesis activity with NleL) and the protein was further purified by cation exchange chromatography.

4.3.2 Thiol-Ene Ub Chain Synthesis. Homotypic Ub chains linked via non-native isopeptide bonds at positions-6 or 48 were synthesized using thiol-ene coupling (TEC) chemistry as previously described.[37, 38]

4.3.3 Native PolyUb Chain Synthesis. To solutions containing reaction buffer A (50 mM Tris-HCl, pH 7.4, 50 mM NaCl, 5 mM $MgCl_2$, 0.1 mM DTT) was added Ub (50 μM), E1 (150 nM), E2 (1 μM of either UBE2D3 or UBE2L3), and E3 (0.05-5 μM of either NleL or IpaH9.8). Reactions were then initiated using ATP (2 mM) and allowed to proceed at 37 ^{0}C. PolyUb chain formation was analyzed by SDS-PAGE and MS (see below).

4.3.4 Chain Elongation Using Thiol-Ene Derived Ub Substrates. To each reaction mixture was added Ub oligomers derived from thiol-ene chemistry (50 μM), Ub (25 μM), E1 (150 nM), UBE2D3 (1 μM), and NleL (5 μM) in buffer A. ATP (2 mM) was then added, and polymerization was allowed to occur for 3 h at 37 °C. It is important to note the concentrations of Ub dimers and tetramers are based on the molecular weight of each chain. By contrast, the concentrations of heterogeneous mixtures of Ub oligomers were measured based on the molecular weight of a single Ub molecule.

4.3.5 Minimal Trypsin Digestion of Ub Chains. After chain synthesis, reactions were concentrated and exchanged into water using Amicon spin filters [0.5mL with 3.5 kiloDalton (kD) molecular weight cutoff (MWCO)]. The enzyme/chain mixture (30 μL or half of the total reaction mixture) was digested with trypsin (0.5 μg; Cal Biochem MS grade) in ammonium bicarbonate buffer at 37 °C for 6 h. Trypsin was deactivated with 10% acetic acid, and the resulting mixtures were dialyzed into water (Slide-A-lyzer® MINI dialysis units, 3.5 kD MWCO) to remove small peptides arising from conjugating enzymes.

4.3.6 Middle-Down Mass Spectrometry Analysis. Samples were dissolved in a solution of water/acetonitrile/acetic acid (45:45:10) and injected into a 7T linear ion trap/Fourier transform ion cyclotron resonance (LTQ/FT-ICR) hybrid mass spectrometer (Thermo Scientific Inc., Bremen, Germany) equipped with an automated chip-based nanoESI source

(Triversa NanoMate, Advion BioSciences, Ithaca, NY) as described previously.[39-41] The resolving power of the FT-ICR mass analyzer was set at 100,000. All FT-ICR spectra were processed with in-house software (MASH Suite[42]) using a signal to noise threshold of 3 and fit factor of 60%, and then validated manually.

4.3.7 Electron Capture Dissociation (ECD) Analysis of Ub Chain Linkages. For tandem mass spectrometry (MS/MS) experiments using ECD, individual charge states of protein molecular ions were first isolated. Then, the ions were dissociated by ECD using 3-4% "electron energy" and a 70 ms duration with a 65 μs delay. All FT-ICR spectra were processed with the MASH software suite using a signal to noise threshold of 3 and fit factor of 60%, and then validated manually. The resulting mass lists were further assigned based on the protein sequence of Ub with or without the di-glycine (GG) modification at each lysine residue using a 10 and 20 ppm tolerance for precursor and fragment ions, respectively. All reported calculated (Calc'd) and experimental (Expt'l) values correspond to the most abundant molecular weights.

4.3.8 Bottom up Mass Spectroscopy Analyses of Ubiquitin Linkages Formed by Bacterial E3 Ligase IpaH9.8

"In gel" digestion and mass spectrometric analysis was performed at the Mass Spectrometry Facility [Biotechnology Center, University of Wisconsin-Madison]. The digestion was performed as outlined on the website: http://www.biotech.wisc.edu/ServicesResearch/MassSpec/ingel.htm. Ub oligomer digests were purified using OMIX C18 SPE cartridges (Agilent, Palo Alto, CA) according to the manufacturer's protocol and eluted in 25μl of 60/40/0.1% ACN/H2O/TFA, dried to completion in the speed-vac, and finally reconstituted in 20μl of 0.05% TFA in H_2O.

Peptides were analyzed by nanoLC-MS/MS using an Agilent 1100 nanoflow system (Agilent, Palo Alto, CA) connected to a hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, Bremen, Germany) equipped with a nanoelectrospray ion source. Chromatography of peptides was accomplished prior to mass spectral analysis using a capillary emitter column (in-house packed with MAGIC C18, 3 μM, 150x0.075mm, Michrom Bioresources, Inc.) onto which 4μl of extracted peptides was automatically loaded. The NanoHPLC system delivered solvents A: 0.1% (v/v) formic acid in water, and B: 95% (v/v) acetonitrile, 0.1% (v/v) formic acid at either 0.5 μL/min, to load sample, or 0.20 μL/min, to elute peptides directly into the nano-electrospray over a 60 minutes 1% (v/v) B to 60% (v/v) B followed by 10 minute 60% (v/v) B to 100% (v/v) B gradient. As peptides eluted from the HPLC-column/electrospray source survey MS scans were acquired in the Orbitrap with a resolution of 100,000 and up to 5 most intense peptides per scan were fragmented and detected in the ion trap over the 300 to 2000 m/z; redundancy was limited by dynamic exclusion. Raw MS/MS data were converted to Mascot generic format (mgf) using Trans Proteomic Pipeline (Seattle Proteome Center, Seattle, WA). Resulting mgf files were used to search against forward/reverse Uniprot human database concatenated with a list of common contaminants (70,133 entries). Methionine oxidation, asparagine/glutamine deamidation, lysine ubiquitination, and cysteine modified with BFA were selected as variable modifications. Peptide mass tolerance was set at 20 ppm and fragment mass at 0.8 Da. Due to low number of protein entries for statistical determination of significance in peptide matches, all relevant hits were manually reviewed. Furthermore all of the predicted ubiquitinated peptides were manually investigated to confirm proper residue assignment. Additional level of confidence in identification and

spectral annotation was done with help of Scaffold software (version 4.0.4, Proteome Software Inc., Portland, OR).

4.4 RESULTS

Middle-Down Analysis of Branched PolyUb Chains. Under conditions in which Ub is completely folded, trypsin only cleaves the peptide bond between arginine-74 and glycine-75, thereby releasing the C-terminal diglycine (GG) motif.[43] By exploiting the minimal trypsinolysis of Ub, we reasoned middle-down MS could be employed to unambiguously characterize the presence of branching within polyUb chains. The general concept is that minimal trypsinolysis furnishes Ub_{1-74} monomers still harboring a GG motif at any lysine previously engaged in the formation of a chain. Since chain branching involves the addition of multiple Ub molecules to a single Ub subunit, minimal trypsinolysis should then afford a Ub_{1-74} derivative harboring two or more GG motifs tethered to specific lysine residues. If there is chain branching, we would expect to observe at least three distinct species by MS: Ub_{1-74}, arising from chain caps/unpolymerized Ub (Calc'd: 8450.57 Da); Ub_{1-74} with a single GG motif (GG-Ub_{1-74}, Calc'd: 8564.62 Da), resulting from the linear portion of the chain; and Ub_{1-74} with two GG motifs (2xGG-Ub_{1-74}, Calc'd: 8678.66 Da), originating from the branch point of the chain (Figure 4.2).

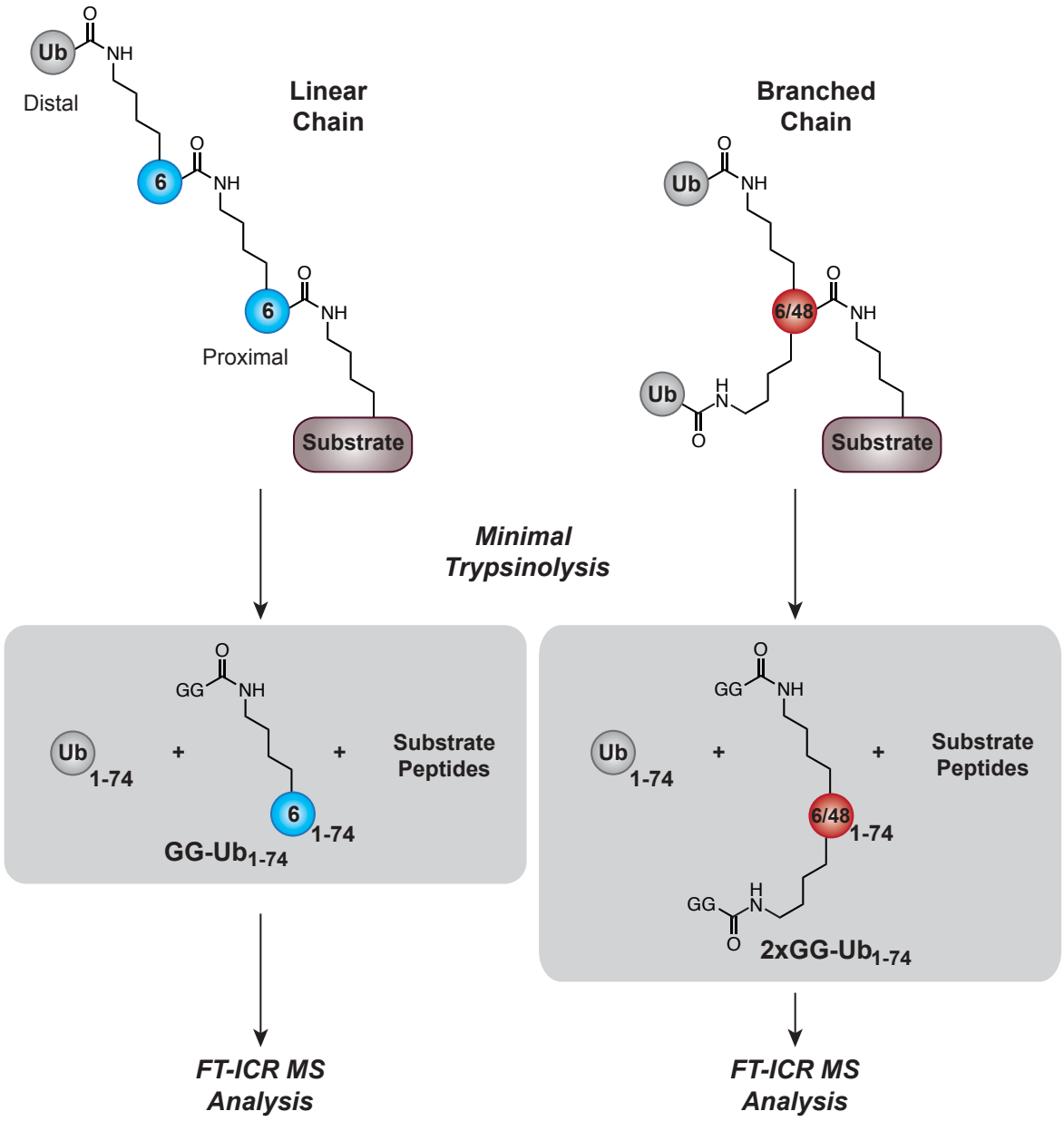


Figure 4.2. Minimal trypsinolysis of substrates modified with polyUb chains of different topology. What differentiates the formation of linear and branched chains is the presence of a Ub_{1-74} derivative harboring a single GG motif (linear; GG-Ub_{1-74}) or two GG motifs (branched; 2xGG-Ub_{1-74}).

To determine whether middle-down MS could be used to detect branched conjugates, we examined polyUb chain formation by the E3 ligase NleL from EHEC

O157:H7. NleL is composed of an N-terminal unstructured region (residues 1-169), a pentapeptide repeat domain (residues 170-370), and a catalytic HECT-like domain (HECT; homologous to E6AP C-terminus) with a conserved cysteine residue (C753) at the C-terminus. Previous studies with NleL have shown that residues 170 to 782 comprise a catalytically active E3 ligase capable of assembling both unanchored/free polyUb chains as well as chains anchored to itself through a process referred to as autoubiquitylation.[31, 33] In accordance, the addition of $NleL_{170-782}$ to reaction mixtures containing human E1, the E2 UBE2D3 or UBE2L3, ATP, and Ub, resulted in the formation of polyUb chains (Figure 4.3A). To characterize the topology of these chains, minimal trypsinolysis was performed in an ammonium bicarbonate buffer to maintain Ub in a native, folded state. Initial experiments focused on determining the amount of trypsin required to completely disassemble chains into Ub_{1-74} fragments. Using a trypsin concentration of 12.5 μg/mL and an enzyme-to-substrate ratio of approximately 1:1, chains were completely digested within 5 h as judged by SDS-PAGE (Figure 4.10). The resulting Ub_{1-74} derivatives were then analyzed using a FT-ICR MS. Although Ub species were detected, low molecular weight peptides from other proteins in the reaction mixture dominated the spectra. This was not entirely unexpected considering that, unlike peptides, the concentration of intact proteins is typically diluted by isotopic distributions and the presence of multiple charge states.[44] To eliminate signals arising from peptides, a post-digest dialysis step was incorporated into the workflow (Figure 4.11). This led to the detection of all three Ub_{1-74} species (Figure 4.3B).

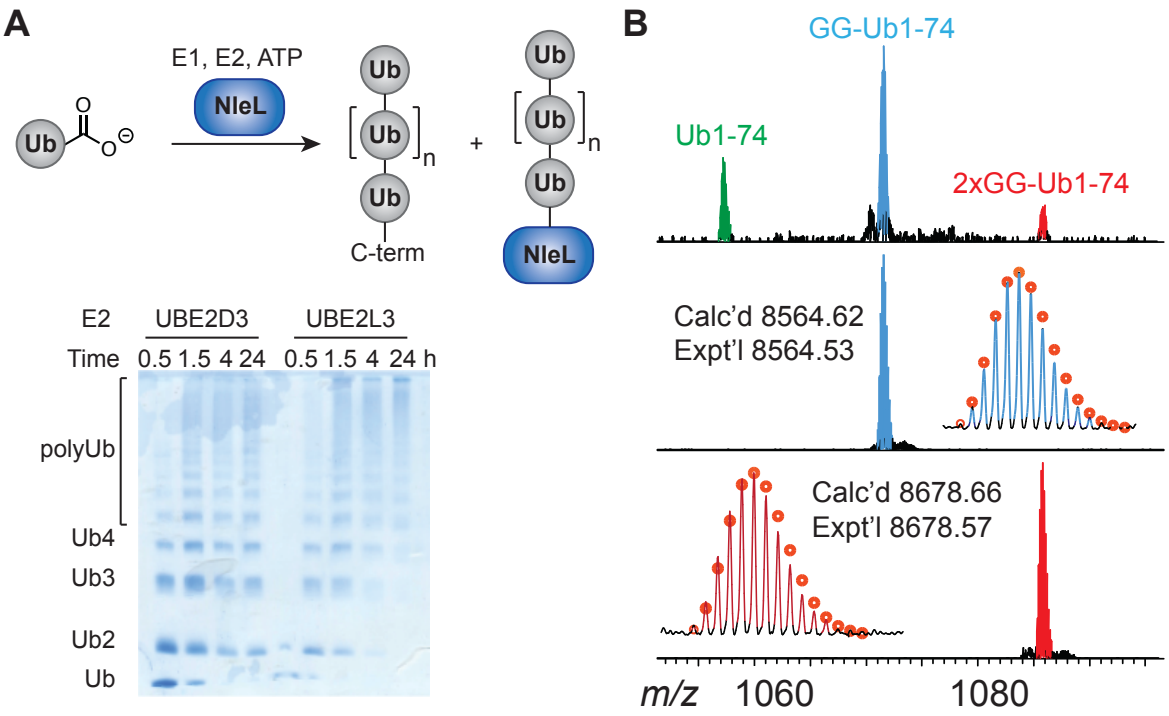


Figure 4.3. Middle-down MS analysis of NleL-catalyzed reactions **A.** Generation of unanchored and NleL-anchored polyUb chains using two different E2s: UBE2D3 and UBE2L3. **B.** FT-ICR MS analysis of NleL-catalyzed reactions after minimal trypsinolysis with UBE2D3 as the E2. The spectra correspond to the Ub^{8+} charge state. The top spectrum shows all three Ub_{1-74} species after restricted trypsin digestion. The middle spectrum displays the isolated GG-Ub_{1-74} M^{8+} parent ion and the bottom spectrum shows the isolated 2xGG-Ub_{1-74} M^{8+} parent ion.

Next, we sought to confirm that 2xGG-Ub_{1-74} was generated from a branch point. We first tested whether there was a correlation between the concentration of NleL and the levels of 2xGG-Ub_{1-74}. Changing the concentration of NleL from 50 nM to 5 μM was commensurate with an increase in the peak intensity of both 2xGG-Ub_{1-74} and GG-Ub_{1-74} (Figure 4.4A), indicating the presence of these Ub species is directly related to the catalytic activity of NleL. Analysis of minimally digested wild type Ub supported this conclusion, as signals corresponding to 2xGG-Ub_{1-74} and GG-Ub_{1-74} were absent in the MS spectra (Figure

4.4B). Because the 2xGG-Ub_{1-74} signal could also reflect a Ub derivative in which the GG motif is not removed from the C-terminal, i.e., proximal, subunit of an unanchored polyUb chain but is cleaved from the adjoining Ub, we wanted to evaluate minimal digestion of different linear chains. Previously, our lab has shown that a bifunctional Ub derivative harboring both a cysteine in lieu of a specific lysine residue and a C-terminal allylamine adduct could be polymerized using thiol-ene coupling chemistry to furnish single linkage polyUb chains of varying length that functionally mimic native oligomers. [37, 38] Thus, we assembled homotypic 6- and 48-linked polyUb chains using thiol-ene chemistry and subjected these polymers to restricted digestion conditions. FT-ICR analysis of the resulting Ub_{1-74} derivatives showed that 2xGG-Ub_{1}-74 is absent (Figure 4.4C). These data indicate that within a complex mixture of homotypic polyUb chains the GG motif on the proximal subunit is always removed by trypsin. Collectively, these results demonstrate that middle-down MS can be used to characterize the presence of branched polyUb chains.

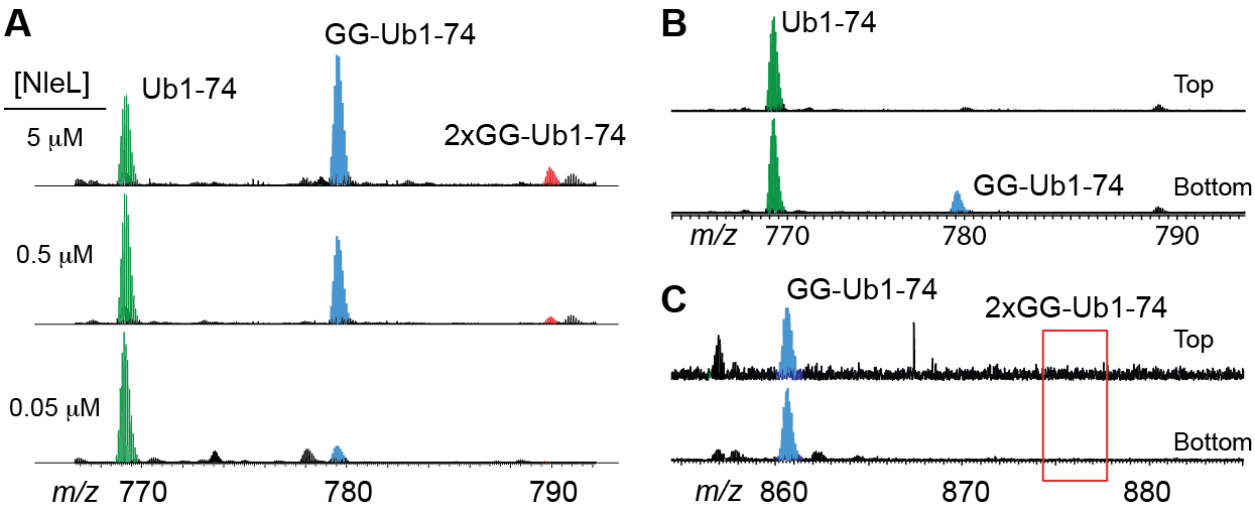


Figure 4.4. Formation of branch points under different conditions. **A.** Generation of 2xGG-Ub_{1-74} and GG-Ub_{1-74} as a function of NleL concentration. Reactions were performed with three different concentrations of NleL and subsequently digested with trypsin. FT-ICR MS shows that each Ub_{1-74} species can be detected at all three NleL concentrations. **B.** FT-ICR MS analysis of minimal trypsin digests of wild type Ub (top spectrum) and NleL-catalyzed

polyUb chain formation (bottom spectrum). GG-modified Ub_{1-74} is only present in reactions with NleL. **C.** Homotypic 6- and 48-linked polyUb chains of varying length were assembled using thiol-ene coupling and subsequently digested with trypsin. FT-ICR MS shows the presence of Ub_{1-74} and GG-Ub_{1-74}, but not 2xGG-Ub_{1-74} in the spectrum for both 6- (top spectrum) and 48-linked (bottom spectrum) polyUb chains. The red box indicates where the signal for 2xGG-Ub_{1-74} should be.

Identification of Modified Lysine Residues by ECD. To identify residues of 2xGG-Ub_{1-74} modified with Ub GG motifs, we used electron capture dissociation (ECD)[45] to induce protein fragmentation, resulting in the cleavage of N-Cα bonds and the formation of *c*- and $z^{\cdot}$-type ions. Analysis of the fragments then informs on the exact residues modified with GG. In the case of the 2xGG-Ub_{1-74} species, the M^{8+} charge state was isolated and fragmented by ECD, resulting in 19 *c* ions and 25 $z^{\cdot}$ ions with a total of 35 out of 73 bond cleavages (Figure 4.5B). The fragmentation patterns around K6 and K48 unambiguously confirmed the presence of isopeptide linkages at these positions (Figure 4.5A, Figure 4.14). ECD analysis was also performed on the M^{8+} charge state of GG-Ub_{1-74} and the fragmentation pattern showed the GG modification largely resides on K6 (Figure 4.13). These results are consistent with previous reports in that NleL assembles K6-linked polymers with greater efficiency than it generates K48 linkages.

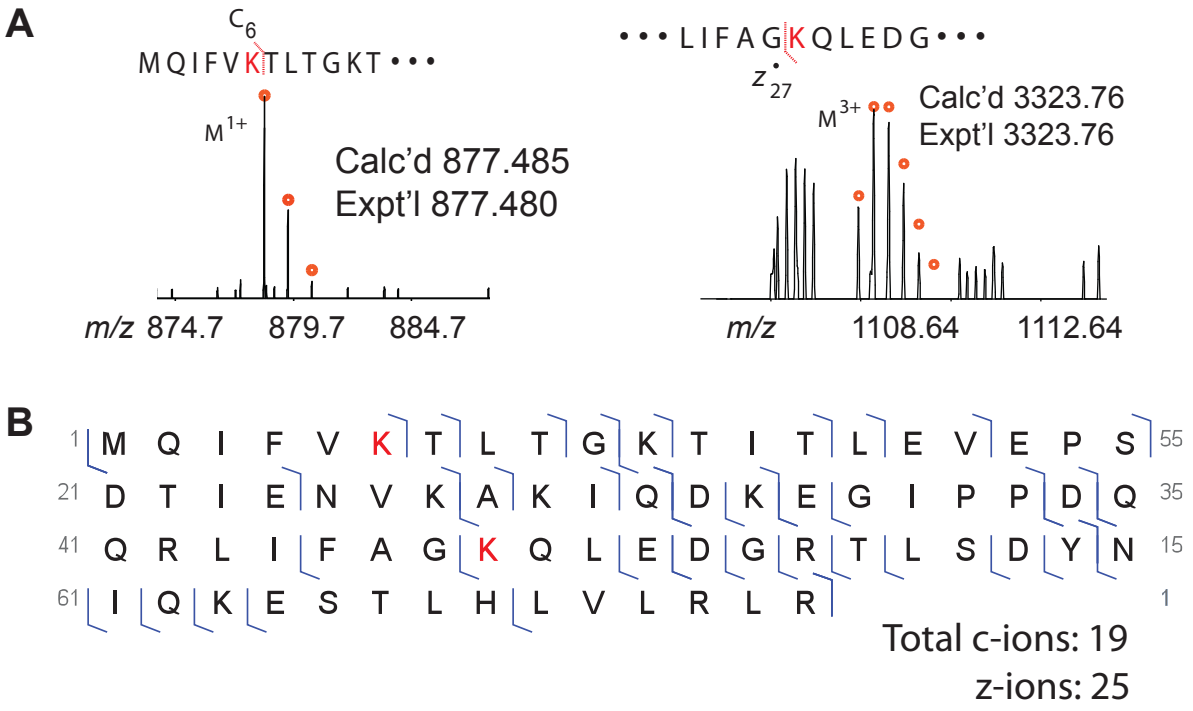


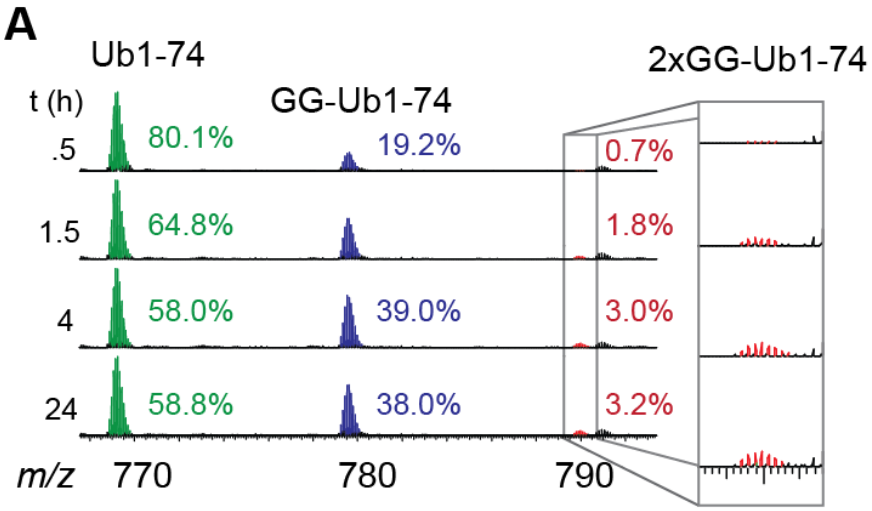
Figure 4.5. ECD analysis of 2xGG-Ub$_{1-74}$ generated from NleL-catalyzed reactions. **A.** ECD fragments of 2xGG-Ub$_{1-74}$ containing modified lysines (red K= lysine harboring GG motif). Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight. **B.** Observable ECD fragments (*c* and *z*· ions) containing GG-modified lysines at positions 6 and 48.

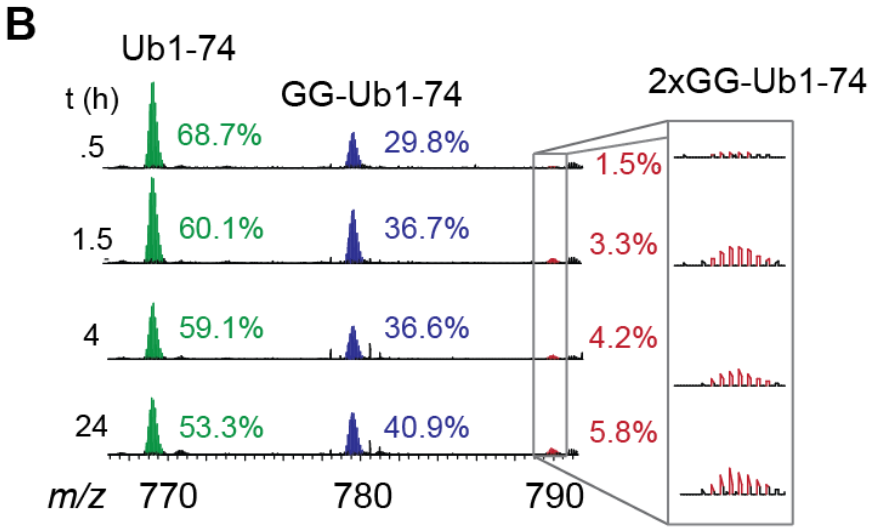
Branching Depends on Chain Length. With a characterization method established, we wanted to gain more insight into the formation of branched polyUb chains. We reasoned that since the Ub sequence remains largely the same between Ub$_{1-74}$, GG-Ub$_{1-74}$, and 2xGG-Ub$_{1-74}$, the effect of the GG motif on ionization efficiency should be negligible. As a result, FT-ICR MS could be used to measure the relative ratios of Ub$_{1-74}$, GG-Ub$_{1-74}$, and 2xGG-Ub$_{1-}$

$_{74}$ and monitor the formation of branch points over time. Indeed, a similar approach has been used to calculate the ratio of Ub_{1-74} and $GG\text{-}Ub_{1-74}$ in order to inform on the length of homotypic polyUb chains [23]. To ensure this strategy would be suitable for relative quantitation, signal ratios were measured over a range of charge states (Figure 4.15) from Ub^{+8} to Ub^{+12}; in each case the ratios remained invariant.

Using ratios of different Ub_{1-74} signals, the extent of chain branching was then examined. E3 ligases catalyze polyUb formation via a step-growth mechanism in which individual Ub subunits are sequentially added to the growing chain.[31, 46, 47] According to this mechanism, the product distribution should be composed of low molecular conjugates such as dimers, trimers, etc., at early time points. As the reaction proceeds, the distribution should shift to high molecular weight species, increasing the number of lysine residues in each chain that can serve as acceptors for another Ub. Accordingly, the probability of building a branched chain depends on the length of the polyUb chain because the internal acceptor sites outnumber the sites at the end of the chain as the chain gets longer. To test this, the population of branch points was evaluated at different times in chain formation. At each time point, signal intensities were measured and divided by the total population to arrive at the amount of individual Ub_{1-74} derivatives. With UBE2D3 as the E2 partner, approximately 3 % of the total chain population ($GG\text{-}Ub_{1-74}$ + $2xGG\text{-}Ub_{1-74}$) contains a branch point at the onset of the reaction (Figure 4.6A). As the reaction progresses the amount of branching doubles, reaching 7 % of the total chain population by 4 h (Figure 4.6A.) Additional branch points were not formed after allowing the reaction to proceed overnight. SDS-PAGE analysis showed that the polyUb chain population is largely composed of dimers, trimers, and tetramers at 0.5 h (Figure 4.3A). However, by 4 h a

significant proportion of the chains exist as high molecular weight species. Using UBE2L3 as the E2, the formation of polyUb chains appeared faster during the initial phases of the reaction relative to polymerizations conducted with UBE2D3 (Figure 4.6B). For instance, at 0.5 h, ~30 % of the Ub_{1-74} derivatives are composed of GG-Ub_{1-74} and 2xGG-Ub_{1-74} compared to only 20 % with UBE2D3 (Figure 4.6C). There is also a slight increase in the total amount of branch points with UBE2L3, as 2xGG-Ub_{1-74} represents 10 % of the chains by 4 h. These results suggest a combination of UBE2L3-NleL affords a more active chain assembly complex than UBE2D3-NleL. Consistent with this notion, SDS-PAGE analysis showed that with UBE2L3 dimers and trimers are almost completely consumed after 4 h, whereas with UBE2D3 short oligomers persist at longer incubation times (Figure 4.3A).


A
Ub1-74
GG-Ub1-74
2xGG-Ub1-74
t (h)
.5
80.1%
19.2%
0.7%
1.5
64.8%
1.8%
4
58.0%
39.0%
3.0%
24
58.8%
38.0%
3.2%
m/z
770
780
790


B
Ub1-74
GG-Ub1-74
2xGG-Ub1-74
t (h)
.5
68.7%
29.8%
1.5%
1.5
60.1%
36.7%
3.3%
4
59.1%
36.6%
4.2%
24
53.3%
40.9%
5.8%
m/z
770
780
790

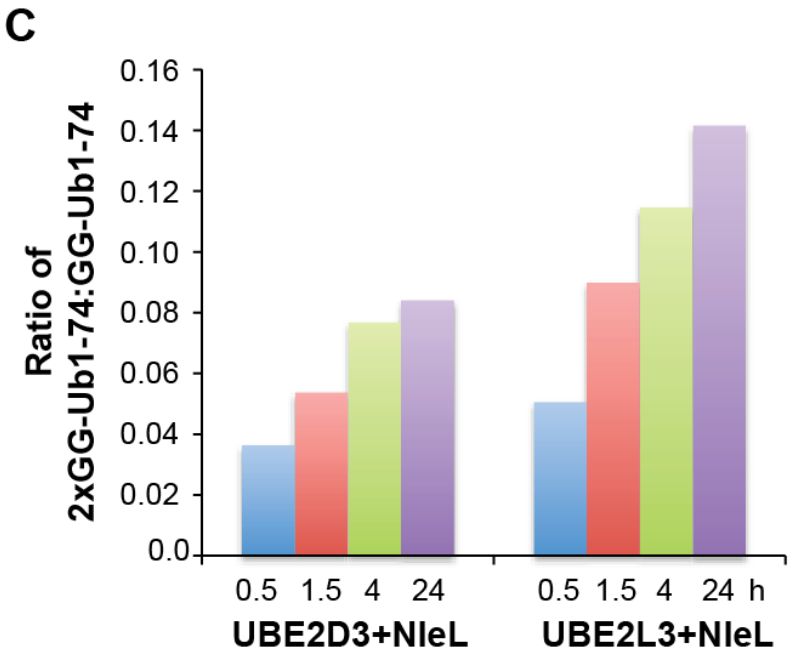
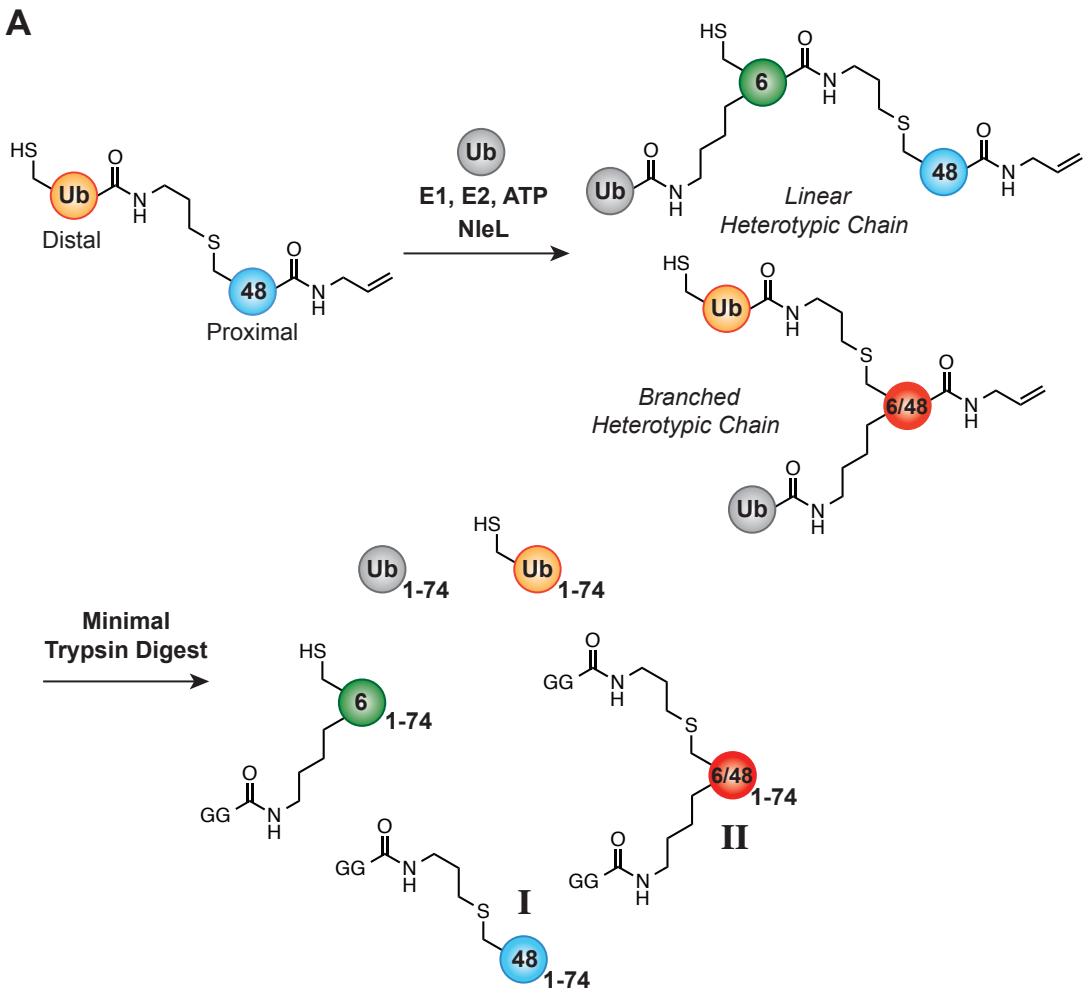

C
Ratio of
2xGG-Ub1-74:GG-Ub1-74
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.0
0.5 1.5 4 24
0.5 1.5 4 24 h
UBE2D3+NIeL
UBE2L3+NIeL

Figure 4.6. Dynamics of branched chain formation using FT-ICR to analyze minimally digested polyUb chains formed by NleL (0.5 uM) over time. **A.** Time course analysis of NleL-catalyzed reactions with UBE2D3 (1uM) as the E2. Percentages are based on the total population of Ub_{1-74} derivatives. **B.** Time course analysis of NleL-catalyzed reactions with UBE2L3 (1uM) as the E2. **C.** Ratio of branch point (2xGG-Ub_{1-74}) to linear (GG-Ub_{1-74}) Ub over time with UBE2D3 and UBE2L3.

To provide additional support for a model in which branching depends on chain length, we assessed the ability of NleL to extend unanchored polyUb chains. Since NleL is related to the HECT family of human E3 ligases and members of this family of enzymes are capable of binding and elongating free polyUb chains,[48, 49] we surmised NleL might also interact with and modify well-defined 6- and 48-linked oligomers. To evaluate this possibility, we again employed homotypic polyUb chains derived from thiol-ene coupling chemistry (Figure 4.7A).[37, 38] These chains contain cysteine residues in the distal subunits, which block extension in the form of homotypic chains and allow us to assess the ability of NleL to exclusively catalyze the formation of linear or branched heterotypic chains as a function of chain length. The thiol-ene-derived chains are also composed of a proximal subunit bearing a C-terminal allylamine adduct, which prevents activation by E1 and subsequent transfer to E2 and NleL. Thus, thiol-ene derived chains can only be modified if (i) Ub is also present in the reaction mixture and NleL catalyzes the transfer of Ub to the synthetic chains, or (ii) the allylamine moiety has been removed from the C-terminus using the yeast Ub C-terminal hydrolase Yuh1[37, 38, 50] and NleL transfers a preformed chain to another chain.[51] As evidenced by a new peak indicative of a branch point (Calc'd: 8710.83 Da), NleL catalyzes the modification of allylamine-capped, 48-linked Ub dimers only in the presence of free Ub (Figure 4.7B). With allylamine removed, the addition of free Ub is not

required for heterotypic chain formation, indicating that 48-linked dimers are shuttled from E1 to E2 and finally to the active site cysteine of NleL.


A
Distal
Proximal
Ub
E1, E2, ATP
NleL
Linear
Heterotypic Chain
Branched
Heterotypic Chain
Minimal
Trypsin Digest
1-74
6/48
I
II

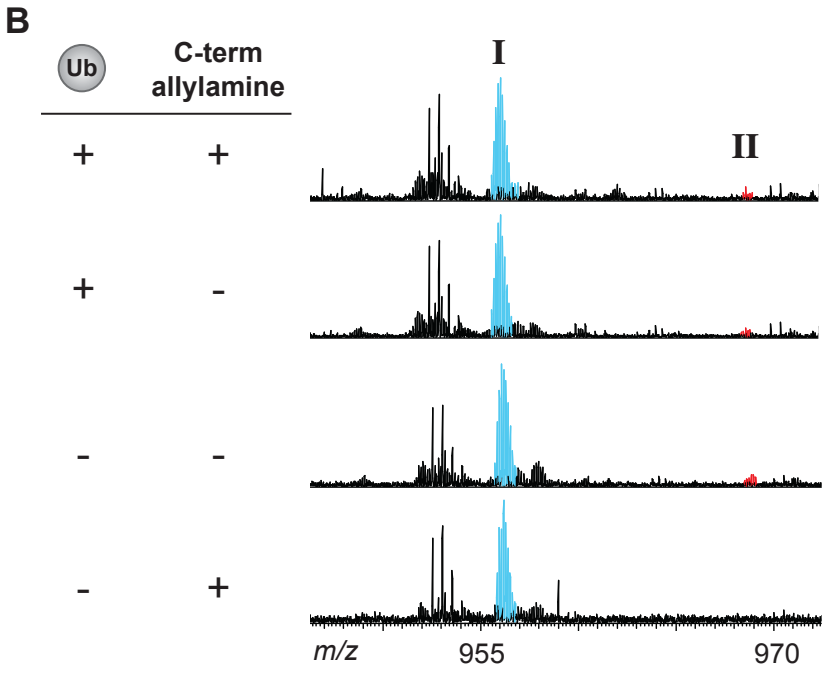

B
Ub
C-term
allylamine
I
II
m/z
955
970

Figure 4.7. Extension of preformed 48-linked Ub dimers using NleL. **A.** Reaction scheme depicting the modification of 48-linked dimers along with the subsequent trypsin digestion. **B.** FT-ICR MS analysis of NleL-catalyzed extension reactions. The resulting branched chains are shown in red. Preformed dimers either retained the allylamine moiety from the thiol-ene reaction or the allylamine was removed prior to chain extension using the C-terminal hydrolase Yuh1.

Encouraged by the results with 48-linked dimers, we sought to evaluate 6-linked oligomers as substrates of NleL. First, we wanted to investigate whether the linkage influences the efficiency of chain extension. Comparing reactions with 48- and 6-linked trimers, the ratio of branched product (II) to unmodified substrate (I) was measured (Figure 4.8A). From these data, we observed that II is formed to a much greater extent with 6-linked chains compared to those bearing 48-linkages, suggesting NleL prefers the former as substrates. Next, we wanted to evaluate the extent of branching as a function of chain length. Starting with 6-linked dimers, two new product peaks were observed: one corresponding to linear heterotypic chain (Calc'd: 8539.83 Da) and another commensurate with a branched chain (Figure 4.8B). With 6-linked trimers and tetramers, however, only 2xGG-Ub_{1-74} could be detected, suggesting linear chains are not formed and branching occurs by placing Ub on one or more of the internal subunits (Figures 4.8A and 4.8B). From these data we conclude that (i) branching depends on the linkage of the linear chain initially assembled by the E3 ligase, and (ii) as chain length increases, it is more favorable to install a branch point rather than extend the chain in a linear fashion.

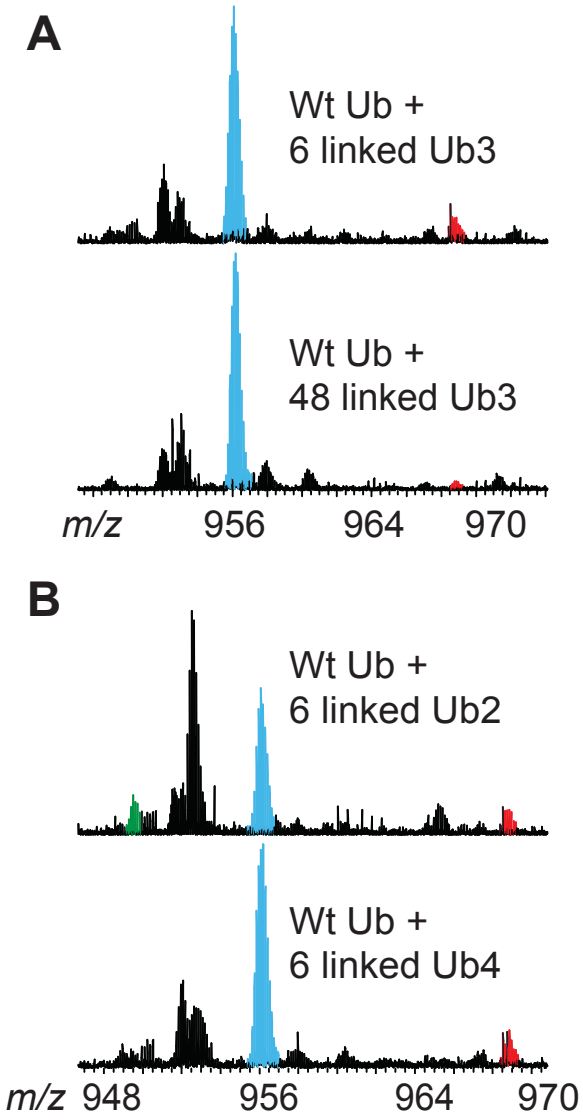


Figure 4.8. Impact of chain length and linkage on the ability of NleL to extend preformed Ub oligomers. **A.** Comparison between the NleL-catalyzed ubiquitylation of preformed 6- and 48-linked Ub trimers. **B.** Comparison between the NleL-catalyzed extension of preformed 6-linked dimers and tetramers. The linear products are shown in green and the branched chains are shown in red.

Formation of Branched Chains with Other Bacterial E3 Ligases. Recently, another class of bacterial effector proteins was discovered that catalyze polyUb chain formation via a thioester intermediate.[30, 52-55] Unlike NleL, these proteins, termed NELs (novel E3 ligases), bear no sequence or structural similarity to eukaryotic HECT or RING (really interesting new gene) ligases. NELs contain an N-terminal leucine-rich repeat and a C-terminal α-helical catalytic domain. Despite differences between bacterial HECT-like ligases and the NELs, we sought to determine whether a representative member of the NEL family, IpaH9.8, assembles branched polyUb chains. IpaH9.8 is produced by the pathogen *Shigella flexneri* to dampen the host inflammatory response during infection.[34] Consistent with previous reports, when IpaH9.8 was added to a reaction mixture containing UBE2D3 and Ub, polyUb chain formation was observed by SDS-PAGE (Figure 4.9B). Chain formation, however, was negligible with UBE2L3 as the E2. Minimal trypsin digest of the resulting Ub conjugates then led to the detection of Ub_{1-74}, GG-Ub_{1-74}, and 2xGG-Ub_{1-74} by FT-ICR MS (Figure 4.9A) and ECD analysis of GG-Ub_{1-74}, revealed that K48 is the predominant linkage (Figure 4.20). Compared to reactions with NleL, branch points are formed to a much lesser extent, making it difficult to characterize the linkages in 2xGG-Ub_{1-74} using ECD. To overcome this problem, we used a bottom-up approach. Complete trypsinolysis of the Ub conjugates followed by MS analysis using an Orbitrap instrument showed that both K6 and K48 linkages are present (Figure 4.21), suggesting IpaH9.8 catalyzes the formation of the same branch points as NleL.

Because the efficiency with which IpaH9.8 assembles branched chains is less than that of NleL, we sought to further investigate the relationship between chain extension and branching. The population of branch points (2xGG-Ub_{1-74}) was assessed at different time

points and compared to the amount of linear polyUb (GG-Ub_{1-74}) (Figure 4.9C). With NleL, branch points are detected within the first 20 % conversion of Ub into polyUb chains. By contrast, IpaH9.8 does not generate branch points until ~50 % of Ub has been transformed into polyUb chains. Analysis of IpaH9.8-catalyzed reactions by SDS-PAGE indicated that while chains form rapidly, they are lower in molecular weight relative to those produced by NleL. These results provide additional support for the conclusion that branching is directly related to chain length. Alternatively, since IpaH9.8 primarily builds chains bearing K48 linkages and our results with NleL show that it is difficult to generate branch points starting from preformed K48-linked chains, the low abundance of 2xGG-Ub_{1-74} could be a consequence of the linkage type within the linear chains.

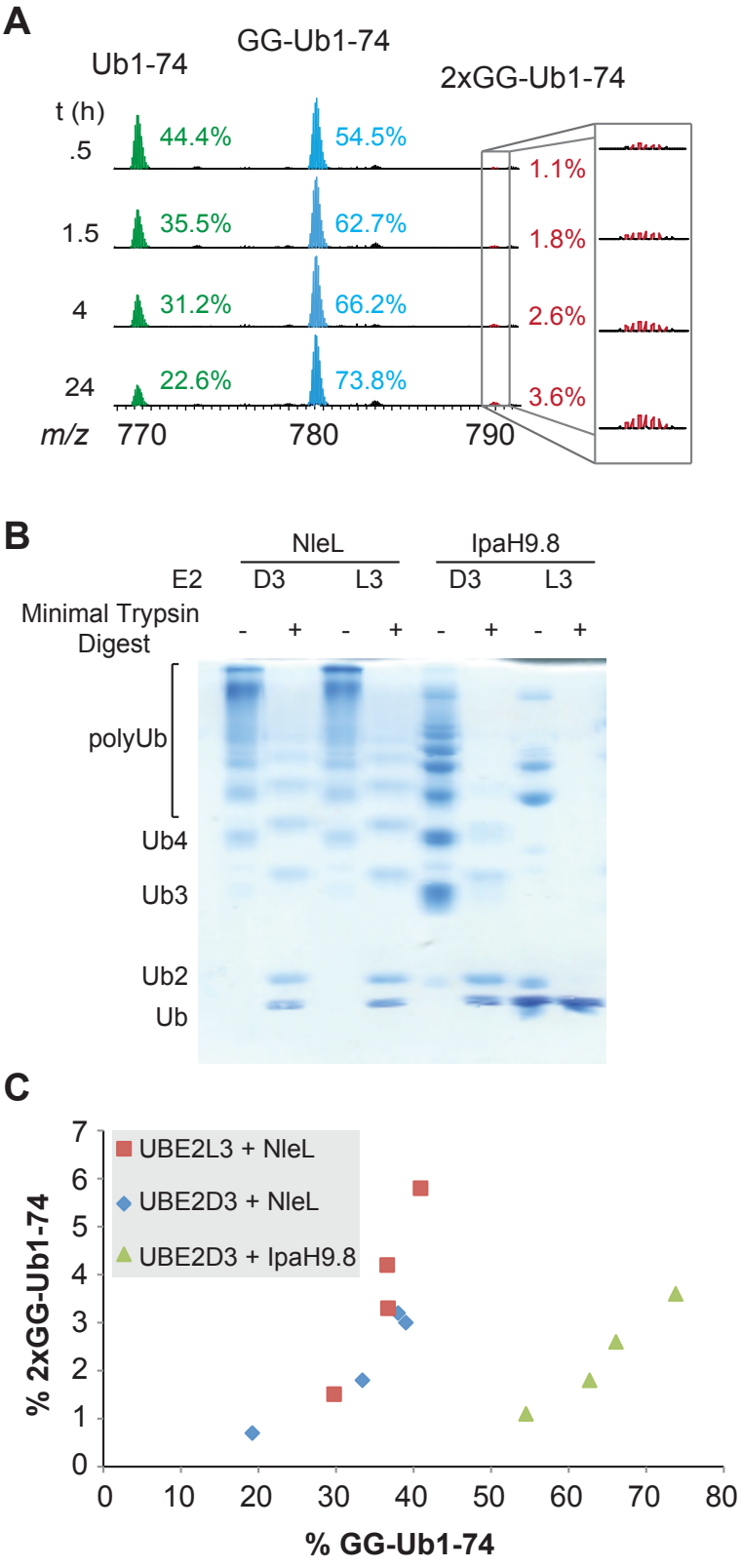

A
Ub1-74
GG-Ub1-74
2xGG-Ub1-74
t (h)
.5
1.5
4
24
44.4%
54.5%
1.1%
35.5%
62.7%
1.8%
31.2%
66.2%
2.6%
22.6%
73.8%
3.6%
m/z
770
780
790
B
NleL
IpaH9.8
E2
D3
L3
D3
L3
Minimal Trypsin Digest
- + - + - + - +
polyUb
Ub4
Ub3
Ub2
Ub
C
UBE2L3 + NleL
UBE2D3 + NleL
UBE2D3 + IpaH9.8
% 2xGG-Ub1-74
% GG-Ub1-74
0
1
2
3
4
5
6
7
0
10
20
30
40
50
60
70
80

Figure 4.9. Branching in IpaH9.8-catalyzed reactions. **A.** Time course analysis of IpaH9.8-catalyzed ubiquitylation reactions. Percentages are based on the total population of Ub_{1-74} derivatives. **B.** SDS-PAGE analysis of NleL- and IpaH9.8-catalyzed reactions. **C.** Examining the amount of 2xGG-Ub_{1-74} as a function of GG-Ub_{1-74} for both NleL- and IpaH9.8-catalyzed reactions.

4.5 DISCUSSION

Bottom-up MS methods have been instrumental in shaping our understanding of the repertoire of polyUb chain linkages in eukaryotic cells.[14, 21, 56] Trypsinolysis of polyUb chains furnishes signature peptides bearing a lysine residue with a GG motif from the C-terminus of Ub. The seven tryptic peptide fragments generated from this process then inform on the sites of polyUb chain formation. Because the signature GG motifs reside in distinct Ub peptides, it remains a formidable challenge to assess chain length and topology, i.e., the degree to which branched chains are formed unless modifications occur on adjacent lysines, e.g., K6+K11, K27+K29, and K29+K33.[22] Characterizing length and topology is important because both factors control the dynamics of biochemical pathways.[4, 5] For instance, chains with a minimum of four subunits provide an effective signal for proteasomal degradation[11] and differences in the intracellular trafficking of major histocompatibility complex class II (MHC II) in professional antigen-presenting cells is a consequence of differences in chain length.[57] With regards to topology, chain branching through K11 and K63 has also been implicated in controlling the rate at which MHC I is internalized by endocytosis,[58] and just recently, the anaphase-promoting complex (APC/C) was found to assemble branched chains containing K11 linkages in order to promote efficient degradation by the proteasome during prometaphase of the cell cycle.[59]

In the present work we exploited middle-down MS to investigate the formation of branched polyUb chains containing isopeptide linkages at non-adjacent lysines. Our studies focused on the bacterial E3 ligase NleL as it has the ability to not only tether heterotypic chains to itself – a process referred to as autoubiquitylation – but also form free, unanchored polyUb chains. The results of these studies demonstrate that NleL constructs branched chains containing isopeptide linkages at K6 and K48 when polyUb formation is completely unrestricted by the presence of K-to-R substitutions. Experiments with well-defined polyUb chains as substrates for NleL showed that it is equally probable to extend a dimer in the form of a linear or branched heterotypic chain. However, when longer chains, such as tetramers, are used as substrates, branching becomes a more prevalent modification. The observation that another bacterial E3 ligase, unrelated to NleL, along with the metazoan APC/C[59] also assemble branched conjugates suggests these atypical chains could be more widespread than currently appreciated.

In light of our results, we speculate that mixtures of linkages, which are often present in high molecular weight Ub conjugates, could represent branched chains both in enzymatic assays and cellular extracts. For example, high molecular weight K48-linked polyUb chains immunoprecipitated from mammalian cells using the K48 linkage-specific antibody typically contain K6, K11, and K63 linkages.[19, 20] Our results could also apply to the action of E4 Ub ligases as these enzymes collaborate with E1, E2, and E3s to catalyze the extension of polyUb chains [60]. The prototype of this activity is the yeast enzyme Ufd2, which elongates existing K29-linked chains by forming K48 linkages.[61] The initial Ufd2-dependent elongation step may occur by extending the chain from the end to form a linear

heterotypic chain. Alternatively, Ufd2 could first build a branch point before continuing to catalyze the sequential addition of Ub through K48.

The distinct Ub topologies created by different chain linkages could also play an important role in the ability to construct branched conjugates. Although NleL generates both K6 and K48 linkages, we found that by using unanchored polyUb chains as substrates K6-linked chains could be elongated in the form of a branched conjugate, whereas K48-linked chains were modified to a much lesser extent. These results, along with observations that long (>2 subunits) K6-linked chains are constructed by NleL at a much faster rate than those bearing K48 linkages,[33] suggest the structural ensemble of K6-linked chains is well-suited for NleL to maintain a persistent and productive interaction. By contrast, the conformations adopted by K48-linked chains may not promote high affinity interactions with NleL, and as a result, processive chain formation does not occur. Although additional binding studies are required to test this hypothesis, we surmise that as K6-linked chains become longer and the number of available K48 residues increases within each chain, entropy becomes an important factor in targeting the placement of a Ub unit at an internal subunit rather than at the chain terminus. This model could pertain to other ligases. For example, the yeast HECT ligase Rsp5 prefers to assemble K63 linkages, but since chains can also be extended through K11, K33, and K48 there could be a preponderance of arborization within these oligomers.[62] Middle-down MS will be instrumental in testing this model and further elucidating the biochemical details of E3 ligases, one of the most predominant classes of enzymes encoded by the human genome. Moreover, middle-down MS can ultimately be combined with affinity chromatography steps to uncover the extent to

which branched chains form under different cellular conditions, thus providing unprecedented insight into how Ub chain topology influences function.

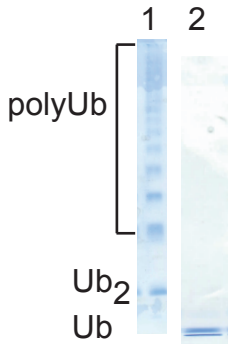


Figure 4.10. Trypsin Digest Optimization. Mass spectra of samples digested by optimized minimal trypsin digest. Gel 1: Ub chains formed from UBE2D3 and NleL. Gel 2: After optimized trypsin digest conditions. Lack of Ub_2 peak suggests complete digestion to Ub monomers.

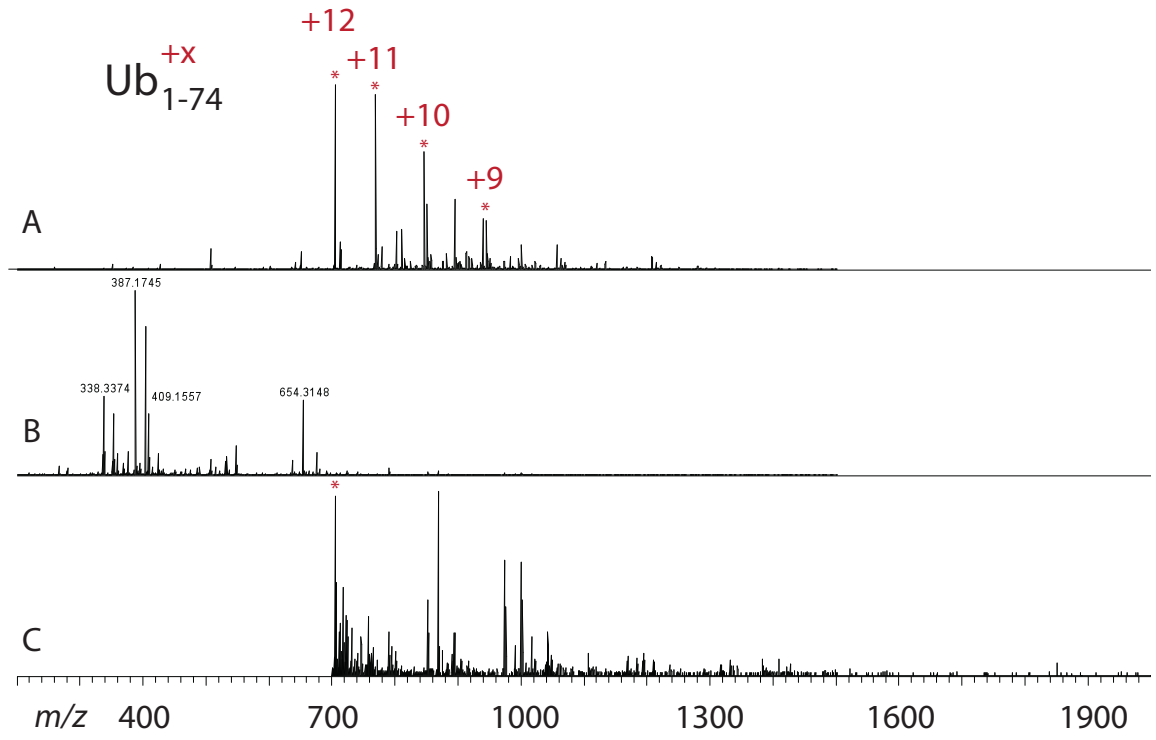


Figure 4.11. Ubiquitin Enrichment Optimization A. Mass spectra of sample processed by fully optimized work flow. * denotes peaks resulting from different charge states of Ub_{1-74} B. Mass spectra of sample processed using previous processing technique. C. Same sample as B, but m/z range is restricted to 700-2000. * denotes peaks resulting from different charge states of Ub_{1-74}

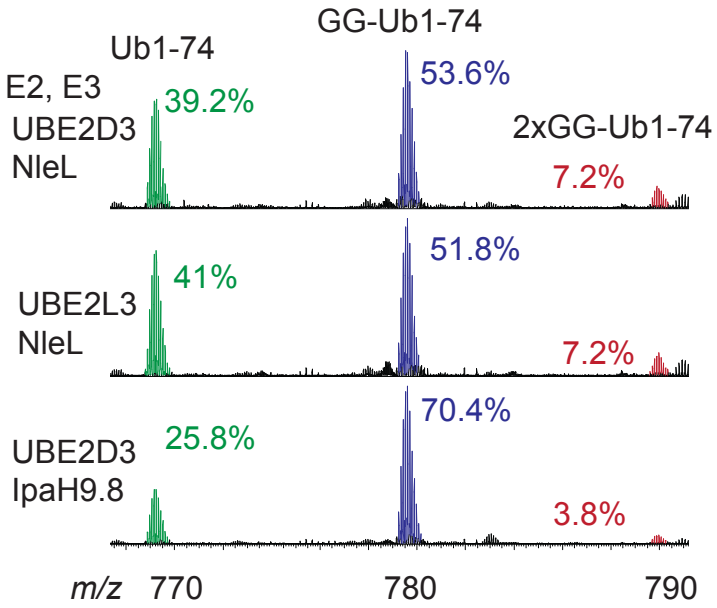


Figure 4.12. Mass Spectra of Ub Chains from Varying E2/E3 Enzymes
Branching varies with enzyme concentration (M^{11+})
Mass spectra of product distribution from chains formed with different E2 (1 uM)/ E3(5uM) combinations.

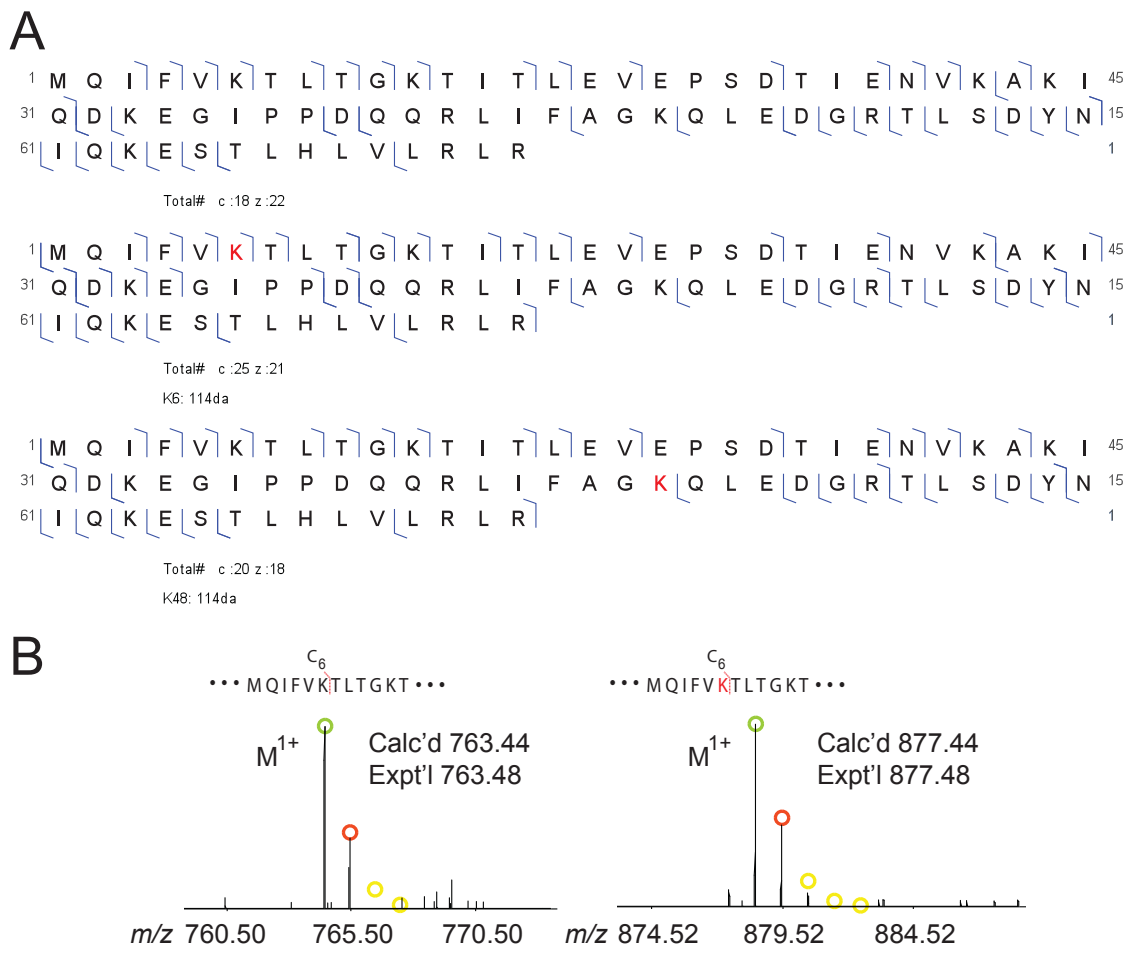
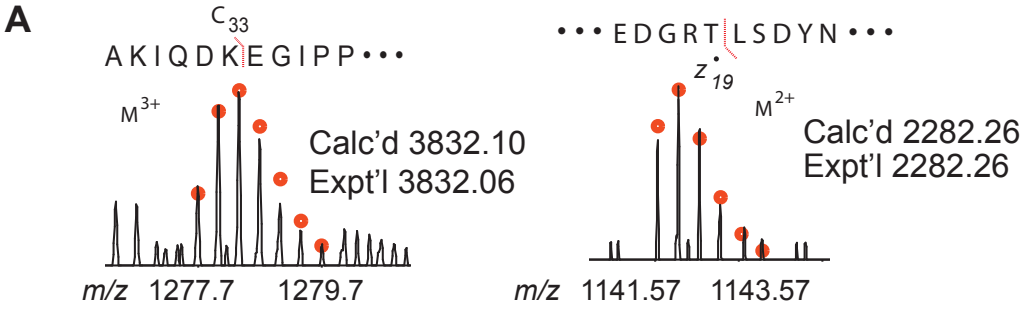
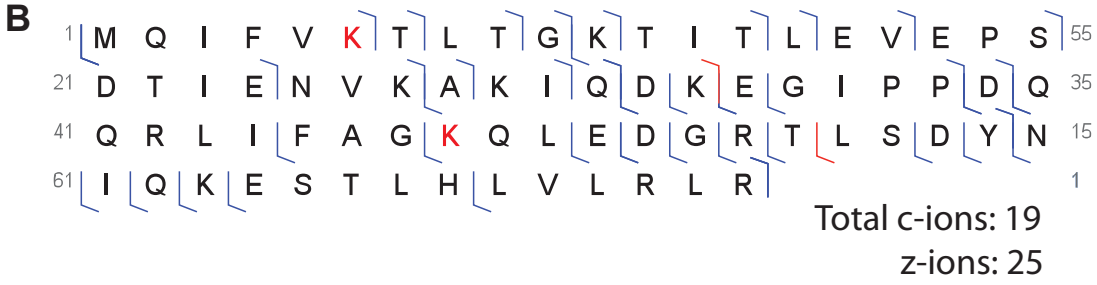


Figure 4.13. ECD Analysis of GG-Ub$_{1-74}$ (E2: UBE2D3, E3: NleL) A. Map of observed fragments. Data analysis for the bottom maps include *N*e-GlyGly-Lys linker modification at lysine 6 or 48 (red) in ***c*** and *z*˙ ion predictions. Top map does not include isopeptide linker modification in theoretical analysis. (Since GG-Ub$_{1-74}$ is a mixture of species modified at different lysines, many fragments are observable with and without *N*e-GlyGly-Lys linker modification, thus only fragment maps comparing branched modification sites were included). B. Key ECD fragment ions for mapping isopeptide linkage site on GG-Ub$_{1-74}$. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.


A
C 33
A K I Q D K E G I P P • • •
M 3+
Calc'd 3832.10
Expt'l 3832.06
m/z 1277.7 1279.7
• • • E D G R T L S D Y N • • •
z 19
M 2+
Calc'd 2282.26
Expt'l 2282.26
m/z 1141.57 1143.57


B
1 M Q I F V K T L T G K T I T L E V E P S 55
21 D T I E N V K A K I Q D K E G I P P D Q 35
41 Q R L I F A G K Q L E D G R T L S D Y N 15
61 I Q K E S T L H L V L R L R 1
Total c-ions: 19
z-ions: 25

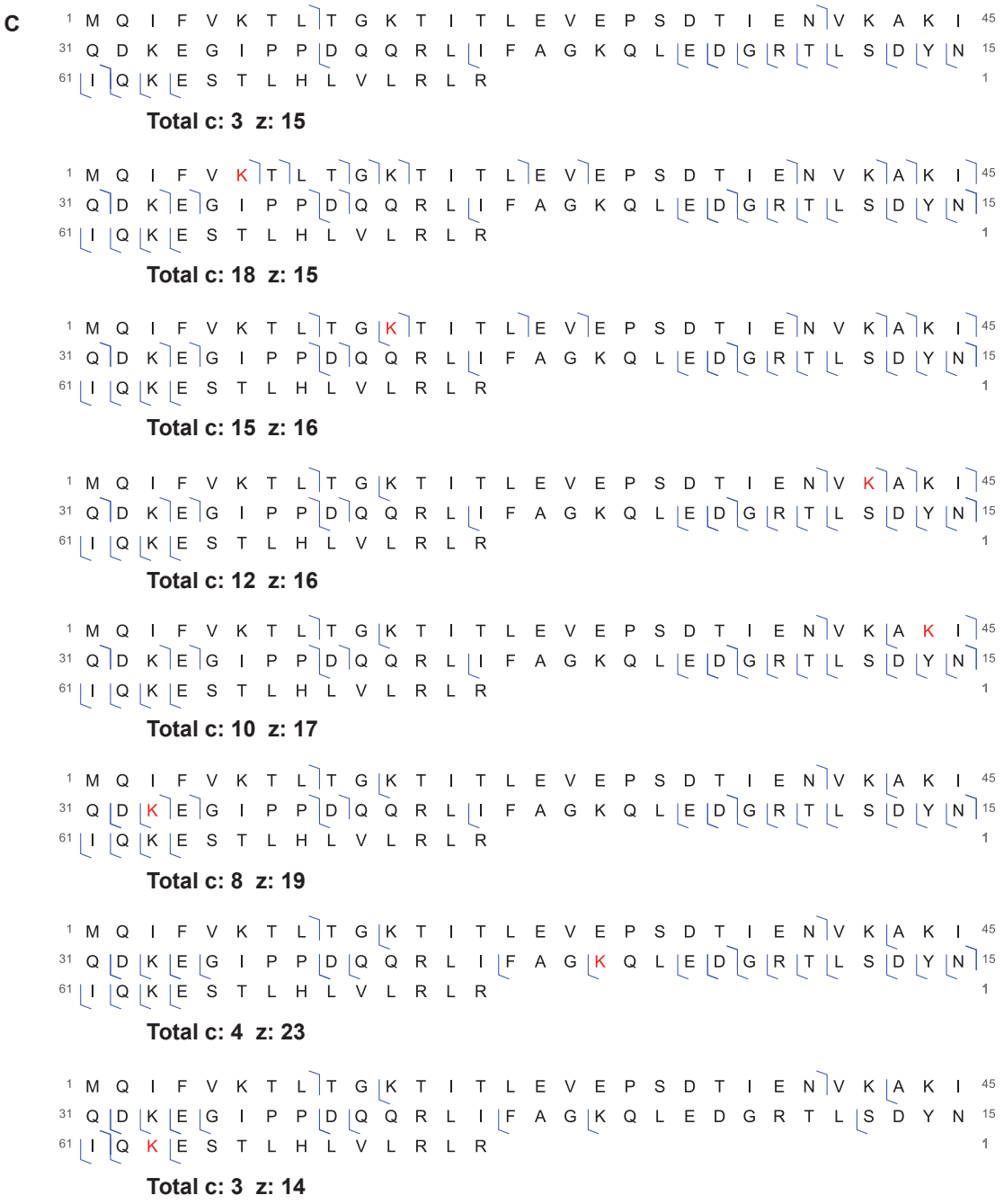


Figure 4.14. ECD Analysis of 2xGG-Ub$_{1-74}$ A. ECD fragments of 2xGG-Ub$_{1-74}$ that contain unmodified lysines at positions 11, 27, 29, 33, and 63. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight. **B.** Observable ECD fragments (*c* and *z*· ions) containing GG-modified lysines at positions 6 and 48 (red K= lysine harboring GG motif, fragments shown in a are highlighted in red). **C.** Observable ECD fragments of 2xGG-Ub$_{1-74}$ varying the location of the GG modification to each lysine showing that K6 and K48 contain the most observable ***c*** and *z*· ions.

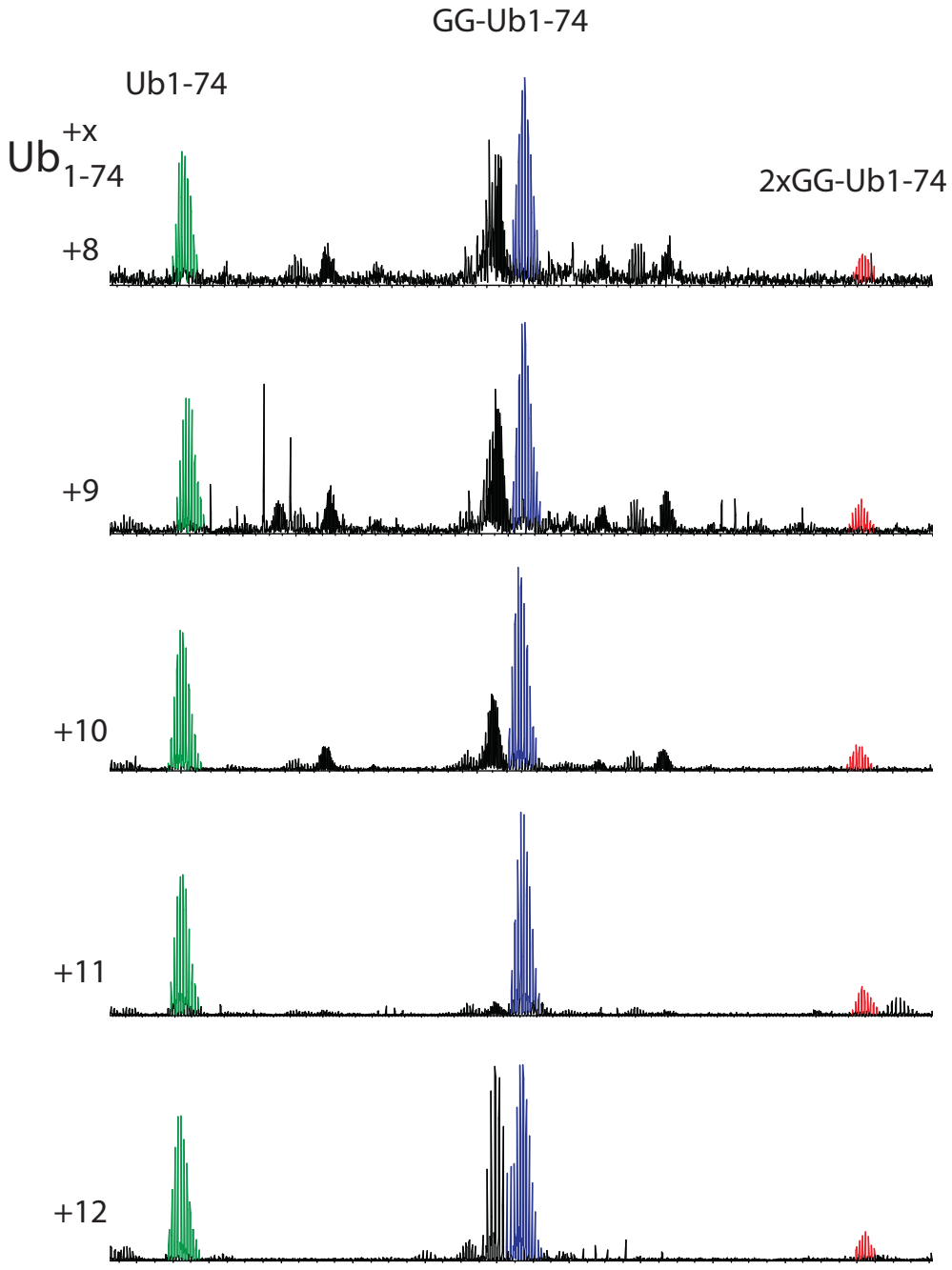


Figure 4.15. Ratios of Ub Species at Charge States 8-12 Ratio of Ub_{1-74} (green):GG-Ub_{1-74} (blue):2xGG-Ub_{1-74} (red) is the same in all charge states (Ub^{8+-12+}).

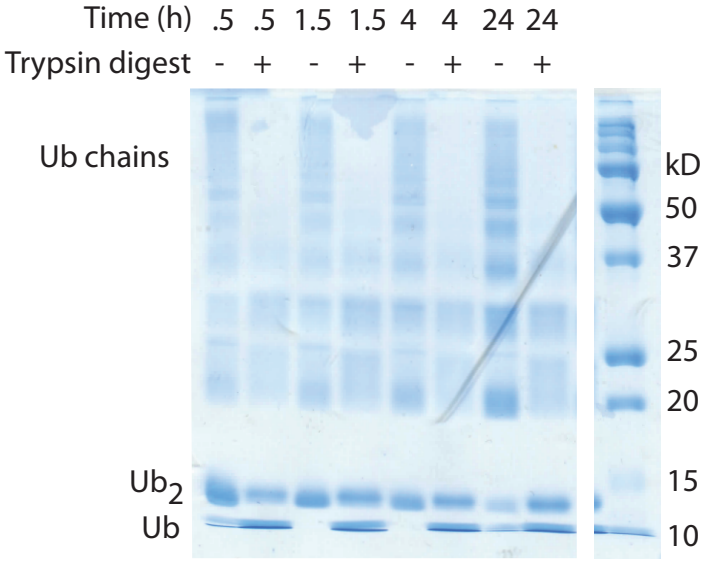
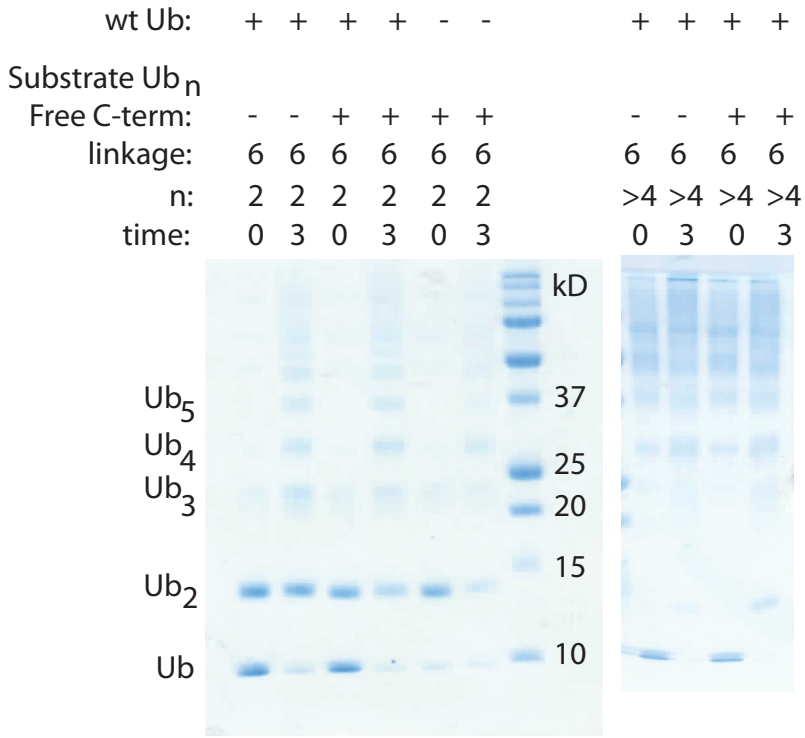


Figure 4.16. SDS-PAGE Analysis of Ub Chain Reactions. Coomassie-stained SDS-PAGE analysis of .5uM IpaH9.8 reactions with either UBE2D3 over time.



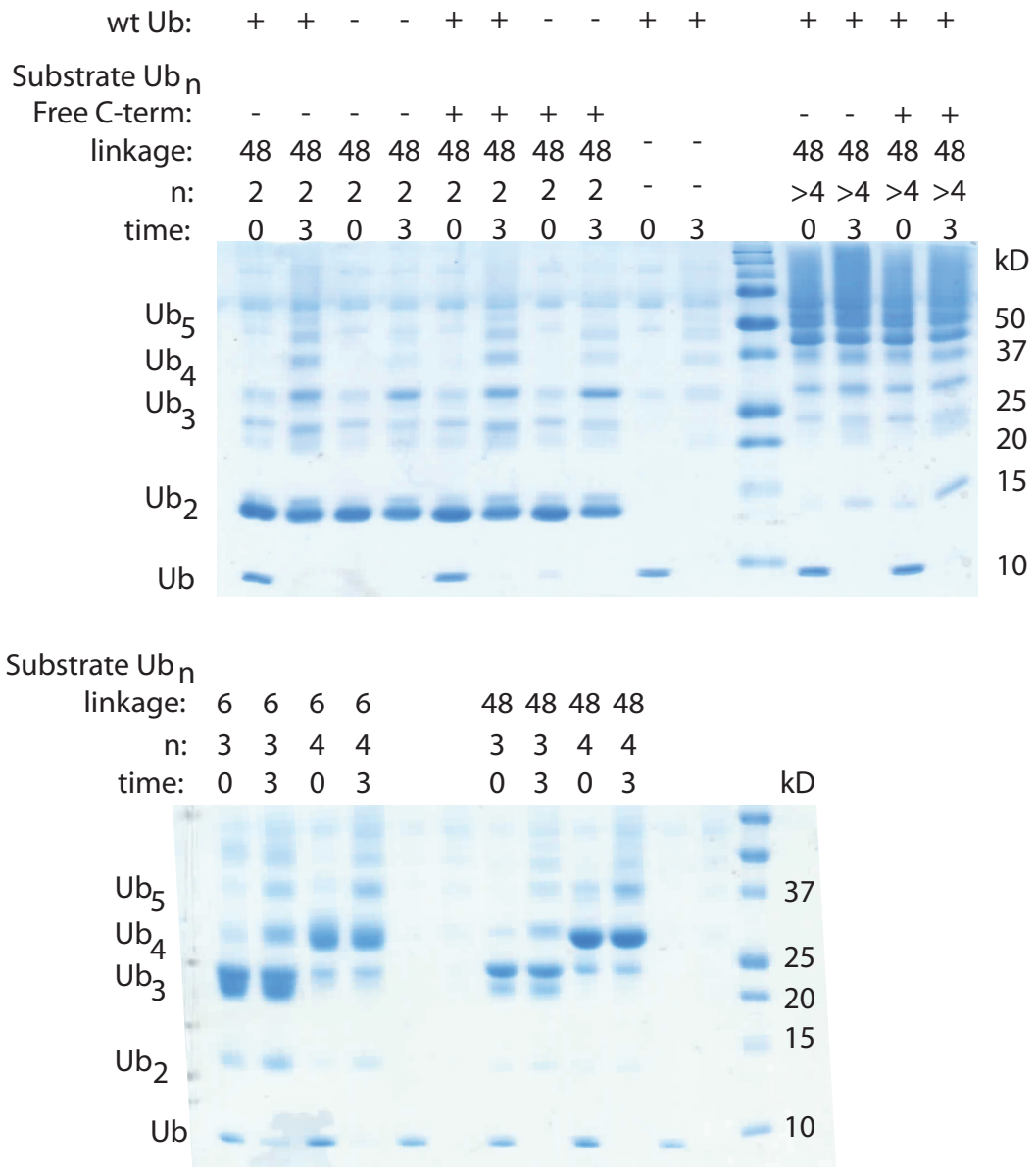


Figure 4.17. Coomassie-stained SDS-PAGE analysis of TEC Ub chains (varying length, linkage, and C-terminus) with wt Ub. T_0 = no ATP.

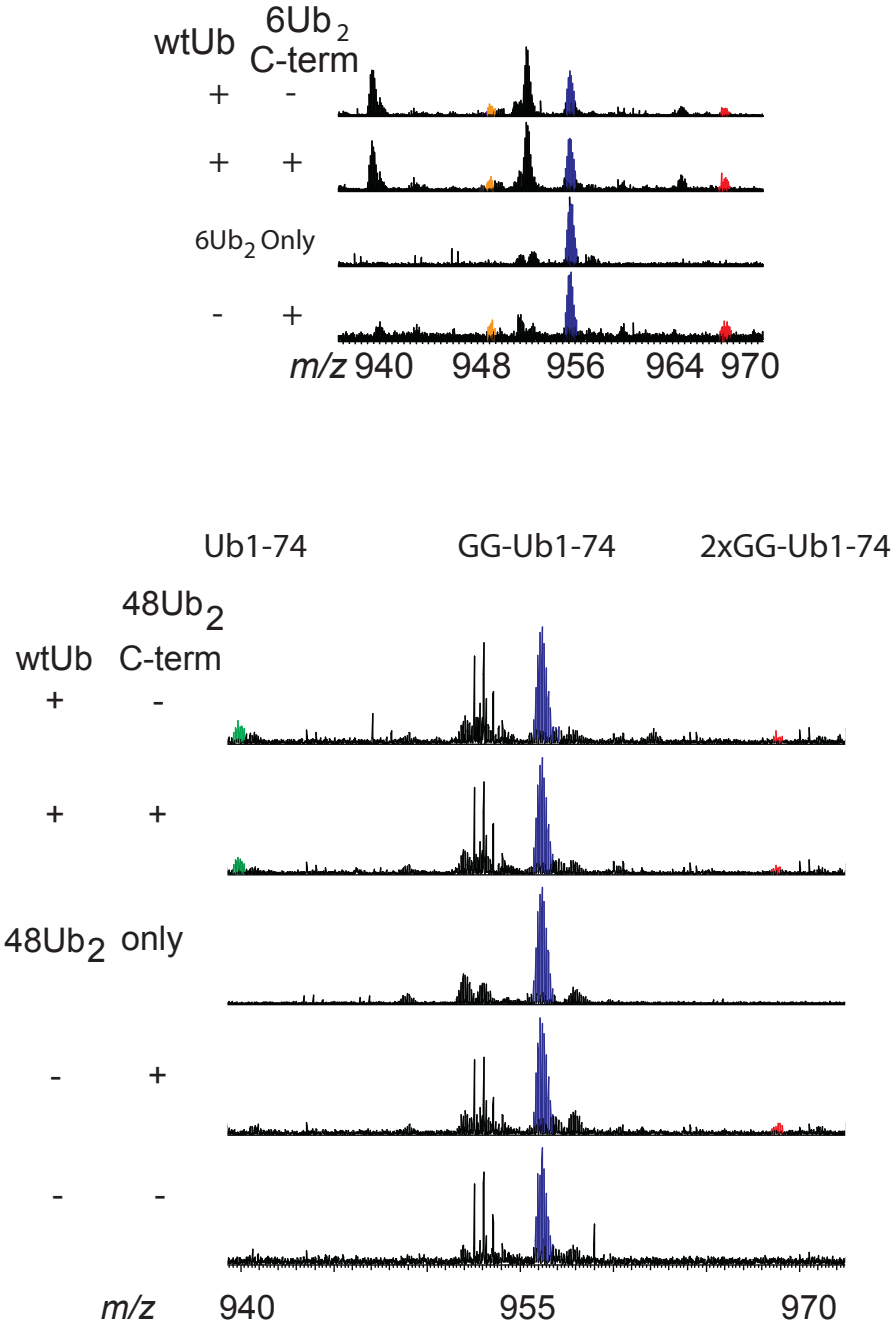


Figure 4.18. Mass Spectra of TEC 48-linked Ub_2 Mass spectra of products of TEC 6- and 48-linked Ub_2 (+/- free C-terminus) reacted with wt Ub (M^{9+} charge state). Peak colors: Yellow = linear heterotypic chain extension product; Blue = TEC linked GG-Ub_{1-74}; Red = branched chain product (2x-GG-Ub_{1-74}); Green = $wtUb_{1-74}$

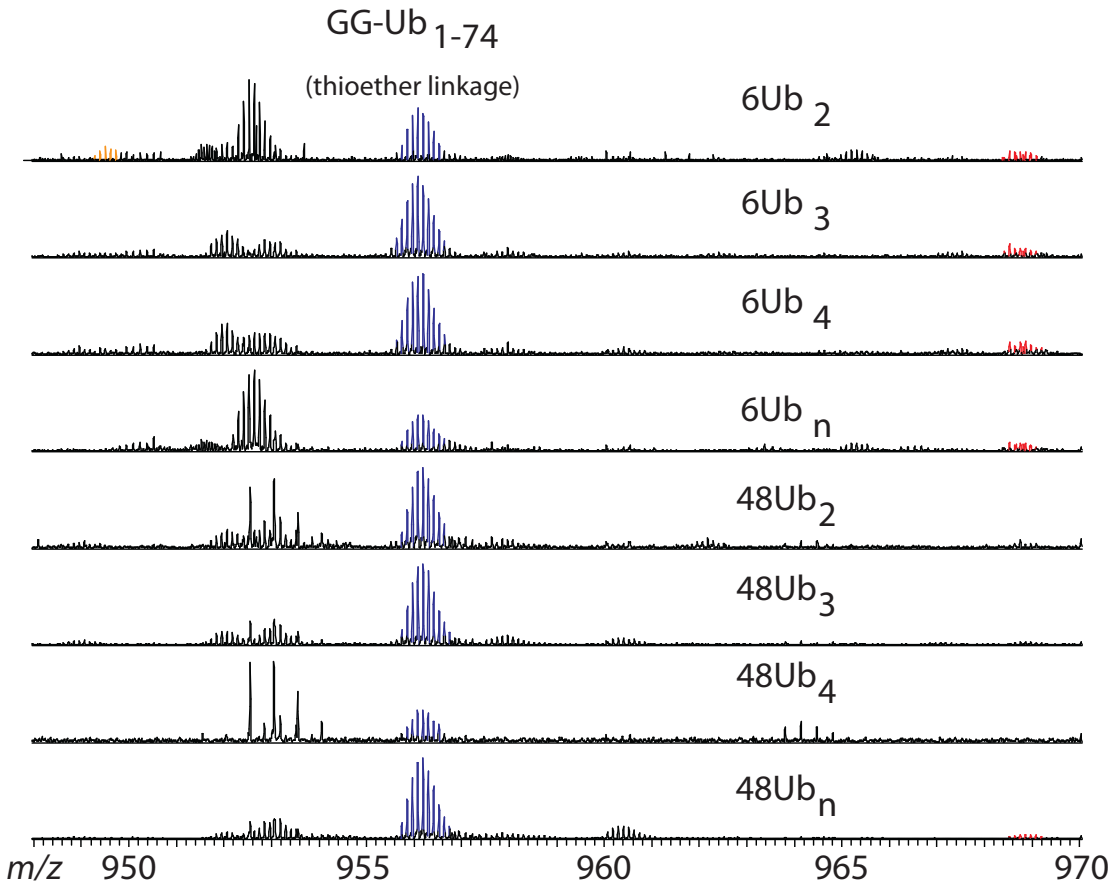


Figure 4.19. **Mass Spectra of TEC Substrates with wt Ub.** Mass spectra of products of TEC Ub chains (varying length, linkage) with wt Ub (M^{9+} charge state). Branched product = red, linear chain extension product = orange

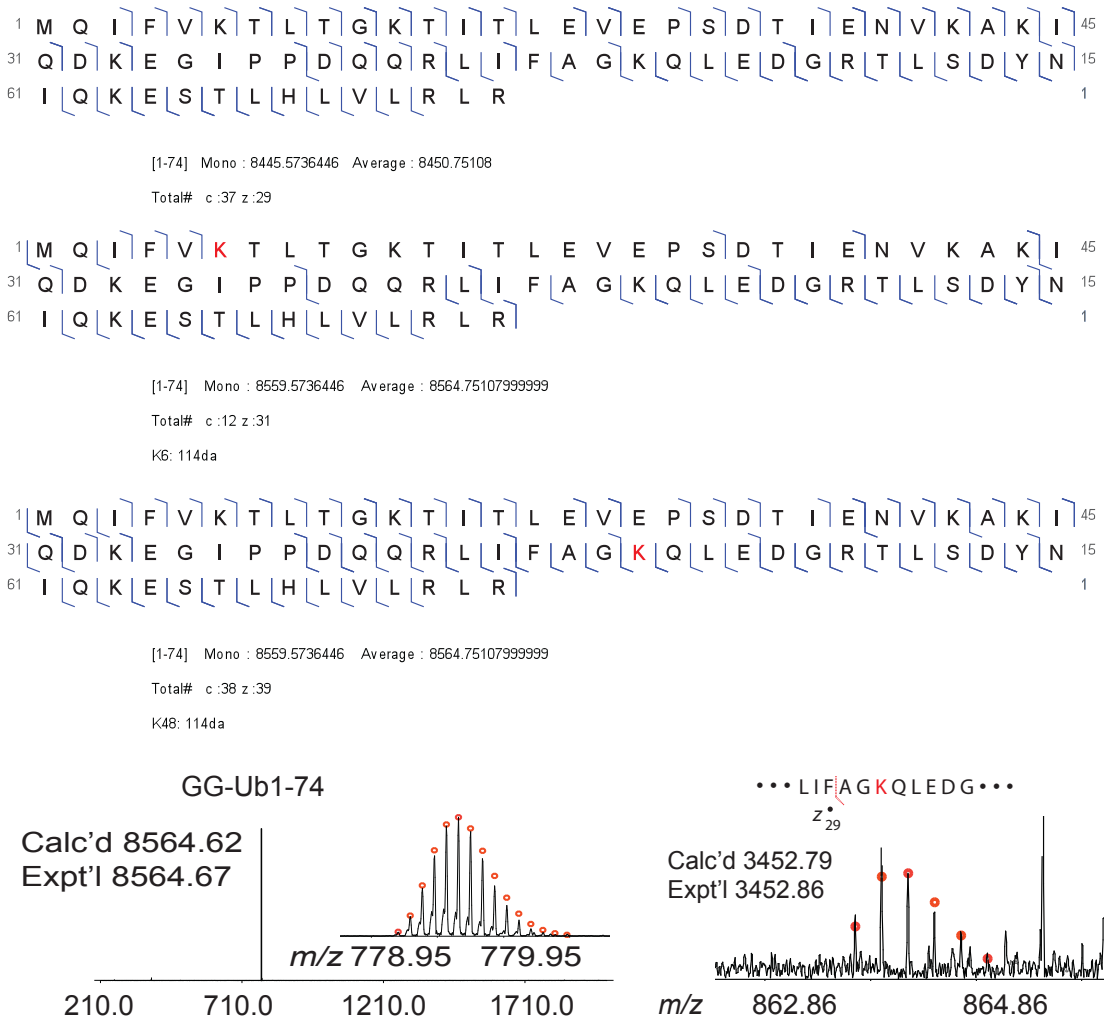


Figure 4.20. ECD Analysis For GG-Ub$_{1-74}$ (IpaH9.8). ECD of GG-Ub$_{1-74}$. (E2: UBE2D3, E3: IpaH9.8) A. Map of observed fragments. Data analysis for the bottom maps include *N*e-GlyGly-Lys linker modification at lysine 6 or 48 (red) in ***c*** and $z^{\cdot}$ ion predictions. Top map does not include isopeptide linker modification in theoretical analysis. B. Key ECD fragment ions for mapping isopeptide linkage site on GG-Ub$_{1-74}$. Circles represent theoretical isotopic abundance distribution of the isotopomer peaks. Calc'd: calculated most abundant molecular weight. Expt'l: experimental most abundant molecular weight.

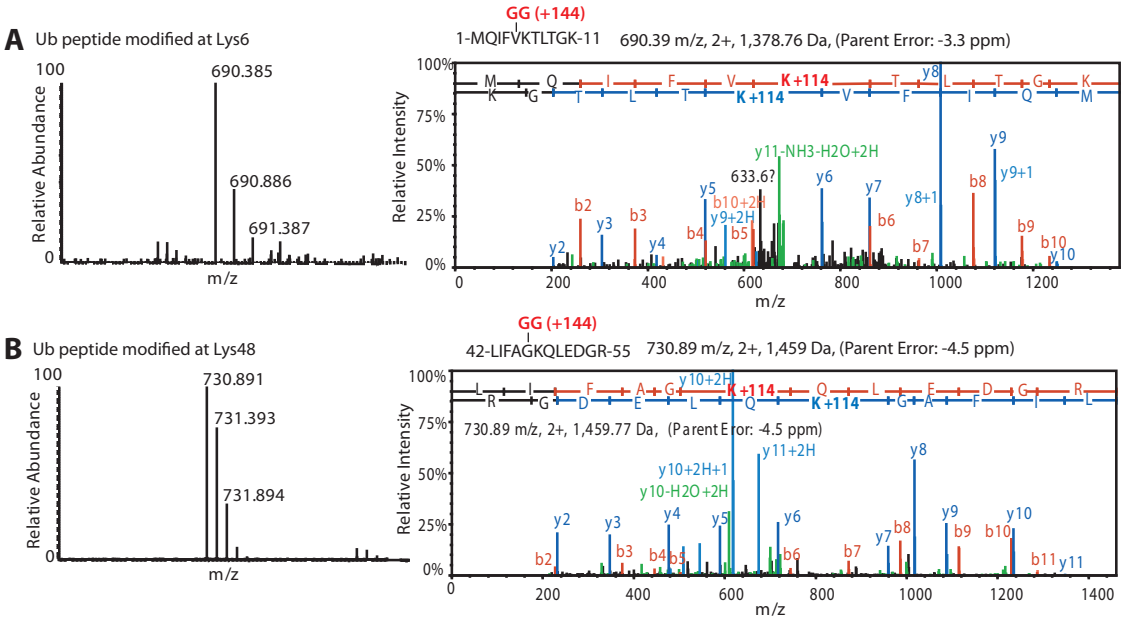


Figure 4.21: MS analysis of Ub linkages formed by IpaH9.8. A) Mass and MS/MS analysis of Ub peptide modified with GG at Lys6. B) Mass and MS/MS analysis of Ub peptide modified with GG at Lys48.

4.6 REFERENCES

[1] Hershko, A., and Ciechanover, A. (1998) The ubiquitin system, Annu. Rev. Biochem. 67, 425-479.

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

[3] Vucic, D., Dixit, V. M., and Wertz, I. E. (2011) Ubiquitylation in apoptosis: a post-translational modification at the edge of life and death, Nat. Rev. Mol. Cell Biol. 12, 439-452.

[4] Grabbe, C., Husnjak, K., and Dikic, I. (2011) The spatial and temporal organization of ubiquitin networks, Nat. Rev. Mol. Cell Biol. 12, 295-307.

[5] Komander, D., and Rape, M. (2012) The ubiquitin code, Annu. Rev. Biochem. 81, 203-229.

[6] Jackson, S. P., and Durocher, D. (2013) Regulation of DNA damage responses by ubiquitin and SUMO, Mol. Cell 49, 795-807.

[7] Pickart, C. M. (2001) Mechanisms underlying ubiquitination, Annu. Rev. Biochem. 70, 503-533.

[8] Dye, B. T., and Schulman, B. A. (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins, Annu. Review Biophys. Biomol. Struct. 36, 131-150.

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

[10] Finley, D., Sadis, S., Monia, B. P., Boucher, P., Ecker, D. J., Crooke, S. T., and Chau, V. (1994) Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant, Mol. Cell. Biol. 14, 5501-5509.

[11] Thrower, J. S., Hoffman, L., Rechsteiner, M., and Pickart, C. M. (2000) Recognition of the polyubiquitin proteolytic signal, EMBO J. 19, 94-102.

[12] Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination, Mol. Cell. Biol. 15, 1265-1273.

[13] Hofmann, R. M., and Pickart, C. M. (1999) Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair, Cell 96, 645-653.

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

[15] Matsumoto, M. L., Dong, K. C., Yu, C., Phu, L., Gao, X., Hannoush, R. N., Hymowitz, S. G., Kirkpatrick, D. S., Dixit, V. M., and Kelley, R. F. (2012) Engineering and structural characterization of a linear polyubiquitin-specific antibody, J. Mol. Biol. 418, 134-144.

[16] Povlsen, L. K., Beli, P., Wagner, S. A., Poulsen, S. L., Sylvestersen, K. B., Poulsen, J. W., Nielsen, M. L., Bekker-Jensen, S., Mailand, N., and Choudhary, C. (2012) Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass, Nat. Cell Biol. 14, 1089-1098.

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

[18] Kulathu, Y., and Komander, D. (2012) Atypical ubiquitylation - the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages, Nat. Rev. Mol. Cell Biol. 13, 508-523.

[19] Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komuves, L., French, D. M., Ferrando, R. E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S. G., Kelley, R. F., and Dixit, V. M. (2008)

Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies, Cell 134, 668-678.

[20] Phu, L., Izrael-Tomasevic, A., Matsumoto, M. L., Bustos, D., Dynek, J. N., Fedorova, A. V., Bakalarski, C. E., Arnott, D., Deshayes, K., Dixit, V. M., Kelley, R. F., Vucic, D., and Kirkpatrick, D. S. (2011) Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals, Mol. Cell. Proteomics 10, M110 003756.

[21] Xu, P., and Peng, J. (2006) Dissecting the ubiquitin pathway by mass spectrometry, Biochim. Biophys. Acta 1764, 1940-1947.

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

[23] Xu, P., and Peng, J. (2008) Characterization of polyubiquitin chain structure by middle-down mass spectrometry, Anal. Chem. 80, 3438-3444.

[24] Strachan, J., Roach, L., Sokratous, K., Tooth, D., Long, J., Garner, T. P., Searle, M. S., Oldham, N. J., and Layfield, R. (2012) Insights into the molecular composition of endogenous unanchored polyubiquitin chains, J. Proteome Res. 11, 1969-1980.

[25] Cannon, J. R., Edwards, N. J., and Fenselau, C. (2013) Mass-biased partitioning to enhance middle down proteomics analysis, J. Mass Spectrom. 48, 340-343.

[26] Chait, B. T. (2006) Chemistry. Mass spectrometry: bottom-up or top-down?, Science 314, 65-66.

[27] Forbes, A. J., Mazur, M. T., Patel, H. M., Walsh, C. T., and Kelleher, N. L. (2001) Toward efficient analysis of >70 kDa proteins with 100% sequence coverage, Proteomics 1, 927-933.

[28] Catherman, A. D., Skinner, O. S., and Kelleher, N. L. (2014) Top Down proteomics: Facts and perspectives, Biochem. Biophys. Res. Comm. 445, 683-693.

[29] Hicks, S. W., and Galan, J. E. (2013) Exploitation of eukaryotic subcellular targeting mechanisms by bacterial effectors, Nat. Rev. Microbiol. 11, 316-326.

[30] Hicks, S. W., and Galan, J. E. (2010) Hijacking the host ubiquitin pathway: structural strategies of bacterial E3 ubiquitin ligases, Curr. Opin. Microbiol. 13, 41-46.

[31] Lin, D. Y., Diao, J., Zhou, D., and Chen, J. (2011) Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7, J. Biol. Chem. 286, 441-449.

[32] Piscatelli, H., Kotkar, S. A., McBee, M. E., Muthupalani, S., Schauer, D. B., Mandrell, R. E., Leong, J. M., and Zhou, D. (2011) The EHEC type III effector NleL is an E3 ubiquitin ligase that modulates pedestal formation, PloS One 6, e19331.

[33] Hospenthal, M. K., Freund, S. M., and Komander, D. (2013) Assembly, analysis and architecture of atypical ubiquitin chains, Nat. Struct. Mol. Biol. 20, 555-565.

[34] Ashida, H., Kim, M., Schmidt-Supprian, M., Ma, A., Ogawa, M., and Sasakawa, C. (2010) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-kappaB-mediated inflammatory response, Nat. Cell Biol. 12, 66-73; sup pp 61-69.

[35] Pickart, C. M., and Raasi, S. (2005) Controlled synthesis of polyubiquitin chains, Meth. Enzymol. 399, 21-36.

[36] Shekhawat, S., Pham, G. H., Prabakaran, J., and Strieter, E. R. (2014) Unpublished results.

[37] Trang, V. H., Valkevich, E. M., Minami, S., Chen, Y. C., Ge, Y., and Strieter, E. R. (2012) Nonenzymatic polymerization of ubiquitin: single-step synthesis and isolation of discrete ubiquitin oligomers, Angew. Chem. 51, 13085-13088.

[38] Valkevich, E. M., Guenette, R. G., Sanchez, N. A., Chen, Y. C., Ge, Y., and Strieter, E. R. (2012) Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases, J. Am. Chem. Soc. 134, 6916-6919.

[39] Ayaz-Guner, S., Zhang, J., Li, L., Walker, J. W., and Ge, Y. (2009) In vivo phosphorylation site mapping in mouse cardiac troponin I by high resolution top-down electron capture dissociation mass spectrometry: Ser22/23 are the only sites basally phosphorylated, Biochemistry 48, 8161-8170.

[40] Ge, Y., Rybakova, I. N., Xu, Q., and Moss, R. L. (2009) Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state, Proc. Nat. Acad. Sci. USA 106, 12658-12663.

[41] Zhang, J., Guy, M. J., Norman, H. S., Chen, Y. C., Xu, Q., Dong, X., Guner, H., Wang, S., Kohmoto, T., Young, K. H., Moss, R. L., and Ge, Y. (2011) Top-down quantitative proteomics identified phosphorylation of cardiac troponin I as a candidate biomarker for chronic heart failure, J. Proteome Res. 10, 4054-4065.

[42] Guner, H., Close, P. L., Cai, W., Zhang, H., Peng, Y., Gregorich, Z. R., and Ge, Y. (2014) MASH Suite: a user-friendly and versatile software interface for high-resolution mass spectrometry data interpretation and visualization, J. Am. Soc. Mass Spectrom. 25, 464-470.

[43] Wilkinson, K. D., and Audhya, T. K. (1981) Stimulation of ATP-dependent proteolysis requires ubiquitin with the COOH-terminal sequence Arg-Gly-Gly, J. Biol. Chem. 256, 9235-9241.

[44] Compton, P. D., Zamdborg, L., Thomas, P. M., and Kelleher, N. L. (2011) On the scalability and requirements of whole protein mass spectrometry, Anal. Chem. 83, 6868-6874.

[45] Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process, J. Am. Chem. Soc. 120, 3265-3266.

[46] Pierce, N. W., Kleiger, G., Shan, S. O., and Deshaies, R. J. (2009) Detection of sequential polyubiquitylation on a millisecond timescale, Nature 462, 615-619.

[47] Kamadurai, H. B., Qiu, Y., Deng, A., Harrison, J. S., Macdonald, C., Actis, M., Rodrigues, P., Miller, D. J., Souphron, J., Lewis, S. M., Kurinov, I., Fujii, N., Hammel, M., Piper, R., Kuhlman, B., and Schulman, B. A. (2013) Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3, eLife 2, e00828.

[48] Wang, M., Cheng, D., Peng, J., and Pickart, C. M. (2006) Molecular determinants of polyubiquitin linkage selection by an HECT ubiquitin ligase, EMBO J. 25, 1710-1719.

[49] Ronchi, V. P., Klein, J. M., and Haas, A. L. (2013) E6AP/UBE3A ubiquitin ligase harbors two E2~ubiquitin binding sites, J. Biol. Chem. 288, 10349-10360.

[50] Johnston, S. C., Riddle, S. M., Cohen, R. E., and Hill, C. P. (1999) Structural basis for the specificity of ubiquitin C-terminal hydrolases, EMBO J. 18, 3877-3887.

[51] Li, W., Tu, D., Brunger, A. T., and Ye, Y. (2007) A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate, Nature 446, 333-337.

[52] Rohde, J. R., Breitkreutz, A., Chenal, A., Sansonetti, P. J., and Parsot, C. (2007) Type III secretion effectors of the IpaH family are E3 ubiquitin ligases, Cell Host Microbe 1, 77-83.

[53] Singer, A. U., Rohde, J. R., Lam, R., Skarina, T., Kagan, O., Dileo, R., Chirgadze, N. Y., Cuff, M. E., Joachimiak, A., Tyers, M., Sansonetti, P. J., Parsot, C., and Savchenko, A. (2008) Structure of the Shigella T3SS effector IpaH defines a new class of E3 ubiquitin ligases, Nat. Struct. Mol. Biol. 15, 1293-1301.

[54] Zhu, Y., Li, H., Hu, L., Wang, J., Zhou, Y., Pang, Z., Liu, L., and Shao, F. (2008) Structure of a Shigella effector reveals a new class of ubiquitin ligases, Nat. Struct. Mol. Biol. 15, 1302-1308.

[55] Quezada, C. M., Hicks, S. W., Galan, J. E., and Stebbins, C. E. (2009) A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases, Proc. Nat. Acad. Sci. USA 106, 4864-4869.

[56] Chen, P. C., Na, C. H., and Peng, J. (2012) Quantitative proteomics to decipher ubiquitin signaling, Amino Acids 43, 1049-1060.

[57] Ma, J. K., Platt, M. Y., Eastham-Anderson, J., Shin, J. S., and Mellman, I. (2012) MHC class II distribution in dendritic cells and B cells is determined by ubiquitin chain length, Proc. Nat. Acad. Sci. USA 109, 8820-8827.

[58] Boname, J. M., Thomas, M., Stagg, H. R., Xu, P., Peng, J., and Lehner, P. J. (2010) Efficient internalization of MHC I requires lysine-11 and lysine-63 mixed linkage polyubiquitin chains, Traffic 11, 210-220.

[59] Meyer, H. J., and Rape, M. (2014) Enhanced protein degradation by branched ubiquitin chains, Cell 157, 910-921.

[60] Hoppe, T. (2005) Multiubiquitylation by E4 enzymes: 'one size' doesn't fit all, Trends Biochem. Sci. 30, 183-187.

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

[62] Kim, H. C., and Huibregtse, J. M. (2009) Polyubiquitination by HECT E3s and the determinants of chain type specificity, Mol. Cell. Biol. 29, 3307-3318.

Chapter 5: Identifying Branched Ubiquitin Chains in Cell Extracts

Contribution:

MS experiments were performed by Mark Scalf (Smith lab)

5.1 Introduction

Ub chain topology dictates the function, and thus biological outcome of ubiquitinated proteins[1]. Knowledge connecting topology to function is uncovered by analytical techniques such as bottom-up MS, which distinguish the location of isopeptide bonds. The scope of this analysis is limited to linear chains because localizing two isopeptide modifications to the same protein is only possible if the modifications are not divided during sample preparation. Top-down MS enables the analysis of multiple PTMs by analyzing intact proteins, but has not been applied extensively to branched Ub chains. We aim to expand MS techniques to monitor branched Ub chains.

In Chapter 4, we demonstrated that middle-down MS can be used to unambiguously identify branched Ub chains from *in vitro* enzymatic reactions[2]. This method works well since Ub peptides are in much greater abundance than peptides from other reaction components. Indeed, the concentration of Ub is 10 to 100 times higher than that of the chain constructing enzymes, and Ub_{1-74}, GG-Ub_{1-74}, and 2xGG-Ub_{1-74} represent the majority of the peptide/protein species that can be analyzed after digestion by trypsin and subsequent dialysis. In order to expand this technique to analyze complex mixtures in which Ub chains are not the most prevalent components, such as in mammalian cells, a Ub enrichment step must precede MS analysis (figure 5.1).

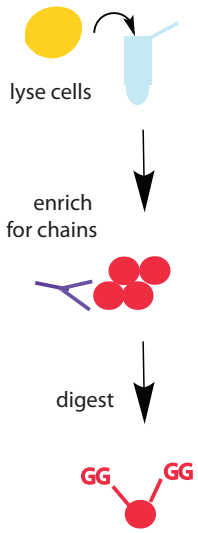


Figure 5.1. Scheme of work flow to enrich for chains from cell lysate**.**

There are a few established methods to enrich for Ub from cells[3,4], but none are specific for native branched Ub chains. Our approach to identifying branched Ub chains in cellular extracts is to enrich for these conjugates using a capture protein that selectively interacts with branched topologies. For this purpose, we intend to use UCHL3, a Ub C-terminal Hydrolase (UCH), which recently has been found to specifically bind and cleave Ub chains branched at K11 and K63. The binding affinity for UCHL3 and a Ub trimer with branch points at K11 and K63 is ten-fold higher than that with other Ub linkage types (unpublished data, Larry Anderson). By developing a novel enrichment strategy for branched Ub, we can expand the utility of middle-down MS to observe chain branching from a cellular context.

5.2 Results and Discussion

5.2.1 UchL3 can pull down branched Ub trimer

We sought to determine whether K11/K63 linked branched Ub trimers can be pulled-down by agarose resin that has been conjugated to UchL3. First, K11/K63 branched trimers were synthesized via TEC between K11C, K63C Ub and aaUb. Trimers were mixed with cell lysate from MEF (mouse embryonic fibroblast) cells[5] to mimic a pull down from cell extract. 10mM chloroacetamide was added to the cell lysate to inhibit deubiquitinating activity from endogenous DUBs.

The catalytic knockout of UchL3 (UchL3C95S) was immobilized on affi gel 10 lysine reactive agarose beads via standard coupling procedure from the manufacturer. Four variations of agarose beads were prepared, each containing a different immobilized protein: 1. UchL3C95S fused to maltose binding protein; 2. free UchL3; 3. free maltose binding protein; 4. Beads lacking protein. Variations 3 and 4 provided negative controls for the experiment. The Ub trimer/cell lysate mixture was incubated with beads, which was then washed with Tris-HCl pH 7.4. Finally, any remaining protein was eluted with 6M GndHCl and visualized by SDS-PAGE analysis. Ub trimer was observed in the elution lane from beads that were conjugated to UchL3C95S fused to maltose binding protein. Trimer did not elute from free quenched agarose beads (figure 5.1) or from immobilized maltose binding protein. These experiments were repeated with UchL3C95A, as an alternative catalytic knockout, and results were similar.

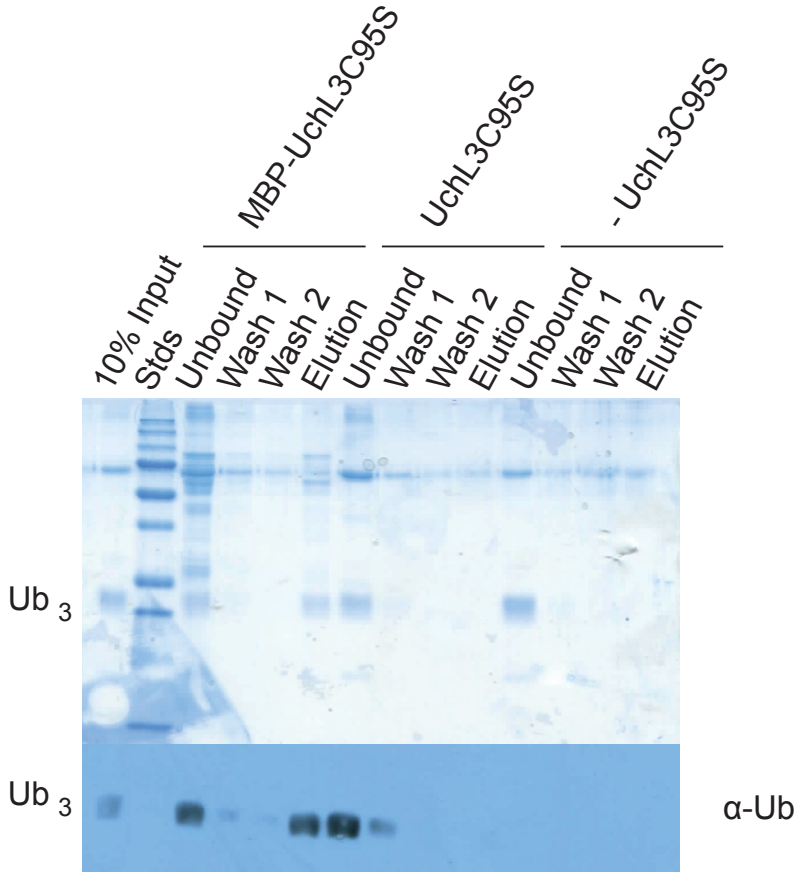


Figure 5.2 11/63 branched trimer pull down using immobilized UchL3. Coomasie stained SDS-PAGE and α-Ub western blot.

The eluted Ub trimer was then exchanged into ammonium bicarbonate buffer for tryptic digest, but only peptide fragments were later observed by MS analysis. We presume Ub was over digested due to improper refolding upon buffer exchange. Slow buffer exchange into tris refolding buffer should allow for better digestion conditions in the future.

Preliminary experiments were also performed to determine whether co-elution of a non-covalently bound UchL3 catalytic knockout and branched Ub chains could be digested and analyzed together by MS as an alternative to eluting trimer from immobilized UchL3.

Wild type Ub was mixed with UchL3 C95S at different ratios, digested with trypsin and analyzed by middle down MS. Ub species were easily observed, but digestion was not complete when the Ub binding partner, UchL3, was present in the digestion mixture, due to the presence of Ub_{1-76} (figure 5.2). We hypothesize that UchL3 binding to Ub protects the latter from efficient trypsin cleavage. For future experiments, digestion conditions will be optimized to match Ub pull down concentrations. The digestion step can be evaluated by monitoring the extent of Ub hydrolysis at R74.

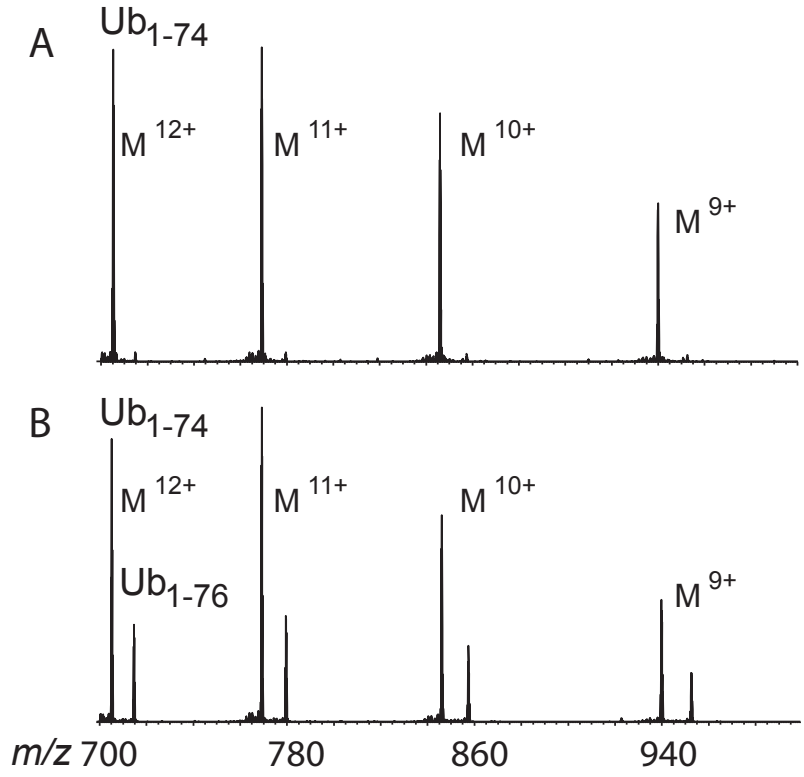


Figure 5.3. MS analysis of Ub_{1-76} digested in the absence (A) and presence (B) of UchL3

5.2.2 Minimal digest optimization

Minimal digestion of Ub chains to monomers, without over digestion to peptides, is crucial for observing branch points. Digestion conditions depend on the concentration of

Ub and other proteins present in the sample, so we sought to monitor trypsin digest of varying concentrations of Ub, both monomer and chains. Either 50, 5, or 0.5 pmols of Ub chains or Ub monomer (1 μM, 100 nM, or 10 nM) were mixed with 0.5 μg trypsin, for varying times (0.5, 3, or 6.5 hours). Total protein in all samples was standardized by adding BSA to 50 pmols. In general, Ub monomer species were observable by MS when starting concentrations of Ub were 1 μM or 100 nM, but not 10 nM. For all samples, trypsin digest did not go to completion, even at longer time points (figure 5.4). We conclude from these studies that it will be necessary to know an approximate concentration of Ub after enrichment, in order to appropriately optimize digestion conditions. Also, Ub concentration before trypsin digest should be at least 100 nM in order to observe Ub species by MS.

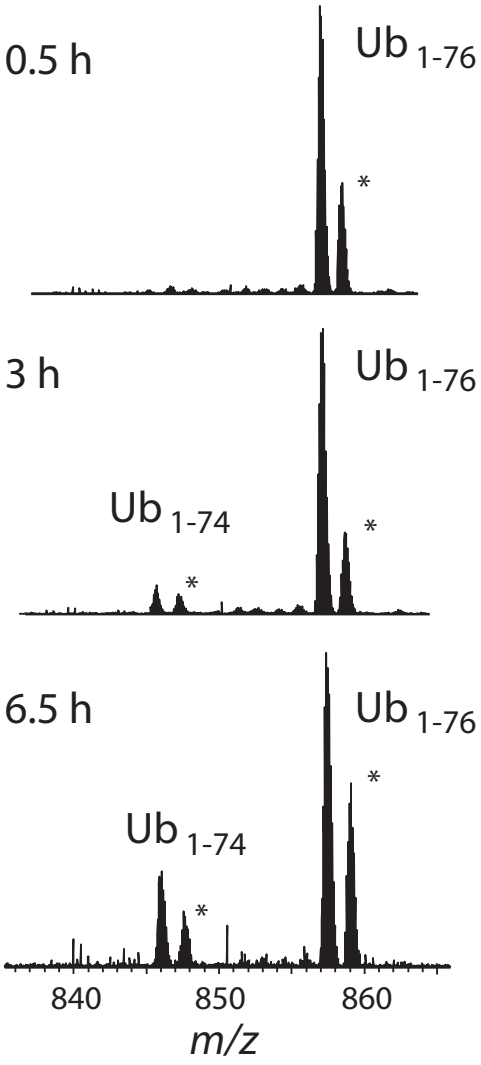


Figure 5.4. MS spectra of 1 μM Ub digested with 0.5 μg trypsin over varying time points. (* denotes oxidized Ub).

5.3 Future Directions

5.3.1 Elution conditions

Our lab has developed a fluorescence anisotropy assay to look at binding of Ub to UchL3 (Larry Anderson, unpublished data). By testing the dissociation constant of branched trimer and mono Ub, we can find the best wash and elution conditions for Ub

enrichment. I propose checking binding with different pH levels and salt concentrations to find wash conditions that allow a large difference in retention of monomer and branched trimer. For elution of trimer, it would be advantageous to find conditions that decrease the affinity between trimer and UchL3 while keeping the native fold of Ub, since retaining native Ub tertiary structure is crucial for the minimal trypsin digest step.

5.3.2 Optimize minimal trypsin digest

To implement this technique for the analysis of Ub chains from cells, trypsin digestion conditions must be optimized for different ratios of Ub to co-eluting proteins and at low Ub concentrations. We have already demonstrated that trypsin digest is influenced by the Ub tertiary structure and the presence of Ub binding domains. A possible strategy to monitor digestion of Ub chains to monomers is to spike N15 labeled wild type Ub into all samples at a low concentration. Since heavy Ub is distinguishable by mass from Ub in the samples of interest, it could serve as an internal standard to report whether all Ub has digested. Complete digestion to monomer is crucial, because undigested $Ub_{1\text{-}76}$ and $Ub_{1\text{-}74}$ modified with GG at a lysine residue have the same molecular weight. Thus, incomplete digestion of mono Ub would lead to misleading ratios of Ub species.

5.3.3 Monitoring branching in cells

Once the enrichment and digestion conditions are optimized, we plan to apply this technique to compare levels of Ub branching from MEF cells in the presence and absence of UchL3. UchL3 is known to promote insulin signaling[5], but its specific substrate *in vivo* is unknown. Our lab has observed that UchL3 cleaves Ub trimers branched at K11 and K63

more rapidly than other Ub trimers *in vitro*. By monitoring chain branching in cells +/- UchL3, we can place its hydrolase activity in a biological context.

5.4 Materials and Methods

Ubiquitin expression, purification, and chain formation was carried out as in chapter 4. Basic digestion conditions were carried out as in chapter 4, with variations on time and concentration of Ub as stated above.

5.4.1 Expression and purification of UchL3

UchL3 C95S and C95A were cloned into pDB.His.MBP with the restriction sites NdeI and XhoI. Briefly, 6xHis-tagged UchL3 was expressed in *E. coli* BL-21 cells grown in LB medium (OD_{600} of 0.6) at 37 ^{0}C, induced with IPTG (0.04 mM) and grown at 16 ^{0}C (16 h). The 6xHis-tagged UchL3 was purified by cobalt resin. From half of the eluted sample, the 6xHis-MBP tag was cleaved with TEV protease (4 °C, 16 h) and the protein (with and without MBP fusion) was further purified by anion exchange chromatography.

5.5 References

[1] Komander, D., & Rape, M. (2012). The ubiquitin code. *Annual Review of Biochemistry*, *81*, 203–229.
[1] Komander, D., & Rape, M. (2012). The ubiquitin code. *Annual Review of Biochemistry*, *81*, 203–229.

[2] Valkevich, E. M., Sanchez, N. A., Ge, Y., & Strieter, E. R. (2014). Middle-down mass spectrometry enables characterization of branched ubiquitin chains. *Biochemistry*, *53*, 4979–4989.

[3] Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., England, P., & Rodriguez, M. S. (2009). Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Reports*, *10*, 1250–1258.

[4] Xu, G., Paige, J. S., & Jaffrey, S. R. (2010). Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature Biotechnology*, *28*, 868–873.

[5] Suzuki, M., Setsuie, R., & Wada, K. (2009). Ubiquitin carboxyl-terminal hydrolase l3 promotes insulin signaling and adipogenesis. *Endocrinology*, *150*, 5230–5239.